TEMAS DE MATEMÁTICAS PARA ECONOMÍA

Héctor Manuel Mora Escobar

hectormora@yahoo.com   hmmorae@unal.edu.co
www.hectormora.info

April 30, 2018

Tabla de contenido

Prólogo

Este documento pretende presentar los temas correspondientes al curso Optimización para los estudiantes de pregrado de la Facultad de Economía de la Universidad Externado de Colombia en Bogotá.

El autor agradecerá los comentarios, sugerencias y correcciones. En particular ahora que es apenas un documento en preparación (empezado en enero 2017) y así, los errores, omisiones, ambigüedades están en su mayor nivel.

El programa oficial del curso sigue en gran parte el libro *Economía Matemática* de Diego Escobar.

En los libros generales de optimización, la mayoría de los problemas son de minimización. Por el contrario, en muchos libros enfocados a economía, generalmente los problemas son de maximización. Obviamente se puede pasar de un problema de maximización a uno de minimización, y viceversa, multiplicando la función objetivo por -1. El autor de este documento es matemático con énfasis en optimización, así, en los capítulos de optimización no lineal, la mayoría de los problemas serán de minimización.

Otra pequeña diferencia de enfoque tiene que ver con lo siguiente. Este texto presenta principalmente los métodos y herramientas sin un enfoque particular, es decir, se espera que el lector adquiera estos conocimientos sin importar que su interés sea la economía, las finanzas, la administración, la ingeniería, ..., con el objetivo de utilizarlos en cualquier área. Habrá de todas maneras algunos ejemplos de problemas de economía. En otros textos los conceptos están casi siempre aplicados a economía.

Conocimientos previos:

Álgebra lineal, obtener la matriz escalonada reducida, resolver un sistema de ecuaciones (inconsistente o una única solución o un número infinito de soluciones) calcular determinantes ($n = 2, 3, 4$), determinar si un conjunto de vectores es linealmente independiente, obtener una base del espacio nulo de una matriz, obtener el polinomio característico ($n = 2, 3$), hallar los valores propios ($n = 2, 3$ con polinomio de raíces enteras), determinar de multiplicidad algebraica y multiplicidad geométrica, obtener una base del espacio propio de un vector propio real (obtener uno o más vectores independientes asociados a un valor propio), obtener un vector propio de un valor propio complejo.

Cálculo vectorial, continuidad de funciones, funciones diferenciables, funciones doblemente diferenciables, obtener derivadas parciales de primer y segundo orden

Ecuaciones diferenciales, obtener la solución general de una ecuación diferencial lineal con coeficientes constantes homogénea, obtener una solución particular de una ecuación diferencial lineal con coeficientes constantes no homogénea,

Capítulo 1

Matrices definidas y semidefinidas positivas

1.1 Factorización de Cholesky

Sea A una matriz simétrica. Bajo ciertas condiciones (como se verá posteriormente, si y solamente si A es definida positiva), existe una matriz U triangular superior invertible tal que

$$U^{\mathrm{T}}U = A. \tag{1.1}$$

Esta factorización, cuando existe, se llama factorización de Cholesky y, en este caso, basta con conocer la matriz U.

El cálculo se puede hacer por filas, es decir, primero se obtienen los elementos de la primera fila de U, en seguida los de la segunda, etc.

Ejemplo 1.1.

$$A = \begin{bmatrix} 4 & -6 & 8 \\ -6 & 34 & -42 \\ 8 & -42 & 55 \end{bmatrix}$$

$$U^{\mathrm{T}}U = A$$

$$\begin{bmatrix} u_{11} & 0 & 0 \\ u_{12} & u_{22} & 0 \\ u_{13} & u_{23} & u_{33} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} = \begin{bmatrix} 4 & -6 & 8 \\ -6 & 34 & -42 \\ 8 & -42 & 55 \end{bmatrix}$$

Primera fila de U^{T} por primera columna de U:

$$\begin{aligned} (U^{\mathrm{T}})_{1\cdot}\, U_{\cdot 1} &= 4 \\ {u_{11}}^2 &= 4 \\ u_{11} &= 2 \end{aligned}$$

Primera fila de U^{T} por segunda columna de U:

$$\begin{aligned} (U^{\mathrm{T}})_{1\cdot}\, U_{\cdot 2} &= -6 \\ u_{11}u_{12} &= -6 \\ 2\,u_{12} &= -6 \\ u_{12} &= -3 \end{aligned}$$

$$\begin{aligned} (U^{\mathrm{T}})_{1\cdot}\, U_{\cdot 3} &= 8 \\ u_{11}u_{13} &= 8 \\ 2\,u_{13} &= 8 \\ u_{13} &= 4 \end{aligned}$$

$$\begin{aligned}
(U^{\mathrm{T}})_{2\cdot}\, U_{\cdot 2} &= 34\\
{u_{12}}^2 + {u_{22}}^2 &= 34\\
(-3)^2 + {u_{22}}^2 &= 34\\
{u_{22}}^2 &= 25\\
u_{22} &= \ 5
\end{aligned}$$

$$\begin{aligned}
(U^{\mathrm{T}})_{2\cdot}\, U_{\cdot 3} &= -42\\
u_{12}u_{13} + u_{22}u_{23} &= -42\\
-3 \times 4 + 5u_{23} &= -42\\
5u_{23} &= -30\\
u_{23} &= -6
\end{aligned}$$

$$\begin{aligned}
(U^{\mathrm{T}})_{3\cdot}\, U_{\cdot 3} &= 55\\
{u_{13}}^2 + {u_{23}}^2 + {u_{33}}^2 &= 55\\
4^2 + (-6)^2 + {u_{33}}^2 &= 55\\
{u_{33}}^2 &= 3\\
u_{33} &= \sqrt{3} \approx 1.7321
\end{aligned}$$

Entonces

$$U = \begin{bmatrix} 2 & -3 & 4\\ 0 & 5 & -6\\ 0 & 0 & 1.7321 \end{bmatrix}$$

y la matriz A tiene factorizacion de Cholesky. $\diamond$

En Scilab o en Matlab, se puede digitar

```
A = [4  -6  8; -6  34  -42; 8  -42  55]
U = chol(A)
```

Ejemplo 1.2.

$$A = \begin{bmatrix} 16 & \text{-12} & 8\\ \text{-12} & 7 & \text{-6}\\ 8 & \text{-6} & 5 \end{bmatrix}$$

$$\begin{aligned}
u_{11} &= 4\\
u_{12} &= \text{-3}\\
u_{13} &= 2
\end{aligned}$$

$$u_{22} = \sqrt{7 - (\text{-3})^2} = \sqrt{-2},$$

luego no existe la factorización de Cholesky para esta matriz A. $\diamond$

Ejemplo 1.3.

$$A = \begin{bmatrix} 16 & \text{-12} & 8\\ \text{-12} & 18 & \text{-6}\\ 8 & \text{-6} & 4 \end{bmatrix}$$

$$\begin{aligned}
u_{11} &= 4\\
u_{12} &= \text{-3}\\
u_{13} &= 2\\
u_{22} &= 3\\
u_{23} &= 0\\
u_{33} &= \sqrt{0} = 0,
\end{aligned}$$

luego, aunque con

$$U = \begin{bmatrix} 4 & \text{-}3 & 2 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

se tiene la igualdad $A = U^{\mathrm{T}}U$, no existe la factorización de Cholesky para esta matriz A puesto que U no es invertible. $\diamond$

Observación: Si la factorización de Cholesky existe, no es única, pero hay únicamente una matriz U con los elementos diagonales positivos.

1.2 Matrices definidas positivas

Definición 1.1. Una matriz A real, simétrica (obviamente cuadrada) es **definida positiva** (o positivamente definida) si:

$$x^{\mathrm{T}}Ax > 0 \qquad \text{para todo } x \neq 0. \tag{1.2}$$

Algunas veces se utiliza la notación

$$A \succ 0$$

para indicar que A es definida positiva.

Ejemplo 1.4. La matriz identidad de orden n.

$$\begin{aligned} x^{\mathrm{T}}Ix &= x^{\mathrm{T}}x \\ &= \|x\|_2^2 \\ &\geq 0 \qquad \text{para todo } x \\ &> 0 \qquad \text{para todo } x \neq 0. \end{aligned}$$

Luego la matriz identidad es definida positiva. $\diamond$

Ejemplo 1.5. La matriz nula de orden n.

$$x^{\mathrm{T}}\mathbf{0}x = 0.$$

luego la matriz nula no es definida positiva. $\diamond$

Para una matriz simétrica A

$$x^{\mathrm{T}}Ax = \sum_{i=1}^{n} a_{ii}x_i^2 \; + 2\sum_{i=1}^{n\text{-}1}\sum_{j=i+1}^{n} a_{ij}x_ix_j.$$

Ejemplo 1.6.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$$

$$\begin{aligned} x^{\mathrm{T}}Ax &= x_1^2 + 5x_2^2 + 4x_1x_2 \\ &= (x_1 + 2x_2)^2 + x_2^2 \\ &\geq 0 \qquad \text{para todo } x \\ &= 0 \qquad \text{sssi} \quad (x_1 + 2x_2)^2 = 0,\ \ x_2^2 = 0 \\ &= 0 \qquad \text{sssi} \quad x_1 + 2x_2 = 0,\ \ x_2 = 0 \\ &= 0 \qquad \text{sssi} \quad x_1 = 0,\ \ x_2 = 0 \\ &= 0 \qquad \text{sssi} \quad x = 0, \end{aligned}$$

luego A es definida positiva. $\diamond$

Ejemplo 1.7.

$$B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

$$\begin{aligned} x^{\mathrm{T}}Bx &= x_1^2 + 4x_2^2 + 4x_1x_2 \\ &= (x_1 + 2x_2)^2 \\ &\geq 0 \quad \text{para todo } x, \\ &= 0 \ , \text{ por ejemplo para } x_1 = 2, \quad x_2 = -1, \end{aligned}$$

luego B no es definida positiva. $\diamond$

Ejemplo 1.8.

$$C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

$$\begin{aligned} x^{\mathrm{T}}Cx &= x_1^2 + 3x_2^2 + 4x_1x_2 \\ &= (x_1 + 2x_2)^2 - x_2^2, \\ &= -1 \ , \text{ por ejemplo para } x_1 = 2, \quad x_2 = -1, \end{aligned}$$

luego C no es definida positiva. $\diamond$

Definición 1.2. Una **submatriz principal** de A es la obtenida al quitar de A algunas (o ninguna) filas y exactamente las columnas correspondientes. Una **submatriz estrictamente principal** A_k es la obtenida al quitar de A las filas y columnas $k+1$, $k+2$, ..., n con $1 \leq k \leq n$, es decir, las n submatrices estrictamente principales de A son:

$$A_1 = \begin{bmatrix} a_{11} \end{bmatrix} \ , \quad A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \ldots, \quad A_n = A.$$

Definición 1.3. El determinante de una submatriz principal se llama un **subdeterminante principal**; el determinante de una submatriz estrictamente principal se llama un **subdeterminante estrictamente principal**. Los n subteterminantes estrictamente principales son:

$$\delta_1 = \det \begin{bmatrix} a_{11} \end{bmatrix} = a_{11}, \quad \delta_2 = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \ldots, \quad \delta_n = \det(A).$$

En ejemplos pequeños, es relativamente fácil aplicar directamente la definición para saber si una matriz simétrica es definida positiva. Para ejemplos grandes, no solo se vuelve difícil, sino casi imposible. La siguiente proposición da condiciones necesarias y suficientes para la caracterización de matrices definidas positivas.

Proposición 1.1. (Condiciones necesarias y suficientes.) *Dada una matriz simétrica las siguientes siete afirmaciones son equivalentes:*

(1) A es definida positiva.

(2) Todos los δ_i, subdeterminantes estrictamente principales, son positivos.

(3) Todos los λ_i, valores propios de A (reales por ser simétrica), son positivos.

(4) A tiene factorización de Cholesky.

(5) Todos los pivotes a_{kk}^k, en el método de eliminación de Gauss sin permutación, son positivos.

(6) Existe una matriz W invertible tal que $A = W^{\mathrm{T}}W$.

(7) Todos los subdeterminantes principales son positivos.

Para matrices grandes, en casos no triviales, la caracterización más usada, más rápida y numéricamente más precisa es la factorización de Cholesky.

En este documento se utilizarán principalmente los criterios (2), (3) y (4). Para matrices pequeñas, generalmente, el criterio más sencillo es el de los subdeterminantes estrictamente principales.

Obviamente, si se tiene una matriz U triangular superior e invertible, se tiene una matriz W invertible. Tratar de averiguar si una matriz es definida positiva mediante la sexta caracterización casi nunca se utiliza, salvo en el caso, en que por anticipado, se sabe que la matriz $A = W^{\mathrm{T}}W$, con W invertible.

En Scilab, el cálculo numérico de los valores propios de puede hacer por medio de la función `spec`:

```
A = [4  -6  8; -6  34  -42; 8  -42  55]
vp = spec(A)
```

En Matlab, usando `eig`

```
A = [4  -6  8; -6  34  -42; 8  -42  55]
vp = eig(A)
```

En Scilab y Matlab el cálculo del determinante de hace mediante la función `det`:

```
A = [1  2  3; 4  5  6; 7  8  9]
det(A)
```

Es importante recalcar que los valores obtenidos son aproximaciones provenientes de un método numérico y no son absolutamente exactos. El cálculo del determinante en el ejemplo anterior da

```
6.661D-16
```

o algo semejante. Esto significa 6.61×10^{-16}, es decir casi cero. El cálculo a mano produce el valor exacto, cero. De manera análoga

```
A = [1  2  3; 4  5  6; 7  8  9]
spec(A)
```

produce

```
  16.116844
- 1.116844
- 1.304D-15
```

o algo parecido. Esto quiere decir que un valor propio de A es aproximadamente, 1.304×10^{-15}, o sea, casi cero. Realmente un valor propio es cero y los otros dos son aproximadamente, 16.116844 y -1.116844.

Ejemplo 1.9. Los valores propios de la matriz identidad son $1, 1, ..., 1$; $\delta_i = 1$ para todo i ; todos los pivotes, en la eliminación gaussiana, tienen el valor 1 ; $I = I^{\mathrm{T}}I$. Como ejemplo de la sexta caracterización:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{3}/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ -1/2 & 0 & \sqrt{3}/2 \end{bmatrix} \begin{bmatrix} \sqrt{3}/2 & 0 & -1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & \sqrt{3}/2 \end{bmatrix}$$

Por cualquiera de los criterios se ve que la matriz identidad es definida positiva. ◇

Ejemplo 1.10. Los valores propios de la matriz nula son 0, 0, ..., 0 ; $\delta_i = 0$ para todo i ; la eliminación gaussiana no se puede efectuar; como la matriz nula no es invertible no se puede expresar como producto de matrices invertibles. Por cualquiera de los criterios se ve que la matriz nula no es definida positiva. ◇

Ejemplo 1.11.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \quad \det(A - \lambda I) = (1-\lambda)(5-\lambda) - 4 = \lambda^2 - 6\lambda + 1,$$

sus raíces (los valores propios) son: $\lambda_1 = 3 + \sqrt{8}$, $\lambda_2 = 3 - \sqrt{8}$; $\delta_1 = 1$, $\delta_2 = 1$; en la eliminación de Gauss sin permutaciones

$$A^{(1)} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \quad A^{(2)} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

luego $a_{11}^{(1)} = 1$, $a_{22}^{(2)} = 1$.

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

Luego, por cualquiera de los criterios, la matriz A es definida positiva. $\diamond$

Ejemplo 1.12.

$$B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad \det(B - \lambda I) = (1-\lambda)(4-\lambda) - 4 = \lambda^2 - 5\lambda,$$

sus raíces (los valores propios) son: $\lambda_1 = 5$, $\lambda_2 = 0$; $\delta_1 = 1$, $\delta_2 = 0$; en la eliminación de Gauss sin permutaciones

$$B^{(1)} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad B^{(2)} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix},$$

luego $b_{11}^{(1)} = 1$, $b_{22}^{(2)} = 0$. Al tratar de encontrar la factorización de Cholesky:

$$B = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}.$$

Como B no es invertible no se puede encontrar W invertible, tal que $B = W^{\mathrm{T}}W$. Luego, por cualquiera de los criterios, la matriz B no es definida positiva. $\diamond$

Ejemplo 1.13.

$$C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad \det(C - \lambda I) = (1-\lambda)(3-\lambda) - 4 = \lambda^2 - 4\lambda - 1,$$

sus raíces (los valores propios) son: $\lambda_1 = 2 + \sqrt{5}$, $\lambda_2 = 2 - \sqrt{5} < 0$; $\delta_1 = 1$, $\delta_2 = -1$; en la eliminación de Gauss sin permutaciones

$$C^{(1)} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad C^{(2)} = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix},$$

luego $c_{11}^{(1)} = 1$, $c_{22}^{(2)} = -1$. Al tratar de encontrar la factorización de Cholesky ésta no se puede realizar, ya que aparece la raíz de un número negativo. Si $C = W^{\mathrm{T}}W$ entonces $\det(C) = \det(W^{\mathrm{T}})\det(W) = \big(\det(W)\big)^2$, pero como $\det(C) = -1$, entonces no existe tal matriz W. Luego, por cualquiera de los criterios, la matriz C no es definida positiva. $\diamond$

Proposición 1.2. (Condiciones necesarias.) *Si A es una matriz simétrica definida positiva entonces:*

(1) $a_{ii} > 0$ *para todo* i *;*

(2) $a_{ij}^2 < a_{ii}a_{jj}$ *para todo* $i \neq j$ *;*

(3) $\max_i a_{ii} = \max_{i,j} |a_{ij}|$, *es decir,* $\max_{i,j} |a_{ij}|$ *es un elemento diagonal;*

(4) $2|a_{ij}| < a_{ii} + a_{jj}$ *para todo* $i \neq j$.

Ejemplo 1.14. Sean:

$$A = \begin{bmatrix} 5 & 6 \\ 6 & \text{-3} \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 4 \\ 4 & 7 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 3 \\ 3 & 3 \end{bmatrix},$$
$$C = \begin{bmatrix} 10 & \text{-3} & 5 \\ \text{-3} & 1 & -1 \\ 5 & \text{-1} & 3 \end{bmatrix}.$$

Para cada criterio se presenta un ejemplo de una matriz que no lo cumple, luego no es definida positiva y un segundo ejemplo de una matriz que cumple el criterio y, sin embargo, no es definida positiva.

(1) A no cumple el criterio, luego no es definida positiva.
   B cumple el criterio y, sin embargo no es definida positiva.

(2) B no cumple el criterio, luego no es definida positiva.
   C cumple el criterio y, sin embargo, no es definida positiva.

(3) A no cumple el criterio, luego no es definida positiva.
   B cumple el criterio y, sin embargo, no es definida positiva.

(4) D no cumple el criterio, luego no es definida positiva.
   B cumple el criterio y, sin embargo, no es definida positiva. ◇

Definición 1.4. Una matriz es de **diagonal estrictamente dominante por filas** si:

$$\sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| < |a_{ii}| \,, \qquad \text{para todo } i.$$

Una matriz es de **diagonal estrictamente dominante por columnas** si:

$$\sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| < |a_{jj}| \,, \qquad \text{para todo } j.$$

Una matriz es de **diagonal estrictamente dominante** si es de diagonal estrictamente dominante por columnas y de diagonal estrictamente dominante por filas.

Es obvio que para matrices simétricas las tres definiciones anteriores coinciden.

Proposición 1.3. (Condiciones suficientes.) *Si A es una matriz simétrica, de diagonal positiva y estrictamente dominante, entonces es definida positiva.*

Ejemplo 1.15.

$$A = \begin{bmatrix} 4 & 1 & \text{-2} \\ 1 & 5 & 3 \\ \text{-2} & 3 & 6 \end{bmatrix}$$

cumple el criterio, luego es definida positiva. ◇

Ejemplo 1.16.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$$

no cumple el criterio y, sin embargo, es definida positiva. ◇

1.3 Matrices definidas negativas

Definición 1.5. Una matriz A real, simétrica (obviamente cuadrada) es **definida negativa** (o negativamente definida) si:

$$x^{\mathrm{T}}Ax < 0 \qquad \text{para todo } x \neq 0. \tag{1.3}$$

Proposición 1.4. *Sea A un matriz simétrica. Las siguientes afirmaciones son equivalentes*

(1) A es definida negativa.

(2) $-A$ es definida positiva.

(3)

$$\begin{aligned} &\delta_1 < 0 \\ &\delta_2 > 0 \\ &\delta_3 < 0 \\ &\vdots \\ &\delta_i < 0 \quad \textit{si } i \textit{ es impar} \\ &\delta_i > 0 \quad \textit{si } i \textit{ es par} \end{aligned}$$

(4) Todos los λ_i, valores propios de A son negativos,

(5) $-A$ tiene factorización de Cholesky.

La caracterización de matrices definidas negativas por medio de los subdeterminantes esctrictamente principales también es puede escribir

$$(-1)^i\delta_i > 0. \tag{1.4}$$

Ejemplo 1.17.

$$A = \begin{bmatrix} -4 & 6 \\ 6 & -10 \end{bmatrix}$$

Por cualquiera de las caracterizaciones de la sección anterior se puede verificar que $-A$ es definida positiva. Luego A es definida negativa.

$$\begin{aligned} \delta_1 &= -4 < 0 \\ \delta_2 &= \ \ 4 > 0 \end{aligned}$$

Luego A es definida negativa.

$$\begin{aligned} \lambda_1 &\approx -13.708204 \\ \lambda_2 &\approx \ \ -0.291796 \end{aligned}$$

Luego A es definida negativa.

Ejemplo 1.18.

$$A = \begin{bmatrix} -2 & 3 & 4 \\ 3 & -6 & 0 \\ 4 & 0 & -1 \end{bmatrix}$$

$$\begin{aligned} \delta_1 &= -2 < 0 \\ \delta_2 &= \ \ 3 > 0 \\ \delta_3 &= \ 93 \not< 0 \end{aligned}$$

Luego A no es definida negativa.

1.4 Matrices semidefinidas positivas

Definición 1.6. Una matriz A real, simétrica (obviamente cuadrada) es **semidefinida positiva** o definida no negativa si:

$$x^{\mathrm{T}}Ax \geq 0 \qquad \text{para todo } x.$$

Algunas veces se utiliza la notación

$$A \succcurlyeq 0$$

para indicar que A es semidefinida positiva.

Es evidente que toda matriz definida positiva es semidefinida positiva.

Algunas veces, no en estas notas, se agrega a la definición de matrices semidefinidas positivas la existencia de un x no nulo tal que $x^{\mathrm{T}}Ax = 0$. Según esta otra definición una matriz definida positiva no sería semidefinida positiva.

Ejemplo 1.19. Los ejemplos anteriores de matrices definidas positivas son matrices semidefinidas positivas. $\diamond$

Ejemplo 1.20. La matriz nula $\mathbf{0}$. Aplicando la definición $x^{\mathrm{T}}\mathbf{0}x = 0$, luego la matriz nula es semidefinida positiva. $\diamond$

Ejemplo 1.21.

$$B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix},$$

$$\begin{aligned} x^{\mathrm{T}}Bx &= x_1^2 + 4x_2^2 + 4x_1x_2 \\ &= (x_1 + 2x_2)^2 \\ &\geq 0 \qquad \text{para todo } x, \end{aligned}$$

luego B es semidefinida positiva. $\diamond$

Ejemplo 1.22.

$$C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix},$$

$$\begin{aligned} x^{\mathrm{T}}Cx &= x_1^2 + 3x_2^2 + 4x_1x_2 \\ &= (x_1 + 2x_2)^2 - x_2^2 \\ &= -1 \qquad \text{por ejemplo para} \quad x_1 = 2, \ \ x_2 = -1, \end{aligned}$$

luego C no es semidefinida positiva. $\diamond$

Ejemplo 1.23.

$$E = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix},$$

$$\begin{aligned} x^{\mathrm{T}}Ex &= -x_2^2 \\ &= -1 \qquad \text{por ejemplo para} \quad x_1 = 0, \ \ x_2 = 1, \end{aligned}$$

luego E no es semidefinida positiva. $\diamond$

Proposición 1.5. *Dada una matriz A simétrica, las siguientes afirmaciones son equivalentes:*

(1) A es semidefinida positiva;

(2) todos los λ_i valores propios de A son no negativos;

(3) todos los subdeterminantes principales son no negativos, es decir, mayores o iguales a cero.

(4) existe una matriz W, posiblemente no invertible, tal que $A = W^{\mathrm{T}}W$;

Es claro que las caracterizaciones para las matrices definidas positivas son más fuertes que las caracterizaciones para las matrices semidefinidas positivas. En los ejemplos de matrices definidas positivas se observa claramente que éstas cumplen con las caracterizaciones de matrices semidefinidas positivas.

Ejemplo 1.24. Los valores propios de la matriz nula son 0, 0, ..., 0 ; todos los subdeterminates principales son nulos; $\mathbf{0} = \mathbf{0}^{\mathrm{T}}\mathbf{0}$. Por cualquiera de los criterios se ve que la matriz nula es semidefinida positiva. ◇

Ejemplo 1.25.

$$B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad \det(B - \lambda I) = (1-\lambda)(4-\lambda) - 4 = \lambda^2 - 5\lambda,$$

sus raíces (los valores propios) son: $\lambda_1 = 5, \ \lambda_2 = 0$;

Los subdeterminantes principales son:

$$\begin{aligned} \det \begin{bmatrix} 1 \end{bmatrix} &= 1 \\ \det \begin{bmatrix} 4 \end{bmatrix} &= 4 \\ \det \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} &= 0 \end{aligned}$$

Al tratar de encontrar la factorización de Cholesky:

$$B = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix};$$

Luego, por cualquiera de los criterios, la matriz B es semidefinida positiva. ◇

Ejemplo 1.26.

$$C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad \det(C - \lambda I) = (1-\lambda)(3-\lambda) - 4 = \lambda^2 - 4\lambda - 1,$$

sus raíces (los valores propios) son: $\lambda_1 = 2 + \sqrt{5}$, $\lambda_2 = 2 - \sqrt{5} < 0$;

al tratar de encontrar la factorización de Cholesky ésta no se puede realizar, ya que aparece la raíz de un número negativo. Si $C = W^{\mathrm{T}}W$ entonces $\det(C) = \det(W^{\mathrm{T}})\det(W) = \big(\det(W)\big)^2$, pero como $\det(C) = -1$, entonces no existe tal matriz W. Los subdeterminantes principales son -1 (sin quitar filas ni columnas), 1 (quitando la segunda fila y la segunda columna) y 3 (quitando la primera fila y la primera columna). Luego, por cualquiera de los criterios, la matriz C no es semidefinida positiva. ◇

Ejemplo 1.27.

$$E = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \det(E - \lambda I) = (0-\lambda)(-1-\lambda) = \lambda^2 + \lambda,$$

sus raíces (los valores propios) son: $\lambda_1 = 0$, $\lambda_2 = -1$;

Al tratar de encontrar la factorización de Cholesky ésta no se puede realizar, ya que aparece la raíz de un número negativo. Tampoco se puede encontrar W tal que $E = W^{\mathrm{T}}W$. Los subdeterminantes principales son 0 (sin quitar filas ni columnas), 0 (quitando la segunda fila y la segunda columna) y -1 (quitando la primera fila y la primera columna). Luego la matriz E no es semidefinida positiva. ◇

Ejemplo 1.28.

$$A = \begin{bmatrix} 4 & -6 & -6 \\ -6 & 9 & 9 \\ -6 & 9 & 1 \end{bmatrix}$$

$$\begin{aligned} \delta_1 &= 4 \\ \delta_2 &= 0 \\ \delta_3 &= 0 \end{aligned}$$

Lo anterior no garantiza que A sea semidefinida positiva. En efecto no lo es ya que por lo menos un subdeterminante principal es negativo:

$$\det \begin{bmatrix} 9 & 9 \\ 9 & 1 \end{bmatrix} = -72$$

Por otro lado, los valores propios son -5.3693, 0 y 19.3693.

Proposición 1.6. (Condiciones necesarias.) *Si A es una matriz simétrica semidefinida positiva entonces:*

(1) $a_{ii} \geq 0$ *para todo* i *;*

(2) $a_{ij}^2 \leq a_{ii}a_{jj}$ *para todo* i, j *;*

(3) $\max a_{ii} = \max |a_{ij}|$, *es decir,* $\max |a_{ij}|$ *es un elemento diagonal;*

(4) $2|a_{ij}| \leq a_{ii} + a_{jj}$ *para todo* i, j *;*

(5) $a_{ii} = 0 \Rightarrow A_{i\cdot} = 0, A_{\cdot i} = 0$ *;*

(6) $\delta_i \geq 0$ *para todo* i*.*

Ejemplo 1.29.

$$\begin{aligned} A &= \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}, \\ D &= \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 8 & 5 & 8 \\ 5 & 4 & 6 \\ 8 & 6 & 9 \end{bmatrix}. \end{aligned}$$

Para cada criterio se presenta un ejemplo de una matriz que no lo cumple, luego no es semidefinida positiva y un segundo ejemplo de una matriz que cumple el criterio y, sin embargo, no es semidefinida positiva.

(1) A no cumple el criterio, luego no es semidefinida positva, sdp.
C cumple el criterio y, sin embargo, no es semidefinida positiva.

(2) B no cumple el criterio, luego no es semidefinida positiva.
E cumple el criterio y, sin embargo, no es semidefinida positiva.

(3) C no cumple el criterio, luego no es semidefinida positiva.
E cumple el criterio y, sin embargo, no es semidefinida positiva.

(4) C no cumple el criterio, luego no es semidefinida positiva.
E cumple el criterio y, sin embargo, no es semidefinida positiva.

(5) D no cumple el criterio, luego no es semidefinida positiva.
A cumple el criterio y, sin embargo, no es semidefinida positiva.

(6) B no cumple el criterio, luego no es semidefinida positiva.
A cumple el criterio y, sin embargo, no es semidefinida positiva. $\diamond$

1.5 Matrices semidefinidas negativas

Definición 1.7. Una matriz A simétrica es **semidefinida negativa** si

$$x^{\mathrm{T}}Ax \leq 0, \quad \text{para todo } x.$$

Proposición 1.7. *Dada una matriz A simétrica, las siguientes afirmaciones son equivalentes:*

(1) A es semidefinida negativa.

(2) todos los λ_i valores propios de A son no positivos (menores o iguales a cero).

(3) todos los subdeterminantes principales de orden impar son menores o iguales a cero y todos los subdeterminantes principales de orden par son mayores o iguales a cero.

(4) $-A$ es semidefinida positiva.

Definición 1.8. Una matriz simétrica es **indefinida** si no es ni semidefinida positiva ni semidefinida negativa, o lo que es lo mismo, si existen x, y tales que $(x^{\mathrm{T}}Ax)(y^{\mathrm{T}}Ay) < 0$, o también, si tiene por lo menos un valor propio positivo y por lo menos un valor propio negativo.

Ejemplo 1.30.

$$\begin{bmatrix} -1 & -2 \\ -2 & -5 \end{bmatrix} \quad \text{es definida negativa.} \quad \diamond$$

Ejemplo 1.31.

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} && \text{no es semidefinida positiva,} \\ -A &= \begin{bmatrix} -1 & -2 \\ -2 & -3 \end{bmatrix} && \text{tampoco es semidefinida positiva,} \end{aligned}$$

luego A es indefinida. Sean: $x = (2, -1)$, $y = (1, 0)$. Entonces $(x^{\mathrm{T}}Ax)\ (y^{\mathrm{T}}Ay) = (-1)(1) = -1$, lo que comprueba que A es indefinida. $\diamond$

Ejemplo 1.32.

$$A = \begin{bmatrix} -9 & 6 & -7 \\ 6 & -5 & -3 \\ -7 & -3 & -1 \end{bmatrix}$$

$$\begin{aligned} \det \begin{bmatrix} -9 \end{bmatrix} &= -9 \leq 0 \\ \det \begin{bmatrix} -5 \end{bmatrix} &= -5 \leq 0 \\ \det \begin{bmatrix} -1 \end{bmatrix} &= -1 \leq 0 \\ \det \begin{bmatrix} -9 & 6 \\ 6 & -5 \end{bmatrix} &= 9 \geq 0 \\ \det \begin{bmatrix} -9 & -7 \\ -7 & -1 \end{bmatrix} &= -40 \end{aligned}$$

Luego A no es semidefinida negativa. Por otro lado, sus valores propios son aproximadamente -14.76, -6.33 y 6.09.

1.6 Matrices diagonales por bloques

Sea A una matriz simétrica diagonal por bloques

????

donde cada bloque diagonal es a su vez una matriz simétrica.

Algunas propiedades útiles (algunas siguen siendo válidas para matrices no simétricas)

1. El determinante de A es el producto de los determinantes de los bloques.
2. El polinomio característico de A es el producto de los polinomios característicos de los bloques.
3. El conjunto de valores propios de A está conformado por los valores propios de los bloques.
4. A es definida positiva sssi todos los bloques son matrices definidas positivas.
5. A es semidefinida positiva sssi todos los bloques son matrices semidefinidas positivas.

1.7 Matrices definidas positivas en un subespacio

Es posible que una matriz no sea definida positiva pero cumpla la propiedad $x^{\mathrm{T}}Ax > 0$ para todos los vectores x en algún conjunto especial. Para subespcios vectoriales se tiene la siguiente definición.

Definición 1.9. Sea $\mathcal{E}$ un subespacio vectorial de $\mathbb{R}^n$. Una matriz simétrica A es definida positiva en $\mathcal{E}$ si

$$x^{\mathrm{T}}Ax > 0, \ \ \forall x \in \mathcal{E}, \ \ x \neq 0.$$

A es semidefinida positiva en $\mathcal{E}$ si

$$x^{\mathrm{T}}Ax \geq 0 \ , \ \ \forall x \in \mathcal{E}.$$

Obviamente si una matriz es definida positiva es definida positiva en cualquier subespacio. Si $\mathcal{E} = \mathbb{R}^n$, se tienen las definiciones usuales de matriz definida positiva y matriz semidefinida positiva. Si $\mathcal{E} = \{\mathbf{0}\}$ cualquier matriz simétrica es semidefinida positiva y definida positiva en $\mathcal{E}$.

La definición de matriz definida positiva en un subespacio vectorial es útil sobre todo cuando la matriz no es definida positiva, y sí es definida positiva en algún subespacio vectorial de $\mathbb{R}^n$.

Ejemplo 1.33. La matriz

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

no es definida positiva ni semidefinida positiva. Tampoco es semidefinida negativa. Sus valores propios son $1, -1$. Sea $x \in \mathcal{E} = \{(x_1, x_2) : 3x_1 - x_2 = 0\}$.

$$\begin{aligned} x^{\mathrm{T}}Ax &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= 2x_1x_2 \\ &= 2x_1(3x_1) \\ &= 6x_1^2 \\ &> 0 \quad \text{si } x_1 \neq 0. \end{aligned}$$

Luego A es definida positiva en $\mathcal{E}$. $\diamond$

Saber si una matriz es definida positiva en un subespacio vectorial aplicando directamente la definición puede ser fácil para matrices pequeñas, sin embargo, es muy útil tener procedimientos precisos para matrices de cualquier tamaño.

Sean: $\mathcal{E}$ un subespacio vectorial de $\mathbb{R}^n$, v^1, v^2, ..., v^k una base de $\mathcal{E}$, E la matriz $n \times k$ cuyas columnas son los vectores v^1, v^2, ..., v^k. Los elementos de la forma $E\xi$ caracterizan completamente a $\mathcal{E}$, es decir, x es un elemento de $\mathcal{E}$ si y solamente si existe $\xi \in \mathbb{R}^k$ tal que $x = E\xi$. Una matriz simétrica A es definida positiva en $\mathcal{E}$ si $x^{\mathrm{T}}Ax > 0$ para todo $x \in \mathcal{E}, x \neq 0$, o sea, si $(E\xi)^{\mathrm{T}}AE\xi = \xi^{\mathrm{T}}E^{\mathrm{T}}AE\xi > 0$ para todo $\xi \in \mathbb{R}^k, \xi \neq 0$.

Proposición 1.8. *Sea A una matriz simétrica. A es definida positiva en $\mathcal{E}$ si y solamente si $E^{\mathrm{T}}AE$ es definida positiva. A es semidefinida positiva en $\mathcal{E}$ si y solamente si $E^{\mathrm{T}}AE$ es semidefinida positiva.*

En esta proposición el criterio no depende de la base escogida. Para saber si la matriz $E^{\mathrm{T}}AE$, de tamaño $k \times k$, es definida positiva se utiliza cualquiera de los criterios vistos anteriormente.

Ejemplo 1.34. Sean:

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\ \mathcal{E} &= \{(x_1, x_2) : 3x_1 - x_2 = 0\}. \end{aligned}$$

La dimensión de $\mathcal{E}$ es 1. Una base de $\mathcal{E}$ es $v^1 = [1 \;\; 3]^{\mathrm{T}}$, entonces

$$E = \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

$$E^{\mathrm{T}} A E = [6] = 6 > 0,$$

luego A es definida positiva en $\mathcal{E}$. $\diamond$

Ejemplo 1.35. Sean:

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}, \\ \mathcal{E} &= \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 0\}. \end{aligned}$$

La dimensión de $\mathcal{E}$ es 2. Una base de $\mathcal{E}$ es $v^1 = [1 \;\; 0 \;\; \text{-}1]^{\mathrm{T}}$, $v^2 = [0 \;\; 1 \;\; \text{-}1]^{\mathrm{T}}$, entonces

$$\begin{aligned} E &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix}, \\ E^{\mathrm{T}} A E &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Luego A es semidefinida positiva en $\mathcal{E}$. Si se escoge otra base la conclusión será la misma. Otra base de $\mathcal{E}$ es $v^1 = [1 \;\; 0 \;\; \text{-}1]^{\mathrm{T}}$, $v^2 = [1 \;\; \text{-}2 \;\; 1]^{\mathrm{T}}$, entonces

$$\begin{aligned} E &= \begin{bmatrix} 1 & 1 \\ 0 & -2 \\ -1 & 1 \end{bmatrix}, \\ E^{\mathrm{T}} A E &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \end{aligned}$$

luego A es semidefinida positiva en $\mathcal{E}$. $\diamond$

1.7.1 En el espacio nulo de una matriz

Con frecuencia un subespacio vectorial está definido como el espacio nulo de una matriz. En este caso hay un procedimiento explícito para encontrar una base.

Definición 1.10. Sea M una matriz $p \times n$. El espacio nulo de M es el conjunto

$$\mathcal{N}(M) = \{x \in \mathbb{R}^n : Mx = 0\}.$$

Este conjunto es un subespacio vectorial de $\mathbb{R}^n$. Si M no es la matriz nula existe una matriz $\bar{M}$ cuyas filas son linealmente independientes y además

$$\mathcal{N}(M) = \mathcal{N}(\bar{M}).$$

En lo que sigue se supone que **las filas de M son linealmente independientes**. Esto implica que $1 \leq p \leq n$ y que el rango de M es p. Si $p = n$, entonces $\mathcal{N}(M) = \{\mathbf{0}\}$. Como generalización se puede decir que si M no tiene filas ($p = 0$), entonces $\mathcal{N}(M) = \mathbb{R}^n$.

El problema que estudiaremos a continuación es el siguiente. Dada A matriz simétrica $n \times n$ y M matriz $p \times n$, rango$(M) = p$, se desea saber si A es definida positiva (o semidefinida positiva) en $\mathcal{N}(M)$.

Si $p = n$, entonces A es definida positiva y semidefinida positiva en $\mathcal{N}(M) = \{\mathbf{0}\}$. Para la generalización $p = 0$, se trata simplemente de averiguar si A es definida positiva en $\mathbb{R}^n$. En consecuencia podemos suponer que $1 \leq p < n$. El espacio $\mathcal{N}(M)$ tiene dimensión $q = n - p$, o sea, cualquier base de $\mathcal{N}(M)$ tiene q elementos.

Para poder aplicar la proposición 1.8 se requiere conocer una matriz E cuyas columnas son los vectores de una base de $\mathcal{N}(M)$. Como rango$(M) = p$, entonces existe B de tamaño $p \times p$, submatriz de M, invertible. Entonces el sistema

$$Mx = 0,$$

es equivalente a

$$B^{-1}Mx = M'x = 0.$$

La matriz M' tiene una característica importante, tomando adecuadamente p columnas se obtiene la matriz identidad de orden p. La matriz M' se puede obtener por multiplicación explícita o también llevando M a la forma escalonada reducida por filas.

Sean: L la submatriz $p \times q$ formada por las demás columnas de M', x_B el vector columna $p \times 1$ construido con las componentes del vector x correspondientes a las columnas de M' que forman la matriz identidad, x_L el vector columna $q \times 1$ construido con las demás componentes del vector x. Las variables x_B se llaman variables básicas. Las variables x_L se llaman variables libres.

n El sistema $M'x = 0$ es equivalente a

$$\begin{aligned} \begin{bmatrix} I_p & L \end{bmatrix} \begin{bmatrix} x_B \\ x_L \end{bmatrix} &= 0, \\ \text{luego} \qquad x_B &= -Lx_L. \end{aligned}$$

Una manera de obtener una base de $\mathcal{N}(M)$ es mediante el siguiente procedimiento: dar a la primera variable libre (componente de x_L) el valor 1 y a las demás variables libres el valor 0. De esa manera $x_B = -L_{\cdot 1}$ y se obtiene el primer elemento de la base. Enseguida, se le da a la segunda variable libre el valor 1 y a las demás variables libres el valor 0. De esa manera $x_B = -L_{\cdot 2}$ y se obtiene el segundo elemento de la base. Y así sucesivamente hasta obtener los q elementos de la base. Entonces la matriz E tiene la forma

$$E = \begin{bmatrix} -L \\ I_q \end{bmatrix},$$

o la matriz obtenida por permutación de filas para concordar con el orden de las variables. De manera más precisa: si las variables básicas son las p primeras (y las variables libres las q últimas), entonces no se requiere hacer ninguna permutación de filas, en caso contrario se reordenan las filas.

Dicho en palabras, en las p filas de E correspondientes a las variables básicas se colocan las p filas de $-L$, y en las q filas de E correspondientes a las variables libres se colocan las q filas de I_q.

Ejemplo 1.36. Encontrar una matriz E para $\mathcal{N}(M)$ subespacio nulo de la matriz $M = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$.

La matriz M ya está en la forma escalonada reducida. Consideremos, por ejemplo, que la primera columna de M constituye I y que las dos últimas forman L

$$\begin{aligned} I &= \begin{bmatrix} 1 \end{bmatrix} \\ L &= \begin{bmatrix} 1 & 1 \end{bmatrix} \\ x_B &= \begin{bmatrix} x_1 \end{bmatrix} \\ x_L &= \begin{bmatrix} x_2 \\ x_3 \end{bmatrix}, \end{aligned}$$

entonces

$$E = \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \ \diamond$$

Ejemplo 1.37. Encontrar una matriz E para $\mathcal{N}(M)$ subespacio nulo de la matriz

$$M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}.$$

Llevándola a la forma escalonada reducida resulta

$$\begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \end{bmatrix}.$$

Las dos primeras columnas de M constituyen I y las dos últimas forman L

$$\begin{aligned} L &= \begin{bmatrix} -1 & -2 \\ 2 & 3 \end{bmatrix} \\ x_B &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ x_L &= \begin{bmatrix} x_3 \\ x_4 \end{bmatrix}, \end{aligned}$$

entonces

$$E = \begin{bmatrix} 1 & 2 \\ -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad \diamond$$

Ejemplo 1.38. Averiguar si A es semidefinida positiva en $\mathcal{N}(M)$ subespacio nulo de la matriz M,

$$\begin{aligned} A &= \begin{bmatrix} -2 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \\ M &= \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 7 & 8 \\ 3 & 6 & 9 & 10 & 12 \end{bmatrix}. \end{aligned}$$

Llevando M a la forma escalonada reducida se obtiene

$$\begin{bmatrix} 1 & 2 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Las columnas 1, 4, 5 constituyen I, la segunda y tercera forman L,

$$\begin{aligned} L &= \begin{bmatrix} 2 & 3 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \\ x_B &= \begin{bmatrix} x_1 \\ x_4 \\ x_5 \end{bmatrix} \\ x_L &= \begin{bmatrix} x_2 \\ x_3 \end{bmatrix}, \end{aligned}$$

entonces

$$\begin{bmatrix} -L \\ I_q \end{bmatrix} = \begin{array}{c} \begin{bmatrix} -2 & -3 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} x_1 \\ x_4 \\ x_5 \\ x_2 \\ x_3 \end{array}$$

En este caso es indispensable reordenar las filas para obtener la matriz

$$E = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Aplicando la proposición 1.8:

$$E^{\mathrm{T}}AE = \begin{bmatrix} -11 & -16 \\ -16 & -23 \end{bmatrix}.$$

Los valores propios de $E^{\mathrm{T}}AE$ son 0.0880, -34.0880, luego A no es semidefinida positiva en $\mathcal{N}(M)$. $\diamond$

EJERCICIOS

1.1 - 1.12 Diga si las siguientes matrices son definidas positivas, semidefinidas positivas, definidas negativas, semidefinidas negativas, o indefinidas.

$$A^1 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 6 & 6 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}, \quad A^3 = \begin{bmatrix} -4 & 5 \\ 5 & -7 \end{bmatrix},$$

$$A^4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix}, \quad A^5 = \begin{bmatrix} 9 & -18 & -15 \\ -18 & 40 & 22 \\ -15 & 22 & 42 \end{bmatrix}, \quad A^6 = \begin{bmatrix} -9 & 6 \\ 6 & -4 \end{bmatrix},$$

$$A^7 = \begin{bmatrix} 10 & 1 & -1 & 2 \\ 1 & 11 & -2 & 3 \\ -1 & -2 & 12 & -3 \\ 2 & 3 & -3 & 13 \end{bmatrix}, \quad A^8 = \begin{bmatrix} 10 & 1 & 1 & 2 \\ 1 & 20 & -2 & 3 \\ 1 & -2 & 30 & 25 \\ 2 & 3 & 25 & 20 \end{bmatrix},$$

$$A^9 = \begin{bmatrix} 10 & 1 & -1 & 2 \\ 1 & 11 & -2 & 3 \\ -1 & -2 & 12 & -14 \\ 2 & 3 & -14 & 13 \end{bmatrix}, \quad A^{10} = \begin{bmatrix} 10 & 1 & 1 & 2 \\ 1 & 20 & -2 & 3 \\ 1 & -2 & 30 & 2 \\ 2 & 3 & 2 & -20 \end{bmatrix},$$

$$A^{11} = \begin{bmatrix} 10 & 1 & 1 & 2 \\ 1 & 11 & 2 & 3 \\ 1 & 2 & 12 & 13 \\ 2 & 3 & 13 & 11 \end{bmatrix}, \quad A^{12} = \begin{bmatrix} 10 & 1 & 1 & 2 \\ 1 & 11 & 2 & 3 \\ 1 & 2 & 12 & 13 \\ 2 & 3 & 13 & 14 \end{bmatrix}.$$

1.13 Diga si las siguientes matrices A son definidas positivas o semidefinidas positivas en el espacio nulo de la matriz M.

$$A^{13} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad M^{13} = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}.$$

$$A^{14} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad M^{14} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}.$$

$$A^{15} = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}, \quad M^{15} = \begin{bmatrix} 3 & 2 & 1 \end{bmatrix}.$$

$$A^{16} = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}, \quad M^{16} = \begin{bmatrix} 2 & 4 & 6 \\ 3 & 6 & 10 \end{bmatrix}.$$

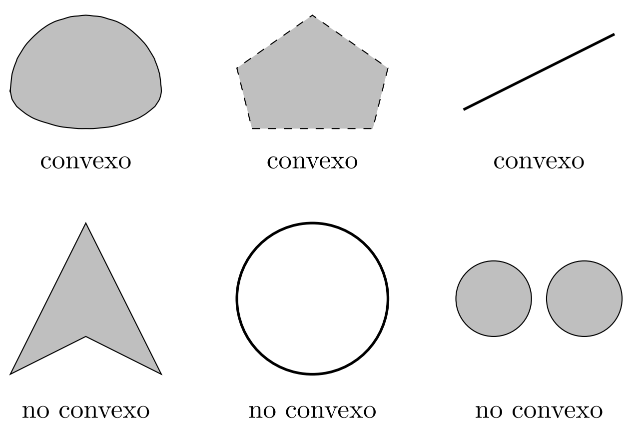
$$A^{17} = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}, \quad M^{17} = \begin{bmatrix} 2 & 4 & 6 \\ 3 & 6 & 10 \\ 10 & 10 & 10 \end{bmatrix}.$$

Capítulo 2

Conjuntos convexos y funciones convexas

Sea V el espacio vectorial $\mathbb{R}^n$. Mientras no se diga lo contrario, todos los conjuntos son subconjuntos de $\mathbb{R}^n$, todos los puntos o vectores son elementos de $\mathbb{R}^n$, todos los números son números reales. La mayoría de las definiciones y resultados que siguen, se pueden generalizar fácilmente a otros espacios vectoriales.

2.1 Conjuntos convexos



Definición 2.1. Sea C un subconjunto de V. Se dice que C es **convexo** si dados x, y en C, t un escalar en el intervalo $[0,1]$, entonces $z = z_{xyt} = (1\text{-}t)x + ty$ también está en C. Gráficamente, un conjunto C es convexo si dados dos puntos x, y en C , cualquier punto del segmento de recta que los une, también está en C.

Ejemplos triviales de conjuntos convexos son: V, $\emptyset$, $\{\bar{x}\}$.

Ejemplo 2.1. $\{(x_1, x_2) \colon x_1^2 + x_2^2 \leq 1\}$ es convexo. $\diamond$

Ejemplo 2.2. La bola (o esfera) cerrada de $\mathbb{R}^n$, con centro en c y radio r denotada por $\bar{B}(c,r) = B[c,r] = \{x \colon ||x - c|| \leq r\}$ y la bola abierta $B(c,r) = \{x \colon ||x - c|| < r\}$, son conjuntos convexos. Dependiendo de la norma $||\ \ ||$ utilizada, cambia la forma de la bola. pero de todas maneras es un conjunto convexo. $\diamond$

Ejemplo 2.3. Sean $\bar{x}$ y $\bar{y}$ dos puntos diferentes. La recta que pasa por $\bar{x}$ y por $\bar{y}$, es decir

$$\{\, \bar{x} + t(\bar{y} - \bar{x}) \,:\, t \in \mathbb{R}\}$$

es un conjunto convexo. Si los dos puntos son iguales, se trata de un solo punto, que también es convexo. $\diamond$

Ejemplo 2.4. Dados $\bar{x}$, $d \neq 0$ elementos de $\mathbb{R}^n$, la recta que pasa por $\bar{x}$ y es paralela a d, o sea, $R(\bar{x}, d) = \{\bar{x} + \mu d \colon \mu \in \mathbb{R}\}$ y la semirrecta que empieza en $\bar{x}$ y va en la dirección de d, o sea , $S(\bar{x}, d) = \{\bar{x} + \mu d \colon \mu \geq 0\}$, son ejemplos de conjuntos convexos. $\diamond$

Ejemplo 2.5. $C = \{(x_1, x_2) : x_2 = x_1^2\}$ no es convexo ya que $(0, 1) = \frac{1}{2}(1, 1) + \frac{1}{2}(-1, 1)$ no está en el conjunto, aunque $(1, 1)$ y $(-1, 1)$ sí están en C. $\diamond$

Ejemplo 2.6. $\{(x_1, x_2) : x_2 \geq x_1^2\}$ sí es convexo. $\diamond$

Definición 2.2. Dados $c \in \mathbb{R}^n$, $c \neq 0$, $\alpha \in \mathbb{R}$, se llama **hiperplano** al siguiente conjunto:

$$H = H_{c,\alpha} = \{x \in \mathbb{R}^n : c^{\mathrm{T}}x = \alpha\}. \tag{2.1}$$

Este hiperplano genera dos semiespacios cerrados:

$$H^+ = \{x \in \mathbb{R}^n : c^{\mathrm{T}}x \geq \alpha\}, \tag{2.2}$$
$$H^- = \{x \in \mathbb{R}^n : c^{\mathrm{T}}x \leq \alpha\}, \tag{2.3}$$

y dos semiespacios abiertos:

$$\overset{\circ}{H}{}^+ = \{x \in \mathbb{R}^n : c^{\mathrm{T}}x > \alpha\},$$
$$\overset{\circ}{H}{}^- = \{x \in \mathbb{R}^n : c^{\mathrm{T}}x < \alpha\}.$$

Ejemplo 2.7. El conjunto $\{(x_1, x_2, x_3) : 2x_1 - 3x_2 + 4x_3 = 5\}$ es un hiperplano de $\mathbb{R}^3$. El conjunto $\{(x_1, x_2, x_3) : 2x_1 - 3x_2 + 4x_3 > 5\}$ es un semiespacio abierto de $\mathbb{R}^3$. $\diamond$

En $\mathbb{R}$ un hiperplano es un punto y los semiespacios semirrectas. En $\mathbb{R}^2$ los hiperplanos son las rectas y los semiespacios los semiplanos. En $\mathbb{R}^3$ los hiperplanos son los planos.

Los conjuntos H, H^+, H^-, $\overset{\circ}{H}{}^+$, $\overset{\circ}{H}{}^-$ son convexos. Veamos que H es convexo. Sean: $x, y \in H$, $t \in [0, 1]$, $z = (1\text{-}t)x + ty$. El punto z está en H si y solamente si $c^{\mathrm{T}}z = \alpha$; efectuando el cálculo:

$$c^{\mathrm{T}}z = c^{\mathrm{T}}((1\text{-}t)x + ty) = (1\text{-}t)c^{\mathrm{T}}x + tc^{\mathrm{T}}y = (1\text{-}t)\alpha + t\alpha = \alpha,$$

luego z está en H, luego H es convexo. En esta demostración no se utilizó que $t \in [0, 1]$, entonces no sólo los puntos del segmento de recta están en H, sino que todos los puntos de la recta que pasa por x y y también están en H, es decir, H es una variedad lineal. Un conjunto L es una variedad lineal o variedad afín si dados x, y en L, t un escalar, entonces $z = z_{xyt} = (1\text{-}t)x + ty$ también está en L.

Veámos ahora que H^+ es convexo. Sean: $x, y \in H^+$, $t \in [0, 1]$, $z = (1\text{-}t)x + ty$. Entonces $c^{\mathrm{T}}x \geq \alpha$, $c^{\mathrm{T}}y \geq \alpha$, t, $1\text{-}t \geq 0$.

$$c^{\mathrm{T}}z = c^{\mathrm{T}}((1\text{-}t)x + ty) = (1\text{-}t)c^{\mathrm{T}}x + tc^{\mathrm{T}}y \geq (1\text{-}t)\alpha + t\alpha = \alpha,$$

entonces H^+ también es convexo, y de manera semejante se comprueba que H^-, $\overset{\circ}{H}{}^+$, $\overset{\circ}{H}{}^-$ son convexos.

Proposición 2.1. *La intersección de dos conjuntos convexos es un convexo.*

La demostración es muy sencilla. Sean: C, D convexos, $x, y \in C \cap D$, $t \in [0, 1]$, $z = (1\text{-}t)x + ty$. Como C es convexo entonces $z \in C$. Como D es convexo $z \in D$. Luego $z \in C \cap D$.

Proposición 2.2. *La intersección de cualquier familia de conjuntos convexos es un convexo, independientemente de que la familia sea finita, infinita, enumerable o no enumerable. Dicho de otra forma, sea $\{C_i\}_{i \in I}$ una familia de conjuntos convexos, entonces*

$$\bigcap_{i \in I} C_i \ \textit{es un conjunto convexo.}$$

En cambio, no se puede afirmar que la unión de dos convexos sea siempre un convexo. Por ejemplo, en $\mathbb{R}^2$ las bolas $B((0,0),1)$, $B((2,0),1)$ son conjuntos convexos, pero su unión no es un convexo.

Ejemplo 2.8. Las restricciones de un problema de programación lineal son igualdades, es decir representan hiperplanos, o bien, son desigualdades y en este caso representan semiespacios. Así cualquier conjunto admisible de un problema de programación lineal es simplemente la intersección de hiperplanos y semiespacios, luego es un conjunto convexo. En particular, dada una matriz real A de tamaño $m \times n$, los siguientes conjuntos son convexos.

$$\begin{aligned}
&\{x \in \mathbb{R}^n : Ax = b\},\\
&\{x \in \mathbb{R}^n : Ax \geq b\},\\
&\{x \in \mathbb{R}^n : Ax = b,\ x \geq 0\},\\
&\{x \in \mathbb{R}^n : Ax \geq b,\ x \geq 0\}. \ \diamond
\end{aligned}$$

Ejemplo 2.9.

$$\bigcap_{0 \leq \theta \leq 2\pi} \{(x_1, x_2) \colon x_1 \cos\theta + x_2 \text{ sen } \theta \leq 1\},$$

es decir, la bola (para la norma $\|\ \|_2$) con centro en $\mathbf{0} = (0,0)$ y radio 1, denotada $B_2[\mathbf{0},1]$, es un conjunto convexo, puesto que es intersección de semiespacios. $\diamond$

Definición 2.3. Se llama un **polítopo** a cualquier conjunto que se pueda expresar como la intersección de un número finito de semiespacios cerrados (o de semiespacios cerrados e hiperplanos).

Directamente de la definición se puede concluir que un polítopo es un conjunto convexo y cerrado.

Definición 2.4. Un **poliedro** es un polítopo acotado.

Ejemplo 2.10. El conjunto admisible de cualquier problema de programación lineal es un polítopo. $\diamond$

Ejemplo 2.11. El subconjunto de $\mathbb{R}^2$ definido por las siguientes restricciones es un poliedro.

$$\begin{aligned}
-x_1 + x_2 &\geq 2\\
x_2 &\geq 3\\
x_2 &\leq 5\\
x &\geq 0. \ \diamond
\end{aligned}$$

Ejemplo 2.12. $B_2[\mathbf{0},1] = \bigcap_{0 \leq \theta \leq 2\pi} \{(x_1, x_2) \colon x_1 \cos\theta + x_2 \ \text{sen}\,\theta \leq 1\}$ no es un polítopo. $\diamond$

Proposición 2.3. *Si C es un conjunto convexo y α un número, entonces*

$$\alpha C = \{\alpha x \colon x \in C\} \tag{2.4}$$

es un conjunto convexo.

Proposición 2.4. *Si C y D son conjuntos convexos, entonces*

$$C + D = \{x + y \colon x \in C, y \in D\} \tag{2.5}$$

es un conjunto convexo.

Corolario 2.1. *Si C, D son conjuntos convexos y α, β son números, entonces*

$$\alpha C + \beta D = \{\alpha x + \beta y \colon x \in C, y \in D\}$$

es un conjunto convexo. En particular $C - D = \{x - y \colon x \in C, y \in D\}$ también es convexo.

Ejemplo 2.13. Partiendo de que $B((0,0),1)$ es un conjunto convexo, se puede afirmar que

$$B((2,-3),4) = \{(2,-3)\} + 4\ B((0,0),1)$$

es también un conjunto convexo. $\diamond$

Proposición 2.5. *Si $C \subseteq \mathbb{R}^m$, $D \subseteq \mathbb{R}^p$ son convexos, entonces $C \times D \subseteq \mathbb{R}^{m+p}$, también es convexo.*

Ejemplo 2.14. El intervalo $[2,3]$ es convexo. También son convexos $B[0,2] \subseteq \mathbb{R}^2$ y $\mathbb{R}^2$. Luego el conjunto $\{(x_1,x_2,x_3)\colon 2 \le x_1 \le 3,\ x_2^2 + x_3^2 \le 4\}$ y el conjunto $\{(x_1,x_2,x_3)\colon 2 \le x_1 \le 3\}$ también son convexos. $\diamond$

Definición 2.5. Se llama **combinación convexa** de x^1, x^2, ... , x^m elementos de V a una combinación lineal en la que todos los escalares son no negativos y además su suma es uno, es decir:

$$x = t_1x^1 + t_2x^2 + \ldots + t_mx^m, \quad t_i \ge 0\ \forall i, \quad \sum_{i=1}^{m} t_i = 1. \tag{2.6}$$

Si todos los escalares son positivos la combinación convexa se llama **estricta**. Se denotará por

$$\mathrm{cc}(A) \tag{2.7}$$

el **conjunto de todas las combinaciones convexas** de elementos de A, es decir, el conjunto de todas las combinaciones convexas de subconjuntos finitos de A.

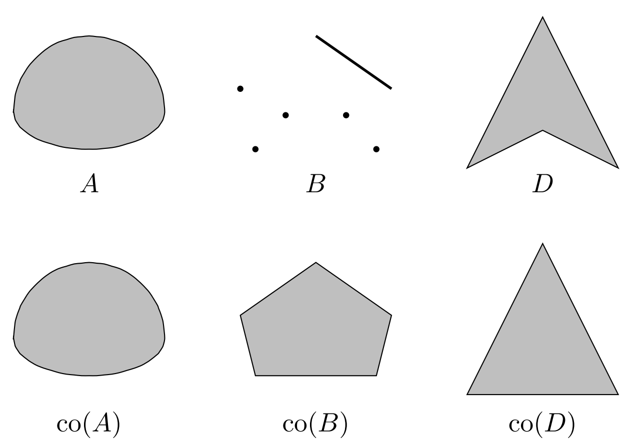
La combinación convexa es la generalización de la expresión (1-t)$x + ty$ con t en el intervalo $[0,1]$.

Ejemplo 2.15. Dados $(1,0)$, $(0,0)$ y $(0,1)$, son ejemplos de combinaciones convexas:

$$\begin{aligned} (\frac{1}{2},\frac{1}{4}) &= \frac{1}{2}(1,0) + \frac{1}{4}(0,0) + \frac{1}{4}(0,1) \\ (0,1) &= 0(1,0) + 0(0,0) + 1(0,1). \quad \diamond \end{aligned}$$

Definición 2.6. Sea A un subconjunto de V. Se llama **envolvente convexa** de A, o **convexo generado** por A, o **casco convexo** de A, denotado $\mathrm{co}(A)$, al conjunto convexo más pequeño que contenga a A. Esto quiere decir que si C es un conjunto convexo que contiene a A , entonces necesariamente $\mathrm{co}(A)$ está contenido en C.

La anterior definición es descriptiva, pero no constructiva.



Proposición 2.6. *El convexo generado por A se puede caracterizar "constructivamente" como la intersección de todos los convexos que contienen a A,*

$$co\,(A) = \bigcap_{\substack{C\ convexo,\\ A \subseteq C}} C. \tag{2.8}$$

Esta intersección está bien definida ya que por lo menos existe un conjunto convexo que contiene a A: el espacio completo $\mathbb{R}^n$.

Proposición 2.7. $co(A) = cc(A)$.

Ejemplo 2.16. $\text{co}\left(\{(1,0),(0,1),(0,0)\}\right) = \{(x_1,x_2) : x_1 + x_2 \leq 1,\ x \geq 0\}$,

$\text{co}\left(\{(1,0),(0,1),(0,0),(0.1,0.2)\}\right) = \{(x_1,x_2) : x_1 + x_2 \leq 1, x \geq 0\}$,

$\text{co}\left(\{(x_1,x_2) : x_1x_2 = 0,\ x_1^2 + x_2^2 \leq 1,\ x \geq 0\}\right) = \{(x_1,x_2) : x_1 + x_2 \leq 1,\ x \geq 0\}$,

$\text{co}\left(\{(x_1,x_2) : x_2 = x_1^2\}\right) = \{(x_1,x_2) : x_2 \geq x_1^2\}$. $\diamond$

2.2 Funciones convexas

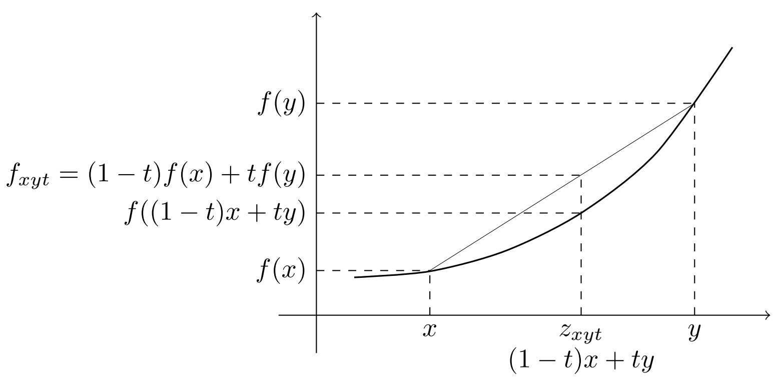
Definición 2.7. Una función f con valores reales, definida en un convexo no vacío C, es **convexa**, si para todo x, y en C y para todo t en $[0,1]$

$$f(z_{xyt}) = f((1-t)x + ty) \leq (1-t)f(x) + tf(y) = f_{xyt}.$$

La función f es **estrictamente convexa** si para todo x, y en C, $x \neq y$ y para todo t en $]0,1[$

$$f(z_{xyt}) = f((1-t)x + ty) < (1-t)f(x) + tf(y) = f_{xyt}.$$

La noción de convexidad se puede interpretar geométricamente de la siguiente manera: una función es convexa si al tomar dos puntos $(x, f(x))$, $(y, f(y))$ (de la "gráfica" de f) y unirlos por medio de un segmento de recta, éste nunca queda por debajo de la gráfica.



Definición 2.8. Una función f definida en un convexo no vacío es **cóncava** (**estrictamente cóncava**) si $-f$ es convexa (estrictamente convexa).

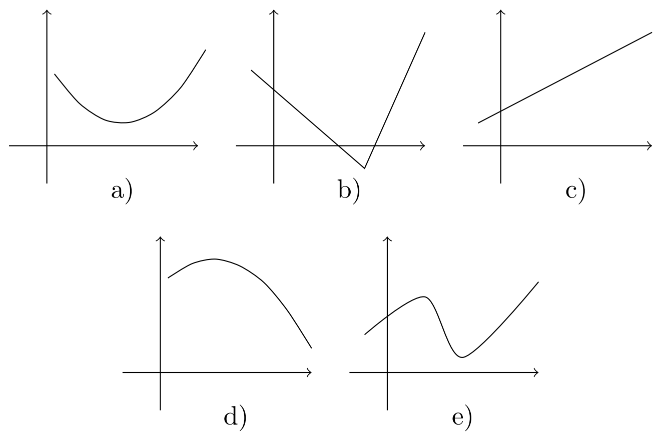
En algunos textos se estudian por aparte las caracterizaciones de funciones convexas y las de funciones cóncavas. En este documento únicamente aparecen las principales caracterizaciones para funciones convexas. Si se desea averiguar si una función es cóncava, basta con averiguar si $-f$ es convexa.

