Math Review

for Microeconomic

April 2, 2008

Joosung Lee

Department of economics, SNU

Contents

Chapter 1

Matrix Algebra

1.1 Terminologies and operations

1.1.1 Basic Matrix Operations

Definition 1.1.1 (Matrix and Vector)**.** An $m \times n$ *matrix* A is a rectangular array of real numbers with m rows and n columns.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

- $m \times n$ is called the dimension or order of A. If $m = n$, the matrix is the square of order n.
- A short hand notation is $A = (a_{ij})$, $i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, n$.
- A vector is a special case of a matrix. A row vector is $1 \times n$ matrix and a column vector is $m \times 1$ matrix.

Check 1.1.1. Basic Matrix Operations

- Equality
- Addition and subtraction
- Scalar multiplication
- Matrix multiplication

Check 1.1.2. Laws of Matrix Operations

- Commutative law of addition
- Associative law of addition
- Associative law of multiplication
- Distributive law

Remark. The commutative law of multiplication is not applicable

Check 1.1.3. Definitions of Some Basic Matrices

- identity (or unit) matrix
- null matrix
- diagonal matrix
- (upper and lower) triangular matrix

1.1.2 Transpose

Definition 1.1.2. We say that $B = (b_{ij})_{n\times m}$ is the *transpose* of $A = (a_{ij})_{m\times n}$ if $a_{ji} = b_{ij}$ for all $i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, n$. We notate B, the transpose of A, as A' or A^T.

Check 1.1.4. Properties of Transpose

- $(A')' = A$
- $(A + B)' = A' + B'$
- $(kA)' = kA'$, where k is a scalar
- $(AB)' = B'A'$

Definition 1.1.3. Special transpose matrices

- If $A' = A$, A is called symmetric.
- If $A'A = I$, A is called orthogonal.
- If $A = A'$ and $AA = A$, A is called idempotent.

1.1.3 Inverse

Definition 1.1.4. An $n \times n$ matrix A is *nonsingular* if there exists an $n \times n$ matrix B such $AB = BA = I$. The Matrix B is called the *inverse* of A and denoted by A^{-1}.

Check 1.1.5. Properties of Inverse

- $(A^{-1})^{-1} = A$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A')^{-1} = (A^{-1})'$
- If $D = diag(d_1, \cdots, d_n)$, then $D^{-1} = diag(d_1^{-1}, \cdots, d_n^{-1})$

Check 1.1.6. Elementary row operation

- *Type 1* interchanging any two rows of A
- *Type 2* multiplying any row of A non-zero scalar

- *Type 3* adding any scalar multiple of a row A to another row

Check 1.1.7 (Gauss elimination method)**.** Algorithm for finding the inverse of a square matrix A:

i. Write the augmented matrix $(A|I)$.

ii. Try to convert A into I by the elementary operations where the rows are from $(A|I)$.

iii. If it is possible to convert A into I then the augmented matrix takes the form $(A|I^{-1})$.

Example 1.1.1. Find the inverses of following matrices, if it exists.

(1) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 5 & 7 \end{pmatrix}$

(2) $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & -1 \\ 1 & 5 & 4 \end{pmatrix}$

Further Study 1.1.1. Partitioned matrix and its inverse

1.1.4 Determinants

Definition 1.1.5. Let A be an $n \times n$ matrix.

i. if $n = 1$, so that $A = (a_{11})$, we define the *determinant* of A to be $|A| = a_{11}$.

ii. if $n \geq 2$, the *determinant* of A is defined to be

$$|A| = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} |A_{1j}|,$$

where A_{ij} is the $(n-1) \times (n-1)$ submatrix obtained by deleting row i and column j from A.

Remark. The determinant can be evaluated by cofactor expansion along *any* row or column.

Check 1.1.8. Properties of determinant 1

- If B is a matrix obtained by interchanging any two rows of A, then $|B| = -|A|$.
- If B is a matrix obtained by adding a multiple of one row of A to another row of A, then $|B| = |A|$.
- If B is a matrix obtained by multiplying a row of A by a non-zero scalar k, then $|B| = k|A|$.
- If A is a triangular matrix, then $|A| = \prod_{i=1}^{n} a_{ii}$

Example 1.1.2. Find the determinants of following matrices, if it exists.

(1) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$

(2) $\begin{pmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{pmatrix}$

Check 1.1.9. Properties of determinant 2

- $|A'| = |A|$.
- $|AB| = |A||B|$.

1.1.5 trace

Definition 1.1.6 (Trace)**.** Let $A = (a_{ij})$ be an $n \times n$ matrix. The *trace* of A, denoted $tr(A)$, is the sum of the diagonal entries of A, that is,

$$tr(A) = \sum_{i}^{n} a_{ii}.$$

Check 1.1.10. Properties of trace

- $tr(A + B) = tr(A) + tr(B)$.
- $tr(A') = tr(A)$.
- $tr(AB) = tr(BA)$,for an $m \times n$ matrix A and an $n \times m$ matrix B.

1.2 Vector Space

1.2.1 vector space and subspace

Definition 1.2.1. A (real) *vector space* is a nonempty set V of objects together with an additive operation $+ : V \times V \to V$,$+(u, v) = u + v$ and a scalar multiplicative operation $\cdot : \mathbb{R} \times V \to V$,$\cdot(a, u) = au$ which satisfies the following axioms for any $u, v, w \in V$ and any $a, b \in \mathbb{R}$:

V1. $(u + v) + w = u + (v + w)$

V2. $u + v = v + u$

V3. $0 + u = u$

V4. $u + (-u) = 0$

V5. $a(u + v) = au + av$

V6. $(a + b)u = au + bu$

V7. $a(bu) = (ab)u$

V8. $1u = u$.

Remark. Some additional definitions

- The objects of a vector space V are called vectors
- The operations $+$ and $\times$ are called vector addition and scalar multiplication, respectively.
- The element $0 \in V$ is the zero vector and $-v$ is the additive inverse of V.
- The n-dimensional vector space $\mathbb{R}^n = \{(u_1, u_2, \cdots, u_n)' | u_i \in \mathbb{R}, i = 1, \cdots, n\}$.

Definition 1.2.2. A subset W of a vector space V is called a *subspace* if W is a vector space.

Example 1.2.1.

- x-axis is a subspace of $\mathbb{R}^2$
- $\{(x, y)|y = x,\ x, y \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^2$
- $\{(x, y)|y = x + 1,\ x, y \in \mathbb{R}\}$ is not a subspace of $\mathbb{R}^2$
- xy-plane is a subspace of $\mathbb{R}^3$

1.2.2 Inner product

Definition 1.2.3. Let V be a vector space. An *inner product* or *scalar product* in V is a function $s : V \times V \to \mathbb{R}$, $s(u, v) = u \cdot v$ which satisfies the following properties:

- $u \cdot v = v \cdot u$
- $u \cdot (v + w) = u \cdot v + u \cdot w$
- $a(u \cdot v) = (au) \cdot v = u \cdot (av)$
- $u \cdot u > 0$ if $u \neq 0$

Remark. Let $u, v \in \mathbb{R}^n$. $u \cdot v = u'v = v'u$

Remark. Let $v \in \mathbb{R}^n$. The norm of v is defined as $||v|| = \sqrt{v \cdot v}$

Remark. Let $u, v \in \mathbb{R}^n$. u and v are orthogonal if $u \cdot v = 0$

Check 1.2.1 (Properties of inner product)**.** Let x,y and z are in $\mathbb{R}^n$ and t be a scalar.

