OR4030 OPTIMIZATION Chapter 6
Optimality Conditions for Constrained Optimization Problems
6.1 Equality Constrained Optimization

6.1.1 Problem Description

Problem:

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & h_i(x) = 0, \quad i = 1, \ldots, m \\ & x \in \Re^n \end{array}$$

where

- $m \leq n$;
- $f, h_1, \ldots, h_m : \Re^n \mapsto \Re$;
- $f, h_1, \ldots, h_m \in C^2(\Re^n)$.

In this chapter, we shall learn necessary or sufficient optimality conditions for above optimization problem without giving proofs.

6.1.2 First-Order Necessary Conditions

Lagrange Theorem

- Let x^* be a local extreme (either a maximum or a minimum) point of f subject to the constraints $h(x) = (h_1(x), \ldots, h_m(x))^T = 0$.
- Assume further that x^* satisfies certain regularity conditions. (e.g., all $\nabla h_i(x^*)$ are linearly independent. In this course it is not required to know these conditions in detail.)
- Then there exists

$$\lambda^* \in \Re^m$$

(called *Lagrange Multipliers*) such that

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0.$$

- A feasible point that satisfies the above condition is called a **stationary point**. Note that such a stationary point may be a constrained minimum point, or a constrained maximum point, or neither (i.e., only a saddle point).
- It should be noted that the first-order necessary conditions

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0 \quad (n \text{ equations})$$

together with the constraints

$$h(x^*) = 0 \quad (m \text{ equations})$$

gives $n + m$ (generally nonlinear) equations in the $n + m$ variables comprising x^* and λ^*.

- Define the *Lagrangian function* as follows:

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x).$$

The first-order necessary conditions together with the constraints can then be expressed in the form:

$$\nabla_x L(x^*, \lambda^*) = 0 \quad \text{and} \quad \nabla_\lambda L(x^*, \lambda^*) = 0,$$

or equivalently,

$$\nabla L(x^*, \lambda^*) = 0.$$

6.1.3 Second-Order Necessary Conditions

- Let x^* be a local minimum point of f subject to $h(x) = 0$. Assume that x^* satisfies certain regularity conditions. Then there exists
$$\lambda^* \in \Re^m$$
such that
$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0,$$
and
$$y^T \left[\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*) \right] y \geq 0$$
for all $y \in \Re^n$ satisfying
$$\begin{aligned} y &\neq 0 \\ \nabla h_i(x^*)^T y &= 0 \qquad \text{for} \quad i = 1, \ldots, m. \end{aligned}$$

- The second-order necessary conditions can be expressed in the form:

$$\nabla L(x^*, \lambda^*) = 0,$$

  and

$$y^T \nabla^2_{xx} L(x^*, \lambda^*) y \geq 0 \tag{1}$$

  for all $y \in \Re^n$ satisfying

$$\begin{aligned} y &\neq 0 \\ \nabla h_i(x^*)^T y &= 0 \qquad \text{for} \quad i = 1, \ldots, m. \end{aligned}$$

  Remark. The above condition for $\nabla^2_{xx} L(x^*, \lambda^*)$ is *weaker* than asking the matrix $\nabla^2_{xx} L(x^*, \lambda^*)$ to be positive semidefinite. Why?
- If x^* is a local maximum point of f subject to $h(x) = 0$, then in the above conditions, (1) is changed to

$$y^T \nabla^2_{xx} L(x^*, \lambda^*) y \leq 0.$$

6.1.4 Second-Order Sufficient Conditions

Just like in the unconstrained case, if, additionally, x^* satisfies

$$y^T \nabla^2_{xx} L(x^*, \lambda^*) y > 0$$

for all $y \in \Re^n$ satisfying

$$\begin{aligned} y &\neq 0 \\ \nabla h_i(x^*)^T y &= 0 \qquad \text{for} \quad i = 1, \ldots, m, \end{aligned}$$

then x^* is a strict local minimum point of f subject to $h(x) = 0$.

For the maximum case, if the above inequality is changed to

$$y^T \nabla^2_{xx} L(x^*, \lambda^*) y < 0$$

for such y, then x^* is a strict local maximum point of f subject to $h(x) = 0$.