Es claro, según la definición, que toda función estrictamente convexa también es convexa. Como para $t = 0$ y para $t = 1$ siempre se tiene $f(x) \leq f(x)$ y $f(y) \leq f(y)$, entonces en la definición de convexidad se puede hacer variar t únicamente en el intervalo abierto $]0,1[$.

Ejemplo 2.17. $f : C \subseteq \mathbb{R} \longrightarrow \mathbb{R}, f(x) = x^2$. Gráficamente se "ve" que f es convexa. Por otro lado, sean x, y en C, t en el intervalo $[0,1]$,

$$\begin{aligned}
f_{xyt} - f(z_{xyt}) &= (1-t)x^2 + ty^2 - \big((1-t)x + ty\big)^2 \\
&= tx^2 - t^2x^2 + ty^2 - t^2y^2 - 2txy + 2t^2xy \\
&= x^2t(1-t) + y^2t(1-t) - 2xyt(1-t) \\
&= (x-y)^2t(1-t) \\
&\geq 0.
\end{aligned}$$

Luego $f(z_{xyt}) \leq f_{xyt}$, es decir, f es convexa. Además si $x \neq y$, entonces $(x-y)^2 > 0$, y si t está en el intervalo abierto $]0,1[$ se tiene que $t(1-t) > 0$, luego $f_{xyt} - f(z_{xyt}) > 0$, o sea, f es estrictamente convexa. $\diamond$



En la figura, las funciones en a), b) y c) son convexas; la función en a) es estrictamente convexa; las funciones en c) y d) son cóncavas; la función en d) es estrictamente cóncava; la función en e) no es ni convexa ni cóncava.

Ejemplo 2.18. Sean: $c \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$, $f : C \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $f(x) = c^{\mathrm{T}}x + \alpha$. Si $n = 2$ la gráfica corresponde a una parte de un plano.

$$\begin{aligned} f_{xyt} - f(z_{xyt}) &= (1-t)(c^{\mathrm{T}}x + \alpha) + t(c^{\mathrm{T}}y + \alpha) \\ &\quad -c^{\mathrm{T}}\big((1-t)x + ty\big) - \alpha = 0 \end{aligned}$$

Esta función es convexa, pero no estrictamente convexa, también es cóncava y no es estrictamente cóncava. $\diamond$

Ejemplo 2.19. $f : C \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}, f(x) = \|x\|$.

$$\begin{aligned} f(z_{xyt}) = f\big((1-t)x + ty\big) &= \|(1-t)x + ty\| \\ &\leq \|(1-t)x\| + \|ty\| \\ &= |1-t|\|x\| + |t|\|y\| \\ &= (1-t)f(x) + tf(y). \end{aligned}$$

Entonces la norma es una función convexa. $\diamond$

Ejemplo 2.20. Sean: $x = (0,0)$, $y = (0,1)$, $t = 1/2$, $z = z_{xyt} = (0, 1/2)$. Entonces $\|z\| = \|0.5y\| = 0.5\|y\|$. Por otro lado (1-$t$)$\|x\| + t\|y\| = 0.5\|y\|$. Luego ninguna norma $\|\|$ es una función estrictamente convexa. Esto no contradice el resultado de Análisis Funcional, según el cual la norma euclidiana es una norma estrictamente convexa. $\diamond$

Una norma es estrictamente convexa si $||x|| = ||y|| = ||x+y||/2$ implica que $x = y$. Otra definición, $||x|| < \max\{||x+y||, ||x-y||\}$ para todo $y \neq 0$. Una caracterización geométrica, la frontera de la bola unitaria (centro en el origen y radio 1) no tiene segmentos de recta. La norma euclidiana es una norma estrictamente convexa, la norma $|| \ \ ||_1$ no lo es.

Ejemplo 2.21. $f : [-2,3] \longrightarrow \mathbb{R}, f(x) = x^3$. Si $x = -1$, $y = 1$, $t = 1/4$, entonces

$$\begin{aligned} f_{xyt} - f(z_{xyt}) &= \frac{3}{4}(-1) + \frac{1}{4}(1) - \big(\frac{3}{4}(-1) + \frac{1}{4}(1)\big)^3 \\ &= -\frac{1}{2} - \frac{-1}{8} \\ &= -\frac{3}{8}. \end{aligned}$$

Luego f no es convexa. Tampoco es cóncava. Sin embargo, para esta función, al cambiar el conjunto de definición, se puede tener convexidad. $\diamond$

Ejemplo 2.22. $f : [0, 3] \longrightarrow \mathbb{R}, f(x) = x^3$ es convexa, más aún, es estrictamente convexa. $\diamond$

Definición 2.9. Sean: $f : A \longrightarrow \mathbb{R}$ una función, α un número real. Se llama un **conjunto inferior de nivel** al subconjunto de A:

$$CI(f, \alpha) = S^-(f, \alpha) = \{x \in A : f(x) \leq \alpha\}.$$

Proposición 2.8. *Sea C un convexo. Si $f : C \longrightarrow \mathbb{R}$ es convexa, entonces todos los conjuntos de nivel son convexos. En particular $\{x : f(x) \leq 0\}$ es un convexo.*

Proposición 2.9. *Sea C un conjunto convexo. Si $f : C \longrightarrow \mathbb{R}$, $g : C \longrightarrow \mathbb{R}$ son funciones convexas y α y β son números no negativos, entonces*

$$\alpha f + \beta g : C \longrightarrow \mathbb{R} \quad \textit{es convexa.}$$

En particular, αf, $f + g$ son convexas.

Obviamente también se tiene la generalización para más de dos funciones. Si $f_1, ..., f_m$ son funciones convexas definidas en un convexo C, y $\alpha_1, ...\alpha_m$ son escalares no negativos, entonces

$$\alpha_1 f_1 + ... + \alpha_m f_m : C \longrightarrow \mathbb{R} \qquad \text{es convexa.}$$

Uno esperaría que el producto de dos funciones convexas fuera una función convexa, pero no es cierto en general. Basta con considerar x, x^2 que son funciones convexas en $\mathbb{R}$, pero $f(x) = x^3$ no es convexa en $\mathbb{R}$.

Proposición 2.10. *Sea C un conjunto convexo. Si $f : C \longrightarrow \mathbb{R}$ es convexa, entonces las siguientes "ampliaciones" de f son funciones convexas:*

$$\begin{array}{lll} f_1 & : \mathbb{R}^m \times C \times \mathbb{R}^p \longrightarrow \mathbb{R}, & f_1(u, x, v) = f(x) \\ f_2 & : \mathbb{R}^m \times C \longrightarrow \mathbb{R}, & f_2(u, x) = f(x) \\ f_3 & : C \times \mathbb{R}^p \longrightarrow \mathbb{R}, & f_3(x, v) = f(x). \end{array}$$

Ejemplo 2.23. Como la función $f(x) = x^2$ es convexa, entonces $f_1(x_1, x_2) = x_1^2$, $f_2(x_1, x_2) = x_2^2$ son convexas, luego $g(x_1, x_2) = x_1^2 + x_2^2$ es una función convexa. $\diamond$

Proposición 2.11. *Sean: C un convexo no vacío, $f : C \longrightarrow \mathbb{R}$. Si f es convexa, entonces es continua en el interior de C, o sea, la continuidad en el interior de C es una condición necesaria para la convexidad.*

Ejemplo 2.24. $f : [0, 1] \longrightarrow \mathbb{R}$ definida por $f(x) = \lfloor x \rfloor$ (parte entera inferior o piso) no es continua en $[0, 1]$ (es discontinua en 1), pero es convexa en $[0, 1]$ y continua en $[0, 1[$. $\diamond$

Ejemplo 2.25.

$$f(x) = \begin{cases} x^2 & \text{si} \quad x \neq 0 \\ 1 & \text{si} \quad x = 0, \end{cases}$$

Como f no es continua en $x = 0$, entonces no es convexa. $\diamond$

Definición 2.10. Sean: $C \subseteq \mathbb{R}^n$ un convexo no vacío, $f : C \longrightarrow \mathbb{R}$. Se llama **epígrafo** de f al subconjunto de $\mathbb{R}^{n+1}$:

$$\text{epi}(f) = \{(x, t) : f(x) \leq t\}.$$

Proposición 2.12. *Sea $C \subseteq \mathbb{R}^n$ un convexo no vacío. $f : C \longrightarrow \mathbb{R}$ es convexa si y solamente si su epígrafo es un conjunto convexo.*

Ejemplo 2.26. Sean: $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = |x|$.

$$\begin{aligned}
\text{epi}(f) &= \{(x_1, x_2) : |x_1| \le x_2\} \\
&= \{(x_1, x_2) : x_1 \le x_2 \text{ si } x_1 \ge 0, -x_1 \le x_2 \text{ si } x_1 \le 0\} \\
&= \{(x_1, x_2) : x_1 - x_2 \le 0 \text{ si } x_1 \ge 0, -x_1 - x_2 \le 0 \text{ si } x_1 \le 0\} \\
&= \{(x_1, x_2) : x_1 - x_2 \le 0, -x_1 - x_2 \le 0\}
\end{aligned}$$

es convexo por ser intersección de dos semiespacios, luego f es convexa. ◇

Ejemplo 2.27. Sea $V = \mathbb{R}^2$, $f : B(\mathbf{0}, 2) \to \mathbb{R}$ definida por $f(x_1, x_2) = x_1^2 + x_2^3$.

$$\begin{aligned}
x &= (0, 0, 0) \in \text{epi}(f), \\
y &= (0, \text{-}1, \text{-}1) \in \text{epi}(f), \\
t &= 1/2, \\
z &= \frac{1}{2}\, x + \frac{1}{2}\, y = (0, \text{-}1/2, \text{-}1/2) \notin \text{epi}(f), \\
& \quad \text{ya que } f(0, \text{-}1/2) = \text{-}1/8 \not\le \text{-}1/2.
\end{aligned}$$

Entonces epi(f) no es convexo, luego f no es convexa. ◇

Proposición 2.13. *Sean: C un convexo no vacío, $f : C \longrightarrow \mathbb{R}$ convexa, $\varphi : f(C) \longrightarrow \mathbb{R}$ convexa y creciente ($u \le v \Rightarrow \varphi(u) \le \varphi(v)$). Entonces $g = \varphi \circ f : C \longrightarrow \mathbb{R}$, definida por $g(x) = \varphi(f(x))$, es convexa.*

Ejemplo 2.28. La función $g(x_1, x_2) = (2x_1 + 3x_2 + 4)^2$ definida para $x \ge 0$ es convexa, pues $f(x_1, x_2) = 2x_1 + 3x_2 + 4$ es convexa y $\varphi(t) = t^2$ es convexa y creciente en $[16, \infty[$. La función g es convexa en todo $\mathbb{R}^2$, pero esto no se deduce de la proposición anterior. ◇

Ejemplo 2.29. La función $g(x_1, x_2) = \exp(2x_1 + 3x_2 + 4) = e^{2x_1+3x_2+4}$, definida en todo $\mathbb{R}^2$, es convexa pues $f(x_1, x_2) = 2x_1 + 3x_2 + 4$ es convexa y $\varphi(t) = \exp(t)$ es convexa y creciente en $\mathbb{R}$. ◇

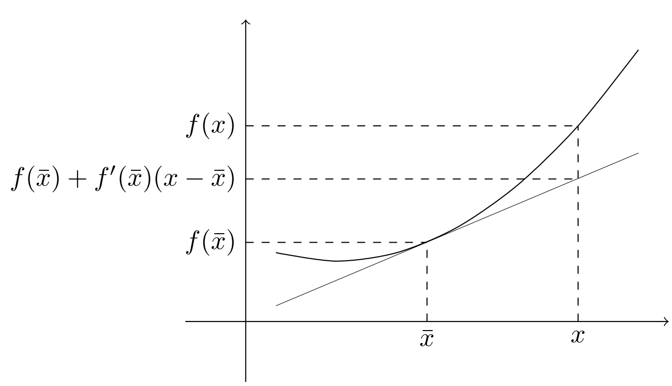
Proposición 2.14. *Sean: $C \subseteq \mathbb{R}^n$ un convexo abierto y no vacío, $f : C \longrightarrow \mathbb{R}$ diferenciable. Entonces f es convexa si y solamente si para todo $x, \bar{x} \in C$*

$$f(x) - f(\bar{x}) \ge f'(\bar{x})^{\mathrm{T}}(x - \bar{x}). \tag{2.9}$$

La anterior desigualdad se puede presentar así:

$$f(x) \ge f(\bar{x}) + f'(\bar{x})^{\mathrm{T}}(x - \bar{x}). \tag{2.10}$$

Si $n = 1$ la expresión del lado derecho representa la recta tangente a la curva en el punto $(\bar{x}, f(\bar{x}))$, o sea, una función derivable es convexa si y solamente si la curva queda por encima de todas las rectas tangentes.



En general, la expresión del lado derecho representa la aproximación lineal de la función en $\bar{x}$, o sea, una función diferenciable es convexa si y solamente si cualquier aproximación lineal subevalúa la función.

Algunas veces para demostrar que una función es convexa, puede ser más sencillo mostrar que

$$f(x) - f(\bar{x}) - f'(\bar{x})^{\mathrm{T}}(x - \bar{x}) \ge 0. \tag{2.11}$$

Ejemplo 2.30. Sea $f(x_1,x_2)=x_1^2+x_2^2$ definida en todo $\mathbb{R}^2$. Como f es diferenciable y además

$$\begin{aligned} f(x)-f(\bar x) &- f'(\bar x)^{\mathrm{T}}(x-\bar x) = \\ &= x_1^2+x_2^2-\bar x_1^2-\bar x_2^2-\begin{bmatrix} 2\bar x_1 & 2\bar x_2 \end{bmatrix}\begin{bmatrix} x_1\text{-}\bar x_1 \\ x_2\text{-}\bar x_2 \end{bmatrix} \\ &= x_1^2+x_2^2-\bar x_1^2-\bar x_2^2-2\bar x_1x_1+2\bar x_1^2-2\bar x_2x_2+2\bar x_2^2 \\ &= x_1^2+x_2^2+\bar x_1^2+\bar x_2^2-2\bar x_1x_1-2\bar x_2x_2 \\ &= (x_1-\bar x_1)^2+(x_2-\bar x_2)^2 \\ &\geq 0. \end{aligned}$$

Entonces f es convexa. ◇

Proposición 2.15. *Sean: $C\subseteq\mathbb{R}^n$ un convexo abierto y no vacío, $f:C\longrightarrow\mathbb{R}$ diferenciable. Entonces f es convexa si y solamente si*

$$\big(f'(y)-f'(x)\big)^{T}(y-x)\geq 0 \qquad \textit{para todo } x,y\in C.$$

Si $n=1$ el gradiente es simplemente la derivada, y así la anterior desigualdad está indicando que $f'(x)$ es creciente.

Ejemplo 2.31. Sea $f(x_1,x_2)=x_1^2+x_2^2$ definida en todo $\mathbb{R}^2$. Como f es diferenciable y además

$$\begin{aligned} \big(f'(y)-f'(x)\big)^{\mathrm{T}}(y-x) &= \big(f'(y)^{\mathrm{T}}-f'(x)^{\mathrm{T}}\big)(y-x) \\ &= \big([2y_1 \quad 2y_2]-[2x_1 \quad 2x_2]\big)\begin{bmatrix} y_1\text{-}x_1 \\ y_2\text{-}x_2 \end{bmatrix} \\ &= [2(y_1-x_1) \quad 2(y_2-x_2)]\begin{bmatrix} y_1\text{-}x_1 \\ y_2\text{-}x_2 \end{bmatrix} \\ &= 2(y_1-x_1)^2+2(y_2-x_2)^2 \\ &\geq 0. \ \diamond \end{aligned}$$

Entonces f es convexa.

Ejemplo 2.32. Sea $f(x_1,x_2)=x_1^3+x_2^2$ definida en todo $\mathbb{R}^2$. Como f es diferenciable y además

$$\begin{aligned} \big(f'(y)-f'(x)\big)^{\mathrm{T}}(y-x) &= \big([3y_1^2 \quad 2y_2]-[3x_1^2 \quad 2x_2]\big)\begin{bmatrix} y_1\text{-}x_1 \\ y_2\text{-}x_2 \end{bmatrix} \\ &= [3(y_1^2-x_1^2) \quad 2(y_2-x_2)]\begin{bmatrix} y_1\text{-}x_1 \\ y_2\text{-}x_2 \end{bmatrix} \\ &= 3(y_1+x_1)(y_1-x_1)^2+2(y_2-x_2)^2 \\ &= -3, \quad \text{si } x=(-1,0),\ y=(0,0). \end{aligned}$$

Entonces f no es convexa. ◇

Ejemplo 2.33. Sea $f(x_1,x_2)=x_1^3+x_2^2$ definida en $C=\{(x_1,x_2): x_1\geq 0\}$. Como f es diferenciable y además

$$\begin{aligned} \big(f'(y)-f'(x)\big)^{\mathrm{T}}(y-x) &= 3(y_1+x_1)(y_1-x_1)^2+2(y_2-x_2)^2 \\ &\geq 0, \end{aligned}$$

entonces, según la última proposición, f es convexa en el interior de C. Faltaría por ver que también es convexa en todo C. ◇

Definición 2.11. Sea $A\subseteq\mathbb{R}^n$ un conjunto abierto. $F:A\longrightarrow\mathbb{R}^n$ se llama **monótona** si

$$\big(F(y)-F(x)\big)^{\mathrm{T}}(y-x)\geq 0 \qquad \text{para todo } x,y\in A.$$

La última proposición se puede enunciar así: una función diferenciable f, definida en un convexo abierto, es convexa si y solamente si su gradiente es monótono.

Una función de una sola variable (definida en un intervalo de $\mathbb{R}$) es **monótona** si es creciente en todo el conjunto de definición o si es decreciente en todo el conjunto de definición.

Proposición 2.16. *Sean: C un convexo abierto y no vacío, $f : C \to \mathbb{R}$ doblemente diferenciable. Entonces f es convexa si y solamente si la matriz hessiana $H_f(x) = f''(x) = \left[\dfrac{\partial^2 f}{\partial x_j \partial x_i}(x)\right]$ es semidefinida positiva en todo punto x de C.*

Recuérdese que para funciones doblemente diferenciables se cumple que $\dfrac{\partial^2 f}{\partial x_j \partial x_i} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}$, luego la matriz hessiana es simétrica.

Ejemplo 2.34. Sea $f(x_1, x_2) = x_1^2 + x_2^2$ definida en todo $\mathbb{R}^2$. Como f es doblemente diferenciable y además

$$f''(x) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

es definida positiva, luego semidefinida positiva, entonces f es convexa. $\diamond$

Una característica de las funciones cuadráticas, de las afines (lineal más una constante) y de las constantes es que su matriz hessiana es constante, es decir, no depende del punto donde se evalúe.

Proposición 2.17. *Sean: C un convexo abierto no vacío, $f : C \longrightarrow \mathbb{R}$ doblemente diferenciable. Si la matriz hessiana $f''(x)$ es definida positiva en todo punto x de C, entonces f es estrictamente convexa.*

Ejemplo 2.35. Sea $f(x_1, x_2) = x_1^2 + x_2^2$ definida en todo $\mathbb{R}^2$. Como f es doblemente diferenciable y además

$$f''(x) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

es definida positiva, entonces f es estrictamente convexa. $\diamond$

Ejemplo 2.36. Sea $f(x_1, x_2) = x_1^4 + x_2^2$ definida en todo $\mathbb{R}^2$. Como f es doblemente diferenciable y además

$$f''(x) = \begin{bmatrix} 12x_1^2 & 0 \\ 0 & 2 \end{bmatrix}$$

es semidefinida positiva para todo x, entonces f es convexa. La matriz hessiana no siempre es definida positiva, por ejemplo para $x_1 = 0$, luego a partir de la última proposición no se puede afirmar que f sea estrictamente convexa, sin embargo, sí lo es. $\diamond$

Ejemplo 2.37. Sea $f : B[(3,2), 1] \to \mathbb{R}$ definida por $f(x_1, x_2) = x_1^3 + x_2^2$. La función f es doblemente diferenciable, además

$$f''(x) = \begin{bmatrix} 6x_1 & 0 \\ 0 & 2 \end{bmatrix}$$

es semidefinida positiva siempre que $x_1 \geq 0$, en particular, en el conjunto abierto definido por la restricción $x_1 > 0$, que contiene la bola $B[(3,2), 1]$, entonces f es convexa, más aún, es estrictamente convexa. $\diamond$

Ejemplo 2.38. Sea $f(x_1, x_2) = x_1^2 x_2^2$ definida en todo $\mathbb{R}^2$. La función f es doblemente diferenciable,

$$f''(x) = \begin{bmatrix} 2x_2^2 & 4x_1x_2 \\ 4x_1x_2 & 2x_1^2 \end{bmatrix}$$

no es semidefinida positiva ya que $\delta_2 = -12x_1^2x_2^2$, entonces f no es convexa. $\diamond$

Ejemplo 2.39. Sea $f(x_1, x_2) = 5(3x_1 - 4x_2 - 2)^2 + 6(x_1 - 2x_2)^2$ definida en todo $\mathbb{R}^2$. Como f es doblemente diferenciable y además

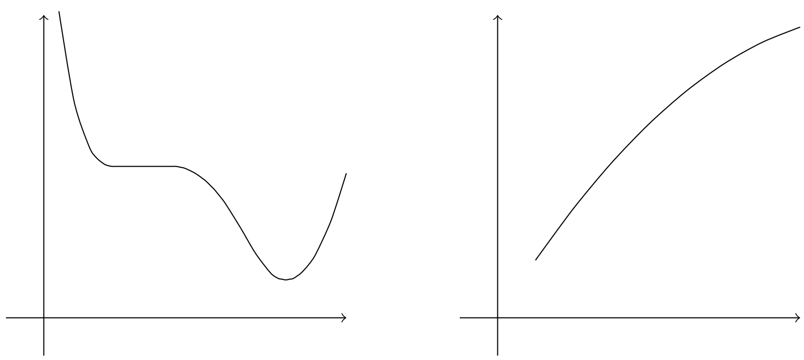
$$f''(x) = \begin{bmatrix} 102 & \text{-144} \\ \text{-144} & 208 \end{bmatrix}, \quad \delta_1 = 102, \quad \delta_2 = 480.$$

Entonces la matriz hessiana es definida positiva y la función f es estrictamente convexa. $\diamond$

2.3 Generalizaciones de funciones convexas

Definición 2.12. Una función f con valores reales, definida en un convexo no vacío C, es **cuasiconvexa**, si para todo $x, y \in C$ y para todo t en el intervalo $[0, 1]$

$$f(z_{xyt}) = f((1-t)x + ty) \leq \max\{f(x), f(y)\}.$$



Definición 2.13. Una función f con valores reales, definida en un convexo no vacío C, es **cuasicóncava**, si $-f$ es cuasiconvexa.

Proposición 2.18. *Toda función convexa es cuasiconvexa.*

Ejemplo 2.40. $f(x_1, x_2) = x_1^2 + x_2^2$ es convexa, luego es cuasiconvexa. $\diamond$

Proposición 2.19. *Sea $C \subseteq \mathbb{R}$ convexo, $f : C \longrightarrow \mathbb{R}$. Cada una de las siguientes propiedades es una condición suficiente para que f sea cuasiconexa.*

- *f es creciente en C*
- *f es decreciente en C.*
- *Existe $\alpha \in C$ tal que f es decreciente en $C \cap]-\infty, \alpha]$ y f es creciente $C \cap [\alpha, +\infty[$*

Ejemplo 2.41. $f(x) = x^3$ siempre es creciente, luego es cuasiconvexa. $\diamond$

Ejemplo 2.42. $f(x) = x^2$ definida en todo $\mathbb{R}$ no es monótona, sin embargo, sí es cuasiconvexa. $\diamond$

Proposición 2.20. *Sean: $C \subseteq \mathbb{R}^n$ convexo, $f : C \longrightarrow \mathbb{R}$. f es cuasiconvexa si y solamente si para todo real α, el conjunto de nivel $CI(f, \alpha)$ es convexo.*

Obviamente la anterior caracterización de funciones cuasiconvexas es una condición necesaria para funciones convexas. Es decir, si f es convexa, entonces todos sus conjuntos inferiores de nivel son conjuntos convexos.

Ejemplo 2.43. $f(x_1, x_2) = (x_1 + 2x_2 - 3)^3 + 4$ no es convexa en $\mathbb{R}^2$, ya que

$$f''(x) = \begin{bmatrix} 6(x_1 + 2x_2 - 3) & 12(x_1 + 2x_2 - 3) \\ 12(x_1 + 2x_2 - 3) & 24(x_1 + 2x_2 - 3) \end{bmatrix}$$

y para $x = (0,0)$, $\delta_1 = -18$, luego la matriz hessiana no es semidefinida positiva. Sin embargo,

$$\begin{aligned} CI(f,\alpha) &= \{(x_1,x_2) : (x_1+2x_2-3)^3+4 \leq \alpha\} \\ &= \{(x_1,x_2) : (x_1+2x_2-3)^3 \leq \alpha - 4\} \\ &= \{(x_1,x_2) : x_1+2x_2-3 \leq \sqrt[3]{\alpha-4}\} \\ &= \{(x_1,x_2) : x_1+2x_2 \leq \sqrt[3]{\alpha-4}+3\} \end{aligned}$$

es convexo (es un semiespacio) para todo α, luego f es cuasiconvexa en todo $\mathbb{R}^2$. $\diamond$

Ejemplo 2.44. $f : (2,\infty) \longrightarrow \mathbb{R}$, $f(x) = \sqrt{x}$ no es convexa, sin embargo,

$$CI(f,\alpha) = \begin{cases} \emptyset & \text{si } \ \alpha < \sqrt{2} \\ [2,\alpha^2] & \text{si } \ \alpha \geq \sqrt{2}, \end{cases}$$

es convexo para todo α, entonces f es cuasiconvexa. Por otro lado, la función es estrictamente creciente, luego es creciente y, por lo tanto, es cuasiconvexa. $\diamond$

Ejemplo 2.45. $f(x_1,x_2) = x_1^2x_2^2$.

$$\begin{aligned} CI(f,0) &= \{(x_1,x_2) : x_1^2x_2^2 \leq 0\} \\ &= \{(x_1,x_2) : x_1^2x_2^2 = 0\} \\ &= \{(x_1,x_2) : x_1x_2 = 0\} \\ &= \{(x_1,x_2) : x_1 = 0, \text{ o }, x_2 = 0\} \\ &= \{ \text{ puntos sobre los ejes } \} \end{aligned}$$

no es un conjunto convexo, entonces f no es cuasiconvexa y mucho menos convexa. $\diamond$

Ejemplo 2.46. Sea $f : [-2,5[\longrightarrow \mathbb{R}$ definida por $f(x) = -|x|$. f no es convexa. Además, si $x = -2$, $y = 3$, $t = 0.2$, entonces

$$\begin{aligned} \max\{f(x),f(y)\}\text{-}f(z_{xyt}) &= \max\{f(x),f(y)\}\text{-}f((1-t)x+ty) \\ &= \max\{\text{-}2,\text{-}3\} - f(\text{-}1) \\ &= -2 - -1 \\ &= -1. \end{aligned}$$

Luego f no es cuasiconvexa. $\diamond$

Proposición 2.21. *Sea $C \subseteq \mathbb{R}^n$, $C \neq \emptyset$, $f : C \to \mathbb{R}$ cuasiconvexa, $f(C) \subseteq D$, $g : D \to \mathbb{R}$ creciente. Entonces $g \circ f$ es cuasiconvexa.*

Ejemplo 2.47.

$$\begin{aligned} & f : \mathbb{R}^2 \to \mathbb{R} \\ & f(x_1,x_2) = 3x_1+4x_2+5 \text{ convexa y cuasiconvexa} \\ & g : \mathbb{R} \to \mathbb{R} \\ & g(t) = t^3 \text{ creciente} \end{aligned}$$

Entonces

$$h(x) = (g \circ f)(x) = (3x_1+4x_2+5)^3 \text{ es cuasiconvexa.}$$

Proposición 2.22. *Sean: C un convexo abierto y no vacío, $f : C \longrightarrow \mathbb{R}$ una función diferenciable. Entonces f es cuasiconvexa si y solamente si para todo $x,y \in C$:*

$$f(y) \leq f(x) \ \Rightarrow \ f'(x)^{\mathrm{T}}(y-x) \leq 0. \tag{2.12}$$

La anterior implicación es obviamente equivalente a

$$f'(x)^{\mathrm{T}}(y-x) > 0 \ \Rightarrow \ f(y) > f(x)\,. \tag{2.13}$$

Ejemplo 2.48. $f(x) = x^3$ es cuasiconvexa ya que: $f(y) \leq f(x)$ implica que $y \leq x$, o sea, $y - x \leq 0$ entonces $\alpha(y - x) \leq 0$ para $\alpha \geq 0$. En particular $3x^2(y - x) \leq 0$. ◇

Ejemplo 2.49. $f(x_1, x_2) = (x_1 + x_2)^3$ es cuasiconvexa ya que: $f'(x)^{\mathrm{T}}(y - x) > 0$ es equivalente a $3(x_1 + x_2)^2((y_1 + y_2) - (x_1 + x_2)) > 0$, lo que implica que $y_1 + y_2 > x_1 + x_2$, luego $f(y) > f(x)$. ◇

Ejemplo 2.50. $f(x_1, x_2) = x_1x_2$ no es cuasiconvexa en $\mathbb{R}^2$ ya que para $x = (1, 0)$, $y = (-1, 1)$, se cumple $f(y) = -1 \leq f(x) = 0$, pero $f'(x)^{\mathrm{T}}(y - x) = [01][-21]^{\mathrm{T}} = 1 \not\leq 0$. ◇

Definición 2.14. Sean: C un conjunto convexo, $f : C \longrightarrow \mathbb{R}$. Se dice que f es **estrictamente cuasiconvexa** si para todo $x \neq y \in C$ y para todo t en el intervalo $]0, 1[$

$$f(z_{xyt}) = f((1 - t)x + ty) < \max\{f(x), f(y)\}. \tag{2.14}$$

Definición 2.15. Sean: C un conjunto convexo, $f : C \longrightarrow \mathbb{R}$. Se dice que f es **semiestrictamente cuasiconvexa** si para todo $x, y \in C$ con $f(x) \neq f(y)$ y para todo t en el intervalo $]0, 1[$

$$f(z_{xyt}) = f((1 - t)x + ty) < \max\{f(x), f(y)\}. \tag{2.15}$$

Estas dos últimas definiciones son las de Avriel et al. 1988, Barten y Böm 1982, Frenk y Kassay 2005. En otros libros, como Avriel 1976, Bazaraa 2006 y Escobar 2005, los nombres son diferentes, la primera se llama fuertemente cuasiconvexa y la segunda estrictamente cuasiconvexa.

Ejemplo 2.51. Sea f definida en todos los reales por

$$f(x) = \begin{cases} 1 & \text{si} \quad x \neq 0 \\ 0 & \text{si} \quad x = 0. \end{cases}$$

f no es convexa, es cuasiconvexa, pero no es estrictamente cuasiconvexa ni semiestrictamente cuasiconvexa. ◇

Ejemplo 2.52. Sea f definida en todos los reales por

$$f(x) = \begin{cases} 1 & \text{si} \quad x = 0 \\ 0 & \text{si} \quad x \neq 0. \end{cases}$$

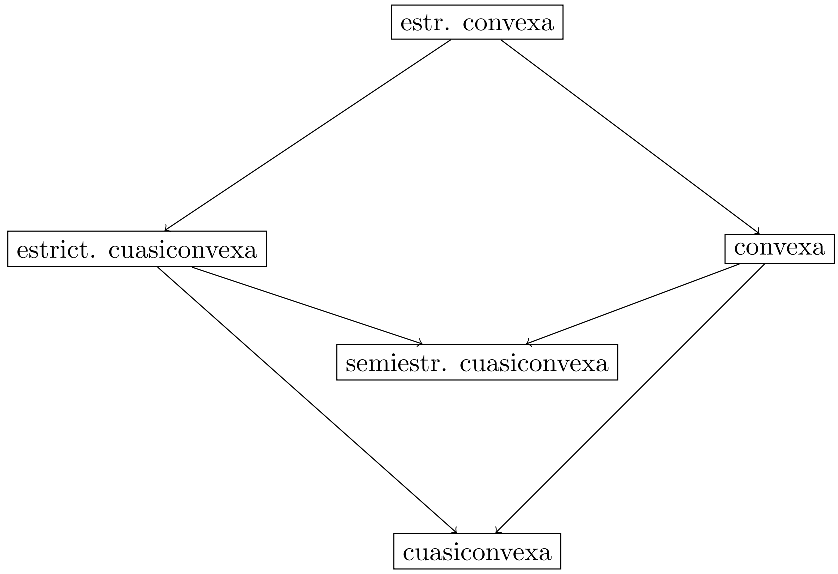
f no es convexa, no es estrictamente cuasiconvexa, es semiestrictamente cuasiconvexa, pero no es cuasiconvexa ya que el conjunto inferior de nivel $CI(f, 0) = \mathbb{R} \setminus \{0\}$ no es convexo. ◇

Ejemplo 2.53. $f(x) = 1 - \exp(\text{-}x^2)$ (su gráfica se parece a una campana de Gauss invertida) no es convexa, es cuasiconvexa, estrictamente cuasiconvexa y semiestrictamente cuasiconvexa. ◇

Ejemplo 2.54. $f(x) = \max\{0, x^2\text{-}1\}$ no es estrictamente convexa, es convexa, no es estrictamente cuasiconvexa, sí es cuasiconvexa y semiestrictamente cuasiconvexa. ◇

Proposición 2.23.

$$\begin{array}{lcl}
\textit{convexidad estricta} & \Longrightarrow & \textit{cuasiconvexidad estricta.} \\
\textit{convexidad} & \Longrightarrow & \textit{cuasiconvexidad semiestricta} \\
\textit{convexidad} & \Longrightarrow & \textit{cuasiconvexidad} \\
\textit{cuasiconvexidad estricta} & \Longrightarrow & \textit{cuasiconvexidad semiestricta} \\
\textit{cuasiconvexidad estricta} & \Longrightarrow & \textit{cuasiconvexidad .}
\end{array}$$



Definición 2.16. Sea $f : A \longrightarrow \mathbb{R}$ doblemente diferenciable en un punto $\bar{x} \in A$. Se llama **hessiano orlado (bordered)** de f en el punto $\bar{x}$ a la matriz de tamaño $(n+1) \times (n+1)$, simétrica

$$B = B(\bar{x}) = B(f, \bar{x}) = \begin{bmatrix} 0 & f'(\bar{x})^{\mathrm{T}} \\ & \\ f'(\bar{x}) & f''(\bar{x}) \\ & \end{bmatrix} \tag{2.16}$$

donde $f'(\bar{x}) = \nabla f(\bar{x}) \in \mathbb{R}^{n\times 1}$ es el gradiente y $f''(x) = \nabla^2 f(\bar{x}) \in \mathbb{R}^{n\times n}$ es la matriz hessiana.

Ejemplo 2.55. $f(x_1, x_2) = x_1^3 + 4x_1x_2 + 5x_2^2$,

$$B(x) = \begin{bmatrix} 0 & 3x_1^2 + 4x_2 & 4x_1 + 10x_2 \\ 3x_1^2 + 4x_2 & 6x_1 & 4 \\ 4x_1 + 10x_2 & 4 & 10 \end{bmatrix}$$

Definición 2.17. Sea M una matriz $p \times p$, su subdeterminante estrictamente principal es

$$\delta_i = \delta_i(M) = \det \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1i} \\ m_{21} & m_{22} & \cdots & m_{2i} \\ \vdots & & & \\ m_{i1} & m_{i2} & \cdots & m_{ii} \end{bmatrix}, \quad i = 1, ..., p$$

Para el hessiano orlado, sea

$$D_k(B) = \delta_{k+1}(B) = \det \begin{bmatrix} 0 & g_1 & \cdots & g_k \\ g_1 & h_{11} & \cdots & h_{1k} \\ \vdots & \vdots & \ddots & \\ g_k & h_{k1} & \cdots & h_{kk} \end{bmatrix}, \ \ k = 1, ..., n \tag{2.17}$$

donde g es el gradiente y $H = [h_{ij}]$ es la matriz hessiana de f.

Para la función f del ejemplo anterior,

$$\begin{aligned} D_1(B) = \delta_2(B) &= -(3x_1^2 + 4x_2)^2 \\ D_2(B) = \delta_3(B) &= \det(B) \\ &= -600x_1x_2^2 + 160x_2^2 - 480x_1^2x_2 + 128x_1x_2 - 90x_1^4 \end{aligned}$$

Proposición 2.24. *(en Bazaraa 2008 p. 771, Avriel 2010 p. 76) Sea f doblemente diferenciable y C un convexo sólido (interior no vacío):*

- *Si f es cuasiconvexa, entonces $D_i(B) \leq 0$, $i = 1, ..., n$.*

- *Si $D_i(B) < 0$, $i = 1, ..., n$, entonces f es cuasiconvexa.*
- *Si f es cuasicóncava, entonces $(-1)^i D_i(B) \geq 0$, $i = 1, ..., n$.*
- *Si $(-1)^i D_i(B) > 0$, $i = 1, ..., n$, entonces f es cuasicóncava.*

Ejemplo 2.56. Sea $f(x_1, x_2) = x_1^3 + x_2^2$ defininida en todo $\mathbb{R}^2$.

$$
\begin{aligned}
B(x) &= \begin{bmatrix} 0 & 3x_1^2 & 2x_2 \\ 3x_1^2 & 6x_1 & 0 \\ 2x_2 & 0 & 2 \end{bmatrix}, \\
D_1(B) &= \det \begin{bmatrix} 0 & 3x_1^2 \\ 3x_1^2 & 6x_1 \end{bmatrix} = -9x_1^4, \\
D_2(B) &= \det(B) = -18x_1^4 - 24x_1x_2^2.
\end{aligned}
$$

Si $x_1 = -1$ y $x_2 = 1$, entonces $\det(B_2(x)) = 6$, luego f no es cuasiconvexa. También se puede ver f no es cuasiconvexa aplicando directamente la definición. Si $x = (-5, 5)$, $y = (1, 0)$, $t = 0.6$, entonces $z = (-1.4, 2)$, $f(x) = -100$, $f(y) = 1$, $f(z) = 1.256$. $\diamond$

Este ejemplo nos permite observar que **la suma de funciones cuasiconvexas no es necesariamente cuasiconvexa**. En efecto, sean

$$
\begin{aligned}
g(x_1, x_2) &= x_1^3 \\
h(x_1, x_2) &= x_2^2
\end{aligned}
$$

definidas en todo $\mathbb{R}^2$. La función g no es convexa, pero se puede comprobar que es cuasiconvexa (por ejemplo, por medio de los conjuntos de nivel). La función h es convexa luego es cuasiconvexa y $f(x) = g(x) + h(x)$.

Para funciones de una variable, la proposición anterior sirve poco pues para toda función doblemente diferenciable

$$
B(x) = \begin{bmatrix} 0 & f'(x) \\ f'(x) & f''(x) \end{bmatrix}
$$

y $D_1(B) = \delta_2(B) = -(f'(x))^2 \leq 0$. Si f no anula su derivada en C, entonces f es cuasiconvexa.

Ejemplo 2.57. Sea $f(x) = x^3$ definida en todo $\mathbb{R}$.

$$
\begin{aligned}
B(x) &= \begin{bmatrix} 0 & 3x^2 \\ 3x^2 & 6x \end{bmatrix}, \\
D_1(B) &= \det(B(x)) = -9x^4.
\end{aligned}
$$

Entonces la proposición apenas permite decir que posiblemente f es cuasiconvexa. Recuérdese que x^3 sí es cuasiconvexa. Por otro lado, si se considera $f(x) = -x^2$ que no es cuasiconvexa, la proposición anterior dice que $-x^2$ cumple esta condición necesaria para cuasiconvexidad. $\diamond$

Ejemplo 2.58. Sea $f(x_1, x_2) = x_1^2 - x_2^2$ definida en $\mathbb{R}^2$.

$$
\begin{aligned}
B(x) &= \begin{bmatrix} 0 & 2x_1 & -2x_2 \\ 2x_1 & 2 & 0 \\ -2x_2 & 0 & -2 \end{bmatrix} \\
D_2(B) = \delta_3(B) &= 8(x_1^2 - x_2^2) \\
&= 8 \quad \text{para } x = (1, 0).
\end{aligned}
$$

Luego la función no es cuasiconvexa. $\diamond$

Ejemplo 2.59. Sea $f(x) = \log x$ con $x \in C =]0, \infty[$. En este libro log indica el logaritmo en base e.

$$
\begin{aligned}
B(x) &= \begin{bmatrix} 0 & \dfrac{1}{x} \\ \dfrac{1}{x} & -\dfrac{1}{x^2} \end{bmatrix} \\
D_1(B) = \delta_2(B) &= -\frac{1}{x^2} < 0 \quad \forall x \in C.
\end{aligned}
$$

Luego $f(x) = \log x$ es cuasiconvexa. $\diamond$

Ejemplo 2.60. Sea $f(x) = x^2$ en $\mathbb{R}$.

$$B(x) = \begin{bmatrix} 0 & 2x \\ 2x & 2 \end{bmatrix},$$
$$D_1(B) = \delta_2(B) = -4x^2.$$

Como $\delta_2(B)$ no es siempre negativo, entonces la proposición anterior no permite garantizar que $f(x) = x^2$ sea cuasiconvexa. Sin embargo, f sí es cuasiconvexa, más aún es estrictamente convexa. ◇

EJERCICIOS

2.1. Determine si $f(x_1, x_2) = x_1^2 + 2x_1x_2 - 10x_1 + 5x_2$ es convexa, o cóncava, o ninguna de las dos. Determine en qué conjunto f es convexa.

2.2. Determine si $f(x_1, x_2) = x_1e^{-x_1-x_2}$ es convexa, o cóncava, o ninguna de las dos. Determine en qué conjunto f es convexa.

2.3. Determine si $f(x_1, x_2) = -x_1^2 - 5x_2^2 + 2x_1x_2 + 10x_1 - 10x_2$ es convexa, o cóncava, o ninguna de las dos. Determine en qué conjunto f es convexa.

2.4. Sea $C = \{(x_1, x_2) : |x_i| \leq 1, \forall i\}$. Determine si $f(x_1, x_2) = 2(x_2 - x_1^2)^2$ es convexa en C, o cóncava, o ninguna de las dos. Determine en qué conjunto f es convexa.

2.5. Sea $f(x_1, x_2) = x_1^2 + x_2^3$. ¿Es f convexa? Determine en qué conjunto f es convexa.

2.6. Sea $f(x_1, x_2) = x_1^4 + x_2^3$. ¿Es f convexa? Determine en qué conjunto f es convexa.

2.7. Sean: C un convexo, $g : C \longrightarrow \mathbb{R}$ convexa. Muestre que $\{x \in C : g(x) \leq 0\}$ es un convexo.

2.8. Sea

$$f(x) = \begin{cases} x^2 & \text{si } x < 0, \\ x^3 & \text{si } x \geq 0. \end{cases}$$

¿Es f continua? ¿Es f diferenciable? ¿Es f doblemente diferenciable? ¿Es f convexa?

2.9. Sean: C un convexo, $g : C \longrightarrow \mathbb{R}$ cóncava. Muestre que $\{x \in C : g(x) > 0\}$ es un convexo.

2.10. Sean: C un convexo, $g : C \longrightarrow \mathbb{R}$ cóncava. Muestre que $f(x) = 1/g(x)$ es convexa en $\{x \in C : g(x) > 0\}$.

2.11. Sean: $c \in \mathbb{R}^n$, $f(x) = (c^{\mathrm{T}}x)^2$. Determine si f es convexa. Determine en qué conjunto f es convexa.

2.12. Sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ continua. Determine las condiciones que debe cumplir f para que $\{x : f(x) = 0\}$ sea convexo.

2.13. Sean: C convexo, $f : C \longrightarrow \mathbb{R}$. Muestre que f es convexa si y solamente si

$$f(\frac{1}{2}x + \frac{1}{2}y) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y) \quad \forall x, y \in C$$

2.14. Sean: f_1, ..., f_m funciones convexas definidas en un convexo C; α_1, ..., $\alpha_m \geq 0$. Muestre que $f(x) = y\alpha_1 f_1(x) + ... + \alpha_m f_m(x)$ es convexa.

2.15. Sean f_1, ..., f_m funciones convexas definidas en un convexo C. Muestre que $f(x) = \max\{f_1(x), ..., f_m(x)\}$ es convexa.

2.16. Sean: $c, d \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$, $C = \{x : d^{\mathrm{T}}x + \beta > 0\}$,

$$f(x) = \frac{c^{\mathrm{T}}x + \alpha}{d^{\mathrm{T}}x + \beta}.$$

¿Es f convexa, cóncava, cuasiconvexa, cuasicóncava en C?

2.17. Sea $a = (1, 2)$, $b = (5, 2)$. Sea A el conjunto de puntos del plano cuya distancia a punto a o al punto b es menor o igual a 1. Sea B el conjunto de puntos del plano cuya distancia al punto a o al punto b es menor o igual a 3. Sea C el conjunto de puntos del plano cuya distancia al punto a y al punto b es menor o igual a 3. Dibuje los conjuntos. ¿Son convexos?

2.18. Sea $A = \{(x_1, x_2) : x_1^2 \leq x_2\}$. Halle $\operatorname{co}(A)$.

2.19. Sea $A = \{(x_1, x_2) : x_1^3 = x_2\}$. Halle $\operatorname{co}(A)$.

2.20. Sea $A = \{(x_1, x_2) : x_1^3 \leq x_2\}$. Halle $\operatorname{co}(A)$.

2.21. Sea $A = \{(x_1, x_2) : x_1 > 0,\ x_1|x_2| \geq 1\}$. ¿$A$ es cerrado? Halle $\operatorname{co}(A)$. ¿$\operatorname{co}(A)$ es cerrado? Dé condiciones suficientes para que si A es cerrado, entonces $\operatorname{co}(A)$ sea cerrado.

2.22. Sea $C = \{(x_1, x_2) : |x_1| \leq x_2\}$. Halle los puntos extremos, las direcciones y las direcciones extremas de C.

2.23. Sea $C = \{(x_1, x_2) : x_1^2 \leq x_2\}$. Halle los puntos extremos, las direcciones y las direcciones extremas de C.

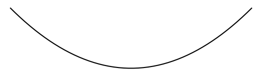
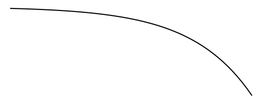
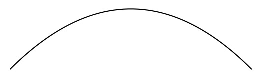
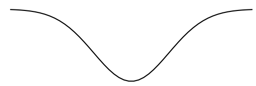
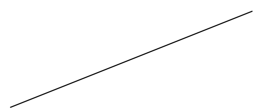
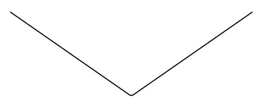
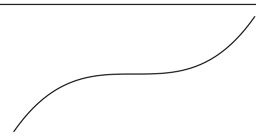
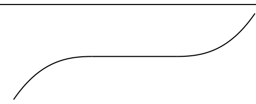
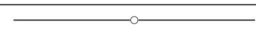
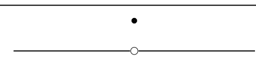
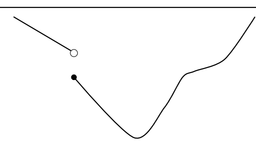
2.24. Sea $C = \{(x_1, x_2, x_3) : x_1^2 \leq x_2,\ x_1 + x_2 + x_3 \leq 1\}$. Halle los puntos extremos, las direcciones y las direcciones extremas de C.

2.25. Sean: C un convexo, A una matriz $m \times n$, $D = \{y : y = Ax, x \in C\}$. Muestre que D es un convexo.

2.26. Sean: A una matriz $m \times n$, $(m \leq n)$ de rango m, $D = \{d : Ad = 0, d \geq 0, d \neq 0\} \neq \emptyset$. Muestre que D es un cono convexo.

2.27. Sean x^1, ..., x^m elementos no nulos de $\mathbb{R}^n$. Una combinación lineal de estos elementos se llama no negativa si todos los escalares son no negativos, y se llama positiva si todos los escalares son positivos. Sea $\operatorname{cnn}(x^1, ..., x^m)$ el conjunto de todas las combinaciones lineales no negativas y $\operatorname{cp}(x^1, ..., x^m)$ el conjunto de todas las combinaciones lineales positivas. Muestre que los conjuntos $\operatorname{cnn}(x^1, ..., x^m)$ y $\operatorname{cp}(x^1, ..., x^m)$ son conos convexos.

2.28. Completar

		EC	C	EcC	sEcC	cC
1						
2		no	no	si	si	si
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						

C = convexa, E = estrictamente, c = cuasi, s = semi

Capítulo 3

Condiciones de optimalidad

Sean: $\mathcal{A}$ un subconjunto abierto de $\mathbb{R}^n$, $\;f : \mathcal{A} \longrightarrow \mathbb{R}$, $\;X \subseteq \mathcal{A}$. En este capítulo se estudiarán condiciones de optimalidad para el problema de minimizar $f(x)$ cuando x varía en X. Este problema se puede denotar

$$\begin{aligned} &\min f(x) \\ \text{sujeto a} \quad & x \in X. \end{aligned}$$

o también

$$\begin{aligned} &\min f(x) \\ \text{tal que} \quad & x \in X. \end{aligned}$$

En este documento se denotará simplemente

$$\begin{aligned} &\min f(x) \\ & x \in X. \end{aligned} \tag{PM}$$

Con frecuencia $\mathcal{A} = \mathbb{R}^n$. Generalmente el conjunto X puede ser todo $\mathbb{R}^n$ o estar definido mediante igualdades o desigualdades.

Las condiciones de optimalidad, de esta sección, se aplican a los puntos interiores de X. En secciones posteriores se estudiará el caso de X definido por desigualdades o igualdades.

Si f y X cumplen ciertas propiedades adicionales, aunque el estudio se haga en el interior de X los resultados pueden ser válidos en todo X.

Cuando $X = \mathcal{A} = \mathbb{R}^n$, se tiene un problema de minimización sin restricciones (irrestricto), que se denotará simplemente

$$\min f(x)$$

sobreentendiendo que el conjunto X es todo $\mathbb{R}^n$.

Definición 3.1. Un punto $x^* \in X$ se llama **minimizador global** o absoluto (o también punto de mínimo global) y $f^* = f(x^*)$ se llama **mínimo global** o absoluto (o valor mínimo global) de PM, si

$$f^* = f(x^*) \leq f(x) \qquad \forall x \in X.$$

Un punto $x^* \in X$ se llama **minimizador global estricto** y $f^* = f(x^*)$ se llama **mínimo global estricto** de PM, si

$$f^* = f(x^*) < f(x) \qquad \forall x \in X,\; x \neq x^*.$$

Un punto $x^* \in X$ se llama **minimizador local** o relativo y $f^* = f(x^*)$ se llama **mínimo local** o relativo de PM, si existe $r > 0$ tal que

$$f^* = f(x^*) \leq f(x) \qquad \forall x \in X \cap B(x^*, r).$$

Un punto $x^* \in X$ se llama **minimizador local estricto** y $f^* = f(x^*)$ se llama **mínimo local estricto** de PM, si existe $r > 0$ tal que

$$f^* = f(x^*) < f(x) \qquad \forall x \in X \cap B(x^*, r), x \neq x^*.$$

Un punto $x^* \in X$ se llama **minimizador global aislado** si existe una vecindad de x^* donde no hay otros minimizadores globales. Un punto $x^* \in X$ se llama **minimizador local aislado** si existe una vecindad de x^* donde no hay otros minimizadores locales.

Obsérvese que todo minimizador global también es minimizador local. Para minimizadores globales, estricto es equivalente a único, y cualquiera de estas dos características implica aislamiento. Es posible que no haya minimizadores locales o que no haya minimizadores globales.

Ejemplo 3.1. Sea $f(x) = \cos x$ en el intervalo $[-20, 20]$. El punto $x = \pi$ es minimizador local estricto y aislado, también es minimizador global aislado. El valor -1 es un mínimo local y global. $\diamond$

Ejemplo 3.2. Sea $f(x) = x^2$ en todos los reales. El punto $x = 0$ es minimizador local y global estricto y aislado. El valor 0 es un mínimo local y global. $\diamond$

Ejemplo 3.3. Sea $f(x) = x^2$ en el intervalo $[3, 9]$. El punto $x = 3$ es minimizador local y global estricto y aislado. $\diamond$

Ejemplo 3.4. Sea $f(x) = \lfloor x \rfloor$ (parte entera, también llamada parte entera inferior). El conjunto de minimizadores locales es $\mathbb{R} \setminus \mathbb{Z}$ (el conjunto de todos los reales no enteros). $\diamond$

Ejemplo 3.5. Sea $f(x) = \lceil x \rceil$ (parte entera superior). Todos los puntos son minimizadores locales. $\diamond$

Ejemplo 3.6. Sea $f(x) = x^2$ en el intervalo $]3, 9]$. Para este problema no hay minimizadores locales ni globales. $\diamond$

Ejemplo 3.7. Sea $f(x) = x^3$ en toda la recta real. Para este problema no hay minimizadores locales ni globales. $\diamond$

Ejemplo 3.8. Sea

$$f(x) = \begin{cases} x^2\big(1 + x^2 + \sin(1/x)\big) & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases}$$

El punto $x = 0$ es un minimizador local y global estricto, pero no es minimizador local aislado. $\diamond$

Si el problema planteado es de maximización, basta con multiplicar por menos uno a la función f y así un minimizador de un problema es maximizador del otro. Obviamente el valor mínimo global de un problema corresponde al valor máximo global del otro problema multiplicado por menos uno.

$$\begin{aligned} \min f(x) \\ x \in X. \end{aligned} \tag{PM}$$

$$\begin{aligned} \max g(x) = -f(x) \\ x \in X. \end{aligned} \tag{PMAX}$$

Sean:

$$\begin{aligned} X^* &: \quad \text{conjunto de minimizadores globales del PM,} \\ f^* &: \quad \text{mínimo global del PM,} \\ \overset{*}{X} &: \quad \text{conjunto de maximizadores globales del PMAX,} \\ g^* &: \quad \text{máximo global del PMAX.} \end{aligned}$$

Entonces

$$\begin{aligned} X^* &= \overset{*}{X}, \\ f^* &= -g^*. \end{aligned}$$

Los resultados anteriores también son ciertos si se cambia la palabra global por la palabra local. El siguiente enunciado es simplemente una manera de decir algo obvio con otras palabras.

Proposición 3.1. *El PM tiene minimizador global si y solamente si el conjunto de imágenes $f(X)$ tiene mínimo.*

La "proposición" anterior, utilizada junto con condiciones suficientes para que $f(X)$ tenga mínimo, permite garantizar la existencia de un minimizador global para el PM.

Proposición 3.2. *Las siguientes son algunas de las condiciones suficientes para que exista minimizador global:*

- $f(X)$ *es cerrado y acotado.*
- $f(X)$ *es cerrado y acotado inferiormente.*
- f *es continua,* X *es cerrado y acotado.*

Ejemplo 3.9. Sea $f(x_1, x_2) = \operatorname{sen} x_1 + \cos x_2$ en todo $\mathbb{R}^2$. El conjunto imagen es simplemente $f(X) = [-2, 2]$ cerrado y acotado, luego existe por lo menos un minimizador global. $\diamond$

Ejemplo 3.10. Sea $f(x_1, x_2) = x_1^2 + x_2^3$ en $X = \{x : x \geq 0\}$. El conjunto imagen es simplemente $f(X) = [0, \infty[$ cerrado y acotado inferiormente, luego existe por lo menos un minimizador global. $\diamond$

Ejemplo 3.11. Sean: $X = \{(x_1, x_2) : x_1 + x_2 = 2, x_1 \geq 0, x_2 \geq 0\}$, $f(x) = x_1^3 + x_2^2$. Como X es cerrado y acotado y f es continua, entonces existe un minimizador global de f en X. $\diamond$

Ejemplo 3.12. Sea $f(x) = 1/(1 + x^2)$ en todo $\mathbb{R}$. Esta función es continua en un cerrado, $f(X)$ es acotado inferiormente ya que $f(x) \geq 0$ para todo x, sin embargo, no existe minimizador. $\diamond$

Definición 3.2. Una función f definida en S subconjunto no acotado de $\mathbb{R}^n$ es **coercitiva** si

$$\lim_{||x|| \to \infty} f(x) = +\infty. \tag{3.1}$$

Esto quiere decir que dado cualquier $M > 0$ existe $R > 0$ tal que si $x \in S$ y $||x|| > R$, entonces $f(x) > M$.

En palabras menos formales, f es coercitiva si siempre que x esté muy alejado del origen, el valor $f(x)$ es positivo y muy grande, y entre más alejado esté x, más grande será $f(x)$.

Ejemplo 3.13. $S = \mathbb{R}^n, f(x) = ||x||$ es coercitiva.

$S = \mathbb{R}^n, f(x) = ||x||^2$ es coercitiva.

$S = \mathbb{R}^n, f(x) = ||x - \bar{x}||^2$ (para un $\bar{x}$ fijo) es coercitiva.

$S = \mathbb{R}^n_+, f(x) = ||x - \bar{x}||^2$ (para un $\bar{x}$ fijo) es coercitiva.

$S = \mathbb{R}^2, f(x) = e^{x_1^2} + e^{x_2^2} - x_1^{100} - x_2^{100}$ es coercitiva.

$S = \mathbb{R}^2, f(x) = a_1x_1 + a_2x_2 + a_3$ no es coercitiva. $\diamond$

Ejemplo 3.14. Sean: $S = \mathbb{R}^2$, $f(x) = (x_1 - x_2)^2$. f no es coercitiva, aunque es coercitiva con respecto a x_1 y con respecto a x_2. $\diamond$

Proposición 3.3. *Sea* X *cerrado y no acotado. Si* $f : X \longrightarrow \mathbb{R}$ *es continua y coercitiva, entonces existe* x^* *minimizador global de* f *en* X*.*

Ejemplo 3.15. Sean: $X = \mathbb{R}^2_+$, $f(x_1, x_2) = (x_1 - 4)^3 + (x_2 - 5)^2$. Como X es cerrado y f es continua y coercitiva, entonces existe por lo menos un minimizador global. $\diamond$

Ejemplo 3.16. Sean: X el interior de $\mathbb{R}^2_+$, $f(x_1, x_2) = (x_1 - 4)^2 + (x_2 - 5)^2$. Como X no es cerrado no se puede aplicar la proposición anterior. Sin embargo, sí existe un minimizador global (el punto $x = (4, 5)$). $\diamond$

De la definición se deduce fácilmente que, si una función es coercitiva en un conjunto no acotado, entonces es coercitiva en cada uno de sus subconjuntos no acotados. En particular si una función es coercitiva en $\mathbb{R}^n$, entonces es coercitiva en cualquier conjunto no acotado.

Definición 3.3.

Problema de optimización convexa:

$$\begin{aligned} \min\ & f(x) \\ & x \in X \end{aligned}$$

donde f es una función convexa y X un conjunto convexo.

Problema de optimización cóncava:

$$\begin{aligned} \max\ & f(x) \\ x \in\ & X \end{aligned}$$

donde f es una función cóncava y X un conjunto convexo.

Problema de optimización cuasiconvexa:

$$\begin{aligned} \min\ & f(x) \\ x \in\ & X \end{aligned}$$

donde f es una función cuasiconvexa y X un conjunto convexo.

Problema de optimización cuasicóncava:

$$\begin{aligned} \max\ & f(x) \\ x \in\ & X \end{aligned}$$

donde f es una función cuasicóncava y X un conjunto convexo.

Proposición 3.4. *En optimización convexa, todo minimizador local es minimizador global, Si la función es estrictamente convexa, hay por mucho un minimizador global.*

En optimización cuasiconvexa, un minimizador local estricto es necesariamente un minimizador global estricto.

Resultados análogos con maximizadores, concavidad y cuasiconcavidad.

En optimización cuasiconvexa, un minimizador local no es necesariamente minimizador global.

$$f(x) = \begin{cases} x^3 & \text{si} \quad x \leq 0 \\ 0 & \text{si} \quad 0 < x \leq 2 \\ (x-2)^3 & \text{si} \quad 2 < x \end{cases}$$

f es cuasiconvexa.

$$\begin{aligned} \min\ & f(x) \\ x \in\ & \mathbb{R} \end{aligned}$$

es un problema de optimización cuasiconvexa, $x = 1$ es minimizador local, pero no es minimizador global.

3.1 Optimalidad en puntos interiores

Proposición 3.5. *Sean: $\bar{x}$ un punto interior de X, f diferenciable en $\bar{x}$. Si $\bar{x}$ es un minimizador local del PM, entonces*

$$f'(\bar{x}) = 0. \tag{3.2}$$

Los puntos donde se anula el gradiente se llaman **puntos críticos**.

Ejemplo 3.17. Sean: $X = \mathbb{R}^2$, $f(x_1, x_2) = (x_1 + x_2 - 5)^2 + (3x_1 - 2x_2)^2$. Entonces f es continua, X es cerrado y f es coercitiva, luego existe por lo menos un minimizador global. Al obtener el gradiente $f'(x) = (20x_1 - 10x_2 - 10, -10x_1 + 10x_2 - 10)$ se deduce que el único punto que cumple la condición necesaria de anular el gradiente es $x = (2, 3)$, luego necesariamente es el minimizador global ya que todos los puntos son interiores. $\diamond$

Ejemplo 3.18. Sea $f(x) = x^3$ en el intervalo $[2, \infty[$. Como X es cerrado, f es continua y coercitiva, entonces existe un minimizador global. Si el minimizador global x^* fuera punto interior, entonces $f'(x^*) = 0$, pero esto no es posible, luego el minimizador tiene que ser el único punto no interior, es decir, $x^* = 2$. $\diamond$

Ejemplo 3.19. Sean: $X = \mathbb{R}^2$, $f(x_1, x_2) = x_1^2 + x_2^3$. En este caso no se puede garantizar que $f(X)$ tenga mínimo. Como $f'(x) = (2x_1, 3x_2^2)$, se deduce que el único punto que cumple la condición necesaria de anular el gradiente es $\bar{x} = (0, 0)$, pero no se puede afirmar que es minimizador local o minimizador global. Más aún, $\bar{x} = (0, 0)$ no es minimizador local -y tampoco global- pues los puntos de la forma $(0, -\varepsilon)$, con $\varepsilon > 0$ y pequeño, son mejores que $\bar{x}$ y están muy cerca de $\bar{x}$. $\diamond$

Ejemplo 3.20. Sean: $X = \{(x_1, x_2) : x_1 + x_2 \geq 1,\ x_1 \geq 0,\ x_2 \geq 0\}$, $f(x_1, x_2) = x_1^2 + x_2^3$. Entonces f es continua, X es cerrado y $f(X)$ es continua y coercitiva, luego existe por lo menos un minimizador global. Al obtener el gradiente $f'(x) = (2x_1, 3x_2^2)$ se deduce que no hay puntos interiores que cumplan la condición necesaria de anular el gradiente, luego el minimizador global debe ser un punto de X que esté en la frontera. $\diamond$

Ejemplo 3.21. $f(x_1, x_2) = x_1^2 - x_2^2$ en todo $\mathbb{R}^2$, $\bar{x} = (0, 0)$ es el único candidato a ser minimizador local. No es minimizador global ya que, por ejemplo, $x = (0, 1)$ es mejor. Tampoco es minimizador local pues , si ε es un número positivo pequeño, los puntos $(0, \varepsilon)$ son puntos vecinos y mejores que $\bar{x}$. $\diamond$

Proposición 3.6. *Sean: $\bar{x}$ un punto interior de X, f doblemente diferenciable en $\bar{x}$. Si $\bar{x}$ es un minimizador local del PM, entonces*

$$f'(\bar{x}) = 0 \tag{3.3}$$

y

$$f''(\bar{x}) \quad \textit{es semidefinida positiva.} \tag{3.4}$$

Ejemplo 3.22. Sean: $X = \mathbb{R}^2$, $f(x_1, x_2) = x_1^2 - x_2^2$. Al calcular el gradiente y la matriz hessiana, se obtiene

$$f'(x) = \begin{bmatrix} 2x_1 \\ -2x_2 \end{bmatrix}, \qquad f''(x) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Se deduce que el único punto que cumple la condición necesaria de anular el gradiente es $x = (0, 0)$, sin embargo, la matriz hessiana en este punto no es semidefinida positiva, luego $x = (0, 0)$ no es minimizador local, luego no existe minimizador local ni global. $\diamond$

Ejemplo 3.23. Sean: $X = \{(x_1, x_2) : x_1 + x_2 > -4\}$, $f(x_1, x_2) = (x_1 + x_2 - 5)^2 + (3x_1 - 2x_2)^2$. Al calcular el gradiente y la matriz hessiana se obtiene

$$f'(x) = \begin{bmatrix} 20x_1 - 10x_2 - 10 \\ -10x_1 + 10x_2 - 10 \end{bmatrix}, \qquad f''(x) = \begin{bmatrix} 20 & -10 \\ -10 & 10 \end{bmatrix}.$$

Se ve que en $x = (2, 3)$ se anula el gradiente y la matriz hessiana es semidefinida positiva, luego $x = (2, 3)$ es candidato a ser minimizador local. $\diamond$

Proposición 3.7. *Sean: $\bar{x}$ un punto interior de X, f doblemente diferenciable en $\bar{x}$. Si $f'(\bar{x}) = 0$ y además $f''(\bar{x})$ es definida positiva, entonces $\bar{x}$ es un minimizador local del PM.*

Ejemplo 3.24. Sean: $X = \{(x_1, x_2) : x_1 + x_2 > -4\}$, $f(x_1, x_2) = (x_1 + x_2 - 5)^2 + (3x_1 - 2x_2)^2$. Considerando el ejemplo anterior se observa que en $\bar{x} = (2, 3)$ se anula el gradiente y la matriz hessiana no sólo es semidefinida positiva sino que también es definida positiva, por lo tanto se concluye que $x = (2, 3)$ es minimizador local. $\diamond$

Ejemplo 3.25. Sean: $X = \mathbb{R}^2$, $f(x_1, x_2) = x_1^4 + x_2^2$. Al calcular el gradiente y la matriz hessiana se obtiene

$$f'(x) = \begin{bmatrix} 4x_1^3 \\ 2x_2 \end{bmatrix}, \qquad f''(x) = \begin{bmatrix} 12x_1^2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Se observa que en $x = (0, 0)$ se anula el gradiente y la matriz hessiana es semidefinida positiva, luego es candidato a minimizador local. Como la matriz hessiana no es definida positiva en $x = (0, 0)$, entonces la proposición anterior no permite garantizar que sea un minimizador local. Sin embargo, sí lo es, ya que es el único punto donde la función vale cero. Más aún, es minimizador global. $\diamond$

Proposición 3.8. *Sean: $\bar{x}$ un punto interior de X, f doblemente diferenciable en $\bar{x}$. Si $f'(\bar{x}) = 0$ y existe $r > 0$ tal que $f''(x)$ es semidefinida positiva para todo $x \in B(\bar{x}, r)$, entonces $\bar{x}$ es un minimizador local del PM.*

Ejemplo 3.26. Sean: $X = \mathbb{R}^2$, $f(x_1, x_2) = x_1^4 + x_2^2$. Al calcular el gradiente y la matriz hessiana se obtiene

$$f'(x) = \begin{bmatrix} 4x_1^3 \\ 2x_2 \end{bmatrix}, \qquad f''(x) = \begin{bmatrix} 12x_1^2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Se observa que en $x = (0, 0)$ se anula el gradiente y para cualquier $r > 0$ la matriz hessiana es semidefinida positiva en $B((0, 0), r)$, luego $x = (0, 0)$ es minimizador local del PM. $\diamond$

Proposición 3.9. *Sean: X un conjunto convexo, abierto, $f : X \longrightarrow \mathbb{R}$ convexa y diferenciable. Un punto $\bar{x}$ elemento de X es minimizador global del PM si y solamente si $f'(\bar{x}) = 0$.*

La proposición anterior dice que en optimización convexa diferenciable, un punto interior es minimizador global si y solamente es un punto crítico. Obsérvese que, en general, el gradiente nulo es una condición necesaria para un minimizador local en puntos interiores.