- $x \cdot y = y \cdot x$
- $(x + y) \cdot z = x \cdot z + y \cdot z$
- $z \cdot (x + y) = z \cdot x + z \cdot y$
- $x \cdot (ty) = t(x \cdot y) = (tx) \cdot y$
- $||x|| \geq 0$ for all $x \in \mathbb{R}^n$
- $||x|| = 0 \Rightarrow x = \mathbf{0}$
- $(x + y) \cdot (x + y) = ||x||^2 + 2(x \cdot y) + ||y||^2$

1.2.3 linearly dependency and dimension

Definition 1.2.4. Let S be a non-empty subset of vectors in $\mathbb{R}^n$. The *span* of S is defined by $span(S) = \{\sum_i a_i v_i | v_i \in S \text{ and } a_i \in \mathbb{R}\}$.

Remark. The span of S is a set consisting of all linear combinations of the vectors in S.

Definition 1.2.5. Let $\{v_1, v_2, \cdots, v_k\}$ be a set of vectors in $\mathbb{R}^n$. Consider the equation $a_1v_1 + a_2v_2 + \cdots + a_kv_k = 0$. If the only solution to this equation is $a_1 = a_2 = \cdots = a_k = 0$, then the elements $v_1, v_2, \cdots, v_k$ are *linearly independent*. Otherwise, they are *linearly dependent*.

Definition 1.2.6. A set $B = \{v_1, v_2, \cdots, v_k\}$ of vectors in $\mathbb{R}^n$ is a *basis* of a subspace V of $\mathbb{R}^n$ if $span(B) = V$ and all vectors in S are linearly independent.

Remark. Let V be a subspace of $\mathbb{R}^n$. Then every basis for V contains the same number of vectors

Definition 1.2.7. The unique number of vectors in each basis is called the *dimension* of V and is denoted by $dim(V)$.

Example 1.2.2.

- $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ is a basis of $\mathbb{R}^2$.
- $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ is also a basis of $\mathbb{R}^2$.
- $dim(\mathbb{R}^2) = 2$

1.3 Systems of linear equations

1.3.1 Existence and Uniqueness of solutions

Definition 1.3.1. For a system of n linear equations for n unknowns $Ax = b$, if $b = \mathbf{O}$ then the system is *homogeneous*.

Check 1.3.1. Homogeneous case

- If $|A| \neq 0$ then the system has a unique trivial(zero) solution.
- If $|A| = 0$ then the system has a infinite number of solutions.

Definition 1.3.2. For a system of n linear equations for n unknowns $Ax = b$, if $b \neq \mathbf{O}$ then the system is *non-homogeneous*.

Check 1.3.2. Non-homogeneous case

- If $|A| \neq 0$ then the system has a unique solution.
- If $|A| = 0$ and
  - If $\text{rank}(A) = \text{rank}(\tilde{A})$ then the system has a infinite number of solutions.

– If $\text{rank}(A) \neq \text{rank}(\tilde{A})$ then the system is inconsistent.

Here $\tilde{A}$ is a so-called augmented matrix,

$$\tilde{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{pmatrix}$$

Check 1.3.3. How to solve the system of linear equations, if $b \neq \mathbf{O}$ and $|A| \neq 0$:

- The inverse matrix method
- Gauss method
- Cramer's rule

1.3.2 Cramer's rule

Theorem 1.3.1 (Cramer's rule)**.** *Let A be an $n \times n$ non-singular matrix. Let $x^\star = (x_1^\star, \cdots, x_n^\star)$ be the unique solution of $Ax = b$. Then $x_i^\star = \frac{|A^i|}{|A|}$, where A^i is the matrix A with b replacing the i'th column of A.*

Example 1.3.1. $A = \begin{pmatrix} 5 & 1 & -1 \\ -2 & 5 & -1 \\ -1 & -1 & 7 \end{pmatrix}$, $b = \begin{pmatrix} 9 \\ 3 \\ 17 \end{pmatrix}$. Solve x.

1.4 Eigenvalue, Eigenvector and Diagonalization

1.4.1 eigenvalue and eigenvector

Definition 1.4.1. For an $n \times n$ matrix A, a non-zero vector x is called *eigenvector* of A if there exists a scalar λ such that $Ax = \lambda x$. The scalar λ is called *eigenvalue* corresponding to the eigenvector x.

Check 1.4.1. How to find the eigenvalues and the eigenvectors of a matrix A.

- Find the solutions of $|A - \lambda I| = 0$, then The solutions $\lambda^\star$are the eigenvalues.
- Find the solution vectors of $(A - \lambda^\star I)x = \mathbf{0}$, then the vectors are the eigenvectors corresponding to the eigenvalue $\lambda^\star$.

Remark. $|A - \lambda I|$ is called the *characteristic polynomial* of A.

1.4.2 diagonalization

Definition 1.4.2. An $n \times n$ matric A is *diagonalizable* if there exists a nonsingular P such that $P^{-1}AP = D$ for a diagonal matrix D.

Remark. An $n \times n$ matric A is diagonalizable if and only if it has a set of n linearly independent eigenvectors.

Remark. Let A be an $n \times n$ symmetric matrix. Then A has n orthogonal eigenvectors and it can be diagonalized as $P'AP = D$ form.

Example 1.4.1. Diagonalize following matrices.

1) $A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 2 \\ 1 & 2 & 2 \end{pmatrix}$

2) $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$

3) $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$

4) $A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{pmatrix}$

1.5 Quadratic forms and Definite matrices

1.5.1 quadratic form

Definition 1.5.1. A *quadratic form* on $\mathbb{R}^n$ is a real-valued function of the form

$$Q(x_1, \cdots, x_n) = \sum_{i \leq j} a_{ij} x_i x_j,$$

in which each term is of degree two.

Remark. A quadratic form on $\mathbb{R}^n$ can be represented by an $n \times n$ matrix A and a vector x in $\mathbb{R}^n$ as $x'Ax$.

Check 1.5.1. Check the condition for followings:

1) for all $x \in \mathbb{R}$, $ax^2 + 2bx + c \geq 0$

2) for all $x, y \in \mathbb{R}$, $ax^2 + 2bxy + cy^2 \geq 0$

1.5.2 definite matrices

Definition 1.5.2. A quadratic form $Q(x_1, \cdots, c_n) = \sum_{i \leq j} a_{ij} x_i x_j$ is called

i. *positive definite*, if $Q(x) > 0$ for all $x \neq 0$.

ii. *negative definite*, if $Q(x) < 0$ for all $x \neq 0$.

iii. *positive semidefinite*, if $Q(x) \geq 0$ for all x.

iv. *negative semidefinite*, if $Q(x) \leq 0$ for all x.

v. *indefinite*, otherwise.

Check 1.5.2. $Q = x^2 + y^2$ is PD, $Q = (x+y)^2$ is PSD, $Q = x^2 - y^2$ is ID.

Definition 1.5.3. Let A be an $n \times n$ matrix. A $k \times k$ submatrix of A obtained by deleting $n-k$ columns, say columns $i_1, i_2, \cdots, i_{n-k}$ and the same $n-k$ rows, say rows $i^1, i^2, \cdots, i^{n-k}$, from A is called a kth order *principal submatrix* of A. The determinant of a $k \times k$ principal submatrix is called a kth order *principal minor* of A.

Definition 1.5.4. The kth order *leading principal submatrix* of A, denoted A_k is the matrix obtained by deleting the last $n-k$ rows and columns from A. The kth order *leading principal minor* of A is $|A_k|$.

Theorem 1.5.1. *Let A be an $n \times n$ symmetric matrix. Then*

i. *A is PD iff $|A_1| > 0, |A_2| > 0, \cdots, |A_n| > 0$*

ii. *A is ND iff $|A_1| < 0, |A_2| > 0, \cdots, (-1)^n|A_n| > 0$*

iii. *A is PSD iff (not only leading principal minors but) every principal minors of A is non-negative.*

iv. *A is NSD iff (not only leading principal minors but) every principal minors of odd order A is non-positive and every principal minors of even order A is non-negative.*

v. *A is ID, otherwise.*

Example 1.5.1. Consider the quadratic form $Q(x, y, z) = 3x^2 + 3y^2 + 5z^2 - 2xy$.

