0.1 Método de los tres puntos

Dados x, d en $\mathbb{R}^n$ fijos,

$$\min \varphi(t) = f(x+td) \\ t \in T$$

Donde $T = \mathbb{R}$ o $T = \mathbb{R}_+$

Supongamos que existe un minimizador t^*. Se busca construir una sucesión $\{t_k\}$ tal que

$$\begin{aligned} t_k &\in T \\ \lim_{k\to\infty} t_k &= t^* \\ \varphi(t_{k+1}) &< \varphi(t_k) \end{aligned}$$

Supongamos que $T = \mathbb{R}$. Sea $t_0 = 0$. En una iteración, dado t_k, hay tres pasos

1. Dado t_k , encontrar $\tilde{t}$ tal que

$$\begin{aligned} &|\tilde{t} - t_k| > \varepsilon \\ &\varphi(\tilde{t}) < \varphi(t_k) \end{aligned}$$

2. Dados t_k , $\tilde{t}$, calcular τ_1, τ_2, τ_3 tales que

$$\begin{aligned} &\tau_1 < \tau_2 < \tau_3 \quad \text{ó} \quad \tau_1 > \tau_2 > \tau_3 \\ &\varphi(\tau_2) < \varphi(t_k) \\ &\varphi(\tau_2) < \varphi(\tau_1), \quad \varphi(\tau_2) < \varphi(\tau_3) \end{aligned}$$

3. Dados τ_1, τ_2, τ_3, calcular t_p^* el minimizador de la parábola que pasa por

$$\begin{aligned} &(\tau_1, \varphi(\tau_1)),\ (\tau_2, \varphi(\tau_3)),\ (\tau_3, \varphi(\tau_3)) \\ &\varphi_i = \varphi(\tau_i) \\ &t_p^* = \tau_2 - \frac{1}{2}\,\frac{(\tau_2-\tau_1)^2(\varphi_2-\varphi_3) - (\tau_2-\tau_3)^2(\varphi_2-\varphi_1)}{(\tau_2-\tau_1)(\varphi_2-\varphi_3) - (\tau_2-\tau_3)(\varphi_2-\varphi_1)} \\ &t_{k+1} = \text{mejor}\{\tau_2,\ t_p^*\} \end{aligned}$$

Como se verá más adelante, si en el paso 1, no se puede encontrar $\tilde{t}$, entonces posiblemente t_k es una buena aproximación de t^*.

En el paso 2, si no se pueden encontrar los tres τ_i, posiblemente no hay minimizador.

Encontrar $\tilde{t}$ **mejor que** t_k

datos: f, x, d, t_k, $\varepsilon > 0$, h_{inic}, $|h_{inic}| > \varepsilon$
$h \leftarrow h_{inic}$, $\varphi_k \leftarrow \varphi(t_k)$
mientras $|h| > \varepsilon$
    si $\varphi(t_k + h) < \varphi_k$
        $\tilde{t} \leftarrow t_k + h$
        salir
    sino
        si $\varphi(t_k - h) < \varphi_k$
            $\tilde{t} \leftarrow t_k - h$
            salir
        fin-si
    fin-si
    $h \leftarrow h/2$
fin-mientras

En el esquema anterior y en el siguiente, se ha supuesto que se minimiza en $\mathbb{R}$. Es necesario modificar ligeramente ambos esquemas para minimizar únicamente en $\mathbb{R}_+$.

Ejemplo

$$
\begin{aligned}
& f(x_1, x_2) = 2x_1^4 + (x_1 + x_2 - 1)^2 \\
& x = (3, 4) \\
& d = (5, -1) \\
& \min f(x + td), \quad t \in \mathbb{R}_+
\end{aligned}
$$

```
x :   3  4
d :   -5  -1
Buscar  t  mejor
tk =   0
```

```
eps =  0.0010
h inic =  10
fi(tk) = 198
fi( 10.0000 ) = 9762278.0000
fi( 5.0000 ) = 469088.0000
fi( 2.5000 ) = 16371.1250
fi( 1.2500 ) = 225.3828
fi( 0.6250 ) = 5.0630
tmejor: fi(0.6250) = 5.0630
```

