0.1 Método del simplejo

También es conocido con el nombre de método del simplex, traducido por simplejo o símplice. Es diferente del método simplex en Optimización Lineal, aunque los dos usan el mismo concepto geométrico pero de diferente manera. Es un método muy popular, particularmente entre los químicos. No requiere derivadas.

Hay varias versiones, las más importantes son la de Spendley Hext y Himsworth SHH[62] que utiliza simplejos equiláteros y la de Nelder y Mead [NeMe65] que va modificando la forma de los simplejos.

Para dos variables su funcionamiento es el siguiente: dado un triángulo equilátero inicial se descarta el peor vértice (el de mayor valor de la función) y se construye un nuevo triángulo equilátero utilizando los otros dos vértices y el punto simétrico al peor vértice con respecto al punto medio del segmento formado por los dos vértices mejores. Se continua este proceso hasta obtener una aproximación de un minimizador. Cuando el punto simétrico obtenido sería el peor vértice del nuevo triángulo, entonces se descarta el segundo peor. Cuando en este proceso, no hay avance, hay dos posibilidades. La primera, la buena, indica que posiblemente hay un mimizador cerca. La segunda, indica que el simplejo está en una zona difícil. De todas maneras, se construye un triángulo equilátero más pequeño y se empieza de nuevo.

Definición 0.1. Dados $m+1$ puntos x^0, x^1, ..., x^m en $\mathbb{R}^n$ tales que $x^1 - x^0$, $x^2 - x^0$, ..., $x^m - x^0$ son linealmente independientes, se llama **m-simplejo** al convexo generado por los $m+1$ puntos. Un simplejo es un n-simplejo. Se puede mostrar que los $m+1$ puntos son los puntos extremos o vértices del m-simplejo.

En $\mathbb{R}^3$ un 1-simplejo es un segmento de recta, un 2-simplejo es un triángulo, un simplejo o 3-simplejo es un tetraedro.

Definición 0.2. Dado un conjunto finito $A \subseteq \mathbb{R}^n$, formado por t puntos x^1, x^2, ..., x^t, se llama **baricentro** de A al punto "promedio" de los puntos,

$$\mathrm{bar}(A) = \frac{1}{t} \sum_{i=1}^{t} x^i.$$

0.1.1 Construcción de un simplejo equilátero

Dado un punto x^0 y $h > 0$, se trata de encontrar otros n puntos de tal forma que x^0, x^1, ..., x^n determinen un simplejo equilátero con arista de longitud h, es decir

$$x^1 - x^0, x^2 - x^0, ..., x^n - x^0$$

deben ser linealmente independientes y

$$||x^i - x^j||_2 = h \quad \forall i \neq j.$$

Basta con buscar un simplejo de arista h y después hacer una traslación.

Sean

$$y^i = h'e^i \ , \quad i = 1, ..., n, \tag{1}$$

con

$$h' = \frac{h\sqrt{2}}{2}, \tag{2}$$

y donde los puntos e^i son los vectores de la base canónica. Estos n puntos son equidistantes entre sí ya que la distancia entra cada pareja de puntos es h. Por ejemplo, $y^1 = (h', 0, 0, ..., 0)$, $y^2 = (0, h', 0, ..., 0)$,

$$||y^1 - y^2||_2 = \sqrt{h'^2 + h'^2 + 0 + ... + 0} = \sqrt{2}\ h' = \sqrt{2}\ \frac{h\sqrt{2}}{2} = h.$$

Sea $e = \dfrac{1}{n}\ [1 \ \ 1 \ \ \ldots \ \ 1]^{\mathrm{T}}$. Entonces $\mathrm{bar}(\{e^1, ..., e^n\}) = e$ y $\mathrm{bar}(\{y^1, ..., y^n\}) = h'e$.

Cualquier punto de la recta que pasa por $h'e$ y por el origen equidista de estos n puntos. Entonces basta con encontrar un punto de la forma

$$y^0 = (\alpha, \alpha, ..., \alpha), \tag{3}$$

tal que

$$||y^0 - y^1||_2 = h.$$

Entonces:

$$
\begin{aligned}
\|y^0 - y^1\|_2^2 &= h^2 \\
\left(\alpha - \frac{h\sqrt{2}}{2}\right)^2 + \alpha^2 \ + ... + \alpha^2 &= h^2 \\
\alpha^2 - \alpha h\sqrt{2} + h^2/2 + (n-1)\alpha^2 &= h^2 \\
n\alpha^2 - \alpha h\sqrt{2} - h^2/2 &= 0.
\end{aligned}
$$