Theorem 1.5.2. *Let A be an $n \times n$ symmetric matrix. Then*

i. *A is PD iff all the eigenvalues of A are positive.*

ii. *A is ND iff all the eigenvalues of A are negative.*

iii. *A is PSD iff all the eigenvalues of A are non-negative.*

iv. *A is NSD iff all the eigenvalues of A are non-positive.*

v. *A is ID, otherwise.*

1.5.3 definite matrices on the linear constraint

Check 1.5.3. Consider the quadratic form $Q(x, y) = x^2 - y^2$. See `Figure` 1.1

- Q on $\mathbb{R}^2$ is ID.
- Q on the x-axis is PD.
- Q on the y-axis is ND.
- Q on the line $x - 2y = 0$ is PD.

Theorem 1.5.3. *Let A be an $n \times n$ symmetric matrix and B be an $m \times n$ matrix. Construct the $(n+m) \times (n+m)$ the bodered matrix $\bar{H} = \begin{pmatrix} \mathbf{0} & B \\ B' & A \end{pmatrix}$.*

i. *If $|\bar{H}|$ and these last $n - m$ leading principal minors all have the same sign as $(-1)^m$, then A is PD on the linear constraint $Bx = \mathbf{0}$.*

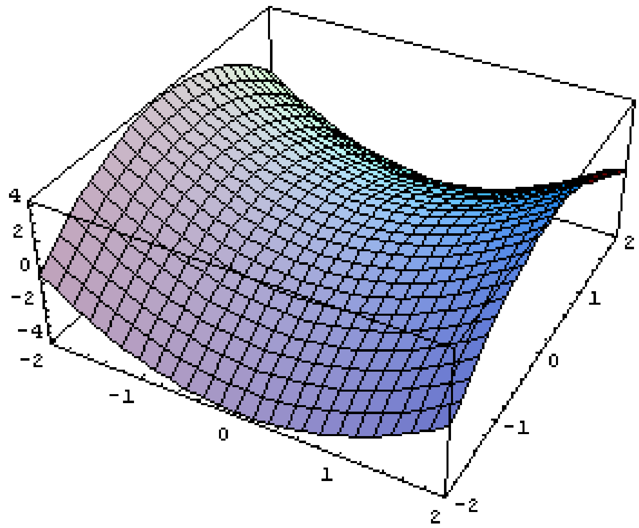


Figure 1.1: $Q(x, y) = x^2 - y^2$

ii. If $|\bar{H}|$ has the same sign as $(-1)^n$ and if these last $n - m$ leading principal minors alternate in sign, then A is ND on the linear constraint $Bx = \mathbf{0}$.

Remark. A is PD on the linear constraint $Bx = \mathbf{0}$ iff $(-1)^m|\bar{H}_{m+n}| > 0, (-1)^m|\bar{H}_{m+n-1}| > 0, \cdots, (-1)^m|\bar{H}_{(m+n)-(n-m-1)}| > 0$.

Remark. A is ND on the linear constraint $Bx = \mathbf{0}$ iff $(-1)^n|\bar{H}_{m+n}| > 0, (-1)^{n-1}|\bar{H}_{m+n-1}| > 0, \cdots, (-1)^{m+1}|\bar{H}_{(m+n)-(n-m-1)}| > 0$.

Example 1.5.2. Chech the definiteness of $Q(x_1, x_2, x_3, x_4) = x_1^2 - x_2^2 + x_3^2 + x_4^2 + 4x_2x_3 - 2x_1x_4$ on the constraint set $x_2 + x_3 + x_4 = 0, x_1 - 9x_2 + x_4 = 0$

Chapter 2

Set, Function and Correspondence

2.1 Set

2.1.1 open set and closed set

Definition 2.1.1. Let $x \in \mathbb{R}^n$ and $\epsilon > 0$. The *ϵ-ball* of x, denoted $B_\epsilon(x)$, is defined to be $B_\epsilon(x) = \{y \in \mathbb{R}^n \big| \ ||x - y|| < \epsilon\}$.

Definition 2.1.2. A set $A \subset \mathbb{R}^n$ is said to be *open* if, for each $x \in A$, there exists $\epsilon > 0$ such that $B_\epsilon(x) \subset A$.

Definition 2.1.3. A set $A \subset \mathbb{R}^n$ is said to be *closed* if $\mathbb{R} \backslash A$ is open.

Example 2.1.1.

1) $\{(x_1, x_2) \in \mathbb{R}^2 \big| \ x_1^2 + x_2^2 < 1\}$ is open.

2) $\{(x_1, x_2) \in \mathbb{R}^2 \big| \ x_1^2 + x_2^2 \leq 1\}$ is closed.

Check 2.1.1. Properties of open set

- $\mathbb{R}^n$ and $\emptyset$ are open.
- The finite intersection of open sets is open.
- Any union of open sets is open.

Check 2.1.2. Properties of closed set

- $\mathbb{R}^n$ and $\emptyset$ are closed.
- Any intersection of closed sets is closed.
- The finite union of closed sets is closed.

Definition 2.1.4. Let $\{x^k\}$ be a sequence of points in $\mathbb{R}^n$. We say that $\{x^k\}$ *converges* to $x \in \mathbb{R}^n$, denoted $x^k \to x$, if, for all $\epsilon > 0$, there exists N such that $||x - x^m|| < \epsilon$ for $m > N$.

Remark. $x^k \to x$ iff $\left[x_i^k \to x_i, \ \forall i = 1, 2, \cdots, n\right]$

Theorem 2.1.1. *A set $A \subset \mathbb{R}^n$ is closed iff, for every sequence $\{x^k\} \subset A$ such that $x^k \to x$, $x \in A$.*

2.1.2 bounded set and compact set

Definition 2.1.5. A set $S \subset \mathbb{R}^n$ is *bounded* if there exists a number B such that $||x|| \leq B$ for all $x \in S$, that is, if S is contained in some ball in $\mathbb{R}^n$.

Definition 2.1.6. Let J be any index set.

- Given a set A, a collection $\{U_\alpha\}_{\alpha \in J}$ of sets whose union contains A is called a *cover* of A.
- It is an *open cover* if each U_α is open.
- A *subcover* of a given cover is merely a subcollection of it whose union also contains A.
- It is a *finite subcover* if the subcover contains only a finite number of sets.

Definition 2.1.7. A set $A \subset \mathbb{R}^n$ is *compact* if every open cover $\{U_\alpha\}_{\alpha \in J}$ of A has a finite subcover $\{U_i\}_{i=1}^m$.

Example 2.1.2. Let $A = (0, 1]$. Find an open cover with no finite subcover.

Theorem 2.1.2. *A set $A \subset \mathbb{R}^n$ is compact iff A is closed and bounded.*

Example 2.1.3.

- $(-\infty, b]$ is closed but not compact.
- (a, b) is bounded but not compact.
- $[a, b]$ is compact.

2.1.3 convex set and hyperplane

Definition 2.1.8. A set $X \subset \mathbb{R}^n$ is *convex* if for each $x_1, x_2 \in X$ and each $\alpha \in (0, 1)$, $\alpha x_1 + (1-\alpha)x_2 \in X$.

Definition 2.1.9. Let $p \in \mathbb{R}^n$ with $p \neq 0$ and c be a real number. The set $H = \{x \in \mathbb{R}^n |\ p \cdot x = c\}$ is called a *hyperplane* in $\mathbb{R}^n$ with respect to p and c.

