Capítulo 2

Conceptos generales

Denotemos con R^n al espacio euclideano de dimensión n y R el espacio euclideano de dimensión 1. Sea $C^k(R^n, R)$ el conjunto de funciones $\gamma\colon R^n \to R$ que son k veces continuamente diferenciables.

Definición: Sea el **conjunto factible** M definido de la siguiente manera:

$$M = \{x \in R^n \mid h_i(x) = 0, \quad i \in I = \{1, ..., m\}$$
$$g_j(x) \geq 0, \quad j \in J = \{1, ..., p\}\}$$

con $h_i, g_j \in C^1(R^n, R)$, $i \in I$, $j \in J$, funciones a las que se les conoce como restricciones.

Definición: Llamaremos **punto factible** a toda $x \in M$.

Definición: Sean $M \subset R^n$ y $f\colon R^n \to R$. Un punto $\bar{x} \in M$ es un **mínimo local** de (P) si existe una vecindad $\mathcal{U}(\bar{x})$ de $\bar{x}$ tal que $f(x) \geq f(\bar{x})$ para toda $x \in \mathcal{U}(\bar{x}) \cap M$.

Definición: Sean M y f como en la definición anterior. Un punto $\bar{x} \in M$ es **mínimo global** de (P) si $f(x) \geq f(\bar{x})$ para toda $x \in M$.

Dada $f \in C^1(R^n, R)$, el problema (P) de encontrar mínimos locales de f en $M \subset R^n$ se expresa de la siguiente manera:

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

Un mínimo local $\bar{x}$ de (P) debe estar en M y entonces $g_j(\bar{x}) \geq 0$, $j \in J$. A partir de esto podemos construir para $\bar{x} \in M$ un conjunto $J_0(\bar{x}) = \{j \in J \mid g_j(\bar{x}) = 0\}$ al cual llamaremos conjunto de índices activos de $\bar{x} \in M$. Ahora es establecemos una condición necesaria para que $\bar{x}$ sea un mínimo local de (P).

Sea $x \in R^n$ un vector columna definido por medio de sus componentes $x_1, ..., x_n$. Dada una función $z \in C^1(R^n, R)$, $Dz(\bar{x})$ será el vector fila $(\frac{\partial z(\bar{x})}{\partial x_1}, \frac{\partial z(\bar{x})}{\partial x_2}, ..., \frac{\partial z(\bar{x})}{\partial x_n})$. Y sea $D^2z(\bar{x}) = D[Dz(\bar{x})]$ la matriz

$$\mathbf{D^2z(\bar{x})} = \begin{pmatrix} \frac{\partial^2 z(\bar{x})}{\partial x_1 x_1} & \frac{\partial^2 z(\bar{x})}{\partial x_1 x_2} & \cdots & \frac{\partial^2 z(\bar{x})}{\partial x_1 x_n} \\ \frac{\partial^2 z(\bar{x})}{\partial x_2 x_1} & \frac{\partial^2 z(\bar{x})}{\partial x_2 x_2} & \cdots & \frac{\partial^2 z(\bar{x})}{\partial x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 z(\bar{x})}{\partial x_n x_1} & \frac{\partial^2 z(\bar{x})}{\partial x_n x_2} & \cdots & \frac{\partial^2 z(\bar{x})}{\partial x_n x_n} \end{pmatrix}$$

Condición necesaria de optimalidad de primer orden:

Si $\bar{x}$ es un mínimo local de (P), entonces existen coeficientes reales $\alpha \geq 0$, λ_i, $i \in I$ y $\mu_j \geq 0$, $j \in J_0(\bar{x})$, no todos cero, tales que:

$$\alpha Df(\bar{x}) \; - \; \sum_{i \in I} \lambda_i Dh_i(\bar{x}) \; - \sum_{j \in J_0(\bar{x})} \mu_j Dg_j(\bar{x}) = 0 \qquad (2.1)$$

La prueba se halla en [2] página 19.

