OPTIMIZATION

Seungjin Han

September, 2017

Optimization Problem

- Maximization Problem

$$\max_x f(x) \quad \text{subject to} \quad x \in C, \tag{1}$$

where C is the constraint set and x is the choice variable.

- Minimization Problem

$$\min_x g(x) \quad \text{subject to} \quad x \in D, \tag{2}$$

where D is the constraint set and x is the choice variable.

- Let x^* be a solution to problem (1). By definition of x^*, $f(x^*) \geq f(x)$ for all $x \in C$. x^* is a (global) maximizer of f subject to $x \in C$ and $f(x^*)$ is the maximum of f subject to $x \in C$.

- x' is a local maximizer of f subject to $x \in C$ if there is a number $\epsilon > 0$ such that $f(x') \geq f(x)$ for all $x \in C$ such that the distance between x and x' is at most ϵ

- Any global maximizers are local maximizers.

- Note that the following two problems are equivalent

$$\begin{aligned} \min_x g(x) \quad \text{subject to} \quad x &\in D \\ &\Leftrightarrow \max_x(-g(x)) \quad \text{subject to} \quad x \in D \end{aligned}$$

- **Extreme Value Theorem**

  A function $f : X \to \mathbb{R}$ has a maximizer and a minimizer if

  1. f is continuous
  2. $X \subset \mathbb{R}^n$ is nonempty and compact

Optimization without Constraint: General Method

- Consider a function $f : X \to \mathbb{R}$ and the maximization problem

$$\max_x f(x)$$

- Suppose that f is differentiable and $X = [\underline{x}, \bar{x}]$
- x is a stationary point x if $f'(x) = 0$
- Being a stationary point is neither a necessary condition nor a sufficient condition for finding the solution
- Suppose that $f : X \to \mathbb{R}$ is differentiable and $X = [\underline{x}, \bar{x}]$. If $x \in \mathsf{Int}[\underline{x}, \bar{x}]$ is a global (or local) maximizer (or minimizer) of f, then $f'(x) = 0$
- General Method for a one-variable function: How to find a solution to $\max_x f(x)$. Assume that $f : X \to \mathbb{R}$ is differentiable and $X = [\underline{x}, \bar{x}]$

  1. Find all stationary points in X and values of f
  2. Find values of f at the endpoints of X
  3. Compare functional values of points in 1 and 2 for global maximizers.

- Example: $y = f(x) = -2(x-1)^2 \qquad \text{on } x \in [0, 2]$

  $f'(x) = -4(x-1) = 0$ so that $x = 1$ is the stationary point. $f(1) = 0, f(0) = -2$ and $f(2) = -2$ so the global maximizer is $x = 1$.

- Suppose that $f : X \to \mathbb{R}$ is differentiable and $X \subset \mathbb{R}^n$ is a compact set. If $x \in \mathsf{Int}X$ is a global (or local) maximizer (or minimizer), then
  $f_1(x) = 0, \quad f_2(x) = 0, \quad \ldots, f_n(x) = 0$
- General Method for multi-variable case: How to find a solution to $\max_x f(x)$. Assume that $f : X \to \mathbb{R}$ is differentiable and $X \subset \mathbb{R}^n$ is a compact set.

  1. Find all stationary points in X and values of f at the stationary points
  2. Find values of f at all the boundary points of X
  3. Compare functional values of points in 1 and 2 for global maximizers

- Note: Suppose X is not a compact set. Then, we may not have a global maximizer even if f is differentiable.

  Example: $f : X \to \mathbb{R}$ where $X = \mathbb{R} = (-\infty, \infty)$ and $f(x) = x^2$ for all $x \in X$

- Sometimes, it is hard to find the values of f at all the boundary points in $X \subset \mathbb{R}^n$

Definition: Concavity/Convexity of a Function

- **Convex Set**

  A set $C \subset \mathbb{R}^n$ is convex if, for all $x, x' \in C$ and all $\lambda \in [0, 1]$

  $$\lambda x + (1 - \lambda)x' \in C$$

  Example: $[0, 1]$ is a convex set

- **Concave Function**

  A function $f : X \to \mathbb{R}$ defined on the convex set $X \subset \mathbb{R}^n$ is concave if, for all $x, x' \in X$ and all $\lambda \in [0, 1]$

  $$f(\lambda x + (1 - \lambda)x') \geq \lambda f(x) + (1 - \lambda)f(x')$$

- **Strictly Concave Function**

  A function $f : X \to \mathbb{R}$ defined on the convex set $X \subset \mathbb{R}^n$ is strictly concave if, for all $x, x' \in X$ such that $x \neq x'$ and all $\lambda \in (0, 1)$

  $$f(\lambda x + (1 - \lambda)x') > \lambda f(x) + (1 - \lambda)f(x')$$

- **Convex Function**

  A function $f : X \to \mathbb{R}$ defined on the convex set $X \subset \mathbb{R}^n$ is convex if, for all $x, x' \in X$ and all $\lambda \in [0, 1]$

  $$f(\lambda x + (1 - \lambda)x') \leq \lambda f(x) + (1 - \lambda)f(x')$$

- **Strictly Convex Function**

  A function $f : X \to \mathbb{R}$ defined on the convex set $X \subset \mathbb{R}^n$ is strictly convex if, for all $x, x' \in X$ such that $x \neq x'$ and all $\lambda \in (0, 1)$

  $$f(\lambda x + (1 - \lambda)x') < \lambda f(x) + (1 - \lambda)f(x')$$

One-Variable Function

- Consider a function $f : X \to R$ with $X \subset \mathbb{R}$
- A twice continuously differentiable function f is (strictly) concave if and only if $f''(x) \leq 0 (f''(x) < 0)$ for all $x \in \mathsf{Int} X$

- A twice continuously differentiable function f is (strictly) convex if and only if $f''(x) \geq 0 (f''(x) > 0)$ for all $x \in \mathsf{Int}X$.

