Capítulo 3

Funciones auxiliares

Tomemos en cuenta el problema donde no hay ecuaciones como restricciones:

$$(Q) \qquad min\,\{f(x) \mid x \in M\}$$

donde

$$M = \{x \in R^n \mid g_j(x) \geq 0 \;\; j \in J\}.$$

Asumamos, para el resto del trabajo que:

a) $\bar{x}$ es mínimo local de (Q).

b) Para cada vecindad de $\bar{x}$ existe un punto x^0 en el cual las restricciones son estrictamente satisfechas, esto es que $g_j(x^0) > 0,\; j \in J$.

c) $\bar{x}$ cumple la condición suficiente de optimalidad de segundo orden con $\bar{v} \in R^p$, $\bar{v} = (\bar{v}_1, \bar{v}_2, ..., \bar{v}_p)$ tal que

$$\bar{v}_j g_j(\bar{x}) = 0$$

tal que $\bar{v}_j > 0$ si $g_j(\bar{x}) = 0$ (propiedad conocida como complementaridad estricta). Si definimos $B(\bar{x})$ y $J_0(\bar{x})$ en forma análoga al capítulo anterior. Por la complementaridad estricta tenemos que $B(\bar{x}) = J_0(\bar{x})$.

Consideraremos una perturbación de la condición suficiente de optimalidad de segundo orden para que $\bar{x}$ sea un mínimo local.

d) Supongamos que para cualquier r las siguientes condiciones se cumplen en algún punto $[x(r), v(r)]$ cercano a $(\bar{x}, \bar{v})$:

$$g_j[x(r)] > 0, \qquad j \in J \tag{3.1}$$

$$v_j(r)g_j[x(r)] = r > 0, \qquad j \in J \tag{3.2}$$

$$v_j(r) \geq 0, \qquad j \in J \tag{3.3}$$

$$D_x\mathcal{L}[x(r), v(r)] = Df[x(r)] - \sum_{j \in J} v_j(r)Dg_j[x(r)] = 0 \tag{3.4}$$

y que para todo $z \in R^n$ tal que $Dg_j[x(r)]z = 0$, $j \in B(\bar{x})$ se cumple:

$$z^T\{D^2 f[x(r)] - \sum_{j \in J_0(\bar{x})} v_j(r)D^2 g_j[x(r)]\}z > 0. \tag{3.5}$$

Despejando $v_j(r)$ de (3.2) y sustituyendo en (3.4) obtenemos:

$$Df[x(r)] - \sum_{j \in j} \frac{r}{g_j[x(r)]} Dg_j[x(r)] = 0 \tag{3.6}$$

que es la derivada en $x(r)$ de la función conocida como **Logarithmic Penalty Function** ($L_r(x)$):

$$L_r(x) = f(x) - r\sum_{j \in J} \ln g_j(x).$$

La condición necesaria para que $x(r)$ sea un mínimo local de $L_r(x)$ queda satisfecha por (3.4).

La matríz de la segunda derivada de $L_r(x)$ sería:

$$D^2L_r(x) = D^2f(x) - \sum_{j\in J}\frac{r}{g_j(x)}D^2g_j(x) + \sum_{j\in J}Dg_j(x)\frac{r}{g_j^2(x)}D^Tg_j(x).$$

e) Supongamos ahora que cuando $r \to 0$, $x(r) \to \bar{x}$.

Esto implicaría que $\frac{r}{g_j[x(r)]}$ y $\frac{r}{g_j^2[x(r)]}$ tendieran ambas a cero para $j \notin J_0(\bar{x})$ pues:

$$\lim_{r\to 0} g_j[x(r)] = g_j(\bar{x}) > 0$$

cuando $j \notin J_0(\bar{x})$.

Entonces la matriz de la segunda derivada evaluada en $x(r)$ se reduce a:

$$D^2L_r[x(r)] = D^2f[x(r)] - \sum_{j\in J_0(\bar{x})}\frac{r}{g_j[x(r)]}D^2g_j[x(r)] + \sum_{j\in J_0(\bar{x})}Dg_j[x(r)]\frac{r}{g_j^2[x(r)]}D^Tg_j[x(r)].$$

Si consideramos aquellos vectores $z \in R^n$ tales que $Dg_j[x(r)]z = 0$, $j \in J_0(\bar{x})$, obtenemos entonces que:

$$z^T\{D^2L_r[x(r)]\}z > 0$$

al aplicar el resultado (3.5) del supuesto d).