Ejemplo 3.27. $X = \mathbb{R}^2$, $f(x_1, x_2) = x_1^2 + x_2^2$ es diferenciable y se vio que es convexa, luego $\bar{x}$ es minimizador global si y solamente si

$$f'(\bar{x}) = \begin{bmatrix} 2\bar{x}_1 \\ 2\bar{x}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Luego el único minimizador global es $\bar{x} = (0, 0)$. $\diamond$

Proposición 3.10. *Sean: C un conjunto convexo, $f : C \longrightarrow \mathbb{R}$ convexa. Si $\bar{x}$ es minimizador local, entonces es minimizador global.*

Si f es estrictamente convexa o estrictamente cuasiconvexa y $\bar{x}$ es un minimizador global entonces es minimizador global único.

Proposición 3.11. *Sea C convexo, $f : C \longrightarrow \mathbb{R}$ estrictamente cuasiconvexa Si x^* es un minimizador local del PM, entonces x^* es el único minimizador global del PM.*

La propiedad anterior se puede presentar también así: al minimizar una función estrictamente cuasiconvexa en un conjunto convexo, el conjunto de minimizadores globales es vacío o tiene un solo elemento.

Si la función es cuasiconvexa, un minimizador local no es necesariamente minimizador global.

Ejemplo 3.28. Sea $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} (x+1)^3 & \text{si} \quad x \leq -1 \\ 0 & \text{si} \quad -1 < x < 1 \\ (x-1)^3 & \text{si} \quad x \geq 1 \end{cases}$$

f es cuasiconvexa, $\bar{x} = -1/2$ es minimizador local pero no global.

Ejemplo 3.29. Sea $X = \mathbb{R}$, $f(x) = [x]$. El punto $x = 3/2$ es minimizador local, f es cuasiconvexa pero no es convexa, $x = 3/2$ no es minimizador global. $\diamond$

Ejemplo 3.30. Sea $X = \{(x_1, x_2) : x_1 + x_2 > -4\}$, $f(x_1, x_2) = (x_1 + x_2 - 5)^2 + (3x_1 - 2x_2)^2$. El punto $x = (2, 3)$ es minimizador local y como f es convexa, entonces es minimizador global. Además, como f es estrictamente convexa, entonces $x = (2, 3)$ es el único minimizador global. $\diamond$

3.2 Optimización con restricciones, generalidades

Considérese inicialmente el mismo problema de minimización visto hasta ahora:

$$\begin{aligned} \min f(x) \\ x \in X. \end{aligned} \qquad \text{(PM)}$$

donde X es un conjunto cualquiera, no necesariamente abierto. Los resultados presentados en este capítulo permiten estudiar las condiciones de optimalidad en puntos cualesquiera (interiores o no interiores).

Definición 3.4. Sean: $d \in \mathbb{R}^n$, $d \neq 0$, $\bar{x} \in X$. d es **dirección** (local) **de descenso** de f en el punto $\bar{x}$, si existe $\varepsilon > 0$ tal que

$$f(\bar{x} + \lambda d) < f(\bar{x}) \qquad \text{para todo } \lambda \in]0, \varepsilon[.$$

Definición 3.5. Sean: $d \in \mathbb{R}^n$, $d \neq 0$, $\bar{x} \in X$. d es **dirección** (local) **de ascenso** de f en el punto $\bar{x}$, si existe $\varepsilon > 0$ tal que

$$f(\bar{x} + \lambda d) > f(\bar{x}) \qquad \text{para todo } \lambda \in]0, \varepsilon[.$$

Ejemplo 3.31. $f(x_1, x_2) = x_1^2 + x_2^2$ en $\mathbb{R}^2$, $\bar{x} = (3, 4)$. En este ejemplo sencillo minimizar f es equivalente a minimizar la distancia de un punto al origen. La distancia del punto $\bar{x}$ al origen es 5 y una dirección de descenso es aquella que permite, a partir de $\bar{x}$, acercarse al origen. El punto $\bar{x}$ está en la circunferencia con centro en el origen y radio 5, entonces las direcciones de descenso son las que permiten entrar al círculo a partir de $\bar{x}$, o sea, $d = (d_1, d_2)$ es dirección de descenso si $3d_1 + 4d_2 < 0$. Obsérvese que las direcciones tangentes a la circunferencia en $\bar{x}$, es decir, $3d_1 + 4d_2 = 0$ con $d \neq 0$, son direcciones de ascenso. Las otras direcciones de ascenso deben cumplir con $3d_1 + 4d_2 > 0$. $\diamond$

Proposición 3.12. *Sean: $d \in \mathbb{R}^n$, $d \neq 0$, $\bar{x} \in \mathbb{R}^n$, f diferenciable en $\bar{x}$.*

- *si $f'(\bar{x})^{\mathrm{T}} d < 0$, entonces d es dirección de descenso.*
- *si $f'(\bar{x})^{\mathrm{T}} d > 0$, entonces d es dirección de ascenso.*
- *si $f'(\bar{x})^{\mathrm{T}} d = 0$, entonces d puede ser dirección de descenso, o de ascenso o ninguna de las dos.*

Además si $f'(\bar{x}) \neq 0$ y f es convexa, entonces $d \neq 0$ es dirección de descenso si y solamente si $f'(\bar{x})^{\mathrm{T}} d < 0$

Definición 3.6. Sean: $d \in \mathbb{R}^n$, $d \neq 0$, $\bar{x} \in X$. d es **dirección realizable** o **admisible** (local) o **factible** en el punto $\bar{x}$ con respecto al conjunto X, si existe $\varepsilon > 0$ tal que

$$\bar{x} + \lambda d \in X \qquad \text{para todo } \lambda \in]0, \varepsilon[.$$

Sea D_r el conjunto de direcciones realizables o admisibles en el punto $\bar{x}$ con respecto al conjunto X. Los tres resultados siguientes son consecuencias directas de las definiciones.

Proposición 3.13. *Si $\bar{x}$ es un punto interior de X, entonces $D_r = \mathbb{R}^n \setminus \{\mathbf{0}\}$. O sea, todo vector no nulo es dirección realizable.*

Proposición 3.14. *Si $\bar{x}$ es un minimizador local de PM, entonces*

$$D_r \cap D_d = \emptyset.$$

3.3 Optimización con desigualdades. Condiciones de Karush-Kuhn-Tucker

La mayoría de los problemas de programación no lineal se plantean como la minimización de una función f en un conjunto definido por desigualdades e igualdades. Para facilitar el estudio de las condiciones de Karush-Kuhn-Tucker, KKT, se presentará primero el problema únicamente con desigualdades y posteriormente con desigualdades e igualdades. Las funciones g_i, h_j están definidas en $\mathcal{A}$ subconjunto abierto de $\mathbb{R}^n$ y tienen valor real.

$$\begin{aligned} \min f(x) & \\ g_1(x) &\leq 0 \\ g_2(x) &\leq 0 \\ &\dots \\ g_m(x) &\leq 0. \end{aligned} \tag{PMD}$$

$$\begin{aligned} \min f(x) & \\ g_1(x) &\leq 0 \\ g_2(x) &\leq 0 \\ &\dots \\ g_m(x) &\leq 0 \\ h_1(x) &= 0 \\ h_2(x) &= 0 \\ &\dots \\ h_s(x) &= 0. \end{aligned} \tag{PMDI}$$

En todos los casos X indica el conjunto de puntos admisibles o realizables, es decir, el conjunto de puntos que cumplen todas las restricciones.

Definición 3.7. Se dice que la desigualdad $g_i(x) \leq 0$ está **activa** o **saturada** en un punto $\bar{x}$ si se cumple exactamente la igualdad, es decir, si $g_i(\bar{x}) = 0$. Se dice que la desigualdad $g_i(x) \leq 0$ está **inactiva** o **no saturada** o **pasiva** en un punto $\bar{x}$ si se cumple estrictamente la desigualdad, es decir, si $g_i(\bar{x}) < 0$. Sea $\bar{x}$ admisible. Se denotará por $\mathcal{I}$ el conjunto de índices de las desigualdades activas o saturadas en $\bar{x}$:

$$\mathcal{I}(\bar{x}) = \mathcal{I} = \{i : g_i(\bar{x}) = 0\}. \tag{3.5}$$

Obsérvese que una desigualdad se cumple o no se cumple en un punto $\bar{x}$. Si se cumple puede estar activa o inactiva.

Las condiciones de Karush-Kuhn-Tucker, también son conocidas con el nombre de Kuhn-Tucker únicamente, sin embargo, se da también crédito al trabajo de Karush en una tesis de maestría de la Universidad de Chicago, en 1939. Las condiciones de KKT, cuando se pueden aplicar, permiten un manejo algorítmico y preciso, aún para problemas con muchas variables.

Definición 3.8. Un punto $\bar{x}$ admisible para el PMD tal que: g_i es diferenciable en $\bar{x}$ para $i \in \mathcal{I}$, se llama **regular** si los gradientes $g_i'(\bar{x})$ para $i \in \mathcal{I}$ son linealmente independientes o si $\mathcal{I} = \emptyset$.

Proposición 3.15. Teorema de KKT. Condiciones necesarias para el PMD. *Sea $\bar{x}$ tal que:*

$\bar{x}$ es un punto admisible,
f es diferenciable en $\bar{x}$,
g_i es diferenciable en $\bar{x}$ para $i \in \mathcal{I}$,
g_i es continua en $\bar{x}$ para $i \notin \mathcal{I}$,
$\bar{x}$ es regular.

Si $\bar{x}$ es un minimizador local del PMD, entonces existen escalares u_i, $i \in \mathcal{I}$, tales que:

$$\begin{aligned} f'(\bar{x}) + \sum_{i \in \mathcal{I}} u_i\, g_i'(\bar{x}) &= 0, \\ u_i \geq 0\ , \quad i &\in \mathcal{I}. \end{aligned} \tag{3.6}$$

Si el punto no es admisible, no puede ser solución del problema de optimización. Si el punto es admisible pero no es regular, entonces no se puede aplicar el teorema; pero aún así, podría ser minimizador local.

Supongamos ahora que se cumplen las condiciones para aplicar el teorema. La igualdad (3.6) es un sistema de ecuaciones, $\sum_{i \in \mathcal{I}} u_i\, g_i'(\bar{x}) = -f'(\bar{x})$, hay n ecuaciones y $\bar{m} = |\mathcal{I}|$ incógnitas u_i.

Si $\bar{m} < n$, este sistema es sobredeterminado, hay más ecuaciones que incógnitas. Si $\bar{m} = n$, se trata de un sistema cuadrado. Como $\bar{x}$ es regular, las columnas de la matriz de coeficientes son linealmente independientes. Para el caso del sistema sobredeterminado hay dos posibiliades: no hay solución o hay una única solución. Para el sistema cuadrado, necesariamente hay una única solución.

Si el sistema no tiene solución, entonces $\bar{x}$ no es minimizador local. Si el sistema tiene solución, pero algún $u_i < 0$, entonces $\bar{x}$ no es minimizador local. Si el sistema tiene solución y todos los $u_i \geq 0$, entonces el punto $\bar{x}$ es candidato a minimizador local.

Ejemplo 3.32. Considere $\bar{x} = (0, 2)$ para el siguiente problema:

$$\begin{aligned} \min f(x) &= (x_1 - 3)^2 + (x_2 - 2)^2 \\ x_1^2 - x_2 + 2 &\leq 0 \\ -x_1 &\leq 0 \\ -x_2 + 5/2 &\leq 0. \end{aligned}$$

El punto $\bar{x}$ es no es factible, luego no puede ser minimizador. $\diamond$

Ejemplo 3.33. Considere $\bar{x} = (1, 3)$ para el siguiente problema:

$$\begin{aligned} \min f(x) &= (x_1 - 3)^2 + (x_2 - 2)^2 \\ x_1^2 - x_2 + 2 &\leq 0 \\ -x_1 &\leq 0 \\ -x_2 + 5/2 &\leq 0. \end{aligned}$$

El punto $\bar{x}$ es factible; $\mathcal{I} = \{1\}$; f, g_1 son diferenciables en $\bar{x}$; g_2, g_3 son continuas en $\bar{x}$; el conjunto formado por el gradiente $g_1'(\bar{x}) = [2 \quad -1]^{\mathrm{T}}$ es linealmente independiente, luego $\bar{x}$ es regular.

$$f'(\bar{x}) + \sum_{i \in \mathcal{I}} u_i g_i'(\bar{x}) = \begin{bmatrix} -4 \\ 2 \end{bmatrix} + u_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

entonces

$$u_1 = 2 \geq 0.$$

Luego $\bar{x} = (1, 3)$ es un buen candidato a minimizador local. $\diamond$

Ejemplo 3.34. Considere $\bar{x} = (0, 5/2)$ para el siguiente problema:

$$\begin{aligned} \min f(x) &= (x_1 - 3)^2 + (x_2 - 2)^2 \\ x_1^2 - x_2 + 2 &\leq 0 \\ -x_1 &\leq 0 \\ -x_2 + 5/2 &\leq 0. \end{aligned}$$

El punto $\bar{x}$ es factible, $\mathcal{I} = \{2, 3\}$, f, g_2, g_3 son diferenciables en $\bar{x}$, g_1 es continua en $\bar{x}$, el conjunto formado por los gradientes $g_2'(\bar{x}) = [-1 \ \ 0]^{\mathrm{T}}$, $g_3'(\bar{x}) = [0 \quad -1]^{\mathrm{T}}$ es linealmente independiente, luego $\bar{x}$ es regular.

$$f'(\bar{x}) + \sum_{i \in \mathcal{I}} u_i g_i'(\bar{x}) = \begin{bmatrix} -6 \\ 1 \end{bmatrix} + u_2 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + u_3 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

entonces

$$u_2 = -6 \not\geq 0 \ , \quad u_3 = 1.$$

Luego $\bar{x} = (0, 5/2)$ no es minimizador local. $\diamond$

Ejemplo 3.35. Considere $\bar{x} = (1, 1)$ para el siguiente problema:

$$\begin{aligned} \min f(x) &= x_1^2 + x_2^2 \\ 2 - x_1 - x_2 &\leq 0 \\ 1 - x_1 &\leq 0 \\ 1 - x_2 &\leq 0. \end{aligned}$$

El punto $\bar{x}$ es factible; $\mathcal{I} = \{1, 2, 3\}$; f, g_1, g_2, g_3 son diferenciables en $\bar{x}$, pero el conjunto formado por los tres gradientes no es linealmente independiente, luego no se puede aplicar el teorema. Sin embargo, el punto $\bar{x} = (1, 1)$ sí es el minimizador global. En este ejemplo sencillo se puede ver que al suprimir la primera restricción el conjunto admisible no cambia. O sea, el problema anterior es exactamente equivalente a:

$$\begin{aligned} \min f(x) &= x_1^2 + x_2^2 \\ 1 - x_1 &\leq 0 \\ 1 - x_2 &\leq 0. \end{aligned}$$

El punto $\bar{x} = (1, 1)$ es factible; $\mathcal{I} = \{1, 2\}$; f, g_1, g_2 son diferenciables en $\bar{x}$; el conjunto formado por los gradientes $g_1'(\bar{x}) = [-1 \ \ 0]^{\mathrm{T}}$, $g_2'(\bar{x}) = [0 \quad -1]^{\mathrm{T}}$ es linealmente independiente, luego $\bar{x}$ es regular.

$$\begin{bmatrix} 2 \\ 2 \end{bmatrix} + u_1 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + u_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

entonces,

$$u_1 = 2 \geq 0 \ , \quad u_2 = 2 \geq 0.$$

Luego $\bar{x} = (1, 1)$ es candidato a minimizador local. $\diamond$

Ejemplo 3.36. Considere $\bar{x} = (0, 1)$ para el siguiente problema:

$$\begin{aligned} \min f(x_1, x_2) &= -x_2 \\ -x_1^2 - x_2^2 + 1 &\leq 0 \\ x_1^2 + (x_2 + 1)^2 - 4 &\leq 0 \\ -x_2 &\leq 0. \end{aligned}$$

El punto $\bar{x}$ es factible; $\mathcal{I} = \{1, 2\}$; f, g_1, g_2 son diferenciables en $\bar{x}$; g_3 es continua en $\bar{x}$; el conjunto formado por los dos gradientes $g_1'(\bar{x}) = [0 \quad -2]^{\mathrm{T}}$, $g_2'(\bar{x}) = [0 \ \ 4]^{\mathrm{T}}$ no es linealmente independiente, luego no se puede aplicar el teorema. Sin embargo, al construir la región admisibles, se "ve" que el punto $\bar{x} = (0, 1)$ sí es el minimizador global. Acá no se puede quitar ninguna de las restricciones pues el conjunto admisible se altera. Habría necesidad de utilizar otros criterios para el estudio de este problema. $\diamond$

En los ejemplos anteriores ha sido relativamente simple saber si un punto $\bar{x}$ cumple o no cumple condiciones de KKT. En un caso general, suponiendo $\bar{x}$ admisible y regular, se trata de resolver un problema de la forma

$$\begin{aligned} Ay &= b \\ y &\geq 0, \end{aligned}$$

donde A es una matriz $n \times \bar{m}$, siendo $\bar{m}$ el número de desigualdades saturadas; $b = -f'(\bar{x})$; y el vector columna $\bar{m} \times 1$ compuesto por las variables $u_i, i \in \mathcal{I}$; las columnas de A son los gradientes de las desigualdades activas evaluados en $\bar{x}$. Como $\bar{x}$ es regular, entonces las columnas de A son linealmente independientes, luego rango$(A) = \bar{m} \leq n$.

Es necesario resolver el sistema $Ay = b$ y ver si hay solución tal que $y \geq 0$.

Un procedimiento seguro es el método de Gauss, construir la matriz ampliada $[A \ \ b]$ y llevarla a la forma escalonada reducida. En Scilab la matriz escalonada reducida se obtiene por medio de la función `rref`.

Como las columnas de A son linealemnte independientes, únicamente se presentan dos casos: sistema inconsistente y sistema con una única solución.

Si el sistema es inconsistente el punto en estudio no es minimizador local. Cuando tenga una única solución y, verificar si $y \geq 0$. Si esto se cumple el punto es candidato a minimizador local. Si algun $y_j < 0$, el punto no es minimizador local.

Ejemplo 3.37. Considere $\bar{x} = (1, 1, 1, 1)$ para el siguiente problema:

$$\min f(x) = x_1^2 + 3x_2^2 + x_3^2 + x_4^2$$

$$\begin{aligned} (x_1 - 4)^2 + x_2^2 + x_3^2 + x_4^2 - 15 &\leq 0 \\ 1 - x_2 &\leq 0 \\ 10 - x_1 - 2x_2 - 3x_3 - 4x_4 &\leq 0 \\ 4 - x_1 - x_2 - x_3 - x_4 &\leq 0. \end{aligned}$$

El punto $\bar{x}$ es factible; $\mathcal{I} = \{2, 3, 4\}$; f, g_2, g_3, g_4 son diferenciables en $\bar{x}$; g_1 es continua en $\bar{x}$; el conjunto formado por los gradientes

$$\begin{aligned} g_2'(\bar{x}) &= [\ \ 0 \ \ -1 \ \ \ 0 \ \ \ 0 \ \ ]^{\mathrm{T}}, \\ g_3'(\bar{x}) &= [\ -1 \ \ -2 \ \ -3 \ \ -4 \ ]^{\mathrm{T}}, \\ g_4'(\bar{x}) &= [\ -1 \ \ -1 \ \ -1 \ \ -1 \ ]^{\mathrm{T}} \end{aligned}$$

es linealmente independiente. El gradiente de f es $f'(\bar{x}) = [2 \ \ 6 \ \ 2 \ \ 2]^{\mathrm{T}}$. El problema que hay que resolver es el siguiente:

$$\begin{bmatrix} 0 & -1 & -1 \\ -1 & -2 & -1 \\ 0 & -3 & -1 \\ 0 & -4 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -2 \\ -6 \\ -2 \\ -2 \end{bmatrix}$$
$$y \geq 0,$$

donde $y_1 = u_2$, $y_2 = u_3$, $y_3 = u_4$.

La matriz ampliada:

$$\begin{bmatrix} 0 & -1 & -1 & -2 \\ -1 & -2 & -1 & -6 \\ 0 & -3 & -1 & -2 \\ 0 & -4 & -1 & -2 \end{bmatrix},$$

y por medio de operaciones elementales sobre las filas se llega a

$$\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

lo cual indica, primero, que el sistema sí tiene solución, y, segundo, que la solución tiene todas sus componentes no negativas, luego $\bar{x}$ es punto de KKT. $\diamond$

Ejemplo 3.38. Considere el problema de minimizar

$$f(x_1, x_2, x_3, x_4) = x_1^2 + 3x_2^2 + x_3^2 + 2x_4^2$$

con las mismas restricciones y el mismo punto $\bar{x} = (1, 1, 1, 1)$

Lo único que cambia es el gradiente $f'(\bar{x}) = (2, 6, 2, 4)$. La matriz ampliada:

$$\begin{bmatrix} 0 & -1 & -1 & -2 \\ -1 & -2 & -1 & -6 \\ 0 & -3 & -1 & -2 \\ 0 & -4 & -1 & -4 \end{bmatrix},$$

y por medio de operaciones elementales sobre las filas se llega a

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

lo cual indica que el sistema no tiene solución (en un paso intermedio del método de Gauss también se pudo haber detectato la inconsistencia), luego el punto no es minimizador local. $\diamond$

Ejemplo 3.39. Considere el problema de minimizar

$$f(x_1, x_2, x_3, x_4) = x_1^2 - 2x_2^2 + x_3^2 + x_4^2$$

con las mismas restricciones y el mismo punto $\bar{x} = (1, 1, 1, 1)$

Lo único que cambia es el gradiente $f'(\bar{x}) = (2, -4, 2, 2)$. La matriz ampliada:

$$\begin{bmatrix} 0 & -1 & -1 & -2 \\ -1 & -2 & -1 & 4 \\ 0 & -3 & -1 & -2 \\ 0 & -4 & -1 & -2 \end{bmatrix},$$

y por medio de operaciones elementales sobre las filas se llega a

$$\begin{bmatrix} 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

lo cual indica que el sistema sí tiene solución, pero como un u_i es negativo el punto no es minimizador local. $\diamond$

3.3.1 Estudio exhaustivo

Para problemas pequeños, sin partir de un punto admisible explícito, se puede tratar de resolver completamente un problema estudiando todas las posibilidades para el conjunto $\mathcal{I}$. Es decir, se buscan todos los posibles puntos de KKT. Para cada posible caso de $\mathcal{I}$, es necesario resolver un sistema de ecuaciones que pueden ser no lineales, Cuando f es cuadrática y todas las restricciones son lineales, el sistema que se va a resolver es un sistema de ecuaciones lineales.

Si todos los puntos factibles son regulares y si se puede garantizar la existencia de un minimizador global, entonces uno de los puntos de KKT debe ser el minimizador global; si hay un solo punto de KKT, este debe ser el minimizador global; si hay varios puntos de KKT, el minimizador global debe ser el mejor punto de KKT. Mejor punto de KKT quiere decir el de menor valor al evaluar f.

Si no se puede garantizar la existencia de un minimizador global y hay uno o varios puntos de KKT, puede ser que uno o algunos de ellos sean minimizadores locales. El mejor de estos puntos puede ser minimizador global, pero también es posible que no haya minimizador local ni global.

Si todos los puntos factibles son regulares y no hay puntos de KKT, se puede concluir que no hay minimizadores locales y tampoco hay minimizador global.

Es posible que para un $\mathcal{I}$ supuesto, resulte un punto x factible y punto de KKT y que para este x obtenido el $\mathcal{I}$ sea diferente del supuesto (más "grande que el supuesto"). En este caso el estudio se hará con el verdadero $\mathcal{I}$

Ejemplo 3.40. Resolver, utilizando condiciones de KKT, el siguiente problema:

$$\begin{aligned} \min \quad f(x_1, x_2) = x_1^2 + 3x_2^2 + 4x_1x_2 + 7x_1 + 5x_2 \\ x_1 - 2x_2 \le 0 \\ 3x_1 + x_2 \le 7 \end{aligned}$$

Reescrito en la forma usual:

$$\begin{aligned} \min \quad f(x_1, x_2) = x_1^2 + 3x_2^2 + 4x_1x_2 + 7x_1 + 5x_2 \\ x_1 - 2x_2 \le 0 \\ 3x_1 + x_2 - 7 \le 0 \end{aligned}$$

En este problema las funciones f, g_1, g_2 son continuas y diferenciables. Todos los puntos factibles son regulares ya que los gradientes de g_1 y g_2 son constantes y sin depender de x, si no hay restricciones activas, si hay una o si hay dos, siempre los gradientes serán linealmente independientes y el punto factible será regular.

Para $\mathcal{I}$ hay cuatro posibilidades: $\emptyset$, $\{1\}$, $\{2\}$, $\{1, 2\}$.

$\mathcal{I} = \emptyset$, $f'(x) = 0$:

$$\begin{bmatrix} 2x_1 + 4x_2 + 7 \\ 4x_1 + 6x_2 + 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Al resolver este sistema de ecuaciones lineales se obtiene $x = (11/2, -9/2)$, punto no factible.

$\mathcal{I} = \{1\}$, $f'(x) + u_1 g_1'(x) = 0$, $g_1(x) = 0$.

$$\begin{bmatrix} 2x_1 + 4x_2 + 7 \\ 4x_1 + 6x_2 + 5 \end{bmatrix} + u_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$x_1 - 2x_2 = 0$$

Se trata de un sistema de tres ecuaciones lineales con tres incógnitas. Su matriz ampliada es

$$\begin{bmatrix} 2 & 4 & 1 & -7 \\ 4 & 6 & -2 & -5 \\ 1 & -2 & 0 & 0 \end{bmatrix}$$

Su solución es $x = (-19/15, -19/30)$, $u_1 = -29/15$. El punto x es factible pero $u_1 < 0$, luego x no es punto de KKT, entonces no minimizador local.

$\mathcal{I} = \{2\}$, $f'(x) + u_2 g_2'(x) = 0$, $g_2(x) = 0$.

La matriz ampliada del sistema es

$$\begin{bmatrix} 2 & 4 & 3 & -7 \\ 4 & 6 & 1 & -5 \\ 3 & 1 & 0 & 7 \end{bmatrix}$$

$x = (53/16, -47/16)$, $u_2 : -5/8$. El punto x no es factible.

$\mathcal{I} = \{1, 2\}$, $f'(x) + u_1 g_1'(x) + u_2 g_2'(x) = 0$, $g_1(x) = 0$, $g_2(x) = 0$.

La matriz ampliada del sistema es

$$\begin{bmatrix} 2 & 4 & 1 & 3 & -7 \\ 4 & 6 & -2 & 1 & -5 \\ 1 & -2 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 7 \end{bmatrix}$$

$x = (2, 1)$, $u_1 = 6$, $u_2 = -7$. El punto x es factible pero $u_2 < 0$, luego x no es minimizador local.

Estudiados todos los casos para $\mathcal{I}$ y como todos los puntos factibles son regulares, se concluye que no hay ningún minimizador local y tampoco minimizador global, el valor mínimo de f no es acotado (inferiormente). $\diamond$

Cuando el número de desigualdades activas es igual a n (en el ejemplo anterior $\mathcal{I} = \{1, 2\}$), se puede resolver directamente el sistema de n ecuaciones con n incógnitas. Supongamos que hay una o varias soluciones. Para cada solución se debe averiguar si el punto es factible. Si el punto es factible, es necesario hacer estudio de condiciones de KKT para el punto.

En el ejemplo anterior, si $\mathcal{I} = \{1, 2\}$), entonces

$$\begin{aligned} x_1 - 2x_2 &= 0 \\ 3x_1 + x_2 - 7 &= 0 \end{aligned}$$

Luego, $x = (2, 1)$, punto que es factible. Al plantear condiciones de KKT,

$$\begin{bmatrix} 15 \\ 19 \end{bmatrix} + u_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} + u_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

La solución es: $u_1 = -6$, $u_2 = 7$. Luego $x = (2, 1)$ no es minimizador local. Con este enfoque, para los casos en que se puede aplicar, se trabaja con sistemas de ecuaciones de tamaño menor, dos sistemas 2×2, en lugar de uno 4×4.

Ejemplo 3.41. Resolver utilizando condiciones de KKT

$$\begin{aligned} f(x_1, x_2) = -3x_1^2 - 4x_2^2 + 7x_1x_2 + 3x_1 + x_2& \\ -3x_1 - 8x_2 &\le -2 \\ x_1 + x_2 &\le 6 \\ 2x_1 - 6x_2 &\le 0 \end{aligned}$$

Reeescribiendo el problema:

$$\begin{aligned} f(x_1, x_2) = -3x_1^2 - 4x_2^2 + 7x_1x_2 + 3x_1 + x_2& \\ -3x_1 - 8x_2 + 2 &\le 0 \\ x_1 + x_2 - 6 &\le 0 \\ 2x_1 - 6x_2 &\le 0 \end{aligned}$$

Hay 8 casos posibles para $\mathcal{I}$: $\emptyset$, $\{1\}$, $\{2\}$, $\{3\}$, $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$, $\{1, 2, 3\}$.

$\mathcal{I} = \emptyset :\ x = (-31, -27)$ no factible.

$\mathcal{I} = \{1\} :\ x = (41/99, 0.094697)$ no factible.

$\mathcal{I} = \{2\} :\ x = (23/719/7)$ factible, $u_2 = -16/7$, x no es minimizador local.

$\mathcal{I} = \{3\} :\ x = (3/2, 1/2)$ factible, $u_3 = 5/4$, x es punto de KKT.

$\mathcal{I} = \{1, 2\} :\ x = (46/5, -16/5)$ no factible.

$\mathcal{I} = \{1, 3\} :\ x = (6/17, 2/17)$ factible, $u_1 = 0.449827$, $u_3 = -0.178201$, x no es minimizador local.

$\mathcal{I} = \{2, 3\} :\ x = (9/2, 3/2)$ factible, $u_1 = 5$, $u_3 = 17/4$, x es punto de KKT.

$\mathcal{I} = \{1, 2, 3\} :$ el sistema de ecuaciones no tiene solución.

En resumen, hay dos puntos candidatos a minimizador local.

Ejemplo 3.42. ????

3.3.2 Lagrangiano

Si las funciones $g_i, i \notin \mathcal{I}$ también son diferenciables en $\bar{x}$, y si se considera que los u_i correspondientes son nulos, la admisibilidad y las condiciones necesarias de KKT se pueden escribir así:

$$
\begin{aligned}
g_i(\bar{x}) &\leq 0\,, \quad i = 1, ..., m && (3.7)\\
f'(\bar{x}) + \sum_{i=1}^{m} u_i g_i'(\bar{x}) &= 0 && (3.8)\\
u_i &\geq 0\,, \quad i = 1, ..., m\\
u_i g_i(\bar{x}) &= 0\,, \quad i = 1, ..., m. && (3.9)
\end{aligned}
$$

La primera condición que debe cumplir un punto $\bar{x}$ para tratar de ser minimizador, es la de ser admisible, o sea, cumplir todas las restricciones $g_i(\bar{x}) \leq 0$. Estas condiciones (3.7) se conocen también con el nombre de **condiciones de factibilidad principal** o **"primal"** (con frecuencia se utiliza en español técnico la palabra "primal", aunque puede ser un anglicismo). La existencia de $u_i \geq 0$, tales que $f'(\bar{x}) + \sum u_i g_i'(\bar{x}) = 0$, se conoce con el nombre de **condiciones de factibilidad dual** (3.8). Las condiciones (3.9) $u_i g_i(\bar{x}) = 0$ se conocen con el nombre de **condiciones de holgura complementaria**.

Si $g(\bar{x})$ denota el vector columna $[g_1(\bar{x}) \;\; g_2(\bar{x}) \;\; ... \;\; g_m(\bar{x})]^{\mathrm{T}}$, $g'(\bar{x})$ denota la matriz $n \times m$ cuyas columnas son los gradientes $g_1'(\bar{x})$, $g_2'(\bar{x})$,..., $g_m'(\bar{x})$, y teniendo en cuenta que $u^{\mathrm{T}} g(\bar{x})$ es suma de números no positivos, entonces la admisibilidad y las condiciones necesarias de KKT se pueden escribir

$$
\begin{aligned}
g(\bar{x}) &\leq 0\\
f'(\bar{x}) + g'(\bar{x})u &= 0\\
u &\geq 0\\
u^{\mathrm{T}} g(\bar{x}) &= 0.
\end{aligned}
$$

Introduciendo la función lagrangiana, o simplemente el lagrangiano, función de $n+m$ variables $x_1, ..., x_n$, $u_1, ..., u_m$

$$
\mathcal{L}(x, u) = f(x) + \sum_{i=1}^{m} u_i g_i(x) = f(x) + u^{\mathrm{T}} g(x), \tag{3.10}
$$

y denotando por $\mathcal{L}_x'(\bar{x}, u)$ las componentes del gradiente $\mathcal{L}'(\bar{x}, u)$ correspondientes a las derivadas parciales de $\mathcal{L}$ con respecto a las variables x_j, y de manera análoga $\mathcal{L}_u'(\bar{x}, u)$, entonces la admisibilidad y las condiciones de KKT se expresan

$$
\begin{aligned}
\mathcal{L}_u'(\bar{x}, u) &\leq 0\\
\mathcal{L}_x'(\bar{x}, u) &= 0\\
u &\geq 0\\
u^{\mathrm{T}} \mathcal{L}_u'(\bar{x}, u) &= 0.
\end{aligned}
$$

Los coeficientes u_i se conocen con el nombre de **coeficientes de KKT** o **coeficientes de Lagrange**. Un punto que cumple las condiciones necesarias de KKT se llama un **punto de KKT**.

Proposición 3.16. Condiciones suficientes para el PMD. *Si el punto $\bar{x}$ cumple las condiciones necesarias de KKT para el PMD y además f es convexa y el conjunto admisible es un conjunto convexo (problema de optimización convexa), entonces $\bar{x}$ es un minimizador global del PMD.*

Hay condiciones suficientes menos exigentes en las cuales no se requiere convexidad, basta con seudoconvexidad de f y cuasiconvexidad de las funciones g_i activas (ver Baz08, p. 195).

Ejemplo 3.43. Considere $\bar{x} = (1, 3)$ para el siguiente problema:

$$
\begin{aligned}
\min f(x) &= (x_1 - 3)^2 + (x_2 - 2)^2\\
x_1^2 - x_2 + 2 &\leq 0\\
-x_1 &\leq 0\\
-x_2 + 5/2 &\leq 0.
\end{aligned}
$$

El punto $\bar{x}$ es factible, $\mathcal{I} = \{1\}$, f, g_1 son diferenciables, g_2 y g_3 son continuas. El punto es regular ya que el conjunto formado por el gradiente $g_1'(\bar{x}) = [2 \ \ -1]$ es linealmente independiente. El sistema $f'(\bar{x})+u_1g_1'(\bar{x}) = 0$,

$$\begin{bmatrix} -4 \\ 2 \end{bmatrix} + u_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

tiene solución, $u_1 = 2$, luego el punto cumple las condiciones de KKT (de primer orden). La función f es convexa ya que su matriz hessiana es definida positiva.

$$f''(x) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

De manera análoga, las funciones g_i son convexas, luego el conjunto factible es convexo, y se concluye que el punto $\bar{x}$ es minimizador global. $\diamond$

Ejemplo 3.44. Considere $\bar{x} = (1, 1)$ para el siguiente problema:

$$\begin{aligned} \min f(x) &= \cos x_1 + \exp(x_1 + x_2) \\ 2 - x_1 - x_2 &\leq 0 \\ x_1^2 + x_2^2 - 2 &\leq 0. \end{aligned}$$

Obviamente este punto es minimizador global puesto que es el único punto admisible. Sin embargo, no es un punto de KKT (no es regular), luego no cumple condiciones suficientes. $\diamond$

3.4 Problemas con desigualdades e igualdades

$$\begin{aligned} &\min f(x) \\ &g_1(x) \leq 0 \\ &g_2(x) \leq 0 \\ &\cdots \\ &g_m(x) \leq 0 \\ &h_1(x) = 0 \\ &h_2(x) = 0 \\ &\cdots \\ &h_s(x) = 0. \end{aligned} \tag{PMDI}$$

Para el PMDI (problema de minimización con desiguladades e igualdades), sea $p = \bar{m} + s$, es decir, p es el número de desigualdades activas más el número de igualdades.

Definición 3.9. Un punto $\bar{x}$ admisible para el PMDI tal que: g_i es diferenciable en $\bar{x}$ para $i \in \mathcal{I}$, h_j es diferenciable en $\bar{x}$ para todo j , se llama **regular** si los gradientes $g_i'(\bar{x})$ con $i \in \mathcal{I}$ y $h_j'(\bar{x})$, $\forall j$ son linealmente independientes, o si $p = 0$, o sea, si $\mathcal{I} = \emptyset$ y $l = 0$.

Proposición 3.17. Condiciones necesarias de Karush-Kuhn-Tucker para el PMDI. *Sea $\bar{x}$ tal que:*

$\bar{x}$ es un punto admisible del PMDI,
f es diferenciable en $\bar{x}$,
g_i es diferenciable en $\bar{x}$ para $i \in \mathcal{I}$,
h_j es diferenciable en $\bar{x}$ para todo j,
g_i es continua en $\bar{x}$ para $i \notin \mathcal{I}$, y
$\bar{x}$ es regular.

Si $\bar{x}$ es un minimizador local del PMDI, entonces existen escalares u_i, $i \in \mathcal{I}$, $v_j \forall j$ tales que:

$$f'(\bar{x}) + \sum_{i \in \mathcal{I}} u_i g_i'(\bar{x}) + \sum_{j=1}^{s} v_j h_j'(\bar{x}) = 0, \tag{3.11}$$

$$u_i \geq 0 \ , \quad i \in \mathcal{I}$$

Ejemplo 3.45. Considere $\bar{x} = (1, 3)$ para el siguiente problema:

$$\begin{aligned} \min f(x) &= (x_1 - 3)^2 + (x_2 - 2)^2 \\ -x_1 &\leq 0 \\ -x_2 + 5/2 &\leq 0 \\ -x_1^2 + x_2 - 2 &= 0. \end{aligned}$$

El punto $\bar{x}$ es factible, $\mathcal{I} = \emptyset$, f, h_1 son diferenciables en $\bar{x}$, g_1, g_2 son continuas en $\bar{x}$, el conjunto formado por el gradiente $h_1'(\bar{x}) = [-2 \;\; 1]^{\mathrm{T}}$ es linealmente independiente, o sea, $\bar{x}$ es regular.

$$f'(\bar{x}) + \sum_{i \in \mathcal{I}} u_i g_i'(\bar{x}) + \sum_{j=1}^{s} v_j h_j'(\bar{x}) = \begin{bmatrix} -4 \\ 2 \end{bmatrix} + v_1 \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

entonces

$$v_1 = -2.$$

Luego $\bar{x} = (1, 3)$ es buen candidato a minimizador local. $\diamond$

Ejemplo 3.46. Considere $\bar{x} = (\sqrt{2}/2, 5/2)$ para el siguiente problema:

$$\begin{aligned} \min f(x) &= (x_1 - 3)^2 + (x_2 - 2)^2 \\ -x_1 &\leq 0 \\ -x_2 + 5/2 &\leq 0 \\ -x_1^2 + x_2 - 2 &= 0. \end{aligned}$$

El punto $\bar{x}$ es factible, $\mathcal{I} = \{2\}$, f, g_2, h_1 son diferenciables en $\bar{x}$, g_1 es continua en $\bar{x}$, el conjunto formado por los gradientes $g_2'(\bar{x}) = [0 \;\; -1]^{\mathrm{T}}$, $h_1'(\bar{x}) = [-\sqrt{2} \;\; 1]^{\mathrm{T}}$ es linealmente independiente, o sea, $\bar{x}$ es regular.

$$\begin{bmatrix} \sqrt{2} - 6 \\ 1 \end{bmatrix} + u_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} + v_1 \begin{bmatrix} -\sqrt{2} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

entonces

$$u_2 = 2 - 3\sqrt{2} \not\geq 0 \;, \quad v_1 = 1 - 3\sqrt{2}.$$

Luego $\bar{x} = (\sqrt{2}/2, 5/2)$ no es minimizador local. $\diamond$

De manera análoga a los problemas con desigualdades, para saber si un punto $\bar{x}$ admisible es punto de KKT, se debe resolver un problema de la forma

$$\begin{aligned} A \begin{bmatrix} y \\ v \end{bmatrix} &= b \\ y &\geq 0. \end{aligned}$$

donde A es una matriz $n \times p$, siendo $p = \bar{m} + s$, $\bar{m}$ el número de desigualdades saturadas; $b = -f'(\bar{x})$; y es el vector columna $\bar{m} \times 1$ compuesto por las variables $u_i, i \in \mathcal{I}$; v el vector columna $s \times 1$ compuesto por las variables v_j ; las columnas de A son los gradientes de las desigualdades activas y de las igualdades evaluados en $\bar{x}$. Como $\bar{x}$ es regular, entonces las columnas de A son linealmente independientes, luego rango$(A) = p = \bar{m}+s \leq n$.

Es necesario resolver el sistema

$$A \begin{bmatrix} y \\ v \end{bmatrix} = b$$

y ver si hay solución tal que $y \geq 0$.

Más precisamente:

1) Construir la matriz ampliada $\hat{A} = [A \;\; b]$ de tamaño $n \times (p + 1)$.

2) Convertirla, mediante operaciones elementales por filas, en una matriz

$$\hat{A}' = \begin{bmatrix} I_{\bar{m}} & 0 & c \\ 0 & I_s & e \\ 0 & 0 & d \end{bmatrix},$$

donde c es un vector columna $\bar{m} \times 1$, e es un vector columna $s \times 1$, d es un vector columna $(n - \bar{m} - s) \times 1$, es decir, se obtiene el siguiente sistema equivalente

$$\begin{bmatrix} I_{\bar{m}} & 0 \\ 0 & I_l \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ v \end{bmatrix} = \begin{bmatrix} c \\ e \\ d \end{bmatrix},$$

o sea,

$$\begin{aligned} I_{\bar{m}} y &= c \\ I_l v &= e \\ 0y + 0v &= d. \end{aligned}$$

El paso 2) siempre es posible puesto que las columnas de A son linealmente independientes.

3) Si $d \neq 0$ el sistema $A \begin{bmatrix} y \\ v \end{bmatrix} = b$ no tiene solución y el punto $\bar{x}$ no es minimizador local. Si $d = 0$ el sistema tiene como única solución $y = c, \;\; v = e$.

4) Si $y = c \geq 0$ el punto $\bar{x}$ es punto de KKT. Si $y = c \not\geq 0$, entonces $\bar{x}$ no es minimizador local.

Ejemplo 3.47. Considere $\bar{x} = (1, 1, 1, 1)$ para el siguiente problema:

$$\min f(x) = x_1^2 + 3x_2^2 + x_3^2 + x_4^2$$

$$\begin{aligned} (x_1 - 4)^2 + x_2^2 + x_3^2 + x_4^2 - 15 &\leq 0 \\ 1 - x_2 &\leq 0 \\ x_1 + 2x_2 + 3x_3 + 4x_4 - 10 &= 0 \\ x_1 + x_2 + x_3 + x_4 - 4 &= 0. \end{aligned}$$

El punto $\bar{x}$ es factible, $\mathcal{I} = \{2\}$, f, g_2, h_1, h_2 son diferenciables en $\bar{x}$, g_1 es continua en $\bar{x}$, el conjunto formado por los gradientes

$$\begin{aligned} g_2'(\bar{x}) &= [\begin{array}{cccc} 0 & -1 & 0 & 0 \end{array}]^{\mathrm{T}}, \\ h_1'(\bar{x}) &= [\begin{array}{cccc} 1 & 2 & 3 & 4 \end{array}]^{\mathrm{T}}, \\ h_2'(\bar{x}) &= [\begin{array}{cccc} 1 & 1 & 1 & 1 \end{array}]^{\mathrm{T}} \end{aligned}$$

es linealmente independiente. El gradiente de f es $f'(\bar{x}) = [2 \;\; 6 \;\; 2 \;\; 2]^{\mathrm{T}}$. El problema que hay que resolver es el siguiente:

$$\begin{bmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ 0 & 3 & 1 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -2 \\ -6 \\ -2 \\ -2 \end{bmatrix}$$
$$y \geq 0,$$

donde $y_1 = u_2$.

See construye la matriz ampliada

$$\begin{bmatrix} 0 & 1 & 1 & -2 \\ -1 & 2 & 1 & -6 \\ 0 & 3 & 1 & -2 \\ 0 & 4 & 1 & -2 \end{bmatrix}.$$

Por medio de operaciones elementales sobre las filas se llega a

$$\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

lo cual indica primero, que el sistema sí tiene solución, y, segundo, que y, una parte de la solución, tiene todas sus componentes no negativas: $u_2 = y_1 = 4$, luego $\bar{x}$ es punto de KKT. $\diamond$

Si las funciones $g_i, i \notin \mathcal{I}$ son diferenciables en $\bar{x}$, y si se considera que los u_i correspondientes son nulos, la admisibilidad y las condiciones necesarias de KKT se pueden escribir así:

$$\begin{aligned} g_i(\bar{x}) &\leq 0\,, \quad i = 1, ..., m \\ h_j(\bar{x}) &= 0\,, \quad j = 1, ..., l \\ f'(\bar{x}) + \sum_{i=1}^{m} u_i g_i'(\bar{x}) + \sum_{j=1}^{s} v_j h_j'(\bar{x}) &= 0 \\ u_i &\geq 0\,, \quad i = 1, ..., m \\ u_i g_i(\bar{x}) &= 0\,, \quad i = 1, ..., m. \end{aligned}$$

Si $g(\bar{x})$ denota el vector columna $[g_1(\bar{x})\ \ g_2(\bar{x})\ \ ...\ \ g_m(\bar{x})]^{\mathrm{T}}$, $g'(\bar{x})$ denota la matriz $n \times m$ cuyas columnas son los gradientes $g_1'(\bar{x})$, $g_2'(\bar{x})$,..., $g_m'(\bar{x})$, $h(\bar{x})$ denota el vector columna $[h_1(\bar{x})\ \ h_2(\bar{x})\ \ ...\ \ h_s(\bar{x})]^{\mathrm{T}}$, $h'(\bar{x})$ denota la matriz $n \times l$ cuyas columnas son los gradientes $h_1'(\bar{x})$, $h_2'(\bar{x})$,..., $h_s'(\bar{x})$, entonces la admisibilidad y las condiciones necesarias de KKT se pueden escribir

$$g(\bar{x}) \leq 0 \tag{3.12}$$
$$h(\bar{x}) = 0 \tag{3.13}$$
$$f'(\bar{x}) + g'(\bar{x})u + h'(\bar{x})v = 0 \tag{3.14}$$
$$u \geq 0 \tag{3.15}$$
$$u^{\mathrm{T}} g(\bar{x}) = 0. \tag{3.16}$$

Utilicemos la función lagrangiana, o simplemente el lagrangiano, función de $n + m + s$ variables $x_1, ..., x_n$, $u_1, ..., u_m$, $v_1, ..., v_s$, definida por:

$$\begin{aligned} \mathcal{L}(x, u, v) &= f(x) + \sum_{i=1}^{m} u_i g_i(x) + \sum_{j=1}^{s} v_j h_j(x) \qquad (3.17) \\ &= f(x) + u^{\mathrm{T}} g(x) + v^{\mathrm{T}} h(x). \end{aligned}$$

Denotemos por $\mathcal{L}_x'(\bar{x}, u, v)$ las componentes del gradiente del lagrangiano $\mathcal{L}'(\bar{x}, u, v)$ correspondientes a las derivadas parciales de $\mathcal{L}$ con respecto a las variables x_j, y de manera análoga $\mathcal{L}_u'(\bar{x}, u, v)$ y $\mathcal{L}_v'(\bar{x}, u, v)$. Entonces la admisibilidad y las condiciones de KKT se expresan así :

$$\mathcal{L}_u'(\bar{x}, u, v) \leq 0 \tag{3.18}$$
$$\mathcal{L}_v'(\bar{x}, u, v) = 0 \tag{3.19}$$
$$\mathcal{L}_x'(\bar{x}, u, v) = 0 \tag{3.20}$$
$$u \geq 0 \tag{3.21}$$
$$u^{\mathrm{T}} \mathcal{L}_u'(\bar{x}, u, v) = 0. \tag{3.22}$$

El enunciado de la proposición (3.16), sigue siendo válido para problemas con desigualdades e igualdades. *Si el punto $\bar{x}$ cumple las condiciones necesarias de KKT para el PMDI y además f es convexa y el conjunto admisible es un conjunto convexo (problema de optimización convexa), entonces $\bar{x}$ es un minimizador global del PMDI.*

Ejemplo 3.48. Considere $\bar{x} = (1, 3)$ para el siguiente problema:

$$\begin{aligned} \min f(x) &= (x_1 - 3)^2 + (x_2 - 2)^2 \\ -x_1 + 1 &\leq 0 \\ -x_2 + 5/2 &\leq 0 \\ -x_1^2 + x_2 - 2 &= 0. \end{aligned}$$

El punto $\bar{x}$ es factible y cumple condiciones necesarias de KKT, $u_1 = 0$, $v_1 = -2$, luego es candidato a minimizador local. Por otro lado, f es convexa, pero el conjunto factible no es convexo. Las únicas igualdades que dan lugar a conjuntos convexos son las que determinan un hiperplano; en este ejemplo la igualdad $-x_1^2 + x_2 - 2 = 0$, no determina un convexo. Luego el conjunto factible no es convexo, luego no se cumplen las condiciones suficientes y no se puede afirmar que el punto sea maximizador global. Para saber si es minimizador es necesario hacer un estudio diferente. $\diamond$

Ejemplo 3.49. Considere $\bar{x} = (1,3)$ para el siguiente problema:

$$\begin{aligned}\min f(x) &= (x_1-3)^2 + x_2^2\\ x_1^2 - x_2 + 2 &\leq 0\\ 3x_1 - x_2 &= 0.\end{aligned}$$

El punto $\bar{x}$ es factible y cumple condiciones necesarias de KKT, $u_1 = 14$, $v_1 = -8$ f es convexa, y el conjunto factible es convexo, luego $(1,3)$ es minimizador global. $\diamond$

Ejemplo 3.50. Resolver, utilizando condiciones de KKT, el siguiente problema:

$$\begin{aligned}\min f(x) &= (x_1-3)^2 + (x_2-2)^2\\ -x_1 &\leq 0\\ -x_2 + 5/2 &\leq 0\\ -x_1^2 + x_2 - 2 &= 0.\end{aligned}$$

En este ejemplo el conjunto admisible X es no acotado y cerrado (pues es intersección de cerrados), la función f es continua y coercitiva, entonces el PMDI tiene por lo menos un minimizador global. Sea $\bar{x}$ un punto admisible. Como no se sabe exactamente qué punto es, entonces es necesario estudiar todas las posibilidades. Estas se pueden agrupar en función de las posibilidades de $\mathcal{I}$: $\mathcal{I} = \emptyset$, $\mathcal{I} = \{1\}$, $\mathcal{I} = \{2\}$, $\mathcal{I} = \{1,2\}$.

i) $\mathcal{I} = \emptyset$:

$$f'(\bar{x}) + \sum_{i\in\mathcal{I}} u_i g_i'(\bar{x}) + \sum_{j=1}^{s} v_j h_j'(\bar{x}) = \begin{bmatrix} 2(\bar{x}_1-3)\\ 2(\bar{x}_2-2)\end{bmatrix} + v_1 \begin{bmatrix} -2\bar{x}_1\\ 1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}.$$

Como se tiene que cumplir la igualdad, entonces

$$-\bar{x}_1^2 + \bar{x}_2 - 2 = 0\ , \qquad \text{luego } \bar{x}_2 = \bar{x}_1^2 + 2,$$

$$\begin{bmatrix} 2(\bar{x}_1-3)\\ 2(\bar{x}_1^2+2-2)\end{bmatrix} + v_1 \begin{bmatrix} -2\bar{x}_1\\ 1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}.$$

De donde

$$\begin{aligned}2(\bar{x}_1-3) - 2v_1\bar{x}_1 &= 0\\ 2\bar{x}_1^2 + v_1 &= 0.\end{aligned}$$

Reemplazando $v_1 = -2\bar{x}_1^2$,

$$\begin{aligned}2(\bar{x}_1-3) + 4\bar{x}_1^3 &= 0\\ 2(2\bar{x}_1^3 + \bar{x}_1 - 3) &= 0\\ (\bar{x}_1-1)(2\bar{x}_1^2 + 2\bar{x}_1 + 3) &= 0.\end{aligned}$$

El polinomio cuadrático no tiene raíces reales ya que su discriminante vale -20, entonces:

$$\begin{aligned}\bar{x}_1 &= 1\\ \bar{x}_2 &= 3\\ v_1 &= -2.\end{aligned}$$

Entonces la suposición $\mathcal{I} = \emptyset$ da lugar a un sistema consistente, $\bar{x} = (1,3)$, factible. Como no hay desigualdades, $\bar{x}$ es un punto de KKT.

ii) $\mathcal{I} = \{1\}$: como está activa la primera desigualdad y se cumple la igualdad, entonces:

$$\begin{aligned}-\bar{x}_1 &= 0\\ -\bar{x}_1^2 + \bar{x}_2 - 2 &= 0.\end{aligned}$$

Luego $\bar{x}_2 = 2$, es decir, la suposición $\mathcal{I} = \{1\}$ conduce al punto $\bar{x} = (0,2)$ que no es realizable.

iii) $\mathcal{I} = \{2\}$: como está activa la segunda desigualdad y se cumple la igualdad, entonces:

$$\begin{aligned}-\bar{x}_2 + 5/2 &= 0\\ -\bar{x}_1^2 + \bar{x}_2 - 2 &= 0.\end{aligned}$$

Luego $\bar{x}_2 = 5/2$, $\bar{x}_1 = \pm\sqrt{2}/2$, es decir, la suposición $\mathcal{I} = \{2\}$ conduce, o bien al punto $\bar{x} = (-\sqrt{2}/2, 5/2)$ que no es admisible (incumple la primera desigualdad), o bien al punto $\bar{x} = (\sqrt{2}/2, 5/2)$ que si es factible. Para este punto las condiciones de KKT dan:

$$\begin{bmatrix} -4.585786 \\ 1 \end{bmatrix} + u_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} + v_1 \begin{bmatrix} -1.414214 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Entonces $u_2 = -2.2426407$, $v_1 = -3.2426407$, luego $\bar{x}$ no es punto de KKT.

iv) $\mathcal{I} = \{1, 2\}$: como están activas la primera y la segunda desigualdad y se cumple la igualdad, entonces:

$$\begin{aligned} -\bar{x}_1 &= 0 \\ -\bar{x}_2 + 5/2 &= 0 \\ -\bar{x}_1^2 + \bar{x}_2 - 2 &= 0. \end{aligned}$$

Pero no existe ningún punto que cumpla las tres igualdades.

En resumen, hay un solo punto que cumple condiciones necesarias de KKT, y como existe por lo menos un minimizador global, entonces este punto $\bar{x} = (1, 3)$ debe ser el minimizador global. $\diamond$

3.5 Condiciones de segundo orden

Así como para puntos interiores hay condiciones necesarias (... y hessiana semidefinida positiva) y condiciones suficientes de segundo orden (... y hessiana definida positiva), es decir, que utilizan derivadas parciales de segundo orden, también hay condiciones de segundo orden para puntos no interiores. La diferencia principal consiste en que no se exige que la hessiana sea semidefinida positiva en todo $\mathbb{R}^n$, sino en un subespacio o en un subconjunto.

3.5.1 Generalidades

Si $\bar{x}$ es un punto de KKT y se trata de un problema de optimización convexa, entonces $\bar{x}$ es minimizador global. No se necesita estudiar las condiciones de segundo orden.

Las condiciones necesarias de segundo orden y las condiciones suficientes de segundo orden se aplican a puntos que cumplen condiciones necesarias de KKT, es decir, para puntos factibles para los que se puede aplicar el teorema de KKT (continuidad, diferenciabilidad, regularidad, ...) y además cumplen las condiciones de KKT. Se requiere también doble diferenciabilidad.

En resumen los requerimientos para poder usar los teoremas de condiciones necesarias de segundo orden o de condiciones suficientes de segundo orden son:

$$\bar{x} \text{ es punto de KKT, } \bar{u},\ \bar{v} \text{ son sus coeficientes} \tag{3.23a}$$

$$f, \quad g_i, i \in \mathcal{I}, \quad h_j, \forall j \quad \text{son doblemente diferenciables} \tag{3.23b}$$

Para un $\bar{x}$ que cumple lo anterior, se define el hessiano del lagrangiano restringido a las variables x así:

$$\mathcal{L}''_x = \mathcal{L}''_x(\bar{x}, \bar{u}, \bar{v}) = f''(\bar{x}) + \sum_{i \in \mathcal{I}} \bar{u}_i g''_i(\bar{x}) + \sum_{j=1}^{s} \bar{v}_j h''_j(\bar{x}) \tag{3.24}$$

Para $\bar{x}$, un punto de KKT, una desigualdad activa, $g_i(x) \leq 0$, es **fuertemente activa** si $\bar{u}_i > 0$. La desigualdad es **débilmente activa** si $\bar{u}_i = 0$. Conjuntos de índices de desigualdades activas, fuertemente activas, débilmente activas:

$$\mathcal{I}(\bar{x}) = \mathcal{I} \quad = \{i : g_i(\bar{x}) = 0\} \tag{3.25}$$

$$\mathcal{I}^+(\bar{x}) = \mathcal{I}^+ = \{i \in \mathcal{I} : \bar{u}_i > 0\} \tag{3.26}$$

$$\mathcal{I}^0(\bar{x}) = \mathcal{I}^0 \ = \{i \in \mathcal{I} : \bar{u}_i = 0\} \tag{3.27}$$

Sea M la matriz cuyas filas son los gradientes de las desigualdades activas y los gradientes de las igualdades. Sea $\tilde{M}$ la matriz cuyas filas son los gradientes de las desigualdades fuertemente activas y los gradientes de las igualdades. Sea D la matriz cuyas filas son los gradientes de las desigualdades débilmente activas.

$$M = \begin{bmatrix} g_i'(\bar{x})^{\mathrm{T}}, \ i \in \mathcal{I} \\ h_j'(\bar{x})^{\mathrm{T}}, \ \forall j \end{bmatrix} \tag{3.28}$$

$$\tilde{M} = \begin{bmatrix} g_i'(\bar{x})^{\mathrm{T}}, \ i \in \mathcal{I}^+ \\ h_j'(\bar{x})^{\mathrm{T}}, \ \forall j \end{bmatrix} \tag{3.29}$$

$$D = \begin{bmatrix} g_i'(\bar{x})^{\mathrm{T}}, \ i \in \mathcal{I}^0 \end{bmatrix} \tag{3.30}$$

El espacio tangente $\mathcal{T}$ es el espacio nulo de la matriz M. El espacio $\tilde{\mathcal{T}}$ es el espacio nulo de la matriz $\tilde{M}$. El cono C es el conjunto definido por $Dx \leq 0$.

$$\mathcal{T} = N_M \tag{3.31}$$

$$\tilde{\mathcal{T}} = N_{\tilde{M}} \tag{3.32}$$

$$C = \{x \in \mathbb{R}^{n\times 1} : Dx \leq 0\} \tag{3.33}$$

Los tres conjuntos son conos convexos, los dos primeros también son subespacios vectoriales. Si una de esas matrices no tiene filas, obviamente el conjunto correspondiente es todo $\mathbb{R}^n$.

El significado de $\mathcal{T}$ puede tener la siguiente interpretación geométrica en $\mathbb{R}^3$ (o en $\mathbb{R}^2$): Si $\bar{x}$ está en la frontera de X, y si se puede construir un plano tangente a la superficie en $\bar{x}$ o una única recta tangente, el espacio tangente es un plano (o una recta) que pasa por el origen y es paralelo al plano tangente (o a la recta tangente).

Ejemplo 3.51. Sean: X en $\mathbb{R}^2$ definido por dos desigualdades

$$\begin{aligned} (x_1+1)^2 + x_2^2 - 25 &\leq 0 \\ -x_1 - x_2 &\leq 0, \end{aligned}$$

$\bar{x} = (2,4)$. El conjunto X es un pedazo de círculo. La recta tangente a X en el punto $\bar{x} = (2,4)$ es

$$R = \{(y_1,y_2) : 3y_1 + 4y_2 = 22\}.$$

El espacio tangente es la recta paralela a R que pasa por el origen:

$$\mathcal{T} = \{(y_1,y_2) : 3y_1 + 4y_2 = 0\}.$$

Si se considera $\bar{x} = (1,3)$, el espacio tangente es todo $\mathbb{R}^2$. Si se considera $\bar{x} = (-4,4)$, no se puede construir "la recta tangente" a X en $(-4,4)$. El espacio tangente es $\mathcal{T} = \{(0,0)\}$. $\diamond$

Ejemplo 3.52. Sean: X en $\mathbb{R}^3$ definido por

$$x_1^2 + x_2^2 + x_3^2 - 20 \leq 0,$$

y el punto $\bar{x} = (2,0,4)$. Por razonamientos geométricos, el plano tangente a X en $\bar{x}$ es

$$P = \{(y_1,y_2,y_3) : y_1 + 2y_3 = 10\}.$$

El espacio tangente es el plano paralelo a P que pasa por el origen:

$$\mathcal{T} = \{(y_1,y_2,y_3) : y_1 + 2y_3 = 0\}. \ \diamond$$

Se puede obtener directamente a partir de la matriz M

$$M = \begin{bmatrix} 4 & 0 & 8 \end{bmatrix}$$

Ejemplo 3.53. Sean: X en $\mathbb{R}^3$ definido por

$$\begin{aligned} x_1^2 + x_2^2 + x_3^2 - 20 &\leq 0, \\ x_1 + 2x_2 + 3x_3 - 14 &= 0, \end{aligned}$$

y el punto $\bar{x} = (2,0,4)$. Por razonamientos geométricos, la recta tangente a X, en $\bar{x}$ es

$$R = \{(2,0,4) + t(-4,-1,2) : t \in \mathbb{R}\}.$$

El espacio tangente es la recta paralela a R que pasa por el origen:

$$\mathcal{T} \quad = \quad \{t(-4,-1,2) : t \in \mathbb{R}\}.$$

Obtenido a partir de M:

$$M = \begin{bmatrix} 4 & 0 & 8 \\ 1 & 2 & 3 \end{bmatrix}$$

Su escalonada reducida es

$$E_M = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1/2 \end{bmatrix}$$

Una base de N_M está formada por un vector, por ejemplo,

$$v^1 = \begin{bmatrix} -2 \\ -1/2 \\ 1 \end{bmatrix}$$

Obviamente

$$\text{gen}(v^1) = \mathcal{T} \quad \diamond$$

Se esta suponiendo que $\bar{x}$ es regular, entonces las filas de M y de $\tilde{M}$ son linealmente independientes. Si

$$\begin{aligned} \text{nf}(M) = n, &\ \text{ entonces } \mathcal{T} = \{\mathbf{0}\} \\ \text{nf}(\tilde{M}) = n, &\ \text{ entonces } \tilde{\mathcal{T}} = \{\mathbf{0}\} \end{aligned}$$

Recuérdese que cualquier matriz simétrica es definida positiva, semidefinida positiva, definida negativa, semidefinida negativa en $\{\mathbf{0}\}$.

Para un punto de KKT, el **cono crítico** $K = K(\bar{x})$ es:

$$\begin{aligned} K &= \{d \in \mathbb{R}^n : \ g_i'(\bar{x})^{\mathrm{T}} d = 0, i \in \mathcal{I}^+, \ g_i'(\bar{x})^{\mathrm{T}} d \le 0, i \in \mathcal{I}^0, \ h_j'(\bar{x})^{\mathrm{T}} d = 0, \forall j\} \\ K &= \tilde{\mathcal{T}} \cap C \end{aligned} \tag{3.34}$$

Por definición,

$$\begin{aligned} &\mathcal{T} \subseteq \tilde{\mathcal{T}} \\ &K \subseteq \tilde{\mathcal{T}} \\ &K \subseteq C \\ &\mathcal{T} \subseteq K \subseteq \tilde{\mathcal{T}} \end{aligned}$$

Obviamente, si no hay desigualdades débilmente activas, entonces $M = \tilde{M}$, $\mathcal{T} = \tilde{\mathcal{T}}$, $\quad K = \mathcal{T} = \tilde{\mathcal{T}}$.