Una vez que tenemos un mínimo local $\bar{x}$ de (P), este puede poseer ciertas cualidades con respecto a las funciones h_i, $i \in I$, g_j, $j \in J$ de M. Llamamos a dichas cualidades **constraint qualifications**. Nos enfocaremos en dos constraint qualifications: **Linear Independence Constraint Qualification** (LICQ) y **Mangasarian-Fromovitz Constraint Qualification** (MFCQ).

Linear Independence Constraint Qualification (LICQ):

Se dice que (LICQ) se cumple en $\bar{x} \in M$ si y solo si los gradientes $\{ Dh_i(\bar{x}),\ Dg_j(\bar{x}) \mid i \in I,\ j \in J_0(\bar{x})\}$ son linealmente independientes.

Es notorio que el hecho de que (LICQ) se cumpla en $\bar{x}$ es una condición muy fuerte. La siguiente constraint qualification exige menos, como veremos en el teorema 1.

Mangasarian-Fromovitz Constraint Qualification (MFCQ):

Se dice que (MFCQ) se cumple en $\bar{x} \in M$ si y solo si:

i) $Dh_i(\bar{x}),\ i \in I$ son linealmente independientes.

ii) Existe un vector $w \in R^n$ tal que:

$$Dh_i(\bar{x})w = 0, \quad i \in I;$$

$$Dg_j(\bar{x})w > 0, \quad j \in J_0(\bar{x}).$$

Para futuras referencias, es conveniente enunciar el siguiente lema.

Lema 1 ([3], teorema 10.4.7):

Dados los vectores $a_i,\ b_j,\ c_k \in R^n$, con $i = 1, ..., m_a$, $j = 1, ..., m_b$ y $k = 1, ..., m_c$. Una y solo una de las siguientes posibilidades se cumple:

i) El sistema

$$\xi^T a_i < 0\ ,\quad i = 1, ..., m_a\ ,\quad m_a \geq 1$$

$$\xi^T b_j \leq 0\ ,\quad j = 1, ..., m_b\ ,$$

$$\xi^T c_k = 0\ ,\quad k = 1, ..., m_c\ .$$

es soluble con $\xi \in R^n$.

ii) Existen números reales $\alpha_i \geq 0$, $i = 1, ..., m_a$, $\beta_j \geq 0$, $j = 1, ..., m_b$ y γ_k, $k = 1, ..., m_c$, de forma que:

$$\sum_{i=1}^{m_a} \alpha_i a_i + \sum_{j=1}^{m_b} \beta_j b_j + \sum_{k=1}^{m_c} \gamma_k c_k = 0$$

con

$$\sum_{i=1}^{m_a} \alpha_i > 0$$

Ahora es posible abordar los siguientes teoremas.

Teorema 1:

Sea $\bar{x} \in M$. Si (LICQ) se cumple en $\bar{x}$, entonces (MFCQ) se cumple en $\bar{x}$.

Demostración:

Si (LICQ) se cumple en $\bar{x}$, entonces $\{Dh_i(\bar{x}), Dg_j(\bar{x}) \mid i \in I,\ j \in J_0(\bar{x})\}$ son linealmente independientes. Luego $\{Dh_i(\bar{x}) \mid i \in I\}$ deben ser linealmente independientes, de lo contrario existiría un conjunto $\{u_i \in R \mid i \in I\}$ donde no todos los valores son cero, en forma que:

$$\sum_{i \in I} u_i Dh_i(\bar{x}) = 0$$

de ahí que

$$\sum_{i \in I} u_i Dh_i(\bar{x}) + \sum_{j \in J_0} 0 \cdot Dg_j(\bar{x}) = 0.$$

Lo cual indicaría que los gradientes $\{Dh_i(\bar{x}), Dg_j(\bar{x}) \mid i \in I,\ j \in J_0(\bar{x})\}$ son linealmente dependientes pero entonces caeríamos en contradicción.

