FACULTAD CS. FÍSICAS Y MATEMÁTICAS — UNIVERSIDAD DE CHILE

MA57B Optimización No Lineal. Semestre 2007-1
Profesor: Héctor Ramírez C. **Auxiliar:** Oscar Peredo A.

Clase Auxiliar #9
Convergencia superlineal y cuadrática del Método PCS

21 de Junio de 2007

En el teorema principal de esta clase, se analiza la convergencia de la sucesión primal-dual generada por el método de Programación Cuadrática Secuencial, cuando el primer punto iterado (d_1, λ_1) está en una vecindad de un punto estacionario (x_*, λ_*) regular que satisface complementaridad estricta:

$$(CE) : c_i(x_*) < 0 \Leftrightarrow (\lambda_*)_i = 0$$

El problema a resolver es el siguiente:

$$(P_{EI}) : \begin{cases} \min_{x\in\Omega} & f(x) \\ \text{s.a} & c_E(x) = 0 \\ & c_I(x) \leq 0 \end{cases}$$

Recordemos que las restricciones de desigualdad activas se denotan por $I_*^0 = \{i \in I : c_i(x^*) = 0\}$. Antes de enunciar el teorema, conviene recordar el *problema cuadrático tangente*:

$$(QP_{EI}(k)) : \begin{cases} \min & \nabla f(x_k)^T d + 1/2 d^T L(x_k, d_k) d \\ \text{s.a} & c_E(x_k) + \nabla c_E(x_k)^T d = 0 \\ & c_I(x_k) + \nabla c_I(x_k)^T d \leq 0 \end{cases}$$

En adelante, cuando se mencione un punto estacionario, se considera regular. Recordemos el Método de Newton para problemas con restricciones de igualdad, el algoritmo es el siguiente:

Problema a ser resuelto:

$$(P_E) : \begin{cases} \min_{x\in\Omega} & f(x) \\ \text{s.a} & c(x) = 0 \end{cases}$$

Escoger punto inicial (x_1, λ_1)
$k \leftarrow 1$
repeat
  if $\nabla_x L(x_k, \lambda_k) < \epsilon$ y $c(x_k) = 0$ **then**
    PARAR
  end if
  Calcular $\nabla^2_{xx} L(x_k, \lambda_k)$ y encontrar un par primal-dual (d_k, λ_k^{QP}) estacionario para el problema:

$$(QP_E) : \begin{cases} \min & \nabla f(x) + 1/2 d^T L(x_k, \lambda_k) d \\ \text{s.a} & c(x) + \nabla c(x)^T d = 0 \end{cases}$$

  $x_{k+1} = x_k + d_k$
  $\lambda_{k+1} = \lambda_k^{QP}$
  $k \leftarrow k + 1$
until ¿Se cumple criterio de parada?

El teorema de convergencia local del método de Newton se utiliza en el principal resultado de esta clase:

Teorema 1 (Convergencia Local de Newton)**.** *Sean f, c de clase C^2 en una vecindad de un punto estacionario x_* del problema (P_E) con multiplicador asociado λ_*. Entonces existe una vecindad V de (x_*, λ_*) tal que si el primer término generado por el método de Newton, (x_1, λ_1), está en V, el método está bien definido y genera una sucesión (x_k, λ_k) que converge a (x_*, λ_*) superlinealmente ($\|x_{k+1} - x_*\| = o(\|x_k - x_*\|)$). Además, si f, c son de clase $C^{2,1}$, la convergencia es cuadrática ($\|x_{k+1} - x_*\| = O(\|x_k - x_*\|)$).*

El resultado principal es el siguiente:

Teorema 2. *Sean f, c de clase C^2 en una vecindad de un punto estacionario x_* del problema (P_{EI}) con multiplicador asociado λ_*. Suponga que se cumple (CE) y $(x_*, (\lambda_*)_{E\cup I_*^0})$ es punto estacionario del problema:*

$$(P'_E) : \begin{cases} min & f(x) \\ s.a & c_i(x) = 0, \qquad i \in E \cup I_*^0 \end{cases}$$

Considere el método PCS donde d_k es punto estacionario de norma mínima de $(QP_{EI}(k))$. Entonces existe una vecindad V de (x_*, λ_*) tal que si el primer término generado por el método PCS, (d_1, λ_1), está en V:

1. *PCS está bien definido y (x_k, λ_k) converge a (x_*, λ_*) superlinelamente.*
2. *Las restricciones activas de $(QP_{EI}(k))$ y (P_{EI}) son las mismas.*
3. *Si f, c son de clase $C^{2,1}$, la convergencia es cuadrática.*

Demostración. *La idea es mostrar que cerca de (x_*, λ_*), el punto estacionario de norma mínima de $(QP_{EI}(k))$ usando PCS y el punto estacionario del problema cuadrático asociado a (P'_E) en la k-ésima iteración usando Newton son iguales. De esta manera, usando el teorema 1, se concluye de inmediato el resultado.*