Si hay desigualdades débilmente activas,

$$\mathcal{T} \subsetneq K \subsetneq \tilde{\mathcal{T}} \tag{3.35}$$

A manera de ilustración, en $\mathbb{R}^3$, si hay dos desigualdades activas y no hay igualdades, entonces M tiene dos filas independientes, $\mathcal{T}$ es una recta que pasa por el origen. Si una de las desigualdades es débilmente activa, $\tilde{M}$ tiene una sola fila, $\tilde{\mathcal{T}}$ es un plano que pasa por el origen, C es un semiespacio cuya frontera (un plano) pasa por el origen y K es un semiplano (contenido en $\tilde{\mathcal{T}}$) cuya frontera, $\mathcal{T}$, pasa por el origen.

3.5.2 Condiciones de segundo orden, $\mathcal{I}^0 = \emptyset$

Proposición 3.18. ***Condiciones necesarias de segundo orden, CN2O***. *Sea $\bar{x}$ que cumple (3.23), $\mathcal{I}^0 = \emptyset$. Si $\bar{x}$ es minimizador local, entonces*

$$\mathcal{L}''_x \text{ es semidefinida positiva en } \mathcal{T}. \tag{3.36}$$

Proposición 3.19. ***Condiciones suficientes de segundo orden, CS2O***. *Sea $\bar{x}$ que cumple* (3.23), $\mathcal{I}^0 = \emptyset$. *Si*

$$\mathcal{L}''_x \text{ es definida positiva en } \mathcal{T}, \tag{3.37}$$

entonces $\bar{x}$ es minimizador local estricto.

Obviamente, si $\mathcal{L}''_x$ es semidefinida positiva (en todo $\mathbb{R}^n$), entonces $\bar{x}$ cumple CN2O.

Así mismo, si $\mathcal{L}''_x$ es definida positiva, entonces $\bar{x}$ cumple CS2O.

Si $\bar{x}$ cumple CN2O y no cumple CS2O, sin usar otros razonamientos, no se puede decir si $\bar{x}$ es o no es minimizador local.

3.5.3 Condiciones de segundo orden, $\mathcal{I}^0 \neq \emptyset$

Proposición 3.20. ***Condiciones necesarias de segundo orden débiles, CN2D***. *Sea $\bar{x}$ que cumple* (3.23). *Si $\bar{x}$ es minimizador local, entonces*

$$\mathcal{L}''_x \text{ es semidefinida positiva en } \mathcal{T} \tag{3.38}$$

Proposición 3.21. ***Condiciones necesarias de segundo orden fuertes, CN2F***. *Sea $\bar{x}$ que cumple* (3.23). *Si $\bar{x}$ es minimizador local, entonces*

$$\mathcal{L}''_x \text{ es semidefinida positiva en } K. \tag{3.39}$$

Proposición 3.22. ***Condiciones suficientes de segundo orden débiles, CS2D***. *Sea $\bar{x}$ que cumple* (3.23). *Si*

$$\mathcal{L}''_x \text{ es definida positiva en } K. \tag{3.40}$$

entonces $\bar{x}$ es minimizador local estricto.

Proposición 3.23. ***Condiciones suficientes de segundo orden fuertes, CS2F***. *Sea $\bar{x}$ que cumple* (3.23). *Si*

$$\mathcal{L}''_x \text{ es definida positiva en } \tilde{\mathcal{T}}. \tag{3.41}$$

entonces $\bar{x}$ es minimizador local estricto.

De nuevo, si $\mathcal{L}''_x$ es semidefinida positiva (en todo $\mathbb{R}^n$), entonces $\bar{x}$ cumple CN2D y CN2F.

Así mismo, si $\mathcal{L}''_x$ es definida positiva, entonces $\bar{x}$ cumple CS2D y CS2F.

Lo ideal sería tener condiciones de segundo orden que fueran al mismo tiempo necesarias y suficientes. Ninguna de las anteriores lo es. Tampoco hay condiciones necesarias y suficientes de aplicación relativamente fácil.

Dado lo anterior, lo deseable es tener condiciones necesarias de segundo orden fuertes y condiciones suficientes de segundo orden débiles. Así el espacio entre las necesarias y las suficientes es más pequeño.

Si $\bar{x}$ cumple CN2F y no cumple CS2D, sin usar otros razonamientos, no se puede decir si $\bar{x}$ es o no es minimizador local.

El teorema CN2F es más fuerte que CN2D. Entonces nace la pregunta, ¿para qué sirve el teorema CN2D? El interés de presentar los dos teoremas es simplemente que la aplicación del teorema CN2D es más sencilla.

De manera análoga, el teorema CS2D es más débil que CS2F. Entonces nace la pregunta, ¿para qué sirve el teorema CS2F? El interés de presentar los dos teoremas es simplemente que la aplicación del teorema CS2F es más sencilla.

Sea $\bar{x}$ un punto de KKT con $\mathcal{I}^0 \neq \emptyset$. El proceso para tratar de saber si es o no es minimizador puede ser el siguiente:

i. Si $\bar{x}$ no cumple CN2D, entonces $\bar{x}$ no es minimizador local.

ii. Si $\bar{x}$ cumple CN2D pero no cumple CN2F, entonces $\bar{x}$ no es minimizador local.

iii. Si $\bar{x}$ cumple CS2F, entonces $\bar{x}$ es minimizador local estricto.

iv. Si $\bar{x}$ no cumple CS2F pero cumple CS2D, entonces $\bar{x}$ es minimizador local estricto.

v. Si $\bar{x}$ cumple CN2F y no cumple CS2D, entonces, a partir de lo estudiado, no se sabe si $\bar{x}$ es o no es minimizador local.

Dado que las condiciones de más fácil aplicación son CN2D y CS2F, el proceso puede ser:

1. Si $\bar{x}$ no cumple CN2D, entonces $\bar{x}$ no es minimizador local.
2. Si $\bar{x}$ cumple CS2F, entonces $\bar{x}$ es minimizador local estricto.
3. Si $\bar{x}$ cumple CN2D, pero no cumple CS2F, entonces, a partir de lo estudiado, no se sabe si $\bar{x}$ es o no es minimizador local.

Ejemplo 3.54. Considere $\bar{x} = (0, 0, 0)$ para el siguiente problema:

$$\begin{aligned} \min f(x_1, x_2, x_3) = x_1^2 - x_2^2 + \frac{1}{3}(x_3 - 1)^3 & \\ -6 - x_1 - x_2 - x_3 &\leq 0 \\ -x_3 &\leq 0. \end{aligned}$$

El punto $\bar{x}$ es factible; $\mathcal{I} = \{2\}$; f, g_2 son diferenciables en $\bar{x}$; g_1 es continua en $\bar{x}$; el conjunto formado por el gradiente $g_2'(\bar{x}) = [0\ \ 0\ \ -1]^{\mathrm{T}}$ es linealmente independiente, o sea, $\bar{x}$ es regular.

$$\begin{aligned} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + u_2 \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \\ u_2 &= 1, \end{aligned}$$

luego $\bar{x} = (0, 0, 0)$ es un punto de KKT.

$$\begin{aligned} \mathcal{L}_x'' &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} + 1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} \\ \mathcal{T} &= \{(y_1, y_2, y_3) :\ -y_3 = 0\}. \end{aligned}$$

$\mathcal{L}_x''$ no es semidefinida positiva, pero podría ser semidefinida positiva en $\mathcal{T}$. Una base de este subespacio puede ser el conjunto formado por los vectores $[1\ \ 0\ \ 0]^{\mathrm{T}}$, $[0\ \ 1\ \ 0]^{\mathrm{T}}$, luego

$$\begin{aligned} E &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \\ E^{\mathrm{T}} \mathcal{L}_x'' E &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \quad \text{no es semidefinida positiva,} \end{aligned}$$

luego $\bar{x}$ no cumple CN20, entonces no es minimizador local ni global. ◇

Ejemplo 3.55. Considere $\bar{x} = (0, 3, 0)$ para el siguiente problema:

$$\begin{aligned} \min f(x_1, x_2, x_3) = x_1^2 - x_2^2 + x_3^2 & \\ -x_2 &\leq 0 \\ x_2 - 3 &\leq 0. \end{aligned}$$

El punto $\bar{x}$ es factible; $\mathcal{I} = \{2\}$; f, g_2 son diferenciables en $\bar{x}$; g_1 es continua en $\bar{x}$; el conjunto formado por el gradiente $g_2'(\bar{x}) = [0 \;\; 1 \;\; 0]^{\mathrm{T}}$ es linealmente independiente, o sea, $\bar{x}$ es regular.

$$\begin{bmatrix} 0 \\ -6 \\ 0 \end{bmatrix} + u_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$
$$u_2 = 6.$$

Luego $\bar{x} = (0, 3, 0)$ es un punto de KKT.

$$\begin{aligned} \mathcal{L}_x'' &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{bmatrix} + 6 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \\ \mathcal{T} &= \{(y_1, y_2, y_3) : y_2 = 0\}. \end{aligned}$$

$\mathcal{L}_x''$ no es semidefinida positiva, pero podría ser semidefinida positiva en $\mathcal{T}$. Una base de este subespacio puede ser el conjunto formado por los vectores $[1 \;\; 0 \;\; 0]^{\mathrm{T}}$, $[0 \;\; 0 \;\; 1]^{\mathrm{T}}$, luego

$$\begin{aligned} E &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \\ E^{\mathrm{T}} \mathcal{L}_x'' E &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{es definida positiva.} \end{aligned}$$

Entonces $\bar{x}$ cumple condiciones necesarias de segundo orden, CN2O. Más aún, no hay desigualdades débilmente activas, entonces $\tilde{M} = M$, $\tilde{\mathcal{T}} = \mathcal{T}$. En consecuencia, $\mathcal{L}_x''$ es definida positiva en $\tilde{\mathcal{T}}$, es decir, $\bar{x} = (0, 3, 0)$ cumple condiciones suficientes CS2O y es minimizador local estricto. $\diamond$

Ejemplo 3.56. Considere $\bar{x} = (0, 0)$ para el siguiente problema:

$$\begin{aligned} \min f(x_1, x_2) = x_1^2 - x_2^2 & \\ -x_1 - x_2 &\leq 0. \end{aligned}$$

El punto $\bar{x}$ es factible; $\mathcal{I} = \{1\}$; f, g_1 son diferenciables en $\bar{x}$; el conjunto formado por el gradiente $g_1'(\bar{x}) = [-1 \;\; -1]^{\mathrm{T}}$ es linealmente independiente, o sea, $\bar{x}$ es regular.

$$\begin{aligned} \begin{bmatrix} 0 \\ 0 \end{bmatrix} + u_1 \begin{bmatrix} -1 \\ -1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ u_1 &= 0, \end{aligned}$$

luego $\bar{x} = (0, 0)$ es un punto de KKT.

$$\begin{aligned} \mathcal{L}_x'' &= \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} + 0 \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \\ \mathcal{T} &= \{(d_1, d_2) : -d_1 - d_2 = 0\}. \end{aligned}$$

$\mathcal{L}_x''$ no es semidefinida positiva, pero podría ser semidefinida positiva en $\mathcal{T}$. Una base de este subespacio puede ser el conjunto formado por el vector $[1 \;\; -1]^{\mathrm{T}}$, luego

$$\begin{aligned} E &= \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ E^{\mathrm{T}} \mathcal{L}_x'' E &= \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ &= 0, \; \text{matriz semidefinida positiva.} \end{aligned}$$

Luego $\bar{x}$ cumple CN2D. La desigualdad activa es débilmente activa, luego $\bar{m}^+ = 0$, $p^+ = 0$, $\tilde{M}$ no tiene filas, $\tilde{\mathcal{T}} = \mathbb{R}^2$. Como $\mathcal{L}''_x$ no es definida positiva en $\mathbb{R}^2$, entonces $\bar{x}$ no cumple CS2F. En resumen, utilizando lo anterior, el camino fácil, se puede decir que $\bar{x}$ es un buen candidato a minimizador local.

Sin embargo, realmente no es minimizador local ya que un punto de la forma $(0, \varepsilon)$, con $\varepsilon > 0$, es admisible, mejor que $\bar{x}$ y está muy cerca de él.

Las condiciones necesarias fuertes, CN2F, en este caso muy sencillo, dan más información:

$$
\begin{aligned}
\tilde{\mathcal{T}} &= \mathbb{R}^2 \\
C &= \{(d_1, d_2) : -d_1 - d_2 \le 0\} \\
K &= C \\
d &= (0, 1) \in K \\
d^{\mathrm{T}} \mathcal{L}''_x d &= -2
\end{aligned}
$$

Es decir, $\mathcal{L}''_x$ no es semidefinida positiva en K, luego $\bar{x}$ no cumple CN2F, luego no es minimizador local. Otro razonamineto alterno es el siguiente, una matriz indefinida nunca es semidefinida positiva en un semiespacio. Ver apéndice. $\diamond$

Ejemplo 3.57.

$$
\begin{aligned}
\min \;\; f(x_1, x_2) = x_1^2 - 8x_1 - x_2^2 + 6x_2 & \\
-3x_1 - 4x_2 + 19 &\le 0 \\
2x_1 + x_2 - 11 &\le 0 \\
x_1 + 3x_2 - 13 &\le 0
\end{aligned}
$$

Hay 8 casos posibles para $\mathcal{I}$: $\emptyset$, $\{1\}$, $\{2\}$, $\{3\}$, $\{1,2\}$, $\{1,3\}$, $\{2,3\}$, $\{1,2,3\}$.

$\mathcal{I} = \emptyset$: $x = (4, 3)$ factible. Para este punto $\mathcal{I} = \{2, 3\}$.

$\mathcal{I} = \{1\}$: $x = (43/7, 1/7)$ no factible.

$\mathcal{I} = \{2\}$: $x = (4, 3)$ factible, $u_2 = 0$, punto de KKT. Para este punto $\mathcal{I} = \{2, 3\}$.

$\mathcal{I} = \{3\}$: $x = (4, 3)$ factible, $u_3 = 0$, punto de KKT. Para este punto $\mathcal{I} = \{2, 3\}$.

$\mathcal{I} = \{1, 2\}$: $x = (5, 1)$ factible, $u_1 = 6/5$, $u_2 = 4/5$, punto de KKT. Condiciones de segundo orden:

$$
\begin{aligned}
\mathcal{I}^0 &= \emptyset \\
\mathcal{L}''_x &= \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \quad \text{no es sdp} \\
M &= \begin{bmatrix} -3 & -4 \\ 2 & 1 \end{bmatrix} \\
N_M &= \{\mathbf{0}\} \\
\mathcal{L}''_x & \quad \text{es sdp en } N_M, \quad x \text{ cumple CN20} \\
\tilde{M} &= M \\
\mathcal{L}''_x & \quad \text{es dp en } N_M, \quad x \text{ cumple CS20} \\
& \quad x \text{ es minimizador local estricto.}
\end{aligned}
$$

$\mathcal{I} = \{1, 3\}$: $x = (1, 4)$ factible, $u_1 = -16/5$, $u_3 = -18/5$, x no es minimizador local.

$\mathcal{I} = \{2,3\}$: $x = (4,3)$ factible, $u_2 = 0$, $u_3 = 0$, x es punto de KKT. Condiciones de segundo orden:

$$\begin{aligned}
\mathcal{I}^0 &= \{1,2\} \\
\mathcal{L}''_x &= \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \quad \text{no es sdp} \\
M &= \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \\
\tilde{M} &= [\ ] \\
D &= \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}
\end{aligned}$$

Salvo equivalencia, las direcciones extremas de $K = C$ son $d^1 = (1,-2)$ y $d^2 = (-3,1)$.

$$\begin{aligned}
{d^1}^{\mathrm{T}} \mathcal{L}''_x d^1 &= -6 \\
\mathcal{L}''_x &\ \text{ no es sdp en } \ K \\
x = (4,3) &\ \text{ no es minimzador local.}
\end{aligned}$$

$\mathcal{I} = \{1,2,3\}$: el sistema de ecuaciones no tiene solución.

En resumen, hay un minimizador local. La región factible es un triángulo cuyos vértices son $(5,1)$, $(1,4)$, $(4,3)$, conjunto cerrado y acotado, luego hay minimizador global, necesariamente el único minimizador local : $(5,1)$. $\diamond$

Ejemplo 3.58. Problema del productor con función de producción de Cobb-Douglas:

$$\begin{aligned}
Q &= A x_1^{\alpha} x_2^{\beta} \\
x_1 = K &= \text{ capital} \\
x_2 = L &= \text{ trabajo} \\
\alpha &> 0 \\
\beta &> 0
\end{aligned}$$

Los valores α y β son valores fijos que dependen de la tecnología disponible. Q es la producción total de un sólo producto. Sean r y w los precios unitarios del capital y del trabajo y sea q el nivel de producción deseado. El problema del productor es:

$$\begin{aligned}
\min\ f(x_1,x_2) &= r x_1 + w x_2 \\
A x_1^{\alpha} x_2^{\beta} &= Q \\
x_1 &> 0 \\
x_2 &> 0
\end{aligned} \tag{3.42}$$

Las restricciones $x_i > 0$ no son como las desigualdades propias de un problema con desigualdades e igualdades. Solamente serán consideradas cuando sea indispensable.

Consideremos un caso particular:

$$\begin{aligned}
\min\ f(x_1,x_2) &= 3x_1 + 5x_2 \\
x_1 x_2 - 10 &= 0 \\
x_1 &> 0 \\
x_2 &> 0
\end{aligned}$$

Al plantear condiciones de KKT (o de Lagrange, solamente igualdades) y factibilidad

$$\begin{bmatrix} 3 \\ 5 \end{bmatrix} + v_1 \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$x_1 x_2 - 10 = 0$$

Entonces

$$
\begin{aligned}
v_1 &= -3/x_2\\
v_1 &= -5/x_1\\
-3/x_2 &= -5/x_1\\
x_2 &= 3x_1/5\\
x_1\,\frac{3}{5}\,x_1 &= 10\\
x_1 &= \sqrt{50/3} \approx 4.08\\
x_2 &= \sqrt{6} \approx 2.45\\
v_1 &= -\sqrt{3/2} \approx -1.22
\end{aligned}
$$

El punto $\bar{x} \approx (4.08, 2.45)$ cumple condiciones necesarias de primer orden para ser minimizador local. La función f es convexa pero el conjunto factible no es convexo, o sea, no es un problema de optimización convexa (luego no se puede deducir que sea minimizador global).

Condiciones de segundo orden:

$$
\begin{aligned}
\mathcal{L}''_x &= \begin{bmatrix} 0 & 0\\ 0 & 0 \end{bmatrix} - 1.22 \begin{bmatrix} 0 & 1\\ 1 & 0 \end{bmatrix}\\
\mathcal{L}''_x &= \begin{bmatrix} 0 & -1.22\\ -1.22 & 0 \end{bmatrix}
\end{aligned}
$$

$\mathcal{L}''_x$ no es semidefinida positiva pero podría serlo en $\mathcal{T}$:

$$
\begin{aligned}
M &= [2.45 \;\; 4.08]\\
E &= \begin{bmatrix} -5\\ 3 \end{bmatrix}\\
E^{\mathrm{T}}\mathcal{L}''_x E &\approx 36.74
\end{aligned}
$$

Luego $\mathcal{L}''_x$ es semifefinida postiva en $\mathcal{T}$, es, decir, $\bar{x}$ cumple condiciones necesarias de segundo orden. Además, no hay desigualdades débilmente activas, luego $\tilde{M} = M$, $\tilde{\mathcal{T}} = \mathcal{T}$, entonces $\bar{x}$ cumple condiciones suficientes de segundo orden CS20, luego es minimizador local estricto.

En este caso "sencillo" se puede analizar el problema sin tener en cuenta condiciones de segundo orden. El conjunto factible X es cerrado no acotado, f es continua y coercitiva en X luego hay un minimizador global. Solamente hay un punto que podría ser minimizador local, el $\bar{x}$ obtenido, luego necesariamente es minimizador global. $\diamond$

Ejemplo 3.59. Pequeña generalización:

$$
\begin{aligned}
\min\;\; f(x_1, x_2) = rx_1 + wx_2&\\
x_1x_2 = q&\\
x_1 > 0&\\
x_2 > 0&
\end{aligned}
\tag{3.43}
$$

Se supone que los datos r, w y q son positivos. Es una generalización del ejemplo anterior.

$$
\begin{aligned}
\begin{bmatrix} r\\ w \end{bmatrix} + v_1 \begin{bmatrix} x_2\\ x_1 \end{bmatrix} &= \begin{bmatrix} 0\\ 0 \end{bmatrix}\\
x_1x_2 - q &= 0
\end{aligned}
$$

Entonces

$$\begin{aligned} v_1 &= -r/x_2 \\ v_1 &= -w/x_1 \\ -r/x_2 &= -w/x_1 \\ x_2 &= rx_1/w \\ x_1 \frac{r}{w} x_1 &= q \\ \bar{x}_1 &= \sqrt{wq/r} \\ \bar{x}_2 &= \sqrt{rq/w} \\ v_1 &= -\sqrt{rw/q} \end{aligned}$$

Condiciones de segundo orden

$$\mathcal{L}''_x = \begin{bmatrix} 0 & v_1 \\ v_1 & 0 \end{bmatrix}$$

no es semidefinida positiva

$$M = \begin{bmatrix} \bar{x}_2 & \bar{x}_1 \end{bmatrix}$$

Base de N_M:

$$\begin{aligned} E &= \begin{bmatrix} -\bar{x}_1/\bar{x}_2 \\ 1 \end{bmatrix} = \begin{bmatrix} -w/r \\ 1 \end{bmatrix} \sim \begin{bmatrix} -w \\ r \end{bmatrix} \\ E^{\mathrm{T}} \mathcal{L}''_x E &= -2v_1 rw = 2\sqrt{\frac{r^3 w^3}{q}} \end{aligned}$$

Luego $\mathcal{L}''_x$ es semidefinida positiva en $\mathcal{T}$. No hay desigualdades débilmente activas, entonces $\mathcal{L}''_x$ es definida positiva en $\tilde{\mathcal{T}}$. Luego $\bar{x}$ es minimizador local.

El razonamiento del ejemplo anterior sobre coercitividad es el mismo para este ejemplo, luego $\bar{x}$ es minimizador global. ◇

Ejemplo 3.60.

$$\begin{aligned} \min \ f(x_1, x_2) &= rx_1 + wx_2 \\ x_1^{\alpha} x_2 &= q \\ x_1 &> 0 \\ x_2 &> 0 \end{aligned} \tag{3.44}$$

Se supone que los datos r, w, α y q son positivos.

$$\begin{aligned} \begin{bmatrix} r \\ w \end{bmatrix} + v_1 \begin{bmatrix} \alpha x_1^{\alpha-1} x_2 \\ x_1^{\alpha} \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ x_1 x_2 - q &= 0 \end{aligned}$$

Entonces ???? ....

Ejemplo 3.61.

$$\begin{aligned} \min \ f(x_1, x_2) &= rx_1 + wx_2 \\ x_1^{\alpha} x_2^{\beta} &= q \\ x_1 &> 0 \\ x_2 &> 0 \end{aligned} \tag{3.45}$$

Se supone que los datos r, w, α, β y q son positivos. Este problema es equivalente al problema

$$\begin{aligned} \min \ f(x_1, x_2) = rx_1 + wx_2 \\ x_1^{\alpha/\beta} x_2 = q^{1/\beta} \\ x_1 > 0 \\ x_2 > 0 \end{aligned} \tag{3.46}$$

que es el ejemplo anterior. Si la restricción es $Ax_1^{\alpha} x_2^{\beta} = Q$, esta es equivalente a $x_1^{\alpha} x_2^{\beta} = Q/A$,

EJERCICIOS

En los ejercicios 1 a 10 estudie el problema propuesto, use condiciones necesarias, suficientes y otros argumentos. Encuentre, si es posible, un punto crítico, los puntos críticos, un minimizador, los minimizadores. ¿Son estos puntos minimizadores globales?

3.1. Minimizar $f(x_1, x_2, x_3) = x_2^2 + 4x_2x_3 + 3x_3^2 - 6x_2 - 10x_3 + 100$ en $\mathbb{R}^3$.

3.2. Minimizar $f(x_1, x_2) = \operatorname{sen}(x_1x_2)$ en $\mathbb{R}^2$.

3.3. Minimizar $f(x_1, x_2) = (x_1 - 1)^2 + e^{x_1} + x_2^2 - 2x_2 + 1$ en $\mathbb{R}^2$.

3.4. Minimizar $f(x_1, x_2) = x_1^3 + 3x_1^2 + x_2^2 + 3x_1 - 2x_2 + 2$ en $\mathbb{R}^2$.

3.5. Minimizar $f(x_1, x_2) = x_1^4 + 4x_1^3 + 6x_1^2 + x_2^2 + 4x_1 - 2x_2 + 2$ en $\mathbb{R}^2$.

3.6. Minimizar $f(x_1, x_2) = 11x_1^2 + 11x_2^2 - 18x_1x_2 + 4x_1 + 4x_2 + 142$ en $\mathbb{R}^2$.

3.7. Minimizar $f(x_1, x_2) = (x_1 + x_2)/(x_1^2 + x_2^2 + 1)$ en $\mathbb{R}^2$.

3.8. Minimizar $f(x_1, x_2) = (x_1^2 - x_2)^2$ en $\mathbb{R}^2$.

3.9. Minimizar $f(x_1, x_2, x_3) = 2x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_2x_3 - 6x_1 - 7x_2 - 8x_3 + 20$ en $\mathbb{R}^3$.

3.10. Minimizar $f(x_1, x_2) = x_1^2 + 2x_1x_2 - 10x_1 + 5x_2 - 8x_3 + 20$ en $\mathbb{R}^2$.

3.11. Sea x^* minimizador global de f en A. Sea B subconjunto de A. Dé condiciones sobre B (por ejemplo, suficientes), para que exista minimizador global de f en B.

En los ejercicios 12 a 26 estudie el problema propuesto, use condiciones necesarias, suficientes, condiciones de segundo orden y otros argumentos. Si es posible, estudie todos los casos para $\mathcal{I}$. Encuentre, si es posible, puntos de KKT. ¿Son estos puntos minimizadores? ¿Son minimizadores globales?

3.12. Minimizar $f(x_1, x_2) = x_1^2 + x_2^2$, sujeto a $3x_1 + 4x_2 = 12$.

3.13. Minimizar $f(x_1, x_2) = x_1^2 + x_2^2$, sujeto a $3x_1 + 4x_2 \geq 12$.

3.14. Minimizar $f(x_1, x_2) = 2x_1^2 + x_2^2$, sujeto a $3x_1 + 4x_2 = 12$.

3.15. Minimizar $f(x_1, x_2) = x_1^2 + x_2^2$, sujeto a $3x_1 + 4x_2 = 12$, $x_1 \leq 0$.

3.16. Minimizar $f(x_1, x_2) = x_1^2 + x_2^2$, sujeto a $3x_1 + 4x_2 = 12$, $x_1 \leq 0$, $x_2 \geq 3$.

3.17. Minimizar $f(x_1, x_2) = x_1^2 - x_2^2$, sujeto a $3x_1 + 4x_2 = 12$.

3.18. Minimizar $f(x_1, x_2) = x_1^2 - x_2^2$, sujeto a $3x_1 + 4x_2 \geq 12$.

3.19. Minimizar $f(x_1, x_2) = x_1^2 - x_2^2$, sujeto a $3x_1 + 4x_2 \leq 12$, $x \geq 0$.

3.20. Minimizar $f(x_1, x_2) = x_1^2 - x_2^2$, sujeto a $3x_1 + 4x_2 \geq 12$, $x \geq 0$.

3.21. Minimizar $f(x) = c^{\mathrm{T}}x$, sujeto a $||x - a||_2 = r$, con $c \neq 0$, $r > 0$.

3.22. Minimizar $f(x) = c^{\mathrm{T}}x$, sujeto a $||x - a||_2 \leq r$, con $c \neq 0$, $r > 0$.

3.23. Minimizar $f(x_1, x_2) = x_1^3 + x_2^2$, sujeto a $3x_1 + 4x_2 = 12$.

3.24. Minimizar $f(x_1, x_2) = x_1^3 + x_2^2$, sujeto a $3x_1 + 4x_2 \geq 12$.

3.25. Minimizar $f(x_1, x_2) = x_1^3 + x_2^2$, sujeto a $3x_1 + 4x_2 \leq 12$, $x \geq 0$.

3.26. Minimizar $f(x_1, x_2) = x_1^3 + x_2^2$, sujeto a $3x_1 + 4x_2 \geq 12$, $x \geq 0$.

Capítulo 4

Sistemas de ecuaciones diferenciales lineales

Realmente el tema principal de este capítulo es sistemas de ecuaciones diferenciales lineales con coeficientes contantes, nombre un poco largo para título del capítulo.

4.1 Introducción

Usualmente una ecuación diferencial ordinaria de primer orden se escribe de la forma

$$y' = f(x, y)$$

donde x es la variable independiente, $y = y(x)$ es la variable que depende de x. Algunas veces no hay condición inicial, otras veces se tiene una condición inicial de la forma

$$y(x_0) = y_0$$

Un ejemplo sencillo puede ser

$$y' = 2x + 3y$$

Sus soluciones son de la forma

$$y = Ce^{3x} - 2x/3 - 2/9$$

Si tuviera la condición inicial

$$y(1) = 4$$

sería necesario calcular el valor de C en la solución general,

$$C = \frac{4 + 2/3 + 2/9}{e^3} \approx 0.243403445$$
$$y \approx 0.243403445\, e^{3x} - 2x/3 - 2/9$$

En un sistema de ecuaciones diferenciales hay una variable independiente, generalmente el tiempo t y dos o más variables dependientes de t, por ejemplo $u(t)$, $v(t)$, ... Para hacerlo más general, supongamos que hay n variables dependientes, $x_1(t)$, $x_2(t)$, ..., $x_n(t)$.

Es muy frecuente denotar la derivada con respecto al tiempo con un punto encima de la variable

$$x_i : \mathbb{R} \to \mathbb{R}$$
$$\dot{x}_i = \frac{dx_i}{dt} \tag{4.1}$$

Así un sistema de ecuaciones diferenciales se puede escribir:

$$\begin{aligned} \dot{x}_1 &= f_1(t, x_1, x_2, ..., x_n) \\ \dot{x}_2 &= f_2(t, x_1, x_2, ..., x_n) \\ &\vdots \\ \dot{x}_n &= f_n(t, x_1, x_2, ..., x_n) \end{aligned} \tag{4.2}$$

Ejemplo 4.1.

$$\begin{aligned} \dot{x}_1 &= t + x_1 x_2 + 5 \\ \dot{x}_2 &= t^2 + x_1^2 + x_2^2 \end{aligned}$$

Por medio de la siguiente notación,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_n \end{bmatrix} \qquad f(t,x) = \begin{bmatrix} f_1(t,x) \\ f_2(t,x) \\ \vdots \\ f_n(t,x) \end{bmatrix} \tag{4.3}$$

El sistema de ecuaciones diferenciales se puede escribir de manera compacta

$$\dot{x} = f(t, x) \tag{4.4}$$

De manera análoga a las ecuaciones diferenciales ordinarias lineales, hay **sistemas de ecuaciones diferenciales lineales**

$$\begin{aligned} \dot{x}_1 &= a_{11}(t)x_1 + a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n + g_1(t) \\ \dot{x}_2 &= a_{21}(t)x_1 + a_{22}(t)x_2 + \cdots + a_{2n}(t)x_n + g_2(t) \\ &\vdots \\ \dot{x}_n &= a_{n1}(t)x_1 + a_{n2}(t)x_2 + \cdots + a_{nn}(t)x_n + g_n(t) \end{aligned} \tag{4.5}$$

que se puede escribir de manera compacta

$$\dot{x} = A(t)\, x + g(t) \tag{4.6}$$

donde $A(t)$ es una matriz $n \times n$, cada entrada es una función de t,

$$A(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \vdots & & & \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{bmatrix}$$

y $g(t)$ es el vector columna

$$g(t) = \begin{bmatrix} g_1(t) \\ g_2(t) \\ \vdots \\ g_n(t) \end{bmatrix}$$

Ejemplo 4.2. El sistema de ecuaciones diferenciales del ejemplo (4.1) no es lineal. El que sigue sí lo es.

$$\begin{aligned} \dot{x}_1 &= t^2 x_1 + t^3 x_2 + t \\ \dot{x}_2 &= 2x_1 + \cos(t) x_2 \end{aligned}$$

$$A(t) = \begin{bmatrix} t^2 & t^3 \\ 2 & \cos(t) \end{bmatrix}, \quad g(t) = \begin{bmatrix} t \\ 0 \end{bmatrix} \quad \diamond$$

Un sistema de ecuaciones diferenciales lineales es **homogéneo** si $g(t) = 0$,

$$\begin{aligned}
\dot{x}_1 &= a_{11}(t)x_1 + a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n \\
\dot{x}_2 &= a_{21}(t)x_1 + a_{22}(t)x_2 + \cdots + a_{2n}(t)x_n \\
&\vdots \\
\dot{x}_n &= a_{n1}(t)x_1 + a_{n2}(t)x_2 + \cdots + a_{nn}(t)x_n
\end{aligned}$$

$$\dot{x} = A(t)\,x \tag{4.7}$$

Cuando todas las funciones $a_{ij}(t)$ y $g_i(t)$ son constantes se tiene un **sistema de ecuaciones diferenciales lineales con coeficientes constantes**.

$$\begin{aligned}
\dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_1 \\
\dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_2 \\
&\vdots \\
\dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + b_n
\end{aligned}$$

$$\dot{x} = Ax + b \tag{4.8}$$

Este tipo de sistemas es el tema principal de este capítulo y, para simplificar, mientras no se diga lo contrario, cuando se diga sistema de ecuaciones diferenciales se debe sobreentender como sistema de ecuaciones diferenciales lineales con coeficientes constantes, sedlcc.

Ejemplo 4.3.

$$\begin{aligned}
\dot{x}_1 &= x_1 + x_2 + x_3 + 4 \\
\dot{x}_2 &= 2x_1 + 3x_2 + 4x_3 - 5 \\
\dot{x}_3 &= x_1 - x_3
\end{aligned}$$

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 1 & 0 & -1 \end{bmatrix}$$
$$b = \begin{bmatrix} 4 \\ -5 \\ 0 \end{bmatrix}$$

Cuando $b = 0$, se tiene un sistema homogéneo (sistema homogéneo de ecuaciones diferenciales lineales con coeficientes constantes):

$$\begin{aligned}
\dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
\dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
&\vdots \\
\dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n
\end{aligned}$$

$$\dot{x} = Ax \tag{4.9}$$

Por ejemplo,

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 1 & 2 & -3 \\ 4 & -5 & 0 \\ -1 & 6 & 7 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$$

Un sistema general de ecuaciones diferenciales (4.4) se llama **autónomo** si no depende explícitamente de t,

$$\dot{x} = f(x) \tag{4.10}$$

Obviamente un sistema de ecuaciones diferenciales con coeficientes constante es autónomo, pero también hay sistemas no lineales autónomos.

Ejemplo 4.4. Sistemas autónomos:

$$\begin{aligned}
\dot{x}_1 &= x_1 x_2 + 5\\
\dot{x}_2 &= x_1^2 + x_2^2
\end{aligned}$$

$$\begin{aligned}
\dot{x}_1 &= x_1 + x_2 + x_3 + 4\\
\dot{x}_2 &= 2x_1 + 3x_2 + 4x_3 - 5\\
\dot{x}_3 &= x_1 - x_3
\end{aligned}$$

Usualmente un sistema general de ecuaciones diferenciales tiene **condiciones iniciales** de la forma

$$\begin{aligned}
x_1(t_0) &= x_1^0\\
x_2(t_0) &= x_2^0\\
&\vdots\\
x_n(t_0) &= x_n^0
\end{aligned}$$

y de manera compacta

$$x(t_0) = x^0 \tag{4.11}$$

Por ejemplo

$$\begin{aligned}
\dot{x}_1 &= x_1 + x_2 + x_3 + 4\\
\dot{x}_2 &= 2x_1 + 3x_2 + 4x_3 - 5\\
\dot{x}_3 &= x_1 - x_3\\
x(1.2) &= \begin{bmatrix} 2\\ 3\\ 4 \end{bmatrix}
\end{aligned}$$

En el sistema autónomo (4.10), $\dot{x} = f(x)$, un punto x^* es un **punto de equilibrio** o **punto estacionario** o **punto fijo** si

$$f(x^*) = 0\,. \tag{4.12}$$

Si la condición inicial fuera $x(t_0) = x^*$, entonces $\dot{x}(t) = 0$, es decir, no habrá cambios en x, luego $x(t) = x^*$ para todo $t \geq t_0$.

Un sistema autónomo puede no tener puntos de equilibrio, puede tener uno solo o puede tener varios.

Ejemplo 4.5.

$$\begin{aligned}
\dot{x}_1 &= x_1^2 + x_2^2 + 1\\
\dot{x}_2 &= x_1 x_2 - 12
\end{aligned}$$

No tiene puntos de equilibrio

$$\begin{aligned}
\dot{x}_1 &= x_1^2 + x_2^2 - 25\\
\dot{x}_2 &= x_1 x_2 - 12
\end{aligned}$$

Tiene cuatro puntos de equilibrio, $(3,4)$, $(4,3)$, $(-3,-4)$ y $(-4,-3)$.

$$\begin{aligned}\dot{x}_1 &= 2x_1 + 3x_2 - 7\\ \dot{x}_2 &= 4x_1 + 5x_2 - 13\end{aligned}$$

Tiene un único punto de equilibrio, $x^* = (2,1)$. $\diamond$

4.2 Resultados generales

Proposición 4.1. *Si $x^1(t)$ y $x^2(t)$ son soluciones del sistema homogéneo* (4.9) *y $c \in \mathbb{R}$, entonces $x^1(t) + x^2(t)$ y $c\,x^1(t)$ también son soluciones de* (4.9).

La proposición anterior también es válida para el sistema homogéneo (no lineal) (4.7).

El objetivo será entonces encontrar n soluciones linealmente independientes $x^1(t)$, $x^2(t)$, ..., $x^n(t)$ y así la solución general de (4.9) se puede escribir

$$x(t) = c_1\, x^1(t) + c_2\, x^2(t) + \cdots + c_n\, x^n(t) \tag{4.13}$$

Supongamos que una solución de (4.9) es de la forma

$$x(t) = e^{\lambda t} v \tag{4.14}$$

Entonces,

$$\dot{x} = \lambda e^{\lambda t} v$$

Al remplazar en (4.9)

$$\begin{aligned}\lambda e^{\lambda t} v &= A(e^{\lambda t} v)\\ e^{\lambda t} \lambda v &= e^{\lambda t} A v\\ \lambda v &= A v\end{aligned}$$

Es decir, v es un vector propio de A asociado al valor propio λ.

Es importante resaltar que **no siempre las soluciones de** (4.9) **son de la forma** (4.14).

En lo que sigue, mientras no se diga lo contrario, se estudiará el sistema de ecuaciones diferenciales

$$\begin{aligned}\dot{x} &= Ax + b\\ x(t_0) &= x^0\end{aligned} \tag{4.15}$$

y su sistema homogéneo asociado

$$\dot{x} = Ax \tag{4.16}$$

Se supondrá que la matriz A **tiene determinante no nulo**, así hay un único punto de equilibrio

$$x^* = -A^{-1} b \tag{4.17}$$

o calculado de manera más eficiente como la solución de

$$A\,x^* = -b \tag{4.18}$$

Sea x^g la solución general del sistema (4.15) y x^{gh} la solución general del sistema homogéneo (4.16). Entonces

$$x^g = x^{gh} + x^* \tag{4.19}$$

Ejemplo 4.6.

$$\dot{x} = \begin{bmatrix} -3 & -4 \\ 3 & 5 \end{bmatrix} x + \begin{bmatrix} 11 \\ -13 \end{bmatrix}$$

El sistema homogéneo asociado es

$$\dot{x} = \begin{bmatrix} -3 & -4 \\ 3 & 5 \end{bmatrix} x$$

El punto de equilibrio, solución de $Ax = -b$, es $x^* = \begin{bmatrix} 1 & 2 \end{bmatrix}^{\mathrm{T}}$ La solución general del sistema homogéneo (más adelante se verán los detalles) es:

$$x^{gh}(t) = C_1\, e^{-t} \begin{bmatrix} -2 \\ 1 \end{bmatrix} + C_2\, e^{3t} \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

y la solución general del sistema no homogéneo

$$x^{g}(t) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + C_1\, e^{-t} \begin{bmatrix} -2 \\ 1 \end{bmatrix} + C_2\, e^{3t} \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

Si hay condiciones iniciales, se calcula el valor de las constantes C_1 y C_2. Sea

$$x(0.5) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Entonces

$$\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + C_1\, e^{-0.5} \begin{bmatrix} -2 \\ 1 \end{bmatrix} + C_2\, e^{1.5} \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

Resulta el sistema de ecuaciones

$$\begin{bmatrix} -1.213061 & -8.963378 \\ 0.606531 & 13.445067 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$$

La solución es:

$$\begin{aligned} C_1 &= 3.297442 \\ C_2 &= -0.223130 \end{aligned}$$

Entonces

$$\begin{aligned} x_1(t) &= 1 - 6.594885\, e^{-t} + 0.44626\, e^{3t} \\ x_2(t) &= 2 + 3.297443\, e^{-t} - 0.66939\, e^{3t} \end{aligned}$$

Así, por ejemplo, $x_2(1) = -10.232006$. $\diamond$

4.3 Representación gráfica de la solución

Consideremos el sistema de n ecuaciones diferenciales

$$\begin{aligned} \dot{x} &= Ax + b \\ x(t_0) &= x^0 \end{aligned}$$

y supongamos que se conoce la solución y se desea graficarla para $t \in [t_0, t_1]$. Lo que realmente se hace es dibujar n gráficas, la de $x_1(t)$, la de $x_2(t)$, ..., la de $x_n(t)$ para t variando en el intervalo $[t_0, t_1]$.

Ejemplo 4.7.

$$\dot{x} = \frac{1}{10}\begin{bmatrix} -9 & 11 & -21 \\ -1 & -21 & 21 \\ -11 & -11 & 1 \end{bmatrix} x + \begin{bmatrix} -6/5 \\ 11/5 \\ 6/5 \end{bmatrix}$$

$$x(1) = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Como se verá más adelante la solución general es

$$x(t) = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + C_1\, e^{-2t}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + C_2\, e^{-t}\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} + C_3\, e^{t/10}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

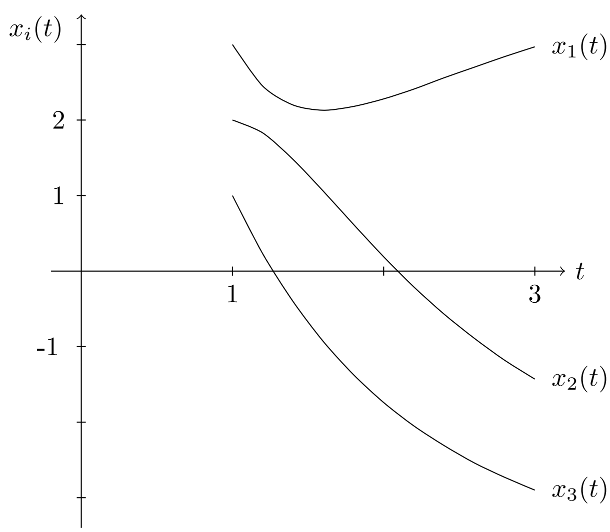
Teniendo en cuenta la condición inicial

$$x(t) = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - 29.5562\, e^{-2t}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + 10.8731\, e^{-t}\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} - 1.8097\, e^{t/10}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Algunos valores son:

t	x1	x2	x3
1.0	3.00	2.00	1.00
1.2	2.45	1.83	0.23
1.4	2.20	1.48	-0.40
1.6	2.13	1.06	-0.93
1.8	2.18	0.62	-1.37
2.0	2.28	0.19	-1.74
2.2	2.41	-0.21	-2.05
2.4	2.56	-0.57	-2.31
2.6	2.70	-0.89	-2.54
2.8	2.84	-1.18	-2.73
3.0	2.97	-1.43	-2.90

Las tres curvas son:



Cuando $n = 2$, también se puede dibujar la curva de $x_1(t)$ y la de $x_2(t)$ o directamente dibujar la curva con los valores $(\,x_1(t),\, x_2(t)\,)$

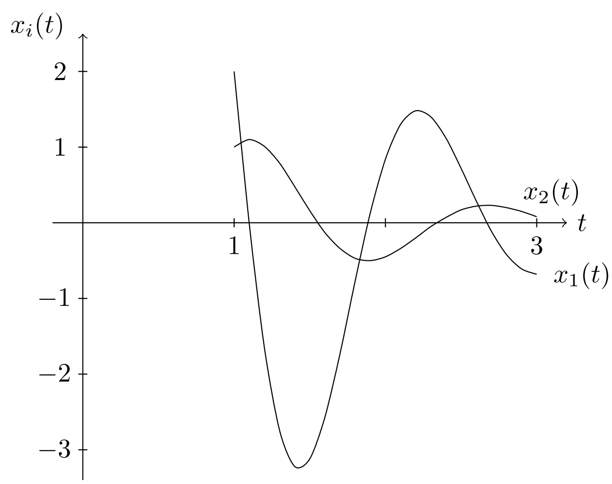
Ejemplo 4.8.

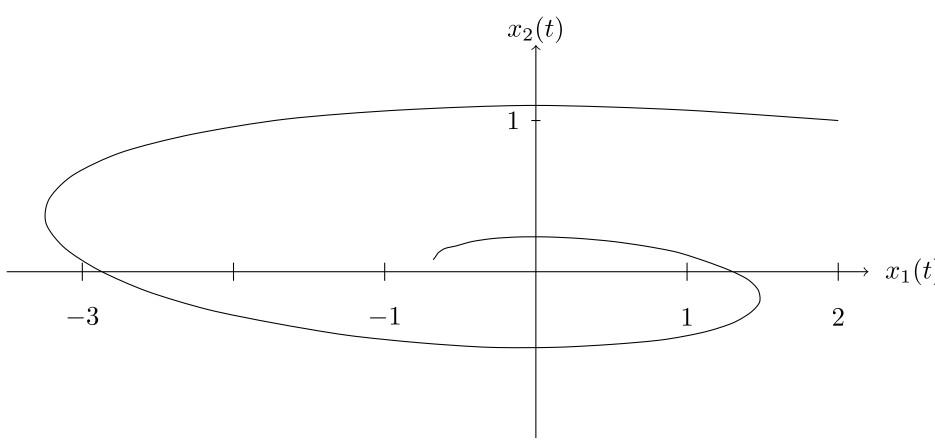
$$\dot{x} = \begin{bmatrix} -2 & -17 \\ 1 & 0 \end{bmatrix} x$$

$$x(1) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Algunos de los valores numéricos de la solución son:

t	x1	x2
1.0	2.00	1.00
1.1	-0.01	1.10
1.2	-1.65	1.01
1.3	-2.74	0.79
1.4	-3.22	0.48
1.5	-3.12	0.16
1.6	-2.57	-0.13
1.7	-1.73	-0.34
1.8	-0.77	-0.47
1.9	0.13	-0.50
2.0	0.84	-0.45
2.1	1.30	-0.34
2.2	1.48	-0.20
2.3	1.40	-0.05
2.4	1.12	0.07
2.5	0.72	0.17
2.6	0.29	0.22
2.7	-0.11	0.23
2.8	-0.42	0.20
2.9	-0.61	0.15
3.0	-0.68	0.08





4.4 Matriz diagonalizable

Una matriz $A \in \mathbb{R}^{n\times n}$ es diagonalizable si existe un matriz $B \in \mathbb{R}^{n\times n}$ invertible tal que $B^{-1}AB$ es una matriz diagonal.

Un resultado de Álgebra Lineal dice que A es diagonalizable si y solamente si **existen** v^1**,** v^2**, ...,** $v^n \in \mathbb{R}^{n\times 1}$**, vectores propios de** A**, linealmente independientes**. En este caso la solución general de (4.9) es

$$x(t) = c_1 e^{\lambda_1 t} v^1 + c_2 e^{\lambda_2 t} v^2 + \cdots + c_n e^{\lambda_n t} v^n \tag{4.20}$$

Una condición necesaria (pero no suficiente) para que una matriz sea diagonalizable es que todos sus valores propios sean reales.

4.4.1 Valores propios reales diferentes

El caso más sencillo de una matriz diagonalizable se da cuando **hay** n **valores propios reales diferentes**, en este caso, sus vectores propios asociados son linealmente independientes.

Ejemplo 4.9.

$$\dot{x} = \begin{bmatrix} -3 & -4 \\ 3 & 5 \end{bmatrix} x$$

Polinomio característico

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I) \\ &= \det \begin{bmatrix} -3-\lambda & -4 \\ 3 & 5-\lambda \end{bmatrix} \\ &= \lambda^2 - 2\lambda - 3 \\ &= (\lambda+1)(\lambda-3) \\ \lambda_1 &= -1 \\ \lambda_2 &= 3 \end{aligned}$$

Vector propio asociado a $\lambda_1 = -1$,

$$\begin{aligned} (A - \lambda_1 I)v^1 &= 0 \\ (A + I)v^1 &= 0 \\ \begin{bmatrix} -2 & -4 \\ 3 & 6 \end{bmatrix} v^1 &= 0 \end{aligned}$$

Matriz escalonada reducida

$$\begin{aligned} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ v_1 &= -2v_2 \end{aligned}$$

Por ejemplo

$$v^1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

o cualquier múltiplo no nulo de este vector.

Vector propio asociado a $\lambda_2 = 3$,

$$\begin{aligned} (A - \lambda_2 I)v^2 &= 0 \\ (A - 3I)v^2 &= 0 \\ \begin{bmatrix} -6 & -4 \\ 3 & 2 \end{bmatrix} v^2 &= 0 \end{aligned}$$

Matriz escalonada reducida

$$\begin{aligned} \begin{bmatrix} 1 & 2/3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ v_1 &= -2v_2/3 \end{aligned}$$

Por ejemplo

$$v^2 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

o cualquier múltiplo no nulo de este vector.

La solución general es

$$x = C_1 \, e^{-t} \begin{bmatrix} -2 \\ 1 \end{bmatrix} + C_2 \, e^{3t} \begin{bmatrix} -2 \\ 3 \end{bmatrix} \quad \diamond$$

Para matrices 2×2, si λ es un valor propio, entonces en la matriz $A - \lambda I$, una fila es múltiplo de la otra. Si $A - \lambda I = 0$, cualquier vector no nulo es vector propio y hay dos vectores propios independientes. Si $A - \lambda I$ tiene una fila nula, basta con utilizar la otra fila para obtener un vector propio. Si $A - \lambda I$ tiene dos filas no nulas, se puede utilizar cualquiera de las dos para obtener el vector propio. En el ejemplo anterior, se buscó la matriz escalonada reducida (procedimiento general) pero no era necesario. Basta con considerar la información proveniente de la primera fila:

$$\begin{aligned} -6v_1 - 4v_2 &= 0 \\ v_1 &= -2v_2/3 \end{aligned}$$

Ejemplo 4.10.

$$\dot{x} = \begin{bmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{bmatrix} x$$

Polinomio característico:

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I) \\ p(\lambda) &= -\lambda^3 + 2\lambda^2 + 5\lambda - 6 \end{aligned}$$

Sus raíces son los valores propios,

$$\lambda_1 = 1, \quad \lambda_2 = -2, \quad \lambda_3 = 3$$

Obtención de vectores propios:

$$\begin{aligned} (A - \lambda_1 I)v^1 &= 0 \\ (A - 1 \cdot I) &= \begin{bmatrix} 0 & -1 & 4 \\ 3 & 1 & -1 \\ 2 & 1 & -2 \end{bmatrix} \\ E_{A-I} &= \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Entonces $v_1 = -v_3$, $v_2 = 4v_3$. Por ejemplo, si $v_3 = 1$, entonces se obtiene $v^1 = \begin{bmatrix} -1 & 4 & 1 \end{bmatrix}^{\mathrm{T}}$ un vector propio asociado a $\lambda_1 = 1$. De manera análoga se construyen v^2 y v^3.

$$v^1 = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}, \quad v^2 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \quad v^3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},$$

$$x(t) = c_1 e^t \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix} + c_2 e^{-2t} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} + c_3 e^{3t} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Si hay condiciones iniciales, es necesario calcular los valores de c_1, c_2, ... Por ejemplo, sea

$$x(1) = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$$

Al remplazar en la solución general se obtiene un sistema de ecuaciones, donde las incógnitas son las constantes c_i,

$$\begin{bmatrix} -2.718282 & -0.135335 & 20.085537 \\ 10.873127 & 0.135335 & 40.171074 \\ 2.718282 & 0.135335 & 20.085537 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$$

La solución es

$$\begin{aligned} c_1 &= 0.245253 \\ c_2 &= -19.704150 \\ c_3 &= 0 \end{aligned}$$

Con los valores de las constantes se puede calcular $x(t)$ para cualquier valor de t. Por ejemplo,

$$x(1.5) = \begin{bmatrix} -0.118136 \\ 3.415578 \\ 0.118136 \end{bmatrix}$$

4.4.2 Multiplicidades algebraica y geométrica iguales

La **multiplicidad algebraica** de un valor propio λ es el número de veces que λ es raíz del polinomio característico.

$$\begin{aligned} p(\lambda) &= (\lambda-4)^3(\lambda+5)^2(\lambda-6) \\ \text{mult.alg}(4) &= 3 \\ \text{mult.alg}(-5) &= 2 \\ \text{mult.alg}(6) &= 1 \end{aligned}$$

La **multiplicidad geométrica** de un valor propio λ se puede presentar de varias maneras, obviamente equivalentes:

- Número de variables libres en $E_{A-\lambda I}$ la matriz escalonada reducida de $A - \lambda I$.
- Número máximo de vectores propios, asociados a λ, linealmente independientes.
- $\dim\big(\{\, v \in \mathbb{R}^{n\times 1} : Av = \lambda v \,\}\big)$

El conjunto $\{\, v \in \mathbb{R}^{n\times 1} : Av = \lambda v \,\}$ recibe el nombre de subespacio propio asociado a λ.

Las multiplicidades están relacionadas por la siguiente desigualdad:

$$1 \le \text{mult.geom}(\lambda) \le \text{mult.alg}(\lambda) \tag{4.21}$$

Una matriz $A \in \mathbb{R}^{n\times n}$ es diagonalizable si y solamente si para todos los valores propios

$$\text{mult.geom}(\lambda) = \text{mult.alg}(\lambda) \tag{4.22}$$

El caso de n valores propios reales diferentes, visto anteriormente, es una caso particular del resultado anterior.

Un caso "extremo" de (4.22) se da cuando

$$\text{mult.geom}(\lambda) = \text{mult.alg}(\lambda) = n$$

Es decir hay un solo valor propio y sus multiplicidades valen n. En este caso la matriz A es un múltiplo de la matriz identidad y, en consecuencia, $\dot{x}_i(t)$ depende únicamente de $x_i(t)$.

Ejemplo 4.11.

$$\dot{x} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} x$$

$$\lambda_1 = \lambda_2 = \lambda_3 = -1$$

$$A - (-1)I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Luego, para el mismo valor propio, hay tres vectores propios independientes. Se puede tomar cualquier conjunto de tres vectores en $\mathbb{R}^{3\times 1}$, linealmente independientes. Por ejemplo, los de la base canónica.

$$x(t) = c_1\, e^{-t} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2\, e^{-t} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3\, e^{-t} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Si

$$x(1) = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

entonces

$$x(t) = 5.436564\, e^{-t} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 8.154845\, e^{-t} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 10.873127\, e^{-t} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

que se puede escribir

$$x(t) = e^{-t} \begin{bmatrix} 5.436564 \\ 8.154845 \\ 10.873127 \end{bmatrix} = e^{-t}\; e \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

Ejemplo 4.12.

$$\dot{x} = \begin{bmatrix} -2 & -8 & -4 \\ 0 & 0 & 1 \\ 0 & 6 & 1 \end{bmatrix} x$$

$$\begin{aligned} p(\lambda) &= -\lambda^3 - \lambda^2 + 8\lambda + 12 \\ \lambda_1 &= -2 \\ \lambda_2 &= -2 \\ \lambda_3 &= \;\;\, 3 \end{aligned}$$

$$\begin{aligned} A + 2I &= \begin{bmatrix} 0 & -8 & -4 \\ 0 & 2 & 1 \\ 0 & 6 & 3 \end{bmatrix} \\ E &= \begin{bmatrix} 0 & 1 & 1/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

$$v^1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad v^2 = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}$$

Para $\lambda_3 = 3$,

$$v^3 = \begin{bmatrix} -4 \\ 1 \\ 3 \end{bmatrix}$$

Entonces, A es diagonalizable y

$$x(t) = c_1 e^{-2t} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 e^{-2t} \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} + c_3 e^{3t} \begin{bmatrix} -4 \\ 1 \\ 3 \end{bmatrix}$$

Si

$$x(0) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

entonces

$$x(t) = 5e^{-2t} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - e^{-2t} \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} + e^{3t} \begin{bmatrix} -4 \\ 1 \\ 3 \end{bmatrix}$$

4.5 Matrices 2×2 no diagonalizables

4.5.1 Valores propios reales iguales

Esto quiere decir que hay un valor propio real doble λ, un vector propio v asociado a λ, pero no fue posible encontrar otro vector propio asociado a λ, linealmente independiente con v. Se puede suponer que una solución de (4.9) es de la forma

$$x = e^{\lambda t}(u + tw) \tag{4.23}$$

Entonces

$$\dot{x} = \lambda e^{\lambda t}(u + tw) + e^{\lambda t} w$$

Remplazando en (4.9)

$$\begin{aligned} \lambda e^{\lambda t}(u + tw) + e^{\lambda t} w &= A(e^{\lambda t}(u + tw)) \\ e^{\lambda t}\big(\lambda u + \lambda tw + w\big) &= e^{\lambda t}\big(Au + tAw\big) \\ \lambda u + \lambda tw + w &= Au + tAw \\ \lambda u + w &= Au \\ \lambda tw &= tAw \\ \lambda w &= Aw \end{aligned}$$

Es decir, w es un vector propio y

$$\begin{aligned} Au - \lambda u &= w \\ (A - \lambda I)u &= w \end{aligned}$$

En resumen

Obtener un vector propio v	$(A - \lambda I)v = 0,$	(4.24)
Resolver	$(A - \lambda I)u = v$	(4.25)

Así,

$$x(t) = c_1\, e^{\lambda t} v + c_2\, e^{\lambda t}\big(tv + u\big) \tag{4.26}$$

Observación. Como $A - \lambda I$ es singular (su determinante es nulo), el sistema $(A - \lambda I)u = v$ tiene un número infinito de soluciones. Basta con tomar una cualquiera.

Ejemplo 4.13. `A :`

```
   -2    3
    0   -2

t0 = 1/2
x0 :    1   2

pol. caract.
              2
    4 + 4x + x

val. pr. :    -2   -2

A - lambda I :
   0   3
   0   0

esc. reducida :
   0   1
   0   0
v :    1   0
u :    0   1/3

matriz para calcular c :
   0.367879   0.183940
          0   0.122626

C :    -5.436564   16.309691
```

4.5.2 Valores propios complejos

Como el polinomio característico tiene coeficientes reales, entonces los valores propios complejos (cuando los hay) vienen por parejas, uno es el conjugado del otro,

$$\lambda_1 = a + i\beta$$
$$\lambda_2 = a - i\beta$$

Sea v^1 un vector propio complejo, asociado a λ_1,

$$v^1 = p + iq$$

donde p y q son vectores reales, en $\mathbb{R}^{n\times 1}$. Entonces

$$v^2 = p - iq$$

es un vector propio asociado a λ_2.

Sean x^1 y x^2 soluciones de (4.9), construidas según (4.14),

$$\begin{aligned}
x^1 &= e^{(\alpha+i\beta)t}(p+iq),\\
x^2 &= e^{(\alpha-i\beta)t}(p-iq)\\
x^1 &= e^{\alpha t}[\cos(\beta t)+i\,\mathrm{sen}\,(\beta t)](p+iq),\\
x^2 &= e^{\alpha t}(\cos(\beta t)-i\,\mathrm{sen}\,(\beta t))(p-iq)\\
x^1 &= e^{\alpha t}[\cos(\beta t)p-\,\mathrm{sen}\,(\beta t)q+i\,\mathrm{sen}\,(\beta t)p+i\cos(\beta t)q]\\
x^2 &= e^{\alpha t}[\cos(\alpha t)p-\,\mathrm{sen}\,(\beta t)q-i\,\mathrm{sen}\,(\beta t)p-i\cos(\beta t)q]\\
x^1+x^2 &= 2e^{\alpha t}(\cos(\beta t)p-\,\mathrm{sen}\,(\beta t)q)\\
x^1-x^2 &= 2ie^{\alpha t}(\,\mathrm{sen}\,(\beta t)p+\cos(\beta t)q)\\
(x^1+x^2)/2 &= e^{\alpha t}(\cos(\beta t)p-\,\mathrm{sen}\,(\beta t)q)\\
(x^1-x^2)/(2i) &= e^{\alpha t}(\,\mathrm{sen}\,(\beta t)p+\cos(\beta t)q)
\end{aligned}$$

Como ya hay dos soluciones reales, linealmente independientes, entonces

$$x(t) = c_1\,e^{\alpha t}[-\,\mathrm{sen}\,(\beta t)q+\cos(\beta t)p]+c_2\,e^{\alpha t}[\,\mathrm{sen}\,(\beta t)p+\cos(\beta t)q] \tag{4.27}$$

Ejemplo 4.14.

$$\begin{aligned}
\dot{x} &= \begin{bmatrix}-9 & 8\\ -9 & 3\end{bmatrix}x\\
p(\lambda) &= \lambda^2+6\lambda+45\\
\lambda_1 &= -3+6i\\
\lambda_2 &= -3-6i\\
A-\lambda_1 I &= \begin{bmatrix}-6-6i & 8\\ -9 & 6-6i\end{bmatrix}
\end{aligned}$$

Aunque no parezca (por ser valores complejos), la primera y la segunda fila de $A-\lambda I$ son equivalentes. Es decir, para obtener un vector propio se puede considerar cualquiera de la dos filas.

$$\begin{aligned}
-9v_1+(6-6i)v_2 &= 0\\
v_1 &= \left(\frac{2}{3}-\frac{2}{3}\,i\right)v_2\\
v^1 &= \begin{bmatrix}\frac{2}{3}-\frac{2}{3}i\\ 1\end{bmatrix}\\
v^1 &= \begin{bmatrix}2/3\\ 1\end{bmatrix}+\begin{bmatrix}-2/3\\ 0\end{bmatrix}i\\
v^1 &= \quad p \quad + \quad q \quad i
\end{aligned}$$

También se puede considerar un múltiplo,

$$\begin{aligned}
v^1 &= \begin{bmatrix}2\\ 3\end{bmatrix}+\begin{bmatrix}-2\\ 0\end{bmatrix}i\\
v^1 &= \; p \; + \; q \; i
\end{aligned}$$

$$\begin{aligned}
\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} &= c_1e^{-3t}\left(\,\mathrm{sen}\,(6t)\begin{bmatrix}2\\ 0\end{bmatrix}+\cos(6t)\begin{bmatrix}2\\ 3\end{bmatrix}\right)+c_2e^{-3t}\left(\,\mathrm{sen}\,(6t)\begin{bmatrix}2\\ 3\end{bmatrix}+\cos(6t)\begin{bmatrix}-2\\ 0\end{bmatrix}\right)\\
&= c_1e^{-3t}\begin{bmatrix}2\,\mathrm{sen}\,(6t)+2\cos(6t)\\ 3\cos(6t)\end{bmatrix}+c_2e^{-3t}\begin{bmatrix}2\,\mathrm{sen}\,(6t)-2\cos(6t)\\ 3\,\mathrm{sen}\,(6t)\end{bmatrix}
\end{aligned}$$

Supongamos que tenemos las condiciones iniciales

$$x(1) = \begin{bmatrix}2\\ 3\end{bmatrix}$$

Con estas condiciones es necesario calcular los valores, c_1, c_2.

$$c_1 \begin{bmatrix} 0.067786 \\ 0.143412 \end{bmatrix} + c_2 \begin{bmatrix} -0.123431 \\ -0.041734 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

La solución del sistema da:

$$\begin{aligned} c_1 &= 19.285536 \\ c_2 &= -5.612210 \end{aligned}$$

Entonces la solución es:

$$\begin{aligned} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} &= 19.285536\ e^{-3t} \begin{bmatrix} 2\,\text{sen}\,(6t) + 2\cos(6t) \\ 3\cos(6t) \end{bmatrix} - 5.612210\ e^{-3t} \begin{bmatrix} 2\,\text{sen}\,(6t) - 2\cos(6t) \\ 3\,\text{sen}\,(6t) \end{bmatrix} \\ \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} &= e^{-3t} \begin{bmatrix} 27.346651\,\text{sen}\,(6t) + 49.795492\cos(6t) \\ -16.836631\,\text{sen}\,(6t) + 57.856607\cos(6t) \end{bmatrix} \end{aligned}$$

Ya con la solución, se pueden calcular los valores de x_1 y x_2 para un t dado. Por ejemplo

$$\begin{bmatrix} x_1(2) \\ x_2(2) \end{bmatrix} = \begin{bmatrix} 0.067786 \\ 0.143412 \end{bmatrix}$$

Veamos ahora el resultado usando la primera fila:

$$\begin{aligned} (-6-6i)v_1 + 8v_2 &= 0 \\ v_2 &= \frac{6+6i}{8} v_1 \\ v &= \begin{bmatrix} 1 \\ \frac{3}{4} + \frac{3}{4}i \end{bmatrix} \sim \begin{bmatrix} 4 \\ 3+3i \end{bmatrix} \\ p &= \begin{bmatrix} 4 \\ 3 \end{bmatrix} \quad q = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} &= c_1 e^{-3t} \left(-\,\text{sen}\,(6t) \begin{bmatrix} 0 \\ 3 \end{bmatrix} + \cos(6t) \begin{bmatrix} 4 \\ 3 \end{bmatrix} \right) + c_2 e^{-3t} \left(\text{sen}\,(6t) \begin{bmatrix} 4 \\ 3 \end{bmatrix} + \cos(6t) \begin{bmatrix} 0 \\ 3 \end{bmatrix} \right) \\ &= c_1\, e^{-3t} \begin{bmatrix} 4\cos(6t) \\ -3\,\text{sen}\,(6t) + 3\cos(6t) \end{bmatrix} + c_2\, e^{-3t} \begin{bmatrix} 4\,\text{sen}\,(6t) \\ 3\,\text{sen}\,(6t) + 3\cos(6t) \end{bmatrix} \end{aligned}$$

solución general que no se parece a la obtenida anteriormente. Para las condiciones iniciales,

$$c_1 = 12.448873, \quad c_2 = 6.836663$$

valores muy diferentes, ni siquiera múltilplos de los obtennidos anteriormente. Sin embargo al utilizar estos valores en la solución general, se obtiene exactamente la misma expresión:

$$\begin{aligned} x(t) &= 12.448873\, e^{-3t} \begin{bmatrix} 4\cos(6t) \\ -3\,\text{sen}\,(6t) + 3\cos(6t) \end{bmatrix} + 6.836663\, e^{-3t} \begin{bmatrix} 4\,\text{sen}\,(6t) \\ 3\,\text{sen}\,(6t) + 3\cos(6t) \end{bmatrix} \\ x(t) &= e^{-3t} \begin{bmatrix} 27.346651\,\text{sen}\,(6t) + 49.795492\cos(6t) \\ -16.836631\,\text{sen}\,(6t) + 57.856607\cos(6t) \end{bmatrix} \end{aligned}$$

4.6 Matrices 3×3 no diagonalizables

4.6.1 Mult.algebraica$(\lambda) = 3$, mult.geométrica$(\lambda) = 1$

Sea λ una valor propio y v un vector propio asociado. Supongamos que

$$\begin{aligned} \text{mult-alg}(\lambda) &= 3 \\ \text{mult-geom}(\lambda) &= 1 \end{aligned}$$

Un proceso análogo al que permite obtener (4.26) dice que la solución general es

$$x(t) = c_1\, e^{\lambda t} v + c_2\, e^{\lambda t}(tv + u) + c_3\, e^{\lambda t}\left(\frac{t^2}{2} v + tu + w\right) \tag{4.28}$$

donde v es un vector propio de A y los vectores u y w se obtienen encontrando una solución cualquiera de

$$(A - \lambda I)u = v \tag{4.29}$$

$$(A - \lambda I)w = u \tag{4.30}$$

Ejemplo 4.15.
```
A :
   -2    1   -1
    0   -2    2
    0    0   -2

t0 = 1
x0 :    1/2   1   1/2

pol. caract.
                2   3
  - 8 - 12x - 6x - x

val. pr. :    -2   -2   -2

A - lambda I :
   0   1   -1
   0   0    2
   0   0    0

esc. reducida :
   0   1   0
   0   0   1
   0   0   0

v :    1   0   0
u :    0   1   0
w :    0   1/2   1/2

matriz para calcular c :
   0.135335   0.135335   0.067668
          0   0.135335   0.203003
          0          0   0.067668

C :    3.694528   -3.694528   7.389056
```

4.6.2 Mult.algebraica$(\lambda_1) = 2$, mult.geométrica$(\lambda_1) = 1$

Ejemplo 4.16.

$$\dot{x} = Ax, \quad A = \begin{bmatrix} 2 & 5 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad x(0.5) = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

$$\det(A - \lambda I) = (2 - \lambda)(2 - \lambda)(3 - \lambda)$$
$$\lambda_1 = 2$$
$$\lambda_2 = 2$$
$$\lambda_3 = 3$$

$\lambda = 2, \quad \text{mult-alg}(2) = 2$

$$A - 2I = \begin{bmatrix} 0 & 5 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$E = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\text{mult-geom}(2) = 1$$
$$v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$
$$(A - 2I)u = v$$
$$\text{matriz ampliada: } \begin{bmatrix} 0 & 5 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$
$$\text{escalonada reducida: } \begin{bmatrix} 0 & 1 & 0 & 1/5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$$u = \begin{bmatrix} 0 \\ 1/5 \\ 0 \end{bmatrix}$$

$\lambda = 3$

$$A - 3I = \begin{bmatrix} -1 & 5 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$v = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$x(t) = c_1\, e^{2t} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2\, e^{2t} \left(t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1/5 \\ 0 \end{bmatrix} \right) + c_3\, e^{3t} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Condiciones iniciales:

$$\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = c_1 \begin{bmatrix} 2.718282 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1.359141 \\ 0.543656 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 0 \\ 4.481689 \end{bmatrix}$$
$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -1.839397 \\ 3.678794 \\ 0.223130 \end{bmatrix}$$
$$x(t) = -1.839397 \begin{bmatrix} e^{2t} \\ 0 \\ 0 \end{bmatrix} + 3.678794 \begin{bmatrix} te^{2t} \\ e^{2t}/5 \\ 0 \end{bmatrix} + 0.223130 \begin{bmatrix} 0 \\ 0 \\ e^{3t} \end{bmatrix}$$
$$x(t) = \begin{bmatrix} (\,3.678794\, t - 1.839397\,)e^{2t} \\ 0.735759\, e^{2t} \\ 0.223130\, e^{3t} \end{bmatrix}$$

4.6.3 Dos valores propios complejos

Cuando hay dos valores propios complejos, necesariamente el otro es real. Sean $\lambda_1 \in \mathbb{R}$, $\lambda_2 = \alpha+\beta i$ y $\lambda_3 = \alpha-\beta i$ los valores propios. Sea v^1 vector propio asociado a λ_1 y $v^2 = p + qi$ vector propio asociado a $\lambda_2 = \alpha + \beta i$.