Multi-Variable Function

- Consider a function $f : X \rightarrow R$ with $X \subset \mathbb{R}^n$

$$y = f(x_1, \ldots, x_n)$$

- Example with two variables; $y = f(x_1, x_2)$

  When the function is twice differentiable, we have

$$\begin{aligned} dy &= f_1 dx_1 + f_2 dx_2 \\ d(dy) &= \frac{\partial dy}{\partial x_1} dx_1 + \frac{\partial dy}{\partial x_2} dx_2 \\ &= (f_{11} dx_1 + f_{21} dx_2) dx_1 + (f_{12} dx_1 + f_{22} dx_2) dx_2 \\ &= f_{11} dx_1^2 + 2 f_{12} dx_1 dx_2 + f_{22} dx_2^2 \\ &= \begin{bmatrix} dx_1 & dx_2 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} \end{aligned}$$

- PD, ND, PSD, and NSD

  1. d^2y is positive definite if $d^2y > 0$ at $dx_1 \neq 0$ and $dx_2 \neq 0$
  2. d^2y is negative definite if $d^2y < 0$ at $dx_1 \neq 0$ and $dx_2 \neq 0$
  3. d^2y is positive semidefinite if $d^2y \geq 0$ at any (dx_1, dx_2)
  4. d^2y is negative semidefinite if $d^2y \leq 0$ at any (dx_1, dx_2)

- In the example, rearranging d^2y yields

$$\begin{aligned} d^2y &= f_{11} \left(dx_1^2 + 2 \frac{f_{12}}{f_{11}} dx_1 dx_2 + \frac{f_{12}^2}{f_{11}^2} dx_2^2 \right) + \left(f_{22} - \frac{f_{12}^2}{f_{11}} \right) dx_2^2 \\ &= f_{11} \left(dx_1 + \frac{f_{12}}{f_{11}} dx_2 \right)^2 + \left(\frac{f_{11} f_{22} - f_{12}^2}{f_{11}} \right) dx_2^2 \end{aligned}$$

Strict Concavity/Convexity of a Multi-Variable Function

- Two Variable Case: Characterization of PD and ND

$$H = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$$

The first leading principle minor is $|D_1| = f_{11}$.

The second leading principle minor $|D_2| = f_{11}f_{22} - f_{12}f_{21}$.

1. d^2y is positive definite iff $f_{11} > 0$ and $f_{11}f_{22} - f_{12}^2 > 0$ at all $(x_1, x_2) \in \mathsf{Int}X$
2. d^2y is negative definite iff $f_{11} < 0$ and $f_{11}f_{22} - f_{12}^2 > 0$ at all $(x_1, x_2) \in \mathsf{Int}X$

- General Case: $y = f(x_1, \ldots, x_n)$

$$H = \begin{bmatrix} f_{11} & \ldots & f_{1n} \\ \vdots & & \vdots \\ f_{n1} & \ldots & f_{nn} \end{bmatrix}$$

$$|D_1| = |f_{11}| = f_{11}$$

$$|D_2| = \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} = f_{11}f_{22} - f_{12}^2$$

$$|D_3| = \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix}$$

$$\vdots$$

$$|D_n| = |H|$$

- In general,

1. d^2y is PD iff $|D_1| > 0, |D_2| > 0, \ldots, |D_n| > 0$ at every $(x_1, \ldots, x_n) \in \mathsf{Int}X$
2. d^2y is ND iff $|D_1| < 0, |D_2| > 0, \ldots, (-1)^n |D_n| > 0$ at every $(x_1, \ldots, x_n) \in \mathsf{Int}X$

- When $y = f(x_1, \ldots, x_n)$ be twice differentiable

1. f is strictly convex iff d^2y is PD at every $(x_1, \ldots, x_n) \in \mathsf{Int}X$.

2. f is strictly concave iff d^2y is ND at every $(x_1, \ldots, x_n) \in \mathsf{Int}X$.
   Example: $f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 + 2x_1x_3$

$$H = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 6 \end{bmatrix}$$

   and dy^2 is PD.

Concavity/Convexity of a Multi-Variable Function

- The kth order leading principle minor of an $n \times n$ symmetric matrix is the determinant of the matrix obtained by deleting the last $n - k$ rows and $n - k$ columns. Consider H with $n \times n$. The kth order leading principal minor is

$$|D_k| = \begin{vmatrix} f_{11} & f_{12} \cdots & f_{1k} \\ \vdots & & \vdots \\ f_{k1} & f_{k2} \cdots & f_{kk} \end{vmatrix}$$

- A kth order principle minor of an $n \times n$ symmetric matrix is the determinant of a $k \times k$ matrix obtained by deleting $n - k$ rows and the corresponding $n - k$ columns

  Example:

$$H = \begin{bmatrix} 0 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & -6 \end{bmatrix}$$

- d^2y is positive semidefinite iff all principal minors are nonnegative
- d^2y is negative semidefinite iff all the kth order principal minors are (i) nonnegative if k is even and (ii) nonpositive if k is odd
- Let $y = f(x_1, \ldots, x_n)$ be twice differentiable.

  1. f is convex iff d^2y is PSD at every $(x_1, \ldots, x_n) \in \mathsf{Int}X$
  2. f is convex iff d^2y is NSD at every $(x_1, \ldots, x_n) \in \mathsf{Int}X$

  Example: principle minors

Optimization without Constraint: Local maximizer/minimizer by using Concavity/Convexity

- (One variable function) Let $f : X \to \mathbb{R}$ with $X \subset \mathbb{R}$ be twice differentiable with continuous f' and f''. Suppose that x^* is a stationary point in $\mathsf{Int}X$ ($f'(x^*) = 0$)
  - If $f''(x^*) < 0$, then x^* is a local maximizer
  - If x^* is a local maximizer, $f''(x^*) \leq 0$
  - If $f''(x^*) > 0$, then x^* is a local minimizer
  - If x^* is a local minimizer, $f''(x^*) \geq 0$
  - If $f''(x^*) = 0$, then we do not know whether x is a local maximizer or minimizer without further investigation.

