Classical Optimization Theory:

Constrained Optimization
(Equality Constraints)

1. Preliminaries

- **Framework:**
  - Consider a set $A \subset \Re^n$.
  - Consider the functions $f : A \to \Re$ and $g : A \to \Re$.
  - **Constraint set:**

$$C = \{x \in A : g(x) = 0\}.$$

  - A typical **optimization problem:**

$$\left.\begin{array}{l} \text{Maximize } f(x), \\ \text{subject to } x \in C. \end{array}\right\} \qquad \text{(P)}$$

- **Local Maximum:**

  A point $x^* \in C$ is said to be a *point of local maximum of* f *subject to the constraints* $g(x) = 0$, if there exists an open ball around x^*, $B_\epsilon(x^*)$, such that $f(x^*) \geq f(x)$ for all $x \in B_\epsilon(x^*) \cap C$.

- **Global Maximum:**

  A point $x^* \in C$ is *a point of global maximum of* f *subject to the constraint* $g(x) = 0$, if x^* solves the problem (P).

- **Local minimum** and **global minimum** can be defined similarly by just reverting the inequalities.

2. Necessary Conditions for Constrained Local Maximum and Minimum

- The basic necessary condition for a constrained local maximum is provided by Lagrange's theorem.

- **Theorem 1(a) (Lagrange Theorem: Single Equality Constraint):**

  Let $A \subset \Re^n$ *be open, and* $f : A \to \Re$, $g : A \to \Re$ *be continuously differentiable functions on* A. *Suppose* x^* *is a point of local maximum or minimum of* f *subject to the constraint* $g(x) = 0$. *Suppose further that* $\nabla g(x^*) \neq 0$. *Then there is* $\lambda^* \in \Re$ *such that*

$$\nabla f(x^*) = \lambda^* \cdot \nabla g(x^*). \tag{1}$$

  – The $(n+1)$ equations given by (1) and the constraint $g(x) = 0$ are called the *first-order conditions* for a constrained local maximum or minimum.

- There is an easy way to remember the conclusion of Lagrange Theorem.
  - Consider the function $L : A \times \Re \rightarrow \Re$ defined by

$$L(x, \lambda) = f(x) - \lambda g(x).$$

  - L is known as the *Lagrangian*, and λ as the *Lagrange multiplier*.
  - Consider now the problem of finding the local maximum (or minimum) in an unconstrained maximization (or minimization) problem in which L is the function to be maximized (or minimized).
  - The first-order conditions are:

$$D_i L(x, \lambda) = 0, \text{ for } i = 1, ..., n+1,$$

which yields

$$D_i f(x) = \lambda D_i g(x), \text{ for } i = 1, ..., n; \text{ and } g(x) = 0.$$

  - The first n equations can be written as $\nabla f(x) = \lambda \cdot \nabla g(x)$ as in the conclusion of Lagrange Theorem.
  - The method described above is known as the "Lagrange multiplier method".

- **The Constraint Qualification:**

  The condition, $\nabla g(x^*) \neq 0$, is known as the *constraint qualification*.

  – It is particularly important to check the constraint qualification before applying the conclusion of Lagrange's Theorem.

  - Without this condition, the conclusion of Lagrange's Theorem would not be valid, as the following example shows.

#1. Let $f : \Re^2 \to \Re$ be given by $f(x_1, x_2) = 2x_1 + 3x_2$ for all $(x_1, x_2) \in \Re^2$, and $g : \Re^2 \to \Re$ be given by $g(x_1, x_2) = x_1^2 + x_2^2$ for all $(x_1, x_2) \in \Re^2$.

– Consider the constraint set

$$C = \left\{ (x_1, x_2) \in \Re^2 : g(x_1, x_2) = 0 \right\}.$$

– Now consider the maximization problem: $\underset{(x_1,x_2)\in C}{\text{Maximize}} \; f(x_1, x_2)$.

(a) Demonstrate that the conclusion of Lagrange's Theorem does not hold here.

(b) What is the solution to the maximization problem?

