Chapter 3

Optimality Conditions

In this chapter we will present the first and second-order necessary and sufficient optimality conditions given by Maurer and Zowe in [13] for infinite-dimensional programming problems.

We consider the mathematical programming problem:

$$\begin{aligned} &\text{minimize } f(x) \qquad\qquad (P)\\ &\text{subject to } \ g(x) \in K \end{aligned}$$

where f is a functional defined on a real Banach space X, g a map from X into a real Banach space Y and K a closed convex cone in Y.

Definition 3.1 *A point $\bar{x} \in X$ is called optimal for (P) if $g(\bar{x}) \in K$ and if restricted to $g^{-1}(K)$ assumes a local minimum at $\bar{x}$.*

From here on, we will assume that the first and second Fréchet derivatives $f'(\bar{x})$, $g'(\bar{x})$, $f''(\bar{x})$, $g''(\bar{x})$ exist and the maps $f''(\bar{x})$ and $g''(\bar{x})$ will be interpreted as bilinear forms on $X \times X$.

The following approximation cones will be necessary for formulating our optimality conditions: the *sequential tangent cone*

$$T(M, \bar{x}) = \{h \in X | h = \lim \frac{(x_n - \bar{x})}{t_n}, x_n \in M, t_n > 0, t_n \to 0\} \tag{3.1}$$

and the *linearizing cone* of M at $\bar{x}$

$$L(M,\bar{x}) = \{h \in X | g'(\bar{x})h \in K_{g(\bar{x})}\} = g'(\bar{x})^{-1}(K_{g(\bar{x})}) \tag{3.2}$$

where $K_{g(\bar{x})}$ stands for the conical hull of $K - g(\bar{x})$.

If $K_{g(\bar{x})}$ is closed, then $T(M,\bar{x}) \subset L(M,\bar{x})$. Since we need $L(M,\bar{x}) \subset T(M,\bar{x})$, we will impose

$$0 \in int(g(\bar{x}) + g'(\bar{x})X - K) \tag{3.3}$$

as a regularity assumption.

3.1 First and second-order necessary conditions

Theorem 3.1 *(First-order necessary condition) Let $\bar{x}$ be optimal for (P) and let $\bar{x}$ be regular. Then,*

$$f'(\bar{x})h \geq 0 \ for \ all \ h \in L(M,\bar{x}). \tag{3.4}$$

Let Y^* stand for the topological dual of Y and let K^+ denote the *dual cone* of K, i.e., $K^+ = \{l \in Y^* | lk \geq 0 \ for \ all \ k \in K\}$.

When $\bar{x}$ is regular, then (3.4) is equivalent to the Karush-Kuhn-Tucker condition below.

Theorem 3.2 *(Karush-Kuhn-Tucker theorem) Let $\bar{x}$ be optimal for (P) and let $\bar{x}$ be regular. Then, there is some $l \in K^+$ such that*

$$f'(\bar{x}) = l \cdot g'(\bar{x}) \ \ and \ l \cdot g(\bar{x}) = 0. \tag{3.5}$$

A functional $l \in K^+$ for which (3.5) holds is called a *Lagrange-multiplier* for (P) at $\bar{x}$, and the function $F(x) = f(x) - lg(x)$ associated with l is called a *Lagrangian* for (P) at $\bar{x}$. Now suppose that (3.5) holds with l and put

$$K^l = K \cap \{y | ly = 0\}, \qquad M^l = g^{-1}(K^l). \tag{3.6}$$

Since $l \cdot g(\bar{x}) = 0$, i.e. $\bar{x} \in M^l$, we can define approximating cones $T(M^l, \bar{x})$ and $L(M^l, \bar{x})$ of M^l at $\bar{x}$ by replacing in (3.1) and (3.2) M by M^l and K by K^l. With these cones we can now state the second-order necessary condition.

Theorem 3.3 *(Second-order necessary condition) Let $\bar{x}$ be optimal for (P) and suppose $F(x) = f(x) - l \cdot g(x)$ is a Lagrangian for (P) at $\bar{x}$. Then $F''(\bar{x})(h,h) \geq 0$ for all $h \in T(M^l, \bar{x})$. Furthermore, if $0 \in int(g(\bar{x}) + g'(\bar{x})X - K^l)$ then $F''(\bar{x})(h,h) \geq 0$ for all $h \in L(M^l, \bar{x})$.*

3.2 First and second-order sufficient conditions

In order to understand the next theorem, we will introduce a formal definition of a "good approximation".

Definition 3.2 *The feasible set $M = g^{-1}(K)$ is said to be* approximated *at $\bar{x} \in M$ by $L(M, \bar{x})$ if there is a map $h : M \to L(M, \bar{x})$ such that $\|h(x) - (x - \bar{x})\| = o(\|x - \bar{x}\|)$ for all $x \in M$.*

Theorem 3.4 *Each of the following conditions implies that M is approximated at $\bar{x}$ by $L(M, \bar{x})$.*

i. $\bar{x}$ is regular, i.e., $0 \in int(g(x) + g'(x)X - K)$ holds.

ii. $\dim X < \infty$ and $K_{g(\bar{x})}$ is closed.

iii. $\dim X < \infty$, $Y = Y_1 \times R^n$ and $K = \{0\} \times R^n_+$, where Y_1 is a Banach space.

Since our purpose is to avoid being limited to finite-dimensional spaces, we will only be interested in (i). The proof to this theorem can be found in [13]. Finally, the first and second-order conditions are stated in the following theorems.