  Example: $f(x) = x^3 - 12x^2 + 36x + 8$

- (Multi variable function) Let $f : X \to \mathbb{R}$ with $X \subset \mathbb{R}^n$ be twice differentiable with continuous f_{ij} for all i, j. Suppose that $x^* \in \mathsf{Int}X$ is a stationary point ($f_i(x^*) = 0$ for all i)
  - If H is negative definite at $x = x^*$, then x^* is a local maximizer
  - If x^* is a local maximizer, then H is negative semidefinite at $x = x^*$
  - If H is positive definite at $x = x^*$, then x^* is a local minimizer
  - If x^* is a local minimizer, then H is positive semidefinite at $x = x^*$

  Example: $y = f(x) = -x_1^3 + 3x_1x_3 + 2x_2 - x_2^2 - 3x_3^2$

Optimization without Constraint: global maximizer/minimizer by using Concavity/Convexity

- (One-variable function) Let $f : X \to \mathbb{R}$ with a convex set $X \subset \mathbb{R}$ be a differentiable function
  1. If f is a concave function and $x^* \in \mathsf{Int}X$ is a stationary point of f, then x^* is a global maximizer
  2. If f is a convex function and $x^* \in \mathsf{Int}X$ is a stationary point of f, then x^* is a global minimizer

- (One-variable function) Let $f : X \to \mathbb{R}$ with $X \subset \mathbb{R}$ be a twice differentiable function.

  1. If $f''(x) \leq 0$ for all $x \in X$ and $x^* \in \mathsf{Int}X$ is a stationary point of f, then x^* is a global maximizer
  2. If $f''(x) \geq 0$ for all $x \in X$ and $x^* \in \mathsf{Int}X$ is a stationary point of f, then x^* is a global minimizer

- Example: $f(x) = -2(x-1)^2$ with the domain $X = \mathbb{R}$

- (Multi-variable function) Let $f : X \to \mathbb{R}$ with a convex set $X \subset \mathbb{R}^n$ be a differentiable function

  1. If f is concave and $x^* \in \mathsf{Int}X$ is a stationary point, then x^* is a global maximizer
  2. If f is convex and $x^* \in \mathsf{Int}X$ is a stationary point, then x^* is a global minimizer

- (Multi-variable case) Let $f : X \to \mathbb{R}$ with a convex set $X \subset \mathbb{R}$ be a twice differentiable function with continuous f_{ij} for all i, j

  1. If f has negative semidefinite H at all $x \in X$ and $x^* \in \mathsf{Int}X$ is a stationary point, then x^* is a global maximizer
  2. If f has positive semidefinite H at all $x \in X$ and $x^* \in \mathsf{Int}X$ is a stationary point, then x^* is a global minimizer

- Example: A firm that produces two goods

  $P_1 = 12, P_2 = 18$

  $r = P_1x_1 + P_2x_2$

  $c(x_1, x_2) = 2x_1^2 + x_1x_2 + 2x_2^2$

  $\pi : X \to \mathbb{R}$ where $X = \mathbb{R}^2_+$

- Example: Firm's profit maximization

  $Q(K, L) = L^\alpha K^\alpha \qquad \alpha < \frac{1}{2}$.

- Example: A monopolist facing the three different markets

  $R = R_1(Q_1) + R_2(Q_2) + R_3(Q_3)$

  $C = C(Q)$ where $Q = Q_1 + Q_2 + Q_3$.

$\pi = R_1(Q_1) + R_2(Q_2) + R_3(Q_3) - C(Q_1 + Q_2 + Q_3)$

Let $P_1 = 63 - 4Q_1$, $P_2 = 105 - 5Q_2$, $P_3 = 75 - 6Q_3$, and $C = 20 + 15Q$.

Optimization with Equality Constraints: Intuition with a Single Constraint

- $f : X \to \mathbb{R}$ where $X \subset \mathbb{R}^n$

$$\begin{aligned} \max_{x \in X} f(x) \qquad \text{subject to } g(x) &= c \\ \min_{x \in X} f(x) \qquad \text{subject to } g(x) &= c \end{aligned}$$

- Consider a two variable function $f : X \to \mathbb{R}$with $X \subset \mathbb{R}^2$ for the maximization problem: $\max_{x \in X} f(x) \qquad$ subject to $g(x) = c$

  Assume f and g are differentiable. Suppose that f is increasing in x. Consider a level curve of f for a

$$L(a) = \{x \in X; f(x) = a\}$$

- Suppose that the maximal value of the function f is a^* at a solution (x_1^*, x_2^*) for the maximization problem. Then, the constraint curve is tangent to $L(a^*)$ at (x_1^*, x_2^*).

$$\begin{aligned} g(x_1, x_2) - c = 0 \\ \Rightarrow g_1(x_1, x_2)dx_1 + g_2(x_1, x_2)dx_2 = 0 \\ \Rightarrow \frac{dx_2}{dx_1} = -\frac{g_1(x_1^*, x_2^*)}{g_2(x_1^*, x_2^*)} \end{aligned}$$

  Furthermore,

$$\begin{aligned} f(x_1, x_2) - a^* &= 0 \\ &\Rightarrow f_1(x_1, x_2)dx_1 + f_2(x_1, x_2)dx_2 = 0 \\ &\Rightarrow \frac{dx_2}{dx_1} = -\frac{f_1(x_1^*, x_2^*)}{f_2(x_1^*, x_2^*)} \end{aligned}$$

- Because the constraint curve is tangent to $L(a^*)$ at (x_1^*, x_2^*), we have

$$-\frac{g_1(x_1^*, x_2^*)}{g_2(x_1^*, x_2^*)} = \frac{dx_2}{dx_1} = -\frac{f_1(x_1^*, x_2^*)}{f_2(x_1^*, x_2^*)}$$