Entonces $\{Dh_i(\bar{x}) \mid i \in I\}$ deben ser linealmente independientes, cumpliéndose $i)$ de (MFCQ).

Para continuar con la demostración, es necesario resaltar antes que si los vectores $w_k \in R^n,\ \ k \in K = \{1, ..., m_w\}$ son linealmente independientes, entonces también los vectores $-w_k$, $k \in K$ deben ser linealmente independientes, de lo contrario existiría un conjunto de números reales, no todos cero $\{c_k\}$ tales que

$$\sum_{k \in K} c_k(-w_k) = 0$$

o lo que es lo mismo

$$\sum_{K \in} -c_k w_k = 0$$

pero como $-c_k,\ \ k \in K$ es un conjunto de reales no todos son cero, la expresión anterior implicaría que los vectores w_k, $k \in K$ son linealmente dependientes, llegando a una contradicción. Por lo tanto, si $w_k \in R^n,\ \ k \in K$ son linealmente independientes, $-w_k$, $k \in K$ deben ser también linealmente independientes.

Sigamos con la demostración.

Sabemos que al ser $\{Dh_i(\bar{x}), Dg_j(\bar{x}) \mid i \in I,\ j \in J_0(\bar{x})\}$ vectores linealmente independientes, entonces también los vectores $\{-Dh_i(\bar{x}), -Dg_j(\bar{x}) \mid i \in I,\ j \in J_0(\bar{x})\}$ son linealmente independientes, por lo tanto no se puede dar el caso de que existan números reales $u_i,\ i \in I,\ v_j \geq 0,\ j \in J_0(\bar{x})$ con

$$\sum_{j \in J_0(\bar{x})} v_j > 0$$

de modo que:

$$\sum_{j \in J_0(\bar{x})} v_j[-Dg_j(\bar{x})] + \sum_{i \in I} u_i[-Dh_i(\bar{x})] = 0.$$

Entonces, por el lema 1, el sistema

$$[-Dh_i(\bar{x})]\xi = 0, \quad i \in I$$

$$[-Dg_j(\bar{x})]\xi < 0, \quad j \in J_0(\bar{x})$$

debe ser soluble con $\xi \in R^n$. Multiplicando en ambos lados de las expresiones anteriores por -1, tenemos que el sistema

$$Dh_i(\bar{x})\xi = 0, \quad i \in I$$

$$Dg_j(\bar{x})\xi > 0, \quad j \in J_0(\bar{x})$$

es, obviamente, soluble; cumpliéndose así la parte *ii*) de (MFCQ) en $\bar{x}$.

Entonces, si (LICQ) se cumple en $\bar{x}$, entonces (MFCQ) se cumple en $\bar{x}$.

Q.E.D.

Pasemos al siguiente teorema que asocia a (MFCQ) con un mínimo local $\bar{x} \in M$.

Teorema 2:

Si se cumple (MFCQ) en $\bar{x}$ y $\bar{x}$ es un mínimo local de (P), entonces existen reales λ_i, $i \in I$ y $\mu_j \geq 0$, $j \in J_0(\bar{x})$ tales que:

$$Df(\bar{x}) - \sum_{i \in I} \lambda_i Dh_i(\bar{x}) - \sum_{j \in J_0(\bar{x})} \mu_j Dg_j(\bar{x}) = 0$$

Demostración:

Al ser $\bar{x}$ mínimo local de (P) sabemos, por la condición necesaria de optimalidad de primer orden, que existen valores reales $\alpha \geq 0$, λ_i, $i \in I$ y $\mu_j \geq 0$, $j \in J_0(\bar{x})$, no todos cero, tales que la ecuación (1) se cumple. Para demostrar el teorema, basta probar que $\alpha \neq 0$.

