MA 5037: Optimization Methods and Applications
The KKT Conditions

Suh-Yuh Yang (楊肅煜)

Department of Mathematics, National Central University
Jhongli District, Taoyuan City 320317, Taiwan

First version: April 20, 2023/Last updated: June 18, 2025

Inequality constrained problems

We begin our exploration into the KKT conditions by analyzing the inequality constrained problem:

$$(P) \qquad \min f(\boldsymbol{x}) \quad \textit{subject to} \quad g_i(\boldsymbol{x}) \leq 0,\ i = 1, 2, \ldots, m,$$

where $f, g_1, \ldots, g_m$ are continuously differentiable functions over $\mathbb{R}^n$.

Definition (feasible descent direction): *Consider the problem*

$$(G) \qquad \min f(\boldsymbol{x}) \quad \textit{subject to} \quad \boldsymbol{x} \in C,$$

where f is continuously differentiable over the set $C \subseteq \mathbb{R}^n$. Then a vector $\boldsymbol{d} \neq \boldsymbol{0}$ is called a feasible descent direction at $\boldsymbol{x} \in C$ if $\nabla f(\boldsymbol{x})^\top \boldsymbol{d} < 0$, and there exists $\varepsilon > 0$ such that $\boldsymbol{x} + t\boldsymbol{d} \in C$ for all $t \in [0, \varepsilon]$.

A necessary local optimality condition and active constraints

Lemma 1: *Consider the minimization problem (G). If $\boldsymbol{x}^*$ is a local minimum of problem (G), then there are no feasible descent directions at $\boldsymbol{x}^*$.*

Proof: Suppose that there is a feasible descent direction $\boldsymbol{d}$ at $\boldsymbol{x}^*$. Then $\nabla f(\boldsymbol{x}^*)^\top \boldsymbol{d} < 0$ and there exists $\varepsilon_1 > 0$ such that $\boldsymbol{x}^* + t\boldsymbol{d} \in C$ for all $t \in [0, \varepsilon_1]$.

By the definition of the descent direction, there is an $\varepsilon_2 > 0$ with $\varepsilon_2 < \varepsilon_1$ such that $f(\boldsymbol{x}^* + t\boldsymbol{d}) < f(\boldsymbol{x}^*)$ for all $t \in (0, \varepsilon_2]$, which is a contradiction to the local optimality of $\boldsymbol{x}^*$. □

The active constraints at $\widetilde{x}$:

Definition: *Assume that we are given a set of inequalities*

$$g_i(\boldsymbol{x}) \le 0, \qquad i = 1, 2, \ldots, m,$$

where $g_i : \mathbb{R}^n \to \mathbb{R}$ are functions, and a vector $\widetilde{\boldsymbol{x}} \in \mathbb{R}^n$. Then

(i) *The active constraints at $\widetilde{\boldsymbol{x}}$ are the constraints satisfied as equalities at $\widetilde{\boldsymbol{x}}$; and*

(ii) *The set of active constraints is denoted by $I(\widetilde{\boldsymbol{x}}) := \{i : g_i(\widetilde{\boldsymbol{x}}) = 0\}$.*

Another necessary local optimality condition

Lemma 2: *Let $\boldsymbol{x}^*$ be a local minimum of the inequality constrained problem*

$$(P) \qquad \min f(\boldsymbol{x}) \quad \textit{subject to} \quad g_i(\boldsymbol{x}) \leq 0,\ i = 1, 2, \ldots, m,$$

where $f, g_1, \ldots, g_m$ are continuously differentiable functions over $\mathbb{R}^n$. Let $I(\boldsymbol{x}^*)$ be the set of active constraints at $\boldsymbol{x}^*$, i.e.,

$$I(\boldsymbol{x}^*) = \{i : g_i(\boldsymbol{x}^*) = 0\}.$$

Then there does not exist a vector $\boldsymbol{d} \in \mathbb{R}^n$ such that

$$\nabla f(\boldsymbol{x}^*)^\top \boldsymbol{d} < 0 \quad \text{and} \quad \nabla g_i(\boldsymbol{x}^*)^\top \boldsymbol{d} < 0 \text{ for } i \in I(\boldsymbol{x}^*). \qquad (1)$$