- Letting $\frac{f_1(x_1^*,x_2^*)}{g_1(x_1^*,x_2^*)} = \frac{f_2(x_1^*,x_2^*)}{g_2(x_1^*,x_2^*)} = \lambda^*$, the (first-order) necessary conditions are

$$\begin{aligned} f_1(x_1^*, x_2^*) - \lambda^* g_1(x_1^*, x_2^*) &= 0 \qquad (3) \\ f_2(x_1^*, x_2^*) - \lambda^* g_2(x_1^*, x_2^*) &= 0 \qquad (4) \\ c - g(x_1^*, x_2^*) &= 0 \qquad (5) \end{aligned}$$

- Set up the Lagrangian function as

$$L(x_1, x_2) = f(x_1, x_2) + \lambda[c - g(x_1, x_2)]$$

- Take the derivatives of $L(x_1, x_2)$ with respect to x_1, x_2 and λ. Their values at the solution must be

$$\begin{aligned} \frac{\partial L}{\partial x_1}(x_1^*, x_2^*) &= f_1(x_1^*, x_2^*) - \lambda^* g_1(x_1^*, x_2^*) = 0 \\ \frac{\partial L}{\partial x_2}(x_1^*, x_2^*) &= f_2(x_1^*, x_2^*) - \lambda^* g_2(x_1^*, x_2^*) = 0 \\ \frac{\partial L}{\partial \lambda}(x_1^*, x_2^*) &= c - g(x_1^*, x_2^*) = 0 \end{aligned}$$

- Interpretation of the Lagrangian multiplier. Let $(x_1^*(c), x_2^*(c))$ be a solution for $\max_{x_1,x_2} f(x_1, x_2)$ subject to $g(x_1, x_2) = c$. Taking the derivative of the maximum value function $f(x_1^*(c), x_2^*(c))$ with respect to c yields

$$\begin{aligned} \frac{df(x_1^*(c), x_2^*(c))}{dc} &= f_1(x_1^*(c), x_2^*(c))\frac{\partial x_1^*}{\partial c} + f_2(x_1^*(c), x_2^*(c))\frac{\partial x_2^*}{\partial c} \\ &= \lambda^* \left[g_1(x_1^*(c), x_2^*(c))\frac{\partial x_1^*}{\partial c} + g_2(x_1^*(c), x_2^*(c))\frac{\partial x_2^*}{\partial c}\right] \\ &= \lambda^* \end{aligned}$$

  The second equation comes from the first-order conditions, $f_1(x_1^*, x_2^*) = \lambda^* g_1(x_1^*, x_2^*)$ and $f_2(x_1^*, x_2^*) = \lambda^* g_2(x_1^*, x_2^*)$. The third equality can be shown by taking the total differential of the constraint:

$$\begin{aligned} & g(x_1^*(c), x_2^*(c)) = c \\ \Rightarrow\ & g_1(x_1^*(c), x_2^*(c))\frac{\partial x_1^*}{\partial c} + g_2(x_1^*(c), x_2^*(c))\frac{\partial x_2^*}{\partial c} = 1 \end{aligned}$$

- λ^* measures the effect of a marginal change in the constraint via c on the optimal value of the objective function.

- Example: $\max f(x_1, x_2)$ subject to $x_1 + x_2 = 6$, where $f(x_1, x_2) = x_1 x_2$.

Optimization with Equality Constraints: Multiple Equality Constraints

- Consider a multi-variable with $f : X \to \mathbb{R}$ with $X \subset \mathbb{R}^n$. Optimization problems with multiple equality constraints are

$$\max_{x \in X} f(x) \text{ or } \min_{x \in X} f(x) \text{ subject to } \begin{cases} g^1(x) = c^1 \\ \vdots \\ g^m(x) = c^m \end{cases} \tag{6}$$

- The Lagrangian function for problem (6) is

$$\begin{aligned} L &= f(x) + \lambda_1[c^1 - g'(x)] + \lambda_2[c^2 - g^2(x)] + \cdots + \lambda_m[c^m - g^m(x)] \quad (7) \\ &= f(x) + \sum_{j=1}^{m} \lambda_j [c^j - g^j(x)] \end{aligned}$$

- The first-order conditions for (7)

$$\frac{\partial L}{\partial x_i} = 0 \qquad \text{for } i = 1, \ldots, n \tag{8}$$

$$\frac{\partial L}{\partial \lambda_i} = 0 \qquad \text{for } j = 1, \ldots, m \tag{9}$$

Optimization with Equality Constraints: Local maximizer (minimizer)

- **Necessary condition for a local maximizer (minimizer)**

  If $x^* = (x_1^*, \ldots, x_n^*) \in \mathsf{Int}X$ is a local max (min), then $x^* = (x_1^*, \ldots, x_n^*)$ and $\lambda^* = [\lambda_1^*, \ldots, \lambda_n^*]$ satisfy (8) and (9).

- **Sufficient condition for a local maximizer (minimizer): Two-Variable and Single-Equality Constraint Case**

  Consider a two-variable objective function and a single constraint:

$$\max_{(x_1, x_2)} f(x_1, x_2) \text{ subject to } g(x_1, x_2) = c.$$

Let $x_2 = h(x_1)$ be the value of the second choice variable such that $g(x_1, h(x_1)) = c$ for all x. Let $F(x_1) = f(x_1, h(x_1))$. Then, the maximization problem becomes $\max_{x_1} F(x_1)$. Taking the first-order derivative yields

$$F'(x_1) = f_1(x_1, h(x_1)) + f_2(x_1, h(x_1))h'(x_1)$$

Taking the second-order derivative yields

$$\begin{aligned} F''(x_1) &= f_{11}(x_1, h(x_1)) + f_{12}(x_1, h(x_1))h'(x_1) + \\ & \quad f_{21}(x_1, h(x_1))h'(x_1) + f_{22}(x_1, h(x_1))(h'(x_1))^2 \\ & \quad + f_2(x_1, h(x_1))h''(x_1) \end{aligned}$$