6.1.5 Examples

Example 6.1 Solve the problem:

$$\begin{array}{llcccccccc} \text{minimize} & f(x) & = & x_1^2 & + & x_2^2 & & & & \\ \text{subject to} & h(x) & = & x_1 & + & x_2 & - & 1 & = & 0. \end{array}$$

Solution: The Lagrangian function is

$$L(x,\lambda) = x_1^2 + x_2^2 + \lambda(x_1 + x_2 - 1).$$

Thus,

$$\nabla L(x,\lambda) = \begin{bmatrix} 2x_1 + \lambda \\ 2x_2 + \lambda \\ x_1 + x_2 - 1 \end{bmatrix} \quad \text{and} \quad \nabla_{xx}^2 L(x,\lambda) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

The first-order necessary conditions are

$$
\begin{aligned}
2x_1^* + \lambda^* &= 0, \\
2x_2^* + \lambda^* &= 0, \\
x_1^* + x_2^* - 1 &= 0.
\end{aligned}
$$

We may solve the system of equations as

$$
\begin{bmatrix} x_1^* \\ x_2^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \\ -1 \end{bmatrix}.
$$

But usually we use the elimination method to solve these equations more quickly. In fact from the first two equations we know that $x_1^* = x_2^*$ (i.e., eliminate λ^*). Substituting this result into the third equation, we obtain $x_1^* = x_2^* = 0.5$. Finally, from the first or the second equation, we obtain $\lambda^* = -1$.

Since

$$\nabla^2_{xx} L(x^*, \lambda^*) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

is positive definite, (x^*, λ^*) satisfies the second-order sufficient conditions. Therefore

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

is a strict local minimum solution to the problem.

Example 6.2 Solve the problem:

$$\begin{array}{lrcl} \text{minimize} & f(x) & = & -x_1 x_2 - x_2 x_3 - x_1 x_3 \\ \text{subject to} & h(x) & = & x_1 + x_2 + x_3 - 3 = 0. \end{array}$$

Solution The Lagrangian function is

$$L(x,\lambda) = -x_1x_2 - x_2x_3 - x_1x_3 + \lambda(x_1 + x_2 + x_3 - 3).$$

Thus,

$$\nabla L(x,\lambda) = \begin{bmatrix} -x_2 - x_3 + \lambda \\ -x_1 - x_3 + \lambda \\ -x_1 - x_2 + \lambda \\ x_1 + x_2 + x_3 - 3 \end{bmatrix}, \nabla^2_{xx} L(x,\lambda) = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}.$$

The first-order necessary conditions are

$$\begin{aligned} -x_2^* - x_3^* + \lambda^* &= 0, \\ -x_1^* - x_3^* + \lambda^* &= 0, \\ -x_1^* - x_2^* + \lambda^* &= 0, \\ x_1^* + x_2^* + x_3^* - 3 &= 0, \end{aligned}$$

which has a solution $\left[x_1^*, \quad x_2^*, \quad x_3^*, \quad \lambda^*\right]^T = \left[1, \quad 1, \quad 1, \quad 2\right]^T$.

Since

$$\nabla^2_{xx} L(x^*, \lambda^*) = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

is not positive definite, we cannot make conclusion immediately.

Let $y \in \Re^3$ satisfying $y \neq 0$ and

$$0 = \nabla h(x^*)^T y = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = y_1 + y_2 + y_3.$$

For such y,

$$\begin{aligned} y^T \nabla^2_{xx} L(x^*, \lambda^*) y &= \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \\ &= -y_1(y_2 + y_3) - y_2(y_1 + y_3) - y_3(y_1 + y_2) \\ &= y_1^2 + y_2^2 + y_3^2 > 0. \end{aligned}$$

So, (x^*, λ^*) satisfies the second order sufficient conditions, and

$$\begin{bmatrix} x_1^*, & x_2^*, & x_3^* \end{bmatrix}^T = \begin{bmatrix} 1, & 1, & 1 \end{bmatrix}^T$$

is a strict local minimum solution to the problem.