(c) What goes wrong? Explain clearly.

- **Several Equality Constraints:**

  In Problem (P) we have considered only one equality constraint, $g(x) = 0$. Now we consider more than one, say m, equality constraints: $g^j : A \to \Re$, such that $g^j(x) = 0$, $j = 1, 2, ..., m$.

  – **Constraint set:**

$$C = \left\{ x \in A : g^j(x) = 0,\ j = 1, 2, ..., m \right\}.$$

  – **Constraint Qualification:**

  The natural generalization of the constraint qualification with single constraint, $\nabla g(x^*) \neq 0$, involves the Jacobian derivative of the constraint functions:

$$Dg\left(x^*\right) = \begin{pmatrix} \dfrac{\partial g^1}{\partial x_1}(x^*) & \dfrac{\partial g^1}{\partial x_2}(x^*) & \cdots & \dfrac{\partial g^1}{\partial x_n}(x^*) \\ \dfrac{\partial g^2}{\partial x_1}(x^*) & \dfrac{\partial g^2}{\partial x_2}(x^*) & \cdots & \dfrac{\partial g^2}{\partial x_n}(x^*) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g^m}{\partial x_1}(x^*) & \dfrac{\partial g^m}{\partial x_2}(x^*) & \cdots & \dfrac{\partial g^m}{\partial x_n}(x^*) \end{pmatrix}.$$

– In general, a point x^* is called a *critical point* of $g = \left(g^1, g^2, ..., g^m\right)$, if the rank of the Jacobian matrix, $Dg\left(x^*\right)$, is $< m$.

– So the natural generalization of the constraint qualification is: *rank* $\left(Dg\left(x^*\right)\right) = m$.

- This version of the constraint qualification is called the **nondegenerate constraint qualification (NDCQ)**.

– The NDCQ is a regularity condition. It implies that the constraint set has a well-defined $(n - m)$-dimensional tangent plane everywhere.

• **Theorem 1(b) (Lagrange Theorem: Several Equality Constraints):**

Let $A \subset \Re^n$ *be open*, *and* $f : A \rightarrow \Re$, $g^j : A \rightarrow \Re$ *be continuously differentiable functions on* A, $j = 1, 2, ..., m$. *Suppose* x^* *is a point of local maximum or minimum of* f *subject to the constraints* $g^j(x) = 0$, $j = 1, 2, ..., m$. *Suppose further that rank* $\left(Dg\left(x^*\right)\right) = m$. *Then there exist* $\left(\lambda_1^*, \lambda_2^*, ..., \lambda_m^*\right) \in \Re^m$ *such that* $\left(x^*, \lambda^*\right) \equiv \left(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, \lambda_2^*, ..., \lambda_m^*\right)$ *is a critical point of the Lagrangian*

$$L\left(x, \lambda\right) \equiv f\left(x\right) - \lambda_1 g^1\left(x\right) - \lambda_2 g^2\left(x\right) - ... - \lambda_m g^m\left(x\right).$$

In other words,

$$\begin{aligned} \frac{\partial L}{\partial x_i}(x^*, \lambda^*) &= 0,\ i = 1, 2, ..., n, \\ &\textit{and} \\ \frac{\partial L}{\partial \lambda_j}(x^*, \lambda^*) &= 0,\ j = 1, 2, ..., m. \end{aligned} \tag{2}$$

– Proof: To be discussed in class (Section 19.6 of textbook).

– The $(n + 1)$ equations given by (2) are called the *first-order conditions* for a constrained local maximum or minimum.

- The following theorem provides a necessary condition involving the second-order partial derivatives of the relevant functions (called “second-order necessary conditions”).