$h''(x_1)$ can be derived from taking the second order derivative of $g(x_1, h(x_1)) = c$ with respect to x_1. It is messy but eventually we can express $F''(x_1^*)$ at (x_1^*, x_2^*) as

$$F''(x_1^*) = \frac{-|\overline{H}|}{g_2(x_1, h(x_1^*))^2},$$

where $\left|\overline{H}\right|$ is the Bordered Hessian of the Lagrangian.

$$\left|\overline{H}\right| = \begin{vmatrix} 0 & g_1(x_1^*, x_2^*) & g_2(x_1^*, x_2^*) \\ g_1(x_1^*, x_2^*) & f_{11}(x_1^*, x_2^*) - \lambda^* g_{11}(x_1^*, x_2^*) & f_{12}(x_1^*, x_2^*) - \lambda^* g_{12}(x_1^*, x_2^*) \\ g_2(x_1^*, x_2^*) & f_{21}(x_1^*, x_2^*) - \lambda^* g_{21}(x_1^*, x_2^*) & f_{22}(x_1^*, x_2^*) - \lambda^* g_{22}(x_1^*, x_2^*) \end{vmatrix}$$

- **Sufficient condition for a local maximizer (minimizer): Two-Variable and Single-Equality Constraint Case**

  Suppose that (x_1^*, x_2^*) and λ^* satisfy the first-order conditions

$$\begin{aligned} f_1(x_1^*, x_2^*) - \lambda^* g_1(x_1^*, x_2^*) &= 0, \\ f_2(x_1^*, x_2^*) - \lambda^* g_2(x_1^*, x_2^*) &= 0, \\ g(x_1^*, x_2^*) &= c. \end{aligned}$$

  If $\left|\overline{H}\right| > 0$, then (x_1^*, x_2^*) is a local maximizer of f subject to $g(x_1, x_2) = c$.

  If $\left|\overline{H}\right| < 0$, then (x_1^*, x_2^*) is a local minimizer of f subject to $g(x_1, x_2) = c$.

- **Sufficient condition for a local maximizer (minimizer): Multi-Variable and Single-Equality Constraint Case**

  The Lagrangian function is $L = f(x) + \lambda[c - g(x)]$.

$$
\left|\overline{H_i}\right| = \begin{vmatrix} 0 & g_1 & g_2 & \dots & g_i \\ g_1 & f_{11}-\lambda g_{11} & f_{12}-\lambda g_{12} & \dots & f_{1i}-\lambda g_{1i} \\ g_2 & f_{21}-\lambda g_{21} & f_{22}-\lambda g_{22} & \dots & f_{2i}-\lambda g_{2i} \\ \vdots & \vdots & \vdots & & \vdots \\ g_i & f_{i1}-\lambda g_{i1} & f_{i2}-\lambda g_{i2} & \dots & f_{ii}-\lambda g_{ii} \end{vmatrix}
$$

$\left|\overline{H_n}\right| = \left|\overline{H}\right|$ (Bordered Hessian of the Lagrangian)

- **Sufficient condition for a local maximizer (minimizer): Multi-Variable and Single-Equality Constraint Case**

  Suppose that $x^* = [x_1^*, \dots, x_n^*]$ and λ^* satisfy the first-order conditions$f$

$$
\begin{aligned} f_i(x^*) - \lambda g_i(x^*) &= 0 \text{ for } i = 1, \dots, n \\ g(x^*) - c &= 0 \end{aligned}
$$

  If $\left|\overline{H_2}\right| > 0, \left|\overline{H_3}\right| < 0, \dots, (-1)^n \left|\overline{H_n}\right| > 0$ at (x^*, λ^*), then x^* is a local maximizer of f subject to $g(x) = c$

  If $\left|\overline{H_2}\right| < 0, \left|\overline{H_3}\right| < 0, \dots, \left|\overline{H_n}\right| < 0$ at (x^*, λ^*), then x^* is a local minimizer of f subject to $g(x) = c$.

  Example: $f(x_1, x_2) = x_1 x_2$ and $g(x_1, x_2) = x_1 + x_2$

Optimization with Equality Constraints: Global Maximizer (Minimizer)

- Maximization problem with a single constraint

$$
\max_{x \in X} f(x) \qquad \text{subject to} \qquad g(x) - c = 0 \tag{10}
$$

- Minimization problem with a single constraint

$$
\min_{x \in X} f(x) \qquad \text{subject to} \qquad g(x) - c = 0 \tag{11}
$$

- Suppose that there exists λ^* such that x^* is a stationary point of $L = f(x) + \lambda^*[c - g(x)]$ and $g(x^*) = c$.

  1. x^* solves problem (10) if L is concave (f is concave and $\lambda^* g$ is convex)
  2. x^* solves problem (11) if L is convex (f is convex and $\lambda^* g$ is concave)

- One useful result: If g is linear in x, then $\lambda^* g$ is both convex and concave. Therefore, if f is concave(convex), any stationary point $x^* \in \mathsf{Int}X$ of L solves problem(10) (problem(11)).

Optimization with Equality Constraints: Example of Cost Minimization

- Firm's production function $Q = Q(x_1, x_2)$ with $Q_1 > 0$ and $Q_2 > 0$
- Cost of $(x_1, x_2) : P_1 x_1 + P_2 x_2$, where P_1 and P_2 are input prices
- Cost minimization problem

$$\min_{x_1, x_2} P_1 x_1 + P_2 x_2 \text{ subject to } Q_0 = Q(x_1, x_2)$$

- Lagrangian function

$$L = P_1 x_1 + P_2 x_2 + \lambda[Q_0 - Q(x_1, x_2)]$$

- First-order conditions are

$$L_1 = P_1 - \lambda Q_1(x_1, x_2) = 0 \quad (12)$$

$$L_2 = P_2 - \lambda Q_2(x_1, x_2) = 0 \quad (13)$$

$$L_\lambda = Q_0 - Q(x_1, x_2) = 0 \quad (14)$$

- From (12) and (13)

$$\frac{P_1}{Q_1} = \frac{P_2}{Q_2} = \lambda$$

  Alternatively, they induce

$$\frac{P_1}{P_2} = \frac{Q_1}{Q_2},$$

  where P_1/P_2 is the negative of the slope of isocosts and Q_1/Q_2 is the absolute value of the slope of an isoquant (i.e., marginal rate of technical substitution of x_1 for x_2).