Supongamos entonces que $\alpha = 0$. Entonces tendríamos que:

$$-\sum_{i \in I} \lambda_i Dh_i(\bar{x}) - \sum_{j \in J_0(\bar{x})} \mu_j Dg_j(\bar{x}) = 0$$

que es equivalente a la expresión

$$\sum_{i \in I} (-\lambda_i) Dh_i(\bar{x}) + \sum_{j \in J_0(\bar{x})} \mu_j [-Dg_j(\bar{x})] = 0 \qquad (2.2)$$

con

$$\sum_{j \in J_0(\bar{x})} \mu_j > 0. \qquad (2.3)$$

La desigualdad (2.3) es cierta porque si

$$\sum_{j \in J_0(\bar{x})} \mu_j = 0$$

entonces $\mu_j = 0, \ \ j \in J_0(\bar{x})$. Y tendríamos que

$$\sum_{i \in I} (-\lambda_i) Dh_i(\bar{x}) = 0$$

pero como ya $\alpha = 0$ y cada $\mu_j = 0, \ \ j \in J_0(\bar{x})$, entonces no pueden ser cero todos los multiplicadores $-\lambda_i, \ \ i \in I$. Luego los vectores $\{Dh_i(\bar{x}) \mid i \in I\}$ serían linealmente dependientes, lo que es una contradicción pues hemos asumido que (MFCQ) se cumple en $\bar{x}$. De manera que

$$\sum_{j \in J_0(\bar{x})} \mu_j > 0.$$

Por el lema 1, al cumplirse (2.3) y la ecuación (2.2), el sistema:

$$Dh_i(\bar{x})\xi = 0, \quad i \in I;$$

$$[-Dg_j(\bar{x})]\xi < 0, \ \ j \in J_0(\bar{x})$$

no sería soluble, y entonces tampoco lo sería el sistema

$$Dh_i(\bar{x})\xi = 0, \quad i \in I;$$

$$Dg_j(\bar{x})\xi > 0, \ \ j \in J_0(\bar{x})$$

lo cual es contradictorio, dado que hemos supuesto que (MFCQ) se cumple en $\bar{x}$.

Dado esto, es imposible que $\alpha = 0$. Luego $\alpha > 0$.

Así que la ecuación

$$\alpha Df(\bar{x}) \ - \ \sum_{i \in I} \lambda_i Dh_i(\bar{x}) \ - \ \sum_{j \in J_0(\bar{x})} \mu_j Dg_j(\bar{x}) = 0$$

puede ser dividida entre α. Resultando de esto:

$$Df(\bar{x}) - \sum_{i \in I} \lambda_i' Dh_i(\bar{x}) - \sum_{j \in J_0(\bar{x})} \mu_j' Dg_j(\bar{x}) = 0$$

donde $\lambda_i' = \frac{\lambda_i}{\alpha} \in R, \ i \in I$ y $\mu_j' = \frac{\mu_j}{\alpha} \geq 0, \ \ j \in J_0(\bar{x})$.

Q.E.D.

Daré ahora unos ejemplos para los cuales únicamente emplearé las funciones $g_j\colon R^2 \to R,\ \ j \in J$.

Ejemplo 2.1:

Consideremos el conjunto factible:

$$M = \{(x,y) \in R^2 \mid g_1(x,y) = y - x \geq 0, \quad g_2(x,y) = x + y \geq 0\}$$

En la Figura 2.1 se pueden observar las gráficas de las siguientes restricciones en una vecindad del punto $(\bar{x},\bar{y}) = \binom{0}{0}$.