Un razonamiento semejante al caso de dos valores propios complejos en matrices 2×2 conduce a la solución general del sistema homogéneo

$$x(t) = C_1 e^{\lambda_1 t} v^1 + c_2\, e^{\alpha t}[-\operatorname{sen}(\beta t) q + \cos(\beta t) p] + c_3\, e^{\alpha t}[\operatorname{sen}(\beta t) p + \cos(\beta t) q] \tag{4.31}$$

Para el valor propio $\lambda_2 = \alpha+\beta i$, la matriz $A-\lambda_2 I$ es singular (determinante nulo), ninguna fila es nula, las tres filas son linealmente dependientes. Para calcular el vector propio correspondiente basta con utilizar dos filas, la otra se puede suprimir. Buscando disminuir el número de divisiones con denominador complejo, se puede usar la tercera fila y la segunda (suprimir la primera e intercambiar las otras dos).

Ejemplo 4.17.

```
A :
   6    0   -5
   0   -2    0
   2    0    0
t0 = 1
x0 :    1/2   1   1/2
pol. caract.
                  2   3
  - 20 + 2x + 4x - x
lambda1 = -2
val. propios compl:  3  + -  i

lambda = -2
A - lambda I :
   8   0   -5
   0   0    0
   2   0    2
parte de A - lambda I :
   8   0   -5
   2   0    2
esc. reducida :
   1   0   0
   0   0   1
v1 :    0   1   0

lambda = 3 + i
alfa = 3
beta = 1
A - z I
    3. - i      0         - 5.
    0         - 5. - i      0
    2.          0         - 3. - i

Quitar la primera fila e intercambiar las otras dos.
    2.          0         - 3. - i
    0         - 5. - i      0
escalonada
    1.    0   - 1.5 - 0.5i
    0     1.    0
v2
    1.5 + 0.5i
    0
    1.
v2
```

```
    3  + 1 i
    0
    2
p :    3   0   2
q :    1   0   0
```

$$x(t) = c_1 e^{-2t}\begin{bmatrix}0\\1\\0\end{bmatrix} + c_2 e^{3t}\left(-\operatorname{sen}(t)\begin{bmatrix}1\\0\\0\end{bmatrix} + \cos(t)\begin{bmatrix}3\\0\\2\end{bmatrix}\right)$$
$$+ c_3 e^{3t}\left(\operatorname{sen}(t)\begin{bmatrix}3\\0\\2\end{bmatrix} + \cos(t)\begin{bmatrix}1\\0\\0\end{bmatrix}\right)$$

Para las condiciones iniciales

$$c_1 = 7.389056$$
$$c_2 = 0.017199$$
$$c_3 = 0.003749$$

Entonces

$$x(t) = e^{-2t}\begin{bmatrix}0\\7.389056\\0\end{bmatrix} + e^{3t}\operatorname{sen}(t)\begin{bmatrix}-0.005953\\0\\0.007497\end{bmatrix} + e^{3t}\cos(t)\begin{bmatrix}0.055344\\0\\0.034397\end{bmatrix}$$

$$x(t) = \begin{bmatrix}e^{3t}\left(-0.005953\operatorname{sen}(t) + 0.055344\cos(t)\right)\\ 7.389056\, e^{-2t}\\ e^{3t}\left(0.007497\operatorname{sen}(t) + 0.034397\cos(t)\right)\end{bmatrix}$$

4.6.4 Mult.algebraica(λ) = 3, mult.geométrica(λ) = 2

Sea λ el valor propio real de multiplicidad algebraica igual a 3 y v^1, v^2 dos vectores propios, asociados a λ, linealmente independientes. Hasta ahora se tienen dos soluciones independientes

$$x^1(t) = e^{\lambda t}v^1$$
$$x^2(t) = e^{\lambda t}v^2$$

Es necesario conseguir otra solución independiente. Al suponer, como en (4.23), $x = e^{\lambda t}(u + tw)$, se llega a la misma conclusión

$$(Av - \lambda I)v = 0$$
$$(A - \lambda I)\, u = v$$

Como v es un vector propio de A, entonces es una combinación lineal no trivial (no nula) de v^1 y v^2,

$$v = \alpha_1 v^1 + \alpha_2 v^2 \tag{4.32}$$
$$(A - \lambda I)\, u = \alpha_1 v^1 + \alpha_2 v^2$$
$$(A - \lambda I)\, u - \alpha_1 v^1 - \alpha_2 v^2 = 0 \tag{4.33}$$

Esta última igualdad es un sistema homogéneo de 3 ecuaciones y 5 incógnitas, u_1, u_2, u_3, α_1, α_2. Se debe encontrar una solución donde no sean nulos al tiempo α_1 y α_2.

La matriz de este sistema de ecuaciones es

$$\begin{bmatrix}A - \lambda I & -v^1 & -v^2\end{bmatrix}$$

Dados λ, v^1 y v^2 el proceso es

1) Resolver (4.33) (obtener u, α_1, α_2) con la condición mencionada anteriormente.
2) $v = \alpha_1 v^1 + \alpha_2 v^2$

La solución general es

$$x(t) = c_1\, e^{\lambda t} v^1 + c_2\, e^{\lambda t} v^2 + c_3\, e^{\lambda t}\,(u + tv) \tag{4.34}$$

Ejemplo 4.18.

$$\dot{x} = \begin{bmatrix} -4 & 3 & 5 \\ 2 & -5 & -5 \\ -2 & 3 & 3 \end{bmatrix} x$$

$$\begin{aligned} p(\lambda) &= -\lambda^3 - 6\lambda^2 - 12\lambda - 8 \\ p(\lambda) &= -(\lambda + 2)^3 \\ \lambda_1 = \lambda_2 = \lambda_3 &= -2 \end{aligned}$$

$$A - (-2)I = \begin{bmatrix} -2 & 3 & 5 \\ 2 & -3 & -5 \\ -2 & 3 & 5 \end{bmatrix}$$

$$E_{A+2I} = \begin{bmatrix} 1 & -3/2 & -5/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Hay dos variables libres, es decir hay dos vectores propios linealmente independientes, por ejemplo,

$$v^1 = \begin{bmatrix} 3/2 \\ 1 \\ 0 \end{bmatrix}, \quad v^2 = \begin{bmatrix} 5/2 \\ 0 \\ 1 \end{bmatrix}$$

Por facilidad, multipliquemos por 2,

$$v^1 = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}, \quad v^2 = \begin{bmatrix} 5 \\ 0 \\ 2 \end{bmatrix}$$

Matriz del sistema, $(A - \lambda I)\, u - \alpha_1 v^1 - \alpha_2 v^2 = 0$,

$$\begin{bmatrix} -2 & 3 & 5 & -3 & -5 \\ 2 & -3 & -5 & -2 & 0 \\ -2 & 3 & 5 & 0 & -2 \end{bmatrix}$$

$$E = \begin{bmatrix} 1 & -3/2 & -5/2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} u_1 &= (3/2)u_2 + (5/2)u_3 - \alpha_2 \\ \alpha_1 &= -\alpha_2 \end{aligned}$$

Por ejemplo

$$\begin{aligned} u_1 &= -1 \\ u_2 &= 0 \\ u_3 &= 0 \\ \alpha_1 &= -1 \\ \alpha_2 &= 1 \end{aligned}$$

$$v = -v^1 + v^2 = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$$

Entonces

$$x(t) = c_1\, e^{-2t} \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} + c_2\, e^{-2t} \begin{bmatrix} 5 \\ 0 \\ 2 \end{bmatrix} + c_3\, e^{-2t} \left(\begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix} \right)$$

Supongamos que hay condiciones iniciales,

$$x(1) = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$$

Entonces,

$$\begin{aligned} c_1 &= 129.308482 \\ c_2 &= -96.057729 \\ c_3 &= 114.530370 \end{aligned}$$

4.7 Clasificación de los puntos de equilibrio

Para el caso de un sistema

$$\dot{x} = Ax + b \tag{4.35}$$

si $\det(A) \neq 0$, entonces existe un único punto de equilibrio

$$x^* = -A^{-1}b. \tag{4.36}$$

Si $\det(A) = 0$ pueder haber un número infinito de puntos de equilibrio o ninguno.

El sistema homogéneo,

$$\dot{x} = Ax \tag{4.37}$$

siempre tiene por lo menos un punto de equilibrio, el origen. Si $\det(A) \neq 0$, entonces el único punto de equilibrio es el origen. Si $\det(A) = 0$ hay un número infinito de puntos de equilibrio.

Una matriz es degenerada si

- $\det(A) = 0$
- $\text{mult.alg}(\lambda_i) > 1$ para algún i.
- $\text{real}(\lambda_i) = 0$ y $\text{imag}(\lambda_i) \neq 0$ para algún i.

Un **punto de equilibrio es degenerado** si la matriz es degenerada.

Consideremos el sistema (4.35) y sea x^* un punto de equilibrio.

- x^* es un punto de **equilibrio estable** o **atractor** o **valle**, también llamado asintóticamente estable, si, independientemente del punto inicial,

  $$\lim_{t\to\infty} x(t) = x^*$$

- x^* es un punto de **equilibrio inestable** si no es estable.

- Un punto de equilibrio inestable x^* es un **punto de silla**, si depediendo del punto inicial diferente de x^*, algunas veces

  $$\lim_{t\to\infty} x(t) = x^*$$

  algunas veces

  $$\lim_{t\to\infty} ||x(t)|| = +\infty$$

- Un punto de equilibrio inestable x^* es un **repulsor** o **fuente**, si independientemente del punto inicial diferente de x^*,

$$\lim_{t\to\infty} ||x(t)|| = +\infty$$

Sea $A \in \mathbb{R}^{2\times 2}$, λ_1, λ_2 sus valores propios, $\lambda_1 < \lambda_2$ cuando son reales, $\lambda_1 = \alpha + \beta i$, $\lambda_2 = \alpha - \beta i$ cuando son (estrictamente) complejos. La solución general del sistema homogéneo está dada por uno de los siguientes casos:

$$x(t) = c_1 e^{\lambda_1 t} v^1 + c_2 e^{\lambda_2 t} v^2$$
$$x(t) = c_1 e^{\lambda_1 t} v^1 + c_2 e^{\lambda_1 t} (tv^1 + u)$$
$$x(t) = c_1\, e^{\alpha t}[-\operatorname{sen}(\beta t) q + \cos(\beta t) p] + c_2\, e^{\alpha t}[\operatorname{sen}(\beta t) p + \cos(\beta t) q]$$

El punto de equilibrio se puede clasificar según los valores propios:

reales	$\lambda_1 < \lambda_2 < 0$	no degenerado	atractor
reales	$0 < \lambda_1 < \lambda_2$	no degenerado	repulsor
reales	$\lambda_1 < 0 < \lambda_2$	no degenerado	punto de silla
complejos	$\alpha < 0$	no degenerado	atractor
complejos	$\alpha > 0$	no degenerado	repulsor
complejos	$\alpha = 0$	degenerado	vórtice
reales	$\lambda_1 = \lambda_2 < 0,$	degenerado	atractor
reales	$\lambda_1 = \lambda_2 > 0,$	degenerado	repulsor
reales	$\lambda_1 < \lambda_2 = 0$	degenerado	atractores
reales	$0 = \lambda_1 < \lambda_2$	degenerado	repulsores

Recuérdese que si 0 es un valor propio de A, entonces $\det(A) = 0$ y el sistema

$$\dot{x} = Ax + b$$

puede no tener puntos de equilibrio o puede tener un número infinito de puntos de equilibrio.

En general, para cualquier valor de n:

- si la parte real de todos los valores propios es negativa, el punto de equilibrio es un atractor.
- si la parte real de todos los valores propios es positiva, el punto de equilibrio es un repulsor.
- Si hay valores propios con parte real negativa y valores propios con parte real positiva, el punto de equilibrio es un punto de silla.

Para matrices 2×2, el punto de equilibrio se puede clasificar sin calcular explícitamente los valores propios, utilizando:

$d = \det(A) = a_{11}a_{22} - a_{12}a_{21}$	determinante
$T = \operatorname{tr}(A) = a_{11} + a_{22}$	traza
$\Delta = T^2 - 4d$	discriminante

valores utilizados en el cálculo de los valores propios:

$$p(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21}$$
$$p(\lambda) = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21}$$
$$p(\lambda) = \lambda^2 - T\lambda + d$$
$$\lambda = \frac{T \pm \sqrt{\Delta}}{2}$$

Ver detalles en [Mon10], página 260 y siguientes.

4.8 Diagramas de fase

Para un sistema de ecuaciones diferenciales, el diagrama de fase es la representación gráfica de la solución. Generalmente se hace para dos variables. Cuando hay tres variables es difícil hacerlo aunque no imposible. Para más variables no se hace.

Permite, en muchos casos, tener una descripción cualitativa de la solución, aún si no se conoce de manera precisa la solución.

4.8.1 Generalidades

Por facilidad supongamos que el sistema de ecuaciones diferenciales

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2) \\ \dot{x}_2 &= f_2(x_1, x_2) \end{aligned} \tag{4.38}$$

tiene un único punto de equilibrio x^*. Hay dos **líneas de fase**, las curvas

$$\begin{aligned} f_1(x_1, x_2) &= 0 \\ f_2(x_1, x_2) &= 0 \end{aligned}$$

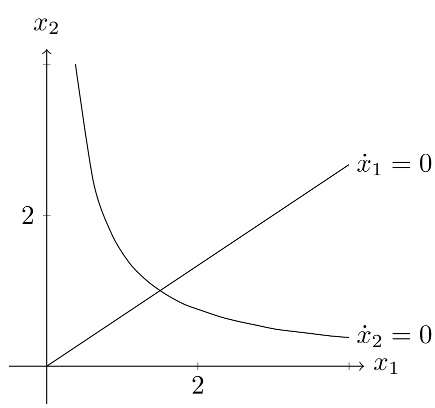
Por la definición, estas dos curvas se cortan justamente en el punto de equilibrio. El siguiente ejemplo permite introducir paso a paso conceptos fundamentales de los diagramas de fase, lineales o no lineales.

Ejemplo 4.19.

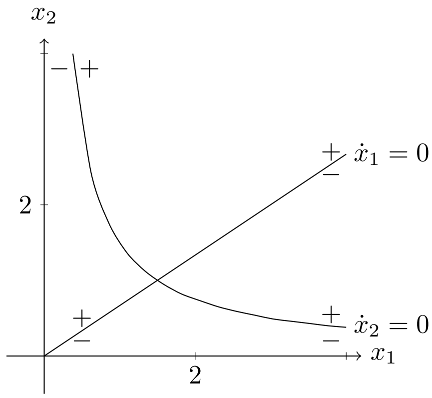
$$\begin{aligned} \dot{x}_1 &= -2x_1 + 3x_2 \\ \dot{x}_2 &= -\frac{3}{x_1} + 2x_2 \\ x_1 &> 0, \;\; x_2 > 0 \end{aligned}$$

El punto de equilibrio es $x^* = (1.5, 1)$. En la siguiente figura están las dos líneas de fase. Ellas son:

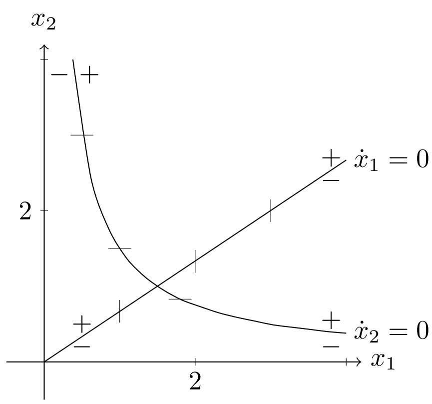
$$\begin{aligned} \dot{x}_1 = 0 : &\qquad x_2 = \frac{2}{3}x_1 \\ \dot{x}_2 = 0 : &\qquad x_2 = \frac{3}{2x_1} \end{aligned}$$



Sobre la primera línea de fase $\dot{x}_1 = 0$, entonces de un lado $\dot{x}_1$ será positivo y del otro lado negativo. Basta con tomar un punto que no esté sobre la línea de fase. Por ejemplo, si $x = (2, 0)$, $\dot{x}_1 = -4$, entonces de ese mismo lado de la línea de fase, $\dot{x}_1$ será negativo y del otro lado $\dot{x}_1$ sera positivo. Esto se indica colocando signos menos y signos más. El mismo punto sirve para la otra línea de fase, para $x = (2, 0)$, $\dot{x}_2 = -3/2$.

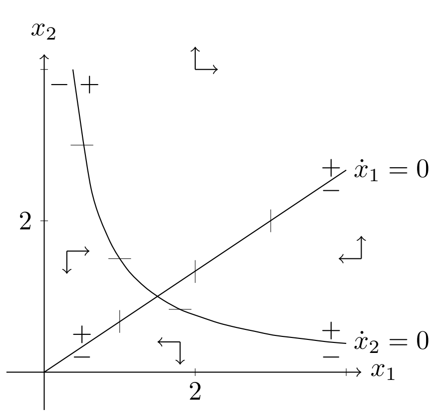


Si la trayectoria o solución del sistema de ecuaciones diferenciales, para una condición inicial dada corta la línea de fase $\dot{x}_1 = 0$, lo hará verticalmente. Si la trayectoria corta la línea de fase $\dot{x}_2 = 0$, lo hará horizontalmente. Esto se indica por medio de pequeños segmentos de recta verticales y horizontales.



En este ejemplo, las dos líneas de fase dividen la región de trabajo, el primer cuadrante, en cuatro sectores. Consideremos el sector que queda encima del punto de equilibrio. En este sector $\dot{x}_1 > 0$ y $\dot{x}_2 > 0$, esto quiere decir que si la trayectoria está en ese sector x_1 está aumentando y x_2 también aumenta. Esto se indica por el símbolo ↳ Es decir, si la trayectoria está en este sector, lleva aproximadamente la dirección nordeste (o noreste), NE. De manera semejante en los otros casos (noroeste, sudeste o sureste, suroeste).

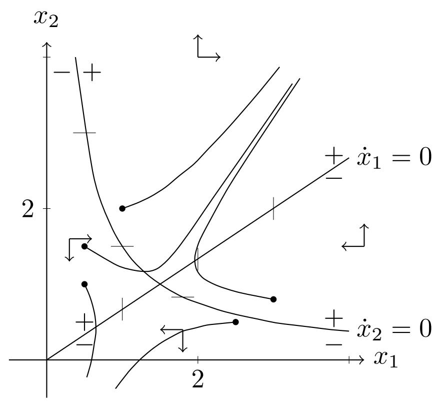
$\dot{x}_1 > 0, \quad \dot{x}_2 > 0$	↳	NE
$\dot{x}_1 > 0, \quad \dot{x}_2 < 0$	↱	SE
$\dot{x}_1 < 0, \quad \dot{x}_2 > 0$	↲	NO
$\dot{x}_1 < 0, \quad \dot{x}_2 < 0$	↴	SO



Las gráficas de la solución del sistema de ecuaciones diferenciales con puntos iniciales diferentes

$$x(0) = (3, 0.8), \qquad (1, 2), \qquad (2.5, 0.5), \qquad (0.5, 1), \qquad (0.5, 1.5)$$

pueden ser semejantes a



◇

4.8.2 Sistemas con coeficientes constantes

Para los sistemas de ecuaciones diferenciales lineales con coeficientes constantes, además de las explicaciones anteriores, hay conceptos adicionales. Supongamos que $\det(A) \neq 0$ (hay un único punto de equilibrio x^*).

- Las líneas de fase son rectas que se cruzan en el punto de equilibrio.
- Si los valores propios son reales, sea

$$\lambda_1 \leq \lambda_2$$

- Si hay dos vectores propios independientes, sean ellos v^1 y v^2.
- Si hay uno solo, sea $v = v^1$ este vector propio, y sea u calculado según (4.25).
- Los vectores propios (o el vector propio) definen dos rectas que pasan por el punto de equilibrio y son paralelas al vector propio

$$R_1 = \{x^* + \sigma v^1 \ : \ \sigma \in \mathbb{R}\} \tag{4.39}$$
$$R_2 = \{x^* + \sigma v^2 \ : \ \sigma \in \mathbb{R}\} \tag{4.40}$$

  Obviamente si los valores propios son complejos, estas dos rectas no están definidas.

- Para matrices diagonalizables, hay cuatro rectas que se cruzan en el punto de equilibrio:

$$\dot{x}_1 = 0, \qquad \dot{x}_2 = 0, \qquad R_1, \qquad R_2$$

En el siguiente ejemplo están estos conceptos. El estudio de la forma de la solución estará detallado en las subsecciones siguientes.

Ejemplo 4.20.

$$\dot{x} = \begin{bmatrix} -2 & -12 \\ -15 & 1 \end{bmatrix} x + \begin{bmatrix} 16 \\ 29 \end{bmatrix}$$

Supongamos que hay tres condiciones iniciales diferentes:

$$x(0) = (3.5, 3), \qquad (0, -0.5), \qquad (0, -1.5)$$

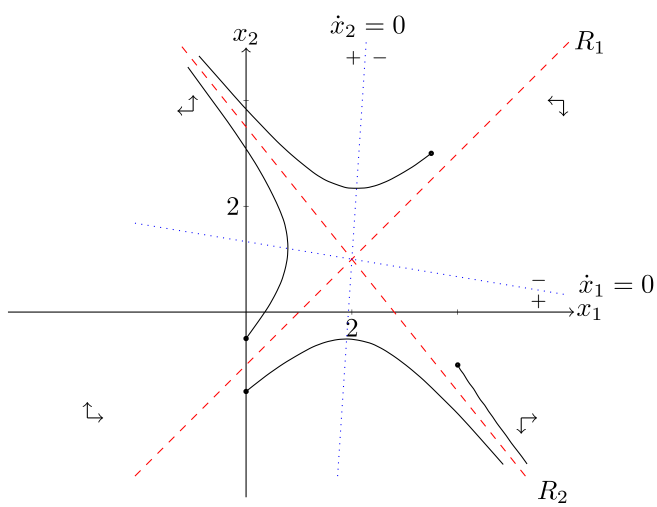
$$
\begin{aligned}
x^* &= \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\
p(\lambda) &= \lambda^2 + \lambda - 182 \\
\lambda_1 &= -14 \\
\lambda_2 &= 13 \\
v^1 &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
v^2 &= \begin{bmatrix} -4 \\ 5 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\dot{x}_1 = 0 &= -2x_1 - 12x_2 + 16 \\
\dot{x}_2 = 0 &= -15x_1 + x_2 + 29 \\
R_1 &= \{(2,1) + \sigma(1,1) : \sigma \in \mathbb{R}\} \\
R_2 &= \{(2,1) + \sigma(-4,5) : \sigma \in \mathbb{R}\}
\end{aligned}
$$

La solución general será

$$
x(t) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_1\, e^{-14t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\, e^{13t} \begin{bmatrix} -4 \\ 5 \end{bmatrix}
$$

Para las tres condiciones iniciales diferentes, la trayectoria puede ser semejante a las de la siguiente gráfica:



Obsérvese que las trayectorias cuando cortan las líneas de fase, lo hacen según lo esperado, verticalmente u horizontalmente a las líneas de fase, y siguen aproximadamente las direcciones NE, NO, SE, SO dependiendo del sector. $\diamond$

Cuando hay dos valores propios diferentes, como se sabe, la solución general es

$$
x(t) = x^* + c_1\, e^{\lambda_1 t} v^1 + c_2\, e^{\lambda_2 t} v^2
$$

y siempre se cumple lo siguiente:

- Si x^0 no está en R_1 ni en R_2, entonces la trayectoria, cuando $t > t_0$, nunca corta las rectas R_1 y R_2.
- Si x^0 está en la recta R_1, entonces $c_2 = 0$ y la trayectoria está contenida completamente en la recta R_1. Si $\lambda_1 < 0$ la trayectoria tiende hacia el punto de equilibrio. Si $\lambda_1 > 0$ la trayectoria se aleja indefinidamente del punto de equilibrio.

- Análogamente, si x^0 está en la recta R_2, entonces $c_1 = 0$ y la trayectoria está completamente contenida en la recta R_2. Si $\lambda_2 < 0$ la trayectoria tiende hacia el punto de equilibrio. Si $\lambda_2 > 0$ la trayectoria se aleja indefinidamente del punto de equilibrio.

En las subsecciones siguientes, se supone que el sistema de ecuaciones diferenciales es homogéneo. Así, las dos líneas de fase y las rectas R_i se cortan en el origen. Entonces para un sistema no homogéneo el diagrama de fase se desplaza "del origen al punto de equilibrio".

$$R_1 = \{\sigma v^1 \, : \, \sigma \in \mathbb{R}\} \tag{4.41}$$
$$R_2 = \{\sigma v^2 \, : \, \sigma \in \mathbb{R}\} \tag{4.42}$$

4.8.3 Atractor $\lambda_1 < \lambda_2 < 0$

Independientemente del punto inicial

$$\lim_{t\to\infty} x(t) = (0,0)$$

- Si x^0 está en R_1, entonces $c_2 = 0$ y la trayectoria se acerca al origen sobre la misma recta.
- Si x^0 está en R_2, entonces $c_1 = 0$ y la trayectoria se acerca al origen sobre la misma recta.
- Si x^0 no está en R_1 ni R_2, la trayectoria tiende hacia el origen, pero trata de acercarse a R_2.

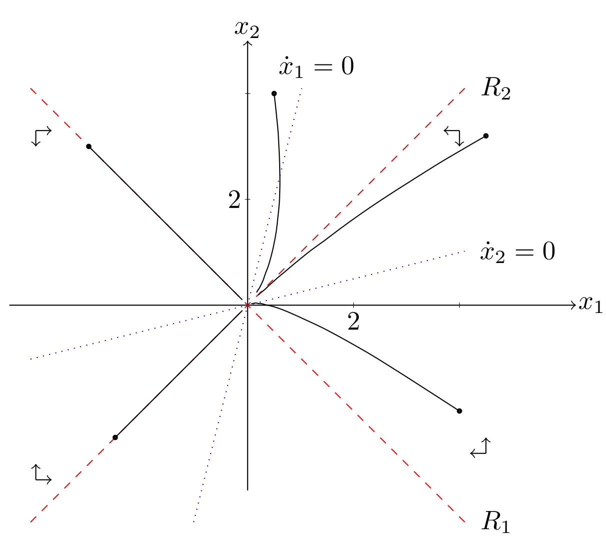
Ejemplo 4.21.

$$\dot{x}_1 = -4x_1 + x_2$$
$$\dot{x}_2 = x_1 - 4x_2$$

$$\begin{aligned}
p(\lambda) &= \lambda^2 + 8\lambda + 15\\
\lambda_1 &= -5\\
\lambda_2 &= -3\\
v^1 &= (-1,1)\\
v^2 &= (1,1)\\
(x_1(t), x_2(t)) &= c_1e^{-5t}(-1,1) + c_2e^{-3t}(1,1)
\end{aligned}$$

Consideremos varias condiciones iniciales

$$x(0) = (0.5,4) \qquad (-2.5,-2.5) \qquad (-3,3) \qquad (4,-2) \qquad (4.5,3.2)$$



4.8.4 Repulsor $0 < \lambda_1 < \lambda_2$

Supongamos que $x^0 \neq x^*$,

$$\lim_{t\to\infty} x(t) = (\pm\infty, \pm\infty)$$

es decir, la trayectoria se aleja indefinidamente del origen.

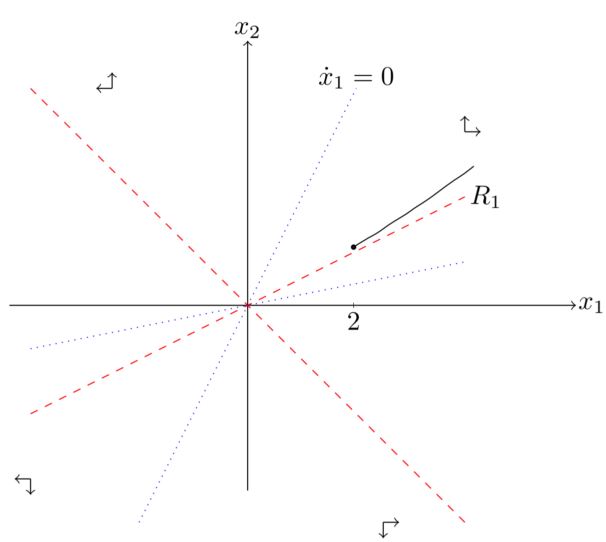
- Si x^0 está en R_1, entonces $c_2 = 0$ y la trayectoria se aleja del origen sobre la misma recta.
- Si x^0 está en R_2, entonces $c_1 = 0$ y la trayectoria se aleja del origen sobre la misma recta.
- Si x^0 no está en R_1 ni en R_2, la trayectoria se aleja del origen y, para t grande, tiende a volverse paralela a la recta R_2. Dependiendo del cociente λ_2/λ_1, la tendencia es más rápida o más lenta.

Ejemplo 4.22. $\dot{x} = Ax$,

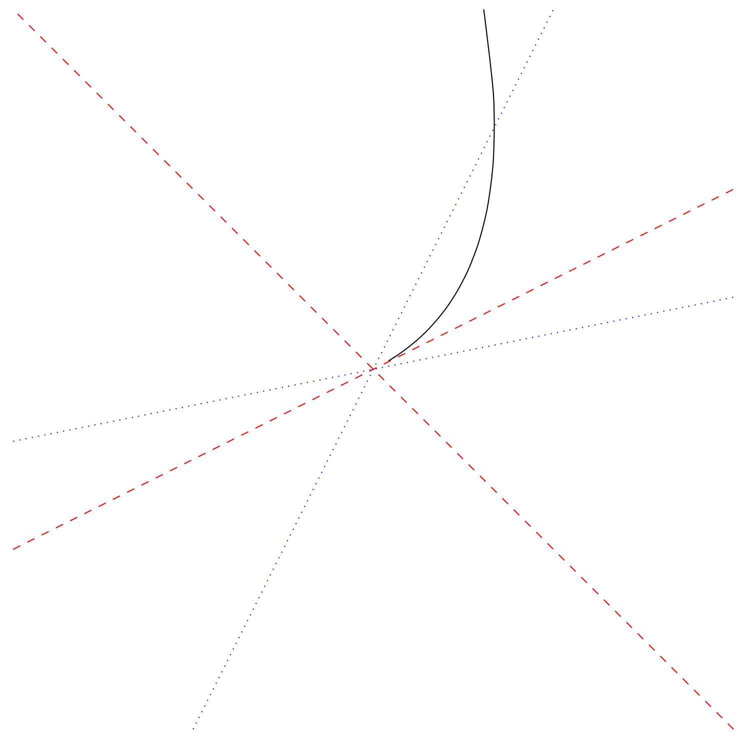
$$A = \begin{bmatrix} 4/3 & -2/3 \\ -1/3 & 5/3 \end{bmatrix}$$
$$x(1) = (2, 1.1)$$

$$p(\lambda) = \lambda^2 - 3\lambda + 2$$
$$\lambda_1 = 1$$
$$\lambda_2 = 2$$
$$v^1 = (2, 1)$$
$$v^2 = (-1, 1)$$
$$c_1 = 0.3801421$$
$$c_2 = 0.0090224$$

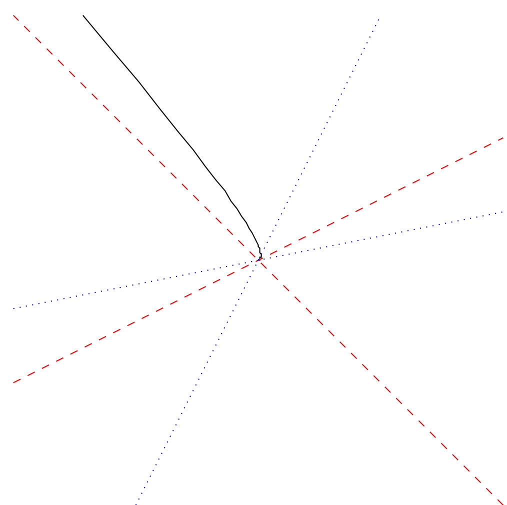
Así, $x(1.8) \approx (4.27, 2.63)$ y la gráfica es aproximadamente,



Para estos valores de t, la trayectoria no parece paralela a R_2. Para valores más grandes de t, $x(4) \approx (14.61, 47.65)$.



Todavía la trayectoria no parece paralela a R_2. Finalmente, $x(6) \approx (-1161, 1621)$,



4.8.5 Punto de silla : $\lambda_1 < 0 < \lambda_2$

- Si x^0 no está en R_1 ni en R_2, para valores suficientemente grandes de t, la trayectoria se aleja indefinidamente del origen,

$$\lim_{t\to\infty} x(t) = (\pm\infty, \pm\infty)$$

  acercándose a la recta R_2.

- Si x^0 está en R_1

$$\lim_{t\to\infty} x(t) = (0, 0)$$

  y la trayectoria permanece en la recta R_1.

- Si x^0 está en R_2, la trayectoria permanece en R_2 alejándose del origen.

- Si x^0 no está en R_1, pero está muy cerca, inicialmente la trayectoria tiende hacia el punto de equilibrio, pero para valores de t más grandes, la trayectoria tenderá a alejarse del punto de equilibrio acercándose a la recta R_2.

Ejemplo 4.23.

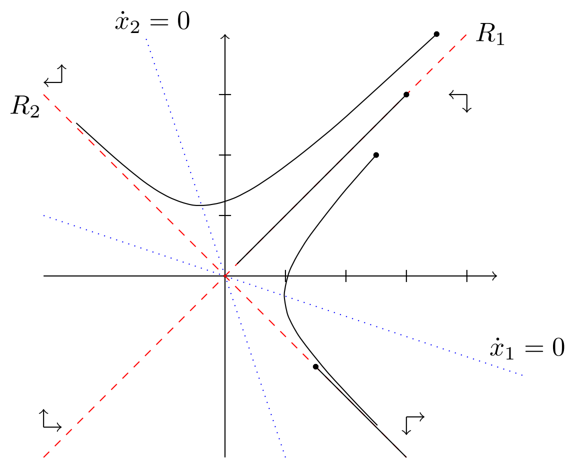
$$A = \begin{bmatrix} -1/2 & -3/2 \\ -3/2 & -1/2 \end{bmatrix}, \quad b = 0$$

con varios puntos iniciales

$$x^0 = (3.5, 4)$$

????

```
pol. caract. = -2 + 1 x + x^2
val. prop :    -2   1
v1 :    1   1
v2 :    -1   1
```



4.8.6 Atractor, valores propios complejos, $\alpha < 0$

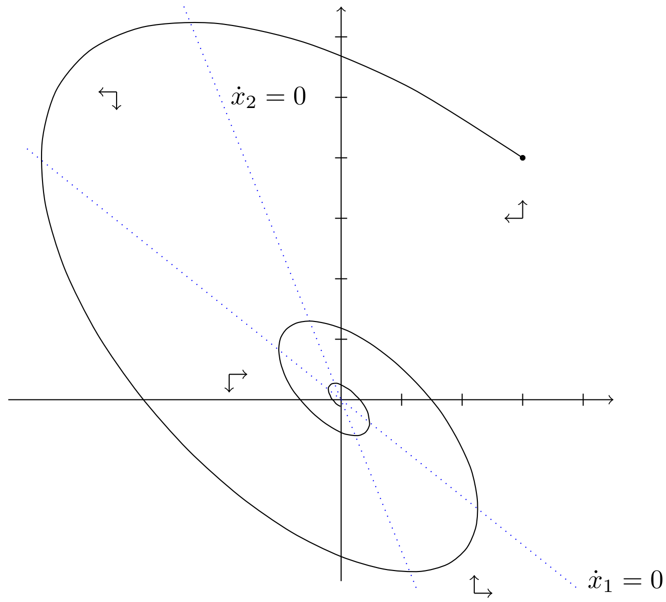
$$\begin{aligned} \lambda_1 &= \alpha + \beta i, \quad \alpha < 0 \\ \lambda_2 &= \alpha - \beta i \end{aligned}$$

La trayectoria forma una espiral convergente hacia el origen (o hacia el punto de equilibrio).

Ejemplo 4.24.

$$A = \begin{bmatrix} -4 & -5 \\ 5 & 2 \end{bmatrix}$$
$$x(0) = (3, 4), \quad 0 \le t \le 4$$

$$p(\lambda) = \lambda^2 + 2\lambda + 17$$
$$\lambda = -1 \pm 4i$$



4.8.7 Repulsor, valores propios complejos, $\alpha > 0$

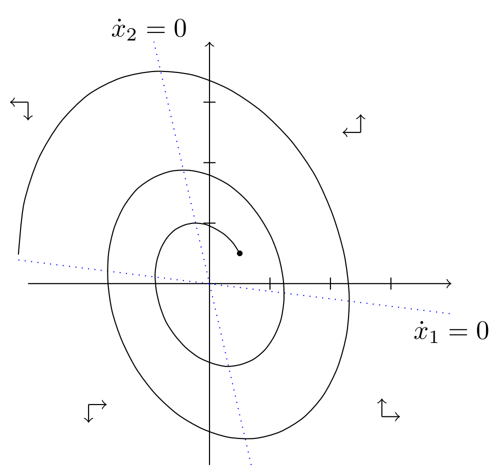
$$\begin{aligned} \lambda_1 &= \alpha + \beta i, \quad \alpha > 0 \\ \lambda_2 &= \alpha - \beta i \end{aligned}$$

La trayectoria forma una espiral divergente alejándose del origen (o del punto de equilibrio).

Ejemplo 4.25.

$$A = \begin{bmatrix} -1 & -8 \\ 13 & 3 \end{bmatrix}$$
$$x(0) = (0.5, 0.5), \quad 0 \leq t \leq 1.5$$

$$\begin{aligned} p(\lambda) &= \lambda^2 - 2\lambda + 101 \\ \lambda &= 1 \pm 10i \end{aligned}$$



4.8.8 Vórtice, valores propios complejos, $\alpha = 0$

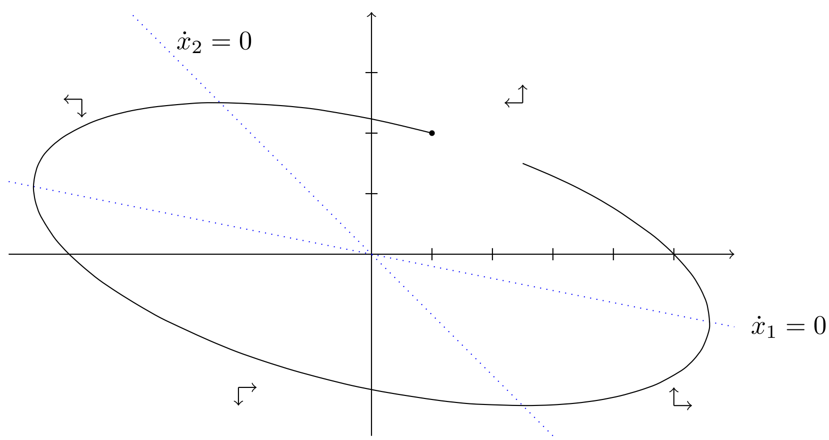
$$
\begin{aligned}
\lambda_1 &= \alpha + \beta i, \quad \alpha = 0 \\
\lambda_2 &= \alpha - \beta i
\end{aligned}
$$

La trayectoria forma una elipse centrada en el origen (o en el punto de equilibrio).

Ejemplo 4.26.

$$
A = \begin{bmatrix} -2 & -10 \\ 2 & 2 \end{bmatrix}
$$
$$
x(0) = (1, 2), \quad 0 \le t \le 1.5
$$

$$
p(\lambda) = \lambda^2 + 16
$$
$$
\lambda = \pm 4i
$$



4.8.9 Atractor, $\lambda_1 = \lambda_2 < 0$, dos vectores propios

Sea $\lambda = \lambda_1 = \lambda_2 < 0$. Si hay dos vectores propios independientes, $E_{A-\lambda I}$ tiene dos variables libres, es decir, ninguna variable básica, o sea, $E_{A-\lambda I} = 0$. Necesariamente $A - \lambda I = 0$, $A = \lambda I$,

$$
A = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}
$$

El sistema es no encadenado o separable,

$$
\begin{aligned}
\dot{x}_1 &= \lambda x_1 \\
\dot{x}_2 &= \lambda x_2
\end{aligned}
$$

o sea, cada ecuación diferencial se puede resolver por aparte.

Sin embargo, continuemos considerándolo como un sistema. Sean v^1, $v^2 \in \mathbb{R}^{2\times 1}$ dos vectores propios independientes, en este caso, dos vectores cualesquiera linealmente independientes. Sea V la matriz cuyas columnas son

v^1 y v^2,

$$\begin{aligned} V &= \begin{bmatrix} v^1 & v^2 \end{bmatrix} \\ x(t) &= c_1 e^{\lambda t} v^1 + c_2 e^{\lambda t} v^2 \\ x(t) &= e^{\lambda t}(c_1 v^1 + c_2 v^2) \\ x(t) &= e^{\lambda t} \begin{bmatrix} v^1 & v^2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\ x(t) &= e^{\lambda t} V c \\ x(t_0) &= x^0 \\ x^0 &= e^{\lambda t_0} V c \\ c &= e^{-\lambda t_0} V^{-1} x^0 \\ x(t) &= e^{\lambda t} V e^{-\lambda t_0} V^{-1} x^0 \\ x(t) &= e^{\lambda (t - t_0)} x^0 \end{aligned}$$

Siempre $x(t)$ es un múltiplo de x^0. Como $\lambda < 0$,

$$\lim_{t \to \infty} x(t) = (0, 0).$$

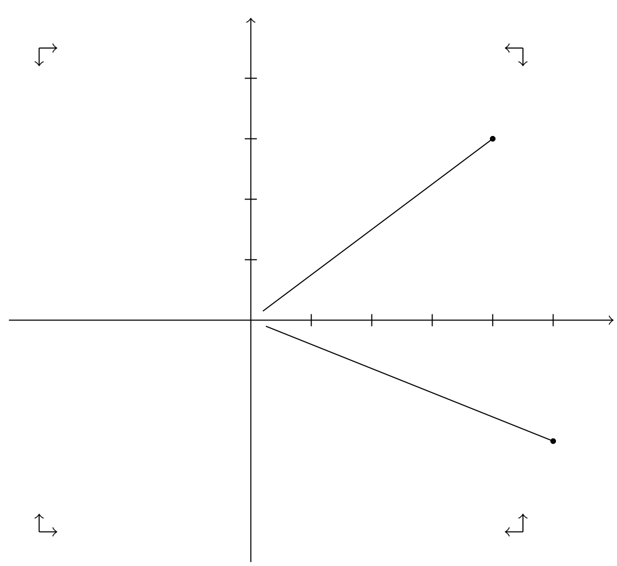
La trayectoria es radial, desde el punto x^0 hacia el origen. Como los vectores propios pueden ser cualquier par independiente, no tiene sentido dibujar las rectas R_1 y R_2. Ademas la línea de fase $\dot{x}_1 = 0$ es simplemente $\lambda x_1 = 0$, $x_1 = 0$, o sea, el eje x_2. De manera análoga, la segunda línea de fase es el eje x_1.

Ejemplo 4.27.

$$\begin{aligned} \dot{x}_1 &= -2x_1 \\ \dot{x}_2 &= -2x_2 \\ x(1) &= \begin{bmatrix} 4 \\ 3 \end{bmatrix} \end{aligned}$$

La solución:

$$\begin{aligned} x(t) &= e^{-2(t-1)} \begin{bmatrix} 4 \\ 3 \end{bmatrix} \\ \lim_{t \to \infty} x(t) &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$



En la gráfica también está la trayectoria si $x(2) = (5, -2)$. $\diamond$

4.8.10 Repulsor, $0 < \lambda_1 = \lambda_2$, dos vectores propios

Con el mismo razonamiento anterior, la solución es

$$x(t) = e^{\lambda(t-t_0)}x^0$$

Como $\lambda > 0$, entonces la trayectoria es radial, alejándose indefinidamente del punto inicial.

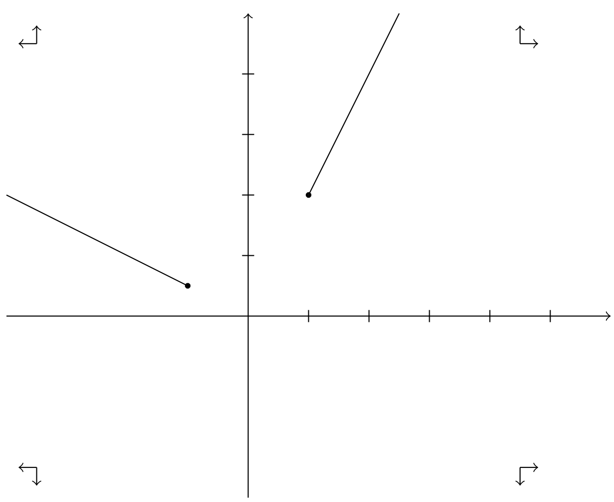
$$\lim_{t\to\infty} x(t) = \begin{bmatrix}\pm\infty\\ \pm\infty\end{bmatrix}$$

Ejemplo 4.28.

$$\begin{aligned}\dot{x}_1 &= 3x_1\\ \dot{x}_2 &= 3x_2\\ x(5) &= \begin{bmatrix}-1\\ 1/2\end{bmatrix}\end{aligned}$$

La solución:

$$\begin{aligned}x(t) &= e^{3(t-5)}\begin{bmatrix}-1\\ 1/2\end{bmatrix}\\ \lim_{t\to\infty} x(t) &= \begin{bmatrix}-\infty\\ \infty\end{bmatrix}\end{aligned}$$



En la gráfica también está la trayectoria si $x(1.5) = (1, 2)$. $\diamond$

4.8.11 Atractor, $\lambda_1 = \lambda_2 < 0$, un vector propio

Sea v el vector propio asociado a $\lambda = \lambda_1 = \lambda_2$. Cualquier otro vector propio es múltipo de v. La solución general es

$$\begin{aligned}x(t) &= c_1e^{\lambda t}v + c_2e^{\lambda t}(tv + u)\\ \text{donde}\quad (A - \lambda I)u &= v\end{aligned}$$

Si x^0 está en la recta $R = \{\sigma v : \sigma \in \mathbb{R}\}$, es decir, $c_2 = 0$, la trayectoria es radial hacia el origen. Si x^0 no está en la recta de v, la trayectoria se acerca hacia el origen y al mismo tiempo hacia la recta R.

Ejemplo 4.29.

$$\begin{aligned}\dot{x}_1 &= -8x_1 - 4x_2\\ \dot{x}_2 &= x_1 - 12x_2\end{aligned}$$

Falta ???? v, u

$$
\begin{aligned}
x(t) &= c_1\, e^{-10t} v + c_2 e^{-10t}(tv + w)\\
x(t) &= e^{-2t}\big(v(c_1 + c_2 t) + c_2 w\big)\\
\lim_{t\to\infty} x(t) &= (0,0)
\end{aligned}
$$

???? Falta gráfica

4.8.12 Repulsor, $0 < \lambda_1 = \lambda_2$, un vector propio

??????????

4.8.13 $\lambda_1 < \lambda_2 = 0$, sistema homogéneo

Supongamos que ninguna de las dos filas de A es nula. Como 0 es un valor propio, la matriz A es singular (no invertible, $\det(A) = 0$), las dos filas son linealmente dependientes, una es múltiplo de la otra, las dos líneas de fase se confunden en una sola que además coincide con R_2. Sea v^1 un vector propio asociado a $\lambda_1 < 0$ y v^2 un vector propio asociado a $\lambda_2 = 0$. La solución general es

$$x(t) = c_1 e^{\lambda_1 t} v^1 + c_2 v^2$$

- Si $x^0 \in R_1$, es decir, x^0 es múltiplo de v^1, entonces $c_2 = 0$, la trayectoria va en línea recta desde x^0 hacia el origen, permaneciendo en la recta R_1.

$$
\begin{aligned}
x(t) &= c_1 e^{\lambda_1 t} v^1\\
x(t) &\to (0,0)
\end{aligned}
$$

- Si $x^0 \in R_2$, es decir, x^0 es múltiplo de v^2, entonces $c_1 = 0$, x^0 es un punto de equilibrio y la trayectoria no se mueve de x^0

$$x(t) = x^0 \qquad \forall t$$

- Si $x^0 \notin R_1$ y $x^0 \notin R_2$, entonces las dos constantes son no nulas,

$$
\begin{aligned}
x(t) &= c_1 e^{\lambda_1 t} v^1 + c_2 v^2\\
x(t) &\to c_2 v^2
\end{aligned}
$$

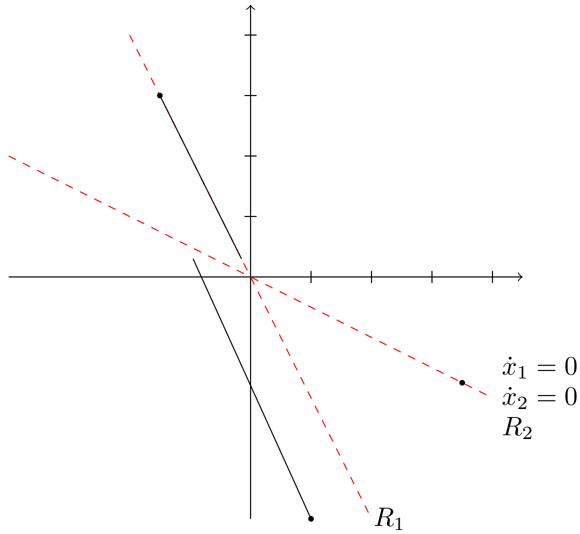
  La trayectoria se hace a lo largo de una línea recta, empezando en el punto inicial, desplazándose de forma paralela a la recta R_1, acercándose a la recta R_2. El límite, cuando t tiende a infinito, será $c_2 v^2$

Ejemplo 4.30.

$$
\begin{aligned}
\dot{x} &= Ax\\
A &= \begin{bmatrix} 1 & 2\\ -2 & -4 \end{bmatrix}\\
x(0.5) &= \begin{bmatrix} -1.5\\ 3 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
p(\lambda) &= \lambda^2 + 3\lambda\\
\lambda_1 &= -3\\
\lambda_2 &= 0\\
v^1 &= \begin{bmatrix} -1\\ 2 \end{bmatrix}\\
v^2 &= \begin{bmatrix} -2\\ 1 \end{bmatrix}
\end{aligned}
$$

x^0 es múltiplo de v^1, luego la trayectoria parte del punto incial y tiende hacia el origen.



Consideremos ahora el mismo sistema de ecuaciones diferenciales pero con otro punto inicial.

$$x(1) = \begin{bmatrix} 3.50 \\ -1.75 \end{bmatrix}$$

Este punto está sobre la recta R_2 que coincide con las dos líneas de fase, $\dot{x}_1 = 0$ y $\dot{x}_2 = 0$, es decir, x^0 es un punto de equilibrio y la trayectoria se queda, para cualquier t, en ese punto.

Consideremos de nuevo el mismo sistema de ecuaciones diferenciales pero con otro punto inicial.

$$x(0.1) = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$$

Este punto no está sobre la recta R_1 ni sobre la recta R_2. La trayectoria se hace a lo largo de una línea recta, empezando en el punto inicial, desplazándose de forma paralela a la recta R_1, acercándose a la recta R_2. El límite, cuando t tiende a infinito, será c_2v^2. En este caso si se hubiera hecho el cálculo de las cosntantes, la trayectoria tiende al punto $(-1.0664, 0.5329)$.

4.8.14 $0 = \lambda_1 < \lambda_2$, sistema homogéneo

Este caso presenta algunos resultados idénticos al caso anterior y otros son análogos.

Supongamos que ninguna de las dos filas de A es nula. Como 0 es un valor propio, la matriz A es singular (no invertible, $\det(A) = 0$), las dos filas son linealmente dependientes, una es múltiplo de la otra, las dos líneas de fase se confunden en una sola que además coincide con R_1. Sea v^1 un vector propio asociado a $\lambda_1 = 0$ y v^2 un vector propio asociado a $\lambda_2 > 0$. La solución general es

$$x(t) = c_1v^1 + c_2e^{\lambda_2 t}v^2$$

- Si $x^0 \in R_1$, es decir, x^0 es múltiplo de v^1, entonces $c_2 = 0$, es decir, x^0 es un punto de equilibrio y la trayectoria no se mueve de x^0

$$x(t) = x^0 \qquad \forall t$$

- Si $x^0 \in R_2$, es decir, $c_1 = 0$, la trayectoria va en línea recta desde x^0, permanece en la recta R_2 pero va alejándose del origen.

$$x(t) = c_2e^{\lambda_2 t}v^2$$
$$x(t) \to (\pm\infty, \pm\infty)$$

- Si $x^0 \notin R_1$ y $x^0 \notin R_2$, entonces las dos constantes son diferentes de 0,

$$x(t) = c_1 v^1 + c_2 e^{\lambda_2 t} v^2$$
$$x(t) \to (\pm\infty, \pm\infty)$$

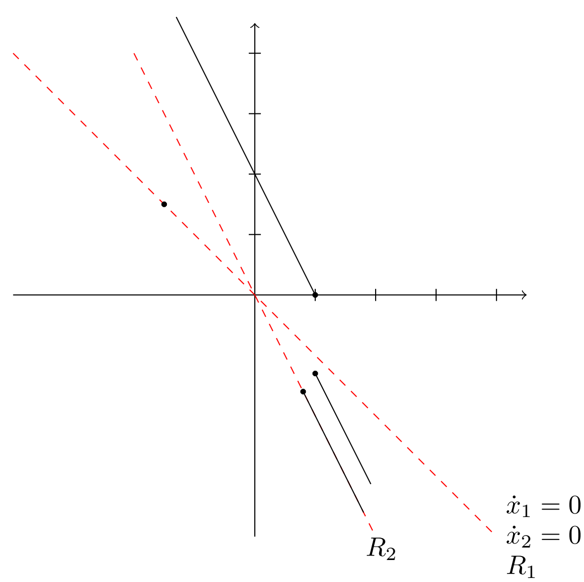
La trayectoria se hace a lo largo de una línea recta, empezando en el punto inicial, desplazándose de forma paralela a la recta R_2, alejándose indefinidamente del punto inicial.

Ejemplo 4.31.

$$\dot{x} = Ax$$
$$A = \begin{bmatrix} -1 & -1 \\ 2 & 2 \end{bmatrix}$$
$$x(0.5) = \begin{bmatrix} -1.5 \\ 1.5 \end{bmatrix}$$

$$p(\lambda) = \lambda^2 - \lambda$$
$$\lambda_1 = 0$$
$$\lambda_2 = 1$$
$$v^1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$
$$v^2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

x^0 es múltiplo de v^1, punto de equilibrio, luego la trayectoria no se mueve del punto inicial.



Consideremos ahora el mismo sistema de ecuaciones diferenciales pero con otro punto inicial.

$$x(1) = \begin{bmatrix} 0.8 \\ -1.6 \end{bmatrix}$$

Este punto está sobre la recta R_2, luego $c_1 = 0$. La trayectoria parte del punto inicial, sigue sobre la recta R_2, alejándose del origen.

Consideremos de nuevo el mismo sistema de ecuaciones diferenciales pero con otro punto inicial.

$$x(2) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Este punto no está sobre la recta R_1 ni sobre la recta R_2. La trayectoria se hace a lo largo de una línea recta, empezando en el punto inicial, desplazándose de forma paralela a la recta R_2 y alejándose indefinidamente del punto inicial. Hay dos sentidos para alejarse del punto inicial. ¿Cuál? En el sentido donde no se corte la recta R_1. Dicho de forma equivalente, alejándose de la recta R_1.

En la gráfica está parte de la trayectoria si $x^0 = (1, -1.3)$. $\diamond$

4.8.15 $\lambda_1 < 0 = \lambda_2$, sistema no homogéneo consistente

Las dos líneas de fase coinciden en una sola, no pasan por el origen. Todos los puntos sobre la línea de fase son puntos de equilibrio. Sean R_1 y R_2 que pasan por el origen (posición arbitraria). Otra opción es hacer coincidir R_2 con las líneas de fase (de todas maneras son rectas paralelas).

Si x^0 está en la línea de fase, la trayectoria no se mueve del punto de equilibrio.

Si x^0 no está en la línea de fase, la trayectoria tiende hacia la línea de fase, donde están los puntos de equilibrio, de manera paralela a la recta R_1.

4.9 Utilización de la exponencial de una matriz

Si $x \in \mathbb{R}$

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{4.43}$$

De manera análoga, si $A \in \mathbb{R}^{n\times n}$, se define la exponencial de la matriz, $e^A \in \mathbb{R}^{n\times n}$:

$$e^A = I + A + \frac{A^2}{2} + \frac{A^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{A^k}{k!} \tag{4.44}$$

Propiedades (algunas son consecuencias triviales de otras):

$$e^{\mathbf{0}} = I \tag{4.45a}$$
$$e^{A+B} = e^A e^B \tag{4.45b}$$
$$e^A e^{-A} = I \tag{4.45c}$$
$$e^A \quad \text{siempre es invertible} \tag{4.45d}$$
$$\left(e^A\right)^{-1} = e^{-A} \tag{4.45e}$$
$$\det(e^A) = e^{\operatorname{tr}(A)} \tag{4.45f}$$
$$e^{(A^{\mathrm{T}})} = \left(e^A\right)^{\mathrm{T}} \tag{4.45g}$$
$$e^{BAB^{-1}} = Be^AB^{-1} \qquad \text{si } B \text{ es invertible} \tag{4.45h}$$
$$e^A e^B = e^B e^A \qquad \text{si } AB = BA \tag{4.45i}$$
$$\text{si } D \text{ es diagonal}, \quad e^D \ \text{ es diagonal, y } \ [e^D]_{ii} = e^{d_{ii}} \tag{4.45j}$$
$$\frac{\mathrm{d}}{\mathrm{dt}} e^{tA} = A\, e^{tA} \tag{4.45k}$$
$$||e^A|| \le e^{||A||} \tag{4.45l}$$

En Scilab y Matlab la exponencial de una matriz se calcula por medio de

```
expm(A)
```

Por ejemplo

$$A = \begin{bmatrix} 2 & -4 \\ -2 & -5 \end{bmatrix}, \qquad e^A = \begin{bmatrix} 17.854086 & -8.925804 \\ -4.462902 & 2.233930 \end{bmatrix}$$

Si se utiliza `exp(A)`, calcula una matriz cuyas entradas son $e^{a_{ij}}$.

4.9.1 Ecuación diferencial lineal con coeficientes constantes

Consideremos la siguiente ecuación diferencial, donde $x = x(t)$, $a \in \mathbb{R}$

$$\dot{x} = ax \tag{4.46}$$
$$x(t_0) = x_0 \tag{4.47}$$

La solución general es

$$x = c\, e^{at} \tag{4.48}$$

Al tener en cuenta la condición inicial,

$$c = e^{-at_0} x_0$$
$$x = x_0\, e^{-at_0}\, e^{at}$$
$$x = x_0\, e^{a(t-t_0)} \tag{4.49}$$

La anterior solución es válida, aún si $a = 0$.

Consideremos ahora la ecuación diferencial no homogénea, con $a \neq 0$,

$$\dot{x} = ax + b \tag{4.50}$$
$$x(t_0) = x_0 \tag{4.51}$$

La solución general es

$$x = c\, e^{at} - \frac{b}{a} \tag{4.52}$$

Teniendo en cuenta la condición inicial,

$$x = (x_0 + \frac{b}{a})\, e^{a(t-t_0)} - \frac{b}{a} \tag{4.53}$$

4.9.2 Sistema homogéneo de ecuaciones diferenciales lineales con coeficientes constantes

$$\dot{x} = Ax \tag{4.54}$$
$$x(t_0) = x^0$$

Aquí, $A \in \mathbb{R}^{n\times n}$, x, $\dot{x}$ y x^0 son vectores columna. De manera análoga a (4.48), la solución general es

$$x = e^{tA} c \tag{4.55}$$

con $c = [c_1 \;\; c_2 \;\; \cdots \;\; c_n]^{\mathrm{T}}$ un vector columna. Al tener en cuenta la condición inicial,

$$c = \left(e^{t_0 A}\right)^{-1} x^0$$
$$c = e^{-t_0 A}\, x^0 \tag{4.56}$$
$$x = e^{tA} e^{-t_0 A}\, x^0$$
$$x = e^{(t-t_0)A} x^0 \tag{4.57}$$

Esta última igualdad es la generalización de (4.49).

La aplicación directa de esta fórmula, con software que permita calcular e^M, permite calcular rápidamente el vector x para un t dado. Puede ser aplicando directamente (4.57), o bien, calculando primero c por (4.56) y después (4.55).

Ejemplo 4.32. Calcular $x(1.5)$ para la solución del sistema

$$\dot{x} = \begin{bmatrix} 2 & -4 \\ -2 & -5 \end{bmatrix} x, \qquad x(1) = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

$$c = \begin{bmatrix} 1031.006822 \\ 2061.963857 \end{bmatrix}, \quad e^{1.5A} = \begin{bmatrix} 80.015242 & -40.007559 \\ -20.003780 & 10.002013 \end{bmatrix}, \quad x(1.5) = \begin{bmatrix} 2.119095 \\ -0.243498 \end{bmatrix} \quad \diamond$$

La fórmula (4.57) permite, para un t dado, calcular $x(t)$ pero no da una expresión analítica de $x(t)$ y no permite hacer un estudio cualitativo de la solución. Por ejemplo no se sabe qué pasa cuando t tiende a infinito.