Remark. For all $x, x' \in H$, $p \cdot x = \ p \cdot x'$. Thus $p \cdot (x - x') = 0$. This means that the normal vector p is orthogonal to H.

Theorem 2.1.3 (Minkowski's seperating hyperplane theorem)**.** *Let X and Y be nonempty, convex and disjoint sets in $\mathbb{R}^n$. Then there exists $p \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ and a real number c such that $p \cdot x \geq c \geq p \cdot y$ for all $x \in X$ and $y \in Y$.*

2.2 Function

2.2.1 continuity

Definition 2.2.1. Let $A \subset \mathbb{R}^n$, $f : A \to \mathbb{R}^n$, and $x^0 \in A$. f is *continuous at* x^0 if, for any $\epsilon > 0$, there exists $\delta > 0$ such that $||x - x^0|| < \delta$ implies $||f(x) - f(x^0)|| < \epsilon$. f is *continuous* if f is continuous as x^0 for all $x^0 \in A$.

Theorem 2.2.1. *$f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous iff for any open set $O \in \mathbb{R}^m$, $f^{-1}(O)$ is open in $\mathbb{R}^n$.*

Theorem 2.2.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a continuous function. If a subset $K \in \mathbb{R}^n$ is compact, then $f(K)$ is compact in $\mathbb{R}^m$.*

Theorem 2.2.3 (Weierstrass Theorem)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function and $K \in \mathbb{R}^n$ be a compact set. Then there exist $M, m \in K$ such that $f(M) = \max_{x \in K} f(x)$ and $f(m) = \min_{x \in K} f(x)$.*

2.2.2 monotonicity

Definition 2.2.2. A function $f : \mathbb{R} \to \mathbb{R}$ is called *monotonic* (also monotonically increasing, increasing, or non-decreasing), if for all x and y such that $x \leq y$ one has $f(x) \leq f(y)$, so f preserves the order.

Definition 2.2.3. A function $f : \mathbb{R} \to \mathbb{R}$ is called *monotonically decreasing* (also decreasing, non-increasing), if for all x and y such that $x \leq y$ one has $f(x) \geq f(y)$, so f reverses the order.

Remark. If the order $\leq$ in the definition of monotonicity is replaced by the strict order $<$, then one obtains a stronger requirement. A function with this property is called *strictly increasing.* Again, by inverting the order symbol, one finds a corresponding concept called *strictly decreasing.*

Remark. Functions that are strictly increasing or decreasing are one-to-one.

2.2.3 linearity

Definition 2.2.4. For any scalar k, a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called *homogeneous of degree k*, if $f(\alpha x) = \alpha^k f(x)$ and all $\alpha > 0$.

Check 2.2.1. If $q = f(x)$ is a production function which is

- homogeneous of degree one, then such a firm said to exhibit constant returns to scale.
- homogeneous of degree $k > 1$, then such a firm said to exhibit increasing returns to scale.
- homogeneous of degree $k < 1$, then such a firm said to exhibit decreasing returns to scale.

Definition 2.2.5. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called *linear*, if it satisfies the following two properties: for $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$,

- Additivity property (or superposition property): $f(x + y) = f(x) + f(y)$.
- Homogeneity property: $f(\alpha x) = \alpha f(x)$ for all α.

Remark. A linear function $f : \mathbb{R}^n \to \mathbb{R}^m$ could be written $y = Ax$, where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ and A is a $m \times n$ matrix.

Definition 2.2.6. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called an *affine* function if it is a polynomial of degree one.

Remark. An affine function $f : \mathbb{R}^n \to \mathbb{R}^m$ could be written $y = Ax + b$, where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, A is some $m \times n$ matrix and b is some m-vector.

Definition 2.2.7. A (utility) function $U : \mathbb{R}^n \to \mathbb{R}$ is called *quasilinear*, if it is linear in one argument, generally the numeraire.

Remark. A quasilinear (utility) function could be written $U(x_1, x_2, \cdots, x_{n-1}, y) = u(x_1, x_2, \cdots, x_{n-1}) + by$, where $u : \mathbb{R}^{n-1} \to \mathbb{R}$ and b is a positive constant.

2.2.4 concavity and convexity

Definition 2.2.8. Let $f : X \to \mathbb{R}$ be a function where $X \subset \mathbb{R}^n$ is a convex set. For all $x_1, x_2 \in X$, and $\alpha \in (0, 1)$,

- f is called *concave*, if $f(\alpha x_1 + (1-\alpha)x_2) \geq \alpha f(x_1) + (1-\alpha)f(x_2)$.
- f is called *strictly concave*, if $f(\alpha x_1 + (1-\alpha)x_2) > \alpha f(x_1) + (1-\alpha)f(x_2)$.
- f is called *convex*, if $f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha)f(x_2)$.
- f is called *strictly convex*, if $f(\alpha x_1 + (1-\alpha)x_2) < \alpha f(x_1) + (1-\alpha)f(x_2)$.

Check 2.2.2. Geographical properties

- $\Big[f \text{ is concave}\Big] \Leftrightarrow \Big[(\text{height of line segment}) \leq (\text{height of arc})\Big]$
- $\Big[f \text{ is convex}\Big] \Leftrightarrow \Big[(\text{height of line segment}) \geq (\text{height of arc})\Big]$
- $\Big[f \text{ is strictly concave}\Big] \Leftrightarrow \Big[(\text{height of line segment}) < (\text{height of arc})\Big]$
- $\Big[f \text{ is strictly convex}\Big] \Leftrightarrow \Big[(\text{height of line segment}) > (\text{height of arc})\Big]$

2.2.5 quasi-concavity and quasi-convexity

Definition 2.2.9. Let $f : X \to \mathbb{R}$ be a function where $X \subset \mathbb{R}^n$ is a convex set.

- f is called *quasi-concave* if, for all $k \in \mathbb{R}$, the set $\{x \in X | \ f(x) \geq k\}$ is convex.
- f is called *quasi-convex* if, for all $k \in \mathbb{R}$, the set $\{x \in X | \ f(x) \leq k\}$ is convex.

Remark. $\{x \in X | \ f(x) \geq k\}$ is called *upper contour set* and $\{x \in X | \ f(x) \leq k\}$ is called *lower contour set.*

Check 2.2.3.

- If f is concave, it is quasi-concave.
- If f is convex, it is quasi-convex.

Theorem 2.2.4. *Let $f : X \to \mathbb{R}$ be a function where $X \subset \mathbb{R}^n$ is a convex set. For all $x_1, x_2 \in X$, and $\alpha \in (0, 1)$,*

- *f is quasi-concave iff* $f(\alpha x_1 + (1-\alpha)x_2) \geq \min\{f(x_1), f(x_2)\}$.

- f *is quasi-convex iff* $f(\alpha x_1 + (1-\alpha)x_2) \leq \max\{f(x_1), f(x_2)\}$.

Definition 2.2.10. Let $f : X \to \mathbb{R}$ be a function where $X \subset \mathbb{R}^n$ is a convex set. For all $x_1, x_2 \in X$, and $\alpha \in (0,1)$,

- f is *strictly quasi-concave* if, $f(\alpha x_1 + (1-\alpha)x_2) > \min\{f(x_1), f(x_2)\}$.
- f is *strictly quasi-convex* if, $f(\alpha x_1 + (1-\alpha)x_2) < \max\{f(x_1), f(x_2)\}$.

Check 2.2.4. We do not define strictly quasi-concave and strictly quasi-convex with the set $\{x \in X |\ f(x) > k\}$. Why?