Proof:
Suppose by contradiction that $\boldsymbol{d}$ satisfies the system of inequalities (1). Then it follows that there exists $\varepsilon_1 > 0$ such that $f(\boldsymbol{x}^* + t\boldsymbol{d}) < f(\boldsymbol{x}^*)$ and $g_i(\boldsymbol{x}^* + t\boldsymbol{d}) < g_i(\boldsymbol{x}^*) = 0$ for any $t \in (0, \varepsilon_1)$ and $i \in I(\boldsymbol{x}^*)$.

Proof of Lemma 2 (cont'd)

For any $i \notin I(\boldsymbol{x}^*)$, we have that $g_i(\boldsymbol{x}^*) < 0$, and hence, by continuity of g_i for all i, it follows that there exists $\varepsilon_2 > 0$ such that $g_i(\boldsymbol{x}^* + t\boldsymbol{d}) < 0$ for any $t \in (0, \varepsilon_2)$ and $i \notin I(\boldsymbol{x}^*)$.

Now, no matter $i \in I(\boldsymbol{x}^*)$ or not, we can conclude that

$$\begin{aligned} & f(\boldsymbol{x}^* + t\boldsymbol{d}) < f(\boldsymbol{x}^*), \\ & g_i(\boldsymbol{x}^* + t\boldsymbol{d}) < 0, \quad i = 1, 2, \ldots, m, \\ & \text{for all } t \in (0, \min\{\varepsilon_1, \varepsilon_2\}). \end{aligned}$$

This is a contradiction to the local optimality of $\boldsymbol{x}^*$. $\square$

Fritz John conditions for inequality constrained problems

Theorem (Fritz John conditions): *Let $\boldsymbol{x}^*$ be a local minimum of the inequality constrained problem*

$$(P) \qquad \min f(\boldsymbol{x}) \quad \text{subject to} \quad g_i(\boldsymbol{x}) \le 0,\ i = 1, 2, \ldots, m,$$

where $f, g_1, \ldots, g_m$ are continuously differentiable functions over $\mathbb{R}^n$. Then there exist multipliers $\lambda_0, \lambda_1, \ldots, \lambda_m \ge 0$, which are not all zeros, such that

$$\begin{aligned} \lambda_0 \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}^*) &= \mathbf{0}, \\ \lambda_i g_i(\boldsymbol{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned} \tag{2}$$

Proof:
Let $I(\boldsymbol{x}^*) = \{i : g_i(\boldsymbol{x}^*) = 0\} := \{i_1, i_2, \ldots, i_k\}$. By Lemma 2, it follows that the following system of inequalities *does not have a solution $\boldsymbol{d}$:*

$$\nabla f(\boldsymbol{x}^*)^\top \boldsymbol{d} < 0, \quad \nabla g_i(\boldsymbol{x}^*)^\top \boldsymbol{d} < 0, \quad i \in I(\boldsymbol{x}^*). \tag{3}$$

System (3) can be rewritten as $\boldsymbol{A}\boldsymbol{d} < \mathbf{0}$, where

$$\boldsymbol{A} := \begin{pmatrix} \nabla f(\boldsymbol{x}^*)^\top \\ \nabla g_{i_1}(\boldsymbol{x}^*)^\top \\ \vdots \\ \nabla g_{i_k}(\boldsymbol{x}^*)^\top \end{pmatrix}.$$

Proof of the theorem on the Fritz John conditions (cont'd)

