

Método de Optimización de Nelder-Mead

Resumen

El método de Nelder-Mead, también conocido como método del simplex descendente, es un algoritmo heurístico ampliamente utilizado para la optimización de funciones no lineales sin restricciones que no requiere el cálculo de derivadas. A continuación se presenta una exposición detallada del método, incluyendo sus fundamentos teóricos, el algoritmo completo y sus propiedades principales.

1. Introducción

1.1. Problema de Optimización

Definición 1.1 (Problema de Optimización No Lineal). Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función objetivo. El problema de optimización sin restricciones consiste en encontrar:

$$\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad (1)$$

donde \mathbf{x}^* es un mínimo local o global de f .

Los métodos de optimización se clasifican en:

- **Métodos con derivadas:** Requieren el cálculo del gradiente ∇f y posiblemente la matriz Hessiana $\nabla^2 f$.
- **Métodos sin derivadas:** Solo evalúan la función objetivo $f(\mathbf{x})$. Son útiles cuando las derivadas no están disponibles o son costosas de calcular.

El método de Nelder-Mead pertenece a la segunda categoría.

2. Conceptos Fundamentales

2.1. Simplex

Definición 2.1 (Símplex). Es el politopo más simple en una dimensión dada y se define como la envoltura convexa de $(n + 1)$ puntos independientes afines en un espacio de dimensión n . Es el conjunto de todos los puntos que se pueden formar como combinaciones convexas de sus vértices. Dados los vértices $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n+1} \in \mathbb{R}^n$, el simplex S está definido como:

$$S = \left\{ \mathbf{x} \in \mathbb{R}^n / \mathbf{x} = \sum_{i=1}^{n+1} \lambda_i \mathbf{x}_i : \sum_{i=0}^n \lambda_i = 1, \lambda_i \geq 0 \right\} \quad (2)$$

Ejemplos:

- En \mathbb{R}^1 : un simplex es un segmento de línea (2 puntos).
- En \mathbb{R}^2 : un simplex es un triángulo (3 puntos).
- En \mathbb{R}^3 : un simplex es un tetraedro (4 puntos).

3. El Método de Nelder-Mead

3.1. Idea Principal

El método de Nelder-Mead mantiene un simplex de $n + 1$ vértices que se mueve, expande y contrae iterativamente en el espacio de búsqueda, intentando desplazarse hacia regiones con menores valores de la función objetivo. En una primera instancia se asume que no existen restricciones respecto al espacio de búsqueda.

3.2. Ordenamiento de Vértices

En cada iteración del algoritmo, los vértices del simplex se ordenan según el valor de la función objetivo:

$$f(\mathbf{x}_1) \leq f(\mathbf{x}_2) \leq \dots \leq f(\mathbf{x}_{n+1}) \quad (3)$$

donde:

- \mathbf{x}_1 es el **mejor vértice** (menor valor de f).
- \mathbf{x}_{n+1} es el **peor vértice** (mayor valor de f).
- \mathbf{x}_n es el **segundo peor vértice**.

3.3. Operaciones Geométricas

El algoritmo utiliza cuatro operaciones geométricas principales, todas basadas en el **centroide** de los n mejores puntos:

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \quad (4)$$

3.3.1. 1. Reflexión

El punto reflejado \mathbf{x}_r se obtiene reflejando el peor punto \mathbf{x}_{n+1} respecto al centroide:

$$\mathbf{x}_r = \bar{\mathbf{x}} + \alpha(\bar{\mathbf{x}} - \mathbf{x}_{n+1}) \quad (5)$$

donde $\alpha > 0$ es el **coeficiente de reflexión** (típicamente $\alpha = 1$).

3.3.2. 2. Expansión

Si la reflexión produce un punto muy bueno ($f(\mathbf{x}_r) < f(\mathbf{x}_1)$), se intenta expandir en esa dirección:

$$\mathbf{x}_e = \bar{\mathbf{x}} + \gamma(\mathbf{x}_r - \bar{\mathbf{x}}) \quad (6)$$

donde $\gamma > 1$ es el **coeficiente de expansión** (típicamente $\gamma = 2$).