Check 2.2.5. For all $x_1, x_2 \in X$, and $\alpha \in (0,1)$,

- $\Big[f \text{ is quasi-concave}\Big]$
  $\Leftrightarrow \Big[f(x_1) \leq f(x_2) \text{ implies } f(\alpha x_1 + (1-\alpha)x_2) \geq f(x_1)\ \Big]$
- $\Big[f \text{ is quasi-convex}\Big]$
  $\Leftrightarrow \Big[f(x_1) \leq f(x_2) \text{ implies } f(\alpha x_1 + (1-\alpha)x_2) \leq f(x_2)\ \Big]$
- $\Big[f \text{ is strictly quasi-concave}\Big]$
  $\Leftrightarrow \Big[f(x_1) \leq f(x_2) \text{ implies } f(\alpha x_1 + (1-\alpha)x_2) > f(x_1)\ \Big]$
- $\Big[f \text{ is strictly quasi-convex}\Big]$
  $\Leftrightarrow \Big[f(x_1) \leq f(x_2) \text{ implies } f(\alpha x_1 + (1-\alpha)x_2) < f(x_2)\ \Big]$

2.3 Correspondence

2.3.1 correspondence and graph

Definition 2.3.1. Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$. A *correspondence* $\Phi : X \rightrightarrows Y$ associates a subset $\Phi(x) \subset Y$ with each $x \in X$.

Remark. A correspondence $\Phi : X \rightrightarrows Y$ may be seen as a function $\Phi : X \to 2^Y$ where 2^Y is the power set of Y.

Definition 2.3.2. The *graph* of a correspondence $\Phi : X \rightrightarrows Y$ is defined to be $G(\Phi) = \{(x,y) \in X \times Y |\ y \in \Phi(x)\}$.

Definition 2.3.3. A correspondence $\Phi : X \rightrightarrows Y$ has a *closed graph* if, for every sequence $\{(x^n, y^n)\} \subset G(\Phi)$ with $(x^n, y^n) \to (x,y)$, we have $(x,y) \in G(\Phi)$.

2.3.2 continuity

Definition 2.3.4. A correspondence $\Phi : X \rightrightarrows Y$ is called *upper hemicontinuous(u.h.c.)* if it has a closed graph and the images of compact sets are bounded, that is, for each compact set $B \subset X$, $\Phi(B)$ is bounded, where $\Phi(B) = \{y \in Y |\ y \in \Phi(x) \text{ for some } x \in B\}$.

Definition 2.3.5. A correspondence $\Phi : X \rightrightarrows Y$ is called *lower hemicontinuous(l.h.c.)* if, for every sequence $\{x^n\} \subset X$ with $x^n \to x \in X$ and every $y \in \Phi(x)$, we can find a sequence $\{y^n\}$ and an integer N such that, for all $n > N$, $y^n \in \Phi(x^n)$ and $y^n \to y$.

Definition 2.3.6. A correspondence $\Phi : X \rightrightarrows Y$ is *continuous* if it is u.h.c. and l.h.c.

Theorem 2.3.1. *Let $\Phi : X \rightrightarrows Y$ be a single valued correspondence.*

- *If Φ is u.h.c., it is a continuous function $X \to Y$.*
- *If Φ is l.h.c., it is a continuous function $X \to Y$.*

2.3.3 Berge's Maximum Theorem

Theorem 2.3.2. *Let $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$, and $f : X \times Y \to \mathbb{R}$. Let $\Phi : X \rightrightarrows Y$ be a continuous correspondence with nonempty and compact values. Then*

- *The correspondence $y^\star : X \rightrightarrows Y$ defined by $y^\star(x) \equiv \text{argmax}_{y\in\Phi(x)} f(x, y)$ is u.h.c.*
- *The function $f^\star : X \to \mathbb{R}$ defined by $f^\star(x) \equiv \max_{y\in\Phi(x)} f(x, y)$ is continuous.*

2.4 Fixed Point Theorem

Definition 2.4.1.

- Let $X \subset \mathbb{R}^n$ and $f : X \to X$. Then $x \in X$ is called a *fixed point* of f if $x = f(x)$.
- Let $X \subset \mathbb{R}^n$ and $\Phi : X \rightrightarrows X$. Then $x \in X$ is called a *fixed point* of Φ if $x \in \Phi(x)$.

Theorem 2.4.1 (Brouwer's Fixed Point Theorem)**.** *Let $X \in \mathbb{R}^n$ be nonempty, compact and convex set and $f : X \to X$ is a continuous function. Then f has a fixed point.*

Theorem 2.4.2 (Kakutani's Fixed Point Theorem)**.** *Let $X \in \mathbb{R}^n$ be nonempty, compact and convex set and $\Phi : X \rightrightarrows X$ is upper hemicontinuous and convex-valued correspondence. Then Φ has a fixed point.*

Chapter 3

Differential Calculus

3.1 Functions from $\mathbb{R}$ to $\mathbb{R}$

Definition 3.1.1. Let $X, Y \in \mathbb{R}$ be open sets. Let $f : X \to Y$. Then the derivative of f at $x^0 \in X$ is defined to be

$$\frac{df}{dx}(x^0) \equiv \lim_{h \to 0} \frac{f(x^0 + h) - f(x^0)}{h}$$

if this limit exists.

Remark. We also denote the derivative of f at x^0 by $\{f(x^0)\}'$ or $f'(x^0)$.

Check 3.1.1. Derivative rules

i. $\{f(x) + g(x)\}' = f'(x) + g'(x)$.

ii. $\{f(x)g(x)\}' = f'(x)g(x) + f(x)g'(x)$.

iii. $\left\{\frac{f(x)}{g(x)}\right\}' = \frac{f'(x)g(x) - f(x)g'(x)}{\{g(x)\}^2}$.

iv. $\{f(g(x))\}' = f'(g(x))g'(x)$.

v. If f is invertible and $f'(x^0) \neq 0$ then $\{f^{-1}(y^0)\}' = \{f'(x^0)\}^{-1}$ where $y^0 = f(x^0)$.

Check 3.1.2. Properties of exponential function

Definition 3.1.2. For $a > 0$, the inverse function of $y = a^x$ is called the base a logarithm and is denoted by $\log_a y$. And the base e logarithm is called the natural logarithm and is written as $\log y$ or $\ln y$.

Remark. We denote as $e \equiv \lim_{n \to \infty} \left\{1 + \frac{1}{n}\right\}^n \fallingdotseq 2.7182...$

Remark. $e = \sum_{i=0}^{\infty} (i!)^{-1} = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \cdots$

Check 3.1.3. Properties of the logarithm

Check 3.1.4. Derivative of exponential function and logarithm.

i. $(e^x)' = e^x$.

ii. $(a^x)' = a^x \ln a$.

iii. $(\ln x)' = \frac{1}{x}$.

iv. $(\ln f(x))' = \frac{f'(x)}{f(x)}$.

3.2 Functions from $\mathbb{R}^n$ to $\mathbb{R}^m$

3.2.1 partial derivative and Jacobian matrix

Definition 3.2.1. Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ be open sets. Let $f : X \to Y$, so that

$$\begin{aligned} y_1 &= f^1(x_1, \cdots, x_n) \\ y_2 &= f^2(x_1, \cdots, x_n) \\ &\vdots \\ y_m &= f^m(x_1, \cdots, x_n). \end{aligned}$$

Then the *partial derivative* of f^i with respect to x_j at $x^0 \in X$ is defined to be:

$$\frac{\partial f^i}{\partial x_j}(x^0) \equiv \lim_{h_j \to 0} \frac{f(x_1^0, \cdots, x_j^0 + h_j, \cdots, x_n^0) - f(x^0)}{h_j}$$

if this limit exists.