By Gordan's alternative theorem, system (3) is *infeasible* if and only if there exists a vector $\boldsymbol{\eta} = (\lambda_0, \lambda_{i_1}, \ldots, \lambda_{i_k})^\top \neq \mathbf{0}$ such that

$$\boldsymbol{A}^\top \boldsymbol{\eta} = \mathbf{0}, \quad \boldsymbol{\eta} \geq \mathbf{0},$$

which is the same as

$$\lambda_0 \nabla f(\boldsymbol{x}^*) + \sum_{i \in I(\boldsymbol{x}^*)} \lambda_i \nabla g_i(\boldsymbol{x}^*) = \mathbf{0}, \quad \lambda_i \geq 0.$$

Define $\lambda_i = 0$ for any $i \notin I(\boldsymbol{x}^*)$. Then we obtain that

$$\lambda_0 \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}^*) = \mathbf{0}$$

and $\lambda_i g_i(\boldsymbol{x}^*) = 0$ for any $i \in \{1, 2, \ldots, m\}$ as required. $\square$

A major drawback of the Fritz John conditions

- The Fritz-John conditions allow λ_0 to be zero, which is not particularly informative since condition (2) then becomes

$$\sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}^*) = \mathbf{0},$$

implying the gradients of the active constraints $\{\nabla g_i(\boldsymbol{x}^*)\}_{i \in I(x^*)}$ are linearly dependent.

This condition has nothing to do with the objective function, implying that there might be a lot of points satisfying the Fritz John conditions which are not local minimum points.

- *If we add an assumption that the gradients of the active constraints are linearly independent at $\boldsymbol{x}^*$, then we can establish the KKT conditions, which are the same as the Fritz John conditions with $\lambda_0 = 1$.*

KKT conditions for inequality constrained problems

Theorem (KKT conditions): *Let $\boldsymbol{x}^*$ be a local minimum of the inequality constrained problem*

$$(P) \qquad \min f(\boldsymbol{x}) \quad \textit{subject to} \quad g_i(\boldsymbol{x}) \leq 0,\ i = 1, 2, \ldots, m,$$

where $f, g_1, \ldots, g_m$ are continuously differentiable functions over $\mathbb{R}^n$. Let $I(\boldsymbol{x}^*)$ be the set of active constraints at $\boldsymbol{x}^*$, i.e.,

$$I(\boldsymbol{x}^*) = \{i : g_i(\boldsymbol{x}^*) = 0\}.$$

Suppose that $\{\nabla g_i(\boldsymbol{x}^*)\}_{i \in I(\boldsymbol{x}^*)}$ are linearly independent. Then there exist multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m \geq 0$ such that

$$\begin{aligned} \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}^*) &= \mathbf{0}, \\ \lambda_i g_i(\boldsymbol{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

Proof of the theorem on the KKT conditions

By the Fritz John conditions, there exist $\tilde{\lambda}_0, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_m \geq 0$, not all zeros, such that

$$\begin{aligned} \tilde{\lambda}_0 \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla g_i(\boldsymbol{x}^*) &= \mathbf{0}, && (4) \\ \tilde{\lambda}_i g_i(\boldsymbol{x}^*) &= 0, \quad i = 1, 2, \ldots, m. && (5) \end{aligned}$$

Moreover, we have $\tilde{\lambda}_0 > 0$, since if $\tilde{\lambda}_0 = 0$, by (4) and (5), it follows that

$$\sum_{i \in I(\boldsymbol{x}^*)} \tilde{\lambda}_i \nabla g_i(\boldsymbol{x}^*) = \mathbf{0},$$

where not all the scalars $\tilde{\lambda}_i, i \in I(\boldsymbol{x}^*)$ are zeros, leading to a contradiction to that $\{\nabla g_i(\boldsymbol{x}^*)\}_{i \in I(\boldsymbol{x}^*)}$ are linearly independent.