- **Theorem 2:**

  Let $A \subset \Re^n$ *be open, and* $f : A \to \Re$, $g : A \to \Re$ *be twice continuously differentiable functions on* A. *Suppose* x^* *is a point of local maximum of* f *subject to the constraint* $g(x) = 0$. *Suppose further that* $\nabla g(x^*) \neq 0$. *Then there is* $\lambda^* \in \Re$ *such that*

  (i) First-Order Condition: $\nabla f(x^*) = \lambda^* \cdot \nabla g(x^*)$,

  (ii) Second-Order Necessary Condition: $y^T \cdot H_L(x^*, \lambda^*) \cdot y \leq 0$, *for all* y *satisfying* $y \cdot \nabla g(x^*) = 0$

  [*where* $L(x, \lambda^*) = f(x) - \lambda^* g(x)$ *for all* $x \in A$, *and* $H_L(x^*, \lambda^*)$ *is the* $n \times n$ *Hessian matrix of* $L(x, \lambda^*)$ *with respect to* x *evaluated at* (x^*, λ^*)].

  – The second-order *necessary* condition for maximization requires that the Hessian is *negative semi-definite* on the linear constraint set $\{y : y \cdot \nabla g(x^*) = 0\}$.

- **Second-Order Necessary Condition for Minimization:** $y^T \cdot H_L(x^*, \lambda^*) \cdot y \geq 0$ (that is, the Hessian is *positive semi-definite*), for all y satisfying $y \cdot \nabla g(x^*) = 0$.

- For m constraints, $g^j(x) = 0$, $j = 1, 2, ..., m$, $\nabla g(x^*) \neq 0$ is replaced by the NDCQ: *rank* $(Dg(x^*)) = m$.

3. Sufficient Conditions for Constrained Local Maximum and Minimum

- **Theorem 3(a) [Single Equality Constraint]:**

  Let $A \subset \Re^n$ *be open, and* $f : A \rightarrow \Re$, $g : A \rightarrow \Re$ *be twice continuously differentiable functions on* A. *Suppose* $(x^*, \lambda^*) \in C \times \Re$ *and*

  (i) First-Order Condition: $\nabla f(x^*) = \lambda^* \cdot \nabla g(x^*)$,

  (ii) Second-Order Sufficient Condition: $y^T \cdot H_L(x^*, \lambda^*) \cdot y < 0$, *for all* $y \neq 0$ *satisfying* $y \cdot \nabla g(x^*) = 0$

  [*where* $L(x, \lambda^*) = f(x) - \lambda^* g(x)$ *for all* $x \in A$, *and* $H_L(x^*, \lambda^*)$ *is the* $n \times n$ *Hessian matrix of* $L(x, \lambda^*)$ *with respect to* x *evaluated at* (x^*, λ^*)]. *Then* x^* *is a point of local maximum of* f *subject to the constraint* $g(x) = 0$.

  – The second-order *sufficient* condition for maximization requires that the Hessian is *negative definite* on the linear constraint set $\{y : y \cdot \nabla g(x^*) = 0\}$.

- **Second-Order Sufficient Condition for Minimization:** $y^T \cdot H_L(x^*, \lambda^*) \cdot y > 0$ (that is, the Hessian is *positive definite*), *for all* $y \neq 0$ *satisfying* $y \cdot \nabla g(x^*) = 0$.

- There is a convenient method of checking the *second-order sufficient condition* stated in Theorem 3(a), by checking the signs of the leading principal minors of the relevant “bordered” matrix. This method is stated in the following Proposition.

- **Proposition 1(a):**

  Let A be an $n \times n$ symmetric matrix and b be an n-vector with $b_1 \neq 0$. Define the $(n+1) \times (n+1)$ matrix S by

$$S = \begin{pmatrix} 0 & b \\ b & A \end{pmatrix}.$$

(a) *If $|S|$ has the same sign as $(-1)^n$ and if the last $(n-1)$ leading principal minors of S alternate in sign, then $y^T Ay < 0$ for all $y \neq 0$ such that $yb = 0$;*

(b) *If $|S|$ and the last $(n-1)$ leading principal minors all have the same negative sign*, *then $y^T Ay > 0$ for all $y \neq 0$ such that $yb = 0$.*

- **Theorem 3(b) [Several Equality Constraints]:**

  Let $A \subset \Re^n$ *be open, and* $f : A \to \Re$, $g^j : A \to \Re$ *be twice continuously differentiable functions on* A, $j = 1, 2, ..., m$. *Suppose* $(x^*, \lambda^*) \in C \times \Re^m$ *and*