Suponagmos en adelante que (x_k, λ_k) **es lo suficientemente cercano a** (x_*, λ_*). *Como $(x_*, (\lambda_*)_{E\cup I_*^0})$ es punto estacionario de (P'_E), $\nabla c_{E\cup I_*^0}$ es sobreyectiva y el problema cuadrático asociado a (P'_E),*

$$(QP'_E(k)) : \begin{cases} min & \nabla f(x_k)^T \tilde{d} + 1/2 \tilde{d}^T L(x_k, \lambda_k)\tilde{d} \\ s.a & c_i(x_k) + \nabla c_i(x_k)^T \tilde{d} = 0, \qquad i \in E \cup I_*^0 \end{cases}$$

tiene solución única. Denotamos por $(\tilde{d}_k, (\tilde{\lambda}_{E\cup I_*^0})_k)$ a la solución de $(QP'_E(k))$ y por $(\tilde{d}_k, \tilde{\lambda}_k)$ a su extensión usando $(\tilde{\lambda}_k)_i = 0, i \in I \backslash I_*^0$.

Mostremos que $(\tilde{d}_k, \tilde{\lambda}_k)$ es punto estacionario de $(QP_{EI}(k))$. Como $(\tilde{d}_k, \tilde{\lambda}_k)$ satisface $(QP_{EI}(k))$, solo falta probar: (1) $c_i(x_k) + \nabla c_i(x_k)^T \tilde{d}_k \leq 0, \qquad i \in I \backslash I_*^0$, (2) $\tilde{\lambda}_i \geq 0, \qquad i \in I_*^0$.

De la convergencia de Newton, se tiene que $(x_{k+1}, \lambda_{k+1}) = (x_k + \tilde{d}_k, \tilde{\lambda}_k)$ con lo cual $\tilde{d}_k$ se puede escoger lo suficientemente chico como para satisfacer la desigualdad del punto 1. Para el punto 2, dado $i \in I_*^0$, se cumple $c_i(x_*) = 0$, y por (CE), $(\lambda_*)_i > 0$. Como $\tilde{\lambda}_k$ es cercano a λ_* (Newton), se puede escoger lo suficientemente cercano tal que $\tilde{\lambda}_i \geq 0$. Por lo tanto $(\tilde{d}_k, \tilde{\lambda}_k)$ es punto estacionario de $(QP_{EI}(k))$.

De esto se deduce que el método PCS está bien definido. Mostremos ahora que el par primal-dual (d_k, λ_k^{QP}) entregado por el método de Newton en cada iteración para el problema (P'_E), coincide con $(\tilde{d}_k, \tilde{\lambda}_k$ cuando (x_k, λ_k) esta suficientemente cerca de (x_*, λ_*). Si $i \in I \backslash I_*^0$, se puede escoger d_k suficientemente chico de manera que se cumpla $c_i(x_k) + \nabla c_i(x_k)^T d_k < 0$. Por (CE), se cumple $(\lambda_k^{QP})_i = 0$, y como $\tilde{\lambda}_i = 0$ para $i \in I \backslash I_*^0$, se cumple $(\lambda_k^{QP})_i = \tilde{\lambda}_i$. Como $(QP_{EI}(k))$ tiene solución única, no hay necesidad de revisar las restricciones de igualdad E, sólo falta probar

$$c_i(x_k) + \nabla c_i(x_k)^T d_k = 0, \qquad i \in I_*^0$$

Supongamos por contradicción que existe un índice $j \in I_*^0$ y una sucesión $(\overline{x}_k, \overline{\lambda}_k^{QP})$ convergente a (x_*, λ_*) tales que

$$c_j(\overline{x}_k) + \nabla c_j(\overline{x}_k)^T d_k < 0$$

Como d_k es chico, se puede escoger tal que $c_j(\overline{x}_k) < 0$ y $(\overline{\lambda}_k^{QP})_j = 0$ por (CE). Además:

$$\nabla f(\overline{x}_k) + L(\overline{x}_k, \overline{\lambda}_k) d_k + \sum_{i \in E\cup I_*^0 \backslash \{j\}} (\overline{\lambda}_k^{QP})_i \nabla c_i(\overline{x}_k) = 0$$

Como $\nabla f(\overline{x}_k) + L(\overline{x}_k, \overline{\lambda}_k)d_k \to \nabla f(x_*)$ *(pues* $d_k \to 0$*) y* $\nabla c_{E \cup I_*^0}$ *es sobreyectiva,* $\overline{\lambda}_k^{QP}$ *converge a algún valor* $\overline{\lambda}^{QP}$*. Tomando límite:*

$$\nabla f(x_*) + \sum_{i \in E \cup I_*^0 \setminus \{j\}} (\overline{\lambda}^{QP})_i \nabla c_i(x_*) = 0$$

Es decir, $\overline{\lambda}^{QP}$ *es multiplicador de* x_**, lo cual contradice la unicidad de* λ_**.*
Pro lo tanto, se tiene el resultado.$\square$