Equality Constrained Problems/Sufficiency Conditions

Convexification Using Augmented Lagrangians

Proof of the Sufficiency Conditions

Sensitivity

Equality constrained problem

minimize \( f(x) \)

subject to \( h_i(x) = 0, \quad i = 1, \ldots, m. \)

where \( f : \mathbb{R}^n \rightarrow \mathbb{R}, \; h_i : \mathbb{R}^n \rightarrow \mathbb{R}, \) are continuously differentiable. To obtain sufficiency conditions, assume that \( f \) and \( h_i \) are \textit{twice} continuously differentiable.
SUFFICIENCY CONDITIONS

Second Order Sufﬁciency Conditions: Let \( x^* \in \mathbb{R}^n \) and \( \lambda^* \in \mathbb{R}^m \) satisfy

\[
\nabla_x L(x^*, \lambda^*) = 0, \quad \nabla_\lambda L(x^*, \lambda^*) = 0,
\]

\[
y' \nabla_{xx}^2 L(x^*, \lambda^*) y > 0, \quad \forall y \neq 0 \text{ with } \nabla h(x^*)' y = 0.
\]

Then \( x^* \) is a strict local minimum.

Example: Minimize \(- (x_1 x_2 + x_2 x_3 + x_1 x_3)\) subject to \(x_1 + x_2 + x_3 = 3\). We have that \( x_1^* = x_2^* = x_3^* = 1 \) and \( \lambda^* = 2 \) satisfy the 1st order conditions. Also

\[
\nabla_{xx}^2 L(x^*, \lambda^*) = \begin{pmatrix} 0 & -1 & -1 \\
-1 & 0 & -1 \\
-1 & -1 & 0 \end{pmatrix}.
\]

We have for all \( y \neq 0 \) with \( \nabla h(x^*)' y = 0 \) or \( y_1 + y_2 + y_3 = 0 \),

\[
y' \nabla_{xx}^2 L(x^*, \lambda^*) y = -y_1(y_2 + y_3) - y_2(y_1 + y_3) - y_3(y_1 + y_2)
= y_1^2 + y_2^2 + y_3^2 > 0.
\]

Hence, \( x^* \) is a strict local minimum.
Lemma: Let $P$ and $Q$ be two symmetric matrices. Assume that $Q \geq 0$ and $P > 0$ on the nullspace of $Q$, i.e., $x'Px > 0$ for all $x \neq 0$ with $x'Qx = 0$. Then there exists a scalar $\bar{c}$ such that

$$P + cQ : \text{positive definite, \quad } \forall \ c > \bar{c}.$$  

Proof: Assume the contrary. Then for every $k$, there exists a vector $x^k$ with $\|x^k\| = 1$ such that

$$x^k'Px^k + kx^k'Qx^k \leq 0.$$  

Consider a subsequence $\{x^k\}_{k \in K}$ converging to some $\bar{x}$ with $\|\bar{x}\| = 1$. Taking the limit superior,

$$\bar{x}'P\bar{x} + \limsup_{k \to \infty, k \in K} (kx^k'Qx^k) \leq 0. \quad (*)$$  

We have $x^k'Qx^k \geq 0$ (since $Q \geq 0$), so $\{x^k'Qx^k\}_{k \in K} \to 0$. Therefore, $\bar{x}'Q\bar{x} = 0$ and using the hypothesis, $\bar{x}'P\bar{x} > 0$. This contradicts $(*)$. 

A BASIC LEMMA
PROOF OF SUFFICIENCY CONDITIONS

Consider the augmented Lagrangian function

\[ L_c(x, \lambda) = f(x) + \lambda' h(x) + \frac{c}{2} \|h(x)\|^2, \]

where \( c \) is a scalar. We have

\[ \nabla x L_c(x, \lambda) = \nabla x L(x, \tilde{\lambda}), \]

\[ \nabla^2 x x L_c(x, \lambda) = \nabla^2 x x L(x, \tilde{\lambda}) + c\nabla h(x)\nabla h(x)' \]

where \( \tilde{\lambda} = \lambda + ch(x) \). If \((x^*, \lambda^*)\) satisfy the suff. conditions, we have using the lemma,

\[ \nabla x L_c(x^*, \lambda^*) = 0, \quad \nabla^2 x x L_c(x^*, \lambda^*) > 0, \]

for suff. large \( c \). Hence for some \( \gamma > 0, \epsilon > 0, \)

\[ L_c(x, \lambda^*) \geq L_c(x^*, \lambda^*) + \frac{\gamma}{2} \|x - x^*\|^2, \quad \text{if} \ |x - x^*| < \epsilon. \]