Ejemplo

$$
\begin{aligned}
&f(x_1, x_2) = 2x_1^4 + (x_1 + x_2 - 1)^2 \\
&x = (3, 4) \\
&d = (5, -1) \\
&\min f(x + td), \quad t \in \mathbb{R}_+
\end{aligned}
$$

```
Buscar  t  mejor
tk =  0.7500
eps =  0.0010
h inic =  0.1000
fi(tk) =  2.8828
fi( 0.8500 ) = 5.6928
fi( 0.6500 ) = 4.4178
fi( 0.8000 ) = 3.4400
fi( 0.7000 ) = 3.3650
fi( 0.7750 ) = 2.9949
fi( 0.7250 ) = 3.0277
fi( 0.7625 ) = 2.9022
fi( 0.7375 ) = 2.9274
fi( 0.7562 ) = 2.8840
fi( 0.7438 ) = 2.8977
fi( 0.7531 ) = 2.8813
```

Encontrar tres puntos

datos: f, x, d, t_k, $\tilde{t}$, $K >> 0$
$\tau_1 \leftarrow t_k, \quad \varphi_1 = \varphi(\tau_1)$
$\tau_2 \leftarrow \tilde{t}, \quad \varphi_2 = \varphi(\tau_2)$
$\tau_3 \leftarrow \tau_2$
mientras $|\tau_3| < K$
    $\tau_3 \leftarrow \tau_2 + 2(\tau_2 - \tau_1)$
    $\varphi_3 \leftarrow \varphi(\tau_3)$
    si $\varphi_3 > \varphi_2$
        salir
    sino
        $\tau_1 \leftarrow \tau_2, \quad \tau_2 \leftarrow \tau_3$
        $\varphi_1 \leftarrow \varphi_2, \quad \varphi_2 \leftarrow \varphi_3$
    fin-si
fin-mientras

Ejemplo

$$f(x_1, x_2) = 2x_1^4 + (x_1 + x_2 - 1)^2$$
$$x = (3, 4)$$
$$d = (5, -1)$$
$$\min f(x + td), \quad t \in \mathbb{R}_+$$

```
Buscar tres puntos
tk =   0
t mejor =  0.1000
fi( 0.0000 ) = 198.0000
fi( 0.1000 ) = 107.2850
fi( 0.3000 ) = 27.7650
fi( 0.7000 ) = 3.3650
fi( 1.5000 ) = 829.1250
      ti        fi(ti) :
      0.3000     27.7650
      0.7000      3.3650
      1.5000    829.1250
```

Ejemplo

$$
\begin{aligned}
&f(x_1, x_2) = x_1^2 - x_2 \\
&x = (2, 3) \\
&d = (0, 1) \\
&\min f(x + td), \quad t \in \mathbb{R}_+
\end{aligned}
$$

```
Buscar tres puntos
tk =   1
t mejor =  1.5000
fi( 1.0000 ) = 0.0000
fi( 1.5000 ) = -0.5000
fi( 2.5000 ) = -1.5000
fi( 4.5000 ) = -3.5000
fi( 8.5000 ) = -7.5000
fi( 16.5000 ) = -15.5000
fi( 32.5000 ) = -31.5000
fi( 64.5000 ) = -63.5000
fi( 128.5000 ) = -127.5000
fi( 256.5000 ) = -255.5000
fi( 512.5000 ) = -511.5000
fi( 1024.5000 ) = -1023.5000
t = 1024.500000  >  cota = 1000
```

Posiblemente no hay minimizador.

Ejemplo. Algoritmo completo.

$$
\begin{aligned}
&f(x_1, x_2) = 2x_1^4 + (x_1 + x_2 - 1)^2 \\
&x = (3, 4) \\
&d = (5, -1) \\
&\min f(x + td), \quad t \in \mathbb{R}_+
\end{aligned}
$$

```
x :    3   4
d :    -5   -1

tk  fi(tk) :   0  198
tm  fi(tm) :   0.1000  107.2850
```

```
Tres puntos: ti  fi(ti) :
      0.3000     27.7650
      0.7000      3.3650
      1.5000    829.1250
tp fi(tp) :   0.5335  7.8596

tk  fi(tk) :   0.7000  3.3650
tm  fi(tm) :   0.7875  3.1706
Tres puntos: ti  fi(ti) :
      0.7000      3.3650
      0.7875      3.1706
      0.9625     21.6351
tp fi(tp) :   0.7465  2.8893

tk  fi(tk) :   0.7465  2.8893
tm  fi(tm) :   0.7581  2.8875
Tres puntos: ti  fi(ti) :
      0.7465      2.8893
      0.7581      2.8875
      0.7813      3.0724
tp fi(tp) :   0.7526  2.8813

tk  fi(tk) :   0.7526  2.8813
No se obtuvo  t mejor.
```

Entonces $t^* \approx 0.7526$.