  (i) First-Order Condition: $\nabla f(x^*) = \lambda^* \cdot \nabla g(x^*)$,

  (ii) Second-Order Sufficient Condition: $y^T \cdot H_L(x^*, \lambda^*) \cdot y < 0$, *for all* $y \neq 0$ *satisfying* $y \cdot \nabla g(x^*) = 0$

  [*where* $L(x, \lambda^*) = f(x) - \lambda_1^* g^1(x) - \lambda_2^* g^2(x) - ... - \lambda_m^* g^m(x)$, *for all* $x \in A$, *and* $H_L(x^*, \lambda^*)$ *is the* $n \times n$ *Hessian matrix of* $L(x, \lambda^*)$ *with respect to* x *evaluated at* (x^*, λ^*)]. *Then* x^* *is a point of local maximum of* f *subject to the constraints* $g^j(x) = 0$, $j = 1, 2, ..., m$.

  – The second-order *sufficient* condition for maximization requires that the Hessian is *negative definite* on the linear constraint set $\{y : y \cdot \nabla g(x^*) = 0\}$.

  – **Second-Order Sufficient Condition for Minimization:** $y^T \cdot H_L(x^*, \lambda^*) \cdot y > 0$ (that is, the Hessian is *positive definite*), *for all* $y \neq 0$ *satisfying* $y \cdot \nabla g(x^*) = 0$.

– Proof: The proof for the case of ‘two variables and one constraint’ will be discussed in class (see Theorem 19.7, pages 461 – 462, of the textbook).

  - For the proof of the general case, refer to Section 30.5 (Constrained Maximization), pages 841 – 844, of the textbook.

- There is a convenient method of checking the *second-order sufficient condition* stated in Theorem 3(b), by checking the signs of the leading principal minors of the relevant “bordered” matrix. This method is stated in the following Proposition.

- **Proposition 1(b):**

  To determine the definiteness of a quadratic form of n variables, $Q(x) = x^T Ax$, *when restricted to a constraint set given by m linear equations $Bx = 0$, construct the $(n+m) \times (n+m)$ matrix S by bordering the matrix A above and to the left by the coefficients B of the linear constraints*:

$$S = \begin{pmatrix} 0 & B \\ B^T & A \end{pmatrix}.$$

  Check the signs of the last $(n-m)$ leading principal minors of S, starting with the determinant of S itself.

(a) I*f $|S|$ has the same sign as $(-1)^n$ and if these last $(n-m)$ leading principal minors alternate in sign*, *then Q is negative definite on the constraint set $Bx = 0$.*

(b) *If $|S|$ and these last $(n-m)$ leading principal minors all have the same sign as $(-1)^m$*, *then Q is positive definite on the constraint set $Bx = 0$.*

- For discussions on Propositions 1(a) and 1(b) refer to Section 16.3 (Linear Constraints and Bordered Matrices) (pages 386-393) of the textbook.

4. Sufficient Conditions for Constrained Global Maximum and Minimum

- **Theorem 4:**

  Let $A \subset \Re^n$ *be an open convex set, and* $f : A \to \Re$, $g^j : A \to \Re$ *be continuously differentiable functions on* A, $j = 1, 2, ..., m$. *Suppose* $(x^*, \lambda^*) \in C \times \Re^m$ *and* $\nabla f(x^*) = \lambda^* \cdot \nabla g(x^*)$. *If* $L(x, \lambda^*) = f(x) - \lambda_1^* g^1(x) - \lambda_2^* g^2(x) - ... - \lambda_m^* g^m(x)$ *is concave* (*respectively*, *convex*) *in* x *on* A, *then* x^* *is a point of global maximum* (*respectively*, *minimum*) *of* f *subject to the constraints* $g^j(x) = 0$, $j = 1, 2, ..., m$.

  – Proof: To be discussed in class.