Now, by defining $\lambda_i := \frac{\tilde{\lambda}_i}{\tilde{\lambda}_0}$ for $i = 0, 1, \ldots, m$, the KKT conditions directly follow from (4) and (5). $\square$

KKT conditions for inequality/equality constrains

Theorem (KKT conditions): *Let $\boldsymbol{x}^*$ be a local minimum of the problem*

$$\min f(\boldsymbol{x}) \quad \text{subject to} \quad \begin{cases} g_i(\boldsymbol{x}) \le 0, & i = 1, 2, \ldots, m, \\ h_j(\boldsymbol{x}) = 0, & j = 1, 2, \ldots, p, \end{cases} \tag{6}$$

where $f, g_1, \ldots, g_m, h_1, h_2, \ldots, h_p$ are continuously differentiable functions over $\mathbb{R}^n$. Suppose that the gradients of the active constraints and the equality constraints

$$\{\nabla g_i(\boldsymbol{x}^*) : i \in I(\boldsymbol{x}^*)\} \cup \{\nabla h_j(\boldsymbol{x}^*) : j = 1, 2, \ldots, p\}$$

are linearly independent. Then there exist multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m \ge 0$ and $\mu_1, \mu_2, \ldots, \mu_p \in \mathbb{R}$ such that

$$\begin{aligned} \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\boldsymbol{x}^*) &= \mathbf{0}, \\ \lambda_i g_i(\boldsymbol{x}^*) &= 0, \ \ i = 1, 2, \ldots, m. \end{aligned}$$

KKT points

Definition (KKT points): *Consider the minimization problem (6), where* $f, g_1, \ldots, g_m, h_1, \ldots, h_p$ *are continuously differentiable functions over* $\mathbb{R}^n$*.*

A feasible point $\boldsymbol{x}^*$ *is called a KKT point if there exist* $\lambda_1, \lambda_2, \ldots, \lambda_m \geq 0$ *and* $\mu_1, \mu_2, \ldots, \mu_p \in \mathbb{R}$ *such that*

$$\begin{aligned} \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\boldsymbol{x}^*) &= \mathbf{0}, \\ \lambda_i g_i(\boldsymbol{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

Regular points

Definition (Regularity): *Consider the minimization problem (6), where* $f, g_1, \ldots, g_m, h_1, \ldots, h_p$ *are continuously differentiable functions over* $\mathbb{R}^n$.

A feasible point $\boldsymbol{x}^*$ *is called regular if the gradients of the active constraints among the inequality constraints and of the equality constraints*

$$\{\nabla g_i(\boldsymbol{x}^*) : i \in I(\boldsymbol{x}^*)\} \cup \{\nabla h_j(\boldsymbol{x}^*) : j = 1, 2, \ldots, p\}$$

are linearly independent.

Note: In the terminologies of the KKT point and regular point:

- *A necessary optimality condition for local optimality of a regular point is that it is a KKT point.*
- *The additional requirement of regularity is not required in the linearly constrained case in which no such assumption is needed; see Chapter 10, Theorem on the KKT conditions for linearly constrained problems.*

KKT conditions for convex optimization problems

When problem is convex, the KKT conditions are always sufficient.

Theorem (sufficiency of the KKT conditions): *Let $\boldsymbol{x}^*$ be a feasible solution of the constrained minimization problem,*

$$\min f(\boldsymbol{x}) \quad \text{subject to} \quad \begin{cases} g_i(\boldsymbol{x}) \le 0, & i = 1, 2, \ldots, m, \\ h_j(\boldsymbol{x}) = 0, & j = 1, 2, \ldots, p, \end{cases} \tag{7}$$

where $f, g_1, \ldots, g_m$ are continuously differentiable convex functions over $\mathbb{R}^n$ and $h_1, \ldots, h_p$ are affine functions. Suppose there exist multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m \ge 0$ and $\mu_1, \mu_2, \ldots, \mu_p \in \mathbb{R}$ such that

$$\begin{aligned} \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\boldsymbol{x}^*) &= \mathbf{0}, \\ \lambda_i g_i(\boldsymbol{x}^*) &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

Then $\boldsymbol{x}^*$ is an optimal solution of problem (7).