0.2 Encajonamiento. Búsqueda secuencial

Problema:

$$\min\ \varphi(t) = f(x + td), \quad t \in [a, b]$$

Si φ es continua, existe t^* minimizador global. Los métodos de encajonamiento crean una sucesión de intervalos buscando reducir el tamaño del intervalo. Sea $[a_0, b_0] = [a, b]$. Los intervalos deben cumplir:

$$t^* \in [a_k, b_k], \quad \forall k \tag{1}$$

$$[a_{k+1}, b_{k+1}] \subseteq [a_k, b_k] \tag{2}$$

$$\lim_{k\to\infty}(b_k - a_k) = 0. \tag{3}$$

En el método de **búsqueda secuencial**, se calculan $m \geq 2$ puntos intermedios igualmente espaciados en cada intervalo. Se busca el mejor punto intermedio y el nuevo intervalo irá desde el vecino izquierdo hasta el vecino derecho.

$$\begin{aligned} h &= \frac{b_k - a_k}{m+1} \\ t_i &= a_k + ih, \quad i = 1, ..., m \\ \tilde{t} &= \text{mejor}\{t_1, t_2, ..., t_m\} \\ [a_{k+1}, b_{k+1}] &= [\tilde{t} - h, \tilde{t} + h] \end{aligned}$$

El proceso termina cunado la longitud del intervalo es suficientemente pequeña.

Si φ es cuasiconvexa, entonces el método cumple las propiedades requeridas.

$$b_k - a_k = \left(\frac{2}{m+1}\right)^k (b_0 - a_0) \to 0$$

Si m es grande, el número de iteraciones necesarias es mucho más pequeño pero en cada iteración hay m utilzaciones de φ. Buscando minimizar el número de utilizaciones de φ, es mejor m pequeño, 2 o 3.

Si no se tiene certeza sobre la cuasiconvexidad de f, puede ser interesante usar m grande.

Ejemplo

$$
\begin{aligned}
&f(x_1, x_2) = 2x_1^4 + (x_1 + x_2 - 1)^2 \\
&x = (3, 4) \\
&d = (5, -1) \\
&\min f(x + td), \quad t \in [0, 1]
\end{aligned}
$$

a	b	t1	t2	fi1	fi2
0.	1.	0.333333	0.666667	22.320988	4.024691
0.333333	1.	0.555556	0.777778	7.115988	3.026368
0.555556	1.	0.703704	0.851852	3.305067	5.819227
0.555556	0.851852	0.654321	0.753086	4.312667	2.881313
0.654321	0.851852	0.720165	0.786008	3.079708	3.144898
0.654321	0.786008	0.698217	0.742112	3.394952	2.904061
0.698217	0.786008	0.727481	0.756744	3.003737	2.884771
0.727481	0.786008	0.74699	0.766499	2.888034	2.923449
0.727481	0.766499	0.740487	0.753493	2.91141	2.881413
0.740487	0.766499	0.749157	0.757828	2.883905	2.886918
0.740487	0.757828	0.746267	0.752048	2.889826	2.881374
0.746267	0.757828	0.750121	0.753975	2.882679	2.881623
0.750121	0.757828	0.75269	0.755259	2.881283	2.882669
0.750121	0.755259	0.751834	0.753546	2.881443	2.881432
0.751834	0.755259	0.752975	0.754117	2.881298	2.881704
0.751834	0.754117	0.752595	0.753356	2.881285	2.881372
0.751834	0.753356	0.752341	0.752849	2.881311	2.881287
0.752341	0.753356	0.752679	0.753018	2.881283	2.881303
0.752341	0.753018	0.752567	0.752792	2.881287	2.881284
0.752567	0.753018	0.752717	0.752867	2.881282	2.881288
0.752567	0.752867	0.752667	0.752767	2.881283	2.881283
0.752567	0.752767	0.752633	0.7527	2.881284	2.881282

0.3 Encajonamiento. Sección dorada

Problema:

$$\min \ \varphi(t) = f(x + td), \quad t \in [a, b]$$

El método de la sección dorada o **número áureo**, también es un método de encajonamiento. Sólamente hay dos puntos intermedios no igualmente espaciados pero simétricos.