Example 6.3 Solve the problem:

$$\begin{array}{lrcrcrcrcrcl} \text{minimize} & f(x) & = & x_1^2 & + & x_2^2 & + & x_3^2 & & & & \\ \text{subject to} & h_1(x) & = & x_1 & + & x_2 & + & 3x_3 & - & 2 & = & 0 \\ & h_2(x) & = & 5x_1 & + & 2x_2 & + & x_3 & - & 5 & = & 0. \end{array}$$

Solution The Lagrangian function is

$$\begin{aligned} L(x, \lambda) \quad = \quad & x_1^2 + x_2^2 + x_3^2 + \lambda_1(x_1 + x_2 + 3x_3 - 2) \\ & + \lambda_2(5x_1 + 2x_2 + x_3 - 5). \end{aligned}$$

Thus,

$$\nabla L(x, \lambda) = \begin{bmatrix} 2x_1 + \lambda_1 + 5\lambda_2 \\ 2x_2 + \lambda_1 + 2\lambda_2 \\ 2x_3 + 3\lambda_1 + \lambda_2 \\ x_1 + x_2 + 3x_3 - 2 \\ 5x_1 + 2x_2 + x_3 - 5 \end{bmatrix} \text{ and } \nabla_{xx}^2 L(x, \lambda) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

The first-order necessary conditions are

$$\begin{aligned}
2x_1^* + \lambda_1^* + 5\lambda_2^* &= 0, \\
2x_2^* + \lambda_1^* + 2\lambda_2^* &= 0, \\
2x_3^* + 3\lambda_1^* + \lambda_2^* &= 0, \\
x_1^* + x_2^* + 3x_3^* - 2 &= 0, \\
5x_1^* + 2x_2^* + x_3^* - 5 &= 0.
\end{aligned}$$

From the first three equations we eliminate variables λ_1^* and λ_2^*, and obtain the equation $5x_1^* - 14x_2^* + 3x_3^* = 0$. Then together with the last two equations we obtain x_1^*, x_2^* and x_3^*. Finally,

$(x_1^*,\ x_2^*,\ x_3^*,\ \lambda_1^*,\ \lambda_2^*) \approx (0.8043, 0.3478,\ 0.2826, -0.0870, -0.3043)$.

Since

$$\nabla^2_{xx} L(x^*, \lambda^*) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

is positive definite, (x^*, λ^*) satisfies the second-order sufficient conditions. Therefore

$$\begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} \approx \begin{bmatrix} 0.8043 \\ 0.3478 \\ 0.2826 \end{bmatrix}$$

is a strict local minimum solution to the problem.

6.2 Equality and Inequality Constrained Optimization

6.2.1 Problem Description

Problem:

$$\begin{array}{lrcll} \min & f(x) & & & \\ \text{s.t.} & h_i(x) & = & 0, & i = 1, \ldots, m \\ & g_j(x) & \leq & 0, & j = 1, \ldots, r \\ & x \in \Re^n & & & \end{array}$$

where

1. $m \leq n$;
2. $f, h_1, \ldots, h_m, g_1, \ldots, g_r : \Re^n \mapsto \Re$;
3. $f, h_1, \ldots, h_m, g_1, \ldots, g_r \in C^2(\Re^n)$.

6.2.2 First-Order Necessary Conditions

Karosh-Kuhn-Tucker (KKT) Conditions

- Let x^* be a local minimum point of f subject to the constraints
  $h(x) = (h_1(x), \dots, h_m(x))^T = 0$ and
  $g(x) = (g_1(x), \dots, g_r(x))^T \leq 0.$
- Assume further that x^* satisfies certain regularity conditions.
- Then there exists

$$\begin{array}{lll} \lambda_i^* \in \Re & \text{for} & i = 1, \dots, m \\ \mu_j^* \geq 0 & \text{for} & j = 1, \dots, r \end{array}$$

(λ^* and μ^* are called *KKT multipliers*) such that

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^{r} \mu_j^* \nabla g_j(x^*) = 0$$
$$\mu_j^* g_j(x^*) = 0 \quad \text{for} \quad j = 1, \dots, r$$

- A feasible point x^* that satisfies the above KKT conditions (together with a set of multipliers λ^* and μ^*) is called a **stationary point**, or a **KKT point**.
- Define the *Lagrangian function* as follows:

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

The first-order necessary conditions together with the equality constraints can then be expressed as the following system of equations:

$$\begin{array}{rcll} \nabla_x L(x^*, \lambda^*, \mu^*) & = & 0 & (n \text{ equations}) \\ h(x^*) & = & 0 & (m \text{ equations}) \\ \mu_j^* g_j(x^*) & = & 0 \ \text{ for } \ j = 1, \ldots, r & (r \text{ equations}) \end{array}$$

which contains $n + m + r$ equations for $n + m + r$ unknown variables $x^* \in R^n, \lambda^* \in R^m$, and $\mu^* \in R^r$.

