8 KKT Conditions and Lagrangian Duality

Consider the general nonlinear programming (NLP)

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h_j(x) = 0\,,\ j = 1,\ldots,p \\ & g_i(x) \le 0\,,\ i = 1,\ldots,m \\ & x \in X\,, \end{aligned}$$

where $X \subseteq \Re^n$ is an open set.

Let

$$h(x) = \begin{pmatrix} h_1(x) \\ \vdots \\ h_p(x) \end{pmatrix} \quad \text{and} \quad g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix}.$$

Then the NLP becomes

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \le 0 \\ & x \in X\,, \end{aligned}$$

Let $S = \{x \in X \,|\, h(x) = 0,\ g(x) \le 0\}$ denote the feasible region of the NLP. If $X = \Re^n$, then

$$S = \{\, x \in \Re^n \,|\, h(x) = 0\,,\ g(x) \le 0 \,\}\,.$$

Theorem 8.1 (Necessary conditions for local minimum of the NLP) *Let $\emptyset \ne X \subseteq \Re^n$ be an open set. Suppose that $\nabla f(x), \nabla h(x)$, and $\nabla g(x)$ are all continuous on X, where $\nabla h(x) = (\nabla h_1(x), \ldots, \nabla h_p(x))$ and $\nabla g(x) = (\nabla g_1(x), \ldots, \nabla g_m(x))$. Suppose that $\bar{x}$ is a local minimum of f on S. Let $I(\bar{x}) = \{\, i \,|\, g_i(\bar{x}) = 0 \,\}$ be the index set of active constraints of g at $\bar{x}$. Suppose further that*

$$\{\, \nabla h_j(\bar{x}) \,\} \cup \{\, \nabla g_i(\bar{x}) \,\}_{i \in I(\bar{x})}$$

are linearly independent, then there exist scalars $\bar{\lambda} \in \Re^p$ and $\bar{\mu} \in \Re^m$ such that

$$\begin{cases} \nabla f(\bar{x}) + \nabla h(\bar{x})\bar{\lambda} + \nabla g(\bar{x})\bar{\mu} = 0 \\ h(\bar{x}) = 0 \\ \bar{\mu} \ge 0,\ g(\bar{x}) \le 0,\ \bar{\mu}^T g(\bar{x}) = 0\,. \end{cases} \tag{8.1}$$

Remarks

1. In Theorem 8.1, $(\bar{\lambda}, \bar{\mu})$ is called multipliers or Lagrangian multipliers and $(\bar{x}, \bar{\lambda}, \bar{\mu})$ or $\bar{x}$ is called a KKT solution or point.

2. The condition that $\{ \nabla h_j(\bar{x}) \} \cup \{ \nabla g_i(\bar{x}) \}_{i \in I(\bar{x})}$ are linearly independent is called a constraint qualification (CQ). This CQ guarantees (8.1) to hold.

3. There are many other CQs. For example, $h(x) = Ax - b$, $\text{rank}(A) = m$ and there exists $y \in S$ such that $g(y) < 0$ and $Ay - b = 0$. [Recall the subproblems involved in the Affine Scaling Algorithm]

4. If S is a polyhedral set, then there is no other CQ required to guarantees (8.1) to hold

Theorem 8.2 (KKT sufficient conditions) *Let $\emptyset \neq X \subseteq \Re^n$ be an open set. Suppose that f and $g_i,\ i = 1, \ldots, m$ are continuously differentiable convex functions and $h_j,\ j = 1, \ldots, p$ are affine functions. Suppose that $(\bar{x}, \bar{\lambda}, \bar{\mu})$ satisfies the KKT conditions*

$$\begin{cases} \nabla f(\bar{x}) + \nabla h(\bar{x})\bar{\lambda} + \nabla g(\bar{x})\bar{\mu} = 0 \\ h(\bar{x}) = 0 \\ \bar{\mu} \geq 0,\, g(\bar{x}) \leq 0,\, \bar{\mu}^T g(\bar{x}) = 0 \,. \end{cases}$$

Then $\bar{x}$ is a global optimal solution to the NLP.

Example 1. The linear programming

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

satisfies all the conditions in Theorem 8.2. So, a KKT solution $\bar{x}$ of the linear programming is actually a global solution of the linear programming.