$$\begin{aligned} s_k &= a_k + \alpha(b_k - a_k) \\ t_k &= b_k - \alpha(b_k - a_k) = a_k + (1 - \alpha)(b_k - a_k) \end{aligned}$$

Se busca que s_{k+1} coincida con t_k o que t_{k+1} coincida con s_{k-1}. De esta manera se efectúa únicamente una evaluación de φ por iteración.

$$\begin{aligned} &\text{Si } \varphi(s_k) < \varphi(t_k), \quad [a_{k+1}, b_{k+1}] = [a_k, t_k] \\ &\text{Si } \varphi(s_k) \geq \varphi(t_k), \quad [a_{k+1}, b_{k+1}] = [s_k, b_k] \end{aligned}$$

Supongamos que $\varphi(s_k) < \varphi(t_k)$

$$\begin{aligned} b_{k+1} - a_{k+1} &= (1 - \alpha)(b_k - a_k) \\ t_{k+1} &= s_k \\ a_{k+1} + (1 - \alpha)(b_{k+1} - a_{k+1}) &= a_k + \alpha(b_k = a_k) \\ a_k + (1 - \alpha)(1 - \alpha)(b_k - a_k) &= a_k + \alpha(b_k - a_k) \\ (1 - \alpha)^2 &= \alpha \\ \alpha &= \frac{3 - \sqrt{5}}{2} \approx 0.381966011 \end{aligned}$$

El valor de α es justamente 2 menos el número áureo.

datos: $a,\ b,\ \varepsilon$
$l = b - a, \quad \delta = \alpha l$
$s = a + \delta, \quad t = b - \delta$
$\texttt{fs} = \varphi(s), \quad \texttt{ft} = \varphi(t)$
mientras $l \geq \varepsilon$
    $l = (1 - \alpha)l, \quad \delta = \alpha l$
    si $\texttt{fs} \leq \texttt{ft}$ **ent**
        $b = t, \quad t = s, \quad s = a + \delta$
        $\texttt{ft} = \texttt{fs}, \quad \texttt{fs} = \varphi(s)$
    fin-ent
    sino
        $a = s, \quad s = t, \quad t = b - \delta$
        $\texttt{fs} = \texttt{ft}, \quad \texttt{ft} = \varphi(t)$
    fin-sino
fin-mientras
$\lambda^* = \text{mejor}\{a, s, t, b\}$

Ejemplo

$$f(x_1, x_2) = 2x_1^4 + (x_1 + x_2 - 1)^2$$
$$x = (3, 4)$$
$$d = (5, -1)$$
$$\min f(x + td), \quad t \in [0, 1]$$

a	b	s	t	fs	ft
0.000000	1.000000	0.381966	0.618034	16.575701	5.252461
0.381966	1.000000	0.618034	5.252461	0.763932	2.908956
0.618034	1.000000	0.763932	2.908956	0.854102	5.977557
0.618034	0.854102	0.708204	3.236567	0.763932	2.908956
0.708204	0.854102	0.763932	2.908956	0.798374	3.399255
0.708204	0.798374	0.742646	2.901865	0.763932	2.908956
0.708204	0.763932	0.729490	2.985765	0.742646	2.901865
0.729490	0.763932	0.742646	2.901865	0.750776	2.882063
0.742646	0.763932	0.750776	2.882063	0.755801	2.883325
0.742646	0.755801	0.747671	2.886534	0.750776	2.882063
0.747671	0.755801	0.750776	2.882063	0.752696	2.881282
0.750776	0.755801	0.752696	2.881282	0.753882	2.881575
0.750776	0.753882	0.751963	2.881399	0.752696	2.881282

0.751963	0.753882	0.752696	2.881282	0.753149	2.881324
0.751963	0.753149	0.752416	2.881300	0.752696	2.881282
0.752416	0.753149	0.752696	2.881282	0.752869	2.881288
0.752416	0.752869	0.752589	2.881285	0.752696	2.881282
0.752589	0.752869	0.752696	2.881282	0.752762	2.881283
0.752589	0.752762	0.752655	2.881283	0.752696	2.881282

0.4 Método de las tangentes paralelas

También es conocido con el nombre Partan (parallel tangents). Este método es una modificación del método del descenso más pendiente. Sirve para evitar o disminuir su zigzagueo.

