

Tercer parcial de Investigación Operativa 2015

1) Plantee la forma general de un problema de minimización no lineal sin restricciones. Describa el método del máximo descenso y el método de Newton para minimizar una función no lineal. Ventajas y desventajas.

RESOLUCION

Un problema de minimización no lineal sin restricciones consiste en hallar el mínimo de una función no lineal $g(x)$. Todos los métodos de minimización no restringidos son naturalmente iterativos: parten de una solución de inicio arbitrario y proceden secuencialmente en busca del mínimo, es decir

$$X_{i+1} = X_i + \lambda^*_i S_i$$

donde X_i es el punto de inicio, S_i es la dirección de búsqueda, λ^*_i es el tamaño de paso óptimo y X_{i+1} es el punto final de la iteración.

El **método del máximo descenso** de Cauchy consta de los siguientes pasos.

- 1) Partida de un punto arbitrario inicial X_1 . Iteración $i = 1$.
- 2) Se busca la dirección $S_i = -\nabla f_i = -\nabla f(X_i)$
- 3) Determinamos el tamaño de paso óptimo λ^* en la dirección S_i y obtenemos $X_{i+1} = X_i + \lambda^*_i S_i = X_i - \lambda^*_i \nabla f_i$
- 4) Probamos la optimalidad de X_{i+1} . Si X_{i+1} es óptimo paramos, de lo contrario pasamos a etapa 5.
- 5) Establecemos la nueva iteración $i = i + 1$ y procedemos desde el paso 2.

El **método de Newton** aproxima cuadráticamente la función usando la serie de Taylor:

$$f(X) = f(X_i) + \nabla f_i^T (X - X_i) + (1/2)(X - X_i)^T [H_i] (X - X_i) \quad (1)$$

donde H_i es la matriz hessiana evaluada en el punto X_i . Estableciendo las derivadas parciales de (1) igual a cero para el mínimo de $f(X)$ obtenemos

$$\frac{\partial f(X)}{\partial x_j} = 0, \quad j = 1, 2, \dots, n \quad (2)$$

Las expresiones (1) y (2) nos dan (derivando respecto a X)

$$\nabla f = \nabla f_i + [H_i](X - X_i) = 0 \quad (3)$$

Si $[H_i]$ es no singular, de (3) se obtiene una aproximación mejorada ($X = X_{i+1}$) como

$$X_{i+1} = X_i - [H_i]^{-1} \nabla f_i \quad (4)$$

en la que se despreciaron los términos de orden superior de la serie. (4) se emplea iterativamente para hallar la solución óptima X^* , constituyendo el método de Newton, que converge para un punto de inicio próximo al óptimo. Para considerar el caso de no cuadráticas y la posibilidad de divergencia, se modifica (4)

$$X_{i+1} = X_i + \lambda^*_i S_i = X_i - \lambda^*_i [H_i]^{-1} \nabla f_i \quad (7)$$

donde λ_i^* es la minimización del tamaño de paso en la dirección $S_i = -[H_i]^{-1} \nabla f_i$.

Esta modificación conlleva aspectos positivos:

- Se hallará el mínimo en menor número de etapas que la versión original.
- Encuentra el mínimo en todos los casos
- Generalmente evita convergencia a punto silla o máximo.

Ahora, la practicidad del método de Newton está cuestionada por hechos como:

- Requiere el almacenado de matrices $[H_i]$ de $n \times n$.
- Se vuelve dificultoso, a veces no posible, calcular los elementos de la matriz $[H_i]$.
- Requiere la inversión de la matriz $[H_i]$ en cada etapa.
- Requiere la evaluación de $[H_i]^{-1} \nabla f_i$ en cada etapa.

2) Sea la función $f(x, y) = -20 + x^2 + y^2$. Aplique el método del máximo descenso para minimizarla (manual). Dibuje las curvas de nivel empleando software Matlab. Resuélvalo con el algoritmo en el ordenador.

RESOLUCION

El primer paso del método del máximo descenso es escoger un punto arbitrario de inicio. Elegimos $X_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

El gradiente de la función es: $\nabla f = \begin{pmatrix} \partial f / \partial x = 2x \\ \partial f / \partial y = 2y \end{pmatrix}$

Entonces S_1 es: $-\nabla f(X_1) = -\begin{pmatrix} 2x \\ 2y \end{pmatrix} = -\begin{pmatrix} 2 \times 1 \\ 2 \times 3 \end{pmatrix} = \begin{pmatrix} -2 \\ -6 \end{pmatrix}$

Para hallar el paso óptimo hay que minimizar $f(X_1 + \lambda^*_{-1} S_1) = f\left(\begin{pmatrix} 1 \\ 3 \end{pmatrix} + \lambda^*_{-1} \begin{pmatrix} -2 \\ -6 \end{pmatrix}\right)$

$$\begin{aligned} f(1 - 2\lambda^*_{-1} ; 3 - 6\lambda^*_{-1}) &= -20 + (1 - 2\lambda^*_{-1})^2 + (3 - 6\lambda^*_{-1})^2 \\ &= -20 + 1 - 4\lambda^*_{-1} + 4\lambda^{*2}_{-1} + 9 - 36\lambda^*_{-1} + 36\lambda^{*2}_{-1} \\ &= -10 - 40\lambda^*_{-1} + 40\lambda^{*2}_{-1} \end{aligned}$$