Remark. $\frac{\partial f^i}{\partial x_j}(x^0)$ is denoted as $f_j^i(x^0)$.

Definition 3.2.2. Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ be open sets.Let h be a n-vector. Let $f : X \to Y$. Then f is called differentiable at $x^0 \in X$ if there exists an $m \times n$ matrix A such that

$$\lim_{h \to 0} \frac{\left\| f(x^0 + h) - [f(x^0) + Ah] \right\|}{\|h\|} = 0.$$

The matrix A is called the *derivative* of or the *Jacobian derivative* of f at x^0.

Remark. The derivative of f at x^0 is denoted $Df(x^0)$

3.2.2 total derivative

Theorem 3.2.1. *Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ be open sets. Let $f : X \to Y$. Suppose that f is differentiable at $x^0 \in X$. Then*

i. For each i and j, the partial derivative of f^i with respect to x_j at $x^0 \in X$, $f_j^i(x^0)$, exists.

ii. The derivative of f at x^0 is equal to $m \times n$ matrix whose entries are the partial derivatives of f at x^0, that is,

$$Df(x^0) = \begin{pmatrix} f_1^1(x^0) & f_2^1(x^0) & \cdots & f_n^1(x^0) \\ f_1^2(x^0) & f_2^2(x^0) & \cdots & f_n^2(x^0) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^m(x^0) & f_2^m(x^0) & \cdots & f_n^m(x^0) \end{pmatrix}.$$

Theorem 3.2.2. *Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ be open sets. Let $f : X \to Y$. If for each i and j, $f^i_j(x^0)$ exists and f^i_j is continuous in the neighborhood of $x^0 \in X$ then f is differentiable at x^0.*

Definition 3.2.3. Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ be open sets. Let $f : X \to Y$. Then the expression $dy = Df(x^0)dx$ is called the *total differential* of f at x^0.

Theorem 3.2.3 (Chain Rule)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^p \to \mathbb{R}^n$. Let $w^0 \in \mathbb{R}^p$ and $x^0 = g(w^0) \in \mathbb{R}^n$. Suppose that f and g are differentiable at x^0 and w^0, respectively. Then $D(f \circ g)(x^0) = Df(x^0)Dg(w^0)$.*

Example 3.2.1. Let $z = x^2 + 3xy + y^3$, $x = 1 + t$, $y = t^2$. find $\frac{dz}{dt}$.

3.3 Functions from $\mathbb{R}^n$ to $\mathbb{R}$

3.3.1 gradient vector

Definition 3.3.1. Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}$ be open sets. Let $f : X \to Y$ be a real-valued function. The *gradient* vector of f at x^0 is defined to be the column vector

$$\nabla f(x^0) \equiv \begin{pmatrix} f_1(x^0) \\ f_2(x^0) \\ \vdots \\ f_n(x^0) \end{pmatrix}.$$

Remark. The derivative of f at x^0 is the row vector, that is,

$$Df(x^0) = \begin{pmatrix} f_1(x^0) & f_2(x^0) & \cdots & f_n(x^0) \end{pmatrix}.$$

Definition 3.3.2. Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}$ be open sets. Let $f : X \to Y$ be a real-valued function. The derivative of f at x^0 in the direction v, denoted $\frac{\partial f}{\partial v}(x^0)$, is defined to be

$$\frac{\partial f}{\partial v}(x^0) \equiv \lim_{t \to 0} \frac{f(x^0 + tv) - f(x^0)}{t}.$$

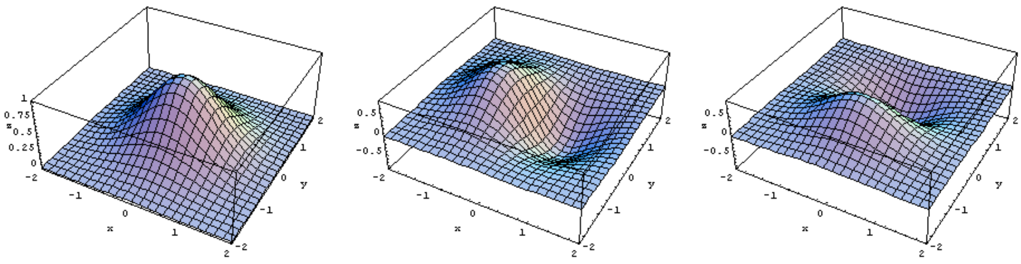
Theorem 3.3.1. *Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}$ be open sets. Let $f : X \to Y$ be a real-valued differentiable function at $x^0 \in X$. Then $\frac{\partial f}{\partial v}(x^0) = Df(x^0)v$.*

Theorem 3.3.2. *Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}$ be open sets. Let $f : X \to Y$ be a real-valued $\mathcal{C}^1$ function. Then, for all $x^0 \in X$ at which $\nabla f(x^0) \neq \mathbf{0}$, the gradient vector $\nabla f(x^0)$ points at x into the direction in which f increases most rapidly.*

Check 3.3.1. If f is k times continuously differentiable, we say that f is $\mathcal{C}^k$.

Example 3.3.1. Find $\nabla f(1, 1)$ and $\nabla f(1, -1)$, where $f(x, y) = e^{-x^2-y^2}$. And check the directions. See `Figure` 3.1.

Example 3.3.2. Find the tangent plane at $(1, 1, 2)$ on the graph $z = x^2 + y^2$. See `Figure` 3.2.



(a) $z = f(x, y) = e^{-x^2-y^2}$ (b) $f_x(x, y) = -2xe^{-x^2-y^2}$ (c) $f_y(x, y) = -2ye^{-x^2-y^2}$

Figure 3.1: Graphics for Example 3.3.1

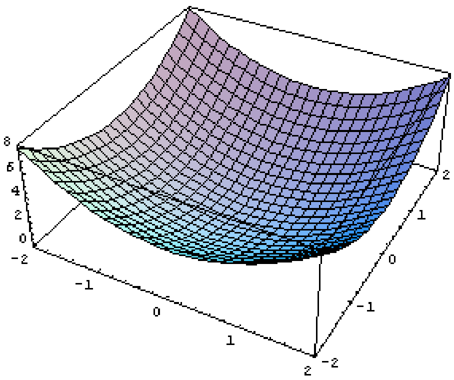


Figure 3.2: $z = x^2 + y^2$

3.3.2 Hessian matrix

Definition 3.3.3. Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}$ be open sets. Let $f : X \to Y$ be a real-valued function. The $x_i x_j$*-second order partial derivative* of f at x^0, denoted $\frac{\partial^2 f}{\partial x_j \partial x_i}(x^0)$, is defined to be $\frac{\partial^2 f}{\partial x_j \partial x_i}(x^0) \equiv \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}(x)\right)\Big|_{x=x^0}$.

Remark. Other frequently used notation for second order partial derivatives includes

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = f_{x_i x_j} = f_{ij} = D_{ij} f.$$

Theorem 3.3.3 (Young's Theorem)**.** *Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}$ be open sets. Let $f : X \to Y$ be a real-valued $\mathcal{C}^2$ function. Then, for all $x \in X$ and for each pair i and j, $f_{ij}(x) = f_{ji}(x)$.*

Definition 3.3.4. Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}$ be open sets. Let $f : X \to Y$ be a real-valued $\mathcal{C}^2$ function. The Hessian matrix of f at $x^0 \in X$, denoted $D^2 f(x^0)$, is defined to be:

$$D^2 f(x^0) \equiv \begin{pmatrix} f_{11}(x^0) & f_{12}(x^0) & \cdots & f_{1n}(x^0) \\ f_{21}(x^0) & f_{22}(x^0) & \cdots & f_{2n}(x^0) \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1}(x^0) & f_{n2}(x^0) & \cdots & f_{nn}(x^0) \end{pmatrix} = \left(f_{ij}(x^0)\right).$$

3.3.3 Implicit function theorem

Theorem 3.3.4. *Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ be open sets. Let $F : X \times Y \to \mathbb{R}^m$ be $\mathcal{C}^1$ function. Consider $F(y, x) = \mathbf{0}$ as possibly defining y as an implicit function of x. Suppose that $F(y^0, x^0) = \mathbf{0}$ and $|D_y F(y^0, x^0)| \neq 0$. Then*

i. There exist $\mathcal{C}^1$ function $y = f(x)$ and a neighborhood $B_\epsilon(x^0)$ of x^0 such that $F(f(x), x) = \mathbf{0}$ for all $x \in B_\epsilon(x^0)$ and $y^0 = f(x^0)$.

ii. The derivative of f at x^0 can be computed as

$$Df(x^0) = -\{D_y F(y^0, x^0)\}^{-1} D_x F(y^0, x^0).$$

Example 3.3.3. For $F(x, y) = x^2 + y^2 - 4$.

