Práctica 3
Métodos de búsqueda para funciones de varias variables

Método de Hooke & Jeeves

Este método data de 1961 y consiste en tres tipos de procedimientos: exploración, cambio de punto base y movimiento de patrón.

Algoritmo de Hooke & Jeeves

1. Elegir un punto inicial $\mathbf{x}$, un valor ϵ de control de fin de algoritmo, una longitud de paso $h \geq \epsilon$ y un factor de aceleración α.

2. **Etapa de Exploración**. Al comienzo de este etapa $\mathbf{x}$ es el punto base. Sea $\overline{\mathbf{x}} = \mathbf{x}$ y $j = 1$. Calcular $f(\overline{\mathbf{x}})$

   a) Evaluar $f(\overline{\mathbf{x}} + h\mathbf{e^j})$, donde $\mathbf{e^j}$ es el vector unitario en la dirección de la variable x_j.

   - Si $f(\overline{\mathbf{x}} + h\mathbf{e^j}) < f(\overline{\mathbf{x}})$, entonces hacer $\overline{\mathbf{x}} = \overline{\mathbf{x}} + h\mathbf{e^j}$.

- En caso contrario, evaluar $f(\overline{\mathbf{x}} - h\mathbf{e^j})$. Si $f(\overline{\mathbf{x}} - h\mathbf{e^j}) < f(\overline{\mathbf{x}})$, entonces hacer $\overline{\mathbf{x}} = \overline{\mathbf{x}} - h\mathbf{e^j}$.

b) Si $j = n$, ir al paso 3 ($\overline{\mathbf{x}}$ es el punto obtenido al final de la etapa de exploración). En caso contrario, hacer $j = j + 1$ y volver al paso 2a).

3. Si $\overline{\mathbf{x}} = \mathbf{x}$, no se ha conseguido reducir el valor de $f(\mathbf{x})$ en la etapa de exploración.

   - Si $h < \epsilon$, PARAR. El punto $\mathbf{x}$ es la aproximación al mínimo de la función.
   - En caso contrario, hacer $h = \frac{h}{10}$ y volver al paso 2 (se reinicia la etapa de exploración con la nueva longitud de paso).

   Si $\overline{\mathbf{x}} \neq \mathbf{x}$ ($f(\overline{\mathbf{x}}) < f(\mathbf{x})$), ir al paso 4.

4. **Movimiento de patrón**. Puesto que hemos mejorado al movernos de $\mathbf{x}$ a $\overline{\mathbf{x}}$, realizamos un nuevo movimiento siguiendo este patrón. Sea

$$\mathbf{y} = \mathbf{x} + \alpha(\overline{\mathbf{x}} - \mathbf{x})$$

   Realizamos una etapa de exploración sobre el punto $\mathbf{y}$, obteniendo un nuevo punto $\overline{\mathbf{y}}$.

   - Si $f(\overline{\mathbf{y}}) > f(\overline{\mathbf{x}})$, entonces el movimiento de patrón ha empeorado el mejor valor obtenido, hacemos un **cambio de punto base**, tomando $\mathbf{x} = \overline{\mathbf{x}}$ y volvemos al paso 2 (obsérvese que esta opción conduce a realizar una nueva etapa de exploración en la que el punto base es $\overline{\mathbf{x}}$).
   - En caso contrario, hacer $\mathbf{x} = \mathbf{y}$, $\overline{\mathbf{x}} = \overline{\mathbf{y}}$ y volver al paso 3 (en este caso, el movimiento de patrón de $\mathbf{x}$ a $\mathbf{y}$ ha sido satisfactorio y en el paso 3 decidimos si hacemos un nuevo movimiento de patrón o volvemos a la etapa de exploración).

Método de Nelder & Mead

Este método, ampliamente utilizado, data de 1965 y se debe a J.A. Nelder y R. Mead. Es una modificación del método simplex introducido por Spendley, Hext y Himsworth en 1962, llamado así por utilizar el concepto de simplex que es un polítopo de $n+1$ vértices en un espacio de dimensión n (un segmento de recta en un dimensión 1, un triángulo en el plano, un tetraedro en el espacio tridimensional, etc).

En el método original de 1962 el simplex inicial debía ser regular (todos los puntos equidistantes entre sí), y se mantenía regular en las siguientes iteraciones. Las modificaciones propuestar por Nelder y Mead permiten que el simplex no sea regular y proporcionan un método de búsqueda muy potente en el caso de funciones de no más de cinco o seis variables.

Como veremos a continuación, el método se basa en tres tipos de operaciones: reflexión, expansión y contracción.