Hay dos soluciones para α. También hay dos posibilidades para y^0, estar de un lado o del otro del hiperplano determinado por y^1, y^2, ..., y^n. Si y^0 está del mismo lado del hiperplano donde está el origen, es necesario tomar α negativo,

$$\alpha = \frac{\sqrt{2}h}{2n}\left(1 - \sqrt{1+n}\right). \tag{4}$$

La translación necesaria es:

$$
\begin{aligned}
d &= x^0 - y^0 && (5)\\
x^i &= y^i + d\ , \quad i = 0, 1, ..., n. && (6)
\end{aligned}
$$

En resumen, dados x^0 y $h > 0$, los cálculos se pueden hacer en el siguiente orden:

- calcular h' según (2)
- calcular α según (4)
- calcular y^i, $i = 1, ..., n$ según (1)
- calcular y^0 según (3)
- calcular d según (5)
- calcular x^i, $i = 0, 1, ..., n$ según (6)

Ejemplo 0.1. Construir un simplejo equilátero en $\mathbb{R}^3$, de arista 0.2, tal que uno de sus vértices sea $(4, 5, 6)$.

$$\begin{aligned}
h' &= 0.1414214, \\
\alpha &= -0.0471405, \\
y^0 &= (-0.0471405,\ \ -0.0471405,\ \ -0.0471405), \\
y^1 &= (0.1414214,\ \ 0,\ \ 0), \\
y^2 &= (0,\ 0.1414214,\ \ 0), \\
y^3 &= (0,\ 0,\ 0.1414214), \\
d &= (-0.0471405,\ -0.0471405,\ -0.0471405) \\
x^0 &= (4,\ 5,\ 6), \\
x^1 &= (4.1885618,\ 5.0471405,\ 6.0471405), \\
x^2 &= (4.0471405,\ 5.1885618,\ 6.0471405), \\
x^3 &= (4.0471405,\ 5.0471405,\ 6.1885618).
\end{aligned}$$

0.1.2 Algoritmo simplificado

Únicamente utiliza simplejos equiláteros. Supongamos que en una iteracion cualquiera se han utilizado o creado, hasta ese momento, los puntos x^0, x^1, x^2, ..., x^m. De estos $m+1$ puntos, los puntos x^{σ_0}, x^{σ_1}, x^{σ_2}, ..., x^{σ_n} determinan el último simplejo. Más aún, supongamos que en todas las iteraciones σ indica el orden de mejor a peor, es decir

$$f(x^{\sigma_0}) \leq f(x^{\sigma_1}) \leq \cdots \leq f(x^{\sigma_n}). \tag{7}$$

El punto $x^p = x^{\sigma_n}$ es el peor para la función f (la función que se desea minimizar). Se desea entonces, obtener el simétrico de x^p con respecto al baricentro de los otros puntos. Sea y el baricentro de los otros n puntos. El simétrico de x^p es:

$$\begin{aligned}
\tilde{x} &= y + (y - x^p) \qquad (8)\\
&= 2y - x^p\\
&= \frac{2}{n}\sum_{i=0}^{n-1} x^{\sigma_i} - x^p\\
&= \frac{2}{n}\sum_{i=0}^{n-1} x^{\sigma_i} + (\frac{2}{n} - \frac{2}{n} - 1)x^p\\
\tilde{x} &= \frac{2}{n}\sum_{i=0}^{n} x^{\sigma_i} - (\frac{2}{n} + 1)x^p.
\end{aligned}$$

Si denotamos el baricentro de todos los $n+1$ puntos por $\bar{x}$, entonces el simétrico de x^p es simplemente

$$\tilde{x} = \frac{2}{n}\bar{x} - (\frac{2}{n} + 1)x^p. \qquad (9)$$

En la nueva iteración, si $\tilde{x}$ es adecuado (más adelante se precisará), entra a hacer parte de los puntos que determinan el simplejo,

$$x^{m+1} = \tilde{x}.$$

El punto $\tilde{x}$ es adecuado si en el nuevo simplejo $\tilde{x}$ no es el peor vértice, o sea, si

$$f(\tilde{x}) < f(x^{\sigma_{n-1}}).$$

En la mayoría de los casos se halla $\tilde{x}$, el simétrico del peor punto, o sea, de x^{σ_n}. Cuando $\tilde{x}$ no es adecuado se halla el simétrico de $x^{\sigma_{n-1}}$. Si este simétrico no es adecuado, se calcula el simétrico de $x^{\sigma_{n-2}}$. Así sucesivamente hasta el simétrico de x^{σ_0}. Si este último tampoco es adecuado, es necesario construir un nuevo simplejo de tamaño menor, a partir del mejor vértice del último simplejo.