En este caso:

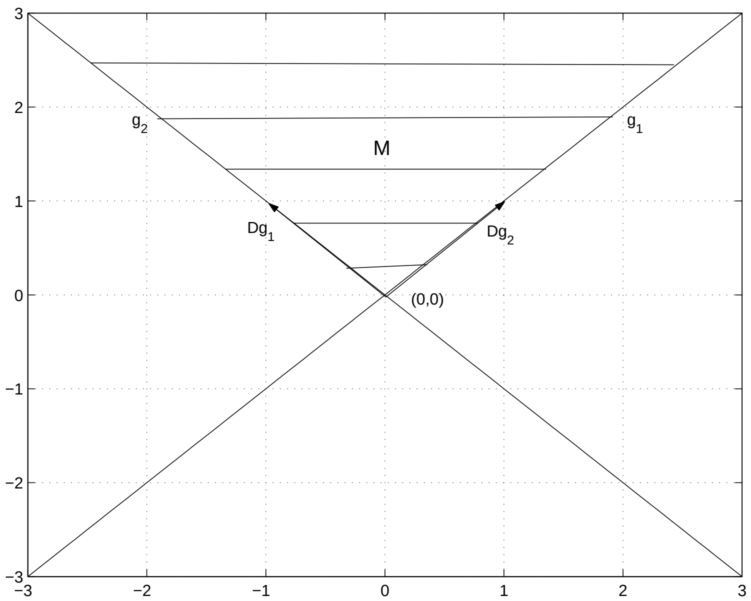

Figura 2.1: La región factible está delimitada por dos rectas.

$Dg_1(\bar{x},\bar{y}) = (-1,1)$

$Dg_2(\bar{x},\bar{y}) = (1,1)$

Como $Dg_1(\bar{x},\bar{y})$ y $Dg_2(\bar{x},\bar{y})$ son linealmente independientes, además de que existe

$\binom{0}{1} = w$ tal que:

$$Dg_j(\bar{x}, \bar{y})w > 0, \quad j \in J_0(\bar{x})$$

Entonces tanto (LICQ) como (MFCQ) se cumplen en $(\bar{x}, \bar{y}) = \binom{0}{0}$.

Ejemplo 2.2:

Definamos nuestro conjunto factible de la siguiente manera:

$$M = \{(x, y) \mid g_1(x, y) = 2y - x^2 - y^2 \geq 0, \;\; g_2(x, y) = y - x^2 \geq 0\}$$

Como se puede apreciar en la Figura 4.1, nos enfocamos en una vecindad del punto $(\bar{x}, \bar{y}) = \binom{0}{0}$.

En este caso:

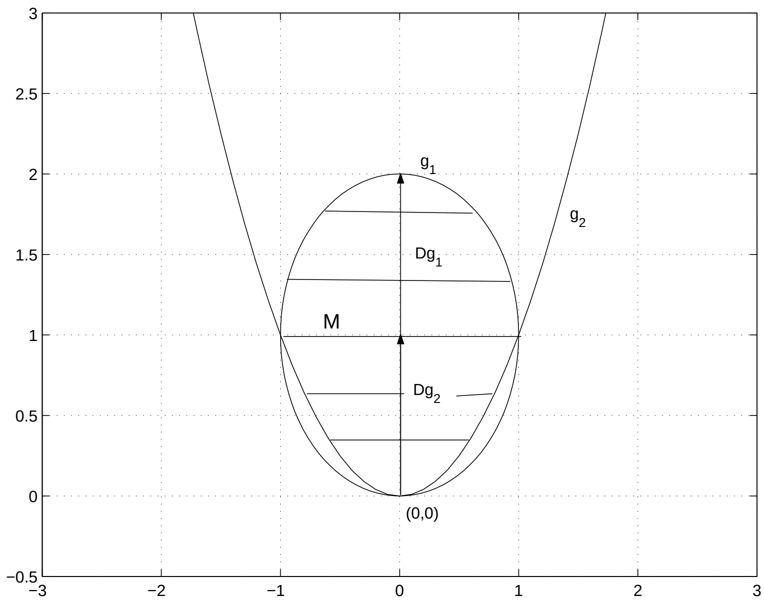

Figura 2.2: En este ejemplo no se cumple (LICQ) en $\binom{0}{0}$.

Los gradientes de las funciones g_1 y g_2 son:

$Dg_1(x,y) = (-2x, 2-2y)$

$Dg_2(x,y) = (-2x, 1)$.

