KKT : Método de Newton

José Luis Morales

Departamento de Matemáticas. ITAM. 2015.

Ejemplo (Cálculo 3)

Optimizar una función de varias variables sujeta a restricciones de igualdad

$$\text{minimizar} \quad f(x) \tag{1}$$

$$\text{sujeta a} \quad c(x) = 0, \tag{2}$$

en donde $f : \mathbb{R}^n \to \mathbb{R}$ y $c : \mathbb{R}^n \to \mathbb{R}^m$ son funciones dos veces continuamente diferenciables; el número de restricciones es menor que el número de variables, i.e. $n > m$.

Función Lagrangiana

$$\mathcal{L}(x, \lambda) = f(x) - \lambda^T c(x)$$

Condiciones de KKT de primer orden

Condiciones necesarias de Karush-Kuhn-Tucker de primer orden para un minimizador local

$$\begin{aligned} \nabla_x \mathcal{L}(x,\lambda) &= 0 \\ \nabla_\lambda \mathcal{L}(x,\lambda) &= 0, \end{aligned}$$

obtenemos el siguiente sistema de $n+m$ ecuaciones no lineales con $n+m$ incógnitas

$$F(w) = \begin{bmatrix} \nabla f(x) - A(x)\lambda \\ c(x) \end{bmatrix}, \quad w = \begin{bmatrix} x \\ \lambda \end{bmatrix}$$

$\Longrightarrow$ *puntos estacionarios.*

Método de Newton

Si utilizamos el método de Newton para obtener una solución aproximada del sistema anterior

$$\begin{bmatrix} \nabla^2_{xx}\mathcal{L}(x_k, \lambda_k) & -A(x_k) \\ A^T(x_k) & 0 \end{bmatrix} h = - \begin{bmatrix} \nabla f(x_k) - A(x_k)\lambda_k \\ c(x_k) \end{bmatrix}$$

$$h = \begin{bmatrix} x_{k+1} - x_k \\ \lambda_{k+1} - \lambda_k \end{bmatrix}$$

Simplificar ...

$h_x = x_{k+1} - x_k$

Matriz de KKT

$$\left[\begin{array}{cc} \nabla^2_{xx}\mathcal{L}(x_k,\lambda_k) & A(x_k) \\ A^T(x_k) & 0 \end{array}\right]\left[\begin{array}{c} h_x \\ -\lambda_{k+1} \end{array}\right] = -\left[\begin{array}{c} \nabla f(x_k) \\ c(x_k) \end{array}\right].$$

Nos interesa investigar las condiciones bajo las cuales el sistema anterior tiene solución única.

Álgebra lineal

Observemos el segundo bloque

$$A^T(x_k)h_x = -c(x_k),$$

es inmediato concluir que si $A^T(x_k)$ es de rango completo, entonces es posible obtener la solución general h_x como

$$h_x = h_x^0 + Z_k y,$$

en donde h_x^0 es una solución particular del sistema completo y Z_k es una base del espacio nulo de $A^T(x_k)$.

Ahora podemos utilizar el primer bloque

$$\nabla^2_{xx}\mathcal{L}(x_k,\lambda_k)h_x - A(x_k)\lambda_{k+1} = -\nabla f(x_k).$$

Si multiplicamos a la izquierda por Z_k^T tenemos

$$Z_k^T[\nabla^2_{xx}\mathcal{L}(x_k,\lambda_k)h_x - A(x_k)\lambda_{k+1}] = -Z_k^T\nabla f(x_k).$$

Notemos que $Z_k^T A(x_k) = 0,\ A^T(x_k)Z_k = 0\ldots$ (Z_k es una base del espacio nulo de $A^T(x_k)$), i.e.

$$Z_k^T\nabla^2_{xx}\mathcal{L}(x_k,\lambda_k)h_x = -Z_k^T\nabla f(x_k).$$

Sustituyendo $h_x = h_x^0 + Z_k y$ en la expresión anterior

$$\begin{aligned} Z_k^T \nabla^2_{xx}\mathcal{L}(x_k,\lambda_k)(h_x^0 + Z_k y) &= -Z_k^T \nabla f(x_k), \\ Z_k^T \nabla^2_{xx}\mathcal{L}(x_k,\lambda_k) Z_k y &= -Z_k^T [\nabla f(x_k) + \nabla^2_{xx}\mathcal{L}(x_k,\lambda_k) h_x^0], \end{aligned}$$

notamos que es posible obtener una solución única y si y sólo si la matriz

$$Z_k^T \nabla^2_{xx}\mathcal{L}(x_k,\lambda_k) Z_k$$

es simétrica positiva definida.

Métodos: Cholesky, GC, GC proyectado

Iterando $k+1$

$$h_x = h_x^0 + Z_k y$$

$$x_{k+1} = x_k + h_x$$

Del primer bloque

$$A(x_k)\lambda_{k+1} = \nabla f(x_k) - \nabla_{xx}^2 \mathcal{L}(x_k, \lambda_k) h_x$$

por lo tanto

$$A(x_k)^T A(x_k)\lambda_{k+1} = A(x_k)^T \left(\nabla f(x_k) - \nabla_{xx}^2 \mathcal{L}(x_k, \lambda_k) h_x \right)$$

$A(x_k)^T A(x_k)$ es es simétrica positiva definida.