1) Find $y = f(x)$ such that $F(f(x), x) = 0$

2) Find $\frac{df(x^0)}{dx}$ at $(x^0, y^0) = (2, 2)$.

Example 3.3.4. Consider a system of equations

$$x^3 + uy^3 = 1, \qquad xy = v.$$

Let its solution (x^0, y^0) be a function of u and v. Find $\frac{\partial x^0}{\partial u}$ and $\frac{\partial x^0}{\partial u}$.

3.4 Forms of Differentiable Function

3.4.1 concavity and convexity for $\mathcal{C}^1$ or $\mathcal{C}^2$ function

Theorem 3.4.1. *Let $f : X \to \mathbb{R}$ be a $\mathcal{C}^1$ function. For all $x^0 \in X$, $x^0 \neq x$,*

- $\left[f \textit{ is concave}\right] \Leftrightarrow \left[f(x) \leq f(x^0) + Df(x^0)(x - x^0)\right]$
- $\left[f \textit{ is convex}\right] \Leftrightarrow \left[f(x) \geq f(x^0) + Df(x^0)(x - x^0)\right]$
- $\left[f \textit{ is strictly concave}\right] \Leftrightarrow \left[f(x) < f(x^0) + Df(x^0)(x - x^0)\right]$
- $\left[f \textit{ is strictly convex}\right] \Leftrightarrow \left[f(x) > f(x^0) + Df(x^0)(x - x^0)\right]$

Theorem 3.4.2. *Let $f : X \to \mathbb{R}$ be a $\mathcal{C}^2$ function. For all $x^0 \in X$,*

- $\left[f \textit{ is concave}\right] \Leftrightarrow \left[D^2 f(x^0) \textit{ is N.S.D}\right]$
- $\left[f \textit{ is convex}\right] \Leftrightarrow \left[D^2 f(x^0) \textit{ is P.S.D}\right]$
- $\left[f \textit{ is strictly concave}\right] \Leftrightarrow \left[D^2 f(x^0) \textit{ is N.D}\right]$
- $\left[f \textit{ is strictly convex}\right] \Leftrightarrow \left[D^2 f(x^0) \textit{ is P.D}\right]$

3.4.2 quasi-concavity and quasi-convexity for $\mathcal{C}^1$ or $\mathcal{C}^2$ function

Theorem 3.4.3. *Let $f : X \to \mathbb{R}$ be a $\mathcal{C}^1$ function. For all $x_1, x_2 \in X$,*

- $\Big[f \text{ is quasi-concave}\Big] \Leftrightarrow \Big[f(x_1) \leq f(x_2) \text{ implies } Df(x^1)(x_2 - x_1) \geq 0\ \Big]$
- $\Big[f \text{ is quasi-convex}\Big] \Leftrightarrow \Big[f(x_1) \leq f(x_2) \text{ implies } Df(x^1)(x_2 - x_1) \leq 0\ \Big]$

Theorem 3.4.4. *Let $f : X \to \mathbb{R}$ be a $\mathcal{C}^2$ function where $X \subset \mathbb{R}^n$. Consider the bodered Hessian B defined as*

$$B \equiv \begin{pmatrix} 0 & f_1 & f_2 & \cdots & f_n \\ f_1 & f_{11} & f_{12} & \cdots & f_{1n} \\ f_2 & f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_n & f_{n1} & f_{n2} & \cdots & f_{nn} \end{pmatrix}.$$

- *If the largest $(n-1)$ leading principal minors of B alternate in sign, for all $x \in X$, with the smallest of these – the third order leading principal minor – positive, then f is strictly quasi-concave.*
- *If these largest $(n-1)$ leading principal minors are all negative for all $x \in X$, then f is strictly quasi-convex.*

Example 3.4.1. Consider $f(x, y) = x^\alpha y^\beta$ $(\alpha, \beta > 0)$ defined on $\mathbb{R}^2_{++}$. Solve the bordered Hessian matrix and Check its quasi-concavity.

Further Study 3.4.1. Pseudoconcave function: see `Simon and Blume (p527)`

Chapter 4

Optimization

In this chapter we consider a function $f : \mathbb{R}^n \to \mathbb{R}$.

4.1 Multivariate optimization

4.1.1 First-order and second-order conditions

Theorem 4.1.1 (F.O.C)**.** *Suppose that $f \in \mathcal{C}^1$ and that $x^0 \in \mathbb{R}^n$ is a local maximizer or local minimizer of f. Then $\nabla f(x^0) = \mathbf{0}$.*

Theorem 4.1.2 (S.O.C)**.** *Suppose that $f \in \mathcal{C}^2$ and that $\nabla f(x^0) = \mathbf{0}$.*

i. Local maximizer

- *If $x^0 \in \mathbb{R}^n$ is a local maximizer, then the (symmetric) $n \times n$ matrix $D^2 f(x^0)$ is N.S.D.*
- *If $D^2 f(x^0)$ is N.D., then x^0 is a local maximizer.*

ii. Local minimizer

- *If $x^0 \in \mathbb{R}^n$ is a local minimizer, then the (symmetric) $n \times n$ matrix $D^2 f(x^0)$ is P.S.D.*
- *If $D^2 f(x^0)$ is P.D., then x^0 is a local minimizer.*

Example 4.1.1. For $f(x, y) = xy$, find the maximum or minimum, if it exists. See `Figure` 4.1.

4.1.2 Envelope theorem

Theorem 4.1.3 (Envelope Theorem)**.** *For $x \in \mathbb{R}^s$ and $q \in \mathbb{R}^t$, consider the maximization problem* $\max_x f(x; q)$ *and define the optimal* value function $F(q) \equiv f(x(q); q)$. *Then*

$$\frac{dF(q)}{dq} = \frac{\partial f(x(q); q)}{\partial q} \equiv \frac{\partial f(x; q)}{\partial q}\Big|_{x=x(q)}.$$

Remark. We usually consider x as endogenous variable and q as exogenous variable.

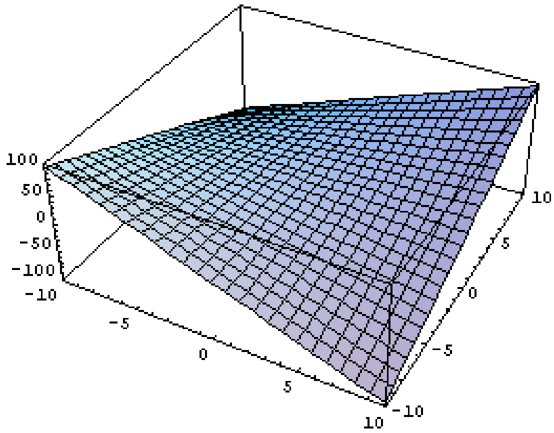


Figure 4.1: $z = xy$

Remark. $x(q)$ is called as the optimal choice function.