La cosnstrucción de un nuevo simplejo de arista más pequeña, a partir de x^{σ_0}, se puede hacer mediante las fórmulas de la sección 0.1.1 o bien guardando las mismas direcciones entre el mejor vértice y los demás.

De manera más precisa, si el último vértice construido es x^m y el mejor vértice es x^{σ_0} y se desea obtener un simplejo tal que el tamaño de las aristas sea ρ veces el tamaño de la anteriores (por ejemplo, $\rho = 1/2$), entonces

$$x^{m+i} = x^{\sigma_0} + \rho(x^{\sigma_i} - x^{\sigma_0}) \ , \quad i = 1, ..., n. \tag{10}$$

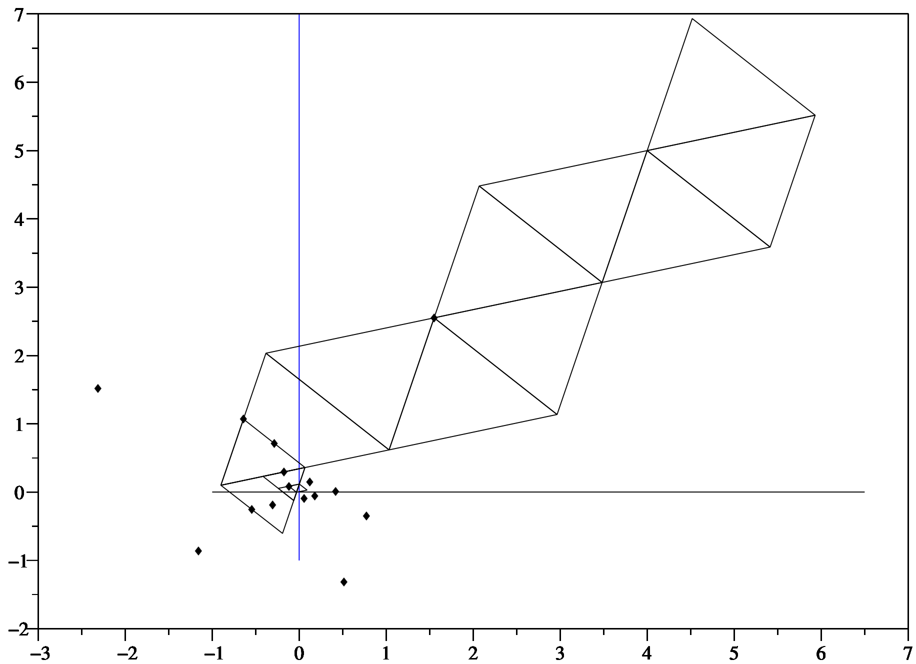
El nuevo simplejo está formado por los vértices

$$x^{\sigma_0},\ x^{m+1},\ x^{m+2},\ ...,\ x^{m+n}$$

Ejemplo 0.2. Minimizar $f(x_1, x_2) = x_1^2 + x_2^2$ a partir de $(4, 5)$, con $h_0 = 2$.

k	x_1^k	x_2^k	$f(x^k)$	σ	$\tilde{x}_1$	$\tilde{x}_2$	$f(\tilde{x})$
0	4.0000	5.0000	41.0000				
1	5.9319	5.5176	65.6312				
2	4.5176	6.9319	68.4596	$0, 1, 2$	5.4142	3.5858	42.1716
3	5.4142	3.5858	42.1716	$0, 3, 1$	3.4824	3.0681	21.5404
4	3.4824	3.0681	21.5404	$4, 0, 3$	2.0681	4.4824	24.3688
5	2.0681	4.4824	24.3688	$4, 5, 0$	1.5505	2.5505	8.9092
6	1.5505	2.5505	8.9092	$6, 4, 5$	2.9647	1.1363	10.0808
7	2.9647	1.1363	10.0808	$6, 7, 4$	1.0329	0.6187	1.4496
8	1.0329	0.6187	1.4496	$8, 6, 7$	-0.3813	2.0329	4.2780
9	-0.3813	2.0329	4.2780	$8, 9, 6$	-0.8990	0.1010	0.8184
10	-0.8990	0.1010	0.8184	$10, 8, 9$	0.5152	-1.3132	1.9899
					-2.3132	1.5152	7.6468
					1.5505	2.5505	8.9092
11	0.0669	0.3598	0.1340				
12	-0.6402	1.0669	1.5482	$11, 10, 12$	-0.1919	-0.6061	0.4042
13	-0.1919	-0.6061	0.4042	$11, 13, 10$	0.7741	-0.3473	0.7198
					-0.6402	1.0669	1.5482
					-1.1578	-0.8649	2.0886
14	-0.0625	-0.1231	0.0191				
15	-0.4160	0.2304	0.2262	$14, 11, 15$	0.4205	0.0063	0.1769
$\vdots$							