Por otro lado, podemos decir que:

$$\lim_{r\to 0}\frac{r}{g_j^2[x(r)]} = +\infty$$

para $j \in J_0(\bar{x})$. Ya que, por la parte (3.2) del supuesto d) podemos despejar r y expresar nuestro límite de la siguiente manera:

$$\lim_{r\to 0}\frac{v_j(r)g_j[x(r)]}{g_j^2[x(r)]} = \lim_{r\to 0}\frac{v_j(r)}{g_j[x(r)]}$$

f) Si $v_j(r) \to \bar{v}_j$, $j \in J_0(\bar{x})$ cuando $r \to 0$, tendríamos:

$$\lim_{r\to 0} \frac{v_j(r)}{g_j[x(r)]} = \frac{\bar{v}_j}{g_j(\bar{x})} = \frac{\bar{v}_j}{0} = +\infty.$$

Por lo tanto, para toda $z \in R^n$ tal que $Dg_j[x(r)]z \neq 0$:

$$z^T\{D^2 L_r[x(r)]\}z > 0. \tag{3.7}$$

Ya tenemos la condición suficiente de optimalidad de segundo orden para que $L_r(x)$ tenga un mínimo local en $x(r)$ mediante la perturbación de la condición suficiente de optimalidad de segundo orden que se cumple en $\bar{x}$.

Sin embargo, hasta aquí no se ha demostrado la existencia de un punto $[x(r), v(r)]$ que cumpla con las condiciones anteriormente exigidas.

Una función similar a $L_r(x)$ se puede derivar del supuesto:

g) Sea $v_j(r) = \lambda_j^2$, $j \in J$. Si $v_j(r)g_j[x(r)] = 0$ entonces, $\lambda_j g_j[x(r)] = 0$. Buscamos λ_j en forma que:

$$\lambda_j g_j[x(r)] = r_\lambda > 0, \quad j \in J.$$

Dado que $\lambda_j = \frac{r_\lambda}{g_j[x(r)]}$ y $v_j(r) = \lambda_j^2$, entonces el enunciado (3.4) del supuesto d) puede ser expresado de la siguiente manera:

$$Df[x(r)] - \sum_{j\in J} \frac{r_\lambda^2}{g_j^2[x(r)]} Dg_j[x(r)] = 0.$$

Esto indica que la derivada de la función:

$$P_{r_\lambda}(x) = f(x) + r_\lambda^2 \sum_{j\in J} \frac{1}{g_j(x)}$$

es cero en $x(r)$.

Obtenemos la segunda derivada de la función P_{r_λ}:

$$D^2P_{r_\lambda}(x) = D^2f(x) - \sum_{j\in J}\frac{r_\lambda^2}{g_j^2(x)}D^2g_j(x) + \sum_{j\in J}Dg_j(x)\frac{2r_\lambda^2}{g_j^3(x)}D^Tg_j(x)$$

y al igual que en $L_r(x)$, el cumplimiento de la condición suficiente de optimalidad de segundo orden para que $\bar{x}$ sea un mínimo local de $f(x)$ implica que $x(r)$, con r pequeña, satisfaga la condición suficiente de optimalidad de segundo orden y sea un mínimo local de la función no restringida $P_{r_\lambda}(x)$.

Los algoritmos de minimización de puntos interiores se basan en encontrar mínimos locales no restringidos para $L_r(x)$ o $P_{r_\lambda}(x)$ en la región donde las restricciones del problema (Q) son estrictamente satisfechas, o sea en:

$$M^0 = \{x \in R^n \mid g_j(x) > 0 \quad j \in J\}.$$

La idea es que mientras r es más pequeña, nuestro valor $x(r)$ se aproximará cada vez más al mínimo local restringido de la función original del problema (Q).

Las técnicas de minimización que usan a las funciones $L_r(x)$ y $P_{r_\lambda}(x)$ pertenecen a la clase de los **algoritmos de minimización de puntos interiores**, pues los puntos $x(r)$ que emplean dichas técnicas, están dentro de M^0 que es el interior de la región M.

En el capítulo 4 explicaré con mayor detalle cómo se relacionan las funciones auxiliares $L_r(x)$ y $P_{r_\lambda}(x)$ con los algoritmos de minimización de puntos interiores.