Theorem 3.5 *(First-order sufficient condition) Let M be approximated at $\bar{x} \in M$ by $L(M, \bar{x})$ and suppose that there is $\beta > 0$ such that $f'(\bar{x})h \geq \beta \|h\|$ for all $h \in L(M, \bar{x})$, then there are $\alpha > 0$ and $\rho > 0$ such that $f(x) \geq f(\bar{x}) + \alpha \|x - \bar{x}\|$ for all $x \in M$ with $\|x - \bar{x}\| \leq \rho$.*

Theorem 3.6 *(Second-order sufficient condition) Let* $\bar{x} \in M$ *and let* $F(x) = f(x) - l \cdot g(x)$ *be a Lagrangian for (P) at* $\bar{x}$*. Suppose that* M *is approximated at* $\bar{x}$ *by* $L(M, \bar{x})$ *and that there are* $\delta > 0$ *and* $\beta > 0$ *such that* $F''(\bar{x})(h,h) \geq \delta \|h\|^2$ *for all* $h \in L(M, \bar{x}) \cap \{h | l \cdot g'(\bar{x})h \leq \beta \|h\|\}$*, then there exist* $\alpha > 0$ *and* $\rho > 0$ *such that* $f(x) \geq f(\bar{x}) + \alpha \|x - \bar{x}\|^2$ *for all* $x \in M$ *with* $\|x - \bar{x}\| \leq \rho$*.*

3.3 Finite Dimensional Case $X = \mathbb{R}^n \quad Y = \mathbb{R}$

For the finite dimensional case, our nonlinear programming problem is translated to:

$$\text{minimize } f(x) \qquad \text{(Pn)}$$
$$\text{subject to } \ x \in M$$

where we define the feasible set as

$$M := \{x \in \mathbb{R}^n | h_i(x) = 0, i = 1, ...m, g_j \leq 0, j = 1, ..., s\}$$

and f, h_i, g_j are continuously differentiable functionals on $\mathbb{R}^n$. Given $\bar{x} \in M$, we can also define the active index set as $J_0(\bar{x}) = \{j | g_j(\bar{x}) = 0\}$.

To establish the necessary conditions for a local minimum, we will begin by presenting the Fritz-John Theorem.

Theorem 3.7 *Let* $\bar{x} \in M$ *be a local minimum of (Pn). Then, there exist* $\lambda_i \in \mathbb{R}$*,* $i = 1, ..., m$ *,* $\mu_j \geq 0$*,* $j = 1, ..., s$ *and* $\lambda_0 \geq 0$ *such that*

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla h_i(\bar{x}) + \sum_{j=1}^{s} \mu_j \nabla g_j(\bar{x}) = 0 \tag{3.7}$$

where $\mu_j g_j(\bar{x}) = 0$ *(refered to as* complementary condition*).*

However, the λ_0 multiplying $\nabla f(\bar{x})$ is not an ideal case. To this matter, we may introduce some regularity conditions stated as follows:

i. The Linearly Independent Constraint Qualification (LICQ) is said to be satisfied on $\bar{x} \in M$ if the vectors $\nabla h_i(\bar{x})$, $i = 1, ..., m$, $\nabla g_j(\bar{x})$, $j \in J_0(\bar{x})$ are linearly independent.

ii. The Mangasarian-Fromoritz Constraint Qualification (MFCQ) is said to be satisfied on $\bar{x} \in M$, if:

(a) $\nabla h_i(\bar{x})$, $i = 1, ..., m$ are linearly independent.

(b) there exists $\xi \in \mathbb{R}^n$ such that $\nabla g_j(\bar{x}) \cdot \xi < 0$, for all $j \in J_0(\bar{x})$.

It can be proven that LICQ$\Longrightarrow$MFCQ. Taking into account these regularity conditions, we get a stronger theorem: the Karush-Kuhn-Tucker Theorem.

Theorem 3.8 *(First-order necessary condition) Let $\bar{x} \in M$ be a local minimum for (Pn) and assume either LICQ or MFCQ is satisfied on $\bar{x}$. Then there exist $\lambda_i \in \mathbb{R}$, $\mu_j \geq 0$, $i = 1, ..., m$, $j = 1, ..., s$, such that*

$$\nabla f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla h_i(\bar{x}) + \sum_{j=1}^{s} \mu_j \nabla g_j(\bar{x}) = 0$$

where $\mu_j g_j(\bar{x}) = 0$, $j = 1, ..., s$.

Given such scalars λ_i, μ_j we define the Lagrangian Function as follows:

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{j=1}^{s} \mu_j \nabla g_j(x) \tag{3.8}$$

which allows us to state the second-order necessary and sufficient conditions.

Theorem 3.9 *(Second-order necessary condition) Assume the above condition holds, and assume also that $f, h_i, g_j \in C^2(\mathbb{R}^n, \mathbb{R})$. Then $\nabla_x^2 L(\bar{x}, \lambda, \mu)$ is a positive semidefinite matrix on the subspace $T_{\bar{x}} := \{\xi \in \mathbb{R}^n | \nabla h_i(\bar{x}) \cdot \xi = 0, i = 1, ..., m, \nabla g_j(\bar{x}) \cdot \xi = 0, j \in J^+(\bar{x})\}$ where $J^+(\bar{x}) = \{j \in J_0(\bar{x}) | \mu_j > 0\}$, i.e.,*

$$\nabla_x^2 L(\bar{x}, \lambda, \mu)|_{T_{\bar{x}}} \succeq 0.$$

Theorem 3.10 *(Second-order sufficient condition) Let $\bar{x} \in M$ and $\nabla L(\bar{x}, \lambda, \mu) = 0$ where $\lambda_i \in \mathbb{R}$, $\mu_j \geq 0$ and $\mu_j g_j(\bar{x}) = 0$. If $\nabla_x^2 L(\bar{x}, \lambda, \mu)|_{T_{\bar{x}}} > 0$ then $\bar{x}$ is a local minimum for (Pn).*