Example 2. Consider the following quadratic programming (QP)

$$\begin{aligned} \min \quad f(x) = \quad & \frac{1}{2} x^T Q x + c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \\ & x \in \Re^n \,, \end{aligned}$$

where $A \in \Re^{m\times n}$ and $Q = Q^T$. Let $g_i(x) = -x_i = -e_i^T x,\ i = 1,\dots,n$ and $h_j(x) = a_j^T x - b_j,\ j = 1,\dots,m$. The QP becomes

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \le 0 \\ & x \in \Re^n . \end{aligned}$$

Since h and g are affine functions, for any local minimum $\bar{x}$ of f on $S = \{x \in \Re^n \,|\, h(x) = 0,\ g(x) \le 0\}$ there must exist $\bar{\lambda} \in \Re^m$ and $\bar{\mu} \in \Re^n$ satisfying the following KKT conditions

$$\begin{cases} \nabla f(\bar{x}) + \nabla h(\bar{x})\bar{\lambda} + \nabla g(\bar{x})\bar{\mu} = 0 \\ h(\bar{x}) = 0 \\ \bar{\mu} \ge 0,\ g(\bar{x}) \le 0,\ \bar{\mu}^T g(\bar{x}) = 0 . \end{cases}$$

By calculation,

$$\nabla g(x) = (-e_1 \dots -e_n) = -I ,$$

$$\nabla h(x) = (a_1 \dots a_m) = A^T$$

and

$$\nabla f(x) = Qx + c .$$

Hence, the KKT conditions for the QP are

$$\begin{cases} Qx + c + A^T\bar{\lambda} - \bar{\mu} = 0 \\ A\bar{x} = b \\ \bar{\mu} \ge 0,\ \bar{x} \ge 0,\ \bar{\mu}^T\bar{x} = 0 . \end{cases} \tag{8.2}$$

Then, if Q is positive semidefinite, i.e., if f is a convex function, any $\bar{x}$ satisfies (8.2) is a global solution to the QP.

Lagrangian Dual

Consider the following problem

$$(P) \qquad \begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \le 0 \\ & x \in X , \end{aligned}$$

where X is a subset of $\Re^n$, $h : \Re^n \to \Re^p$ and $g : \Re^n \to \Re^m$. Then the Lagrangian dual problem of (P) is defined as

$$(D) \qquad \begin{array}{ll} \max & \theta(u,v) \\ \text{s.t.} & v \geq 0 \,, \end{array}$$

where

$$\theta(u,v) \;=\; \inf\{f(x) \,+\, u^T h(x) \,+\, v^T g(x) \,|\, x \in X\}\,.$$

Let

$$L(x,u,v) \;=\; f(x) \,+\, u^T h(x) \,+\, v^T g(x)\,.$$

Then

$$\theta(u,v) \;=\; \inf\{L(x,u,v) \,|\, x \in X\}\,.$$

Relationship between Saddle Point Optimality Conditions and KKT Conditions

Theorem 8.3 **(a)** *Let $S = \{x \in X \,|\, h(x) = 0,\; g(x) \leq 0\}$. Consider the primal problem*

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in S\,. \end{array}$$

Suppose that $(\bar{x}, \bar{u}, \bar{v}) \in S \times \Re^p \times \Re^m$ satisfy the KKT conditions

$$\begin{cases} \nabla f(\bar{x}) \,+\, \nabla h(\bar{x})\bar{u} \,+\, \nabla g(\bar{x})\bar{v} \;=\; 0 \\ h(\bar{x}) \;=\; 0 \\ \bar{v} \;\geq\; 0,\; g(\bar{x}) \;\leq\; 0,\; \bar{v}^T g(\bar{x}) \;=\; 0\,. \end{cases}$$

Suppose that f, and g_i, $i = 1, \ldots, m$ are convex functions, and h_j, $j = 1, \ldots, p$ are affine functions. Then $(\bar{x}, \bar{u}, \bar{v})$ is a saddle point of $L(x,u,v)$, i.e.,

$$L(\bar{x}, u, v) \;\leq\; L(\bar{x}, \bar{u}, \bar{v}) \;\leq\; L(x, \bar{u}, \bar{v}) \;\forall x \in X,\; (u,v) \text{ with } v \geq 0\,.$$

(b) *Conversely, suppose that $(\bar{x}, \bar{u}, \bar{v})$ with $\bar{x} \in \operatorname{int}(X)$, $\bar{v} \geq 0$ is a saddle point of $L(x,u,v)$. Then $(\bar{x}, \bar{u}, \bar{v})$ satisfies the KKT conditions.*

Remarks.

