MAS242 해석학 II
Notes

한승우

September 19, 2023

Contents

Chapter 1

Differentiation

1.1 Higher order partial derivatives

Definition 1.1.1

Given $f : U \to \mathbb{R}$ where U is an open set in $\mathbb{R}^m$, define $\partial_{ij} \triangleq \partial_i(\partial_j f)(x)$ for each $i, j \in [m]$ to be *2nd order partial derivatives*. Any higher order partial derivatives can be defined inductively.

Definition 1.1.2: C^k-regularity

$f : U \to \mathbb{R}$ is C^k-regular if all partial derivatives up to order k and they are continuous.

Theorem 1.1.1

$f : U(\subseteq \mathbb{R}^2) \to \mathbb{R}$ is C^2 at a point $c \in U$, i.e., $\exists \delta > 0$, f is C^2 in $B_\delta(c)$. Then, $\partial_{12} f(c) = \partial_{21} f(c)$.

Proof. Let $|h| < \delta$. Define $A(h) \triangleq f(c_1 + h_1, c_2 + h_2) - f(c_1 + h_1, c_2) - f(c_1, c_2 + h_2) + f(c_1, c_2)$. Define $u(x_1) \triangleq f(x_1, c_2 + h_2) - f(x_1, c_2)$ and $v(x_2) \triangleq f(c_1 + h_1, x_2) - f(c_1, x_2)$. Note that u and v are differentiable.

Then, $A(h) = u(c_1 + h_1) - u(c_1)$ and $A(h) = v(c_2 + h) - v(c_2)$. By MVT, $\exists c_1^* \in (c_1, c_1 + h_1)$ and $c_2^* \in (c_2, c_2 + h_2)$ s.t. $A(h) = u'(c_1^*)h_1 = h_1\big(\partial_1 f(c_1^*, c_2 + h) - \partial_1 f(c_1^*, c_2)\big) = h_1 h_2 \partial_{21} f(c_1^*, c_2^*)$

Similarly, $\exists c_1^{**}, c_2^{**}$ such that $A(h) = h_1 h_2 \partial_{12} f(c_1^{**}, c_2^{**})$. $\partial_{21} f(c_1^*, c_2^*) = \partial_{12} f(c_1^{**}, c_2^{**})$. Hence, as $|h| \to 0$, due to the continuity, $\partial_{21}(c) = \partial_{12}(c)$. $\square$

Corollary 1.1.1

Suppose $f : U(\subseteq \mathbb{R}^m) \to \mathbb{R}$ is C^k at $c \in U$. Then $\partial_{j_1 j_2 \cdots j_k} f(c) = \partial_{j_1' j_2' \cdots j_k'}$ where $j_1' \cdots$ are a permutation of $j_1 \cdots$.

1.2 Extreme Values of differentiable Functions

Definition 1.2.1: Hessian

Let $f : U(\subseteq \mathbb{R}^m) \to \mathbb{R}$ be C_2 in U. Suppose $p \in U$ is a critical point of f, i.e., $\nabla f(p) = 0$. Define

$$\mathcal{H}f(x) \triangleq \begin{pmatrix} \partial_{11}f(x) & \partial_{21}f(x) & \cdots & \partial_{m1}f(x) \\ \partial_{12}f(x) & \partial_{22}f(x) & \cdots & \partial_{m2}f(x) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{1m}f(x) & \partial_{2m}f(x) & \cdots & \partial_{mm}f(x) \end{pmatrix}.$$

(Sometimes $\mathcal{H}f(x) = D^2 f(x)$.)
Define $D(x) = \det \mathcal{H}f(x)$. (Note that $\mathcal{H}f(x)$ is symmetric when f is C^2 by the theorem above.)

Theorem 1.2.1 2nd-order derivative test for two variable functions.

When $m = 2$ and f is C^2, a critical point p is
- a local maximum if $D(p) > 0$ and $\partial_{11}f(p) > 0$ (or $\partial_{22}f(p) > 0$).
- a local minimum if $D(p) > 0$ and $\partial_{11}f(p) < 0$ (or $\partial_{22}f(p) < 0$).
- a saddle point if $D(p) < 0$.

The test fails when $D(p) = 0$.

Proof. Given a unit vector $\mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$, $D_{\mathbf{u}} f = \nabla f \cdot \mathbf{u} = u_1 \partial_1 f + u_2 \partial_2 f$, and thus

$$D^2_{\mathbf{u}} f = (u_1\partial_1 + u_2\partial_2)(u_1\partial_1 f + u_2\partial_2 f) = u_1^2 \partial_{11} f + u_1 u_2 (2\partial_{12} f) + u_2^2 \partial_{22} f.$$

WLOG, $u_1 \neq 0$. Set $z = u_2/u_1$. Then,

$$D^2_{\mathbf{u}} f(p) = u_1^2\big(\partial_{11}f(p) + 2\partial_{12}f(p)z + \partial_{22}f(p)z^2\big).$$

Note that, if $D(p) > 0$, $D^2_{\mathbf{u}} f(p)$ has no real root.
- If $D(p) > 0$ and $\partial_{11}f(p) < 0$, Then, $D^2\mathbf{u} < 0$ for all unit vector $\mathbf{u}$.
- If $D(p) > 0$ and $\partial_{11}f(p) > 0$, Then, $D^2\mathbf{u} > 0$ for all unit vector $\mathbf{u}$.
- If $D(p) < 0$, $D^2_{\mathbf{u}} f(p)$ has different signs depending on $\mathbf{u}$.

For general m?

$$D_{\mathbf{u}}(D_{\mathbf{u}} f) = D_{\mathbf{u}} \sum_{j=1}^{m} \partial_j f u_j = \sum_{j=1}^{m} \big((\nabla \partial_j f) \cdot \mathbf{u}\big) u_j = \sum_{j=1}^{m}\sum_{k=1}^{m} u_k u_j \partial_{kj} f.$$

Hence,

$$D^2_{\mathbf{u}} f(p) = \mathbf{u}^{\mathrm{T}} \cdot D^2 f(p) \cdot \mathbf{u}$$

Since $D^2 f(p)$ is symmetric, its eigenvalues $\lambda_1, \cdots, \lambda_m$ exists and they are real numbers. Also, there exists an $m \times m$ orthogonal matrix $\mathcal{O}$ such that $D^2 f(p) = \mathcal{O}\Lambda(p)\mathcal{O}^{\mathrm{T}}$ where $\Lambda(p)$ is the diagonal matrix with entries are the eigenvalues.

Then, we can write $D^2_{\mathbf{u}} f(p) = \mathbf{u}\mathcal{O}\Lambda(p)\mathcal{O}^{\mathrm{T}}\mathbf{u}^{\mathrm{T}} = (\mathbf{u}\mathcal{O})\Lambda(p) = (\mathbf{u}\mathcal{O})^{\mathrm{T}}$. Since $\mathcal{O}$ is orthogonal, $\mathbf{u}\mathcal{O}$ is another arbitrary unit vector. □

Theorem 1.2.2 Generalized 2nd order partial derivatives test

When f is C^2, a critical point p is
- a local maximum if all eigenvalues of $D^2 f(p)$ are negative.

- a local minimum if all eigenvalues of $D^2 f(p)$ are positive.
- a saddle point if there are both negative eigenvalues and positive eigenvalues.

The test fails when there are zero eigenvalues.

Chapter 2

Inverse Function Theorem

2.1 Jacobian

Definition 2.1.1: Jacobian

Let $\mathbf{f}\colon U(\subseteq \mathbb{R}^m) \to \mathbb{R}^n$ be differentiable. The function $J_f : U \to \mathbb{R}$ defined by

$$J_{\mathbf{f}}(\mathbf{x}) = \det \begin{bmatrix} \partial_1 f_1(\mathbf{x}) & \cdots & \partial_n f_1(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \partial_1 f_n(\mathbf{x}) & \cdots & \partial_n f_n(\mathbf{x}) \end{bmatrix}$$

is called the *Jacobian* of $\mathbf{f}$ at $\mathbf{x}$.

Lemma 2.1.1

If $f : V(\subseteq \mathbb{R}^n) - \mathbb{R}$ and $\mathbf{g}\colon U \to V$ are differentiable, then

$$J_{f\circ \mathbf{g}}(\mathbf{x}) = J_f\big(\mathbf{g}(\mathbf{x})\big) \cdot J_{\mathbf{g}}(\mathbf{x}).$$

Note:-

The linear mapping $\mathrm{d}\mathbf{f}(c)$ is invertible if and only if $J_{\mathbf{f}}(c)$ is nonzero.

2.2 The Inverse Function Theorem

Lemma 2.2.1 Contraction Mapping Principle

Let (X, d) be a complete metric space. Let $\varphi\colon X \to X$. Suppose that there exists $M \in [0,1)$ such that $d\big(\varphi(x_1), \varphi(x_2)\big) \le M d(x_1, x_2)$. (We call it a *contraction mapping*.) Then, there uniquely exists $x_* \in X$ such that $\varphi(x_*) = x_*$.

Proof. Fix any $x_0 \in X$. Since $\{x_j\}_{j\in\mathbb{Z}_+}$, where $x_j = \varphi(x_{j-1})$ for each $j \in \mathbb{Z}_+$, is continuous. It converges to some x_*. As φ is continuous, we have $\varphi(x_*) = x_*$. The uniqueness follows trivially. □

Note:-

- For each $v \in \mathbb{R}^n \setminus \{0\}$, $|Av| = |v| \cdot \left|A\frac{v}{|v|}\right| \le \|A\|_L \cdot |v|$. The result is trivial when $v = 0$.
- For each $u \in \mathbb{R}^n$ with $|u| = 1$, $|ABu| \le \|A\|_L |Bu| \le \|A\|_L \|B\|_L$. Hence, $\|AB\|_L = \|A\|\|B\|$.
- Given invertible $A \in L(\mathbb{R}^n.\mathbb{R}^n)$, $A^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is linear. Moreover, $\|A\|_L > 0$.

Lemma 2.2.2

Given two linear mappings $A, B : \mathbb{R}^n \to \mathbb{R}^n$ with invertibility of A,

$$\|A - B\|_L \|A^{-1}\|_L < 1 \implies B \text{ is invertible.}$$

Proof. (Hint: show that $B\mathbf{x} = 0$ has only the trivial solution, i.e., if $\mathbf{x} \neq 0$, then $B\mathbf{x} \neq 0$.) □

Theorem 2.2.1 Inverse Function Theorem

Let $\mathbf{f} : U(\subseteq \mathbb{R}^n) \to \mathbb{R}^n$ be C^1 in U, $\mathbf{a} \in U$, and $\mathbf{b} = f(a)$. Suppose that $J_{\mathbf{f}}(a) \neq 0$. Then,

$$\exists \delta \in \mathbb{R}_+,\ \mathbf{f}\big|_{B_\delta(a)} : B_\delta(a) \to \mathbf{f}\big(B_\delta(a)\big) \text{ is invertible.}$$

Moreover, $\mathbf{f}\big(B_\delta(a)\big)$ is an open set, and $\left(\mathbf{f}\big|_{B_\delta(a)}\right)^{-1}$ is C^1.

Proof. Let $A \triangleq \mathrm{d}\mathbf{f}(\mathbf{c})$. Define λ by $\lambda \triangleq \dfrac{1}{2\|A^{-1}\|_L} > 0$ so $2\lambda \left\|A^{-1}\right\|_L = 1$. Since $\mathrm{d}\mathbf{f}$ is continuous, there exists $\delta \in \mathbb{R}_+$ such that $\|\mathrm{d}\mathbf{f}(\mathbf{x}) - \mathrm{d}\mathbf{f}(\mathbf{c})\|_L < \lambda$ for each $B_\delta(\mathbf{c})$.

Given a point $\mathbf{y} \in \mathbb{R}^n$, we define $\varphi(\cdot; \mathbf{y})$ by

$$\begin{aligned} \varphi(\cdot; \mathbf{y}) : B_\delta(\mathbf{c}) &\longrightarrow \mathbb{R}^n \\ \mathbf{x} &\longmapsto \varphi(\mathbf{x}; \mathbf{y}) = \mathbf{x} + A^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{x})) \end{aligned}$$

Note that $\mathbf{x}$ is a fixed point of $\varphi(\cdot; \mathbf{y})$ if and only if $A^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{x})) = 0$, i.e., $\mathbf{y} = \mathbf{f}(\mathbf{x})$. Note also that φ is differentiable and $\mathrm{d}\varphi(\mathbf{x}; \mathbf{y}) = \mathrm{Id} - A^{-1}\,\mathrm{d}\mathbf{f}(\mathbf{x}) = A^{-1}\big(A - \mathrm{d}\mathbf{f}(\mathbf{x})\big)$ for each $x \in B_\delta(\mathbf{c})$.

Hence, for all $x \in B_\delta(\mathbf{c})$,

$$\|\mathrm{d}\varphi(\mathbf{x}; \mathbf{y})\|_L = \left\|A^{-1}\big(A - \mathrm{d}\mathbf{f}(\mathbf{x})\big)\right\|_L \leq \|A^{-1}\|_L \cdot \|A - \mathrm{d}\mathbf{f}(\mathbf{x})\|_L < 1/(2\lambda) \cdot \lambda = 1/2.$$

Now, fix any $\mathbf{y} \in \mathbb{R}^n$. Fix $\mathbf{x}_1, \mathbf{x}_2 \in B_\delta(\mathbf{c})$. Define $\Psi : [0,1] \to \mathbb{R}$ by $t \mapsto \varphi(t\mathbf{x}_1 + (1-t)\mathbf{x}_2; \mathbf{y})$. $\Psi(0) = \varphi(\mathbf{x}_2; \mathbf{y})$ and $\Psi(1) = \varphi(\mathbf{x}_1; \mathbf{y})$. Note that Ψ is differentiable on $(0,1)$. By MVT, there exists $t_* \in (0,1)$ such that $\Psi(1) - \Psi(0) = \Psi'(t_*)$. The chain rule gives

$$\Psi'(t_*) = \mathrm{d}\varphi(t_*\mathbf{x}_1 + (1-t_*)\mathbf{x}_2)(\mathbf{x}_1 - \mathbf{x}_2)$$

. Hence,

$$|\varphi(\mathbf{x}_1) - \varphi(\mathbf{x}_2)| = |\mathrm{d}\Psi(t_*\mathbf{x}_1 + (1-t_*)\mathbf{x}_2)| \cdot |(\mathbf{x}_1 - \mathbf{x}_2)| < |\mathbf{x}_1 - \mathbf{x}_2|/2.$$

□

End.