Proof of the sufficiency theorem of the KKT conditions

Let $\boldsymbol{x}$ be a feasible solution of problem (7). It suffices to show that $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$. Note that the following function is convex:

$$s(\boldsymbol{x}) := f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) + \sum_{j=1}^{p} \mu_j h_j(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \mathbb{R}^n.$$

Since $\nabla s(\boldsymbol{x}^*) = \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\boldsymbol{x}^*) = 0$, it follows that $\boldsymbol{x}^*$ is a minimizer of s, i.e., $s(\boldsymbol{x}^*) \leq s(\boldsymbol{x})$, $\forall\, \boldsymbol{x} \in \mathbb{R}^n$. We can conclude that

$$\begin{aligned}
f(\boldsymbol{x}^*) &= f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}^*) + \sum_{j=1}^{p} \mu_j h_j(\boldsymbol{x}^*) \quad (\because \lambda_i g_i(\boldsymbol{x}^*) = 0 = h_j(\boldsymbol{x}^*)) \\
&= s(\boldsymbol{x}^*) \leq s(\boldsymbol{x}) \\
&= f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) + \sum_{j=1}^{p} \mu_j h_j(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \mathbb{R}^n \\
&\leq f(\boldsymbol{x}), \quad \forall\, \boldsymbol{x} \textit{ feasible solution.} \quad \square
\end{aligned}$$

Slater's condition

Definition: *We say that Slater's condition is satisfied for a set of convex inequalities*

$$g_i(\boldsymbol{x}) \leq 0, \quad i = 1, 2, \ldots, m,$$

where $g_1, g_2, \ldots, g_m$ *are given convex functions, if there exists* $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ *such that*

$$g_i(\hat{\boldsymbol{x}}) < 0, \quad i = 1, 2, \ldots, m.$$

Note: Slater's condition requires that there exists a point that strictly satisfies the constraints, and does not require, like in the regularity condition, an a priori knowledge on the point that is a candidate to be an optimal solution.

This is the reason why checking the validity of Slater's condition is usually a much easier task than checking regularity.

Necessity of the KKT conditions under Slater's condition

Theorem: *Let $\boldsymbol{x}^*$ be an optimal solution of the problem*

$$\min f(\boldsymbol{x}) \quad \textit{subject to} \quad g_i(\boldsymbol{x}) \le 0,\ i = 1, 2, \ldots, m,$$

where $f, g_1, \ldots, g_m$ are continuously differentiable functions over $\mathbb{R}^n$. In addition, $g_1, g_2, \ldots, g_m$ are convex functions over $\mathbb{R}^n$. Suppose that there exists $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ such that

$$g_i(\hat{\boldsymbol{x}}) < 0,\ \ i = 1, 2, \ldots, m,$$

Then there exist multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m \ge 0$ such that

$$\begin{aligned} \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}^*) &= \mathbf{0}, \qquad (8) \\ \lambda_i g_i(\boldsymbol{x}^*) &= 0,\ \ i = 1, 2, \ldots, m. \end{aligned}$$

Proof of the theorem

Since $\boldsymbol{x}^*$ is an optimal solution of $f(\boldsymbol{x})$, then the Fritz John conditions are satisfied. That is, there exist $\tilde{\lambda}_0, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_m \geq 0$, which are not all zeros, such that

$$\begin{aligned} \tilde{\lambda}_0 \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla g_i(\boldsymbol{x}^*) &= \mathbf{0}, \\ \tilde{\lambda}_i g_i(\boldsymbol{x}^*) &= 0, \ \ i = 1, 2, \ldots, m. \end{aligned} \tag{9}$$

All that we need to show is that $\tilde{\lambda}_0 > 0$, and then the conditions (8) will be satisfied with $\lambda_i = \frac{\tilde{\lambda}_i}{\tilde{\lambda}_0}$, $i = 1, 2, \ldots, m$. To prove that $\tilde{\lambda}_0 > 0$, assume in contradiction that it is zero. Then we have

$$\sum_{i=1}^{m} \tilde{\lambda}_i \nabla g_i(\boldsymbol{x}^*) = \mathbf{0}. \tag{10}$$

Proof of the theorem (cont'd)

