1 Proof of the Second Order Sufficient Conditions (June 11)

1.1 Second Order Sufficient Conditions

Theorem 1.1. *(Second order sufficient conditions for a minimizer under inequality constraints) Suppose* $\Omega \subseteq \mathbb{R}^n$ *is open and* $f, h_1, \ldots, h_k, g_1, \ldots, g_l \in C^2(\Omega)$*. Consider the minimization problem*

$$\begin{aligned} &\textit{minimize } f(x) \\ &\textit{subject to } h_1(x) = \cdots = h_k(x) = 0 \\ &\qquad\qquad g_1(x) \leq 0, \ldots, g_l(x) \leq 0 \end{aligned}$$

Suppose x_0 *is a feasible point of the constraints. If the following three conditions are satisfied:*

1. *There exist* $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ *and* $\mu_1, \ldots, \mu_l \geq 0$ *such that*

$$\nabla f(x_0) + \sum_i \lambda_i \nabla h_i(x_0) + \sum_j \mu_j \nabla g_j(x_0) = 0,$$

2. $\mu_j g_j(x_0) = 0$ *for each* j*.*

3. *The matrix*

$$L(x_0) = \nabla^2 f(x_0) + \sum_i \lambda_i \nabla^2 h_i(x_0) + \sum_j \mu_j \nabla^2 g_j(x_0)$$

   is positive definite on the tangent space to the "strongly active constraints" at x_0*. That is, it is positive definite on the space*

$$\widetilde{\widetilde{T_{x_0}}} = \{v \in \mathbb{R}^n : \nabla h_i(x_0) = 0 \textit{ for all } i, \textit{ and } \nabla g_j(x_0) = 0 \textit{ for all } 1 \leq k \leq l''\},$$

   where $\{1, \ldots, l''\}$ *is the set of all indices of active constraints whose Lagrange multipliers are positive.*

then x_0 *is a strict local minimizer of* f*.*

Proof. Suppose x_0 is not a strict local minimizer of f. We claim that there then exists a unit vector $v \in \mathbb{R}^n$ such that

(a) $\nabla f(x_0) \cdot v \leq 0$.

(b) $\nabla h_i(x_0) \cdot v = 0$ for each $i = 1, \ldots, k$.

(c) $\nabla g_j(x_0) \cdot v \leq 0$ for all the active constraints (hereafter labelled by $j = 1, \ldots, l'$).

Intuitively, (a) says that f is non-increasing in the direction of $v \neq 0$, and (b) and (c) say that v is a feasible direction. Let us prove the claim.

Since x_0 is not a strict local minimizer, there exists a sequence x_k of feasible points unequal to x_0 converging to x_0 such that $f(x_k) \leq f(x_0)$. Then $f(x_k) - f(x_0) \leq 0$ for each k. Let $v_k = \frac{x_k - x_0}{\|x_k - x_0\|}$, and let $s_k = \|x_k - x_0\|$. Then $x_k = x_0 + s_k v_k$, with which we may rewrite the inequality as $f(s_k v_k + x_0) - f(x_0) \leq 0$. Since each $v_k \in S^1$, we may assume that the sequence v_k is convergent and that it converges to some $v \in S^1$. We claim that this vector v has the three desired properties.

By Taylor's theorem we have

$$0 \geq f(s_k v_k + x_0) - f(x_0) = s_k \nabla f(x_0) \cdot v_k + o(s_k) \tag{A}$$
$$0 = h_i(s_k v_k + x_0) - h_i(x_0) = s_k \nabla h_i(x_0) \cdot v_k + o(s_k) \tag{B}$$
$$0 \geq g_k(s_k v_k + x_0) - g_j(x_0) = s_k \nabla g_j(x_0) \cdot v_k + o(s_k) \tag{C}$$

(The last equation is ≤ 0 because $g_j(x_0) = 0$.) Divide everything by s_k and take the limit as $k \to \infty$. Then

$$0 \geq \nabla f(x_0) \cdot v \tag{a}$$
$$0 = \nabla h_i(x_0) \cdot v \tag{b}$$
$$0 \geq \nabla g_j(x_0) \cdot v, \tag{c}$$

which proves the earlier claim.