4.9.3 Matriz diagonalizable

Sea A diagonalizable, sea V la matriz cuyas n columnas son vectores propios linealmente independientes y sea D la matriz diagonal cuyas entradas diagonales son los valores propios correspondientes, en el mismo orden, a las columnas de V.

$$A = V\,D\,V^{-1}$$

Por facilidad para la presentación supongamos que $A \in \mathbb{R}^{2\times 2}$

$$\begin{aligned}
V &= \begin{bmatrix} v^1 & v^2 \end{bmatrix} \\
D &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \\
x &= e^{tA}c \\
x &= e^{V\,tD\,V^{-1}}c \\
x &= Ve^{tD}V^{-1}c \\
x &= \begin{bmatrix} v^1 & v^2 \end{bmatrix} \begin{bmatrix} e^{t\lambda_1} & 0 \\ 0 & e^{t\lambda_2} \end{bmatrix} (V^{-1}c) \\
x &= \begin{bmatrix} e^{\lambda_1 t}v^1 & e^{\lambda_2 t}v^2 \end{bmatrix} \tilde{c} \\
x &= \begin{bmatrix} e^{\lambda_1 t}v^1 & e^{\lambda_2 t}v^2 \end{bmatrix} \begin{bmatrix} \tilde{c}_1 \\ \tilde{c}_2 \end{bmatrix} \\
x &= \tilde{c}_1 e^{\lambda_1 t}v^1 + \tilde{c}_2 e^{\lambda_2 t}v^2
\end{aligned} \tag{4.58}$$

Esta última igualdad es exactamente la misma (4.20) cuando $n = 2$. Cuando A no es diagonalizable es necesario utilizar la forma canónica de Jordan, tema fuera del alcance de este documento.

4.10 Linealización

Sea $f : \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$, es decir,

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{bmatrix}$$
$$f_i : \mathbb{R}^{n\times 1} \to \mathbb{R}$$

Si todas las funciones f_i son diferenciables, la matriz jacobiana de f en el punto $\bar{x}$ es:

$$J_f(\bar{x}) = \begin{bmatrix}
\dfrac{\partial f_1}{\partial x_1}(\bar{x}) & \dfrac{\partial f_1}{\partial x_2}(\bar{x}) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\bar{x}) \\
\dfrac{\partial f_2}{\partial x_1}(\bar{x}) & \dfrac{\partial f_2}{\partial x_2}(\bar{x}) & \cdots & \dfrac{\partial f_2}{\partial x_n}(\bar{x}) \\
\vdots & & & \\
\dfrac{\partial f_n}{\partial x_1}(\bar{x}) & \dfrac{\partial f_n}{\partial x_2}(\bar{x}) & \cdots & \dfrac{\partial f_n}{\partial x_n}(\bar{x})
\end{bmatrix}$$

La aproximación de primer orden de f en el punto $\bar{x}$, llamada también aproximación afín o aproximación lineal es

$$f(x) \approx L(x) = f(\bar{x}) + J_f(\bar{x})(x - \bar{x}) \tag{4.59}$$

En particular si se tiene un sistema de ecuaciones diferenciales autónomo,

$$\dot{x} = f(x) \tag{4.60}$$

se puede aproximar por el sistema de ecuaciones diferenciales con coeficientes constantes

$$\begin{aligned}\dot{x} &= f(\bar{x}) + J_f(\bar{x})(x - \bar{x})\\ \dot{x} &= J_f(\bar{x})x + c\\ c &= f(\bar{x}) - J_f(\bar{x})\,\bar{x}\end{aligned} \tag{4.61}$$

Si se considera el (un) punto de equilibrio

$$\dot{x} = J_f(x^*)x - \ J_f(x^*)\,x^* \tag{4.62}$$

Si el punto de equilibrio es el origen, la aproximación de (4.60)

$$\dot{x} = J_f(0)x \tag{4.63}$$

Ejemplo 4.33.

$$\begin{aligned}\dot{x}_1 &= -2x_1 + 3x_2\\ \dot{x}_2 &= -\frac{3}{x_1} + 2x_2\\ x_1 &> 0, \ \ x_2 > 0\end{aligned}$$

El punto de equilibrio es $x^* = (1.5, 1)$.

$$\begin{aligned}J_f(x) &= \begin{bmatrix} -2 & 3\\ \dfrac{3}{x_1^2} & 2\end{bmatrix}\\ J_f(x^*) &= \begin{bmatrix} -2 & 3\\ 4/3 & 2\end{bmatrix}\end{aligned}$$

Sistema aproximado (válido suficientemente cerca al punto de equilibrio)

$$\begin{aligned}\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} &= \begin{bmatrix} -2 & 3\\ 4/3 & 2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} - \begin{bmatrix} -2 & 3\\ 4/3 & 2\end{bmatrix}\begin{bmatrix}1.5\\ 1\end{bmatrix}\\ \begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} &= \begin{bmatrix} -2 & 3\\ 4/3 & 2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix} 0\\ -4\end{bmatrix}\end{aligned}$$

$$\begin{aligned}x^* &= (1.5, 1) \quad \text{obviamente es el mismo.}\\ p(\lambda) &= \lambda^2 - 8\\ \lambda_1 &= -\sqrt{8}\\ \lambda_2 &= \ \ \sqrt{8}\\ v^1 &= (-3.6213, 1)\\ v^2 &= (0.6213, 1)\end{aligned}$$

Para este sistema aproximado, el punto de equilibrio es un punto de silla. Salvo si el punto inicial está sobre la recta R_1, $x(t)$ se aleja del punto de equlibrio acercándose a la recta R_2.

4.11 Métodos numéricos

4.11.1 Método de Euler

$$\begin{aligned}\dot{x} &= f(t, x)\\ x(t_0) &= x^0\end{aligned}$$

$$\begin{aligned}
x(t_0+h) &\approx x(t_0) + hf(t_0, x^0)\\
x(t_0+h) &\approx x^1\\
x^1 &= x^0 + hf(t_0, x^0)\\
t_1 &= t_0 + h\\
x(t_1+h) &\approx x(t_1) + hf(t_1, x^1)\\
x^2 &= x^1 + hf(t_1, x^1)\\
t_i &= t_0 + i\,h\\
x(t_i) &\approx x^i\\
x^{i+1} &= x^i + hf(t_i, x^i)
\end{aligned}$$

$$\begin{aligned}
\dot{x} &= Ax + b\\
x(t_0) &= x^0
\end{aligned}$$

$$t_i = t_0 + i\,h \tag{4.64}$$

$$x(t_i) \approx x^i \tag{4.65}$$

$$x^{i+1} = x^i + h(Ax^i + b) \tag{4.66}$$

4.12 Ejemplos en economía

Ejemplo 4.34. Tomado de [Esc04]

$$\begin{aligned}
\dot{K}(t) &= I(t) - \delta K(t)\\
\dot{m}(t) &= (r+\delta)m(t) - \frac{1}{K(t)}
\end{aligned}$$

donde

$$\begin{aligned}
K(t) &= \text{stock de capital}\\
I(t) &= \text{inversión bruta}\\
\delta &= \text{tasa de depreciación del capital}\\
r &= \text{tasa de interés}\\
m(t) &= aI(t), \quad a > 0 \ \text{ función auxiliar,}\\
& \text{función de valoración marginal corriente.}
\end{aligned}$$

Remplazando

$$\begin{aligned}
\dot{K}(t) &= -\delta K(t) + \frac{1}{a}\,m(t)\\
\dot{m}(t) &= -\frac{1}{K(t)} + (r+\delta)m(t)
\end{aligned}$$

El anterior es un sistema de ecuaciones diferencialese no lineal, por el término $1/K(t)$.

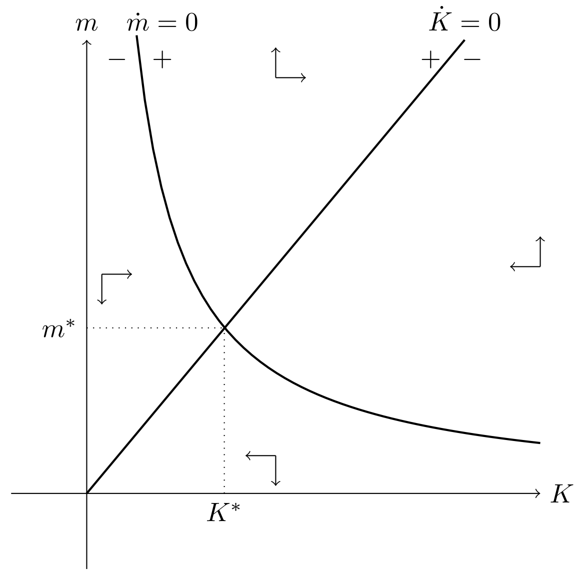
En el espacio (K, m) (K en el eje horizontal, m en el eje vertical), las líneas de fase son:

$$\dot{K}(t) = 0, \qquad m(t) = a\delta K(t) \qquad \text{una recta que pasa por el origen} \tag{4.67}$$

$$\dot{m}(t) = 0, \qquad m(t) = \frac{1}{(r+\delta)K(t)} \qquad \text{una hipérbola}$$

El punto de equilibrio (K^*, m^*) corresponde al corte de estas dos líneas de fase, es decir, a la solución del sistema de dos ecuaciones con dos incógnitas (4.67).

$$
\begin{aligned}
a\delta K(t) &= \frac{1}{(r+\delta)K(t)} \\
K^* &= \sqrt{\frac{1}{a\delta(r+\delta)}} \\
m^* &= a\delta\sqrt{\frac{1}{a\delta(r+\delta)}} = \sqrt{\frac{a\delta}{r+\delta}}
\end{aligned}
$$



Capítulo 5

Optimización dinámica

5.1 Control óptimo

5.1.1 Introducción

Problema general de control óptimo

$$\max_{x,u} \; J(x,u) = \int_{t_0}^{T} f(t, x(t), u(t))\, dt \tag{5.1a}$$
$$\dot{x}(t) = g(t, x(t), u(t)) \tag{5.1b}$$
$$x(t_0) = x_0 \tag{5.1c}$$
$$x(T) = x_T \quad \text{u otra condición terminal} \tag{5.1d}$$

Datos:

valores reales: $t_0, \;\; T, \;\; x_0, \;\; [\, x_T\,]$,

funciones: $f, \;\; g$.

Variables o incógnitas:

$x = x(t)$, variable o función de estado,

$u = u(t)$, variable o función de control.

J es una función que depende de las funciones x y u. Las funciones definidas en un conjunto de funciones (las funciones cuyos parámetros son funciones) reciben el nombre de **funcional**, es decir, J es un funcional.

Ejemplo 5.1.

$$\max \int_0^1 (x+u)dt$$
$$\dot{x} = 1 - u^2$$
$$x(0) = 0.9$$

En este ejemplo sencillo $t_0 = 0$, $T = 1$, $x_0 = 0.9$. Como no está definido el valor $x(T) = x(1)$ se dice que $x(T)$ es libre. $\diamond$

Ejemplo 5.2 (Escobar p. 208)**.** Maximizar la ganancia total **a tiempo presente**, en un periodo $[t_0, T]$. La ganancia en el instante t está dada por

$$\Pi(t) = \log(K(t)) - \frac{1}{2}a(I(t))^2, \quad a > 0$$

donde

$$\begin{aligned} K(t) &= \text{ stock de capital} \\ I(t) &= \text{ inversión} \end{aligned}$$

Se supone que el capital se deprecia a una tasa δ, es decir,

$$\dot{K}(t) = I(t) - \delta K(t)$$

En resumen,

$$\begin{aligned} &\max \int_0^T e^{-rt}\Pi(t)\,dt \\ &\dot{K}(t) = I(t) - \delta K(t) \\ &K(0) = 0 \\ &K(T) \geq 0 \end{aligned}$$

La expresión e^{-rt} en la integral se utiliza para hacer el cálculo a tiempo presente, siendo

$$r = \text{ tasa de interés de descuento}$$

5.1.2 Principio del máximo de Pontriaguin

Para el enunciado de este principio se requiere la definición del **hamiltoniano** del problema:

$$H(t, x(t), u(t), \lambda(t)) = f(t, x(t), u(t)) + \lambda(t)g(t, x(t), u(t)) \tag{5.2}$$

análogo al lagrangiano. La función $\lambda = \lambda(t)$ recibe, algunas veces, el nombre de variable de **coestado**.

Notación:

$$\begin{aligned} H_x &= \frac{\partial H}{\partial x} \\ H_u &= \frac{\partial H}{\partial u} \end{aligned}$$

Teorema 5.1. ***Principio del máximo de Pontriaguin (Pontryagin)***. *Condiciones necesarias de primer orden:*

Sean

$U \subseteq \mathbb{R}$, $u : U \to \mathbb{R}$, *continua por partes en* $[0, T]$,
f, g *funciones* C^1.

Si $x^*(t)$ *es una trayectoria óptima y* $u^*(t)$ *es un control óptimo, entonces existe* $\lambda(t)$, *función* C^1, *tal que*

- $$\frac{\partial H}{\partial x}(t, x^*(t), u^*(t), \lambda(t)) = -\dot{\lambda}(t)\,, \quad \textit{salvo en las discontinuidades} \tag{5.3}$$
- $u^*(t)$ *maximiza* $H(t, x^*(t), u(t), \lambda(t))$, *luego para muchos casos:*

$$\frac{\partial H}{\partial u}(t, x^*(t), u^*(t), \lambda(t)) = 0 \tag{5.4}$$

Las igualdades (5.3) y (5.4) se pueden escribir simplemente

$$H_x = -\dot{\lambda} \tag{5.5}$$

$$H_u = 0 \tag{5.6}$$

teniendo en cuenta que las derivadas parciales de H con respecto a x y a u se calculan en $(t, x^*(t), u^*(t), \lambda(t))$.

La restricción (5.1b) se puede escribir $\dot{x} = H_\lambda$ y así se tienen tres igualdades

$$\begin{aligned} H_x &= -\dot{\lambda} \\ H_u &= 0 \\ H_\lambda &= \dot{x} \end{aligned}$$

Condiciones de transversalidad o condiciones terminales. Con los supuestos del principio del máximo de Pontriaguin:

		CONDICIÓN
a)	T fijo, X_T fijo	
b)	T fijo, X_T libre	$\lambda(T) = 0$
c)	T libre, X_T fijo	$H(T, x^*, u^*, \lambda) = 0$
d)	T libre, X_T libre	$\lambda(T) = 0,\quad H(T, x^*, u^*, \lambda) = 0$
e)	T fijo, $x_T \geq \bar{x}$	$\lambda(T) \geq 0,\quad \lambda(T)(x_T - \bar{x}) = 0$
f)	$T \leq \bar{T}$, x_T fijo,	$H(T, x^*, u^*, \lambda) \geq 0,\quad H(T, x^*, u^*, \lambda)(\bar{T} - T) = 0$

En resumen, (en la mayoría de los casos)

$$\boxed{\begin{array}{rcl} \dot{x} &=& g(t, x, u) \\ x(t_0) &=& x_0 \\ H_x &=& -\dot{\lambda} \\ H_u &=& 0 \\ \lambda(T) &=& 0 \quad \text{u otra cond. de transv.} \end{array}}$$

5.1.3 Ejemplos de aplicación del principio del máximo

Ejemplo 5.3. Cerdá, ej. 4.6, p. 120

$$\max \int_0^1 (x + u)dt$$
$$\dot{x} = 1 - u^2$$
$$x(0) = 1$$

Solución:

$$\begin{aligned} H(t, x, u, \lambda) &= x + u + \lambda(1 - u^2) \\ H_x &= -\dot{\lambda} \quad \text{condición necesaria} \\ 1 &= -\dot{\lambda} \\ \dot{\lambda} &= -1 \\ \lambda(t) &= -t + C \end{aligned}$$

Por las condiciones de transversalidad, T es fijo ($T = 1$) y x_T es variable, luego $\lambda(T) = 0$.

$$\begin{aligned} 0 &= -1 + C \\ C &= 1 \\ \lambda^*(t) &= -t + 1 \end{aligned}$$

$$\begin{aligned} H_u &= 0 \quad \text{condición necesaria} \\ 1 - 2u\lambda &= 0 \\ u &= \frac{1}{2\lambda} \\ u^*(t) &= \frac{1}{2(1 - t)} \end{aligned}$$

$$
\begin{aligned}
\dot{x} &= 1 - u^2 \quad \text{restricción del problema} \\
\dot{x}(t) &= 1 - \frac{1}{4(1-t)^2}
\end{aligned}
$$

integrando

$$
\begin{aligned}
x(t) &= t - \frac{1}{4(1-t)} + D \\
x(0) &= 1 \quad \text{condición inicial} \\
-\frac{1}{4} + D &= 1 \\
D &= \frac{5}{4} \\
x^*(t) &= t - \frac{1}{4(1-t)} + \frac{5}{4}
\end{aligned}
$$

En resumen, las funciones $x(t) = t - 1/(4(1-t)) + 5/4$ y $u(t) = 1/(2(1-t))$ cumplen condiciones necesarias para ser la solución, es decir, son simplemente candidatas a ser la solución del problema de control óptimo. $\diamond$

Ejemplo 5.4.

$$
\begin{aligned}
\max \int_{0.5}^{1} &(ux - u^2 - x^2)dt \\
\dot{x} &= x + u \\
x(0.5) &= 2
\end{aligned}
$$

Solución

$$
\begin{aligned}
H(t, x, u, \lambda) &= ux - u^2 - x^2 + \lambda(x+u) \\
H_x &= -\dot{\lambda} \quad \text{condición necesaria} \\
u - 2x + \lambda &= -\dot{\lambda} \\
\lambda(1) &= 0 \quad \text{por cond. de transversalidad} \\
H_u &= 0 \quad \text{condición necesaria} \\
x - 2u + \lambda &= 0 \\
u &= \frac{x+\lambda}{2} \\
\frac{x+\lambda}{2} - 2x + \lambda &= -\dot{\lambda} \\
\dot{\lambda} &= \frac{3}{2}x - \frac{3}{2}\lambda, \quad \lambda(5) = 0 \\
\dot{x} &= x + \frac{x+\lambda}{2} \\
\dot{x} &= \frac{3}{2}x + \frac{1}{2}\lambda, \quad x(1) = 2 \\
\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} &= \begin{bmatrix} 3/2 & 1/2 \\ 3/2 & -3/2 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix}
\end{aligned}
$$

$p(\lambda) = \lambda^2 - 3$, valores propios: $-\sqrt{3}$, $\sqrt{3}$. Vectores propios

$$
v^1 = \begin{bmatrix} -0.1547005 \\ 1 \end{bmatrix}, \qquad v^2 = \begin{bmatrix} 2.1547005 \\ 1 \end{bmatrix}
$$

$$
\begin{bmatrix} x \\ \lambda \end{bmatrix} = C_1 \, e^{-\sqrt{3}\,t} \begin{bmatrix} -0.1547005 \\ 1 \end{bmatrix} + C_2 \, e^{\sqrt{3}\,t} \begin{bmatrix} 2.1547005 \\ 1 \end{bmatrix}
$$

Condiciones, $x(0.5) = 2$, $\lambda(1) = 0$

$$
\begin{aligned}
2 &= C_1 e^{-\sqrt{3}/2}(-0.1547005) + C_2\, e^{\sqrt{3}/2} \times 2.1547005 \\
0 &= C_1 e^{-\sqrt{3}} + C_2\, e^{\sqrt{3}} \\
\\
2 &= -0.0650701\, C_1 + 5.122677\, C_2 \\
0 &= 0.1769212\, C_1 + 5.6522337\, C_2 \\
C_1 &= -8.872499 \\
C_2 &= 0.277719 \\
x(t) &= 1.3725804\, e^{-\sqrt{3}\,t} + 0.5984015\, e^{\sqrt{3}\,t} \\
\lambda(t) &= -8.8724989\, e^{-\sqrt{3}\,t} + 0.2777191\, e^{\sqrt{3}\,t} \\
u(t) &= (x(t) + \lambda(t))/2 \\
u(t) &= -3.7499593\, e^{-\sqrt{3}\,t} + 0.4380603\, e^{\sqrt{3}\,t}
\end{aligned}
$$

Ejemplo 5.5. Modelo de Ramsey [Escobar p. 205 y p. 215]

- Dado K un capital, sea $q = F(K)$ lo producido.
- La función de producción F es de clase C^2, cóncava y creciente.
- Lo producido, q, puede ser consumido o invertido:

$$q = F(K) = C + \dot{K} \tag{5.7}$$

- Se desa maximizar la utilidad por consumo en un periodo determinado.
- Sea ρ la tasa de interés de descuento.
- Sea U la función de utilidad, creciente en C y cóncava (utilidad marginal decreciente), $U'(C) > 0$, $U''(C) < 0$.
- Intervalo de tiempo $[0, T]$

$$\max \int_0^T e^{-\rho t}\, U(C(t))\, dt$$

por (5.7)

$$
\begin{aligned}
&\max \int_0^T e^{-\rho t}\, U(F(K) - \dot{K})\, dt \qquad (5.8)\\
&K(0) = K_0 \\
&K(T) = K_T
\end{aligned}
$$

La utilización de $e^{-\rho t}$ indica que el cálculo se hace a tiempo presente.

Supongamos que la utilidad y la función de producción están dadas por

$$U(C) = \frac{C^{1-\mu}}{1-\mu} \tag{5.9}$$

$$f(K) = bK \tag{5.10}$$

La función (5.9) se conoce como función CRRA, Constant Relative Risk Aversion. Asi

$$\begin{aligned} \max \int_0^T e^{-\rho t}\frac{C^{1-\mu}}{1-\mu}\,dt \\ \dot{K} = bK - C \\ K(0) = K_0 \\ K(T) = K_T \end{aligned} \tag{5.11}$$

donde $K = K(t)$, $C = C(t)$. La variable de estado es K, la de control es C.

$$H(t, K, C, \lambda) = e^{-\rho t}\frac{C^{1-\mu}}{1-\mu} + \lambda(bK - C)$$

Por el principio del máximo

$$H_C = 0 = e^{-\rho t}C^{-\mu} - \lambda \tag{5.12}$$

$$H_K = -\dot{\lambda} = \lambda b \tag{5.13}$$

Sea $m(t)$ la **función de valor marginal corriente**,

$$m(t) = e^{\rho t}\lambda(t) \tag{5.14}$$

$$\lambda(t) = e^{-\rho t}m(t) \tag{5.15}$$

$$\dot{\lambda} = -\rho e^{-\rho t}m + e^{-\rho t}\dot{m}$$

De (5.13)

$$\begin{aligned} -\lambda b &= -\rho e^{-\rho t}m + e^{-\rho t}\dot{m} \\ -e^{-\rho t}mb &= e^{-\rho t}(-\rho m + \dot{m}) \\ -mb &= -\rho m + \dot{m} \end{aligned}$$

$$\dot{m} = (\rho - b)m \tag{5.16}$$

De (5.12) y (5.15)

$$\begin{aligned} e^{-\rho t}C^{-\mu} - me^{-\rho t} &= 0 \\ C^{-\mu} - m &= 0 \end{aligned}$$

$$m = C^{-\mu} \tag{5.17}$$

$$\dot{m} = -\mu C^{-\mu-1}\dot{C}$$

Usando (5.16)

$$(\rho - b)m = \mu C^{-\mu-1}\dot{C}$$

Usando (5.17)

$$(\rho - b)C^{-\mu} = -\mu C^{-\mu-1}\dot{C}$$

$$\dot{C} = \frac{b-\rho}{\mu}C \tag{5.18}$$

En resumen, usando (5.11)

$$\begin{aligned}
\dot{K} &= bK - C \\
\dot{C} &= \frac{b-\rho}{\mu}C \\
\begin{bmatrix}\dot{K}\\ \dot{C}\end{bmatrix} &= \begin{bmatrix} b & -1 \\ 0 & \frac{b-\rho}{\mu}\end{bmatrix}\begin{bmatrix}K\\ C\end{bmatrix} \\
\text{valores propios: } &\frac{b-\rho}{\mu}, \quad b \\
v^1 &= \begin{bmatrix} 1 \\ b - \frac{b-\rho}{\mu}\end{bmatrix}, \quad v^2 = \begin{bmatrix}1\\0\end{bmatrix}
\end{aligned}$$

La solución

$$\begin{bmatrix}K\\C\end{bmatrix} = c_1 e^{\frac{b-\rho}{\mu}t}\begin{bmatrix} 1 \\ b - \frac{b-\rho}{\mu}\end{bmatrix} + c_2 e^{bt}\begin{bmatrix}1\\0\end{bmatrix}$$

Con las condiciones sobre K, se calculan los valores de las constantes,

$$\begin{aligned}
K_0 &= c_1 + c_2 \\
K_T &= c_1 e^{\frac{b-\rho}{\mu}T} + c_2 e^{bT}
\end{aligned}$$

Si $b < \rho$ hay un punto de silla en el origen.
Si $b \geq \rho$, la trayectoria "tiende" al infinito, el origen es un repulsor.

Ejemplo 5.6 (Lomelí, p. 395)**.**

$$\begin{aligned}
\max \int_0^{10} &(2xu + u^2)\, dt \\
\dot{x} &= u \\
x(0) &= 10
\end{aligned}$$

Hamiltoniano

$$H = 2xu + u^2 + \lambda u$$

Principio del máximo

$$\begin{aligned}
H_x = 2u &= -\dot{\lambda} \\
H_u = 2x + 2u + \lambda &= 0
\end{aligned}$$

Condición de transversalidad

$$\lambda(10) = 0$$

Resumen de condiciones:

$$\dot{x} = u \tag{5.19}$$

$$2u = -\dot{\lambda} \tag{5.20}$$

$$2x + 2u + \lambda = 0 \tag{5.21}$$

$$x(0) = 10 \tag{5.22}$$

$$\lambda(10) = 0 \tag{5.23}$$

Derivando (5.21)

$$2\dot{x} + 2\dot{u} + \dot{\lambda} = 0$$

Utilizando (5.19) y (5.20)

$$\begin{aligned}
2u + 2\dot{u} - 2u &= 0 \\
2\dot{u} &= 0 \\
\dot{u} &= 0
\end{aligned}$$

$$u = C \tag{5.24}$$

Utilizando (5.19)

$$\begin{aligned} \dot{x} &= C \\ x &= Ct + D \end{aligned} \tag{5.25}$$

De (5.20)

$$\begin{aligned} \dot{\lambda} &= -2C \\ \lambda &= -2Ct + E \end{aligned}$$

Utilizando (5.23)

$$\begin{aligned} 0 &= -20C + E \\ E &= 20C \\ \lambda &= -2Ct + 20C \end{aligned} \tag{5.26}$$

De (5.25) y (5.22)

$$\begin{aligned} 10 &= 0 + D = D \\ x &= Ct + 10 \end{aligned}$$

Usando (5.21), (5.24), (5.26)

$$\begin{aligned} 2(Ct + 10) + 2C - 2Ct + 20C &= 0 \\ 2Ct + 20 + 2C - 2Ct + 20C &= 0 \\ 20 + 22C &= 0 \\ C &= -\frac{10}{11} \end{aligned}$$

Resumen de resultados

$$\begin{aligned} x &= -\frac{10}{11}\,t + 10 \\ u &= -\frac{10}{11} \\ \lambda &= \frac{20}{11}\,t - \frac{200}{11} \end{aligned}$$

Se puede verificar que cumplen (5.19), ..., (5.23).

Otra forma

$$\begin{aligned} \dot{x} &= u \\ \dot{\lambda} &= -2u \end{aligned}$$

Como

$$\begin{aligned} u &= -x - \lambda/2 \\ \dot{x} &= -x - \lambda/2 \\ \dot{\lambda} &= 2x + \lambda \end{aligned}$$

Sistema de ecuaciones diferenciales

$$
\begin{aligned}
A &= \begin{bmatrix} -1 & -1/2 \\ 2 & 1 \end{bmatrix} \\
p(t) &= t^2 \\
\text{valores propios: } & \ 0,\ 0 \\
v &= \begin{bmatrix} -1 \\ 2 \end{bmatrix} \\
(A - 0I)u &= v \\
u &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
\begin{bmatrix} x \\ \lambda \end{bmatrix} &= C_1 \begin{bmatrix} -1 \\ 2 \end{bmatrix} + C_2 \left(t \begin{bmatrix} -1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)
\end{aligned}
$$

Condiciones $x(0) = 10$, $\lambda(10) = 0$

$$
\begin{aligned}
10 &= -C_1 + C_2 \\
0 &= 2C_1 + 20C_2 \\
C_1 &= -100/11 \\
C_2 &= 10/11 \\
x(t) &= 100/11 - \frac{10}{11}t + 10/11 = 10 - \frac{10}{11}t \\
\lambda(t) &= -\frac{200}{11} + \frac{20}{11}t \\
u(t) &= -\frac{10}{11}
\end{aligned}
$$

Ejemplo 5.7. Chiang p. 173

$$
\begin{aligned}
\max & \int_0^2 (2x - 3u)dt \\
& \dot{x} = x + u \\
& x(0) = 4 \\
& x(2) \text{ libre} \\
& 0 \le u(t) \le 2
\end{aligned}
$$

$$
\begin{aligned}
H &= 2x - 3u + \lambda(x + u) \\
H &= (2 + \lambda)x + (\lambda - 3)u
\end{aligned}
$$

OJO, demasiado rápido

$$
\begin{aligned}
H_u &= (\lambda - 3) = 0 \\
\lambda &= 3 \quad \text{(no cumple condición} \quad \lambda(2) = 0)
\end{aligned}
$$

Retomando el principio del máximo de Pontriaguin

$$
\begin{aligned}
&u \quad \text{maximiza} \quad H \\
&\text{si} \quad \lambda - 3 > 0, \quad \text{entonces} \quad u^* = 2 \\
&\text{si} \quad \lambda - 3 < 0, \quad \text{entonces} \quad u^* = 0 \\
&H_x = -\dot{\lambda} \\
&2 + \lambda = -\dot{\lambda} \\
&\dot{\lambda} + \lambda = -2 \\
&\lambda_{gh} = Ce^{-t} \\
&\lambda_p = -2 \\
&\lambda(t) = Ce^{-t} - 2 \\
&\lambda(2) = 0 \\
&C = 2e^2 \\
&\lambda(t) = 2e^2 e^{-t} - 2
\end{aligned}
$$

¿Cuando $\lambda(t) = 3$?

$$
\begin{aligned}
&2e^2 e^{-t} - 2 = 3 \\
&t = 1.0837
\end{aligned}
$$

Sea $0 \leq t \leq 1.0837$, $u^* = 2$

$$
\begin{aligned}
&\dot{x} = x + u \\
&\dot{x} = x + 2 \\
&x = Ce^t - 2 \\
&x(0) = 4 \\
&C = 6 \\
&x(t) = 6e^t - 2
\end{aligned}
$$

Sea $1.0837 \leq t \leq 2$, $u^* = 0$

$$
\begin{aligned}
&\dot{x} = x + u \\
&\dot{x} = x \\
&x = De^t
\end{aligned}
$$

¿Qué condición para calcular D? Continuidad en $t = 1.0837$

$$
\begin{aligned}
&6\, e^{1.0837} - 2 = De^{1.0837} \\
&D = 5.3233
\end{aligned}
$$

En resumen:

$$
\begin{aligned}
&\text{Si} \quad t \in [0, 1.0837] \qquad && x(t) = 6e^t - 2 \\
& && u(t) = 2 \\
& && \lambda(t) = 2e^2 e^{-t} - 2 \\
&\text{Si} \quad t \in [1.0837,\ 2] \qquad && x(t) = 5.3233\, e^t \\
& && u(t) = 0 \\
& && \lambda(t) = 2e^2 e^{-t} - 2
\end{aligned}
$$

Problemas de este tipo reciben el nombre "*bang-bang*".

5.1.4 Condiciones suficientes

Teorema 5.2. ***Teorema de Mangasarian.*** *Si $x^*(t)$ y $u^*(t)$ cumplen las condiciones necesarias para el problema de control óptimo* (5.1) *y además f y g son diferenciables y cóncavas en las variables (x,u) y si además, $\lambda(t)\geq 0$ en $[t_0,T]$ cuando f no es lineal en x o en u, entonces $x^*(t)$ y $u^*(t)$ son solución del problema de control óptimo.*

En los ejemplos resueltos, cuando se aplicó el principio del máximo, las funciones obtenidas, $x(t)$ y $u(t)$, son simplemente funciones que cumplen condiciones necesarias, es decir, son funciones candidatas a solución del problema (5.1). Ahora bien, si esas funciones cumplen las condiciones suficientes, entonces hay certeza de que son la solución del problema.

Ejemplo 5.8. Retomemos el ejemplo 5.3.

$$\begin{aligned} \max \int_0^1 (x+u)dt \\ \dot{x} = 1-u^2 \\ x(0)=1 \end{aligned}$$

Claramente las funciones f y g son diferenciables. La función f es lineal en x y u, luego es cóncava (también convexa). El hessiano de g, con respecto a x y u es

$$\begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}$$

Esta matriz es semidefinida negativa, luego g es cóncava. Como f es lineal en x y u, no se requiere verificar que $\lambda(t)\geq 0$. Entonces las funciones obtenidas $x(t)$ y $u(t)$ son la solución del problema. $\diamond$

5.1.5 Generalización

El problema de control óptimo, en un caso más general, puede tener n variables de estado y k controles

$$\begin{aligned} \max \int_{t_0}^{T} f(t,x_1(t),...,x_n(t),u_1(t),...,u_k(t))\,dt \\ \dot{x}_1(t) &= g_1(t,x_1(t),...,x_n(t),u_1(t),...,u_k(t)) \\ &\vdots \\ \dot{x}_n(t) &= g_n(t,x_1(t),...,x_n(t),u_1(t),...,u_k(t)) \\ x_1(t_0) &= x_1^0 \\ &\vdots \\ x_n(t_0) &= x_n^0 \end{aligned} \tag{5.27}$$

Hamiltoniano

$$H(t,x_1,...,x_n,u_1,...,u_k,\lambda_1,...,\lambda_n) = f(t,x_1,...,x_n,u_1,...,u_k) + \sum_{i=1}^{n} \lambda_i g_i(t,x_1,...,x_n,u_1,...,u_k)$$

Teorema 5.3. ***Principio del máximo de Pontriaguin****: Si x_1^*, ..., x_n^* son trayectorias óptimas y u_1^*, ..., u_k^* son controles óptimos para* (5.27)*, entonces existen λ_1, ..., λ_n tales que*

$$\frac{\partial H}{\partial x_i} = -\dot{\lambda}_i, \quad i=1,...,n$$

$(u_1^*,...,u_k^*)$ *maximiza H, en muchos casos* $\dfrac{\partial H}{\partial u_j}=0, \quad j=1,...,k$

condiciones de transversalidad análogas a las del caso simple.

Ejemplo 5.9.

$$\begin{aligned}
\max \ J &= \int_0^1 (-u^2)dt \\
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= u \\
x_1(0) &= 0 \\
x_1(1) &= 2 \\
x_2(0) &= 0 \\
x_2(1) &= 5
\end{aligned}$$

Solución:

$$H(t, x_1, x_2, u, \lambda_1, \lambda_2) = -u^2 + \lambda_1 x_2 + \lambda_2 u$$

Condiciones necesarias:

$$\dot{x}_1 = x_2 \tag{5.28}$$
$$\dot{x}_2 = u \tag{5.29}$$
$$x_1(0) = 0 \tag{5.30}$$
$$x_1(1) = 2 \tag{5.31}$$
$$x_2(0) = 0 \tag{5.32}$$
$$x_2(1) = 5 \tag{5.33}$$
$$H_{x_1} = 0 = -\dot{\lambda}_1 \tag{5.34}$$
$$H_{x_2} = \lambda_1 = -\dot{\lambda}_2 \tag{5.35}$$
$$H_u = -2u + \lambda_2 = 0 \tag{5.36}$$

De (5.34) $\lambda_1 = A$

De (5.35) $\lambda_2 = -At + B$

De (5.36) $u = -At/2 + B/2$

De (5.29) $x_2 = -At^2/4 + Bt/2 + C$

De (5.28) $x_1 = -At^3/12 + Bt^2/4 + Ct + D$

De (5.30) a (5.33) $D = 0, \quad C = 0, \quad A = -12, \quad B = 4$

Luego

$$\begin{aligned}
\lambda_1(t) &= -12 \\
\lambda_2(t) &= 12t + 4 \\
x_1(t) &= t^3 + t^2 \\
x_2(t) &= 3t^2 + 2t \\
u(t) &= 6t + 2
\end{aligned}$$

5.1.6 Horizonte infinito

????

5.2 Cálculo de variaciones

En el problema de control óptimo, si la variable de control es la derivada, $u(t) = \dot{x}(t)$, se tiene el problema de cálculo de variaciones.

$$\begin{aligned} &\max \int_0^T f(y, x(t), \dot{x}(t))\, dt \\ &x(0) = x_0 \\ &x(T) = x_T \end{aligned} \tag{5.37}$$

Ejemplo 5.10.

$$\begin{aligned} &\max \int_0^1 (tx' + {x'}^2)\, dt \\ &x(0) = 1 \\ &x(1) = 1 \end{aligned}$$

Ejemplo 5.11. Problema isoperimétrico a). Dada una curva cerrada, de longitud fija L, definida por

$$\begin{aligned} x &= x(s) \\ y &= y(s) \end{aligned}$$

donde s es el parámetro arco, se busca maximizar el área encerrada por esa curva

$$\begin{aligned} &\max \ \frac{1}{2}\int_0^L (x(s)y'(s) - x'(s)y(s))\, ds \\ &x(0) = x(L) \\ &y(0) = y(L) \end{aligned}$$

Como s es el parámetro arco entonces se debe también cumplir

$$(x'(s))^2 + (y'(s))^2 = 1, \quad \forall\, s \in [0, L]$$

Ejemplo 5.12. Problema isoperimétrico b). Dados dos puntos $(0,0)$ y $(a,0)$ y un valor $L > a$, encontrar la curva que va de $(0,0)$ a $(a,0)$, de longitud L, tal que el área entra la curva y el eje x, sea máxima.

$$\begin{aligned} \max \int_0^a f(x)\, dx& \\ \int_0^a \sqrt{1 + (f'(x))^2}\, dx = L& \\ f(0) = 0& \\ f(a) = 0& \end{aligned}$$

Ejemplo 5.13. Braquistocrona. Dados dos puntos (a, y_a) y (b, y_b), con $y_a > y_b$, se busca encontrar la funcion continua y diferenciable $y = y(x)$ tal que al soltar un objeto en (a, y_a) llegue (siguiendo la curva de $y(x)$) lo más pronto posible a (b, y_b).

La deducción física da lugar a

$$\begin{aligned} &\min T(y(x)) = \frac{1}{\sqrt{2g}} \int_a^b \sqrt{\frac{1 + {y'}^2}{x}}\, dx \\ &y(a) = y_a \\ &y(b) = y_b \end{aligned}$$

Ejemplo 5.14 (Cerdá p. 28)**.** Una persona desea saber cual deber ser su consumo $C(t)$ en el periodo $[0, T]$ para maximizar su flujo de utilidad descontada con tasa de descuento r. Se supone conocidos $K_0 = K(0)$ y $K_T = K(T)$ donde $K(t)$ es el capital. Es posible prestar o tomar prestado capital a una tasa de interés i. Los ingresos provienen de un salario (exógeno) $v(t)$ o del rendimiento del capital prestado. La utilidad U depende de C, $U = U(C(t))$, es una función conocida de clase C^2.

La utilidad descontada está dada por

$$\int_0^T e^{-rt} U(C(t))\, dt$$

Ademas

$$\dot{K} = iK + v - C$$

luego

$$C = iK + v - \dot{K}$$

Entonces

$$\begin{aligned} \max \int_0^T & e^{-rt} U(\, iK + v - \dot{K}\,)\, dt \\ K(0) &= K_0 \\ K(T) &= K_T \end{aligned}$$

5.2.1 Condiciones necesarias, ecuación de Euler

Teorema. Si $x^*(t)$ es una función, solución local del problema de cálculo de variaciones (5.37), entonces

$$f_x = \frac{d}{dt} f_{\dot{x}} \tag{5.38}$$

De manera más precisa,

$$\frac{\partial f}{\partial x}(\, t, x^*(t), \dot{x}^*(t)\,) = \frac{d}{dt}\left[\frac{\partial f}{\partial \dot{x}}(\, t, x^*(t), \dot{x}^*(t)\,)\right] \tag{5.39}$$

Condiciones de transversalidad

		CONDICIÓN
a)	T fijo, X_T fijo	
b)	T fijo, X_T libre	$[f_{\dot{x}}]_{t=T} = 0$
c)	T fijo, $x_T \geq \bar{x}$	$[f_{\dot{x}}]_{t=T} \leq 0$, $[f_{\dot{x}}]_{t=T}(x_T - \bar{x}) = 0$
d)	T libre, X_T libre	$[f_{\dot{x}}]_{t=T} = 0$, $[F - \dot{x} f_{\dot{x}}]_{t=T} = 0$
e)	T libre, X_T fijo	$[F - \dot{x} f_{\dot{x}}]_{t=T} = 0$
f)	T libre, $x(T) = \varphi(T)$	$[F + (\varphi' - \dot{x}) f_{\dot{x}}]_{t=T} = 0$

Ejemplo 5.15 (Cerdá p. 34)**.**

$$\begin{aligned} &\max \int_0^2 (-12tx - \dot{x}^2) dt \\ &x(0) = 2 \\ &x(2) = 12 \end{aligned}$$

$$
\begin{aligned}
f_x &= -12t \\
f_{\dot{x}} &= -2\dot{x} \\
\frac{d}{dt}f_{\dot{x}} &= -2\ddot{x} \\
-12t &= -2\ddot{x} \\
\ddot{x} &= 6t \\
\dot{x} &= 3t^2 + C \\
x &= t^3 + Ct + D \\
x(0) &= 2 = D \\
x(2) &= 12 = 8 + 2C + 2 \\
C &= 1 \\
x(t) &= t^3 + t + 2
\end{aligned}
$$

Ejemplo 5.16 (Cerdá p. 35)**.**

$$
\begin{aligned}
&\max \int_o^1 (tx + x^2 - 2x^2\dot{x})dt \\
&x(0) = 2 \\
&x(1) = 5
\end{aligned}
$$

$$
\begin{aligned}
f_x &= t + 2x - 4x\dot{x} \\
f_{\dot{x}} &= -2x^2 \\
\frac{d}{dt}f_{\dot{x}} &= -4x\dot{x} \\
t + 2x - 4x\dot{x} &= -4x\dot{x} \\
t + 2x &= 0 \\
x &= t/2 \\
x(0) &= 2 \quad ???? \\
&\text{No hay solución}
\end{aligned}
$$

Ejemplo 5.17 (Lomelí p. 325)**.** Una empresa fabrica un producto. Sea x el nivel de producción y

$$c(x, \dot{x}) = x^2 + \dot{x}^2$$

la función de costo durante el periodo $[0, 1]$. El precio del producto es 4 y se mantiene fijo. La tasa de descuento es 0.1. Sea $x(0) = 0$ y se busca que $x(1) = 10$. La empresa desea conocer la producción que maximice la ganancia acumulada en valor presente.

$$
\begin{aligned}
&\max \int_0^1 e^{-0.1t}(4x - x^2 - \dot{x}^2)dt \\
&x(0) = \ \ 0 \\
&x(1) = 10
\end{aligned}
$$

$$
\begin{aligned}
e^{-0.1t}(4 - 2x) &= \frac{d}{dt}\left[-2e^{-0.1t}\dot{x}\right] \\
e^{-0.1t}(4 - 2x) &= 0.2e^{-0.1t}\dot{x} - 2e^{-0.1t}\ddot{x} \\
4 - 2x &= 0.2\dot{x} - 2\ddot{x} \\
2\ddot{x} - 0.2\dot{x} - 2x &= -4 \\
\ddot{x} - 0.1\dot{x} - x &= -2
\end{aligned}
$$

Para resolver la ec. dif. homogénea

$$\begin{aligned}
m^2 - 0.1m - 1 &= 0 \\
m &= \frac{0.1 \pm \sqrt{0.01+4}}{2} \\
m &= 1.0512, \quad -0.9512
\end{aligned}$$

Para la solución particular

$$\begin{aligned}
x_p &= A \\
\dot{x}_p &= 0 \\
\ddot{x}_p &= 0 \\
0 - 0.1 \cdot 0 - A &= -2 \\
A &= 2 \\
x &= c_1\, e^{-0.9512\, t} + c_2\, e^{1.0512\, t} + 2 \\
x(0) = 0 &= c_1 + c_2 + 2 \\
x(1) = 10 &= c_1\, e^{-0.9512} + c_2\, e^{1.0512} + 2 \\
c_1 &= -5.5445 \\
c_2 &= 3.5445 \\
x(t) &= -5.5445\, e^{-0.9512\, t} + 3.5445\, e^{1.0512\, t} + 2
\end{aligned}$$

5.2.2 Condición necesaria de Legendre

Si x^* es una solución de (5.37), entonces

$$f_{\dot{x}\dot{x}}(t, x^*, \dot{x}^*) \leq 0 \quad \forall t \in [t_0, T] \tag{5.40}$$

Una versión que no exige la existencia de $f_{\dot{x}\dot{x}}$ es la siguiente.

Si x^* es una solución de (5.37), entonces f es cóncava en $\dot{x}$.

5.2.3 Condiciones suficientes

Sea x^* que satisface la ecuación de Euler. Si f es cóncava en $(x, \dot{x})$ entonces x^* es solución del problema de maximización.

Ejemplo 5.18. ????

5.2.4 Varias variables

En esta primera generalizaci'on del problema de calcula variacional, f depende de t, de las funciones x_1, ..., x_n y de sus derivadas de primer orden $\dot{x}_1$, ..., $\dot{x}_n$. Hay condiciones iniciales para las funciones x_1, ..., x_n y, dependiendo el caso, condiciones finales o condiciones de transversalidad.

$$\begin{aligned}
&\max \int_{t_0}^{T} f(t, x_1, ..., x_n, \dot{x}_1, ..., \dot{x}_n)\, dt \\
&x_1(t_0) = x_1^0\,, \qquad x_1(T) = x_1^T \\
&\quad\vdots \\
&x_n(t_0) = x_n^0\,, \qquad x_n(T) = x_n^T
\end{aligned} \tag{5.41}$$

Para este problema, la ecuación de Euler, condición necesaria, se convierte en

$$\begin{aligned} f_{x_1} &= \frac{d}{dt} f_{\dot{x}_1} \\ &\vdots \\ f_{x_n} &= \frac{d}{dt} f_{\dot{x}_n} \end{aligned} \tag{5.42}$$

Ejemplo 5.19.

$$\begin{aligned} \max \int_0^1 &(-x_1 - 2x_2 - \dot{x}_1^2 - 4\dot{x}_2^2)\, dt \\ x_1(0) = 1, &\qquad x_1(1) = 9/4 \\ x_2(0) = 0, &\qquad x_2(1) = 9/8 \end{aligned}$$

Al aplicar la ecuación de Euler

$$\begin{aligned} -1 &= \frac{d}{dt}(-2\dot{x}_1) \\ -2 &= \frac{d}{dt}(-8\dot{x}_2) \end{aligned}$$

Entonces,

$$\begin{aligned} -1 &= -2\ddot{x}_1 \\ -2 &= -8\ddot{x}_2 \\ \dot{x}_1 &= \frac{1}{2}t + A \\ x_1 &= \frac{1}{4}t^2 + At + B \\ \dot{x}_2 &= \frac{1}{4}t + C \\ x_2 &= \frac{1}{8}t^2 + Ct + D \end{aligned}$$

Aplicando las condiciones en 0 y en 1,

$$\begin{aligned} x_1(t) &= \frac{1}{4}t^2 + t + 1 \\ x_2(t) &= \frac{1}{8}t^2 + t \qquad \diamond \end{aligned}$$

5.2.5 Derivadas de orden superior

Hay una sola función $x(t)$ pero la función f depende de t, x, $\dot{x}$, $\ddot{x}$, $\dddot{x}$, ..., $x^{(m)}$ y hay condiciones, en t_0 y T, para x, $\dot{x}$, $\ddot{x}$, ..., $x^{(m-1)}$.

$$\begin{aligned} \max \int_{t_0}^{T} &f(t, x, \dot{x}, \ddot{x}, ..., x^{(m)})\, dt \\ x(t_0) = x_0, &\quad x(T) = x_T \\ \dot{x}(t_0) = \dot{x}_0, &\quad \dot{x}(T) = \dot{x}_T \\ &\vdots \\ x^{(m\text{-}1)}(t_0) = x_0^{(m-1)}, &\quad x^{(m\text{-}1)}(T) = x_T^{(m-1)} \end{aligned} \tag{5.43}$$

Condición necesaria. La ecuación de Euler se generaliza a la ecuación de **Euler-Poisson**

$$f_x - \frac{d}{dt} f_{\dot{x}} + \frac{d^2}{dt^2} f_{\ddot{x}} - \frac{d^3}{dt^3} f_{\dddot{x}} + \frac{d^m}{dt^m} f_{x^{(m)}} = 0 \tag{5.44}$$

Ejemplo 5.20.

$$\max \int_0^1 (x^2 - \ddot{x}^2)dt$$
$$x(0) = 6, \quad x(1) = 8.4408$$
$$\dot{x}(0) = 3, \quad \dot{x}(1) = 1.6193$$

$$\begin{aligned}
f_x - \frac{d}{dt}f_{\dot{x}} + \frac{d^2}{dt^2}f_{\ddot{x}} &= 0\\
2x - \frac{d}{dt}0 + \frac{d^2}{dt^2}(-2\ddot{x}) &= 0\\
2x - 2x^{(4)} &= 0\\
x^{(4)} - x &= 0\\
m^4 - 1 &= 0\\
(m^2-1)(m^2+1) &= 0\\
m &= 1, \quad -1, \quad i, \quad -i\\
x(t) &= c_1e^t + c_2e^{-t} + c_3\cos t + c_4 \operatorname{sen} t
\end{aligned}$$

Utilizando las condiciones de x y $\dot{x}$ en 0 y 1,

$$\begin{aligned}
c_1 &= 1\\
c_2 &= 2\\
c_3 &= 3\\
c_4 &= 4\\
x(t) &= e^t + 2e^{-t} + 3\cos t + 4 \operatorname{sen} t \quad \diamond
\end{aligned}$$

5.3 Programación dinámica

También conocida con el nombre de **optimización dinámica en tiempo discreto**. Esta sección, pretende presentar la principal idea de la programación dinámica: *toda subpolítica de una política óptima también debe ser óptima*. El anterior principio se conoce como el principio de optimalidad de Richard Bellman.

El nombre de programación dinámica se debe a que inicialmente el método se aplicó a la optimización de algunos sistemas dinámicos, es decir, sistemas que evolucionan con el tiempo, cuando el tiempo es discreto o cuando se puede aproximar adecuadamente por un tiempo discreto. Sin embargo, el tiempo no es indispensable, se requiere simplemente que los sistemas se puedan expresar por etapas o por fases.

Dicho de otra forma, la idea básica de la programación dinámica consiste en convertir un problema de n variables en una sucesión de problemas más simples, por ejemplo, de una variable, y para más sencillez, una variable discreta.

Desde un punto de vista recurrente[1]: un problema complejo se resuelve mediante el planteamiento de problemas más sencillos pero análogos al problema general. Estos problemas más sencillos se resuelven mediante el planteamiento de problemas aún más sencillos pero que siguen guardando la misma estructura. Este proceso recurrente se aplica hasta encontrar problemas de solución inmediata. Una vez resueltos estos problemas super-sencillos se pasa a la solución de los problemas un poquito mas complejos y así sucesivamente hasta calcular la solución del problema general.

Aunque el principio es muy sencillo y aplicable a muchos problemas, se aplica de manera específica a cada problema. Es decir, no existe un algoritmo (o un programa de computador) único que se pueda aplicar a todos los problemas.

En lugar de presentar fórmulas, definiciones o conceptos generales pero abstractos, se presentan ejemplos típicos con sus soluciones. Todos los ejemplos presentados son deterministas, es decir, se supone que todos los datos del problema son conocidos de manera precisa.

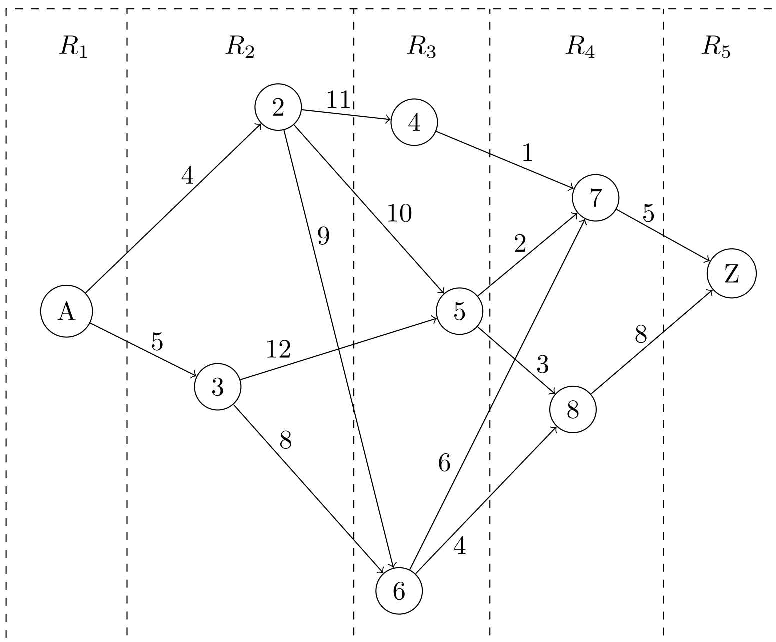
[1]Con cierta frecuencia, en lugar de recurrente, se utiliza el término "recursivo", anglicismo usado para la traduccion de "recursive".

5.3.1 Problema de la ruta más corta

Enunciado del problema

El Ministerio de Obras desea construir una autopista entre la ciudad de Girardot, que denotaremos simplemente por A y la ciudad de Barranquilla denotada por$Z$. Globalmente, la autopista sigue la dirección del río Magdalena. El valle del río y su zona de influencia directa se dividió en una sucesión de n regiones adyacentes que, por facilidad, llamaremos R_1, R_2, ... R_n. La ciudad A está en R_1 y Z está en R_n. La autopista pasa por todas las n regiones, pero está previsto que pase solamente por una ciudad de cada región. Sin embargo en algunas regiones hay varias ciudades importantes y cada una de ellas podría ser la ciudad de la región por donde pasa la autopista. Con este esquema, el Ministerio está estudiando muchos tramos de autopista, cada tramo va de una ciudad en una región, a otra ciudad en la región siguiente. Sin embargo, de cada ciudad en una región no hay necesariamente tramos a todas las ciudades de la siguiente región. Pero por otro lado, para cada ciudad de la regiones intermedias, de la 2 a la $n-1$, hay por lo menos un tramo proveniente de una ciudad de la región anterior y por lo menos un tramo que va hasta una ciudad de la siguiente región.

Para cada uno de estos tramos posibles el Ministerio ha calculado un costo total que tiene en cuenta, entre otros aspectos, la distancia, las dificultades específicas de la construcción, el sobrecosto del transporte desde otras ciudades de cada región hasta la ciudad por donde pasa la autopista.



El objetivo del Ministerio es encontrar una sucesión de tramos concatenados, que van desde A hasta Z, con costo mínimo.

Las condiciones del problema se pueden formalizar de la siguiente manera:

- $N = \{A, ..., Z\}$ conjunto de ciudades o nodos. N es simplemente un conjunto finito cualquiera (no necesariamente un subconjunto del abecedario), en el cual están A y Z.
- R_1, R_2, ..., R_n forman una partición de N, es decir,

$$R_1 \cup R_2 \cup ... \cup R_n = N,$$
$$R_i \neq \emptyset \ , \quad i = 1, ..., n,$$
$$R_i \cap R_j = \emptyset \ \text{ si } \quad i \neq j.$$

- $R_1 = \{A\}$.
- $R_n = \{Z\}$.

- $\Gamma \subseteq N \times N$ conjunto de tramos posibles ("flechas" del grafo[2]).
- Si $(i,j) \in \Gamma$ entonces existe $1 \leq k \leq n-1$ tal que $i \in R_k$ y $j \in R_{k+1}$.
- Para $1 \leq k \leq n-1$, si $i \in R_k$ entonces existe $j \in R_{k+1}$ tal que $(i,j) \in \Gamma$.
- Para $2 \leq k \leq n$, si $i \in R_k$ entonces existe $j \in R_{k-1}$ tal que $(j,i) \in \Gamma$.
- Si $(i,j) \in \Gamma$ entonces se conoce $c(i,j) = c_{ij} > 0$.

En teoría de grafos se habla de predecesor, sucesor, conjunto de predecesores, conjunto de sucesores. La utilización de estos conceptos, con sus respectivas notaciones, facilita y hace más compacto el planteamiento del problema. Si la flecha (i,j) está en el grafo se dice que i es un **predecesor** de j, y que j es un **sucesor** de i. Se denota por $\Gamma^-(i)$ el conjunto de predecesores de i y por $\Gamma^+(i)$ el conjunto de sus sucesores. Entonces:

- $\Gamma^-(i) \neq \emptyset,\ \forall i \neq A$.
- $\Gamma^+(i) \neq \emptyset,\ \forall i \neq Z$.
- $i \in R_k \ \Rightarrow\ \Gamma^-(i) \subseteq R_{k-1},\ k = 2, ..., n$.
- $i \in R_k \ \Rightarrow\ \Gamma^+(i) \subseteq R_{k+1},\ k = 1, ..., n-1$.

Planteamiento del problema de optimización

El problema se puede presentar de la siguiente manera: encontrar i_1, i_2, ..., i_n para minimizar una función con ciertas restricciones:

$$\begin{aligned} \min c(i_1,i_2) + c(i_2,i_3) + ... + c(i_{n-1},i_n) &= \sum_{k=1}^{n-1} c(i_k,i_{k+1}) \\ (i_k,i_{k+1}) &\in \Gamma,\ k = 1, ..., n-1 \\ i_k &\in R_k,\ k = 1, ..., n. \end{aligned}$$

En realidad el problema depende únicamente de $n-2$ variables: i_2, i_3, ..., i_{n-1} ya que $i_1 = A$, $i_n = Z$. Además la última condición se puede quitar puesto que ya está implícita en las propiedades de Γ.

A cada una de las ciudades de R_2 llega por lo menos un tramo desde una ciudad de R_1, pero como en R_1 solo hay una ciudad entonces: a cada una de las ciudades de R_2 llega por lo menos una ruta desde A. A su vez, a cada una de las ciudades de R_3 llega por lo menos un tramo desde una ciudad de R_2 y como a cada ciudad de R_2 llega una ruta desde A entonces: a cada una de las ciudades de R_3 llega por lo menos una ruta desde A. Repitiendo este proceso se puede deducir que a cada una de las ciudades llega por lo menos una ruta desde A. En particular existe por lo menos una ruta desde A hasta Z.

Una manera de resolver este problema es utilizando la **fuerza bruta**: hacer una lista de todas las rutas posibles desde A hasta Z, para cada ruta evaluar el costo (sumando los costos de cada tramo), y buscar la ruta (o una de las rutas) de menor costo.

De ahora en adelante, cuando se hable del mejor (camino, procedimiento, política, ...) se entenderá que es el mejor, en el sentido estricto cuando hay uno solo, o uno de los mejores cuando hay varios.

Solución por programación dinámica

La forma recurrente de resolver este problema es muy sencilla. Para conocer la mejor ruta de A hasta Z basta con conocer la mejor ruta a cada una de las ciudades de R_{n-1}. Al costo de cada ruta óptima hasta una de ciudad de R_{n-1} se le agrega el costo del tramo entre esa ciudad de R_{n-1} y Z en R_n y finalmente se escoge la menor suma. Definir algunas funciones permite escribir el razonamiento anterior de manera más corta y precisa.

Sean :

[2] Un grafo G es una pareja $G = (N, \Gamma)$, donde N es el conjunto finito de vértices y $\Gamma \subseteq N \times N$ es el conjunto de flechas

$$\begin{aligned} C_n^*(Z): &\ \text{costo mínimo de todas las rutas desde } A \text{ hasta } Z. \\ C_{n-1}^*(i): &\ \text{costo mínimo de todas las rutas desde } A \text{ hasta la} \\ &\ \text{ciudad } i \in R_{n-1}. \end{aligned}$$

Entonces la solución recurrente dice:

$$C_n^*(Z) = \min_{i \in R_{n-1}} \{C_{n-1}(i) + c(i, Z)\}.$$

Obviamente esto presupone que se conocen todos los valores $C_{n-1}^*(i)$, y entonces la pregunta inmediata es: *¿Cómo se calculan los valores $C_{n-1}^*(i)$?*

La respuesta es de nuevo recurrente: *Utilizando los costos de las rutas mínimas desde A hasta las ciudades de R_{n-2}.*

Sea :

$$\begin{aligned} C_{n-2}^*(i): &\ \text{costo mínimo de todas las rutas desde } A \text{ hasta la} \\ &\ \text{ciudad } i \in R_{n-2}. \end{aligned}$$

Entonces la solución recurrente dice:

$$C_{n-1}^*(j) = \min\{C_{n-2}(i) + c(i, j) : i \in \Gamma^-(j)\}, \quad j \in R_{n-1}.$$

Este proceso se repite hasta poder utilizar valores "inmediatos" o de muy fácil obtención. Para este problema de la autopista, puede ser

$$C_2^*(j) = c(A, j), \quad j \in R_2.$$

donde $C_2^*(j)$indica costo mínimo de todas las rutas (hay una sola) desde A hasta la ciudad $j \in R_2$.

La deducción de la solución recurrente se hizo hacia atrás (desde n hasta 2) pero el cálculo se hace hacia adelante (de 2 hasta n). En resumen,

- definir una función que permita la recurrencia,
- definir el objetivo final,
- definir las condiciones iniciales (fáciles de evaluar),
- definir una relación o fórmula recurrente,
- calcular los valores iniciales,
- hacer cálculos recurrentes hasta encontrar la solución

Para este problema :

$$\begin{aligned} C_k^*(i) &= \text{costo mínimo de todas las rutas desde } A \text{ hasta la} \\ &\quad \text{ciudad } i \in R_k, \ \ k = 2, ..., n. \\ C_n^*(Z) &= \ ? \\ C_2^*(j) &= c(A, j), \quad j \in R_2 \\ C_{k+1}^*(j) &= \min\{C_k^*(i) + c(i, j) : i \in \Gamma^-(j)\}, \ \ 2 \le k \le n-1, \ \ j \in R_{k+1} \end{aligned}$$

Un planteamiento ligeramente diferente podría ser:

$$
\begin{aligned}
C_k^*(i) &= \text{costo mínimo de todas las rutas desde } A \text{ hasta la} \\
&\quad\ \text{ciudad } i \in R_k,\ \ k = 2, ..., n. \\
C_n^*(Z) &= \ ? \\
C_1^*(A) &= 0 \\
C_{k+1}^*(j) &= \min\{C_k^*(i) + c(i,j) : i \in \Gamma^-(j)\},\ \ 1 \le k \le n-1,\ \ j \in R_{k+1}
\end{aligned}
$$

La diferencia está simplemente en el "sitio inicial".

Resultados numéricos

Los valores iniciales se obtienen inmediatamente:

$$
\begin{aligned}
C_2^*(2) &= c_{A2} = 4, \\
C_2^*(3) &= c_{A3} = 5.
\end{aligned}
$$

El proceso recurrente empieza realmente a partir de las ciudades de R_3:

$$
\begin{aligned}
C_3^*(4) &= \min_{i \in \Gamma^-(4)} \{C_2^*(i) + c_{i4}\}, \\
&= \min\{C_2^*(2) + c_{24}\}, \\
&= \min\{4 + 11\}, \\
C_3^*(4) &= 15.
\end{aligned}
$$

$$
\begin{aligned}
C_3^*(5) &= \min_{i \in \Gamma^-(5)} \{C_2^*(i) + c_{i5}\}, \\
&= \min\{C_2^*(2) + c_{25}, C_2^*(3) + c_{35}\}, \\
&= \min\{4 + 10, 5 + 12\}, \\
C_3^*(5) &= 14.
\end{aligned}
$$

$$
\begin{aligned}
C_3^*(6) &= \min_{i \in \Gamma^-(6)} \{C_2^*(i) + c_{i6}\}, \\
&= \min\{C_2^*(2) + c_{26}, C_2^*(3) + c_{36}\}, \\
&= \min\{4 + 9, 5 + 8\}, \\
C_3^*(6) &= 13.
\end{aligned}
$$

En resumen para la región R_3:

$$
\begin{aligned}
C_3^*(4) &= 15, \\
C_3^*(5) &= 14, \\
C_3^*(6) &= 13.
\end{aligned}
$$

Para la región R_4:

$$
\begin{aligned}
C_4^*(7) &= \min_{i \in \Gamma^-(7)} \{C_3^*(i) + c_{i7}\}, \\
&= \min\{C_3^*(4) + c_{47}, C_3^*(5) + c_{57}, C_3^*(6) + c_{67}\}, \\
&= \min\{15 + 1, 14 + 2, 13 + 6\}, \\
C_4^*(7) &= 16.
\end{aligned}
$$

$$
\begin{aligned}
C_4^*(8) &= \min_{i\in\Gamma^-(8)} \{C_3^*(i) + c_{i8},\} \\
&= \min\{C_3^*(5) + c_{58}, C_3^*(6) + c_{68}\}, \\
&= \min\{14 + 3, 13 + 4\}, \\
C_4^*(8) &= 17.
\end{aligned}
$$

En resumen para la región R_4:

$$
\begin{aligned}
C_4^*(7) &= 16, \\
C_4^*(8) &= 17.
\end{aligned}
$$

Finalmente para Z en R_5:

$$
\begin{aligned}
C_5^*(Z) &= \min_{i\in\Gamma^-(Z)} \{C_4^*(i) + c_{iZ}\}, \\
&= \min\{C_4^*(7) + c_{7Z}, C_4^*(8) + c_{8Z}\}, \\
&= \min\{16 + 5, 17 + 8\}, \\
C_5^*(Z) &= 21.
\end{aligned}
$$

Ya se obtuvo el costo mínimo, pero es necesario conocer también las ciudades por donde debe pasar una autopista de costo mínimo. Con la información que se tiene no se puede reconstruir la mejor ruta. Entonces es necesario volver a resolver el problema, pero esta vez es necesario, para cada ciudad intermedia j, no solo conocer $C_k^*(j)$, sino también saber desde que ciudad i^* de la región R_{k-1} se obtiene este costo mínimo.