Después de la iteración 10 se construyó un nuevo simplejo de arista 1.0. Después de la iteración 13 se construyó un nuevo simplejo de arista 0.5.



0.2 Método de Nelder y Mead

Esta versión del método de Nelder y Mead es una adaptación de la presentada en [Kel99].
Además de usar triángulos no necesariamente equiláteros, utiliza varias clases de puntos: ξ^r el punto reflejado (el simétrico con respecto al baricentro de los otros puntos), ξ^e el punto producido por una expansión, ξ^c el punto producido por una contracción externa, y ξ^i el punto producido por una contracción interna.
La expresión (8) se puede reescribir como

$$\xi = y + \mu(y - x^p),$$

con $\mu = \mu_r = 1$. El subíndice r indica que se trata de una reflexión. El punto obtenido se denotará ξ^r. Para otros valores de μ, los puntos obtenidos corresponden a desplazamientos a partir de y en la dirección $y - x^p$. Si $\mu_e > 1$ se tiene ξ^e, una expansión. Si $0 < \mu_c < 1$ se tiene ξ^c, una contracción externa (ξ^c queda fuera del simplejo). Si $-1 < \mu_i < 0$ se tiene un contracción interna (ξ^i queda dentro del simplejo).

$$-1 < \mu_i < 0 < \mu_c < \mu_r = 1 < \mu_e.$$

Unos valores típicos son:

$$\begin{aligned} \mu_i &= -\frac{1}{2}, \\ \mu_c &= \frac{1}{2}, \\ \mu_r &= 1, \\ \mu_e &= 2. \end{aligned}$$

Para los valores de la función f se pueden utilizar los mismos subíndices:

$$\begin{aligned} f_i &= f(\xi^i), \\ f_c &= f(\xi^c), \\ f_r &= f(\xi^r), \\ f_e &= f(\xi^e). \end{aligned}$$

En cada iteración se usan los valores

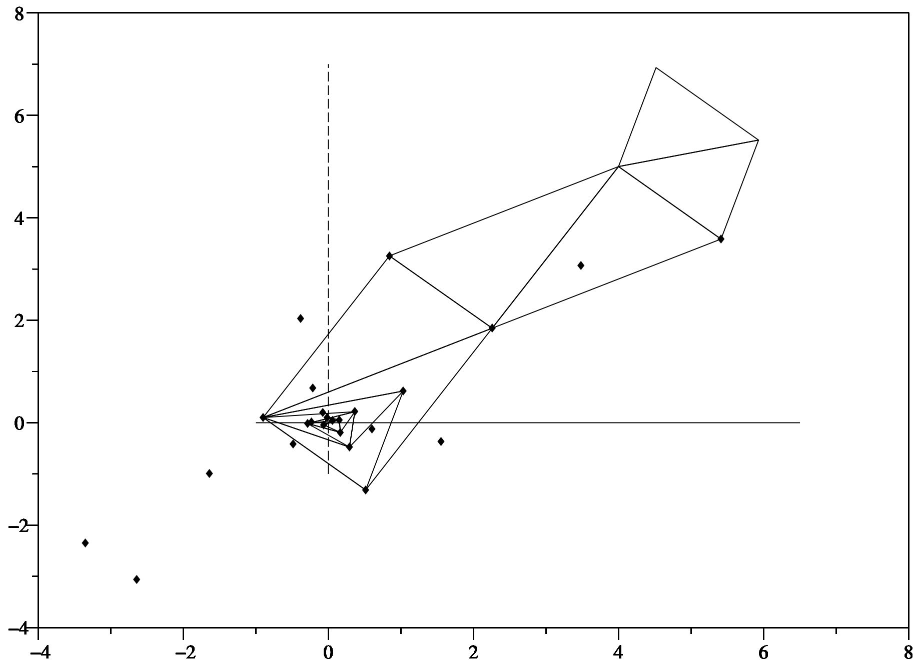
$$\begin{aligned} f_0 = f(x^{\sigma_0}) &= \text{ mejor valor de } f, \\ f_p = f(x^{\sigma_n}) &= \text{ peor valor de } f, \\ f_q = f(x^{\sigma_{n-1}}) &= \text{ segundo peor valor de } f. \end{aligned}$$