Example 4.1.2. Consider $f(x;q) = -2x^2 + 8qx + q^2$.

1) Find $x^0(q)$ such that maximize $f(x;q)$.

2) Find maximized $f(x)$, that is, $F(q) \equiv f(x^0(q);q)$.

3) Find $\frac{d}{dq}F(q)$ and $\frac{\partial}{\partial q}f(x^0(q);q)$. Compare.

4) Explain the envelope theorem with the `Figure` 4.2

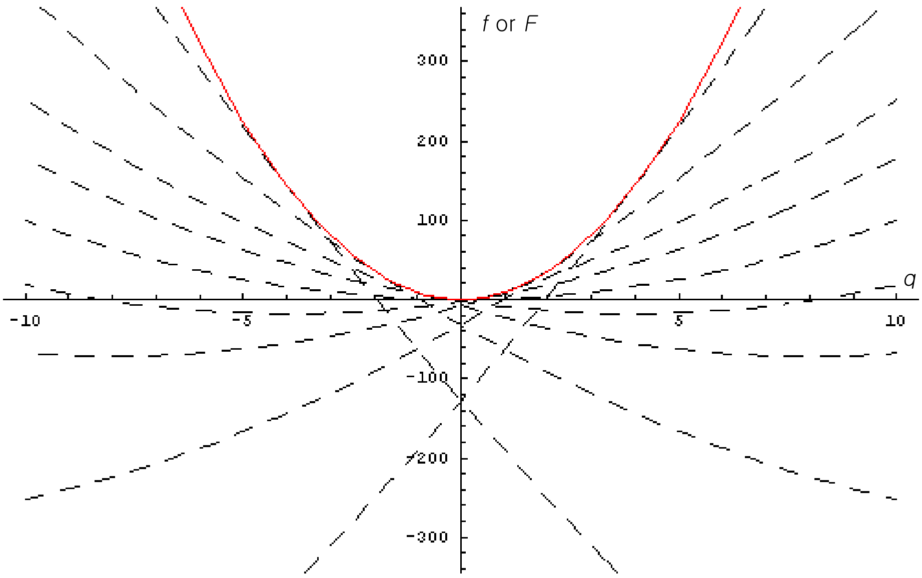


Figure 4.2: $f = -2x^2 + 8qx + q^2$ and $F = 9q^2$ on q-f or q-F plane

4.2 Constrained maximization

4.2.1 Lagrangian multiplier method

Definition 4.2.1. Let $f, h_i, h_2, \cdots, h_m$ be real-valued $\mathcal{C}^1$ functions. For $x \in \mathbb{R}^s$ and $q \in \mathbb{R}^t$, consider the following maximization problem:

$$\max_x f(x;q) \quad \text{s.t.} \qquad \begin{aligned} h_1(x;q) &= \bar{b}_1 \\ &\vdots \\ h_m(x;q) &= \bar{b}_m \end{aligned}$$

The *Lagrangian* function is defined as $\mathcal{L}(x,\lambda;q) \equiv f(x;q)+\sum_{i=1}^m \lambda_i(h_i(x;q)-\bar{b}_i)$.

Theorem 4.2.1 (F.O.C)**.** *Let $h(x;q) \equiv [h_1(x;q), \cdots, h_m(x;q)]'$. Suppose that x^0 is a local maximizer of the problem such that $|D_x h(x^0)| \neq 0$. Then there exists $\lambda^0 \equiv [\lambda_1^0, \cdots, \lambda_m^0]'$ which satisfies the following first-order condition*

$$D_{(\lambda,x)}\mathcal{L}(x^0,\lambda^0;q) = \mathbf{0}.$$

Theorem 4.2.2 (S.O.C)**.** *Suppose that x^0 and λ^0 satisfies the F.O.C.. If the Hessian matrix $D_x^2\mathcal{L}(x^0,\lambda^0;q)$ is N.D. on the linear constraint $D_x h(x^0;q)v = 0$, then x^0 is a strict local maximizer of the maximization problem.*

Remark. For (x^0, λ^0) satisfying the first-order condition of a maximization problem, let us construct the Bordered Hessian matrix

$$\bar{H} \equiv \begin{pmatrix} \mathbf{0} & D_x h(x^0;q) \\ D_x h(x^0;q)' & D_x^2\mathcal{L}(x^0,\lambda^0;q) \end{pmatrix} = D^2_{(\lambda,x)}\mathcal{L}(x^0,\lambda^0;q).$$

If $(-1)^s|\bar{H}_{m+s}| > 0, (-1)^{s-1}|\bar{H}_{m+s-1}| > 0, \cdots, (-1)^{m+1}|\bar{H}_{(m+s)-(s-m-1)}| > 0$ then x^0 is a strict local maximizer of the maximization problem.

Example 4.2.1. Solve a following maximization problem:

$$\max_{x,y} f(x,y) = x^2y \quad \text{s.t.} \quad 2x^2 + y^2 = 3$$

See `Figure` 4.3.

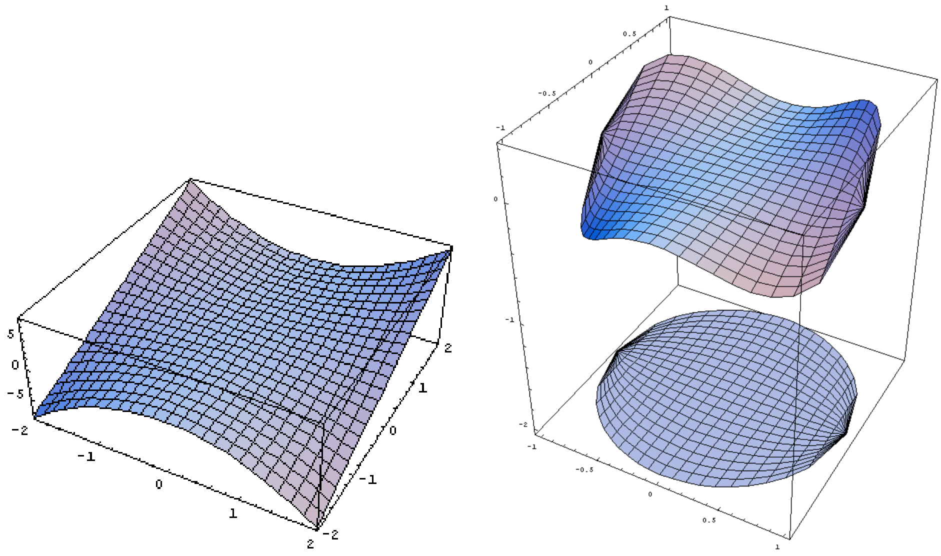
4.2.2 Envelope theorem with linear constraints

Theorem 4.2.3. *Let $x(q)$ be the solution of the maximization problem given q. Then*

$$\begin{aligned} \frac{df(x(q);q)}{dq_k}\Big|_{dq_j=0,\forall j\neq k} &= \frac{\partial \mathcal{L}}{\partial q_k}(x(q),\lambda(q),q) \\ &= \frac{\partial f}{\partial q_k}(x(q);q) + \sum_{i=1}^m \lambda_i(q)\frac{\partial h_i}{\partial q_k}(x(q),q). \end{aligned}$$

Further Study 4.2.1. Comparative statistics: see `Varian (ch.27)`

Further Study 4.2.2. Kuhn-Tucker theorem: see `Varian (ch.27)`



(a) $z = f(x, y) = x^2y$ (b) $z = f(x, y) = x^2y$ on $2x^2 + y^2 = 3$

Figure 4.3: Graphics for Example 4.2.1