- Note that the conditions also ask that

$$g_j(x^*) \leq 0, \;\; \mu_j^* \geq 0 \;\; \text{and} \;\; \mu_j^* g_j(x^*) = 0, \quad \text{for} \;\; j = 1, \ldots, r. \tag{2}$$

  The last request is called the *complementary slackness condition* which asks that in each pair of μ_j^* and $g_j(x^*)$, at least one of them must be 0.
  (This condition is a natural extension of the complementary slackness condition in linear programming.)
- If in (2), it is requested further that in each pair of μ_j^* and $g_j(x^*)$, *exactly one* is zero, it is called the *strict complementary slackness condition*. That is, if $g_j(x^*) < 0$, then $\mu_j^* = 0$, and if $g_j(x^*) = 0$, then $\mu_j^* > 0$.

6.2.3 Second-Order Necessary Conditions

- Suppose that x^* is a local minimum point of f subject to the constraints $h(x) = 0$ and $g(x) \leq 0$.
- Assume that x^* satisfies certain regularity conditions.
- Then there exists

$$\begin{array}{lll} \lambda_i^* \in \Re & \text{for} & i = 1, \dots, m, \\ \mu_j^* \geq 0 & \text{for} & j = 1, \dots, r, \end{array}$$

such that

$$\begin{array}{rcll} \nabla_x L(x^*, \lambda^*, \mu^*) & = & 0 & & (n \text{ equations}) \\ h(x^*) & = & 0 & & (m \text{ equations}) \\ \mu_j^* g_j(x^*) & = & 0 & \text{for } j = 1, \dots, r & (r \text{ equations}) \end{array}$$

and

$$y^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) y \geq 0 \qquad (3)$$

for all $y \in \Re^n$ satisfying

$$\begin{aligned} y &\neq 0 \\ \nabla h_i(x^*)^T y &= 0 \qquad \text{for} \qquad i = 1, \ldots, m \\ \nabla g_j(x^*)^T y &= 0 \qquad \text{for} \qquad j \in A(x^*) \end{aligned}$$

where $A(x^*) = \{j : g_j(x^*) = 0\}$.

We call $g_j(x) \le 0$ an *active or binding constraint* at x^* if the index $j \in A(x^*)$. For example, suppose the problem has two inequality constraints:

$$
\begin{aligned}
g_1(x) &= x_1 + 2x_2 - 3 \le 0, \\
g_2(x) &= 2x_1 + x_2 - 5 \le 0.
\end{aligned}
$$

At point $x^* = (1, 1)$,

$$
\begin{aligned}
g_1(x^*) &= x_1^* + 2x_2^* - 3 = 0, \\
g_2(x^*) &= 2x_1^* + x_2^* - 5 < 0.
\end{aligned}
$$

We see that at x^*, the first constraint is active, but the second one is not. Hence $A(x^*) = \{1\}$.

6.2.4 Second-Order Sufficient Conditions

The above conditions become sufficient if

1. the inequality (3) is further strengthened to

$$y^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) y > 0$$

   for all $y \in \Re^n$ satisfying

$$\begin{aligned} y &\neq 0 \\ \nabla h_i(x^*)^T y &= 0 \qquad \text{for} \qquad i = 1, \ldots, m \\ \nabla g_j(x^*)^T y &= 0 \qquad \text{for} \qquad j \in A(x^*), \end{aligned}$$

   and

2. $\mu_j^* > 0$ for $j \in A(x^*)$.

Note that Condition 2 is in fact the strict complementary slackness condition.

6.2.5 Examples

Example 6.4 Solve the problem:

$$\begin{array}{lcccccccccc} \text{minimize} & f(x) & = & x_1^2 & + & x_2^2 & & & & \\ \text{subject to} & g(x) & = & -x_1 & - & x_2 & + & 1 & \leq & 0. \end{array}$$

Solution The Lagrangian function is

$$L(x, \mu) = x_1^2 + x_2^2 + \mu(-x_1 - x_2 + 1).$$

Thus,

$$\nabla_x L(x, \mu) = \begin{bmatrix} 2x_1 - \mu \\ 2x_2 - \mu \end{bmatrix} \quad \text{and} \quad \nabla_{xx}^2 L(x, \mu) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

The KKT necessary conditions are

$$\begin{aligned} 2x_1^* - \mu^* &= 0, & (4) \\ 2x_2^* - \mu^* &= 0, & (5) \\ \mu^*(-x_1^* - x_2^* + 1) &= 0, & (6) \\ \mu^* &\geq 0. & (7) \end{aligned}$$

We now consider the following two cases.

