

MÉTODOS INDIRECTOS – PRIMER ORDEN

Una buena dirección de búsqueda debe reducir la función objetivo tal que si x^0 es el punto original y x^1 es el nuevo punto entonces:

$$f(x^1) \leq f(x^2)$$

Una dirección s se llama dirección de descenso y satisface para cualquier punto que

$$\nabla^T f(x) * s < 0$$

De la figura vemos que el ángulo entre los vectores gradiente y s^k es θ , entonces

$$\nabla^T f(x) \cdot s^K = |\nabla^T f(x)| \cdot |s^K| \cdot \cos \theta$$

Si $0 \leq \theta \leq 90^\circ$ no mejora la función objetivo

Si $\theta > 90^\circ$ las direcciones de búsqueda producirán disminuciones en el valor de $f(x)$ entonces

$$\nabla^T f(x) * s < 0$$

MÉTODO DEL GRADIENTE

Este método utiliza en sus cálculos únicamente la derivada primera de la función. Para maximización la dirección de búsqueda es simplemente el gradiente, para minimización la dirección de búsqueda es el negativo del gradiente

$$s^K = -\nabla f(x)$$

Entonces para minimización, en una etapa k la transición desde el punto x^K a otro punto x^{K+1} puede verse por la siguiente expresión

$$x^{k+1} = x^k + \Delta x^k = x^k + \lambda^k \cdot s^k = x^k - \lambda^k \cdot \nabla f(x^k)$$

donde

Δx^k = distancia desde x^k a x^{k+1}

s^k = dirección de búsqueda

λ^k = escalar que determina el largo del paso en la dirección s^k

El negativo del vector gradiente da la dirección para la minimización, pero no la magnitud del paso, por lo que son posibles varios procedimientos dependiendo de cómo se elige λ^k . En este método el óptimo no se alcanza en una etapa por lo que la ecuación anterior debe aplicarse repetidamente para alcanzar el mínimo. En el mínimo, el valor de los elementos del vector gradiente debe ser cero.

Existen dos métodos para seleccionar el valor de λ . Un método emplea una búsqueda unidimensional a lo largo del gradiente negativo, el segundo especifica el tamaño del paso a un valor constante y utiliza este valor en cada iteración.

El algoritmo del gradiente puede ser resumido como sigue:

1. Elegir un punto inicial de búsqueda x^0
2. Calcular (analítica ó numéricamente) las derivadas parciales $\frac{\partial f}{\partial x_j}$ $j = 1, \dots, n$
3. Calcular la dirección de búsqueda $s = -\nabla f(x^k)$
4. Usar la relación $x^{k+1} = x^k + \lambda^k \cdot s^k$ para obtener el valor x^{k+1} (para encontrar λ^k use la ecuación 6-5 ó minimice $f(x)$ numéricamente)
5. Compare $f(x^{k+1})$ con $f(x^k)$, si el cambio en $f(x)$ es menor que alguna tolerancia, pare. En caso contrario regrese a la etapa 2, haga $k=k+1$. La terminación puede también especificarse fijando una tolerancia sobre el tamaño de λ ó especificando alguna tolerancia sobre la norma.

Inconvenientes: Este método puede finalizar en cualquier tipo de punto estacionario, esto es, un punto donde los elementos del vector gradiente son cero. Los puntos estacionarios se testean examinando la matriz del Hessiano. Si esta no es definida positiva, el punto estacionario es un punto de ensilladura.

MÉTODO DEL GRADIENTE CONJUGADO (Fletcher y Reeves)

El método combina información actual del vector gradiente con los vectores gradientes de iteraciones previas para obtener una nueva dirección de búsqueda. Los pasos son los siguientes:

1. En x^0 calcular $f(x^0)$. Sea $s^0 = -\nabla f(x^0)$

2. Guardar $\nabla f(x^0)$ y calcular $x^1 = x^0 + \lambda^0 s^0$
minimizando $f(x)$ con respecto a λ en la dirección

s^0 (Esto es, llevando a cabo una búsqueda unidimensional para λ)

3. Calcular $f(x^1)$, $\nabla f(x^1)$, la nueva dirección de búsqueda es una combinación lineal de s^0 y $\nabla f(x^1)$

$$s^1 = -\nabla f(x^1) + s^0 \cdot \frac{\nabla^T f(x^1) \nabla f(x^1)}{\nabla^T f(x^0) \nabla f(x^0)}, \text{ para la iteración } k,$$

la relación es:

$$s^{k+1} = -\nabla f(x^{k+1}) + s^k \cdot \frac{\nabla^T f(x^{k+1}) \nabla f(x^{k+1})}{\nabla^T f(x^k) \nabla f(x^k)}$$

4. El procedimiento finaliza cuando s sea menor a la tolerancia

MÉTODOS INDIRECTOS DE SEGUNDO ORDEN

La dirección de búsqueda del método del gradiente puede interpretarse que es ortogonal a una aproximación lineal (tangente a) de la función objetivo en el punto x^k .

Si hacemos una aproximación cuadrática de $f(x)$ en x^k

$$f(x) = f(x^k) + \nabla^T f(x^k) \cdot \Delta x^k + \frac{1}{2} (\Delta x^k)^T \cdot H(x^k) \cdot \Delta x^k \text{ donde}$$

$H(x^k$

) es el Hessiano de $f(x)$. Entonces será posible tomar en cuenta la curvatura de $f(x)$ en x^k en la determinación de una dirección de búsqueda como se describe para el método de Newton.

MÉTODO DE NEWTON

Este método al tomar en cuenta la curvatura de la función en cada punto, identifica mejores direcciones de búsqueda.

El mínimo de $f(x)$ en x^k es obtenido diferenciando la aproximación cuadrática de $f(x)$ con respecto de cada uno de los componentes de x e igualando la expresión resultante a cero

$$\nabla f(x) = \nabla f(x^k) + H(x^k) \cdot \Delta x^k = 0$$

ó

$x^{k+1} - x^k = \Delta x^k = -[H(x^k)]^{-1} \cdot \nabla f(x^k)$ estando en este método tanto la dirección de búsqueda como el tamaño del paso especificados. Si la función es cuadrática, se requiere solo un paso para alcanzar el óptimo. Sin embargo, para una función objetivo no lineal el mínimo no será alcanzado en una etapa por lo que la ecuación se puede modificar introduciendo el tamaño del paso:

$$x^{k+1} - x^k = -\lambda^k \cdot [H(x^k)]^{-1} \cdot \nabla f(x^k)$$

$$\text{con } s^k = -[H(x^k)]^{-1} \cdot \nabla f(x^k)$$

La ecuación se aplica iterativamente hasta satisfacer algún criterio de terminación (considerar $\lambda=1$ en cada etapa)

Bibliografía

"Optimization of Chemical Process" , Edgar T.F. and Himmelblau D.M., McGraw-Hill New York, 1988.

"Engineering Optimization - Methods and Applications" , Reklaitis G.V., Ravindran A. and Ragsdell K.M., Wiley, New York, 1983.