Algoritmo de Nelder & Mead

1. Consideremos $\mathbf{x_1} \in \mathbb{R}^n$ y n direcciones $\mathbf{d_1}, \mathbf{d_2}, \ldots, \mathbf{d_n}$. Entonces, $\{\mathbf{x_1}, \mathbf{x_2} = \mathbf{x_1} + \Delta\mathbf{d_1}, \ldots, \mathbf{x_{n+1}} = \mathbf{x_1} + \Delta\mathbf{d_n}\}$ forman un simplex. Tomamos $\epsilon > 0$ para el criterio de parada del test de convergencia del paso 10.

2. Evaluar la función f en todos los vértices del simplex, $\{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_{n+1}}\}$.

3. Sean $\mathbf{x_{max}}$, $\mathbf{x_{2max}}$, y $\mathbf{x_{min}}$ los vértices en donde se alcanza el valor máximo de la función, el siguiente valor máximo y el valor mínimo. Es decir,

$$f(\mathbf{x_{max}}) = \max_{1 \le i \le n+1} f(\mathbf{x_i}) = f_{max}$$

$$f(\mathbf{x_{2max}}) = \max_{\substack{1\le i\le n+1\\ \mathbf{x_i}\neq\mathbf{x_{max}}}} f(\mathbf{x_i}) = f_{2max}$$

$$f(\mathbf{x_{min}}) = \min_{1\le i\le n+1} f(\mathbf{x_i}) = f_{min}$$

4. Eliminar $\mathbf{x_{max}}$ y calcular $\mathbf{x_0}$ el baricentro de los vértices restantes

$$\mathbf{x_0} = \frac{1}{n}\sum_{\mathbf{x_i}\neq\mathbf{x_{max}}} \mathbf{x_i}.$$

5. **Reflexión**. Parece razonable alejarse de $\mathbf{x_{max}}$. Tomamos $\mathbf{x_r} = \mathbf{x_0} + \alpha(\mathbf{x_0} - \mathbf{x_{max}})$, donde $\alpha > 0$ es el factor de reflejo (usualmente $\alpha = 1$). $\mathbf{x_r}$ es el reflejo de $\mathbf{x_{max}}$ en $\mathbf{x_0}$.

6. Evaluar $f_r = f(\mathbf{x_r})$.

   - Si $f_r < f_{min}$, ir al paso 7.
   - Si $f_r \geq f_{min}$, pero $f_r < f_{2max}$, entonces $\mathbf{x_r}$ mejora los dos peores puntos del simplex. Reemplazar $\mathbf{x_{max}}$ por $\mathbf{x_r}$ e ir al paso 10.
   - Si $f_r < f_{max}$, pero $f_r \geq f_{2max}$, hacer $\mathbf{x_{max}} = \mathbf{x_r}$ y $f_{max} = f_r$. Ir al paso 8.
   - Si $f_r \geq f_{max}$, ir al paso 8.

7. **Expansión**. Parece adecuado moverse en la dirección de $\mathbf{x_0}$ a $\mathbf{x_r}$. Tomamos $\mathbf{x_e} = \mathbf{x_0} + \beta(\mathbf{x_r} - \mathbf{x_0})$, donde $\beta > 1$ es el factor de expansión. Calcular $f_e = f(x_e)$.

   - Si $f_e < f_{min}$, reemplazar $\mathbf{x_{max}}$ por $\mathbf{x_e}$. Ir al paso 10
   - Si $f_e \geq f_{min}$, rechazamos $\mathbf{x_e}$ (parece que nos hemos movido demasiado lejos en la dirección $\mathbf{x_0}$ a $\mathbf{x_r}$). Reemplazar $\mathbf{x_{max}}$ por $\mathbf{x_r}$ er ir al paso 10

8. **Contracción**. Puesto que $f_r > f_{max}$, parece que nos hemos desplazado demasiado en la dirección de $\mathbf{x_{max}}$ a $\mathbf{x_0}$. Intentaremos rectificar este movimiento contrayendo la longitud de paso. Consideremos

$$\mathbf{x_c} = \mathbf{x_0} + \gamma(\mathbf{x_{max}} - \mathbf{x_0}),$$

donde $0 < \gamma < 1$ es el factor de contracción. Sea $f_c = f(x_c)$.

- Si $f_c < f_{max}$, reemplazamos $\mathbf{x_{max}}$ por $\mathbf{x_c}$, ir al paso 12
- Si $f_c \geq f_{max}$, ir al paso 9.

9. Puesto que nuestros esfuerzos por encontrar un valor menor que f_{max} han fracasado, reducimos el tamaño del simplex generando nuevos vértices que se encuentran a la mitad de la distancia del punto $\mathbf{x_{min}}$ que sus antecesores. Por lo tanto, cada punto $\mathbf{x_i}$ es reemplazado por

$$\mathbf{x_{min}} + \frac{1}{2}(\mathbf{x_i} - \mathbf{x_{min}}) = \frac{1}{2}(\mathbf{x_i} + \mathbf{x_{min}}).$$

Calculamos los nuevos valores $f_i = f(\mathbf{x_i})$ y pasamos al paso 10.