Since \( L_c(x, \lambda^*) = f(x) \) when \( h(x) = 0, \)

\[ f(x) \geq f(x^*) + \frac{\gamma}{2} \|x - x^*\|^2, \quad \text{if} \ h(x) = 0, \ |x - x^*| < \epsilon. \]
Sensitivity theorem for the problem \( \min_{a'x=b} f(x) \). If \( b \) is changed to \( b + \Delta b \), the minimum \( x^* \) will change to \( x^* + \Delta x \). Since \( b + \Delta b = a'(x^* + \Delta x) = a'x^* + a'\Delta x = b + a'\Delta x \), we have \( a'\Delta x = \Delta b \). Using the condition \( \nabla f(x^*) = -\lambda^* a \),

\[
\Delta \text{cost} = f(x^* + \Delta x) - f(x^*) = \nabla f(x^*)'\Delta x + o(\|\Delta x\|)
\]

\[
= -\lambda^* a'\Delta x + o(\|\Delta x\|)
\]

Thus \( \Delta \text{cost} = -\lambda^* \Delta b + o(\|\Delta x\|) \), so up to first order

\[
\lambda^* = -\frac{\Delta \text{cost}}{\Delta b}.
\]

For multiple constraints \( a'_i x = b_i, \ i = 1, \ldots, n \), we have

\[
\Delta \text{cost} = -\sum_{i=1}^{m} \lambda_i^* \Delta b_i + o(\|\Delta x\|).
\]
SENSITIVITY THEOREM

Sensitivity Theorem: Consider the family of problems

$$\min_{h(x)=u} f(x) \quad (*)$$

parameterized by $u \in \mathbb{R}^m$. Assume that for $u = 0$, this problem has a local minimum $x^*$, which is regular and together with its unique Lagrange multiplier $\lambda^*$ satisfies the sufficiency conditions.

Then there exists an open sphere $S$ centered at $u = 0$ such that for every $u \in S$, there is an $x(u)$ and a $\lambda(u)$, which are a local minimum-Lagrange multiplier pair of problem $(*)$. Furthermore, $x(\cdot)$ and $\lambda(\cdot)$ are continuously differentiable within $S$ and we have $x(0) = x^*$, $\lambda(0) = \lambda^*$. In addition,

$$\nabla p(u) = -\lambda(u), \quad \forall \ u \in S$$

where $p(u)$ is the primal function

$$p(u) = f(x(u)).$$
EXAMPLE

Illustration of the primal function \( p(u) = f(x(u)) \) for the two-dimensional problem

\[
\begin{align*}
\text{minimize} & \quad f(x) = \frac{1}{2} (x_1^2 - x_2^2) - x_2 \\
\text{subject to} & \quad h(x) = x_2 = 0.
\end{align*}
\]

Here,

\[
p(u) = \min_{h(x)=u} f(x) = -\frac{1}{2} u^2 - u
\]

and \( \lambda^* = -\nabla p(0) = 1 \), consistently with the sensitivity theorem.

- Need for regularity of \( x^* \): Change constraint to \( h(x) = x_2^2 = 0 \). Then \( p(u) = -u/2 - \sqrt{u} \) for \( u \geq 0 \) and is undefined for \( u < 0 \).
PROOF OUTLINE OF SENSITIVITY THEOREM

Apply implicit function theorem to the system

\[ \nabla f(x) + \nabla h(x) \lambda = 0, \quad h(x) = u. \]

For \( u = 0 \) the system has the solution \((x^*, \lambda^*)\), and the corresponding \((n + m) \times (n + m)\) Jacobian

\[
J = \left( \begin{array}{c}
\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*) \\
\n\nabla h(x^*)
\end{array} \right)
\]

is shown nonsingular using the sufficiency conditions. Hence, for all \( u \) in some open sphere \( S \) centered at \( u = 0 \), there exist \( x(u) \) and \( \lambda(u) \) such that \( x(0) = x^* \), \( \lambda(0) = \lambda^* \), the functions \( x(\cdot) \) and \( \lambda(\cdot) \) are continuously differentiable, and

\[
\nabla f(x(u)) + \nabla h(x(u)) \lambda(u) = 0, \quad h(x(u)) = u.
\]

For \( u \) close to \( u = 0 \), using the sufficiency conditions, \( x(u) \) and \( \lambda(u) \) are a local minimum-Lagrange multiplier pair for the problem \( \min_{h(x)=u} f(x) \).

To derive \( \nabla p(u) \), differentiate \( h(x(u)) = u \), to obtain \( I = \nabla x(u) \nabla h(x(u)) \), and combine with the relations \( \nabla x(u) \nabla f(x(u)) + \nabla x(u) \nabla h(x(u)) \lambda(u) = 0 \) and \( \nabla p(u) = \nabla u \{ f(x(u)) \} = \nabla x(u) \nabla f(x(u)) \).