Case 1: Suppose

$$-x_1^* - x_2^* + 1 < 0. \tag{8}$$

(8) and (6) imply

$$\mu^* = 0. \tag{9}$$

Substituting (9) into (4) and (5) gives

$$(x_1^*, x_2^*) = (0, 0). \tag{10}$$

Since (10) contradicts (8), case 1 is rejected.

Case 2: Suppose

$$-x_1^* - x_2^* + 1 = 0. \tag{11}$$

Rewrite (4), (5) and (11) as

$$\begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \\ \mu^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

We thus have

$$(x_1^*, x_2^*, \mu^*) = (0.5, 0.5, 1). \tag{12}$$

Since $\mu^* \geq 0$,

$$(x_1^*, x_2^*) = (0.5, 0.5)$$

is a KKT point.

This point may be a potential minimum solution to the problem.

Since

$$\nabla^2_{xx} L(x^*, \mu^*) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

is positive definite and $\mu^* > 0$, (x^*, μ^*) satisfies the second-order sufficient conditions. Therefore

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

is a strict local minimum solution to the problem.

Example 6.5 Solve the problem:

$$\begin{array}{lrcrcrcrcl} \text{minimize} & f(x) & = & x_1^3 & + & x_2^2 & & & & \\ \text{subject to} & g(x) & = & -x_1 & & & - & 1 & \leq & 0. \end{array}$$

Solution The Lagrangian function is

$$L(x,\mu) = x_1^3 + x_2^2 + \mu(-x_1 - 1).$$

Thus,

$$\nabla_x L(x,\mu) = \begin{bmatrix} 3x_1^2 - \mu \\ 2x_2 \end{bmatrix} \quad \text{and} \quad \nabla_{xx}^2 L(x,\mu) = \begin{bmatrix} 6x_1 & 0 \\ 0 & 2 \end{bmatrix}.$$

The KKT necessary conditions are

$$\begin{aligned} 3(x_1^*)^2 - \mu^* &= 0, && (13) \\ x_2^* &= 0, && (14) \\ \mu^*(-x_1^* - 1) &= 0, && (15) \\ \mu^* &\geq 0. && (16) \end{aligned}$$

We now consider the following two cases.

Case 1: Suppose

$$-x_1^* - 1 < 0. \tag{17}$$

(17) and (15) imply

$$\mu^* = 0. \tag{18}$$

Substituting (18) into (13) gives

$$x_1^* = 0. \tag{19}$$

As (19) satisfies (17), $(x_1^*, x_2^*) = (0, 0)$ with $\mu^* = 0$ is a potential minimum solution to the problem.

Since at (x^*, y^*), there is no any active constraint, the second order optimality condition can be simplified.

- necessary condition: $\nabla_{xx}^2 L(x^*, \mu^*)$ is p.s.d.;
- sufficient condition: $\nabla_{xx}^2 L(x^*, \mu^*)$ is p.d.

We consider

$$\nabla^2_{xx} L(x^*, \mu^*) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}. \tag{20}$$

As (20) is positive semi-definite, (x^*, μ^*) satisfies the second-order necessary conditions but does not satisfy the sufficient conditions to be a strict local minimizer, the identity of the point $(x_1^*, x_2^*) = (0, 0)$ is unknown from the above theorems. But in this particular example, it is easy to see that x^* is not a local minimizer. In fact if we take $\bar{x} = (-\epsilon, 0)^T$ with very small positive value ϵ, it is obvious that $\bar{x}$ is very close to x^* and feasible, but $f(\bar{x}) < f(x^*)$.

Case 2: Suppose

$$x_1^* = -1. \tag{21}$$

Substituting (21) into (13) gives

$$\mu^* = 3. \tag{22}$$

Since (22) satisfies (16),

$$(x_1^*, x_2^*) = (-1, 0)$$

with $\mu^* = 3$ is a potential minimum solution to the problem.

Note that $\mu^* > 0$ satisfies the sufficient condition to be a strict local minimum solution to the problem.

Consider

$$\nabla_{xx}^2 L(x^*, \mu^*) = \begin{bmatrix} -6 & 0 \\ 0 & 2 \end{bmatrix}. \tag{23}$$

Since (23) is indefinite, we cannot make conclusion immediately.