3.3.3. 3. Contracción

Si la reflexión no mejora suficientemente, se contrae el simplex:

- **Contracción externa** (si $f(\mathbf{x}_r) < f(\mathbf{x}_{n+1})$):

$$\mathbf{x}_c = \bar{\mathbf{x}} + \beta(\mathbf{x}_r - \bar{\mathbf{x}}) \quad (7)$$

- **Contracción interna** (si $f(\mathbf{x}_r) \geq f(\mathbf{x}_{n+1})$):

$$\mathbf{x}_c = \bar{\mathbf{x}} - \beta(\bar{\mathbf{x}} - \mathbf{x}_{n+1}) \quad (8)$$

donde $0 < \beta < 1$ es el **coeficiente de contracción** (típicamente $\beta = 0,5$).

3.3.4. 4. Reducción

Si ninguna de las operaciones anteriores mejora la situación, se reduce el simplex completo, acercando los $(\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_{n+1})$ puntos hacia el mejor punto:

$$\mathbf{x}_i = \mathbf{x}_1 + \delta(\mathbf{x}_i - \mathbf{x}_1), \quad i = 2, \dots, n+1 \quad (9)$$

donde $0 < \delta < 1$ es el **coeficiente de reducción** (típicamente $\delta = 0,5$).

3.4. Algoritmo Completo

Algorithm 1 Método de Nelder-Mead

```

1: Entrada: Punto inicial  $\mathbf{x}_0$ , parámetros  $\alpha, \gamma, \beta, \delta$ , tolerancia  $\epsilon$ 
2: Inicialización: Construir simplex inicial con  $n + 1$  vértices alrededor de  $\mathbf{x}_0$ 
3: repeat
4:   Ordenar vértices:  $f(\mathbf{x}_1) \leq f(\mathbf{x}_2) \leq \dots \leq f(\mathbf{x}_{n+1})$ 
5:   Calcular centroide:  $\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i$ 
6:   Reflexión:  $\mathbf{x}_r = \bar{\mathbf{x}} + \alpha(\bar{\mathbf{x}} - \mathbf{x}_{n+1})$ 
7:   if  $f(\mathbf{x}_1) \leq f(\mathbf{x}_r) < f(\mathbf{x}_n)$  then
8:     Aceptar reflexión:  $\mathbf{x}_{n+1} = \mathbf{x}_r$ 
9:   else if  $f(\mathbf{x}_r) < f(\mathbf{x}_1)$  then
10:    Expansión:  $\mathbf{x}_e = \bar{\mathbf{x}} + \gamma(\mathbf{x}_r - \bar{\mathbf{x}})$ 
11:    if  $f(\mathbf{x}_e) < f(\mathbf{x}_r)$  then
12:      Aceptar expansión:  $\mathbf{x}_{n+1} = \mathbf{x}_e$ 
13:    else
14:      Aceptar reflexión:  $\mathbf{x}_{n+1} = \mathbf{x}_r$ 
15:    end if
16:  else
17:    Contracción:
18:    if  $f(\mathbf{x}_r) < f(\mathbf{x}_{n+1})$  then
19:      Contracción externa:  $\mathbf{x}_c = \bar{\mathbf{x}} + \beta(\mathbf{x}_r - \bar{\mathbf{x}})$ 
20:    else
21:      Contracción interna:  $\mathbf{x}_c = \bar{\mathbf{x}} - \beta(\bar{\mathbf{x}} - \mathbf{x}_{n+1})$ 
22:    end if
23:    if  $f(\mathbf{x}_c) < f(\mathbf{x}_{n+1})$  then
24:      Aceptar contracción:  $\mathbf{x}_{n+1} = \mathbf{x}_c$ 
25:    else
26:      Reducción:  $\mathbf{x}_i = \mathbf{x}_1 + \delta(\mathbf{x}_i - \mathbf{x}_1)$  para  $i = 2, \dots, n + 1$ 
27:    end if
28:  end if
29: until criterio de parada satisfecho
30: Salida:  $\mathbf{x}^* \approx \mathbf{x}_1$ 