10. **Test de Convergencia**. Este test está basado en la desviación estándar de las $(n+1)$ evaluaciones de la función f en los vértices del simplex. Calcular

$$\sigma^2 = \frac{1}{n+1}\sum_{i=1}^{n+1}(f_i - \overline{f})^2$$

donde $\overline{f} = \sum_{i=1}^{n} \frac{f_i}{n+1}$.

- Si $\sigma < \epsilon$, significa que los valores de la funciones están muy próximos y posiblemente cerca del mínimo. PARAR
- En caso contrario, ir al paso 3.

Basándose en diversas pruebas del método, combinando muchas funciones diferentes, Nelder y Mead recomiendan tomar $\alpha = 1$, $\beta = 2$ y $\gamma = 0{,}5$.

Algoritmo de Rosenbrock

1. Elegir un punto inicial $\mathbf{x}_1$, un criterio de parada, $\epsilon > 0$, un factor de expansión, $\alpha > 1$, y un factor de contracción $\beta \in (-1, 0)$. Tomar $\mathbf{d}_j = \mathbf{e}^j$, $j = 1, \ldots, n$, el vector unitario en la dirección del eje de la variable x_j. Sean $h_1, h_2, \ldots, h_n$ las longitudes de paso a lo largo de cada una de las dirección. Tomar $\mathbf{y}_1 = \mathbf{x}_1$, hacer $k = j = 1$ e ir al paso 2.

2. Calcular $f(\mathbf{y}_j + h_j \mathbf{d}_j)$.
   - Si $f(\mathbf{y}_j + h_j \mathbf{d}_j) < f(\mathbf{y}_j)$, tomar $\mathbf{y}_{j+1} = \mathbf{y}_j + h_j \mathbf{d}_j$, $h_j = \alpha h_j$. Ir al paso 3.
   - Si $f(\mathbf{y}_j + h_j \mathbf{d}_j) \geq f(\mathbf{y}_j)$, tomar $\mathbf{y}_{j+1} = \mathbf{y}_j$, $h_j = \beta h_j$. Ir al paso 3.

3. Si $j < n$, hacer $j = j + 1$ e ir al paso 2. En caso contrario, ir al paso 4.

4. Calcular $f(\mathbf{y}_{n+1})$.
   - Si $f(\mathbf{y}_{n+1}) < f(\mathbf{y}_1)$, entonces tomar $\mathbf{y}_1 = \mathbf{y}_{n+1}$, hacer $j = 1$ y volver al paso 2.
   - Si $f(\mathbf{y}_{n+1}) < f(\mathbf{y}_1)$, pero $f(\mathbf{y}_{n+1}) < f(\mathbf{x}_k)$, ir al paso 5.
   - Si $f(\mathbf{y}_{n+1}) < f(\mathbf{y}_1)$ y $f(\mathbf{y}_{n+1}) = f(\mathbf{x}_k)$, distinguimos dos casos:
     - Si $|h_j| < \epsilon$, $\mathbf{x}_k$ es una aproximación al óptimo de la función. FIN.
     - En caso contrario, tomar $\mathbf{y}_1 = \mathbf{y}_{n+1}$, hacer $j = 1$ y volver al paso 2.

5. Tomar $\mathbf{x}_{k+1} = \mathbf{y}_{n+1}$.
   - Si $\| x_{k+1} - x_k \| < \epsilon$, tomar $\mathbf{x}_{k+1}$ como una aproximación al óptimo de la función. FIN.
   - Calcular $\lambda_1, \lambda_2, \ldots, \lambda_n$ de la relación

$$\mathbf{x}_{k+1} - \mathbf{x}_k = \sum_{j=1}^{n} \lambda_j \mathbf{d}_j$$

Formar un nuevo conjunto de direcciones $\overline{\mathbf{d}}_1, \ldots, \overline{\mathbf{d}}_n$ de la siguiente forma

$$\mathbf{w}_j = \begin{cases} \mathbf{d}_j & \text{si } \lambda_j = 0 \\ \sum_{i=j}^{n} \lambda_i \mathbf{d}_i & \text{si } \lambda_j \neq 0 \end{cases}$$

$$\mathbf{y}_j = \begin{cases} \mathbf{w}_j & j = 1 \\ \mathbf{w}_j - \sum_{i=1}^{j-1} (\mathbf{w}_j^t \overline{\mathbf{d}}_i) \overline{\mathbf{d}}_i & j \geq 2 \end{cases}$$

Tomar $\overline{\mathbf{d}}_j = \frac{\mathbf{y}_j}{\|\mathbf{y}_j\|}$. Considerar las longitudes de paso iniciales $h_1, \ldots, h_n$. Tomar $\mathbf{y}_1 = \mathbf{x}_{k+1}$, hacer $k = k + 1$, $j = 1$ e ir al paso 2.