Obviamente

$$
\begin{aligned}
C_2^*(2) &= 4, \;\; i^* = A, \\
C_2^*(3) &= 5, \;\; i^* = A.
\end{aligned}
$$

$C_3^*(4)$ se obtiene viniendo de la única ciudad predecesora: 2.

$$
C_3^*(4) = 15, \;\; i^* = 2.
$$

$$
\begin{aligned}
C_3^*(5) &= \min\{C_2^*(2) + c_{25}, C_2^*(3) + c_{35}\}, \\
&= \min\{\overline{4 + 10}, 5 + 12\}, \\
C_3^*(5) &= 14, \;\; i^* = 2.
\end{aligned}
$$

Para $C_3^*(7)$ hay empate ya que se obtiene el costo mínimo viniendo de 2 o viniendo de 3, sin embargo basta con tener información sobre una de las mejores ciudades precedentes, por ejemplo, la primera encontrada.

$$
\begin{aligned}
C_3^*(6) &= \min\{C_2^*(2) + c_{26}, C_2^*(3) + c_{36}\}, \\
&= \min\{\overline{4 + 9}, \overline{5 + 8}\}, \\
C_3^*(6) &= 13, \;\; i^* = 2.
\end{aligned}
$$

Toda la información necesaria para poder calcular el costo mínimo desde A hasta Z y poder reconstruir una ruta mínima es:

j	$C_2^*(j)$	i^*	j	$C_3^*(j)$	i^*	j	$C_4^*(j)$	i^*	j	$C_5^*(j)$	i^*
2	4	A	4	15	2	7	16	4	Z	21	7
3	5	A	5	14	2	8	17	5			
			6	13	2						

Luego según la tabla, el costo mínimo es 21. Además a Z se llega proveniente de 7, a 7 se llega proveniente de 4, a 4 se llega proveniente de 2, y finalmente a 2 se llega proveniente de A. Entonces la autopista de costo mínimo (o una de las autopistas de costo mínimo) es: $(A, 2, 4, 7, Z)$

Solución hacia atrás

La solución presentada en los dos numerales anteriores se conoce como la solución **hacia adelante**. También se tiene una solución análoga hacia atrás. Se busca el costo mínimo desde cada ciudad hasta Z empezando con el costo desde las ciudades en la región R_{n-1}.

$$\begin{aligned} C_k^*(i) = {} & \text{costo mínimo de todas las rutas desde la ciudad } i \in R_k \\ & \text{hasta } Z,\ k = n-1, n-2, ..., 1. \end{aligned}$$

Objetivo final:

$$C_1^*(A) = \ ?$$

Condiciones iniciales:

$$C_{n-1}^*(j) = c(j, Z), \ \ j \in R_{n-1}.$$

Relación recurrente:

$$C_{k-1}^*(i) = \min_{j \in \Gamma^+(i)} \{c(i,j) + C_k^*(j)\}, \ \ n-1 \geq k \geq 2, \ \ i \in R_{k-1}.$$

Obviamente

$$\begin{aligned} C_4^*(7) &= 5, \ \ j^* = Z, \\ C_4^*(8) &= 8, \ \ j^* = Z. \end{aligned}$$

$$\begin{aligned} C_3^*(4) &= \min_{j \in \Gamma^+(4)} \{c(4,j) + C_4^*(j)\}, \\ &= \min\{c(4,7) + C_4^*(7)\}, \\ &= \min\{\overline{1+5}\}, \\ C_3^*(4) &= 6, \ \ j^* = 7. \end{aligned}$$

$$\begin{aligned} C_3^*(5) &= \min_{j \in \Gamma^+(5)} \{c(5,j) + C_4^*(j)\}, \\ &= \min\{c(5,7) + C_4^*(7), c(5,8) + C_4^*(8)\}, \\ &= \min\{\overline{2+5}, 3+8\}, \\ C_3^*(5) &= 7, \ \ j^* = 7. \end{aligned}$$

y así sucesivamente.

i	$C_4^*(i)$	j^*	i	$C_3^*(i)$	j^*	i	$C_2^*(i)$	j^*	i	$C_1^*(i)$	j^*
7	5	Z	4	6	7	2	17	4	A	21	2
8	8	Z	5	7	7	3	19	5			
			6	11	7						

Luego según la tabla, el costo mínimo es 21. Además se llega desde A pasando por 2, se llega desde 2 pasando por 4, se llega desde 4 pasando por 7, y finalmente se llega desde 7 pasando por Z. Entonces la autopista de costo mínimo (o una de las autopistas de costo mínimo) es: $(A, 2, 4, 7, Z)$

5.3.2 Problema de asignación de médicos

Enunciado del problema

(Adaptación de un ejemplo de Hillier y Lieberman). La OMS (Organización Mundial de la Salud) tiene un equipo de m médicos especialistas en salud pública y los desea repartir en n paises P_1, P_2, ..., P_n para que desarrollen campañas educativas tendientes a disminuir la mortalidad infantil. De acuerdo con las condiciones específicas de cada país, de la tasa de natalidad, de la mortalidad antes de los dos años, del número de habitantes, la OMS posee evaluaciones bastante precisas de los valores $b(i,j) \equiv b_{ij}$, $i = 0, ..., m$, $j = 1, ..., n$ que indican el beneficio de asignar i médicos al país P_j. Este beneficio b_{ij} indica (en cientos de mil) la disminución en el número de niños muertos antes de los dos años de vida, durante los próximos 5 años. Así por ejemplo, $b_{23} = 6$ indica que si se asignan 2 médicos al país P_3 se espera que en los próximos 5 años haya una disminución de 600000 en el número de niños muertos antes de los dos años de vida.

Los beneficios no son directamente proporcionales al número de médicos, es decir, si $b_{23} = 6$ no se cumple necesariamente que $b_{43} = 12$.

Se supone además que no es obligatorio asignar médicos en cada uno de los paises y también se supone que es posible asignar todos los médicos a un solo país.

También se supone que al aumentar el número de médicos asignados a un país el beneficio no disminuye. Pensando en un problema más general se podría pensar que cuando hay demasiadas personas asignadas a una labor, se obstruye el adecuado funcionamiento y el resultado global podría disminuir. Sin embargo se puede suponer que b_{ij} indica el mayor beneficio obtenido en el país P_j al asignar a los más i médicos.

La OMS desea saber cuantos médicos debe asignar a cada país para maximizar el beneficio total, o sea, para maximizar la disminución total en la mortalidad infantil en los n paises.

Planteamiento del problema de optimización

Se necesita conocer el número de médicos que se asigna a cada país para maximizar el beneficio total, sin sobrepasar el número de médicos disponibles. Si x_j indica el número de médicos que se asignan al pais P_j, entonces el problema de optimización es:

$$\max \sum_{j=1}^{n} b(x_j, j)$$
$$\sum_{j=1}^{n} x_j \le m,$$
$$0 \le x_j \le m, \;\; j = 1, ..., n,$$
$$x_j \in \mathbb{Z}, \;\; j = 1, ..., n.$$

Este problema se puede resolver por la fuerza bruta construyendo todas las combinaciones, verificando si cada combinación es factible ($\sum_{j=1}^{n} x_j \le m$) y escogiendo la mejor entre las factibles. De esta forma cada variable puede tomar $m + 1$ valores: 0, 1, 2, ..., m, o sea, es necesario estudiar $(m + 1)^n$ combinaciones. Es claro que para valores grandes de m y n el número de combinaciones puede ser inmanejable.

Solución recurrente

Para asignar optimamente m médicos a los n paises hay que asignar una parte de los médicos a los paises P_1, P_2, ..., P_{n-1} y el resto al país P_n. Supongamos que sabemos asignar óptimamente 0, 1, 2, 3, ..., m médicos a los paises P_1, P_2, ..., P_{n-1}. Entonces para asignar óptimamente m médicos a los n paises hay que considerar $m+1$ posibilidades:

- 0 médicos a los paises P_1,..., P_{n-1} y m médicos a P_n
- 1 médico a los paises P_1,..., P_{n-1} y $m-1$ médicos a P_n
- 2 médicos a los paises P_1,..., P_{n-1} y $m-2$ médicos a P_n

  ...
- m médicos a los paises P_1,..., P_{n-1} y 0 médicos a P_n

Al escoger la mejor combinación se tiene la solución del problema. Obviamente para conocer las soluciones óptimas en los primeros $n-1$ paises se requiere conocer las soluciones óptimas en los primeros $n-2$ paises y así sucesivamente. O sea, primero se resuelve el problema de asignar óptimamente médicos al primer país, con estos resultados se puede obtener la asignación óptima de médicos a los 2 primeros paises, y así sucesivamente hasta obtener la solución global.

$$\begin{aligned} B_k^*(i) = {} & \text{beneficio máximo obtenido al asignar } i \text{ médicos} \\ & \text{a los paises } P_1,\ P_2,\ ...,\ P_k,\quad k = 1, ..., n,\ \ i = 0, ..., m. \end{aligned}$$

Objetivo final:

$$B_n^*(m) = \ ?$$

Condiciones iniciales:

$$B_1^*(i) = b_{i1},\ \ i = 0, ..., m.$$

Relación recurrente:

$$\begin{aligned} B_{k+1}^*(i) = \max_{0 \le j \le i} \{B_k^*(i-j) + b_{j,k+1}\},\ \ & k = 1, ..., n\text{-}1, \\ & i = 0, ..., m. \end{aligned} \tag{5.45}$$

Esta relación recurrente dice que la mejor manera de asignar i médicos a los paises P_1, P_2, ..., P_{k+1} es estudiando todas las posibilidades consistentes en asignar j médicos al país P_{k+1} y el resto, $i-j$, a los paises P_1, P_2, ..., P_k.

Resultados numéricos

Consideremos los siguientes datos: $m = 5$, $n = 4$ y los siguientes beneficios:

i	b_{i1}	b_{i2}	b_{i3}	b_{i4}
0	0	0	0	0
1	2	2	1	4
2	4	3	4	5
3	6	4	7	6
4	8	8	9	7
5	10	12	11	8

Condiciones iniciales:

$$
\begin{aligned}
B_1^*(0) &= 0, \;\; j^* = 0, \\
B_1^*(1) &= 2, \;\; j^* = 1, \\
B_1^*(2) &= 4, \;\; j^* = 2, \\
B_1^*(3) &= 6, \;\; j^* = 3, \\
B_1^*(4) &= 8, \;\; j^* = 4, \\
B_1^*(5) &= 10, \;\; j^* = 5.
\end{aligned}
$$

$$
\begin{aligned}
B_2^*(0) &= \max_{0\le j\le 0}\{B_1^*(0-j) + b_{j2}\}, \\
&= \max\{B_1^*(0) + b_{02}\}, \\
&= \max\{0+0\}, \\
B_2^*(0) &= 0, \;\; j^* = 0.
\end{aligned}
$$

$$
\begin{aligned}
B_2^*(1) &= \max_{0\le j\le 1}\{B_1^*(1-j) + b_{j2}\}, \\
&= \max\{B_1^*(1) + b_{02}, B_1^*(0) + b_{12}\}, \\
&= \max\{\overline{2+0}, 0+2\}, \\
B_2^*(1) &= 2, \;\; j^* = 0.
\end{aligned}
$$

$$
\begin{aligned}
B_2^*(2) &= \max_{0\le j\le 2}\{B_1^*(2-j) + b_{j2}\}, \\
&= \max\{B_1^*(2) + b_{02}, B_1^*(1) + b_{12}, B_1^*(0) + b_{22}\}, \\
&= \max\{\overline{4+0}, 2+2, 0+3\}, \\
B_2^*(2) &= 4, \;\; j^* = 0.
\end{aligned}
$$

Para los dos primeros paises se tiene:

$$
\begin{aligned}
B_2^*(0) &= 0, \;\; j^* = 0, \\
B_2^*(1) &= 2, \;\; j^* = 0, \\
B_2^*(2) &= 4, \;\; j^* = 0, \\
B_2^*(3) &= 6, \;\; j^* = 0, \\
B_2^*(4) &= 8, \;\; j^* = 0, \\
B_2^*(5) &= 12, \;\; j^* = 5.
\end{aligned}
$$

Un ejemplo de un cálculo para los tres primeros paises es:

$$
\begin{aligned}
B_3^*(4) &= \max_{0\le j\le 4}\{B_2^*(4-j) + b_{j3}\}, \\
&= \max\{B_2^*(4) + b_{03}, B_2^*(3) + b_{13}, B_2^*(2) + b_{23}, \\
&\qquad B_2^*(1) + b_{33}, B_2^*(0) + b_{43}\}, \\
&= \max\{8+0, 6+1, 4+4, \overline{2+7}, 0+9\}, \\
B_3^*(4) &= 9, \;\; j^* = 3.
\end{aligned}
$$

En la siguiente tabla están los resultados importantes:

i	$B_1^*(i)$	j^*	$B_2^*(i)$	j^*	$B_3^*(i)$	j^*	$B_4^*(i)$	j^*
0	0	0	0	0	0	0		
1	2	1	2	0	2	0		
2	4	2	4	0	4	0		
3	6	3	6	0	7	3		
4	8	4	8	0	9	3		
5	10	5	12	5	12	0	13	1

Se observa que no es necesario calcular $B_4^*(0)$, $B_4^*(1)$, ..., $B_4^*(4)$, ya que no se utilizan para calcular $B_4^*(5)$. El beneficio máximo es 13. Este beneficio máximo se obtiene asignando 1 médico al cuarto país ($j^* = 1$). Entonces quedan 4 médicos para los primeros 3 paises. El j^* correspondiente a $B_3^*(4)$ es 3, esto indica que hay que asignar 3 médicos al tercer país. Entonces queda 1 médico para los primeros 2 paises. El j^* correspondiente a $B_2^*(1)$ es 0, esto indica que hay que asignar 0 médicos al segundo país. Entonces queda 1 médico para el primer país.

$$
\begin{aligned}
x_1^* &= 1,\\
x_2^* &= 0,\\
x_3^* &= 3,\\
x_4^* &= 1,\\
B_4^*(5) &= 13.
\end{aligned}
$$

Para estos datos, hay otras soluciones (óptimas) posibles, debido a que cuando se usa la fórmula de recurrencia (5.45), para estos datos, el mínimo se puede obtener con diferentes valores de j. Son soluciones:

$$(0,0,4,1)$$
$$(0,1,3,1)$$

5.3.3 Problema de médicos con cotas inferiores y superiores

El planteamiento de este problema es una generalización del anterior, ahora para cada país P_j hay una cota inferior u_j y una cota superior v_j para el número de médicos que se pueden asignar allí. Obviamente se debe cumplir que $0 \leq u_j \leq v_j \leq m$.

Los datos para este problema son:

- m número de médicos,
- n número de paises,
- u_1, ..., u_n cotas inferiores para el número de médicos,
- v_1, ..., v_n cotas superiores para el número de médicos,
- para cada país P_j los valores de los beneficios $b_{ij} \equiv b(i,j)$: $b_{u_j j}$, $b_{u_j+1,j}$, ..., $b_{v_j j}$.

Para que no haya información redundante, los datos deben cumplir la siguiente propiedad: la cota superior para el número de médicos en el país P_j debe permitir asignar el número mínimo de médicos en los otros paises, es decir:

$$v_j \leq m - \sum_{\substack{i=1\\ i\neq j}}^{n} u_i.$$

En caso contrario

$$v_j' = \min\{v_j, m - \sum_{\substack{i=1\\ i\neq j}}^{n} u_i\}.$$

El problema tiene solución si y solamente si el número total de médicos disponibles alcanza para cumplir con las cotas inferiores:

$$\sum_{j=1}^{n} u_j \le m.$$

El planteamiento del problema de optimización es el siguiente: encontrar x_1, x_2, ..., x_n para maximizar el beneficio total con ciertas restricciones:

$$\begin{aligned}
\max \sum_{j=1}^{n} b(x_j, j) & \\
\sum_{j=1}^{n} x_j &\le m, \\
u_j \le x_j &\le v_j, \ \ j = 1, ..., n, \\
x_j &\in \mathbb{Z}, \ \ j = 1, ..., n.
\end{aligned}$$

Antes de plantear la solución recurrente es necesario considerar lo siguiente: el número de médicos que se asignan al conjunto de paises P_1, P_2, ..., P_k, no puede ser inferior a la suma de sus cotas inferiores. Por otro lado, no puede ser superior a m, tampoco puede ser superior a la suma de sus cotas superiores y debe permitir satisfacer las cotas mínimas para el resto de paises, es decir, para los paises P_{k+1}, P_{k+2}, ..., P_n. Para esto se introducen unos nuevos valores

$$\begin{aligned}
U_k &= \sum_{j=1}^{k} u_j \qquad\qquad\qquad\qquad , \ \ k = 1, ..., n \\
V_k &= \min\{m, \ \sum_{j=1}^{k} v_j, \ \ m - \sum_{j=k+1}^{n} u_j\}, \ \ k = 1, ..., n \\
V_k &= \min\{\sum_{j=1}^{k} v_j, \ \ m - \sum_{j=k+1}^{n} u_j\}
\end{aligned}$$

Aquí se sobreentiende que el valor de una sumatoria es cero, cuando el límite inferior es más grande que el límite superior, por ejemplo,

$$\sum_{i=4}^{2} a_i = 0.$$

$$\begin{aligned}
B_k^*(i) = & \text{ beneficio máximo obtenido al asignar } i \text{ médicos} \\
& \text{ a los paises } P_1, \ P_2, \ ..., \ P_k, \ \ k = 1, ..., n, \ \ i = U_k, ..., V_k.
\end{aligned}$$

Objetivo final:

$$B_n^*(m) = \ ?$$

Condiciones iniciales:

$$B_1^*(i) = b_{i1}, \ \ i = U_1, ..., V_1.$$

Relación recurrente:

$$B_{k+1}^*(i) = \max_{\substack{0 \le j \le i \\ u_{k+1} \le j \le v_{k+1} \\ U_k \le i-j \le V_k}} \{B_k^*(i-j) + b_{j,k+1}\}, \quad k = 1, ..., n-1, \ \ i = U_{k+1}, ..., V_{k+1}.$$

En resumen, la variación de j en la relación recurrente está dada por:

$$\max\{i - V_k, u_{k+1}\} \leq j \leq \min\{i - U_k, v_{k+1}\}.$$

Consideremos los siguientes datos: $m = 5$, $n = 4$, $u_j : 0, 0, 1, 2$, $v_j : 3, 4, 2, 3$ y los siguientes beneficios:

i	b_{i1}	b_{i2}	b_{i3}	b_{i4}
0	0	0		
1	2	2	1	
2	4	3	4	5
3	6	4		6
4		8		

Si se considera el segundo país, para los demás paises en total se necesitan por lo menos 0+1+2=3 médicos, luego en el segundo país no se pueden asignar más de $5 - 3 = 2$ médicos, entonces los valores $b_{32} = 4$ y $b_{42} = 8$ nunca se van a utilizar. En realidad los datos con los cuales se va a trabajar son:

$$u_j : 0, 0, 1, 2, \quad v_j : 2, 2, 2, 3.$$

i	b_{i1}	b_{i2}	b_{i3}	b_{i4}
0	0	0		
1	2	2	1	
2	4	3	4	5
3				6

$$\begin{aligned}
U_1 &= 0, \ \ V_1 = \min\{2, 5 - 3\} = 2,\\
U_2 &= 0, \ \ V_2 = \min\{4, 5 - 3\} = 2,\\
U_3 &= 1, \ \ V_3 = \min\{6, 5 - 2\} = 3,\\
U_4 &= 3, \ \ V_4 = \min\{9, 5 - 0\} = 5.
\end{aligned}$$

Condiciones iniciales:

$$\begin{aligned}
B_1^*(0) &= 0, \ \ j^* = 0,\\
B_1^*(1) &= 2, \ \ j^* = 1,\\
B_1^*(2) &= 4, \ \ j^* = 2.
\end{aligned}$$

$$\begin{aligned}
B_2^*(1) &= \max_{0\leq j\leq 1} \{B_1^*(1 - j) + b_{j2}\}\\
&= \max\{B_1^*(1) + b_{02}, B_1^*(0) + b_{12}\}\\
&= \max\{\overline{2 + 0}, 0 + 2\}\\
B_2^*(1) &= 2, \ \ j^* = 0.
\end{aligned}$$

$\vdots$

$$\begin{aligned}
B_3^*(3) &= \max_{1\leq j\leq 2} \{B_2^*(3 - j) + b_{j3}\}\\
&= \max\{B_2^*(2) + b_{13}, B_2^*(1) + b_{23}\}\\
&= \max\{4 + 1, \overline{2 + 4}\}\\
B_3^*(3) &= 6, \ \ j^* = 2.
\end{aligned}$$

⋮

i	$B_1^*(i)$	j^*	$B_2^*(i)$	j^*	$B_3^*(i)$	j^*	$B_4^*(i)$	j^*
0	0	0	0	0				
1	2	1	2	0	1	1		
2	4	2	4	0	4	2		
3					6	2		
4								
5							11	2

El beneficio máximo es 11. Este beneficio máximo se obtiene asignando 2 médicos al cuarto país ($j^* = 2$). Entonces quedan 3 médicos para los primeros 3 paises. El j^* correspondiente a $B_3^*(3)$ es 2, esto indica que hay que asignar 2 médicos al tercer país. Entonces queda 1 médico para los primeros 2 paises. El j^* correspondiente a $B_2^*(1)$ es 0, esto indica que hay que asignar 0 médicos al segundo país. Entonces queda 1 médico para el primer país.

$$
\begin{aligned}
x_1^* &= 1,\\
x_2^* &= 0,\\
x_3^* &= 2,\\
x_4^* &= 2,\\
B_4^*(5) &= 11.
\end{aligned}
$$

El problema anterior también se puede resolver como un problema con cotas superiores (sin cotas inferiores) considerando unicamente los datos por encima de lo exigido por las cotas mínimas, es decir, de los m médicos disponibles en realidad hay únicamente

$$m' = m - \sum_{i=1}^{n} u_i$$

médicos para distribuir, pues de todas formas hay que asignar $\sum_{i=1}^{n} u_i$ para satisfacer las cotas mínimas.

La cota máxima para el número de médicos adicionales en el país P_j es naturalmente

$$v'_j = v_j - u_j, \;\; j = 1, ..., n.$$

De manera análoga se puede pensar en un beneficio b'_{ij} correspondiente al beneficio adicional al asignar i médicos, por encima de de la cota mínima u_j en el país P_j

$$b'_{ij} = b(i + u_j, j) - b(u_j, j), \;\; 1 \le j \le n, \;\; 0 \le i \le v'_j.$$

Al aplicar estos cambios a los datos del problema se tiene:

$$m' = 5 - 3 = 2, \;\; n = 4, \;\; v'_j : 2, 2, 1, 1.$$

i	b'_{i1}	b'_{i2}	b'_{i3}	b'_{i4}
0	0	0	0	0
1	2	2	3	1
2	4	3		

La solución de este problema modificado es:

$$
\begin{aligned}
x_1^{*'} &= 1,\\
x_2^{*'} &= 0,\\
x_3^{*'} &= 1,\\
x_4^{*'} &= 0,\\
B_4^{*'}(2) &= 5.
\end{aligned}
$$

Entonces

$$
\begin{aligned}
x_1^* &= 1 + u_1 = 1 + 0 = 1,\\
x_2^* &= 0 + u_2 = 0 + 0 = 0,\\
x_3^* &= 1 + u_3 = 1 + 1 = 2,\\
x_4^* &= 0 + u_4 = 0 + 2 = 2,\\
B_4^*(5) &= 5 + \sum_{i=1}^{n} b(u_i, i) = 5 + 6 = 11.
\end{aligned}
$$

Como observación final sobre **estos** datos numéricos, se puede ver que para cada uno de los tres problemas hay varias soluciones, en particular para este último problema hay otra solución:

$$
\begin{aligned}
x_1^* &= 0,\\
x_2^* &= 1,\\
x_3^* &= 2,\\
x_4^* &= 2,\\
B_4^*(5) &= 11.
\end{aligned}
$$

5.3.4 Problema del morral (knapsack)

Enunciado del problema

Un montañista está planeando una excursión muy especial. Evaluando la capacidad de su morral, la dificultad de la excursión, algunos implementos indispensables y sus fuerzas, cree que tiene en su morral una capacidad de $C \in \mathbb{Z}$ kilos (u otra unidad de peso, o de manera más precisa, de masa) disponibles para alimentos. De acuerdo con su experiencia, sus necesidades y sus gustos ha escogido n tipos de alimentos A_1, A_2, ..., A_n, todos más o menos equilibrados. Estos alimentos vienen en paquetes indivisibles (por ejemplo en lata) y $p_i \in \mathbb{Z}$ indica el peso de cada paquete del alimento A_i. Teniendo en cuenta la composición de cada alimento, las calorías, las vitaminas, los minerales, el sabor, el contenido de agua, etc., el montañista asignó a cada paquete del alimento A_i un beneficio global b_i.

El montañista desea saber la cantidad de paquetes de cada alimento que debe llevar en su morral, de tal manera que maximice su beneficio, sin sobrepasar la capacidad destinada para alimentos.

En este problema se supone que no es obligación llevar paquetes de cada uno de los alimentos. También se supone que no hay cotas inferiores ni superiores para el número de paquetes de cada alimento.

Tal vez ningún montañista ha tratado de resolver este problema para organizar su morral, seguramente ni siquiera ha tratado de plantearlo. Lo que si es cierto es que hay muchos problemas, de gran tamaño y de mucha importancia, que tienen una estructura análoga. Hay libros y muchos artículos sobre este problema.

Planteamiento del problema de optimización

Si x_j indica el número de paquetes del alimento A_j que el montañista debe llevar en su morral, entonces se debe maximizar el beneficio, bajo ciertas restricciones:

$$
\begin{aligned}
\max \sum_{j=1}^{n} b_j x_j&\\
\sum_{j=1}^{n} p_j x_j \leq C.&\\
x_j \in \mathbb{Z},\ \ j = 1, ..., n.&
\end{aligned}
$$

Este problema se puede resolver por la fuerza bruta construyendo todas las combinaciones, haciendo variar x_j entre 0 y $\lfloor C/p_j \rfloor$, verificando si cada combinación es factible ($\sum_{j=1}^{n} p_j x_j \leq C$) y escogiendo la mejor entre las factibles. El significado de $\lfloor\, t\, \rfloor$ es simplemente la parte entera inferior, o parte entera usual, es decir, el mayor entero menor o igual a t. Es claro que para valores grandes de n el número de combinaciones puede ser inmanejable.

La función objetivo (la función que hay que maximizar) es lineal, la restricción también es lineal, las variables deben ser enteras y se puede suponer que los coeficientes b_j y p_j también son enteros. Entonces este problema también se puede resolver por métodos de programación entera (programación lineal con variables enteras).

Solución recurrente

Para conocer optimamente el número de paquetes de cada uno de los n alimentos, se divide el morral con capacidad $i = C$ en dos "submorrales", uno con capacidad j utilizado únicamente para el alimento A_n y otro submorral con capacidad $i-j$ utilizado para los alimentos A_1, A_2,, ..., A_{n-1}. Si se conoce la solución optima de un morral con capacidad $c = 0, ..., C$ para los alimentos A_1, A_2,, ..., A_{n-1} entonces basta con estudiar todas las posibilidades de variación de j y escoger la mejor para obtener la respuesta global. Las posibles combinaciones son:

- C kilos para A_1,...,A_{n-1}, 0 kilos para A_n
- $C-1$ kilos para A_1,...,A_{n-1}, 1 kilo para A_n

  $\vdots$
- 0 kilos para A_1,...,A_{n-1}, C kilos para A_n

El razonamiento anterior se puede aplicar para los primeros $n-1$ alimentos con un morral de i kilos de capacidad, y así sucesivamente. Precisando más:

$$\begin{aligned} B_k^*(i) = {} & \text{beneficio máximo utilizando los primeros } k \text{ alimentos en un} \\ & \text{morral con capacidad de } i \text{ kilos, } 1 \le k \le n,\ 0 \le i \le C. \end{aligned}$$

Objetivo final:

$$B_n^*(C) = \ ?$$

Condiciones iniciales:

$$B_1^*(i) = b_1 \left\lfloor \frac{i}{p_1} \right\rfloor, \ \ 0 \le i \le C$$

Relación recurrente:

$$B_{k+1}^*(i) = \max_{0\le j\le i} \left\{ B_k^*(i-j) + b_{k+1} \left\lfloor \frac{j}{p_{k+1}} \right\rfloor \right\}, \ \ 1 \le k \le n-1, \ \ 0 \le i \le C.$$

Esta relación recurrente dice que la mejor manera de conocer el número de paquetes de los alimentos A_1, A_2, ..., A_{k+1} en un morral con una capacidad de i kilos es estudiando todas las posibilidades consistentes en dejar j kilos para el alimento A_{k+1} y el resto, $i-j$ kilos, para los alimentos A_1, A_2, ..., A_k.

i	$B_1^*(i)$	j^*	$B_2^*(i)$	j^*	$B_3^*(i)$	j^*	$B_4^*(i)$	j^*
0	0	0	0	0	0	0		
1	0	1	0	0	0	0		
2	0	2	0	0	6	2		
3	0	3	10	3	10	0		
4	14	4	14	0	14	0		
5	14	5	14	0	16	2		
6	14	6	20	6	20	0		
7	14	7	24	3	24	0		
8	28	8	28	0	28	0		
9	28	9	30	9	30	0		
10	28	10	34	6	34	0		
11	28	11	38	3	38	0	38	0

Resultados numéricos

Consideremos los siguientes datos: $C = 11,\ \ n = 4,\quad p_i : 4,\ 3,\ 2,\ 5,\quad b_i : 14, 10, 6, 17.8.$

$$
\begin{aligned}
B_1^*(0) &= 0,\ \ j^* = 0,\\
B_1^*(1) &= 0,\ \ j^* = 1,\\
B_1^*(2) &= 0,\ \ j^* = 2,\\
B_1^*(3) &= 0,\ \ j^* = 3,\\
B_1^*(4) &= 14,\ \ j^* = 4,\\
B_1^*(5) &= 14,\ \ j^* = 5,\\
B_1^*(6) &= 14,\ \ j^* = 6,\\
B_1^*(7) &= 14,\ \ j^* = 7,\\
B_1^*(8) &= 28,\ \ j^* = 8,\\
B_1^*(9) &= 28,\ \ j^* = 9,\\
B_1^*(10) &= 28,\ \ j^* = 10,\\
B_1^*(11) &= 28,\ \ j^* = 11.
\end{aligned}
$$

$\vdots$

$$
\begin{aligned}
B_2^*(3) &= \max_{1\le j\le 3}\left\{B_1^*(3-j) + b_2\left\lfloor \frac{j}{p_2}\right\rfloor\right\},\\
&= \max_{1\le j\le 3}\left\{B_1^*(3-j) + 10\left\lfloor \frac{j}{3}\right\rfloor\right\},\\
&= \max\{0+0, 0+0, 0+0, \overline{0+10}\},\\
B_2^*(3) &= 10,\ \ \ j^* = 3.
\end{aligned}
$$

En la tabla está el resumen de los resultados. El beneficio máximo es 38. Este beneficio máximo se obtiene asignando 0 kilos al cuarto alimento ($j^* = 0$). Entonces quedan 11 kilos para los primeros 3 alimentos. El j^* correspondiente a $B_3^*(11)$ es 0, esto indica que hay que asignar 0 kilos al tercer alimento. Entonces quedan 11 kilos para los primeros 2 alimentos. El j^* correspondiente a $B_2^*(11)$ es 3, esto indica que hay que asignar 3 kilos al segundo alimento, o sea, 1 paquete del segundo alimento. Entonces quedan 8 kilos para el primer alimento, o sea, 2 paquetes.

$$
\begin{aligned}
x_1^* &= 2,\\
x_2^* &= 1,\\
x_3^* &= 0,\\
x_4^* &= 0,\\
B_4^*(11) &= 38.
\end{aligned}
$$

Uno podría pensar que una manera simple de resolver este problema es buscar el alimento con mayor beneficio por kilo, y asignar la mayor cantidad posible de este alimento. Para los datos anteriores los beneficios por kilo son: 3.5, 3.33, 3 y 3.56 . Entonces se deberían llevar dos paquetes del cuarto alimento para un beneficio de 35.6 . Es claro que esta no es la mejor solución.

El problema del morral se puede convertir en uno análogo al problema de los médicos, introduciendo b_{ij} el beneficio obtenido con i kilos dedicados al alimento A_j

$$b_{ij} = b_j \left\lfloor \frac{i}{p_j} \right\rfloor$$

Para los datos anteriores se tendría la tabla

i	b_{i1}	b_{i2}	b_{i3}	b_{i4}
0	0	0	0	0
1	0	0	0	0
2	0	0	6	0
3	0	10	6	0
4	14	10	12	0
5	14	10	12	17.8
6	14	20	18	17.8
7	14	20	18	17.8
8	28	20	24	17.8
9	28	30	24	17.8
10	28	30	30	35.6
11	28	30	30	35.6

y la solución óptima

$$\begin{aligned} x_1^{*'} &= 8, \\ x_2^{*'} &= 3, \\ x_3^{*'} &= 0, \\ x_4^{*'} &= 0, \\ B_4^*(11) &= 38. \end{aligned}$$

Es conveniente recordar que x_i' indica el número de kilos dedicados al producto i, luego $x_1 = 2$, $x_2 = 1$, $x_3 = 0$, $x_4 = 0$, $B_4^*(11) = 38$.

Este paso por el problema de los médicos, presenta un inconveniente cuando C tiene un valor grande: es necesario construir una tabla muy grande de datos (los beneficios), cuando en realidad el conjunto de datos es pequeño: n, C, los p_i, los b_i.

El problema del morral también se puede generalizar al caso con cotas inferiores y superiores para el número de paquetes de cada alimento.

5.3.5 Problema de un sistema eléctrico

Enunciado del problema

Un sistema eléctrico está compuesto por n partes. Para que el sistema funcione se requiere que cada parte funcione. En cada parte hay que colocar por lo menos una unidad, pero se pueden colocar varias unidades para aumentar la probabilidad de que esa parte funcione. La probabilidad de que todo el sistema funcione es igual al producto de las probabilidades de que cada parte funcione. Los datos del problema son:

- n : número de partes
- v_i : número máximo de unidades que se pueden colocar en la parte i, $1 \leq i \leq n$
- $p_{ij} \equiv p(i,j)$: probabilidad de que la parte i funcione si se colocan j unidades, $1 \leq i \leq n$, $1 \leq j \leq v_i$

- $c_{ij} \equiv c(i,j)$: costo de colocar j unidades en la parte i del sistema, $1 \le i \le n$, $1 \le j \le v_i$
- $C \in \mathbb{Z}$: cantidad disponible para la construcción del sistema eléctrico

Se desea conocer el número de unidades que hay que colocar en cada parte de manera que se maximice la probabilidad de que todo el sistema funcione si se dispone de un capital de C pesos para la fabricación.

Se supone, lo cual es totalmente acorde con la realidad, que p_{ij}, c_{ij} son crecientes con respecto a j, es decir, si el número de unidades aumenta entonces ni el costo ni la probabilidad pueden disminuir, o sea, $p_{ij} \le p_{i,j+1}$, $c_{ij} \le c_{i,j+1}$ para $1 \le j \le v_i - 1$. También se supone que $1 \le v_i$. Además el dinero disponible C debe alcanzar para colocar en cada parte el número máximo de unidades posible, o sea, $c(i, v_i) \le C$ Si esto no es así, se puede modificar el valor v_i de la siguiente manera:

$$v_i' = j_i \quad \text{donde} \quad c(i, j_i) = \max_{1 \le j \le v_i} \{c_{ij} \mid c_{ij} \le C\}.$$

Además se debería cumplir

$$c(i, v_i) \le C - \sum_{\substack{j=1 \\ j \ne i}}^{n} c(j, 1), \quad \forall i.$$

El problema tiene solución si y solamente si

$$\sum_{i=1}^{n} c_{i1} \le C$$

Planteamiento del problema de optimización

Sea x_i el número de unidades colocadas en la parte i

$$\begin{aligned} \max f(x_1, x_2, ..., x_n) &= \prod_{i=1}^{n} p(i, x_i) \\ \sum_{i=1}^{n} c(i, x_i) &\le C, \\ 1 \le x_i &\le v_i, \quad i = 1, ..., n, \\ x_i &\in \mathbb{Z}, \quad i = 1, ..., n. \end{aligned}$$

Este problema se puede resolver con fuerza bruta, construyendo todas las posibilidades realizables y escogiendo la que maximice la probabilidad de que todo el sistema funcione.

Solución recurrente

La idea recurrente es la misma de los problemas anteriores, el problema se puede resolver dinámicamente, considerando inicialmente las mejores políticas para la primera parte, luego para la primera y segunda partes, en seguida para la primera, segunda y tercera partes y así sucesivamente hasta considerar todo el sistema eléctrico.

La cantidad de dinero t que se gasta en las primeras k partes tiene que ser suficiente para colocar por lo menos una unidad en cada una de las primeras k partes, además debe permitir que con el resto, $C - t$, se pueda colocar por lo menos una unidad en las restantes $n - k$ partes y no debe ser superior a la cantidad necesaria para colocar el número máximo de unidades en cada una de las primeras k partes. Entonces los límites de variación serán:

$$C_k = \sum_{i=1}^{k} c_{i1} \qquad\qquad 1 \le k \le n$$

$$D_k = \min \left\{ \sum_{i=1}^{k} c(i, v_i), \;\; C - \sum_{i=k+1}^{n} c_{i1} \right\}$$

Se puede considerar que, para el conjunto de todos las n partes, $C_n = D_n = C$.

$$P_k^*(t) = \text{probabilidad máxima de que el subsistema, formado por las primeras } k \text{ partes, funcione, si se gastan } t \text{ pesos en su construcción, donde } 1 \le k \le n,\ C_k \le t \le D_k.$$

Objetivo final:

$$P_n^*(C) = \ ?$$

El planteamiento puede resultar más claro si se define π_{it} como la máxima probabilidad de que la parte i funcione si se dispone de t pesos para su construcción,

$$\pi(i,t) = \pi_{it} := \max_{1\le j\le v_i} \{p_{ij} \mid c_{ij} \le t\}, \quad 1 \le i \le n,\ u_i \le t \le w_i.$$

π se puede considerar como una función o como una tabla. Las cotas para la variación de t son:

$$u_i = c(i,1)$$
$$w_i = \min\{c(i,v_i)\ ,\ C - \sum_{\substack{k=1\\k\ne i}}^{n} c(k,1)\ \}$$

Lo que se gasta en la parte i debe ser menor o igual al costo del mayor número de unidades en esa parte y también debe ser menor o igual a C menos lo que se requiere para colocar una unidad en las otras partes.

Condiciones iniciales:

$$P_1^*(t) = \ \pi_{1t}\,, \quad C_1 \le t \le D_1\,.$$

Relación recurrente: Si se dispone de t pesos para el subsistema formado por las partes 1, 2, ..., k, $k+1$, entonces se puede disponer de s pesos para la parte $k+1$ y, el resto, $t-s$, para las primeras k partes. Haciendo variar s se escoge la mejor posibilidad. Entonces la relación recurrente es:

$$P_{k+1}^*(t) = \max_{\substack{0\le s\le t\\ u_{k+1}\le s\le w_{k+1}\\ C_k\le t-s\le D_k}} \{\, P_k^*(t-s)\cdot \pi(k+1,s)\,\}, \quad C_{k+1} \le t \le D_{k+1}.$$

En esta fórmula resultan varias cotas inferiores y superiores para s. Algunas resultan redundantes. Sin embargo, ante la duda, es preferible escribir dos veces la misma cosa, que olvidar alguna restricción importante. Entonces s varía entre la mayor cota inferior y la menor cota superior.

Resultados numéricos

Consideremos los siguientes datos: $C = 10$, $n = 4$, v_i : 3, 3, 2, 3

$j \to$	1	2	3
p_{1j}	0.4	0.6	0.8
p_{2j}	0.5	0.6	0.7
p_{3j}	0.7	0.8	
p_{4j}	0.4	0.6	0.8

$j \to$	1	2	3
c_{1j}	2	3	5
c_{2j}	2	4	5
c_{3j}	2	5	
c_{4j}	1	3	6

Entonces:

$$\begin{aligned} u_1 &= 2, \quad w_1 = 5 \\ u_2 &= 2, \quad w_2 = 5 \\ u_3 &= 2, \quad w_3 = 5 \\ u_4 &= 1, \quad w_4 = 4 \end{aligned}$$

Tabla de π

$t \rightarrow$	1	2	3	4	5
π_{1t}		0.40	0.60	0.60	0.80
π_{2t}		0.50	0.50	0.60	0.70
π_{3t}		0.70	0.70	0.70	0.80
π_{4t}	0.40	0.40	0.60	0.60	

$$\begin{aligned} C_1 &= 2, \quad D_1 = 5, \\ C_2 &= 4, \quad D_2 = 7, \\ C_3 &= 6, \quad D_3 = 9, \\ C_4 &= 10, \quad D_4 = 10. \end{aligned}$$

Condiciones iniciales: con 2 pesos se puede colocar una unpdad en la primera parte y la probabilidad de que la primera parte funcione es 0.4, con 3 pesos se pueden colocar dos unidades en la primera parte y la probabilidad de que esta parte funcione es 0.6, con 4 pesos se pueden colocar únicamente dos unidades en la primera parte y la probabilidad sigue siendo 0.6, con 5 pesos se pueden colocar tres unidades en la primera parte y la probabilidad de que funcione la primera parte pasa a ser 0.6.

$$\begin{aligned} P_1^*(2) &= \pi_{12} = 0.4, \quad s^* = 2, \\ P_1^*(3) &= \pi_{13} = 0.6, \quad s^* = 3, \\ P_1^*(4) &= \pi_{14} = 0.6, \quad s^* = 4, \\ P_1^*(5) &= \pi_{15} = 0.8, \quad s^* = 5. \end{aligned}$$

Para el cálculo de $P_2^*(6)$ la variación de s está dada por

$$\begin{aligned} 0 \leq s \leq 6 \\ 2 \leq s \leq 5 \\ 2 \leq 6 - s \leq 5 \end{aligned}$$

En resumen, $2 \leq s \leq 4$.

$$\begin{aligned} P_2^*(6) &= \max_{2 \leq s \leq 4} \{P_1^*(6-s)\pi_{2s}\} \\ &= \max\{P_1^*(4)\pi_{22},\ P_1^*(3)\pi_{23},\ P_1^*(2)\pi_{24}\} \\ &= \max\{0.6 \times 0.5,\ 0.6 \times 0.5,\ 0.4 \times 0.6\} \\ P_2^*(6) &= 0.3, \quad s^* = 2. \end{aligned}$$

La tabla con el resumen de los reultados es la siguiente:

i	$P_1^*(i)$	s^*	$P_2^*(i)$	s^*	$P_3^*(i)$	s^*	$P_4^*(i)$	s^*
2	.4	2						
3	.6	3						
4	.6	4	.2	2				
5	.8	5	.3	2				
6			.3	2	.14	2		
7			.4	2	.21	2		
8					.21	2		
9					.28	2		
10							.126	3

Con 10 pesos disponibles, la probabilidad máxima de que el sistema funcione es 0.126, asignando 3 pesos para la cuarta parte, o sea, con 2 unidades en la cuarta parte. Quedan 7 pesos y el s^* correspondiente a $P_3^*(7)$ es 2, luego con 2 pesos se coloca una unidad en la tercera parte. Quedan 5 pesos y el s^* correspondiente a $P_2^*(5)$ es 2, luego con 2 pesos se coloca una unidad en la segunda parte. Quedan 3 pesos y el s^* correspondiente a $P_1^*(5)$ es 3, luego con 3 pesos se colocan dos unidades en la primera parte.

$$
\begin{aligned}
x_1^* &= 2,\\
x_2^* &= 1,\\
x_3^* &= 1,\\
x_4^* &= 2,\\
P_4^*(10) &= 0.126\,.
\end{aligned}
$$

5.3.6 Problema de mantenimiento y cambio de equipo por uno nuevo

Enunciado del problema

Hoy 31 de diciembre del año 0, el señor Tuta y su socio el señor Rodríguez, propietarios de un bus de e_0 años de edad, desean planificar su política de mantenimiento y compra de su equipo de trabajo, es decir de su bus, durante n años. La decisión de seguir con el mismo equipo o comprar uno nuevo se toma una vez al año, cada primero de enero. El señor Tuta tiene mucha experiencia y puede evaluar de manera bastante precisa los siguientes valores.

- u : vida útil del equipo, en años,
- p_i : precio de un equipo nuevo al empezar el año i, $1 \le i \le n$,
- $m_{ij} = m(i,j)$: precio del mantenimiento durante el año i, desde el 2 de enero hasta el 31 de diciembre, de un equipo que al empezar ese año (el 2 de enero), tiene j años de edad, $1 \le i \le n$, $0 \le j \le u-1$,
- $v_{ij} = v(i,j)$: precio de venta, el 1 de enero del año i, de un equipo que en esta fecha tiene j años de edad, $1 \le i \le n+1$, $1 \le j \le u$.

La decisión de comprar un equipo nuevo o guardar el que se tiene, se toma y se lleva a cabo el 1 de enero de cada año. Al final de los n años la compañía vende el equipo que tenga.

Solución recurrente

Sea E_k el conjunto de valores correspondientes a la edad que puede tener el equipo al finalizar el año k. Si $e_0 < u$ entonces la edad que puede tener el equipo al finalizar el primer año es 1 (si se compra uno nuevo) o $e_0 + 1$ (si se continua con el mismo equipo durante el primer año), entonces $E_0 = \{1, e_0 + 1\}$. Si $e_0 = u$ entonces la edad del equipo al finalizar el primer año es necesariamente 1, luego $E_0 = \{1\}$. De manera análoga, dependiendo de e_0 las edades al finalizar el segundo año pueden ser: $E_2 = \{1, 2, e_0 + 2\}$ o $E_2 = \{1, 2\}$. En general

$$
\begin{aligned}
E_0 &= \{e_0\},\\
E_{k+1} &= \{\,1\,\} \cup \{j+1 : j \in E_k,\ j < u\}.
\end{aligned}
$$

Sea

$C_k^*(e) =$ costo mínimo de la política de compra y mantenimiento del equipo desde el 31 de diciembre del año 0 hasta el 31 de diciembre del año k, de tal manera que el 31 de diciembre del año k el equipo tiene e años de edad, $1 \le k \le n$, $e \in E_k$.

Si el 31 de diciembre del año n el equipo tiene e años entonces hay que vender el 1 de enero del año $n+1$ a un precio $v(n+1, e)$. Entonces se debe escoger la mejor opción entre las posibles. O sea, el objetivo final es conocer el valor:

$$\min_{e \in E_n} \{C_n^*(e) - v_{n+1,e}\} = \,?$$

Condiciones iniciales: Si al empezar el primer año se compra un bus nuevo, bien sea porque $e_0 = u$ o bien sea porque se tomó esa decisión, se recupera el dinero de la venta, se compra un bus nuevo y durante el primer año se gasta en el mantenimiento de un bus con 0 años, así al final del primer año el bus tiene 1 año. Si no se compra bus entonces el único gasto es el manteniemiento de un bus de e_0 años, y al final del primer año el bus tiene $e_0 + 1$ años.

$$C_1^*(e) = \begin{cases} m(1, e_0) & \text{si } e = e_0 + 1, \\ -v(1, e_0) + p_1 + m_{1,0} & \text{si } e = 1. \end{cases}$$

Relación de recurrencia: Si e indica la edad del bus al final del año $k+1$, entonces $e = 1$ o e toma otros valores en E_{k+1}. Si $e > 1$ entonces al costo de la política óptima de los primeros k años se le aumenta el costo del mantenimiento de un bus con $e - 1$ años. Si $e = 1$ entonces al final del año k el bus podía tener d años, luego para cada edad d se toma el valor $C_k^*(d)$, se le resta lo de la venta, se le suma el valor de la compra y se le agrega el costo del mantemiento, y finalmente se escoge el menor valor.

$$C_{k+1}^*(e) = \begin{cases} \min\limits_{d \in E_k} \{C_k^*(d) - v_{k+1,d}\} + p_{k+1} + m_{k+1,0} & \text{si } e = 1, \\ C_k^*(e-1) + m_{k+1,e-1} & \text{si } e > 1. \end{cases}$$
$$k = 1, ..., n-1, \;\; e \in E_{k+1}.$$

Resultados numéricos

Consideremos los siguientes datos: $n = 4$, $u = 4$, $e_0 = 2$, p_i : 10, 12, 14, 15.

i	m_{i0}	m_{i1}	m_{i2}	m_{i3}
1	1	3	4	6
2	1	3	4	5
3	1	2	4	6
4	2	3	4	5

i	v_{i1}	v_{i2}	v_{i3}	v_{i4}
1	6	4	3	2
2	6	4	3	1
3	7	5	3	2
4	7	4	3	2
5	6	4	2	0

Entonces:

$$\begin{aligned} E_1 &= \{1, 3\}, \\ E_2 &= \{1, 2, 4\}, \\ E_3 &= \{1, 2, 3\}, \\ E_4 &= \{1, 2, 3, 4\}. \end{aligned}$$

Condiciones iniciales:

$$\begin{aligned} C_1^*(1) &= -4 + 10 + 1 = 7, \\ C_1^*(3) &= 4. \end{aligned}$$

$$\begin{aligned} C_2^*(1) &= \min_{d \in E_1} \{C_1^*(d) - v_{2d}\} + p_2 + m_{20} \\ &= \min\{C_1^*(1) - v_{21}, C_1^*(3) - v_{23}\} + p_2 + m_{20} \\ &= \min\{\overline{7-6}, 4 - 3\} + 12 + 1 \\ C_2^*(1) &= 14, \;\; d^* = 1. \end{aligned}$$

$$
\begin{aligned}
C_2^*(4) &= C_1^*(3) + m_{23} \\
&= 4 + 5 = 9, \;\; d^* = 3.
\end{aligned}
$$

$\vdots$

La tabla con el resumen de los reultados es la siguiente:

e	$C_1^*(e)$	$s*$	$C_2^*(e)$	s^*	$C_3^*(e)$	s^*	$C_4^*(e)$	s^*
1	7	2	14	1	20	2	28	3
2			10	1	16	1	23	1
3	4	2			14	2	20	2
4			9	3			19	3

Ahora es necesario encontrar $\min_e \{C_4^*(e) - v_{5e}\}$, es decir, el mínimo de $28-6$, $23-4$, $20-2$, $19-0$. Este valor mínimo es 18 y se obtiene para $e = 3$. Si e_k indica la edad del bus al finalizar el año k, entonces $e_4 = 3$, $e_3 = 2$, $e_2 = 1$. Al mirar en la tabla, para $C_2^*(1)$ se tiene que $d^* = 1$, o sea, $e_1 = 1$.

Si x_k indica la decisión tomada al empezar el año k, con la convención

$x_k = 0$: se compra un bus nuevo,

$x_k = 1$: se mantiene el bus que se tiene,

entonces se puede decir que $e_k = 1 \;\Rightarrow\; x_k = 0$ y que $e_k > 1 \;\Rightarrow\; x_k = 1$.

$$
\begin{aligned}
x_1^* &= 0, \\
x_2^* &= 0, \\
x_3^* &= 1, \\
x_4^* &= 1.
\end{aligned}
$$

5.3.7 Problema de producción y almacenamiento

Enunciado del problema

Considere una compañía que fabrica un bien no perecedero. Esta compañía estimó de manera bastante precisa las demandas d_1, d_2, ..., d_n de los n períodos siguientes. La producción en el período i, denotada por p_i, puede ser utilizada, en parte para satisfacer la demanda d_i, o en parte puede ser almacenada para satisfacer demandas posteriores. Para facilitar la comprensión del problema, supóngase que la demanda de cada período se satisface en los últimos dias del período. Sea x_i el inventario al final del período $i-1$ despues de satisfacer la demanda d_{i-1}, es decir, el inventario al empezar el período i. El costo $c_i(x_i, p_i)$ de almacenar x_i unidades y producir p_i unidades durante el período i se supone conocido. También se conoce x_1 el inventario inicial y x_{n+1} el inventario deseado al final de los n períodos.

Se desea planear la producción y el almacenamiento de cada período, de manera que permitan cumplir con las demandas previstas y se minimice el costo total de almacenamiento y producción.

Para facilitar el planteamiento se puede suponer que el inventario deseado al final de los n periodos se puede incluir en la demanda del último periodo d_n, o sea, hacer $d_n \leftarrow d_n + x_{n+1}$, $x_{n+1} \leftarrow 0$.

Planteamiento del problema de optimización

Las variables de este problema son: x_2, x_3, ..., x_n, p_1, p_2, ..., p_n. Recordemos que x_1, x_{n+1} son datos del problema. Las variables están relacionadas por las siguientes igualdades

$$
\begin{aligned}
x_1 + p_1 - d_1 &= x_2 \\
x_2 + p_2 - d_2 = x_1 + p_1 - d_1 + p_2 - d_2 &= x_3 \\
x_3 + p_3 - d_3 = x_1 + p_1 - d_1 + p_2 - d_2 + p_3 - d_3 &= x_4
\end{aligned}
$$

En general,

$$x_i + p_i - d_i = \sum_{j=1}^{i} p_j - \sum_{j=1}^{i} d_j + x_1 = x_{i+1}, \ \ i = 1, ..., n. \tag{5.46}$$

Es claro que para $i = n$,

$$x_n + p_n - d_n = \sum_{j=1}^{n} p_j - \sum_{j=1}^{n} d_j + x_1 = 0. \tag{5.47}$$

Sea

$$D_i = \sum_{j=1}^{i} d_j - x_1, \ \ \ i = 1, ..., n, \tag{5.48}$$

es decir, la demanda neta acumulada de los primeros i períodos, descontando el inventario inicial. En particular, toda la producción esta dada por $\sum_{j=1}^{n} p_j = D_n$. Entonces las igualdades que relacionan las variables son:

$$\sum_{j=1}^{i-1} p_j - D_{i-1} = x_i, \ \ i = 2, ..., n, \tag{5.49}$$

$$\sum_{j=1}^{n} p_j - D_n = 0. \tag{5.50}$$

El problema de optimización es entonces:

$$\begin{aligned} \min \ \ & \sum_{i=1}^{n} c_i(x_i, p_i) \\ & \sum_{j=1}^{i-1} p_j - D_{i-1} = x_i, \ \ \ i = 2, \dots, n \\ & \sum_{j=1}^{n} p_j - D_n = 0, \\ & x_2, \dots, x_n, p_1, \dots, p_n \geq 0, \\ & x_2, \dots, x_n, p_1, \dots, p_n \in \mathbb{Z}. \end{aligned}$$

El problema se puede plantear únicamente con las variables p_1, ..., p_{n-1}. A partir de (5.47) y (5.49) se tiene

$$\begin{aligned} -x_n + d_n &= p_n, \\ -\sum_{j=1}^{n-1} p_j + D_{n-1} + d_n &= p_n, \\ D_n - \sum_{j=1}^{n-1} p_j &= p_n. \end{aligned}$$

La función objetivo se puede agrupar en tres partes:

$$\min \ \ c_1(x_1, p_1) + \sum_{i=2}^{n-1} c_i(x_i, p_i) + c_n(x_n, p_n)$$

Entonces el problema, expresado únicamente con las variables p_1, ..., p_{n-1}, es :

$$\begin{aligned}
\min\ & c_1(x_1,p_1) + \sum_{i=2}^{n-1} c_i\big(\sum_{j=1}^{i-1} p_j - D_{i-1}, p_i\big) \\
& \quad + c_n\big(\sum_{j=1}^{n-1} p_j - D_{n-1}, D_n - \sum_{j=1}^{n-1} p_j\big) \\
& \sum_{j=1}^{i-1} p_j - D_{i-1} \geq 0, \quad i = 2, ..., n, \\
& D_n - \sum_{j=1}^{n-1} p_j \geq 0, \\
& p_i \geq 0,\ p_i \in \mathbb{Z}, \quad i = 1, ..., n-1.
\end{aligned}$$

Para resolver este problema por la fuerza bruta, estudiando todas las posibilidades, basta con considerar que el mayor valor de p_i se tiene cuando esta producción satisface por sí sola toda la demanda de los períodos i, $i+1$, ..., n. En ese caso no habría producción en los períodos $i+1$, ..., n. Dicho de otra forma, la producción en el período i no debe ser mayor que la demanda acumulada de los períodos i, ..., n.

Sea

$$E_i = \sum_{j=i}^{n} d_j, \quad i = 1, ..., n,$$

entonces los siguientes son límites para la variación de las variables p_i

$$0 \leq p_i \leq E_i, \quad i = 1, ..., n-1$$

Obviamente los valores p_1, ..., p_{n-1} deben cumplir las otras restricciones. Igualmente es claro que para las variables x_i también se tiene la misma restricción, o sea, ni el inventario al empezar el período i, ni la producción durante el período i pueden ser superiores a E_i De la definición de D_i y E_i se deduce inmediatamente que

$$D_i + E_{i+1} = D_n, \quad i = 1, ..., n-1$$

Solución recurrente

Sea

$$C_k^*(q) = \text{costo mínimo de producir en total } q \text{ unidades durante los primeros } k \text{ períodos tal manera que se satisfagan las demandas de estos períodos, } 1 \leq k \leq n.$$

La producción acumulada de los primeros k períodos debe ser suficiente para satisfacer la demanda total de estos k períodos (descontando x_1) y no debe sobrepasar la demanda total de los n períodos, entonces en la definición de $C_k^*(q)$ la variación de q está dada por

$$D_k \leq q \leq D_n.$$

Objetivo final:

$$C_n^*(D_n) = ?$$

Condiciones iniciales:

$$C_1^*(q) = c_1(x_1, q), \quad D_1 \leq q \leq D_n.$$

Si en los primeros $k+1$ períodos la producción es de q unidades y la producción en el período $k+1$ es de y unidades, entonces la producción en los primeros k períodos es de $q-y$ unidades y el inventario durante el período $k+1$ es $q-y-D_k$.

Relación recurrente:

$$C^*_{k+1}(q) = \min_{\substack{0\le q-y\le q \\ D_k\le q-y\le D_n \\ 0\le q-y-D_k\le E_{k+1} \\ 0\le\, y\,\le E_{k+1}}} \{C^*_k(q-y) + c_{k+1}(q-y-D_k, y)\},$$

definida para $k = 1, \ldots, n-1$, $D_{k+1} \le q \le D_n$. De las anteriores cotas inferiores y superiores para y y para $q-y$ se deduce fácilmente que $0 \le y \le q - D_k$. En resumen

$$C^*_{k+1}(q) = \min_{0\le y\le q-D_k} \{C^*_k(q-y) + c_{k+1}(q-y-D_k, y)\}.$$

Resultados numéricos

Consideremos los siguientes datos: $n = 4$, $d_i = 2, 3, 2, 5$, $c_i(x_i, p_i) = hx_i + \gamma_i p_i$, $h = 2$, $\gamma_i = 1, 4, 3, 6$, $x_1 = 1$, $x_5 = 0$.

Entonces

$$D_1 = 1,\ D_2 = 4,\ D_3 = 6,\ D_4 = 11.$$

$$\begin{aligned} C^*_1(q) &= c_1(x_1, q) \\ &= 2x_1 + 1q \\ &= 2 + q, \quad 1 \le q \le 11. \end{aligned}$$

$$\begin{aligned} C^*_2(q) &= \min_{0\le y\le q-1} \{C^*_1(q-y) + c_2(q-y-, y)\} \\ &= \min_{0\le y\le q-1} \{2 + (q-y) + 2(q-y-1) + 4y\} \\ &= \min_{0\le y\le q-1} \{3q + y\}, \quad y^* = 0 \\ &= 3q, \quad 4 \le q \le 11. \end{aligned}$$

$$\begin{aligned} C^*_3(q) &= \min_{0\le y\le q-4} \{C^*_2(q-y) + c_3(q-y-4, y)\} \\ &= \min_{0\le y\le q-4} \{3(q-y) + 2(q-y-4) + 3y\} \\ &= \min_{0\le y\le q-4} \{5q - 2y - 8\}, \quad y^* = q - 4 \\ &= 3q, \quad 6 \le q \le 11. \end{aligned}$$

$$\begin{aligned} C^*_4(11) &= \min_{0\le y\le 11-6} \{C^*_3(11-y) + c_4(11-y-6, y)\} \\ &= \min_{0\le y\le 5} \{3(11-y) + 2(5-y) + 6y\} \\ &= \min_{0\le y\le 5} \{43 + y\}, \quad y^* = 0 \\ &= 43. \end{aligned}$$

La producción óptima en el cuarto período es 0, luego la producción en los tres primeros períodos es 11. La producción óptima en el tercer período es $q - 4 = 11 - 4 = 7$, luego la producción en los dos primeros períodos es 4. La producción óptima en el segundo período es 0, entonces la producción óptima para el primer período es 4

$$\begin{aligned} p_4^* &= 0, \\ p_3^* &= 7, \\ p_2^* &= 0, \\ p_1^* &= 4, \\ C_4^*(11) &= 43. \end{aligned}$$

EJERCICIOS

5.1. Una corporación tiene n plantas de producción, P_1, P_2, ..., P_n. De cada una de ellas recibe propuestas para una posible expansión de las instalaciones. La corporación tiene un presupuesto de D billones de pesos para asignarlo a las n plantas. El gerente de la planta P_j envía m_j propuestas, indicando c_{ij} el costo de su propuesta i-ésima, $1 \le i \le m_j$, y g_{ij} la ganancia adicional total acumulada al cabo de 10 años, descontando la inversión. Obviamente en cada planta se lleva a cabo una sola propuesta. Una propuesta válida para cada una de las plantas consiste en no invertir en expansión, siendo su costo y ganancia nulos. Más aún, se podría pensar que en este caso la ganancia podría ser negativa. El gerente de cada planta envió las propuestas ordenadas por costo en orden creciente es decir, $c_{ij} \le c_{i+1,j}$, $i = 1, ..., m_j - 1$. Se supone que a mayor costo, también mayor ganancia.

Plantee el problema de optimización. Resuelva el problema por PD: defina una función que permita la recurrencia, dé las condiciones iniciales, la relación de recurrencia.

Resuelva el problema para los siguientes valores numéricos: $n = 3$, $D = 5$,

	Planta 1		Planta 2		Planta 3	
	c_{i1}	g_{i1}	c_{i2}	g_{i2}	c_{i3}	g_{i3}
Prop. 1	0	0	0	0	0	0
Prop. 2	2	8	1	5	1	3
Prop. 3	3	9	2	6		
Prop. 4	4	12				

5.2. El proceso de manufactura de un bien perecedero es tal que el costo de cambiar el nivel de producción de un mes al siguiente es $\$a$ veces el cuadrado de la diferencia de los niveles de producción. Cualquier cantidad del producto que no se haya vendido al final del mes se desperdicia con un costo de $\$b$ por unidad. Si se conoce el pronóstico de ventas d_1 , d_2 , ..., d_n para los próximos n meses y se sabe que en el último mes (el mes pasado) la producción fue de x_0 unidades.

Plantee el problema de optimización. Resuelva el problema por PD: defina una función que permita la recurrencia, dé las condiciones iniciales, la relación de recurrencia.

Resuelva el problema para los siguientes valores numéricos: $n = 4$, $a = 1$, $b = 2$, $d_i = 42$, 44, 39,36, $x_0 = 40$.

5.3. Considere el problema del morral con cotas inferiores y superiores con los siguientes datos: C = capacidad en kilos del morral (entero posotivo); n número de alimentos; p_1, p_2, ..., p_n, donde p_i es un entero positivo que indica el peso, en kilos, de un paquete del alimento i; b_1, b_2, ..., b_n, donde b_i indica el beneficio de un paquete del alimento i; u_1, u_2, ..., u_n, donde u_i es un entero no negativo que indica el número mínimo de paquetes del alimento i que el montañista debe llevar; v_1, v_2, ..., v_n, donde v_i es un entero que indica el número máximo de paquetes del alimento i que el montañista debe llevar. Los datos cumplen con la condición $u_i < v_i \ \forall i$.

Plantee el problema de optimización. Resuelva el problema por PD: defina una función que permita la recurrencia, dé las condiciones iniciales, la relación de recurrencia.

5.4. A partir del lunes próximo usted tiene n exámenes finales y $m \ge n$ días para prepararlos. Teniendo en cuenta el tamaño y la dificultad del tema usted a evaluado c_{ij} la posible nota o calificación que usted obtendría si dedica i días para estudiar la materia j. Usted ha estudiado regularmente durante todo el semestre y en todas sus materias tiene buenas notas de tal forma que aún obteniendo malas notas en los exámenes finales usted aprobará todas las materias. En consecuencia su único interés es obtener el mejor promedio posible. De todas maneras piensa dedicar por lo menos un día para cada materia.

Plantee el problema de optimización. Resuelva el problema por PD: defina una función que permita la recurrencia, dé las condiciones iniciales, la relación de recurrencia.

Resuelva el problema para los siguientes valores numéricos: $n = 4$ materias, $m = 7$ días,

c_{0j}	10	10	05	05
c_{1j}	20	15	35	10
c_{2j}	23	25	30	20
c_{3j}	25	30	40	35
c_{4j}	40	35	42	40
c_{5j}	45	40	45	45
c_{6j}	48	48	45	48
c_{7j}	50	48	48	50

5.5. Una empresa agrícola tiene actualmente x_0 empleados y conoce de manera bastante precisa las necesidades de mano de obra para las siguientes n semanas de cosecha, es decir, conoce los valores e_i, $i = 1, ..., n$, donde e_i es el número de empleados necesarios durante la semana i. También conoce: $c_i(j)$ el precio de contratar j empleados nuevos al empezar la semana i, $i = 1, ..., n$; $d_i(j)$ el costo de despedir j empleados al finalizar la semana i, $i = 0, ..., n$; y $m_i(j)$ el costo de mantener sin trabajo (pero con sueldo) j empleados durante la semana i, $i = 1, ..., n$. Después de las n semanas de cosecha, la empresa únicamente necesita e_{n+1}, un número pequeño de empleados que se quedan trabajando por varias semanas. La empresa desea saber cuantos empleados debe tener durante cada una de las n semanas.

Plantee el problema de optimización. Resuelva el problema por PD: defina una función que permita la recurrencia, dé las condiciones iniciales, la relación de recurrencia.

Resuelva el problema para los siguientes valores numéricos: $n = 4$, $x_0 = 10$, $e_i = 30, 45, 40, 25$, $e_{n+1} = 5$, $c_i(j) = 10 + 3j$, $d_i(j) = 6j$, $m_i(j) = 8j$.