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

Dado un y^j, por medio del descenso más pendiente se obtiene z^j. En este punto z^j se utiliza la dirección $\delta^j = z^j - y^{j-1}$ para obtener y^{j+1}. Así sucesivamente hasta obtener y^{n+1}. Con este punto se vuelve a empezar. Obviamente, en la primera subiteración, y^2 se obtiene a partir de y^1 por medio del decenso más pendiente.

Método de las tangentes paralelas

datos: x^1, ε_g, MAXIT
para $k = 1, ...,$MAXIT
    $y^1 = x^k$
    si $||f'(y^1)|| \leq \varepsilon_g$ **parar**
    $d^1 = -f'(y^1)$
    $t_1 = \text{argmin}\ \ f(y^1 + td),\ \ t \geq 0$
    $y^2 = y^1 + t_1 d^1$
    para $j = 2, ..., n$
        si $||f'(y^j)|| \leq \varepsilon_g$ **parar**
        $d^j = -f'(y^j)$
        $t_j = \text{argmin}\ \ f(y^j + td),\ \ t \geq 0$
        $z^j = y^j + t_j d$
        $\delta^j = z^j - y^{j-1}$
        $\mu_j = \text{argmin}\ \ f(z^j + \mu\delta^j),\ \ \mu \in \mathbb{R}$
        $y^{j+1} = z^j + \mu_j \delta^j$
    fin-para
    $x^{k+1} = y^{n+1}$
fin-para

Durante el desarrollo del algoritmo basta con tener en memoria los vectores x, x^{ant}, y, y^{ant}, d, z.

Ejemplo Aplicar el método de tangentes paralelas para minimizar $f(x) = 0.1x_1^2 + x_2^2 + 10x_3^2 + 100x_4^2 - 0.2x_1 - 2x_2 - 20x_3 - 200x_4$, a partir del punto $x^1 = (2, 3, 4, 5)$.

j	y^j	$f'(y^j)$	d^j	t_j	z^j	δ^j	μ_j
1	2	0.2	-0.2	0.005025			
	3	4	-4				
	4	60	-60				
	5	800	-800				
2	1.9990	0.1998	-0.1998	0.047812	1.9894	-0.0106	0.045515
	2.9799	3.9598	-3.9598		2.7906	-0.2094	
	3.6985	53.9695	-53.9695		1.1181	-2.8819	
	0.9797	-4.0676	4.0676		1.1741	-3.8259	
3	1.9890	0.1978	-0.1978	0.478224	1.8944	-0.1046	0.043371
	2.7810	3.5621	-3.5621		1.0776	-1.9023	
	0.9869	-0.2620	0.2620		1.1122	-2.5863	
	1.0000	0.0018	-0.0018		0.9992	0.0195	
4	1.8898	0.1780	-0.1780	4.865467	1.0239	-0.9650	0.024810
	0.9951	-0.0099	0.0099		1.0431	-1.7379	
	1.0000	0.0001	-0.0001		0.9997	0.0128	
	1.0000	0.0000	0.0000		1.0000	0.0000	
5	1.0000						
	1.0000						
	1.0000						
	1.0000						

Sigue una nueva iteración:

$$x^2 = y^5, \qquad y^1 = x^2, \qquad f'(y^1) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Un resultado interesante y útil es el siguiente (ver la demostración en [Lue89]):

Para funciones cuadráticas estrictamente convexas, el método de tangentes paralelas es equivalente al método del gradiente conjugado. En particular, en ausencia de errores de redondeo y truncamiento, se obtiene la solución en no más de n subiteraciones.

Por ejemplo, al aplicar el método del gradiente conjugado a la función anterior se obtienen los siguientes resultados parciales

j	y^j	$f'(y^j)$	α_j	d^j	t_j
1	2	0.2		-0.2	0.005025
	3	4		-4	
	4	60		-60	
	5	800		-800	
2	1.9990	0.1998	0.004576	-0.2007	0.049988
	2.9799	3.9598		-3.9781	
	3.6985	53.9695		-54.2440	
	0.9797	-4.0676		0.4070	
3	1.9890	0.1978	0.004345	-0.1987	0.498965
	2.7810	3.5621		-3.5794	
	0.9869	-0.2620		0.0263	
	1.0000	0.0018		-0.0000	
4	1.8898	0.1780	0.002483	-0.1785	4.986180
	0.9951	-0.0099		0.0010	
	1.0000	0.0001		0.0000	
	1.0000	0.0000		0.0000	
5	1.0000				
	1.0000				
	1.0000				
	1.0000				

$$x^2 = y^5, \qquad y^1 = x^2, \qquad f'(y^1) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