Evaluamos en $(\bar{x},\bar{y}) = \binom{0}{0}$:

$Dg_1(\bar{x},\bar{y}) = (0,2)$

$Dg_2(\bar{x},\bar{y}) = (0,1)$.

Estos vectores son linealmente dependientes. No obstante, si existe $w = \binom{0}{1}$ tal que:

$$Dg_j(\bar{x},\bar{y})w > 0, \;\; j \in J_0(\bar{x})$$

Por lo tanto en $(\bar{x},\bar{y})$ se cumple (MFCQ) pero no (LICQ).

Ejemplo 2.3:

Consideremos al conjunto factible de la siguiente forma:

$$M = \{(x,y) \;\mid\; g_1(x,y) = y - x^2 \geq 0, \;\; g_2(x,y) = -x^2 - y \geq 0\}$$

Veamos Si (LICQ) y (MFCQ) se cumplen para $(\bar{x},\bar{y}) = \binom{0}{0}$. Ver la Figura 2.3.

Los gradientes de g_1 y g_2 evaluados en $\binom{0}{0}$ son:

$Dg_1(\bar{x},\bar{y}) = (0,1)$

$Dg_2(\bar{x},\bar{y}) = (0,-1)$

Como $Dg_1(\bar{x},\bar{y})$ y $Dg_2(\bar{x},\bar{y})$ son linealmente dependientes y no existe $w \in R^2$ tal que $Dg_j(\bar{x},\bar{y})w > 0$, $j = 1,2$. Entonces no se cumplen (LICQ) ni (MFCQ) en $\bar{x} = \binom{0}{0}$.