1. Since

   $$\theta(u,v) \;\leq\; L(x,u,v) \quad \forall\, x \in X\,,$$

   we have, in particular,

   $$\theta(u,v) \;\leq\; L(\bar{x},u,v)\,.$$

From

$$L(\bar{x},\bar{u},\bar{v}) \le L(x,\bar{u},\bar{v}) \quad \forall x \in X,$$

we obtain

$$\theta(\bar{u},\bar{v}) = L(\bar{x},\bar{u},\bar{v}).$$

Hence, part (a) of Theorem 8.3 implies

$$\theta(u,v) \le \theta(\bar{u},\bar{v}).$$

That is, $(\bar{u},\bar{v})$ solves

$$\begin{array}{rl} \max & \theta(u,v) \\ \text{s.t.} & v \ge 0. \end{array}$$

[Lagrangian multipliers solve the Lagrangian Dual]

2. Consider the linear programming

$$(P) \qquad \begin{array}{rl} \min & c^T x \\ \text{s.t.} & Ax = b \\ & x \ge 0. \end{array}$$

Let

$$L(x,u,v) = c^T x + u^T(Ax - b) - v^T x.$$

The KKT conditions are

$$\begin{cases} c + A^T\bar{u} - \bar{v} = 0 \\ A\bar{x} = b \\ \bar{v} \ge 0,\ \bar{x} \ge 0,\ \bar{v}^T\bar{x} = 0. \end{cases}$$

The Lagrangian dual is

$$(D) \qquad \begin{array}{rl} \max & \theta(u,v) \\ \text{s.t.} & v \ge 0, \end{array}$$

where

$$\begin{aligned} \theta(u,v) &= \inf_{x\in\Re^n} \{ L(x,u,v) \} \\ &= \inf_{x\in\Re^n} \{ (c + A^T u - v)^T x - u^T b \} \\ &= \begin{cases} -u^T b & \text{if } c + A^T u - v = 0 \\ -\infty & \text{if } c + A^T u - v \ne 0. \end{cases} \end{aligned}$$

Then the Lagrangian dual (D) becomes

$$(D1) \qquad \begin{array}{rl} \max & (-u)^T b \\ \text{s.t.} & c + A^T u - v = 0 \\ & v \ge 0, \end{array}$$

which, by letting $y = -u$, is equivalent to

$$(D2) \qquad \begin{array}{ll} \max & y^T b \\ \text{s.t.} & A^T y \leq c\,. \end{array}$$

Suppose that $(\bar{x}, \bar{u}, \bar{v})$ satisfies the KKT conditions. Then $\bar{x}$ solves (P), $(\bar{u}, \bar{v})$ solves (D1) and $\bar{y} = -\bar{u}$ solves (D2).

Example 3. Consider

$$(P) \qquad \begin{array}{lll} \min & f(x) = & x_1^2 + x_2^2 = \frac{1}{2} x^T H x \\ \text{s.t.} & & -x_1 - x_2 + 4 \leq 0 \\ & & x_1 \geq 0 \\ & & x_2 \geq 0\,, \end{array}$$

where $H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. Let $g_1(x) = -x_1 - x_2 + 4$, $g_2(x) = -x_1, g_3(x) = -x_2$ and $X = \Re^2$. The Lagrangian dual problem is

$$(D) \qquad \begin{array}{ll} \max & \theta(v) \\ \text{s.t.} & v_1, v_2, v_3 \geq 0 \\ & v \in \Re^3\,, \end{array}$$

where

$$\theta(v) = \inf_{x \in \Re^2} L(x, v)$$

and

$$L(x, v) = f(x) + v^T g(x) = \frac{1}{2} x^T H x + v_1 g_1(x) + v_2 g_2(x) + v_3 g_3(x)\,.$$