5. How to Solve Optimization Problems

- Two Routes:
  - **Route 1 (Sufficiency Route):** Use the *sufficient conditions*.
    - These will involve the concavity (convexity) and/or the second-order conditions.
  - **Route 2 (Necessary Route):** Use the *necessary conditions* PLUS the *Weierstrass Theorem*.
    - This route is useful when there is not enough information about the second-order conditions (bordered Hessian) or concavity/quasiconcavity.

#2. **An Example of the Sufficiency Route:**

Let $f : \Re^2 \rightarrow \Re$ be given by $f(x,y) = \left(1 - x^2 - y^2\right)$ and $g : \Re^2 \rightarrow \Re$ be given by $g(x,y) = x + 4y - 2.$

(a) Set up the Lagrangian and find out the values of (x^*, y^*, λ^*) satisfying the first-order conditions and the constraint $g(x,y) = 0$.

(b) Set up the appropriate bordered Hessian matrix and check whether (x^*, y^*) is a point of *local* maximum of f subject to the constraint $g(x,y) = 0.$

(c) Check whether is also a point of *global* maximum of f subject to the constraint $g(x,y) = 0.$

- **More Examples of the Sufficiency Route:** Examples 19.7 and 19.8 of the textbook.

#3. **An Example of the Necessary Route:**

Consider the following constrained maximization problem:

$$\left.\begin{array}{ll}\text{Maximize} & \prod_{i=1}^{n} x_i \\ \text{subject to} & \sum_{i=1}^{n} x_i = n, \\ \text{and} & x_i \geq 0,\ i = 1, 2, ...n.\end{array}\right\} \text{(P)}$$

[Note that we have not yet encountered the inequality constraints of the "≥ 0" type. We will see how to handle them in this specific context.]

(a) Step I: Define $C = \left\{x \in \Re^n_+ : \sum_{i=1}^{n} x_i = n\right\}$. Apply Weierstrass Theorem carefully to show that there exists $x^* \in C$ such that x^* solves (P).

(b) Step II: Convert the problem "suitably" so that Lagrange Theorem is applicable. (Note that Lagrange Theorem is applicable on an *open* set whereas Weierstrass Theorem on a *closed* set.)

- Since x^* solves (P), $x_i^* > 0,\, i = 1, 2, ..., n$. We can therefore conclude that x^* also solves the following problem:

$$\left.\begin{array}{ll} \text{Maximize} & \prod_{i=1}^{n} x_i \\ \text{subject to} & \sum_{i=1}^{n} x_i = n, \\ \text{and} & x_i > 0,\, i = 1, 2, ...n. \end{array}\right\} \text{(Q)}$$

- Define $A = \Re^n_{++}$, so that A is an open subset of $\Re^n$. Define $f : A \to \Re$ by $f(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} x_i$, and $g : A \to \Re$ by $g(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} x_i - n$.

(c) Step III: Apply Lagrange Theorem to find x^*.

- **More Examples of the Necessary Route:**

  Examples 18.4, 18.5 and 18.6 of the textbook.

References

- Must read the following sections from the textbook:
  - Section 16.3 (pages 386 – 393): Linear Constraints and Bordered Matrices,
  - Section 18.1 and 18.2 (pages 411 – 424): Equality Constraints,
  - Section 19.3 (pages 457 – 469): Second-Order Conditions (skip the subsection on Inequality Constraints for the time-being),
  - Section 30.5 (pages 841 – 844): Constrained Maximization.
- This material is based on

1. Bartle, R., *The Elements of Real Analysis*, (chapter 7),
2. Apostol, T., *Mathematical Analysis: A Modern Approach to Advanced Calculus*, (chapter 6).
3. Takayama, A., *Mathematical Economics*, (chapter 1).

- A proof of the important Proposition 1 can be found in

4. Mann, H.B., “Quadratic Forms with Linear Constraints”, *American Mathematical Monthly* (1943), pages 430 – 433, and also in
5. Debreu, G., “Definite and Semidefinite Quadratic Forms”, *Econometrica* (1952), pages 295 – 300.