Since the constraint is active, we consider $y = (y_1, y_2) \neq (0, 0)$ such that

$$0 = \nabla g(x^*)^T y = \begin{bmatrix} -1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = -y_1. \tag{24}$$

This implies

$$y_2 \neq 0. \tag{25}$$

For such y,

$$\begin{aligned} y^T \nabla^2_{xx} L(x^*, \mu^*) y &= \begin{bmatrix} 0 & y_2 \end{bmatrix} \begin{bmatrix} -6 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ y_2 \end{bmatrix} \\ &= 2y_2^2 > 0. \end{aligned}$$

Therefore $(x_1^*, x_2^*) = (-1, 0)$ is a strict local minimum solution to the problem.

Example 6.6 Solve the problem:

$$\begin{array}{lrclclcll}
\text{minimize} & f(x) & = & x_1^2 & + & x_2^2 & & & \\
\text{subject to} & g_1(x) & = & -x_1 & - & x_2 & + & 1 & \leq & 0, \\
& g_2(x) & = & -x_1 & & & + & 2 & \leq & 0.
\end{array}$$

Solution The Lagrangian function is

$$L(x,\mu) = x_1^2 + x_2^2 + \mu_1(-x_1 - x_2 + 1) + \mu_2(-x_1 + 2).$$

Thus,

$$\nabla_x L(x,\mu) = \begin{bmatrix} 2x_1 - \mu_1 - \mu_2 \\ 2x_2 - \mu_1 \end{bmatrix} \quad \text{and} \quad \nabla_{xx}^2 L(x,\mu) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

The KKT necessary conditions are

$$\begin{aligned}
2x_1^* - \mu_1^* - \mu_2^* &= 0, && (26)\\
2x_2^* - \mu_1^* &= 0, && (27)\\
\mu_1^*(-x_1^* - x_2^* + 1) &= 0, && (28)\\
\mu_2^*(-x_1^* + 2) &= 0, && (29)\\
\mu_1^* &\geq 0, && (30)\\
\mu_2^* &\geq 0. && (31)
\end{aligned}$$

We now consider the following four cases.

Case 1: Suppose that

$$
\begin{aligned}
-x_1^* - x_2^* + 1 &< 0, \text{ and} \qquad (32)\\
-x_1^* + 2 &< 0. \qquad (33)
\end{aligned}
$$

(32) and (28) imply

$$\mu_1^* = 0. \tag{34}$$

(33) and (29) imply

$$\mu_2^* = 0. \tag{35}$$

Rewrite (26), (27), (34) and (35) as

$$
\begin{bmatrix} 2 & 0 & -1 & -1 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} x_1^* \\ x_2^* \\ \mu_1^* \\ \mu_2^* \end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},
$$

and we thus have

$$(x_1^*, x_2^*, \mu_1^*, \mu_2^*) = (0, 0, 0, 0). \tag{36}$$

Since (36) contradicts (32) and (33), case 1 is rejected.

Case 2: Suppose that

$$
\begin{aligned}
-x_1^* - x_2^* + 1 &= 0, \text{ and} && (37)\\
-x_1^* + 2 &< 0. && (38)
\end{aligned}
$$

(38) and (29) imply

$$\mu_2^* = 0. \tag{39}$$

Rewrite (26), (27), (37) and (39) as

$$
\begin{bmatrix} 2 & 0 & -1 & -1 \\ 0 & 2 & -1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} x_1^* \\ x_2^* \\ \mu_1^* \\ \mu_2^* \end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.
$$

We thus have

$$(x_1^*, x_2^*, \mu_1^*, \mu_2^*) = (0.5, 0.5, 1, 0). \tag{40}$$

Since (40) contradicts (38), case 2 is rejected.

Case 3: Suppose that

$$\begin{aligned} -x_1^* - x_2^* + 1 &< 0, \text{ and} && (41) \\ x_1^* &= 2. && (42) \end{aligned}$$

(41) and (28) imply

$$\mu_1^* = 0. \tag{43}$$

Rewrite (26), (27), (42) and (43) as

$$\begin{bmatrix} 2 & 0 & -1 & -1 \\ 0 & 2 & -1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \\ \mu_1^* \\ \mu_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 0 \end{bmatrix}.$$

We thus have

$$(x_1^*, x_2^*, \mu_1^*, \mu_2^*) = (2, 0, 0, 4). \tag{44}$$

Since (44) satisfies (31) and (41),

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} \mu_1^* \\ \mu_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$

is a potential minimum solution to the problem.