0.5 Modificación del método de Newton

El método de Newton, a partir de un punto inicial x^1, busca un punto crítico, es decir, un punto x^K tal que $f'(x^K) \approx 0$. Esto hace que para ciertas funciones este método no sea un método de minimización y entonces es posible que el punto crítico obtenido sea peor que el punto inicial, o sea,

$$\begin{aligned} f'(x^K) &\approx 0, \\ f(x^K) &> f(x^1). \end{aligned}$$

Por ejemplo, al tratar de minimizar

$$\begin{aligned} f(x_1, x_2) =& (x_1 + 1.5)(x_1 + 0.5)(x_1 - 0.5) + (x_2 - 0.5)(x_2 - 1.5)(x_2 - 2.5) \\ &+ 0.3x_1x_2 + 0.01(x_1 - 3)^4 + 0.01(x_2 - 4)^4 \end{aligned}$$

por el método de Newton a partir de $x^1 = (-2.5, 1)$, se obtiene:

k	x^k	$f(x^k)$	$f'(x^k)$	$f''(x^k)$		d^k
1	-2.5	3.5856	4.6450	-8.3700	0.3000	0.5190
	1.0		-2.0800	0.3000	-1.9200	-1.0022
2	-1.9810	5.0674	0.6360	-5.9086	0.3000	0.1266
	-0.0022		2.6115	0.3000	-7.0913	0.3736
3	-1.8544	5.6411	0.0386	-5.2984	0.3000	0.0112
	0.3714		0.3538	0.3000	-5.1916	0.0688
4	-1.8432	5.6538	0.0003	-5.2443	0.3000	0.0002
	0.4402		0.0122	0.3000	-4.8381	0.0025
5	-1.8430	5.6538	0.0000			
	0.4427		0.0000			

La modificación busca convertir el método en uno de descenso, es decir, siempre se debe tener que

$$f(x^{k+1}) < f(x^k).$$

En el método de Newton se obtiene d^k solución de $f''(x^k)\, d^k = -f'(x^k)$. Si $f(x^k + d^k) \geq f(x^k)$, entonces se averigua si d^k es dirección de descenso.

Si lo es, se obtiene t_k minimizando $f(x^k + td^k)$ para $t \geq 0$ y se construye $x^{k+1} = x^k + t_k d^k$.

Si d^k no es dirección de descenso, en lugar de utilizar la matriz hessiana se utiliza

$$M = f''(x^k) + \lambda I, \tag{4}$$
$$Md^k = -f'(x^k). \tag{5}$$

Es claro que para $\lambda > 0$ suficientemente grande la matriz M es de diagonal positiva estrictamente dominante y, por lo tanto, definida positiva.

Recordemos que una matriz A simétrica es definida positiva si y solamente si su inversa existe y es definida positiva (los valores propios de A^{-1} son los inversos multiplicativos de los valores propios de A.)

La dirección d^k así obtenida es de descenso si

$$\begin{aligned} f'(x^k)^{\mathrm{T}} d^k &< 0, \\ f'(x^k)^{\mathrm{T}}(-M^{-1} f'(x^k)) &< 0, \\ -f'(x^k)^{\mathrm{T}} M^{-1} f'(x^k) &< 0. \end{aligned}$$

En consecuencia, si M es definida positiva, entonces d^k, obtenida según (**??**), es dirección de descenso.

Se podría pensar que la mejor estrategia es tomar inicialmente un valor λ grande para garantizar que M sea definida positiva. Esto hace que desde el principio M se aleje del hessiano, o sea, el método se aleja del método de Newton, perdiendo desde el comienzo las posibles ventajas del método de Newton, en particular la convergencia cuadrática.

Además, para valores grandes de λ, se tiene que $M \approx \lambda I$, o sea el método se aproximaría al método del descenso más pendiente y a su lentitud debida al zigzagueo.

Por eso se empieza con $M = f''(x^k)$ y, cuando no se tiene una dirección de descenso, se modifica M.

El siguiente valor de λ garantiza que $M = H + \bar{\lambda} I$ es definida positiva ($H = f''(x^k)$).

$$
\begin{aligned}
\sigma_i &= \sum_{\substack{j=1 \\ j \neq i}}^{n} |h_{ij}|, \quad i = 1, ..., n \\
\sigma &= \max_{1 \le i \le n} \{\sigma_i - h_{ii}\} \\
\lambda' &> \max\{\sigma, 0\}
\end{aligned}
$$

Si $\sigma < 0$, entonces H es definida positiva y no necesita modificaciones.