We now claim that equality actually holds in (c). Suppose for the sake of contradiction that there is some $1 \leq k \leq l'$ such that $\nabla g_j(x_0) \cdot v < 0$ for some j for which g_j is strongly active at x_0. By the first condition of the theorem,

$$0 \geq \underbrace{\nabla f(x_0) \cdot v}_{\geq 0 \text{ by (a)}} = \underbrace{-\sum \lambda_i \nabla h_i(x_0) \cdot v}_{= 0 \text{ by (b)}} \underbrace{- \sum \mu_j \nabla g_j(x_0) \cdot v}_{\geq 0 \text{ by (c)}},$$

and so the right hand side is strictly greater than zero, because we only considered strongly active constraints. This is a contradiction, so we conclude that $\nabla g_j(x_0) = 0$ for all j such that g_j is strongly active at x_0. Therefore $v \in \widetilde{\widetilde{\widetilde{T}}}_{x_0}$.

Again, by Taylor's theorem

$$0 \geq f(s_k v_k + x_0) - f(x_0) = s_k \nabla f(x_0) \cdot v_k + \frac{1}{2} s_k^2 v_k^T \nabla^2 f(x_k) \cdot v_k + o(s_k^2)$$
$$0 = h_i(s_k v_k + x_0) - h_i(x_0) = s_k \nabla h_i(x_0) \cdot v_k + \frac{1}{2} s_k^2 v_k^T \nabla^2 h_i(x_k) \cdot v_k + o(s_k^2)$$
$$0 \geq g_k(s_k v_k + x_0) - g_j(x_0) = s_k \nabla g_j(x_0) \cdot v_k + \frac{1}{2} s_k^2 v_k^T \nabla^2 g_j(x_k) \cdot v_k + o(s_k^2)$$

Multiply the second line by λ_i and the third by μ_j and add everything up to get

$$0 \geq s_k \underbrace{\left[\nabla f(x_0) + \sum \lambda_i \nabla h_i(x_0) + \sum \mu_j \nabla g_j(x_0)\right]}_{= 0 \text{ by condition 1}} v_k + \frac{s_k^2}{2} v_k^T \underbrace{\left[\nabla^2 f(x_0) + \sum \lambda_i \nabla^2 h_i(x_0) + \sum \mu_j \nabla^2 g_j(x_0)\right]}_{= L(x_0)} v_k + o(s_k^2$$

Divide everything by s_k^2 to get

$$0 \geq \frac{1}{2} v_k^T \left[\nabla^2 f(x_0) + \sum \lambda_i \nabla^2 h_i(x_0) + \sum \mu_j \nabla^2 g_j(x_0) \right] \cdot v_k + \frac{o(s_k^2)}{s_k^2}$$

Taking the limit $k \to \infty$ gives

$$0 \leq v^T L(x_0) \cdot v,$$

which violates condition 3 of the theorem. We have a contradiction, so we conclude that x_0 must be a strict local minimizer. $\square$

1.2 A Quick Example

Consider the example from last class:

$$\begin{aligned} \text{minimize } & f(x, y) = -x \\ \text{subject to } & g_1(x, y) = x^2 + y^2 - 1 \leq 0 \\ & g_2(x, y) = y + x - 1 \leq 0 \end{aligned}$$

We found that $(1, 0)$ was a good candidate: that it satisfied the necessary conditions. Recall that $\mu_1 = 1/2$, $g_1(1, 0) = 0$ and $\mu_2 = 0$, $g_2(1, 0) = 0$. Therefore the first constraint is strongly active. The Lagrangian is the identity matrix, so the second order sufficient conditions are satisfied. Therefore $(1, 0)$ is a strict local minimizer of f.