Esta función se minimiza haciendo

$$\begin{aligned} \partial f / \partial \lambda^*_{-1} &= 0 \\ -40 + 80\lambda^*_{-1} &= 0 \\ \lambda^*_{-1} &= 0,5 \end{aligned}$$

De esta forma se tiene $X_2 = X_1 + \lambda^*_{-1} S_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix} + 0,5 \begin{pmatrix} -2 \\ -6 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} + \begin{pmatrix} -1 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Para saber si X_2 es el óptimo evaluamos el gradiente de la función en ese punto:

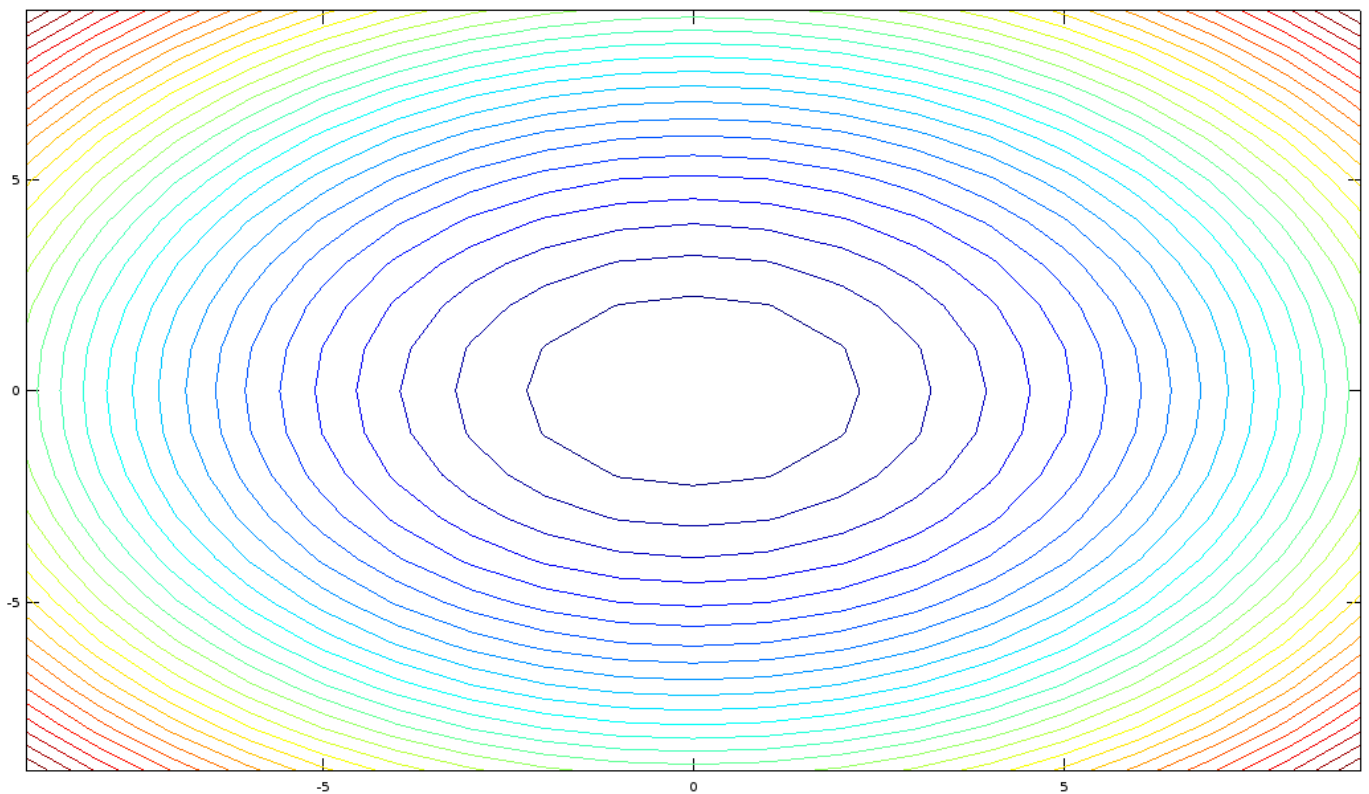
$$\nabla f(X_2) = \begin{pmatrix} 2x \\ 2y \end{pmatrix} = \begin{pmatrix} 2 \times 0 \\ 2 \times 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Efectivamente X_2 es el óptimo buscado.

Para dibujar las curvas de nivel, en Octave se usan los siguientes comandos:

```
[X, Y] = meshgrid(-9:9, -9:1:9);
Z = -20 + X.^2 + Y.^2;
contour(X, Y, Z, 30);
```

El primer comando crea la rejilla mediante dos matrices, X e Y y las llena con valores desde -9 hasta 9. El segundo calcula los valores de Z para cada combinación de valores de X e Y. El tercero dibuja las curvas de nivel con los valores de X, Y y Z (30 indica la cantidad de niveles a dibujar). El resultado es este:



Lo cual se entiende mejor si vemos la envoltura de la función:

