Econ 703 TA Note 8

Chia-Min Wei *

October 23, 2025

1 Directional Derivatives and Differentiability on $\mathbb{R}^n$

Definition 1.1: Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, $c \in \mathbb{R}^n$ a point, and $u \in \mathbb{R}^n$ a direction vector with unit length. We say that f is differentiable in the direction of u if

$$D_u f(c) = \lim_{h \to 0+} \frac{f(c + hu) - f(x)}{h}$$

exists. $D_u f(c)$ is called the (one-sided) directional derivative of f at c in the direction of u.

Among all directional derivatives, those in the coordinate directions $e_i = (0, \cdots, 1, \cdots, 0)$ have a special notation:

$$D_{e_i} f(c) = \frac{\partial f}{\partial x_i}(c).$$

Definition 1.2: A function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at c if

(i) f is differentiable at c in the direction of all unit vectors.

(ii) For any unit vector $u = (u_1, \cdots, u_n) \in \mathbb{R}^n$,

$$D_u f(c) = \sum_{i=1}^{n} u_i \frac{\partial f}{\partial x_i}(c).$$

Example 1.1: Let $f : \mathbb{R} \to \mathbb{R}$. There are only two possible directions: 1 and -1. Recall the definition of right and left derivatives:

$$f'(c+) = \lim_{h \to 0+} \frac{f(c + h) - f(c)}{h}, \quad f'(c-) = \lim_{h \to 0-} \frac{f(c + h) - f(c)}{h}.$$

The directional derivatives have the following relationships with the right and left derivatives.

$$D_1 f(c) = f'(c+), \quad D_{-1} f(c) = -f'(c-).$$

When f is differentiable at c, since $-1 = (-1) \times 1$, by the definition above,

$$D_{-1} f(c) = -D_1 f(x) = -f'(c+).$$

Hence, f is differentiable at c if and only if $f'(c-) = f'(c+)$.

*This TA note is based on Prof. John Kennan's math camp lecture taught in 2025 at UW-Madison. All errors are mine.

Example 1.2: For multivariate functions, $D_{-u}f(c) = -D_u f(c)$ for all unit vector u does **not** guarantee that f is differentiable at c. Consider the following example:

$$f(x,y) = \begin{cases} \frac{x^2 y}{x^2+y^2}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0). \end{cases}$$

For any unit vector $u = (u_1, u_2)$,

$$\begin{aligned} D_u f(0,0) &= \lim_{t \to 0+} \frac{f(tu_1, tu_2) - f(0,0)}{t} \\ &= \frac{(t^3 u_1^2 u_2)/t^2 - 0}{t} = u_1^2 u_2. \end{aligned}$$

It is clear that $D_{-u}f(0,0) = -u_1^2 u_2 = -D_u f(0,0)$. Note that

$$\frac{\partial f}{\partial x}(0,0) = 0, \quad \frac{\partial f}{\partial y}(0,0) = 0.$$

However, if we take $u = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, then

$$D_u f(0,0) = \frac{1}{2\sqrt{2}} \neq \frac{1}{\sqrt{2}} \frac{\partial f}{\partial x}(0,0) + \frac{1}{\sqrt{2}} \frac{\partial f}{\partial y}(0,0).$$

Therefore, f is not differentiable at $(0,0)$.

Definition 1.3: Let $f : \mathbb{R}^n \to \mathbb{R}$, and f is differentiable at c. The $\mathbb{R}^{n\times 1}$ matrix

$$\nabla f(c) = \left[\frac{\partial f}{\partial x_1}(c), \cdots, \frac{\partial f}{\partial x_n}(c)\right]^\intercal$$

is called the gradient of f at c.

Therefore, for any direction u, $D_u f(c) = \nabla f(c) \cdot u$.

Theorem 1.1: Let $f : \mathbb{R}^n \to \mathbb{R}$. f is differentiable at c with gradient $\nabla f(c)$ if and only if

$$f(c+v) = f(c) + \nabla f(c) \cdot v + r(v)$$

where $r(v) = o(\|v\|)$ as $v \to 0$.

The hyperplane $h(x) = f(c) + \nabla f(c) \cdot (x - c)$ is a good approximation of $f(x)$ near $x = c$.

2 First Order Condition

Definition 2.1 (Local maximal/minimal point): Let $f : A \to \mathbb{R}$ and A is an open subset of $\mathbb{R}^n$. We say that $c \in A$ is a local maximal (minimal) point of f if there exists $\delta > 0$ such that

$$f(c) \geq (\leq) f(x)$$

for all $\|x - c\| < \delta$.

Theorem 2.1 (First Order Condition): Let $f : A \to \mathbb{R}$ be differentiable where A is an open set in $\mathbb{R}^n$. If c is a local maximal or minimal point, then

$$\nabla f(c) = \mathbf{0}.$$

Proof. We will prove the case when c is a local maximal point. Assume the contrary that $\nabla f(c) \neq 0$. Then there exists a unit vector u such that $D_u f(c) = \nabla f(c) \cdot u > 0$. But since c is a local maximal point, for any $x = c + tu$ with small enough $t > 0$,

$$\frac{f(c+tu) - f(c)}{t} < 0.$$

Therefore,

$$D_u f(c) = \lim_{t \to 0+} \frac{f(c+tu) - f(c)}{t} \leq 0.$$

A contradiction. $\square$

The first order condition is a **necessary** condition for a point to be a local maximal or minimal point. A point c with $\nabla f(c) = \mathbf{0}$ but is neither local maximal nor minimal is called a **saddle point**.

3 Second Order Condition

Definition 3.1 (Twice Differentiability): Let $f : A \to \mathbb{R}$ be differentiable where A is an open set in $\mathbb{R}^n$. We say that f is twice differentiable at c if $\frac{\partial f}{\partial x_i}(x)$ is differentiable at c for all coordinate x_i.

The matrix

$$\mathrm{H}(c) = \begin{bmatrix} \frac{\partial f}{\partial x_1 \partial x_1}(c) & \frac{\partial f}{\partial x_1 \partial x_2}(c) & \cdots & \frac{\partial f}{\partial x_1 \partial x_n}(c) \\ \frac{\partial f}{\partial x_2 \partial x_1}(c) & \frac{\partial f}{\partial x_2 \partial x_2}(c) & \cdots & \frac{\partial f}{\partial x_2 \partial x_n}(c) \\ \vdots & & & \\ \frac{\partial f}{\partial x_n \partial x_1}(c) & \frac{\partial f}{\partial x_n \partial x_2}(c) & \cdots & \frac{\partial f}{\partial x_n \partial x_n}(c) \end{bmatrix} = \begin{bmatrix} \nabla \frac{\partial f}{\partial x_1}(c)^\intercal \\ \nabla \frac{\partial f}{\partial x_2}(c)^\intercal \\ \vdots \\ \nabla \frac{\partial f}{\partial x_n}(c)^\intercal \end{bmatrix}$$

is called the Hessian of f at c. The Hessian is symmetric when the cross derivatives are continuous at c (Schwartz's Theorem).

Theorem 3.1: Let $f : A \to \mathbb{R}$ be differentiable on $A \subset \mathbb{R}^n$. f is twice differentiable at c with hessian $\mathrm{H}(c)$ if and only if

$$f(c+u) = f(c) + \nabla f(c)^\intercal u + \frac{1}{2} u^\intercal \mathrm{H}(c) u + r(u)$$

with $r(u) = o(\|u\|^2)$ as $u \to 0$.

Definition 3.2 (Positive-Definite, Negative-Definite and Indefinite Matrix): Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. If for all $v \in \mathbb{R}^n$,

$$v^\intercal A v > (<) 0,$$

then we say that A is a positive-definite (negative-definite) matrix. If there exists $u, v \in \mathbb{R}^n$ such that

$$u^\intercal A u > 0, \quad v^\intercal A v < 0,$$

then we say that A is an indefinite matrix.

Theorem 3.2 (Second Order Condition): Let $f : A \to \mathbb{R}$ be twice continuously differentiable on A with hessian $\mathrm{H}(c)$. Suppose $\nabla f(c) = \mathbf{0}$.

1. If $\mathrm{H}(c)$ is negative-definite, then c is a local maximal point.
2. If $\mathrm{H}(c)$ is positive-definite, then c is a local minimal point.
3. If $\mathrm{H}(c)$ is indefinite, then c is a saddle point.

Proof. Suppose $\mathrm{H}(c)$ is a negative-definite matrix. By Theorem 3.1, we can write

$$f(c+v) = f(c) + \frac{1}{2} v^\intercal \mathrm{H}(c) v + r(v),$$

where $r(v) = o(\|v\|^2)$. Since the function $k(v) = v^\intercal \mathrm{H}(c) v$ is continuous on the unit sphere (which is closed and bounded),

$$m = \max_{\|v\|=1} v^\intercal \mathrm{H}(c) v < 0$$

exists. There exists $\delta > 0$ such that for all $\|v\| < \delta$,

$$|r(v)| < \left|\frac{m}{4}\right| \|v\|^2 .$$

Observe that for all v with $\|v\| < \delta$,

$$\begin{aligned} f(c+v) &= f(c) + \frac{1}{2} v^\intercal \mathrm{H}(c) v + r(v) \\ &\le f(c) + \frac{m}{2} \|v\|^2 - \frac{m}{4} \|v\|^2 \\ &< f(c). \end{aligned}$$

$\square$

The second order condition is a **sufficient** condition for a point to be a local maximal/minimal or a saddle point.

4 Sylvester's Criterion

Definition 4.1 (Principal Minor): A principal minor of an $n \times n$ matrix A is the determinant of some smaller square matrix formed by removing the same rows and columns of A.

Definition 4.2 (Leading Principal Minor): The k^{th} leading principal minor of an $n \times n$ matrix A is the determinant of the smaller square matrix formed by keeping the first k rows and columns of A.

Example 4.1:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}.$$

A principal minor is obtained by deleting the same rows and columns. For example, if we delete the 2nd and 4th rows and columns, we get the submatrix

$$\begin{bmatrix} 1 & 3 \\ 9 & 11 \end{bmatrix},$$

whose determinant is 74. Thus, 74 is one of the principal minors of A. The 3^{rd} leading principal minor of A is the determinant of

$$\begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{bmatrix}.$$

Theorem 4.1 (Sylvester's criterion): Let M be an $n \times n$ symmetric matrix. Write Δ_k as M's k^{th} leading principal minor.

(i) M is positive-definite if and only if $\Delta_k > 0$ for all $1 \leq k \leq n$.

(ii) M is negative-definite if and only if $(-1)^k \Delta_k > 0$ for all $1 \leq k \leq n$.

(iii) M is indefinite if and only if the first Δ_k that breaks both pattern above is nonzero.

(iv) Sylvester criterion is inconclusive if the first Δ_k that breaks both pattern is zero.