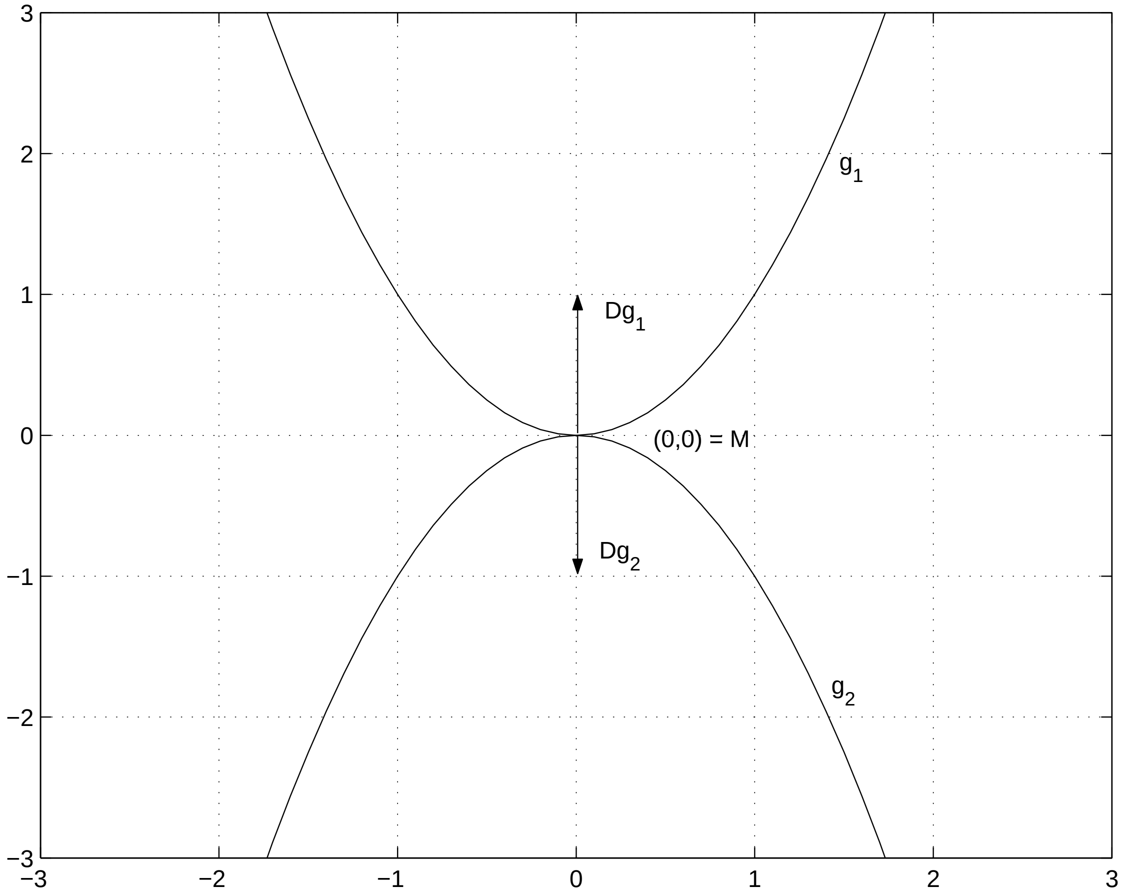

Figura 2.3: En este caso, M consiste en un solo punto.

Para hacer más sencillas las fórmulas en el futuro, definiré la función Lagrangiana $\mathcal{L}(x,a,b)$ asociada al problema (P) donde $a \in R^m$ y $b \in R^p$:

$$\mathcal{L}(x,a,b) = f(x) - \sum_{i\in I} a_i h_i(x) - \sum_{j\in J} b_j g_j(x).$$

Y sean $D_x\mathcal{L}(x,a,b)$ y $D^2_{xx}\mathcal{L}(x,a,b)$ tales que:

$$D_x\mathcal{L}(x,a,b) = Df(x) - \sum_{i\in I} a_i Dh_i(x) - \sum_{j\in J} b_j Dg_j(x)$$

$$D^2_{xx}\mathcal{L}(x,a,b) = D^2 f(x) - \sum_{i\in I} a_i D^2 h_i(x) - \sum_{j\in J} b_j D^2 g_j(x).$$

Ya con esta definición, podemos pasar al teorema 3.

Teorema 3 (Condición suficiente de optimalidad de segundo orden):

Sean las funciones del problema (P) $f,\ h_i,\ g_j \in C^2(R^n, R)$, $i \in I$, $j \in J$. Supongamos que existen vectores $\bar{v} \in R^p$ y $\bar{u} \in R^m$ tales que $(\bar{x}, \bar{v}, \bar{u})$ cumplan:

$$g_j(\bar{x}) \geq 0, \qquad j \in J,$$

$$h_i(\bar{x}) = 0, \qquad i \in I,$$

$$\bar{v}_j g_j(\bar{x}) = 0, \quad j \in J,$$

$$\bar{v}_j \geq 0, \qquad j \in J,$$

$$D_x\mathcal{L}(\bar{x}, \bar{v}, \bar{u}) = 0$$

y que para cada $z \in R^n$ diferente de cero con:

i) $Dg_j(\bar{x})z = 0, \;\; j \in B(\bar{x})$ donde $B(\bar{x}) = \{j \in J \mid \bar{v}_j > 0\}$,

ii) $Dg_j(\bar{x})z \geq 0, \;\; j \in J_0(\bar{x}) \setminus B(\bar{x})$ y

iii) $Dh_i(\bar{x})z = 0, \;\; i \in I$,

tenemos que:

$$z^T[D^2_{xx}\mathcal{L}(\bar{x}, \bar{v}, \bar{u})]z > 0.$$

Entonces $\bar{x}$ es un mínimo local de (P).

La prueba de este teorema se halla en ([5], pp. 30, 31).

En el siguiente capítulo nos enfocaremos en los problemas de optimización cuyas restricciones no contienen funciones del tipo $h_i(x) = 0$, $i \in I$ y modificaremos las condiciones del teorema 3, para obtener en base a esto ciertas funciones auxiliares que nos ayudan a traducir el problema original a uno sin restricciones.