By the gradient inequality for convex functions, we have that for all $i = 1, 2, \ldots, m$,

$$0 > g_i(\hat{\boldsymbol{x}}) \geq g_i(\boldsymbol{x}^*) + \nabla g_i(\boldsymbol{x}^*)^\top (\hat{\boldsymbol{x}} - \boldsymbol{x}^*).$$

Multiplying the ith inequality by $\tilde{\lambda}_i$ and summing over $i = 1, 2, \ldots, m$, we obtain

$$0 > \sum_{i=1}^{m} \tilde{\lambda}_i g_i(\boldsymbol{x}^*) + \left(\sum_{i=1}^{m} \tilde{\lambda}_i \nabla g_i(\boldsymbol{x}^*) \right)^\top (\hat{\boldsymbol{x}} - \boldsymbol{x}^*), \tag{11}$$

where the inequality is strict since not all the $\tilde{\lambda}_i$ are zero. Plugging the identities (9) and (10) into (11), we obtain the impossible statement that $0 > 0$, thus establishing the result. $\square$

Generalized Slater's condition (GSC)

Definition: *Consider the system*

$$\begin{aligned} g_i(\boldsymbol{x}) &\le 0, & i &= 1,2,\ldots,m, \\ h_j(\boldsymbol{x}) &\le 0, & j &= 1,2,\ldots,p, \\ s_k(\boldsymbol{x}) &= 0, & k &= 1,2,\ldots,q, \end{aligned}$$

where $g_i, i = 1,2,\ldots,m$, are convex functions and $h_j, s_k, j = 1,2,\ldots,p$, $k = 1,2,\ldots,q$, are affine functions. Then we say that the generalized Slater's condition (GSC) is satisfied if there exists $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ for which

$$\begin{aligned} g_i(\hat{\boldsymbol{x}}) &< 0, & i &= 1,2,\ldots,m, \\ h_j(\hat{\boldsymbol{x}}) &\le 0, & j &= 1,2,\ldots,p, \\ s_k(\hat{\boldsymbol{x}}) &= 0, & k &= 1,2,\ldots,q. \end{aligned}$$

Necessity of the KKT conditions under the GSC

Theorem: *Let $\boldsymbol{x}^*$ be an optimal solution of the problem*

$$\begin{aligned} \min f(\boldsymbol{x}) \quad \text{subject to} \quad & g_i(\boldsymbol{x}) \le 0, && i = 1,2,\ldots,m, \\ & h_j(\boldsymbol{x}) \le 0, && j = 1,2,\ldots,p, \\ & s_k(\boldsymbol{x}) = 0, && k = 1,2,\ldots,q, \end{aligned}$$

where $f, g_1, \ldots, g_m$ are continuously differentiable convex functions over $\mathbb{R}^n$, and $h_j, s_k, j = 1,2,\ldots,p, k = 1,2,\ldots,q$, are affine. Suppose that there exists $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ such that

$$\begin{aligned} g_i(\hat{\boldsymbol{x}}) < 0, && i = 1,2,\ldots,m, \\ h_j(\hat{\boldsymbol{x}}) \le 0, && j = 1,2,\ldots,p, \\ s_k(\hat{\boldsymbol{x}}) = 0, && k = 1,2,\ldots,q. \end{aligned}$$

Then there exist multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m \ge 0$, $\eta_1, \eta_2, \ldots, \eta_p \ge 0$, and $\mu_1, \mu_2, \ldots, \mu_q \in \mathbb{R}$ such that

$$\begin{aligned} & \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}^*) + \sum_{j=1}^{p} \eta_j \nabla h_j(\boldsymbol{x}^*) + \sum_{k=1}^{q} \mu_k \nabla s_k(\boldsymbol{x}^*) = \mathbf{0}, \\ & \lambda_i g_i(\boldsymbol{x}^*) = 0, \quad i = 1,2,\ldots,m, \\ & \eta_j h_j(\boldsymbol{x}^*) = 0, \quad i = 1,2,\ldots,p. \end{aligned}$$