Since $L(x, v)$ is convex on x, any $\bar{x}$ such that $\nabla_x L(\bar{x}, v) = 0$ satisfies

$$L(\bar{x}, v) \leq L(x, v).$$

Let $\bar{x}$ be such that $\nabla_x L(\bar{x}, v) = 0$. Then

$$H\bar{x} + v_1 \nabla g_1(\bar{x}) + v_2 \nabla g_2(\bar{x}) + v_3 \nabla g_3(\bar{x}) = 0\,,$$

which, gives

$$\bar{x}_1 = \frac{v_1 + v_2}{2}, \quad \bar{x}_2 = \frac{v_1 + v_3}{2}\,.$$

Therefore,

$$\theta(v) = -\frac{v_1^2}{2} - \frac{1}{4} v_2^2 - \frac{1}{4} v_3^2 - \frac{v_1 v_2}{2} - \frac{v_1 v_3}{2} + 4 v_1\,.$$

This gives an explicit form of the dual problem

$$(D)\qquad \begin{array}{lll} \max & \theta(v) = & -\dfrac{v_1^2}{2} - \dfrac{1}{4}v_2^2 - \dfrac{1}{4}v_3^2 - \dfrac{v_1 v_2}{2} - \dfrac{v_1 v_3}{2} + 4v_1 \\ \text{s.t.} & & v_1, v_2, v_3 \geq 0 \\ & & v \in \Re^3 . \end{array}$$

We can see easily that $\bar{x} = (2,2)^T$ solves (P) and $\bar{v} = (4,0,0)^T$ solves (D) with

$$f(\bar{x}) = 8 = \theta(\bar{v}) .$$

The KKT conditions for (P) are

$$\begin{cases} H\bar{x} + \bar{v}_1 \nabla g_1(\bar{x}) + \bar{v}_2 \nabla g_2(\bar{x}) + \bar{v}_3 \nabla g_3(\bar{x}) = 0 \\ \bar{v}_1 \geq 0 ,\ g_1(\bar{x}) \leq 0 ,\ \bar{v}_1 g_1(\bar{x}) = 0 \\ \bar{v}_2 \geq 0 ,\ g_2(\bar{x}) \leq 0 ,\ \bar{v}_2 g_2(\bar{x}) = 0 \\ \bar{v}_3 \geq 0 ,\ g_3(\bar{x}) \leq 0 ,\ \bar{v}_3 g_3(\bar{x}) = 0 , \end{cases}$$

which, by eliminating $\bar{x}$, gives

$$\begin{cases} \bar{v}_1 \geq 0 ,\ \bar{v}_1 + \dfrac{\bar{v}_2 + \bar{v}_3}{2} - 4 \geq 0 ,\ \bar{v}_1(\bar{v}_1 + \dfrac{\bar{v}_2 + \bar{v}_3}{2} - 4) = 0 \\ \bar{v}_2 \geq 0 ,\ \dfrac{\bar{v}_1 + \bar{v}_2}{2} \geq 0 ,\ \bar{v}_2(\dfrac{\bar{v}_1 + \bar{v}_2}{2}) = 0 \\ \bar{v}_3 \geq 0 ,\ \dfrac{\bar{v}_1 + \bar{v}_3}{2} \geq 0 ,\ \bar{v}_3(\dfrac{\bar{v}_1 + \bar{v}_3}{2}) = 0 . \end{cases}$$

Case A) If $\bar{v}_1 > 0$, then

$$4 = \bar{v}_1 + \frac{\bar{v}_2 + \bar{v}_3}{2} .$$

A1) If $\bar{v}_2 > 0$, then

$$\frac{\bar{v}_1 + \bar{v}_2}{2} = 0 ,$$

which leads to a contradiction.

A2) If $\bar{v}_2 = 0$, then we have two cases.

A2I) $\bar{v}_3 > 0$, which cannot hold.

A2II) $\bar{v}_3 = 0$, which implies that $\bar{v}_1 = 4$. A solution found.

Case B) If $\bar{v}_1 = 0$, then

$$\bar{v}_2 = 0 = \bar{v}_3 , \text{and } \bar{v}_1 \geq 4 ,$$

which is impossible

Therefore, by solving the KKT conditions, we find

$$\bar{x} = (2,2)^T , \quad \bar{v} = (4,0,0)^T .$$