As

$$\nabla_{xx}^2 L(x^*, \mu^*) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

is positive definite, and only the second constraint is active which corresponds to the multiplier $\mu_2^* > 0$, we see that (x^*, μ^*) satisfies the sufficient conditions. Therefore $(x_1^*, x_2^*) = (2, 0)$ is a strict local minimum solution to the problem.

Case 4: Suppose that

$$
\begin{aligned}
-x_1^* - x_2^* + 1 &= 0, \text{ and} \qquad (45)\\
x_1^* &= 2. \qquad (46)
\end{aligned}
$$

Rewrite (26), (27), (45) and (46) as

$$
\begin{bmatrix} 2 & 0 & -1 & -1 \\ 0 & 2 & -1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}
\begin{bmatrix} x_1^* \\ x_2^* \\ \mu_1^* \\ \mu_2^* \end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}.
$$

We thus have

$$
(x_1^*, x_2^*, \mu_1^*, \mu_2^*) = (2, -1, -2, 6). \qquad (47)
$$

As (47) violates (30), case 4 is rejected.

To summarize, this example has a unique local minimum solution $x^* = (2, 0)^T$.

So far we assume that all inequality constraints are $g_j(x) \le 0$ type. If all inequality constraints are $g_j(x) \ge 0$ type, i.e., the problem is

$$\begin{array}{llcll} \min & f(x) & & & \\ \text{s.t.} & h_i(x) & = & 0, & i = 1, \ldots, m, \\ & g_j(x) & \ge & 0, & j = 1, \ldots, r, \\ & x \in \Re^n, & & & \end{array}$$

how to revise the optimality conditions? We may rewrite the constraints as

$$-g_j(x) \le 0, \quad j = 1, \ldots, r,$$

then use the result of this section. In fact now the first order necessary condition asks that

$$\nabla f(x^*) - \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) - \sum_{j=1}^{r} \mu_j^* \nabla g_j(x^*) = 0$$

$$\mu_j^* g_j(x^*) = 0 \quad \text{for} \quad j = 1, \ldots, r.$$

Here the sign of the first sum can be written either as "+", or "-", because the signs of multipliers λ_i^* have no restriction. Or we can define the *Lagrangian function* as follows:

$$\bar{L}(x, \lambda, \mu) = f(x) - \sum_{i=1}^{m} \lambda_i h_i(x) - \sum_{j=1}^{r} \mu_j g_j(x),$$

then in the first order necessary condition, $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$ should be replaced by:

$$\nabla_x \bar{L}(x^*, \lambda^*, \mu^*) = 0,$$

and in the second order conditions, the matrix $\nabla^2_{xx} L$ should be changed to $\nabla^2_{xx} \bar{L}(x^*, \lambda^*, \mu^*)$.

Suppose a minimization problem has inequality constraints only, for example,

$$\begin{array}{ll} \min & f(x) \\ s.t. & g_1(x) \leq 0; \\ & g_2(x) \leq 0. \end{array}$$

Here the feasible region is

$$S = \{x \mid g_1(x) \leq 0;\ g_2(x) \leq 0\}.$$

If at the optimal solution x^*, there is no active constraint, i.e.,

$$g_1(x^*) < 0,\ g_2(x^*) < 0,$$

then by KKT condition, we know that

$$\mu_1^* g_1(x^*) = 0 \Longrightarrow \mu_1^* = 0,$$

$$\mu_2^* g_2(x^*) = 0 \Longrightarrow \mu_2^* = 0.$$

So,

$$\nabla f(x^*) + \sum_{j=1}^{2} \mu_j^* \nabla g_j(x^*) = 0 \Longrightarrow \nabla f(x^*) = 0.$$

This result is correct because in this case there is a neighborhood $N(x^*, \epsilon)$ of x^* such that

$$x \in N(x^*, \epsilon) \Longrightarrow g_1(x) < 0,\ g_2(x) < 0,$$

that is,

$$N(x^*, \epsilon) \subset S.$$

Since x^* is a minimum point of f over the feasible region S,

$$f(x) \geq f(x^*) \text{ for all } x \in N(x^*, \epsilon).$$

This means that x^* is an local **unconstrained** minimum point. Therefore,

$$\nabla f(x^*) = 0.$$

It tells us that for an inequality constrained minimization problem, if at the optimal solution x^*, $\nabla f(x^*) \neq 0$, then there must be some active constraints, i.e., x^* is on some boundary of the feasible region.