Método de Newton modificado

```
datos: x^1, ε_g, MAXIT
para k = 1,...,MAXIT
    si ||f'(x^k)|| ≤ ε_g parar
    λ = 0
    fink = 0
    mientras fink = 0
        M = f''(x^k) + λI
        resolver M d^k = −f'(x^k)
        si f(x^k + d^k) < f(x^k)
            x^{k+1} = x^k + d^k
            fink = 1
        sino
            si f'(x^k)^T d^k < 0
                t_k = argmin f(x^k + t d^k),  t > 0
                x^{k+1} = x^k + t_k d^k
                fink = 1
            sino
                incrementar λ
            fin-si
        fin-si
    fin-mientras
fin-para
```

La terminación esperada del algoritmo anterior se tiene cuando $||f'(x^k)|| \le \varepsilon_g$. Pero hay otras terminaciones posibles:

- demasiadas iteraciones
- $f(x^k + td^k)$ no tiene minimizador (no está acotada inferioremente).
- M no es invertible. En este caso se puede incrementar λ.

Ejemplo Minimizar

$$\begin{aligned} f(x_1, x_2) =&(x_1 + 1.5)(x_1 + 0.5)(x_1 - 0.5) + (x_2 - 0.5)(x_2 - 1.5)(x_2 - 2.5) \\ &+ 0.3x_1x_2 + 0.01(x_1 - 3)^4 + 0.01(x_2 - 4)^4 \end{aligned}$$

por el método de Newton modificado a partir de $x^1 = (-2.5, 1)$.

k	x^k	$f(x^k)$	$f'(x^k)$	$f''(x^k)$		λ'
1	-2.5	3.5856	4.6450	-8.3700	0.3000	8.6700
	1.0		-2.0800	0.3000	-1.9200	

λ	M		d	$x^k + d$	$f(x^k + d)$	$f'(x^k)^{\mathrm{T}}\, d$	t_k
0	-8.37	0.30	0.5190	-1.9810	5.0674	4.4956	
	0.30	-1.92	1.0022	-0.0022			
2.89	-5.48	0.30	0.9490	-1.5510	3.9849	0.5581	
	0.30	0.97	1.8508	2.8508			
5.78	-2.59	0.30	1.8393	-0.6607	2.2385		
	0.30	3.86	0.3559	1.3959			

k	x^k	$f(x^k)$	$f'(x^k)$	$f''(x^k)$		λ'
2	-0.6607	2.2385	-2.4660	0.6439	0.3000	0.1108
	1.3959		-1.8721	0.3000	0.1892	

λ	M		d	$x^k + d$	$f(x^k + d)$	$f'(x^k)^{\mathrm{T}}\, d$	t_k
0	0.6439	0.3000	-2.9857	-3.6464	3232.0679	-20.0217	0.047556
	0.3000	0.1892	14.6278	16.0237			

k	x^k	$f(x^k)$	$f'(x^k)$	$f''(x^k)$		λ'
3	-0.8027	1.6105	-2.2972	-0.0809	0.3000	0.3809
	2.0915		-0.4691	0.3000	3.9863	

λ	M	d	$x^k + d$	$f(x^k + d)$	$f'(x^k)^{\mathrm{T}} d$	t_k
0	-0.0809 0.3000	-21.8620	-22.6647	-6543.7093		
	0.3000 3.9863	1.7629	3.8545			

Así se obtiene $x^4 = \begin{bmatrix}-22.6647 & 3.8545\end{bmatrix}^{\mathrm{T}}$. Después de otras iteraciones

k	x^k	$f(x^k)$	$f'(x^k)$	$f''(x^k)$	λ'
6	-64.4160	-54550.7914	-0.0018	161.8936 0.3000	0
	4.1038		0.0140	0.3000 15.6239	

λ	M	d	$x^k + d$	$f(x^k + d)$	$f'(x^k)^{\mathrm{T}} d$	t_k
0	161.8936 0.3000	0.0000	-64.4159	-54550.7914		
	0.3000 15.6239	-0.0009	4.1029			

k	x^k	$f(x^k)$	$f'(x^k)$	$f''(x^k)$	λ'
7	-64.4159	-54550.7914	-0.0000		
	4.1029		0.0000		