5.6. Un zoocriadero de chigüiros tiene actualmente x_0 animales y tiene capacidad para una gran cantidad de ellos. En un año el número de chiguiros se multiplica por $a > 1$. Al principio de cada año (del año i) el gerente toma la decisión de vender algunos chigüiros al precio unitario p_i, $i = 1, ..., n + 1$. Después de n años se venden todos los chigüiros. El gerente desea saber cuantos chigüiros debe vender al comienzo de cada uno de los n años.

Plantee el problema de optimización. Resuelva el problema por PD: defina una función que permita la recurrencia, dé las condiciones iniciales, la relación de recurrencia.

Resuelva el problema para los siguientes valores numéricos: $n = 4$, $x_0 = 5$, $a = 3$, $p_i = 50, 10, 60, 25, 45$.

5.7. Resuelva por PD el siguiente problema de optimización:

$$\min \quad f(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 4)^2 + (x_1 - x_2)^2$$

Sugerencia. Fije una variable y halle la solución (en función de la variable fija). Haga variar la variable que estaba fija.

5.8. Resuelva por PD el siguiente problema de PL:

$$\begin{aligned} \max \quad z = x_1 + 1.4x_2 & \\ x_1 + x_2 &\leq 40 \\ x_1 + 2x_2 &\leq 58 \\ x &\geq 0. \end{aligned}$$

5.9. Una corporación tiene tres plantas de producción. De cada una de ellas recibe propuestas para una posible expansión de las instalaciones. La corporación tiene un presupuesto de 6 mil millones de pesos para asignarlo a las tres plantas. Cada planta envía sus propuestas, indicando la ganancia adicional (en miles de millones de pesos en valor presente) durante los próximos 20 años al invertir i miles de millones de pesos.

i	P_1	P_2	P_3
0	0	0	0
1	2	3	7
2	7	7	10
3	14	16	12
4	19	20	17
5	26	25	23
6	27	29	28

El gerente de la corporación resolvió el problema y los resultados no le parecieron adecuados, pues había plantas sin inversión. Por eso, colocó como condición que había que invertir por lo menos un millón, en cada una de las tres plantas.

5.10. Problema de compra y almacenamiento. Una imprenta conoce, d_1, ..., d_n sus necesidades de papel (en toneladas) para los siguientes n meses. También puede predecir de manera bastante aproximada los valores p_1, ..., p_n, precio (por tonelada) del papel al comienzo de cada uno de estos n meses. La imprenta paga inmediatamente sus pedidos y su proveedor de papel lo entrega también al comienzo de cada mes. La imprenta tiene cuatro áreas, una de oficinas, otra de máquinas, otra donde va almacenando el material impreso antes de entregarlo a sus clientes y una cuarta área de almacenamiento de papel, con muy buenas condiciones de seguridad, humedad y temperatura. Esta cuarta área tiene un capacidad de C toneladas. El área de máquinas tiene suficiente espacio para tener allí el papel que se va a utilizar durante cada mes. De igual manera, el área para material impreso tiene suficiente capacidad para tener temporalmente el material impreso mientras se envía a los clientes. Actualmente hay y_0 toneladas de papel. Al final de los n meses el stock final de papel debe ser nulo. El gerente de la imprenta necesita conocer x_1, x_2, ..., x_n, número de toneladas que debe comprar el comienzo de cada mes, de manera que el costo total sea mínimo. Los datos son: $C = 9$, $y_0 = 2$.

mes	1	2	3	4	5	6
d_i	5	8	2	4	7	5
p_i	11	13	18	17	10	20

Capítulo 6

Apéndice

6.1 Matriz semidefinida positiva en un cono

Sea $H \in \mathbb{R}^{n\times n}$, simétrica. Se desea saber si H es semidefinida positiva en el cono

$$K = \{x \in \mathbb{R}^n : Mx = 0,\ Dx \leq 0\} \tag{6.1}$$

Si H es semidefinida positiva (en todo $\mathbb{R}^n$), entonces H es semidefinida positiva en cualquier subconjunto.

Si H es semidefinida negativa, entonces H no es semidefinida positiva en K, salvo si $K = \{\mathbf{0}\}$.

En lo que sigue supondremos que H es indefinida y que $K \neq \{\mathbf{0}\}$.

El caso en que D "no tiene filas", es decir, $K = \{x \in \mathbb{R}^n : Mx = 0\}$, ya se estudió en la sección 1.7. Entonces se supondrá que la matriz D es una matriz $q \times 1$, con $q \geq 1$.

En cambio supondremos que en la definición de K, puede no estar $Mx = 0$, esto lo denotaremos diciendo que M puede no tener filas. Es decir, M es una matriz $p \times n$ con $p \geq 0$.

También se supone que las filas de M son linealmente independientes y que las filas de D también son linealmente independientes.

Averiguar si H es semidefinida positiva en K puede provenir de condiciones de segundo orden en un problema de optimización no lineal. En este caso, se puede suponer que **el conjunto formado por las filas de M y de D es un conjunto linealmente independiente**. Dicho de otra forma, las filas de la matriz

$$\begin{bmatrix} M \\ D \end{bmatrix}$$

son linealmente independientes.

6.1.1 Generalidades

Sea S un subconjunto de $\mathbb{R}^n$. Se dice que S es un **cono** si para todo $x \in S$ y $\mu > 0$, entonces μx también está en S.

K es un conjunto convexo, es intersección de hiperplanos y semiespacios. También se comprueba fácilmente que es un cono. O sea, K es un cono convexo.

Sea C un conjunto convexo, $d \in \mathbb{R}^n$, $d \neq 0$ es **dirección** de C si para todo $x \in C$ y todo $\mu > 0$, se cumple que $x + \mu d$ también está en C.

En un cono todos los puntos no nulos del cono son direcciones. Un conjunto acotado no tiene direcciones.

Dos direcciones de C, d^1 y d^2 son **equivalentes** si una es un múltiplo positivo de la otra:

$$d^1 = \mu d^2, \quad \text{con} \quad \mu > 0 \tag{6.2}$$

Sea d una dirección de C. Se dice que d es una **dirección extrema** de C si no existen d^1 y d^2, direcciones de C, no equivalentes, tales que $d = \mu_1 d^1 + \mu_2 d^2$ con $\mu_1, \mu_2 > 0$.

Una combinación lineal $\alpha_1 x^1 + \alpha_2 x^2 + \cdots + \alpha_k x^k$ se llama **cónica** si $\alpha_i \geq 0$ para todo i.

Una combinación cónica se llama **positiva** si todos los escalares son positivos.

Dados k puntos de $\mathbb{R}^n$, x^1, x^2, ..., x^k, el **cono generado** por estos k puntos es el conjunto de todas las combinaciones cónicas de estos puntos:

$$\text{cono}(x^1, x^2, ..., x^k) = \{\alpha_1 x^1 + \alpha_2 x^2 + \cdots + \alpha_k x^k \, : \, \alpha_i \geq 0, \ \forall i\} \tag{6.3}$$

Para saber si un punto está en el cono generado por x^1, x^2, ..., x^k es necesario encontrar los escalares. Si todos son no negativos, entonces el punto si está en el cono.

Por ejemplo, averiguar si $x = (21, 6)$ está en cono$((2, 4), (5, 1))$.

$$\begin{aligned} (21, 6) &= \alpha_1(2, 4) + \alpha_2(5, 1) \\ 21 &= 2\alpha_1 + 5\alpha_2 \\ 6 &= 4\alpha_1 + \alpha_2 \\ \alpha_1 &= 0.5 \\ \alpha_2 &= 4 \end{aligned}$$

Luego x está en cono$((2, 4), (5, 1))$.

Ejemplo 6.1. Sea $D = I_2$, la matriz identidad de orden 2,

$$K = \{x \in \mathbb{R}^2 : Dx \leq 0\}$$

K es el tercer cuadrante, $d^1 = (-1, 0)$, $d^2 = (0, -1)$, $d^3 = (-2, -3)$ son direcciones de K.

$d^4 = (1, 0)$ no es dirección.

d^3 es equivalente a $(-10, -15)$.

d^1 y d^2 son direcciones extremas. $(-0.1, 0)$ también es dirección extrema.

d^3 no es dirección extrema, $d^3 = 2d^1 + 3d^2$.

d^1 y d^2 son, salvo equivalencia, las únicas direcciones extremas.

Cualquier dirección es combinación cónica de las direcciones extremas. $\diamond$

Ejemplo 6.2.

$$K = \{x \in \mathbb{R}^2 : x_1 + 2x_2 = 0\}$$

$(2, -1)$ y $(-2, 1)$ son, salvo equivalencia, las direcciones extremas de K.

Ejemplo 6.3.

$$K = \{x \in \mathbb{R}^2 : x_1 + 2x_2 \leq 0\}$$

$(2, -1)$ y $(-2, 1)$ son, salvo equivalencia, las direcciones extremas de K.

6.1.2 Cono donde H es semidefinida positiva, $n = 2$

Sea $H \in S_n$ y sea P el conjunto de puntos donde H es semidefinida positiva,

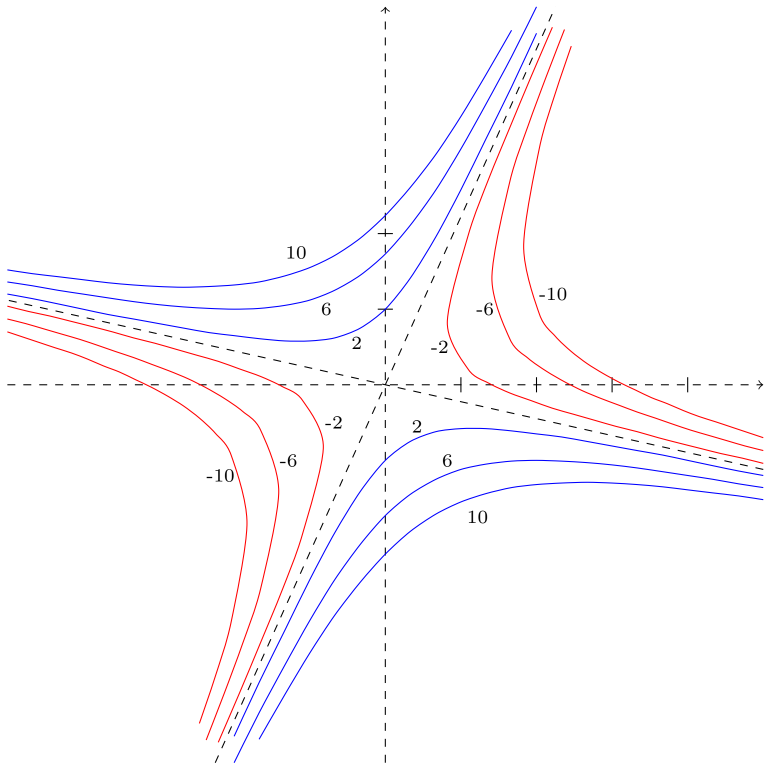
$$P = \{x \in \mathbb{R}^{n\times 1} : x^{\mathrm{T}} Hx \geq 0\} \tag{6.4}$$

Claramente este conjunto es un cono, no siempre convexo. Tiene una cierta simetría, si $x \in P$, también $-x \in P$. Esto se puede denotar, $-P = P$.

Sea $H \in S_2$ indefinida. Sean $\lambda_1 < 0 < \lambda_2$ los valores propios de H, v^1 y v^2 vectores propios asociados ortonormales (ortonormales quiere decir que son ortogonales dos a dos y que cada vector tiene norma 1): $Hv^i = \lambda_i v^i$, ${v^1}^{\mathrm{T}} v^2 = {v^2}^{\mathrm{T}} v^1 = 0$, ${v^i}^{\mathrm{T}} v^i = 1$. Sea $\varphi : \mathbb{R}^2 \to \mathbb{R}$,

$$\varphi(x) = x^{\mathrm{T}} Hx$$

La gráfica de φ es una "silla de montar", las curvas de nivel, $\varphi(x) = c$, para $c \neq 0$, son hipérbolas. Las asíntotas son las dos rectas tales $\varphi(x) = 0$. Estas dos rectas se cortan en el origen. Son de la forma tu, $t \in \mathbb{R}$, donde u es un vector fijo.



Sea $u \neq 0$, un punto en una asíntota. Entonces u se puede expresar como combinación lineal de v^1 y v^2:

$$u = \alpha_1 v^1 + \alpha_2 v^2$$

Como $\{v^1,\, v^2\}$ es una base ortonormal,

$$\begin{aligned}
\alpha_1 &= u^{\mathrm{T}} v^1, \quad \alpha_2 = u^{\mathrm{T}} v^2 \\
\varphi(u) &= (\alpha_1 v^1 + \alpha_2 v^2)^{\mathrm{T}} H (\alpha_1 v^1 + \alpha_2 v^2) \\
&= (\alpha_1 v^1 + \alpha_2 v^2)^{\mathrm{T}} (\alpha_1 H v^1 + \alpha_2 H v^2) \\
&= (\alpha_1 v^1 + \alpha_2 v^2)^{\mathrm{T}} (\alpha_1 \lambda_1 v^1 + \alpha_2 \lambda_2 v^2) \\
&= \alpha_1^2 \lambda_1 {v^1}^{\mathrm{T}} v^1 + \alpha_1 \alpha_2 \lambda_2 {v^1}^{\mathrm{T}} v^2 + \alpha_2 \alpha_1 \lambda_1 {v^2}^{\mathrm{T}} v^1 + \alpha_2^2 \lambda_2 {v^2}^{\mathrm{T}} v^2 \\
&= \alpha_1^2 \lambda_1 + \alpha_2^2 \lambda_2 = 0
\end{aligned}$$

Supongamos un valor fijo para una de las coordenadas, por ejemplo, para α_1,

$$\begin{aligned}
\alpha_1 &= 1 \\
0 &= \lambda_1 + \alpha_2^2 \lambda_2 \\
\alpha_2^2 &= -\frac{\lambda_1}{\lambda_2} \\
\alpha_2 &= \pm r \\
r &= \sqrt{-\frac{\lambda_1}{\lambda_2}}
\end{aligned} \tag{6.5}$$

Así, una pareja de vectores directores de las asíntotas es

$$\begin{aligned}
u^1 &= v^1 + r v^2 \\
u^2 &= v^1 - r v^2
\end{aligned}$$

Las asíntotas puede estar definidas por u^i o por $-u^i$. Se puede escoger el signo para que v^2 quede en el cono(u^1, u^2).

$$\begin{aligned}
u^1 &= -v^1 + r v^2 \\
u^2 &= v^1 + r v^2
\end{aligned} \tag{6.6}$$

Así, $v^2 = \frac{1}{2r}u^1 + \frac{1}{2r}u^2$. Ambos escalares son positivos y $v^2 \in \text{cono}(u^1, u^2)$.

Asíntotas:

$$S^1 = \{tu^1 : t \in \mathbb{R}\}$$
$$S^2 = \{tu^2 : t \in \mathbb{R}\}$$

Las asíntotas S_1 y S_2 dividen el plano en cuatro regiones de forma angular, R_1, R_2, R_3 y R_4, opuestas dos a dos. En un par de regiones opuestas $\varphi(x) \geq 0$; en el otro par $\varphi(x) \leq 0$. La matriz H será semidefinida postiva en las regiones donde está v^2 o donde está $-v^2$. Denotemos

$$R_2 = \text{cono}(u^1, u^2) \tag{6.7}$$
$$R_4 = \text{cono}(-u^1, -u^2) \tag{6.8}$$
$$P = R_2 \cup R_4 \tag{6.9}$$

Ejemplo 6.4. Determinar la región donde H es semidefinida positiva

$$H = \begin{bmatrix} -1 & -2 \\ -2 & 2 \end{bmatrix}$$

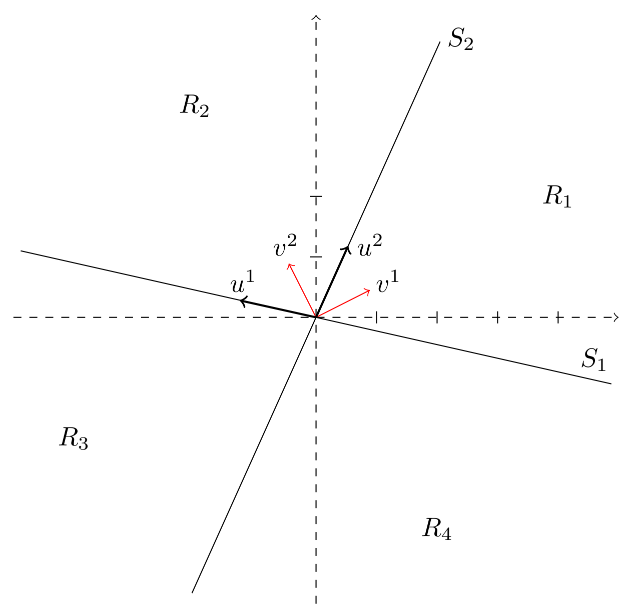
Los valores propios de H son: $\lambda_1 = -2$ y $\lambda_2 = 3$

$$v^1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad v^2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

normalizados

$$v^1 = \begin{bmatrix} 0.894427 \\ 0.447214 \end{bmatrix}, \quad v^2 = \begin{bmatrix} -0.447214 \\ 0.894427 \end{bmatrix}$$

$$r = \sqrt{-\frac{-2}{3}} = 0.816497$$
$$u^1 = \begin{bmatrix} -1.259576 \\ 0.283083 \end{bmatrix}, \quad u^2 = \begin{bmatrix} 0.529279 \\ 1.177510 \end{bmatrix}$$



H es semidefinida positiva en $R_2 \cup R_4$.

6.1.3 $n = 2$, K es un semiplano

Sea $K = \{x \in \mathbb{R}^2 : c^{\mathrm{T}}x \leq 0\}$, es decir, D tiene una sola fila y no hay matriz M.

Si H es indefinida, entonces H no es semidefinida positiva en K. No es posible que todo el semiplano quede contenido en la región R_2 o en la región R_4.

6.1.4 K es un semiespacio de $\mathbb{R}^n$

Sea $K = \{x \in \mathbb{R}^n : c^{\mathrm{T}}x \leq 0\}$, es decir, D tiene una sola fila y no hay matriz M.

Se puede mostrar que *si H es indefinida tampoco es semidefinida positiva en K.*

Sea Q el hiperplano definido por c,

$$Q = \{x \in \mathbb{R}^n : c^{\mathrm{T}}x = 0\}$$

Q es también un subespacio de dimensión $n-1$.

Sea $\lambda_1 < 0$ el valor propio mínimo y sea v^1 un vector propio de H asociado a λ_1, de norma 1. Supongamos inicialmente que $v^1 \notin Q$. Como $-v^1$ también es vector propio, se puede suponer que $c^{\mathrm{T}}v^1 < 0$, luego $v^1 \in K$.

$$\begin{aligned}
{v^1}^{\mathrm{T}} H v^1 &= {v^1}^{\mathrm{T}} \lambda_1 v^1 \\
&= \lambda_1 {v^1}^{\mathrm{T}} v^1 \\
&= \lambda_1 ||v^1||_2^2 \\
&= \lambda_1 < 0
\end{aligned}$$

Luego H no es semidefinida positiva en K.

Supongamos ahora que $v^1 \in Q$. Existe por lo menos otro vector propio v^2, normalizado, ortogonal a v^1 que no esté en Q, o sea, $c^{\mathrm{T}}v^2 \neq 0$, asociado a un valor propio α_2. Se puede escoger v^2 tal que $c^{\mathrm{T}}v^2 < 0$. Sea $\varepsilon > 0$,

$$\begin{aligned}
v &= v^1 + \varepsilon v^2 \\
c^{\mathrm{T}}v &= c^{\mathrm{T}}v^1 + \varepsilon\, c^{\mathrm{T}}v^2 \\
c^{\mathrm{T}}v &= \varepsilon\, c^{\mathrm{T}}v^2 < 0 \\
v &\in K \\
v^{\mathrm{T}}Hv &= (v^1 + \varepsilon\, v^2)^{\mathrm{T}} H (v^1 + \varepsilon\, v^2) \\
&= {v^1}^{\mathrm{T}}Hv^1 + \varepsilon {v^1}^{\mathrm{T}}Hv^2 + \varepsilon {v^2}^{\mathrm{T}}Hv^1 + \varepsilon^2 {v^2}^{\mathrm{T}}Hv^2 \\
&= \lambda_1 + \varepsilon {v^1}^{\mathrm{T}}\lambda_2 v^2 + \varepsilon {v^2}^{\mathrm{T}}\lambda_1 v^1 + \varepsilon^2 {v^2}^{\mathrm{T}}\lambda_2 v^2 \\
v^{\mathrm{T}}Hv &= \lambda_1 + \varepsilon^2\lambda_2
\end{aligned}$$

Si $\lambda_2 \leq 0$ entonces $v^{\mathrm{T}}Hv < 0$, luego H no es semidefinida positiva en K. Si $\lambda_2 > 0$, se puede escoger $\varepsilon > 0$ tal que $\lambda_1 + \varepsilon^2\lambda_2 < 0$. Luego, en resumen, H no es semidefinida positiva en K.

6.1.5 $n = 2$, K es una región angular

Sea $K = \{x \in \mathbb{R}^2 : Dx \leq 0\}$, es decir, no hay matriz M y D tiene por lo menos dos filas. Supondremos que ninguna fila es múltiplo de otra.

En un caso extremo, K puede ser $\{(0,0)\}$ o ser una semirecta. Supongamos que no es así.

K es intersección de semiplanos. Está completamente determinado conociendo las dos direcciones extremas. Es necesario encontrarlas. Cada semiplano determina dos direcciones extremas. Las dos direcciones extremas de K serán aquellas direcciones de los semiplanos que cumplan $Dd \leq 0$.

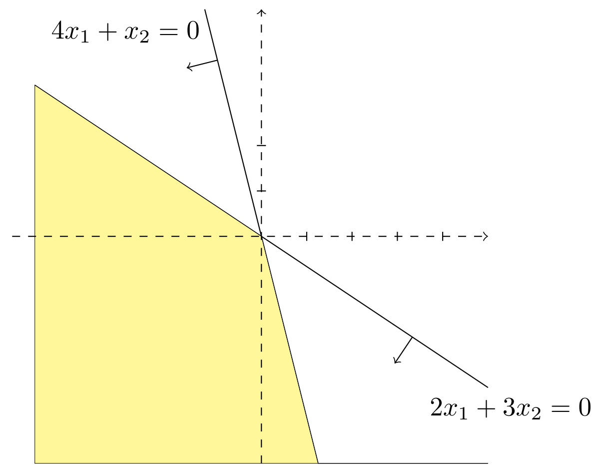
Ejemplo 6.5. Dibujar los conos $K_i = \{x \in \mathbb{R}^2 : D_ix \leq 0\}$

$$D_1 = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$$

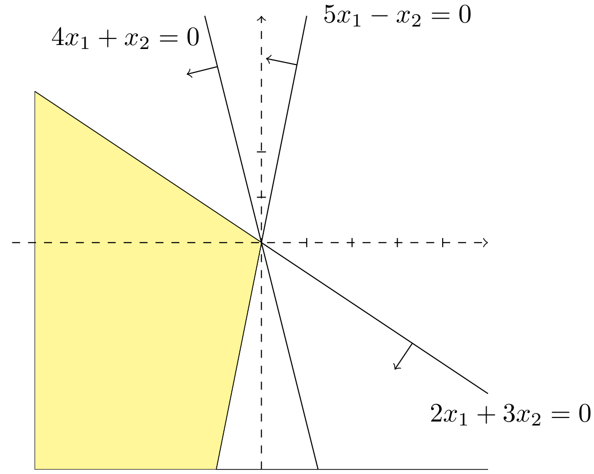
$$D_2 = \begin{bmatrix} 2 & 3 \\ 4 & 1 \\ 5 & -1 \end{bmatrix}$$

$$D_3 = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}$$

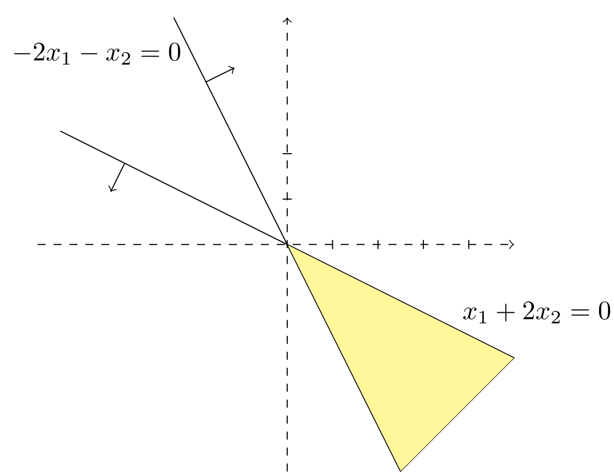
Cono K_1:



Cono K_2:



Cono K_3:



En el ejemplo anterior, salvo equivalencia, cada cono tiene únicamente dos direcciones extremas

En muchos casos, el cono $K \subseteq \mathbb{R}^n$ se puede expresar como combinación cónica de sus direcciones extremas. Los conos del ejemplo anterior cumplen esta propiedad.

Algunos de los casos donde la afirmación anterior no es cierta.

1. Un semiplano (en $\mathbb{R}^2$) cuya frontera pasa por el origen tiene únicamente dos direcciones extremas. Sin embargo el cono generado por ellas es simplemente la recta frontera.

2. $\mathbb{R}^n$, con $n \geq 2$, no tiene direcciones extremas.

Ejemplo 6.6. Hallar las direcciones extremas de los conos del ejemplo anterior.

$$D_1 = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$$

Direcciones extremas de $2x_1 + 3x_2 \leq 0$: $d = (-3, 2)$, $\quad d = (3, -2)$.

Direcciones extremas de $4x_1 + x_2 \leq 0$: $d = (-1, 4)$, $\quad d = (1, -4)$.

$$D_1 \begin{bmatrix} -3 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ -10 \end{bmatrix}, \quad D_1 \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 10 \end{bmatrix}, \quad D_1 \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ 0 \end{bmatrix}, \quad D_1 \begin{bmatrix} 1 \\ -4 \end{bmatrix} = \begin{bmatrix} -10 \\ 0 \end{bmatrix}$$

Luego las direcciones extremas son $d = (-3, 2)$ y $d = (1, -4)$.

$$D_2 = \begin{bmatrix} 2 & 3 \\ 4 & 1 \\ 5 & -1 \end{bmatrix}$$

Las 6 direcciones de los tres semiplanos son: $(-3, 2)$, $\ (3, -2)$, $\ (-1, 4)$, $\ (1, -4)$, $\ (1, 5)$, $\ (-1, -5)$.

Las direcciones extremas son $d = (-3, 2)$ y $d = (-1, -5)$, que cumplen $D_2 d \leq 0$.

$$D_3 = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}$$

Las direcciones extremas son $d = (2, -1)$ y $d = (1, -2)$. $\quad \diamond$

Sea $K \subseteq \mathbb{R}^2$ un cono generado por sus dos direcciones extremas d^1 y d^2. H es semidefinida positiva en K si y solamente si d^1 y d^2 están ambas en el cono(u^1, u^2) o ambas en el cono$(-u^1, -u^2)$

Supongamos que H es semidefinida positiva en $R_2 \cup R_4$, las direcciones d^1 y d^2 deben estar ambas en R_2 o ambas en R_4.

El proceso para averiguar si H es semidefinida positiva en K, tiene los siguientes pasos:

1. Obtener d^1 y d^2 direcciones extremas de K.
2. Si ${d^1}^{\mathrm{T}} H d^1 < 0$ o ${d^2}^{\mathrm{T}} H d^2 < 0$, entonces H no es semidefinida positiva en K. Fin del proceso.
3. Calcular los valores propios de H, $\lambda_1 < 0 < \lambda_2$.
4. Calcular v^1 vector propio asociado a λ_1 y v^2 vector propio asociado a λ_2.
5. Normalizar v^1 y v^2.
6. Calcular u^1 y u^2 utilizando (6.5) y (6.6).
7. Calcular α_1 y α_2 tales que
$$d^1 = \alpha_1 u^1 + \alpha_2 u^2$$
8. Calcular β_1 y β_2 tales que
$$d^2 = \beta_1 u^1 + \beta_2 u^2$$
9. Si el signo de α_1 y α_2 es igual al de β_1 y β_2 (d^1 y d^2 están en R^2 o d^1 y d^2 están en R^4), entonces H es semidefinida positiva en K.
10. Si H es semidefinida positiva y además ninguno de los escalares α_i , β_i, es nulo, entonces H es definida positiva en K.

Si el proceso va mas allá del paso 2, entonces d^1 y d^2 están en $R_2 \cup R_4$ conjunto donde H es semidefinida positiva. Por esa razón, depués de calcular α_1 y α_2, no es necesario verificar que tienen el mismo signo. Lo mismo para β_1 y β_2.

Ejemplo 6.7. Averiguar si H es sdp en $K = \{x \in \mathbb{R}^2 : Dx \leq 0\}$.

$$H = \begin{bmatrix} -1 & -2 \\ -2 & 2 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & 2 \\ -5 & 4 \end{bmatrix}$$

$$\begin{aligned} d^1 &= \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \quad d^2 = \begin{bmatrix} -4 \\ -5 \end{bmatrix} \\ {d^1}^{\mathrm{T}} H d^1 &= 38 \\ {d^2}^{\mathrm{T}} H d^2 &= -46 \end{aligned}$$

Luego H no es sdp en K. $\diamond$

Ejemplo 6.8. Averiguar si H es sdp en $K = \{x \in \mathbb{R}^2 : Dx \leq 0\}$.

$$H = \begin{bmatrix} -1 & -2 \\ -2 & 2 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 2 \\ -3 & -1 \end{bmatrix}$$

$$\begin{aligned} d^1 &= \begin{bmatrix} 2 \\ -4 \end{bmatrix}, \quad d^2 = \begin{bmatrix} 1 \\ -3 \end{bmatrix} \\ {d^1}^{\mathrm{T}} H d^1 &= 60 \\ {d^2}^{\mathrm{T}} H d^2 &= 29 \\ \lambda_1 &= -2, \quad \lambda_2 = 3 \\ v^1 &= \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad v^2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \\ v^1 &= \begin{bmatrix} 0.894427 \\ 0.447214 \end{bmatrix}, \quad v^2 = \begin{bmatrix} -0.447214 \\ 0.894427 \end{bmatrix} \\ r &= 0.81649658 \\ u^1 &= \begin{bmatrix} -1.259576 \\ 0.283083 \end{bmatrix}, \quad u^2 = \begin{bmatrix} 0.529279 \\ 1.177510 \end{bmatrix} \\ \alpha_1 &= -2.738613, \quad \alpha_2 = -2.738613 \\ \beta_1 &= -1.693422, \quad \beta_2 = -2.140636 \end{aligned}$$

Como los signos coinciden, se concluye que H es sdp en K. Más aún, H es definida positiva en K. $\diamond$

Uso del **coseno entre dos vectores**.

Si x es un punto de $\mathbb{R}^n$, también se puede interpretar como el vector que va desde el origen $\mathbf{0}$ hasta x. Si x, y están en $\mathbb{R}^n$ y no son nulos, entonces

$$\cos(\theta) = \frac{x^{\mathrm{T}} y}{||x||\,||y||} \tag{6.10}$$

donde θ es el ángulo formado por los dos vectores, o lo que es lo mismo, el ángulo determinado por los puntos x, $\mathbf{0}$, y.

Si dos vectores son paralelos y tienen la misma dirección, $\cos(\theta) = 1$. Si dos vectores son paralelos y tienen dirección opuesta, $\cos(\theta) = -1$.

Si $\cos(\theta) > 0$, los dos vectores tienen una dirección más o menos parecida. Si $\cos(\theta) < 0$, los dos vectores tienen una dirección más o menos opuesta, es decir, x y $-y$ tienen una dirección más o menos parecida.

Si ${d^1}^{\mathrm{T}} H d^1 \geq 0$ y ${d^2}^{\mathrm{T}} H d^2 \geq 0$, entonces d^1 y d^2 estan en el conjunto donde H es senidefinida positiva, es decir, en $P = R_2 \cup R_4$. El coseno permite averiguar en cual de las dos regiones está cada una de las dos direcciones.

$$\begin{aligned} c_1 &= \cos\big(\theta(d^1, v^2)\big) = \frac{{d^1}^{\mathrm{T}} v^2}{||d^1||\,||v^2||} \\ c_2 &= \cos\big(\theta(d^2, v^2)\big) = \frac{{d^2}^{\mathrm{T}} v^2}{||d^2||\,||v^2||} \end{aligned}$$

Si $c_1 > 0$ y $c_2 > 0$, entonces $d^1, d^2 \in R_2$, es decir, H es semidefinida positiva en K.

Si $c_1 < 0$ y $c_2 < 0$, entonces $d^1, d^2 \in R_4$, es decir, H es semidefinida positiva en K.

Si $c_1c_2 < 0$, entonces d^1 está en una región y d^2 está en la otra, es decir, H no es semidefinida positiva en K.

Como relamente lo que importa es el signo de c_1 y c_2, basta con calcular el producto ${d^i}^{\mathrm{T}} v^2$.

$$c_1 = {d^1}^{\mathrm{T}} v^2 \tag{6.11}$$
$$c_2 = {d^2}^{\mathrm{T}} v^2 \tag{6.12}$$

El proceso para saber si H es semidefinida positiva en K tiene los siguientes pasos:

a) Obtener d^1 y d^2 direcciones extremas de K.

b) Si ${d^1}^{\mathrm{T}} H d^1 < 0$ o ${d^2}^{\mathrm{T}} H d^2 < 0$, entonces H no es semidefinida positiva en K. Fin del proceso.

c) Calcular los valores propios de H, $\lambda_1 < 0 < \lambda_2$.

d) Calcular v^2 vector propio asociado a λ_2.

e) Calcular c_1 y c_2 usando (6.11) y (6.12).

f) Si $c_1c_2 < 0$, entonces H no es semidefinida positiva en K. Fin del proceso.

g) Si H es sdp en K y $\min\{ {d^1}^{\mathrm{T}} H d^1, {d^2}^{\mathrm{T}} H d^2 \} > 0$, entonces H es definida positiva en K. Si el mínimo es cero, entonces H es semidefinida positiva en K.

Ejemplo 6.9. Averiguar si H es sdp en $K = \{x \in \mathbb{R}^2 : Dx \leq 0\}$.

$$H = \begin{bmatrix} -1 & -2 \\ -2 & 2 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$$

$$\begin{aligned}
d^1 &= \begin{bmatrix} -3 \\ 2 \end{bmatrix}, \quad d^2 = \begin{bmatrix} 1 \\ -4 \end{bmatrix} \\
{d^1}^{\mathrm{T}} H d^1 &= 23 \\
{d^2}^{\mathrm{T}} H d^2 &= 47 \\
\lambda_1 &= -2, \quad \lambda_2 = 3 \\
v^2 &= \begin{bmatrix} -1 \\ 2 \end{bmatrix} \\
c_1 &= 7 \\
c_2 &= -9
\end{aligned}$$

Luego H no es sdp en K. $\diamond$

Ejemplo 6.10. Averiguar si H es sdp en $K = \{x \in \mathbb{R}^2 : Dx \leq 0\}$.

$$H = \begin{bmatrix} -1 & -2 \\ -2 & 2 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 2 \\ -3 & -1 \end{bmatrix}$$

$$\begin{aligned}
d^1 &= \begin{bmatrix} 2 \\ -4 \end{bmatrix}, \quad d^2 = \begin{bmatrix} 1 \\ -3 \end{bmatrix} \\
{d^1}^{\mathrm{T}} H d^1 &= 60 \\
{d^2}^{\mathrm{T}} H d^2 &= 29 \\
\lambda_1 &= -2, \quad \lambda_2 = 3 \\
v^2 &= \begin{bmatrix} -1 \\ 2 \end{bmatrix} \\
c_1 &= -10 \\
c_2 &= -7
\end{aligned}$$

Luego H es sdp en K. Además es definida positiva en K. $\diamond$

6.1.6 $n = 2$, K es una semirecta

$$K = \{x \in \mathbb{R}^2 : Mx = 0,\ Dx \leq 0\}, \quad M \in \mathbb{R}^{1\times 2}, \quad D \in \mathbb{R}^{1\times 2}$$
$$K = \{t\,v : t \geq 0\}, \quad v \neq 0, \quad Mv = 0, \quad Dv \leq 0$$

Observaciones: Si $Mv = 0$, entonces $M(-v) = 0$. Si $Dv \leq 0$, entonces $D(-v) \geq 0$. Además $(-v)^{\mathrm{T}}H(-v) = v^{\mathrm{T}}Hv$. El vector v genera N_M, el espacio nulo de M.

En consecuencia, basta con escoger $v \neq 0$ tal que $Mv = 0$. No importa si v cumple o incumple $Dv \leq 0$, es decir, la matriz D no se utiliza. Dicho de otra forma:

$$H \text{ es semidefinida positiva en } K \text{ sssi } H \text{ es semidefinida positiva en } N_M.$$

Entonces H es semidefinida positiva en K si y solamente si $v^{\mathrm{T}}Hv \geq 0$.

Ejemplo 6.11. Averiguar si H es semidefinida positiva en $K = \{x \in \mathbb{R}^2 : Mx = 0,\ Dx \leq 0\}$.

$$H = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad M = \begin{bmatrix} 4 & 5 \end{bmatrix}, \quad D = \begin{bmatrix} 6 & 7 \end{bmatrix}$$

$$v = \begin{bmatrix} -5 \\ 4 \end{bmatrix}$$
$$v^{\mathrm{T}}Hv = -7$$

Luego H no es sdp en K. $\diamond$

6.1.7 K es una semirecta de $\mathbb{R}^n$

$$K = \{x \in \mathbb{R}^n : Mx = 0,\ Dx \leq 0\}, \quad M \in \mathbb{R}^{(n-1)\times n}, \quad D \in \mathbb{R}^{1\times n}$$
$$K = \{t\,v : t \geq 0\}, \quad v \neq 0, \quad Mv = 0, \quad Dv \leq 0$$

Razonamientos análogos a los del caso $n = 2$ permiten afirmar:

$$H \text{ es semidefinida positiva en } K \text{ sssi } H \text{ es semidefinida positiva en } N_M$$

La única diferencia es que el número de filas de M es $n - 1$, generalización del caso anterior.

Ejemplo 6.12. Averiguar si H es semidefinida positiva en $K = \{x \in \mathbb{R}^3 : Mx = 0,\ Dx \leq 0\}$.

$$H = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}, \quad M = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$$

$$v = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$$
$$v^{\mathrm{T}}Hv = 4$$

Luego H es sdp en K. $\diamond$

6.1.8 Otros casos

Consideremos el caso general, no considerado en los casos anteriores:

$$K = \{x \in \mathbb{R}^n : Mx = 0, \quad Dx \leq 0\}$$

El conjunto K es el conjunto factible de un problema de OL, optimización lineal. Mediante resultados de OL, es posible encontrar las direcciones extremas de K (proceso fuera del alcance de este documento). Supongamos que d^1, d^2, ..., d^m son las direcciones extremas de K.

El siguiente proceso permite detectar muchos de los casos donde H no es sdp en K.

i) Si para algún $i = 1, ..., m$,

$$d^{i\,\mathrm{T}} H d < 0$$

entonces H no es sdp en K.

ii) Para todos los casos posibles, $i \neq j$, construir

$$d = d^i + d^j$$

Si

$$d^{\mathrm{T}} H d < 0$$

entonces H no es sdp en K.

iii) Para todos los casos posibles, i, j, k diferentes, construir

$$d = d^i + d^j + d^k$$

Si

$$d^{\mathrm{T}} H d < 0$$

entonces H no es sdp en K.

Ejemplo 6.13. ??????

6.2 Nociones elementales de análisis en $\mathbb{R}^n$

6.2.1 Normas

Sea V un espacio vectorial real (los escalares son números reales). Una norma es una función que “mide” el “tamaño” de los vectores (elementos de V).

Definición 6.1. Sea V un espacio vectorial, una **norma** es una función $\mu : V \to \mathbb{R}$ que cumple las siguientes propiedades:

$$\begin{aligned}
\mu(x) &\geq 0, \quad \forall x \in V \\
\mu(x) &= 0 \ \text{ sssi } \ x = \mathbf{0} \\
\mu(\alpha\, x) &= |\alpha|\, \mu(x), \quad \forall \alpha \in \mathbb{R},\ \forall x \in V \\
\mu(x+y) &\leq \mu(x) + \mu(y),\ \forall x, y \in V \qquad \text{desigualdad triangular}
\end{aligned}$$

Usualmente, la norma de x, es decir, $\mu(x)$, se denota por $||x||$. Así las propiedades se escriben:

$$\begin{aligned}
||x|| &\geq 0, \quad \forall x \in V \\
||x|| &= 0 \ \text{ sssi } \ x = \mathbf{0} \\
||\alpha\, x|| &= |\alpha|\, ||x||, \quad \forall \alpha \in \mathbb{R},\ \forall x \in V \\
||x+y|| &\leq ||x|| + ||y|| \qquad \text{desigualdad triangular}
\end{aligned}$$

Un espacio vectorial puede tener diferentes normas. Sea $V = \mathbb{R}^n$, las siguientes son algunas de las normas más conocidas.

- $||x||_2 = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}$, la norma euclidiana.
- $||x||_1 = \sum_{i=1}^n |x_i|$, norma del taxista.

- $||x||_{\max} = ||x||_\infty = \max\limits_{1 \le i \le n} |x_i|$.

- $||x||_p = \left(\sum\limits_{i=1}^{n} |x_i|^p \right)^{1/p}$, para $p \ge 1$, llamada norma de Hölder de orden p, o norma l_p

- $||x||_A = (x^{\mathrm{T}} A x)^{1/2}$ norma generada por la matriz A simétrica y definida positiva.

La norma euclidiana es un caso particular de la norma de Hölder para $p = 2$. La norma $||\ \ ||_1$ es un caso particular de la norma de Hölder para $p = 1$. La norma $||\ \ ||_\infty$ es el límite de la norma de Hölder cuando p tiende a infinito. La norma euclidiana es un caso particular de la norma $||\ \ ||_A$ cuando A es la matriz identidad. Si no se especifica la norma, se puede suponer que se trata de la norma euclidiana. Para $n = 2$ o $n = 3$, la norma euclidiana coincide exactamente con la distancia del punto x al origen.

$$\begin{aligned} ||\,(3, 0, -4)\,||_1 &= 7 \\ ||\,(3, 0, -4)\,||_1 &= 5 \\ ||\,(3, 0, -4)\,||_\infty &= 4 \end{aligned}$$

Sea $V = C_{[a,b]}$ el conjunto de funciones continuas en el intervalo $[a, b]$. Dos normas muy usadas son:

$$\begin{aligned} ||f|| &= \max_{a \le x \le b} |f(x)| \\ ||f|| &= \int_a^b |f(x)|\, dx \end{aligned}$$

6.2.2 Bolas abiertas y bolas cerradas

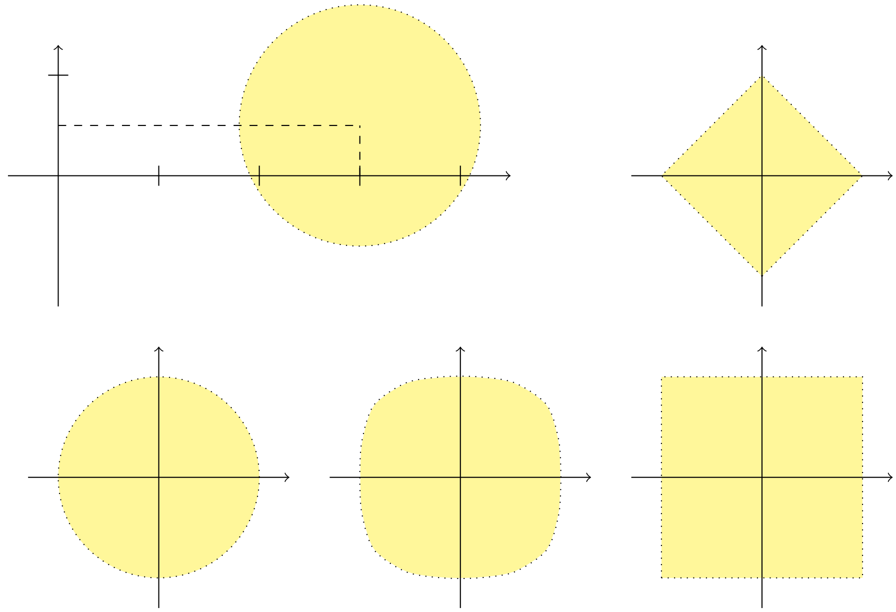
Si x, z están en $V = \mathbb{R}^n$, el valor $||x - z||$ mide la "distancia" entre x y z. Obviamente esa distancia depende de la norma utilizada. En $\mathbb{R}^2$ o en $\mathbb{R}^3$, el valor $||x - z||_2$ es exactamente la distancia usual entre los puntos x y z, por ejemplo, $||(1, 2) - (2, 4)||_2 = \sqrt{5}$, que es exactamente la distancia entre estos dos puntos.

Definición 6.2. Sea $V = \mathbb{R}^n$, $||\ \ ||$ una norma, $c \in V$, $r \in \mathbb{R}$, $r > 0$. La **bola abierta** con centro en c y radio r, para la norma $||\ \ ||$ es el conjunto de puntos x de V tales que su distancia a c, es decir, $||x - c||$, es menor que r,

$$B_{||\ \ ||}(c, r) = B(c, r) = \{x \in V : ||x - c|| < r\} \tag{6.13}$$

Sean $B_1(c, r)$, $B_2(c, r)$ y $B_\infty(c, r)$ la bolas para las normas $||\ \ ||_1$, $||\ \ ||_2$ y $||\ \ ||_\infty$. En $\mathbb{R}^2$, la bola $B_2(c, r)$ es exactamente el círculo con centro en c y radio r, excluyendo los puntos de la frontera. En $\mathbb{R}^3$, la bola $B_2(c, r)$ es exactamente la esfera[1] sólida con centro en c y radio r, excluyendo los puntos de la frontera.

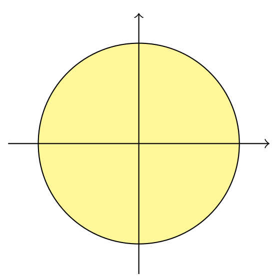
[1]En varias áreas de matemáticas, esfera significa únicamente la frontera de la bola o cáscara esférica, es decir, $\{x \in \mathbb{R}^3 : ||x - c||_2 = r\}$.



$B_2(\,(3,0.5),\,1.2),\quad B_1(\mathbf{0},1),\quad B_2(\mathbf{0},1),\quad B_3(\mathbf{0},1),\quad B_\infty(\mathbf{0},1)$

Definición 6.3. Sea $V=\mathbb{R}^n$, $||\ \ ||$ una norma, $c\in V$, $r\in\mathbb{R}$, $r>0$. La **bola cerrada** con centro en c y radio r, para la norma $||\ \ ||$ es el conjunto de puntos x de V tales que su distancia a c, es decir, $||x-c||$, es menor o igual que r,

$$\bar{B}(c,r)=B[c,r]=\{x\in V:||x-c||\le r\} \tag{6.14}$$



$\bar{B}_2(\mathbf{0},1)=B_2[\mathbf{0},1]$

6.2.3 Conjuntos abiertos y cerrados

Definición 6.4. Sea $A\subseteq\mathbb{R}^n$, $\bar{x}\in A$. El punto $\bar{x}$ es un **punto interior** de A si existe una bola abierta centrada en $\bar{x}$ contenida en A, es decir, existe $r>0$ tal que

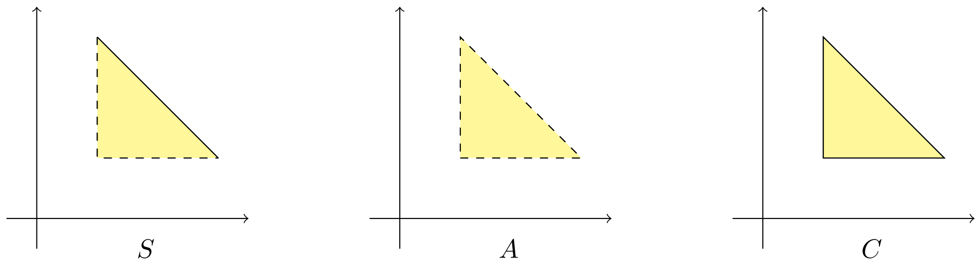
$$B(\bar{x},r)\subseteq A$$

Puesto que $B[\bar{x},r/2]\subseteq B(\bar{x},r)$, en la definición se puede cambiar la bola abierta por una bola cerrada.

En $\mathbb{R}^n$ en particular, no importa cual norma se use para encontrar una bola centrada en $\bar{x}$ contenida en A. Dicho de otra forma, si $\bar{x}$ es un punto interior de A usando una norma, también será punto interior usando cualquier otra norma. Este resultado es conocido como "en $\mathbb{R}^n$ todas las normas son equivalentes".

Ejemplo 6.14. Sean

$$\begin{aligned}
S&=\{x\in\mathbb{R}^2: x_1>1,\ \ x_2>1,\ \ x_1+x_2\le 4\}\\
A&=\{x\in\mathbb{R}^2: x_1>1,\ \ x_2>1,\ \ x_1+x_2< 4\}\\
C&=\{x\in\mathbb{R}^2: x_1\ge 1,\ \ x_2\ge 1,\ \ x_1+x_2\le 4\}
\end{aligned}$$



$\bar{x} = (1.1, 1.1)$ es punto interior de los tres conjuntos, ya que la bola $B(\,(1.1, 1.1),\ 0.01)$ está contenida en los tres conjuntos.

$\bar{x} = (2, 1)$ no está en S ni en A, luego no es punto interior de ellos. Este punto está en C pero cualquier bola centrada en $(2, 1)$, aunque el radio sea muy pequeño, tiene puntos que no están en C, es decir, no hay bolas centradas en $\bar{x}$ contenidas en C, luego no es punto interior de C.

Definición 6.5. Dado A subconjunto de $\mathbb{R}^n$ se define el **interior de** A como el conjunto de todos los puntos interiores de A

$$\text{int}A = \text{int}(A) = \mathring{A} = \{x : \ x \text{ es punto interior de } A\} \tag{6.15}$$

Siempre, $\mathring{A} \subseteq A$.

Definición 6.6. Un subconjunto de $\mathbb{R}^n$ es **abierto** si es igual a su interior, es decir, si todos sus puntos son puntos interiores.

Ejemplo 6.15. Del ejemplo anterior:

$$\begin{aligned}
&\mathring{S} = A\\
&\mathring{A} = A\\
&\mathring{C} = A\\
&S \text{ no es abierto}\\
&A \text{ es abierto}\\
&C \text{ no es abierto}
\end{aligned}$$

Otros ejemplos:

$$\begin{aligned}
&\mathbb{R}^n \text{ es abierto}\\
&\emptyset \ \text{ es abierto}\\
&\{\bar{x}\} \text{ no es abierto}\\
&B(c, r) \text{ es abierto}\\
&B[c, r] \text{ no es abierto}\\
&]3, 6[\text{ es abierto}\\
&[3, 6[\text{ no es abierto}
\end{aligned}$$

Sea $c \in \mathbb{R}^n$, $c \neq 0$, $\alpha \in \mathbb{R}$

$$\begin{aligned}
&\{x \in \mathbb{R}^{n\times 1} : c^{\mathrm{T}}x = \alpha\} \text{ no es abierto}\\
&\{x \in \mathbb{R}^{n\times 1} : c^{\mathrm{T}}x \leq \alpha\} \text{ no es abierto}\\
&\{x \in \mathbb{R}^{n\times 1} : c^{\mathrm{T}}x > \alpha\} \text{ es abierto}\\
&\{x \in \mathbb{R}^{n\times 1} : c^{\mathrm{T}}x < \alpha\} \text{ es abierto.}
\end{aligned}$$

Proposición 6.1. *La unión de dos conjuntos abiertos es un conjunto abierto. La unión de cualquier familia de conjuntos abiertos es un conjunto abierto. La intersección de dos conjuntos abiertos es un conjunto abierto. La intersección de cualquier familia finita de conjuntos abiertos es un conjunto abierto.*

El contraejemplo clásico relacionado con la última afirmación es:

$$]\text{-}1,1[\,\cap\,]\text{-}1/2,1/2[\,\cap\,]\text{-}1/3,1/3[\,\cap\cdots=\bigcap_{k=1}^{\infty}\,]\text{-}1/k,\ 1/k[\,=\{0\}$$

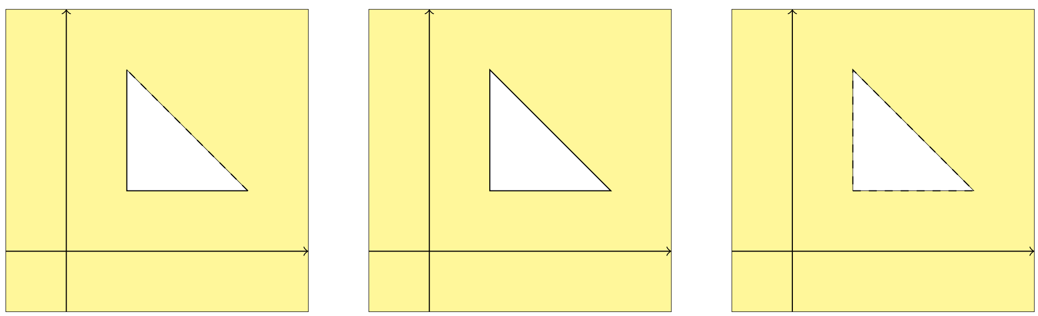
La intersección de esos intervalos abiertos no es un conjunto abierto.

Considerando como conjunto universal todo $\mathbb{R}^n$, el complemento de un conjunto es el conjunto de elementos de $\mathbb{R}^n$ que no están en A,

$$\complement A=\{x\in\mathbb{R}^n\ :\ x\notin A\} \tag{6.16}$$

Definición 6.7. Sea $A\subseteq\mathbb{R}^n$. Se dice que A es un conjunto **cerrado** si $\complement A$ es un conjunto abierto.

Ejemplo 6.16. Para los conjuntos del ejemplo 6.14:



$\complement S$ $\complement A$ $\complement C$

Entonces, C es un conjunto cerrado, S y A no son cerrados.

$\mathbb{R}^n$ es cerrado
$\emptyset$ es cerrado
$\{\bar{x}\}$ es cerrado
$B(c,r)$ no es cerrado
$B[c,r]$ es cerrado
$]3,6[$ no es cerrado
$[3,6[$ no es cerrado
$[3,6]$ es cerrado

Hay conjuntos que son abiertos y no son cerrados, hay conjuntos cerrados que no son abiertos, hay conjuntos que no son ni cerrados ni abiertos, hay conjuntos que son al tiempo abiertos y cerrados, $\mathbb{R}^n$, $\emptyset$.

Proposición 6.2. *La intersección de dos conjuntos cerrados es un conjunto cerrado. La intersección de cualquier familia de conjuntos cerrados es un conjunto cerrado. La unión de dos conjuntos cerrados es un conjunto cerrado. La unión de cualquier familia finita de conjuntos cerrados es un conjunto cerrado.*

Proposición 6.3. *Sea $f:\mathbb{R}^n\to\mathbb{R}$ continua, $\alpha\in\mathbb{R}$. Entonces*

$$\begin{aligned}
&\{x: f(x)<\alpha\} && \textit{es abierto}\\
&\{x: f(x)>\alpha\} && \textit{es abierto}\\
&\{x: f(x)\le\alpha\} && \textit{es cerrado}\\
&\{x: f(x)\ge\alpha\} && \textit{es cerrado}\\
&\{x: f(x)=\alpha\} && \textit{es cerrado}
\end{aligned}$$

6.2.4 Conceptos adicionales

Es posible definir conjuntos cerrados sin utilizar el concepto de conjunto abierto. Obviamente las dos definiciones deben ser, y son, equivalentes.

Definición 6.8. Un punto $x \in \mathbb{R}^n$ es **punto adherente** de A o **punto de adherencia** de A, si $\forall r > 0$, $B(x, r) \cap A \neq \emptyset$, es decir, si toda bola centrada en x tiene por lo menos un punto de A.

Definición 6.9. Un punto $x \in \mathbb{R}^n$ es **punto de acumulación** de A si $\forall r > 0$, $B(x, r) \cap (A \setminus \{x\}) \neq \emptyset$, es decir, si toda bola centrada en x tiene por lo menos un punto de A diferente de x.

Un punto de adherencia de A no necesariamente está en A. Un punto de acumulación de A no necesariamente está en A. Los puntos de A son puntos de adherencia de A. Los puntos de A no son necesariamente punto de acumulación. Un conjunto finito no tiene puntos de acumulación.

En este documento, acum(A) denotará el conjunto de todos los puntos de acumulación de A,

$$\text{acum}(A) = \{x : x \text{ es punto de acumulación de } A\} \tag{6.17}$$

El anterior conjunto se conoce como el derivado de A y con bastante frecuencia lo denotan con A'.

Definición 6.10. La **adherencia**, **cerradura** o **clausura** de un conjunto A es el conjunto de todos los puntos de adherencia de A, o también, es el conjunto formado por los puntos de A y sus puntos de acumulación.

$$\bar{A} = \{x : x \text{ es punto de adherencia de } A\} = A \cup \text{acum}(A) \tag{6.18}$$

Siempre se cumple que

$$A \subseteq \bar{A}$$

Definición 6.11. Un conjunto es **cerrado** si contiene todos su puntos de acumulación,

$$\text{acum}(A) \subseteq A \tag{6.19}$$

También, si y solamente si,

$$\bar{A} = A \tag{6.20}$$

Definición 6.12. Un punto de A que no sea punto de acumulación se llama **punto aislado** de A.

Definición 6.13. Un punto x es **punto de frontera** de A si toda bola centrada en x tiene puntos de A y puntos que no están en A. La **frontera** de un conjunto es el conjunto de todos los puntos de frontera,

$$\partial A = \{x : x \text{ es punto de frontera de } A\} \tag{6.21}$$

Un punto de frontera no necesariamente está en A. Un punto aislado de A está en la frontera de A.

Notación. Sean x, y dos puntos de $\mathbb{R}^n$. El segmento de recta que une x con y se denota

$$[x, y] = \{(1 - t)x + ty : t \in [0, 1]\}$$

Ejemplo 6.17. Sean S, A, C como en el ejemplo 6.14,

$$\begin{aligned}
S &= \{x \in \mathbb{R}^2 : x_1 > 1,\ \ x_2 > 1,\ \ x_1 + x_2 \leq 4\} \\
A &= \{x \in \mathbb{R}^2 : x_1 > 1,\ \ x_2 > 1,\ \ x_1 + x_2 < 4\} \\
C &= \{x \in \mathbb{R}^2 : x_1 \geq 1,\ \ x_2 \geq 1,\ \ x_1 + x_2 \leq 4\} \\
D &= S \cup \{(0, 0)\} \\
E &= A \cup \{(0, 0)\} \\
K &= C \cup \{(0, 0)\} \\
F &= [(1, 1), (3, 1)] \cup [(3, 1), (1, 3)] \cup [(1, 3), (1, 1)] \\
G &= F \cup \{(0, 0)\}
\end{aligned}$$

Entonces,

$$
\begin{aligned}
&\bar{S} = \bar{A} = \bar{C} = C\\
&\bar{D} = \bar{E} = \bar{K} = K\\
&(2,1) \in \text{acum}(S)\\
&(2,1) \in \text{acum}(A)\\
&(2,1) \in \text{acum}(C)\\
&(2,1) \in \text{acum}(D)\\
&(2,1) \in \text{acum}(K)\\
&(0,0) \notin \text{acum}(K)\\
&(0,0) \text{ es punto aislado de } K\\
&\partial S = \partial A = \partial C = F\\
&\partial D = \partial E = \partial K = G
\end{aligned}
$$

Caracterizaciones "constructivas" de interior y de adherencia. Sea $S \subseteq \mathbb{R}^n$.

$$\mathring{S} = \bigcup_{\substack{A \subseteq S\\ A \text{ abierto}}} A \tag{6.22}$$

$$\bar{S} = \bigcap_{\substack{S \subseteq C\\ C \text{ cerrado}}} C \tag{6.23}$$

6.3 Factorización de Cholesky

Sea A una matriz simétrica. Bajo ciertas condiciones (como se verá posteriormente, si y solamente si es definida positiva), existe una matriz U triangular superior invertible tal que

$$U^{\mathrm{T}}U = A. \tag{6.24}$$

Esta factorización, cuando existe, se llama factorización de Cholesky y, en este caso, basta con conocer la matriz U.

El cálculo se puede hacer por filas, es decir, primero se obtienen los elementos de la primera fila de U, en seguida los de la segunda, etc.

Supongamos conocidos los elementos de las filas $1, 2, ..., k$-1 de la matriz U. O sea, se conocen los elementos u_{ij} para $i = 1, 2, ..., k$-1 y para $j = i, i + 1, ..., n$. Como U es triangular superior, se sabe también que $u_{ij} = 0$ para $i > j$. Al multiplicar la fila k de la matriz U^{T} por la columna k de la matriz U se tiene:

$$
\begin{aligned}
a_{kk} &= (U^{\mathrm{T}})_{k\cdot}U_{\cdot k}\\
&= (U_{\cdot k})^{\mathrm{T}}U_{\cdot k}\\
&= \sum_{i=1}^{n} u_{ik}^2\\
&= \sum_{i=1}^{k} u_{ik}^2 + \sum_{i=k+1}^{n} u_{ik}^2
\end{aligned}
$$

$$
\begin{aligned}
a_{kk} &= \sum_{i=1}^{k} u_{ik}^2 + \sum_{i>k} u_{ik}^2 \\
&= \sum_{i=1}^{k} u_{ik}^2 \\
&= \sum_{i=1}^{k\text{-}1} u_{ik}^2 + u_{kk}^2.
\end{aligned}
$$

De la última igualdad todo se conoce salvo u_{kk}, entonces

$$
u_{kk} = \sqrt{a_{kk} - \sum_{i=1}^{k\text{-}1} u_{ik}^2}\ , \quad k = 1, ..., n. \tag{6.25}
$$

0

Para que tenga sentido la raíz cuadrada se necesita que la cantidad bajo el radical sea no negativa. Como además, U es invertible (y su determinante es igual al producto de los elementos diagonales), entonces $u_{kk} \neq 0$. Luego para poder obtener U de manera adecuada se necesita que

$$
a_{kk} - \sum_{i=1}^{k\text{-}1} u_{ik}^2 > 0\ , \quad k = 1, ..., n. \tag{6.26}
$$

Al multiplicar la fila k de la matriz U^{T} por la columna j de la matriz U, con $k < j$, se tiene:

$$
\begin{aligned}
a_{kj} &= (U^{\mathrm{T}})_{k\cdot} U_{\cdot j} \\
&= (U_{\cdot k})^{\mathrm{T}} U_{\cdot j} \\
&= \textstyle\sum_{i=1}^{n} u_{ik} u_{ij}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{i=1}^{k} u_{ik} u_{ij} + \sum_{i=k+1}^{n} u_{ik} u_{ij} \\
&= \sum_{i=1}^{k} u_{ik} u_{ij} + \sum_{i>k} u_{ik} u_{ij} \\
&= \sum_{i=1}^{k} u_{ik} u_{ij} \\
a_{kj} &= \sum_{i=1}^{k\text{-}1} u_{ik} u_{ij} + u_{kk} u_{kj}.
\end{aligned}
$$

De la última igualdad todo se conoce salvo u_{kj}, entonces

$$
u_{kj} = \Big(a_{kj} - \sum_{i=1}^{k\text{-}1} u_{ik} u_{ij}\Big) \frac{1}{u_{kk}}\ , \qquad \begin{aligned} k &= 1, ..., n\text{-}1 \\ j &= k+1, ..., n. \end{aligned} \tag{6.27}
$$

En resumen, si siempre se cumple la condición (6.26), mediante las fórmulas (6.25) y (6.27) se puede obtener la matriz U.

6.4 algoritmo de la factorización de Cholesky

La factorización de Cholesky es muy empleada en la solución de sistemas de ecuaciones lineales cuando la matriz de coeficientes es simétrica y definida positiva, lo cual es bastante frecuente. Por ejemplo, las matrices de rigidez de los problemas estructurales son definidas positivas. Dado el sistema $Ax = b$, si se conoce la factorización, se tiene entonces $U^{\mathrm{T}}Ux = b$. Si se toma $y = Ux$ el problema se convierte en $U^{\mathrm{T}}y = b$. Este sistema de ecuaciones lineales es muy fácil de resolver, puesto que es triangular inferior. Una vez obtenido y, se resuelve $Ux = y$, también fácil de efectuar puesto que es triangular superior, obteniéndose así x solución del sistema inicial. El método de Cholesky es más rápido que el método de Gauss.

Otra de las ventajas importantísimas del método de la factorización de Cholesky, cuando se hace utilizando el computador, es que la matriz U se puede almacenar exactamente donde estaba la matriz A, o sea, si el programa se elabora de forma adecuada, a medida que se obtiene un elemento u_{ij}, éste se puede almacenar justamente donde estaba el elemento a_{ij}. Obviamente al finalizar el proceso los valores de la matriz A se han perdido, y en su lugar está la matriz U. En la mayoría de los casos esto no es ningún problema, por el contrario, es una gran ventaja: únicamente se requiere espacio para almacenar una matriz y no para dos.

En el algoritmo presentado a continuación se hace uso de esta ventaja, así al comienzo los valores a_{ij} corresponden, en realidad a los elementos de A, pero al final estarán los valores de la matriz U. Si al final la variable `error` vale 0, se está indicando que se pudo efectuar la factorización de Cholesky, si vale 1 no hay factorización de Cholesky para esta matriz, o sea, A no es definida positiva. La variable de entrada ε indica la tolerancia o precisión en el siguiente sentido: los valores positivos menores que ε se pueden considerar como nulos.

```
datos: A, n, ε
resultados: A, error
error = 0
para k = 1,...,n  mientras error= 0
    s = a_kk
    para i = 1,...,k − 1
        s = s − a_ik^2
    fin-para i
    si s < ϵ  ent
        error= 1
        parar
    fin-ent
    sino
        a_kk = √s
        para  j = k + 1,...,n
            t = a_kj
            para  i = 1,...,k-1
                t = t − a_ik a_ij
            fin-para  i
            a_kj = t/a_kk
        fin-para j
    fin-sino
fin-para k
```

Normas

Bolas

DIBUJOS

Funciones homógeneas

Punto de acumulacion o punto limite

clausura o cerradura

Conjuntos cerrados abierto frontera puntos aislados

ÍNDICE ANALÍTICO