- Bordered Hessian

$$\left|\overline{H}\right| = \begin{vmatrix} 0 & Q_1(x_1^*, x_2^*) & Q_2(x_1^*, x_2^*) \\ Q_1(x_1^*, x_2^*) & -\lambda Q_{11}(x_1^*, x_2^*) & -\lambda Q_{12}(x_1^*, x_2^*) \\ Q_2(x_1^*, x_2^*) & -\lambda Q_{21}(x_1^*, x_2^*) & -\lambda Q_{22}(x_1^*, x_2^*) \end{vmatrix}$$

  If $\left|\overline{H}\right| < 0$, then (x_1^*, x_2^*) is a **local minimizer**.

- If there exists λ^* such that $P_1x_1 + P_2x_2 + \lambda^*[Q_0 - Q(x_1, x_2)]$ is a convex function and $Q_0 = Q(x_1^*, x_2^*)$, then (x_1^*, x_2^*) is a global minimizer
- Since $P_1x_1 + P_2x_2$ is linear, it is a convex function. Also, note that $\lambda^* = \frac{P_1}{Q_1} = \frac{P_2}{Q_2} > 0$. Therefore, if $Q(x_1, x_2)$ is a concave function, then (x_1^*, x_2^*) is a global minimizer.

Quasiconcavity/Quasiconvexity of a Function

- Consider a function $g : X \to \mathbb{R}$ with $X \subset \mathbb{R}^n$. The upper level set of g for any $a \in \mathbb{R}$ is defined as

$$P(a) = \{x \in X; g(x) \geq a\}$$

- Definition: A function $g : X \to \mathbb{R}$ with a convex set $X \subset \mathbb{R}^n$ is quasiconcave if $P(a)$ is convex for every a.
- The lower level set of g for any $a \in \mathbb{R}$ is defined as

$$L(a) = \{x \in X; g(x) \leq a\}$$

- Definition: A function $g : X \to \mathbb{R}$ with a convex set $X \subset \mathbb{R}^n$ is quasiconvex if $L(a)$ is convex for every a.
- Alternative Definitions
  - A function $g : X \to \mathbb{R}$ with a convex set $X \subset \mathbb{R}^n$ is quasiconcave if, for all $x, x' \in X$ and all $\lambda \in [0, 1]$

$$g\left(\lambda x + (1-\lambda)x'\right) \geq \min[g(x), g(x')]$$

  - A function $g : X \to \mathbb{R}$ with a convex set $X \subset \mathbb{R}^n$ is quasiconvex if, for all $x, x' \in X$ and all $\lambda \in [0, 1]$

$$g\left(\lambda x + (1-\lambda)x'\right) \leq \max[g(x), g(x')]$$

  - A function $g : X \to \mathbb{R}$ with a convex set $X \subset \mathbb{R}^n$ is strictly quasiconcave if, for all $x, x' \in X$ $(x \neq x')$ and all $\lambda \in (0, 1)$

$$g\left(\lambda x + (1-\lambda)x'\right) > \min[g(x), g(x')]$$

– A function $g : X \to \mathbb{R}$ with a convex set $X \subset \mathbb{R}^n$ is strictly quasi-convex if, for all $x, x' \in X$ $(x \neq x')$ and all $\lambda \in (0, 1)$

$$g\left(\lambda x + (1-\lambda)x'\right) < \max[g(x), g(x')]$$

- Useful properties

  A concave function is quasiconcave.

  A convex function is quasiconvex.

- Consider a twice differentiable function $g : X \to \mathbb{R}$ with a convex set $X \subset \mathbb{R}^n$.For $k = 1, \ldots n$,

$$B_k = \begin{bmatrix} 0 & g_1 & g_2 & \cdots & g_k \\ g_1 & g_{11} & g_{12} & \cdots & g_{1k} \\ g_2 & g_{21} & g_{22} & \cdots & g_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ g_k & g_{k1} & g_{k2} & \cdots & f_{kk} \end{bmatrix}$$

- If $|B_1| < 0, |B_2| < 0, \ldots, |B_n| < 0$, then f is strictly quasiconvex
- If $|B_1| < 0, |B_2| > 0, \ldots, (-1)^n |B_n| > 0$, then f is strictly quasiconcave
- If f is quasiconvex, then $|B_1| \leq 0, |B_2| \leq 0, \ldots, |B_n| \leq 0$
- If f is quasiconcave, then $|B_1| \leq 0, |B_2| \geq 0, \ldots, (-1)^n |B_n| \geq 0$

Optimization with Non-negativity Restrictions

- Consider a function $f : X \to \mathbb{R}$ with $X \subset \mathbb{R}$.
- Maximization problem with non-negativity restriction

$$\max_x f(x) \text{ subject to } x \geq 0. \tag{15}$$

- If a solution x^* for problem (15) exists, then

$$\text{(i) } f'(x^*) \leq 0, \quad \text{(ii) } x^* \geq 0, \quad \text{(iii) } x^* \times f'(x^*) = 0.$$

  The last condition means that at least one of x^* and $f'(x^*)$ must be zero (complementary slackness between x and $f'(x)$).