```

3.5. Criterios de Parada

El algoritmo se detiene cuando se cumple alguno de los siguientes criterios:

1. **Tamaño del simplex:**

$$\max_{i=2, \dots, n+1} \|\mathbf{x}_i - \mathbf{x}_1\| < \epsilon \quad (10)$$

2. **Variación de la función objetivo:**

$$\sqrt{\frac{1}{n+1} \sum_{i=1}^{n+1} (f(\mathbf{x}_i) - \bar{f})^2} < \epsilon \quad (11)$$

donde $\bar{f} = \frac{1}{n+1} \sum_{i=1}^{n+1} f(\mathbf{x}_i)$

3. Número máximo de iteraciones alcanzado.

Convergencia: La convergencia a un mínimo global no está garantizada, dado que puede converger a un mínimo local. La relación entre la curvatura de la función objetivo y el tamaño del simplex pueden inducir problemas de convergencia.

4. Algunas características del Método

4.1. Ventajas

- **No requiere derivadas:** Especialmente útil cuando el gradiente no está disponible o es costoso de calcular.
- **Simplicidad conceptual:** Fácil de entender e implementar.
- **Robusto para funciones ruidosas:** Puede manejar funciones con ruido o discontinuidades menores.
- **Bajo costo computacional por iteración:** El número de evaluaciones de función por iteración es típicamente $O(1)$ en el mejor caso y $O(n)$ en el peor caso (cuando se requiere reducción).

4.2. Desventajas

- **No hay garantías de convergencia:** El método puede converger a mínimos locales o colapsar en regiones de poca variación de la función objetivo.
- **Degeneración del simplex:** El simplex puede colapsar a una dimensión menor.
- **Convergencia lenta:** Especialmente cerca del óptimo, comparado con métodos de segundo orden.
- **Sensibilidad al simplex inicial:** La elección inicial puede afectar significativamente el resultado.

4.3. Restricciones

En caso de existir restricciones en el espacio de búsqueda, éstas se pueden imponer *a posteriori* en la coordenada correspondiente, luego de cada operación geométrica.

5. Parámetros Recomendados

Los valores estándar propuestos por Nelder y Mead son:

Estos parámetros deben satisfacer:

$$\alpha > 0, \quad \gamma > 1, \quad \gamma > \alpha, \quad 0 < \beta < 1, \quad 0 < \delta < 1 \quad (12)$$

Parámetro	Símbolo	Valor
Reflexión	α	1
Expansión	γ	2
Contracción	β	0.5
Reducción	δ	0.5

Cuadro 1: Parámetros estándar del método de Nelder-Mead

6. Construcción del Símplex Inicial

Una estrategia común para construir el símplex inicial a partir de un punto \mathbf{x}_0 es:

$$\mathbf{x}_i = \mathbf{x}_0 + \lambda_i \mathbf{e}_i, \quad i = 1, \dots, n \quad (13)$$

donde \mathbf{e}_i es el i -ésimo vector canónico y λ_i es un paso apropiado (por ejemplo, $\lambda_i = 0,05$ o el 5 % del valor de $x_{0,i}$). También se puede hacer una generación de m puntos aleatorios ($m > n + 1$) y seleccionar los $n + 1$ mejores, verificando su independencia.

7. Aplicaciones

El método de Nelder-Mead se utiliza ampliamente en:

- Optimización de modelos donde el gradiente no está disponible
- Ajuste de parámetros en simulaciones o funciones
- Optimización de funciones objetivo con evaluación costosa