• Si $f_0 \le f_r < f_q$, entonces ξ^r es un buen punto y se convierte en el nuevo punto.
• Si $f_r < f_0$, el punto ξ^r es muy bueno y se hace una expansión (se calcula ξ^e). Se escoge como nuevo punto el mejor entre ξ^r y ξ^e.
• Si $f_q \le f_r < f_p$, se hace una contracción externa (se calcula ξ^c). Si $f_c \le f_r$, el nuevo punto es ξ^c. En caso contrario se reduce el tamaño del simplejo.
• Si $f_p \le f_r$, entonces se hace una contracción interna (se construye ξ^i). Si $f_i < f_p$, el nuevo punto es ξ^i. En caso contrario se reduce el tamaño del simplejo.

```
datos: : x^0, h_0, ε, maxit
h = h_0
construir x^1, ..., x^n
m = n
para k = 1, ..., maxit
    recalcular σ
    si f_p − f_0 ≤ ε parar
    calcular ξ^r
    si f_0 ≤ f_r < f_q
        m = m + 1,   x^m = ξ^r
        pasar a la iteración siguiente
    fin-si
    si f_r < f_0
        calcular ξ^e
        m = m + 1,   x^m = mejor{ξ^e, ξ^r}
        pasar a la iteración siguiente
    fin-si
    si f_q ≤ f_r < f_p
        calcular ξ^c
        si f_c ≤ f_r
            m = m + 1,   x^m = ξ^c
            pasar a la iteración siguiente
        sino
            reducir el simplejo
        fin-si
    fin-si
    si f_p ≤ f_r
        calcular ξ^i
        si f_i < f_p
            m = m + 1,  x^m = ξ^i
            pasar a la iteración siguiente
        sino
            reducir el simplejo
        fin-si
    fin-si
fin-para k
```

Ejemplo 0.3. Minimizar $f(x_1, x_2) = x_1^2 + x_2^2$ a partir de $(4, 5)$, con $h_0 = 2$.

	x_1	x_2	$f(x)$	σ
x^0	4.0000	5.0000	41.0000	
x^1	5.9319	5.5176	65.6312	
x^2	4.5176	6.9319	68.4596	$0, 1, 2$
ξ^r	5.4142	3.5858	42.1716	
x^3	5.4142	3.5858	42.1716	$0, 3, 1$
ξ^r	3.4824	3.0681	21.5404	
ξ^e	2.2576	1.8434	8.4950	
x^4	2.2576	1.8434	8.4950	$4, 0, 3$
ξ^r	0.8434	3.2576	11.3234	
x^5	0.8434	3.2576	11.3234	$4, 5, 0$
ξ^r	-0.8990	0.1010	0.8184	
ξ^e	-3.3485	-2.3485	16.7276	
x^6	-0.8990	0.1010	0.8184	$6, 4, 5$
ξ^r	0.5152	-1.3132	1.9899	
x^7	0.5152	-1.3132	1.9899	$6, 7, 4$
ξ^r	-2.6414	-3.0556	16.3133	
ξ^i	1.0329	0.6187	1.4496	
x^8	1.0329	0.6187	1.4496	$6, 8, 7$
ξ^r	-0.3813	2.0329	4.2780	
ξ^i	0.2911	-0.4767	0.3120	
x^9	0.2911	-0.4767	0.3120	$9, 6, 8$
ξ^r	-1.6408	-0.9943	3.6808	
ξ^i	0.3645	0.2154	0.1792	
x^{10}	0.3645	0.2154	0.1792	$10, 9, 6$



8
6
4
2
0
−2
−4
−4
−2
0
2
4
6
8

Bibliografía

[Kel99] Kelley C.T., *Iterative Methods for Optimization*, SIAM, Philadelphia, 1999.

[NeMe65] Nelder J.A., Mead R., *A simplex method for function minimization*, Comput. J., vol. 7 (1965), p. 308-313

[SHH62] Spendley W.G., Hext G.R., Himsworth F.R., *Sequential application of simplex designs in optimization and evolutionary operation*, Technometrics, vol 4, 1962, p. 441-461.