- In general, maximization problem with non-negativity restrictions for a multi-variable function $f : X \to \mathbb{R}$ with $X \subset \mathbb{R}^n$

$$\max_x f(x) \text{ subject to } x_i \geq 0 \text{ forall } i = 1, \ldots, n, \tag{16}$$

where $x = [x_1, \ldots, x_n]$. First order conditions are

(i) $f_i(x^*) \leq 0$, (ii) $x_i^* \geq 0$, (iii) $x_i^* \times f_i(x^*) = 0$ for all $i = 1, \ldots, n$.

for all $i = 1, \ldots, n$

Optimization with Inequality Constraints and Non-negativity Restrictions

- Consider an example with two inequality constraints and three choice variables (i.e., $x = [x_1, x_2, x_3]$)

$$\max_x f(x) \text{ subject to } \begin{array}{c} g^1(x) \leq c^1 \\ g^2(x) \leq c^2 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{array} \tag{17}$$

- Set up the Lagrangian function for problem (17) as

$$L = f(x) + \lambda_1[c^1 - g^1(x)] + \lambda_2[c^2 - g^2(x)] \tag{18}$$

- Problem (17) can be transformed to

$$\max_x f(x) \text{ subject to } \begin{array}{c} g^1(x) + s^1 = c^1 \\ g^2(x) + s^2 = c^2 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, s^1 \geq 0, s^2 \geq 0 \end{array} \tag{19}$$

- Set up the Lagrangian function for problem (18) as

$$\bar{L} = f(x) + \lambda_1[c^1 - s^1 - g^1(x)] + \lambda_2[c^2 - s^2 - g^2(x)] \tag{20}$$

- First order conditions for a solution for (20) is

$$\frac{\partial \bar{L}}{\partial x_i} \leq 0, \quad x_i^* \geq 0, \quad x_i^* \times \frac{\partial \bar{L}}{\partial x_i} = 0 \quad \text{for all } i = 1, 2, 3 \tag{21}$$

$$\frac{\partial \bar{L}}{\partial s^j} \leq 0, \quad s^j \geq 0, \quad s^j \times \frac{\partial \bar{L}}{\partial s^j} = 0 \quad \text{for all } j = 1, 2 \tag{22}$$

$$\frac{\partial \bar{L}}{\partial \lambda_j} = c^j - s^j - g^j(x^*) = 0 \quad \text{for all } j = 1, 2 \tag{23}$$

- From (22) and (23), we have

$$\begin{aligned}\frac{\partial \bar{L}}{\partial s^j} &= -\lambda_j^* \leq 0 \Leftrightarrow \lambda_j^* \geq 0 \\ s^j &= c^j - g^j(x^*) \geq 0\end{aligned}$$

Hence the first order conditions, (21) to (23), can be rewritten as

$$\frac{\partial \bar{L}}{\partial x_i} \leq 0, \quad x_i^* \geq 0, \quad x_i^* \times \frac{\partial \bar{L}}{\partial x_i} = 0 \quad \text{for all } i = 1, 2, 3 \tag{24}$$

$$\lambda_j^* \geq 0, \; c^j - g^j(x^*) \geq 0, \; \lambda_j^* \times \left[c^j - g^j(x^*)\right] = 0 \text{ for all } j = 1, 2 \tag{25}$$

- (24) and (25) can be equivalently expressed as the following first-order conditions for the Lagrangian function in (18)

$$\begin{aligned}\frac{\partial L}{\partial x_i} &\leq 0, \quad x_i^* \geq 0, \quad x_i^* \times \frac{\partial L}{\partial x_i} = 0 \quad \text{for all } i = 1, 2, 3 \\ \frac{\partial L}{\partial \lambda_j} &\geq 0, \quad \lambda_j^* \geq 0, \quad \lambda_j^* \times \frac{\partial L}{\partial \lambda_j} = 0 \quad \text{for all } j = 1, 2\end{aligned}$$

- Generally, consider the following maximization problem

$$\max_x f(x) \text{ subject to } \begin{array}{c} g^j(x) \leq c^j \text{ for all } j = 1, \ldots, m \\ x_i \geq 0 \text{ for all } i = 1, \ldots, n \end{array} \tag{26}$$

The Lagrangian function for problem (26) is

$$L = f(x) + \sum_{j=1}^{m} \lambda_j [c^j - g^j(x)]$$

and the Kuhn-Tucker Conditions are

$$\begin{aligned}\frac{\partial L}{\partial x_i} &\leq 0, \qquad x_i^* \geq 0, \qquad x_i^* \times \frac{\partial L}{\partial x_i} = 0 \qquad \text{for all } i = 1, \ldots, n \\ \frac{\partial L}{\partial \lambda_j} &\geq 0, \qquad \lambda_j^* \geq 0, \qquad \lambda_j^* \times \frac{\partial L}{\partial \lambda_j} = 0 \qquad \text{for all } j = 1, \ldots, m\end{aligned}$$

- **Arrow-Enthoven Sufficiency Theorem: Quasiconcave Programming**

If the following conditions are satisfied:

(a) x^* satisfies the Kuhn-Tucker conditions

(b) each g^j is differentiable and quasiconvex in $\mathbb{R}^n_+$

(c) f is differentiable and it is [concave] or [quasiconcave in $\mathbb{R}^n_+$ and the n derivatives $f_i(x^*)$ are not all zero and f is twice differentiable in the neighborhood of x^*] or [quasiconcave in $\mathbb{R}^n_+$ and $f_i(x^*) < 0$ for at least one x_i] or [quasiconcave in $\mathbb{R}^n_+$ and $f_i(x^*) > 0$ for some x_j that can take on a positive value without violating the constraints],

then x^* is a solution to the maximization problem.

