Optimización II.
Programación Cuadrática Secuencial

November 11, 2005

El Método de Programación Cuadrática Secuencial es uno de los más efectivos para problemas de optimización con restricciones no lineales.

- Es apropiado para pequeños y grandes problemas
- Es muy bueno resolviendo problemas con no-linealidades significativas.

La idea general del método consiste en, a partir de un problema dado, construir una serie de problemas cuadráticos equivalente al primero, por lo que resolver el problema se reduciría a resolver esta secuancia de problemas QP.

<u>MÉTODO LOCAL SQP</u>:

Comencemos por el problema:

$$\min f(x) \tag{1}$$
$$sa : c(x) \quad = \quad 0 \tag{2}$$

donde $f : R^n \to R$ y $c : R^n \to R^m$ son funciones suaves.

Este problema solo contiene restricciones de igualdad, el cual no es muy común, pero la comprensión de (**??**) es muy importante en el diseño de SQP para problemas con restricciones más generales.
La idea de SQP es modelar (1) en la iteración x_k como un subproblema cuadrático y usar el minimizador de este subproblema para definir la próxima iteración.

Se conoce que el Lagrangiano para este problema es: $\mathcal{L}(x, \lambda) = f(x) - \lambda^T c(x)$.

Denotemos por $A(x)$ el jacobiano de las restricciones, es decir:

$$A(x) = [\nabla c_1(x), \nabla c_2(x), ..., \nabla c_m(x)]^T \tag{3}$$

donde $c_i(x)$ son las componentes del vector $c(x)$.

Especificando las restricciones de KKT para el caso de las restricciones de igualdad, obtenemos el siguiente sistema de ecuaciones de$n + m$ ecuaciones con $n + m$ incógnitas x y λ.

$$F(x, \lambda) = \begin{bmatrix} \nabla f(x) - A(x)^T \lambda \\ c(x) \end{bmatrix} = 0 \tag{4}$$

Si A_* es de rango completo, cualquier solución (x^*, λ^*) del problema (1) satisface (4).

La expresión (4) sugiere usar el Método de Newton.

El Jacobiano de (4) esta dado por:

$$\begin{bmatrix} W(x, \lambda) & -A(x)^T \\ A(x) & 0 \end{bmatrix}, \tag{5}$$

donde W denota el Hessiano del Lagrangiano,

$$W(x, \lambda) = \nabla^2_{xx} \mathcal{L}(x, \lambda) \tag{6}$$

El paso de Newton para la iteración (x_k, λ_k) es dada por:

$$\begin{bmatrix} x_{k+1} \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ \lambda_k \end{bmatrix} + \begin{bmatrix} p_k \\ p_\lambda \end{bmatrix} \tag{7}$$

donde $p_k = x_{k+1} - x_k$ y $p_\lambda = \lambda_{k+1} - \lambda_k$ resuelven el sistema de KKT,

$$\begin{bmatrix} W_k & -A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} p_k \\ p_\lambda \end{bmatrix} = \begin{bmatrix} -\nabla f_k + A^T \lambda_k \\ -c_k \end{bmatrix} \tag{8}$$

Esta iteración, la cual es llamada a veces como Método de Newton-Lagrange, está bien definida cuando la matriz de KKT es no singular.
La no singularidad es consecuancia de las siguientes condiciones:

a El Jacobiano de las restricciones A_k es de rango completo por columnas.

b La matriz W_k es positiva definida en el espacio tangente de las restricciones, $d^T W_k d > 0$ para todo $d \neq 0$ tal que $A_k d = 0$.

La primera condición nos garantiza LICQ para los gradientes de las restricciones y la segunda garantiza que se cumplan las condiciones suficientes de segundo orden en la solución.

Hay un enfoque alternativo para las iteraciones (7) y (8). Supongamos que en la iteración x_k, λ_k se define el problema.

$$\min_p \frac{1}{2} p^T W_k p + \nabla f_k^T p \tag{9}$$

$$s.a : A_k p + c_k \quad = \quad 0 \tag{10}$$

Nota: El problema anterior se obtiene de la aproximación de 2do orden del lagrangiano $\mathcal{L}(x_k + p)$.

$$\begin{aligned}
\mathcal{L}(x_k + p) &\approx \mathcal{L}(x_k) + \nabla\mathcal{L}(x_k)^T p + \frac{1}{2} p^T W_k p \\
&= f_k - \lambda^T c_k + (\nabla f_k - \nabla c_k \lambda)^T p + \frac{1}{2} p^T W_k p \\
&= f_k - \lambda^T (c_k + \nabla c_k^T p) + \nabla f_k^T p + \frac{1}{2} p^T W_k p \\
&= f_k - \lambda^T (c_k + A_k p) + \nabla f_k^T p + \frac{1}{2} p^T W_k p
\end{aligned}$$

Si se cumplen las condiciones anteriores para el problema (1) entonces el problema anterior tiene la única solución (x_k, λ_k).

$$W_k p_k + \nabla f_k - A_k^T \mu_k = 0 \tag{11}$$

$$A_k p_k + c_k = 0 \tag{12}$$

lo anterior lo podemos escribir:

$$\begin{bmatrix} W_k & -A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} p_k \\ \mu_k \end{bmatrix} = \begin{bmatrix} -\nabla f_k \\ -c_k \end{bmatrix} \tag{13}$$

Desarrollando (8) tenemos que:

$$\begin{bmatrix} W_k & -A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} p_k \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} -\nabla f_k \\ -c_k \end{bmatrix} \tag{14}$$

Como la matriz es no singular, entonces la solución es única y por tanto $\lambda_{k+1} = \mu_k$

Algoritmo 18.1(Algoritmo Local SQP)

Selecionar un punto inicial (x_0, λ_0)
for k = 0, 1, 2, ...
    Evaluar $f_k, \nabla f_k, W_k = W(x_k, \lambda_k), c_k$ y A_k;
    Resolver (9) para obtener p_k y μ_k;
    $x_{k+1} \leftarrow x_k + p_k$;$\lambda_{k+1} \leftarrow \lambda_k + p_k$;
    Si *se cumple la prueba de convergencia*
        PARAR con la solución aproximada (x_{k+1}, λ_{k+1});
end (for)

RESTRICCIONES DE DESIGUALDAD

El modelo SQP puede ser extendido fácilmente al problema general de programación no lineal:

$$\begin{aligned} \min f(x) & & \quad (15) \\ s.a: c_i(x) &= 0,\ i \in E \quad (16) \\ c_i(x) &\geq 0,\ i \in I \quad (17) \end{aligned}$$

Para definir el subproblema se linealizan tanto las restricciones de igualdad como de desigualdad, y obtenemos:

$$\begin{aligned} \min \frac{1}{2}p^T W_k p + \nabla f_k^T p & & \quad (18) \\ s.a: \nabla c_i(x_k)^T p + c_i(x_k) &= 0,\ i \in E \quad (19) \\ \nabla c_i(x_k)^T p + c_i(x_k) &\geq 0,\ i \in I \quad (20) \end{aligned}$$

Y para resolver este problema se pueden usar los algoritmos de programación cuadrática vistos en secciones anteriores.