- Generally, consider the following maximization problem

$$\min_x f(x) \text{subject to } \begin{array}{c} g^j(x) \geq c^j \text{ for all } j = 1, \ldots, m \\ x_i \geq 0 \text{ for all } i = 1, \ldots, n \end{array} \tag{27}$$

The Lagrangian function for problem (26) is

$$L = f(x) + \sum_{j=1}^{m} \lambda_j [c^j - g^j(x)]$$

and the Kuhn-Tucker Conditions are

$$\begin{array}{lll} \frac{\partial L}{\partial x_i} & \geq & 0, \quad x_i^* \geq 0, \quad x_i^* \times \frac{\partial L}{\partial x_i} = 0 \quad \text{for all } i = 1, \ldots, n \\ \frac{\partial L}{\partial \lambda_j} & \leq & 0, \quad \lambda_j^* \geq 0, \quad \lambda_j^* \times \frac{\partial L}{\partial \lambda_j} = 0 \quad \text{for all } j = 1, \ldots, m \end{array}$$

- **Arrow-Enthoven Sufficiency Theorem: Quasiconvex Programming**

If the following conditions are satisfied:

(a) x^* satisfies the Kuhn-Tucker conditions

(b) each g^j is differentiable and quasiconcave in $\mathbb{R}^n_+$

(c) f is differentiable and it is [convex] or [quasiconvex in $\mathbb{R}^n_+$ and the n derivatives $f_i(x^*)$ are not all zero and f is twice differentiable in the neighborhood of x^*] or [quasiconvex in $\mathbb{R}^n_+$ and $f_i(x^*) > 0$ for at least one x_i] or [quasiconvex in $\mathbb{R}^n_+$ and $f_i(x^*) < 0$ for some x_j that can take on a positive value without violating the constraints],

then x^* is a solution to the minimization problem.

Optimization with Inequality Constraints and Non-negativity Constraints: Examples

- Example 1: A decision maker wants to find out (x_1, x_2) that maximizes $x_1 + x_2^{1/2}$ subject to (i) $x_1 + bx_2 \leq 1$, (ii) $x_1 \geq 0$, (iii) $x_2 \geq 0$. Assume that $b > 0$.

  The Lagrangean function is

  $$L = x_1 + x_2^{1/2} + \lambda[1 - x_1 - bx_2]$$

  Kuhn-Tucker conditions are

  $$\begin{aligned} L_1 &= 1 - \lambda \leq 0, \quad x_1 \geq 0, \quad L_1 x_1 = 0 \\ L_2 &= \frac{1}{2} x_2^{-1/2} - \lambda b \leq 0, \quad x_2 \geq 0, \quad L_2 x_2 = 0 \\ L_\lambda &= 1 - x_1 - bx_2 \geq 0, \quad \lambda \geq 0, \quad L_\lambda \lambda = 0 \end{aligned}$$

  Derive the solution to the problem.

  Case 1) $x_1 = 0$ and $x_2 = 0$. Then $L_\lambda = 1 > 0$ and $L_2 = -\lambda b \leq 0$. Therefore, $\lambda \geq 1/b$. So, this violates the condition $L_\lambda \lambda = 0$. This means that $x_1 = 0$ and $x_2 = 0$ is not part of a solution.

  Case 2) $x_1 > 0$ and $x_2 = 0$. Then, $L_1 = 1 - \lambda = 0$. Therefore, $\lambda = 1$. Note that as $x_2 \to 0$, $L_2 \to \infty$. Therefore, this is not part of a solution.

  Case 3) $x_1 = 0$ and $x_2 > 0$. Then, $L_1 = 1 - \lambda \leq 0$. So, $\lambda \geq 1$. Then, $L_\lambda = 1 - bx_2 = 0$. $x_2 = 1/b \geq 0$. Then, $L_2 = \frac{1}{2} b^{1/2} - \lambda b = 0$. $\lambda = \frac{1}{2} b^{-1/2} \geq 1$. So, $0 < b \leq 1/4$.

  Case 4) $x_1 > 0$ and $x_2 > 0$. Then, $\lambda = 1$ from $L_1 = 1 - \lambda = 0$. We have $x_2 = 1/4b^2$ from $L_2 = \frac{1}{2} x_2^{-1/2} - \lambda b = 0$. $x_1 = (4b - 1)/4b > 0$. Therefore, $b > 1/4$.

  From all the possible cases, we have

  $$(x_1^*, x_2^*) = \begin{cases} \left(0, \frac{1}{b}\right) & \text{if } 0 < b \leq 1/4 \\ \left(\frac{4b-1}{4b}, \frac{1}{4b^2}\right) & \text{if } b > 1/4 \end{cases}$$

  Note that the objective function is a concave function and the function appeared in the constraint is a convex function. Furthermore, at least one of the first-order derivatives of the objective function evaluated at (x_1^*, x_2^*) is not zero. Therefore, (x_1^*, x_2^*) is the solution to the problem.

- Example 2: War time rationing.

  A consumer's utility function is $U[x_1, x_2] = x_1 x_2^2$ in a two-good economy with $p_1 = 1$, $p_2 = 1$ and $I = 100$. Rationing constraint is $2x_1 + x_2 \leq 120$. The consumer's maximization problem is

$$\max U[x_1, x_2] \text{ subject to } \begin{array}{c} x_1 + x_2 \leq 100 \\ 2x_1 + x_2 \leq 120 \\ x_1 \geq 0, x_2 \geq 0 \end{array}$$

  The Lagrangian function is then

$$L = x_1 x_2^2 + \lambda_1[100 - x_1 - x_2] + \lambda_2[120 - 2x_1 - x_2]$$

  The Kuhn-Tucker conditions are

$$\begin{aligned} L_1 &= x_2^2 - \lambda_1 - 2\lambda_2 \leq 0, \quad x_1 \geq 0, \quad x_1 \times L_1 = 0 \\ L_2 &= 2x_1 x_2 - \lambda_1 - \lambda_2 \leq 0, \quad x_2 \geq 0, \quad x_2 \times L_2 = 0 \\ L_{\lambda_1} &= 100 - x_1 - x_2 \geq 0, \quad \lambda_1 \geq 0, \quad \lambda_1 \times L_{\lambda_1} = 0 \\ L_{\lambda_2} &= 120 - 2x_1 - x_2 \geq 0, \quad \lambda_2 \geq 0, \quad \lambda_2 \times L_{\lambda_2} = 0 \end{aligned}$$