Review on Analysis

김지수 (Jisu KIM)

통계적 기계학습 (Statistical Machine Learning), 2026년 1학기 (spring)

Calculus

Taylor Theorem

1. For $\alpha = (\alpha_1, \ldots, \alpha_k)$, $x = (x_1, \ldots, x_k)$, we define the notations

$$|\alpha| := \alpha_1 + \ldots + \alpha_k$$
$$\alpha! := \alpha_1! \times \ldots \times \alpha_k!$$
$$x^\alpha := x_1^{\alpha_1} \ldots x_k^{\alpha_k}.$$

2. (Clairaut's theorem) If f's all k-th partial derivatives are continuous at a, then we can change the order of derivatives. So we can define the notation below:

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_p^{\alpha_p}}, \quad |\alpha| \le k.$$

Theorem. *Multivariate Taylor Expansion (다변량 테일러 전개)*

If a function $f : \mathbb{R}^k \longrightarrow \mathbb{R}$ has continuous rth partial derivatives on an open set B containing a, then for $x \in B$, there exists $\xi \in (a, x) := \{\lambda a + (1-\lambda)x : \lambda \in (0,1)\}$ such that

$$f(x) = \sum_{|\alpha| \le r-1} \frac{D^\alpha f(a)}{\alpha!}(x-a)^\alpha + \sum_{|\alpha| = r} \frac{D^\alpha f(\xi)}{\alpha!}(x-a)^\alpha.$$

We can alternatively express as follows: there exist functions h_α's so that

$$f(x) = \sum_{|\alpha| \le r} \frac{D^\alpha f(a)}{\alpha!}(x-a)^\alpha + \sum_{|\alpha| = r} h_\alpha(x)(x-a)^\alpha, \ \lim_{x \to a} h_\alpha(x) = 0, \forall \alpha : |\alpha| = r.$$

A probability space is a triple $(\Omega, \mathcal{F}, P)$ where Ω is a set of "outcomes," $\mathcal{F}$ is a set of "events," and $P : \mathcal{F} \to [0,1]$ is a function that assigns probabilities to events.

Fundamental Theorem of Calculus

Theorem. *If f is integrable on $[a, b]$, then the function $F : [a, b] \to \mathbb{R}$ defined as*

$$F(x) = \int_a^x f(y)dy,$$

satisfies that

(a) F is absolutely continuous on $[a, b]$.

(b) F is differentiable a.e. on (a, b), and

$$\frac{d}{dx}F(x) = f(x) \qquad \text{for a.e. } x,$$

(c) In particular, when f is continuous, then

$$\frac{d}{dx}F(x) = f(x) \qquad \text{for all } x \in (a, b).$$

Theorem. *Suppose $f : \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}$ is such that f is integrable. Fix $x_0 \in \mathbb{R}^d$, and suppose $\frac{\partial f}{\partial x}$ exists at $x \in B(x_0, \epsilon) := \left\{x \in \mathbb{R}^d : \|x - x_0\| < \epsilon\right\}$ for small neighborhood of x_0. Suppose there exists $g : \mathbb{R}^k \to \mathbb{R}$ such that $\left|\frac{\partial f}{\partial x}(x, t)\right| \le g(t)$ for all t, and g is integrable. Then,*

$$\frac{d}{dx}\int f(x_0, t)dt = \int \frac{\partial}{\partial x}f(x_0, t)dt.$$

Convexity

Convex set

S : convex set in $\mathbb{R}$-vector space (e.g., $\mathbb{R}^d$)

$$\Leftrightarrow \forall x, y \in S, \{\lambda x + (1 - \lambda)y \in S \text{ for any } \lambda : 0 \le \lambda \le 1\}$$

Convex set in $\mathbb{R}$:
S : convex set in $\mathbb{R} \Leftrightarrow S$: an interval
By an interval, we mean

$$(a, b), [a, b), (a, b] \text{ or } [a, b], -\infty \le a \le b \le +\infty$$

Convex function

Let f be a real-valued function defined on a convex set S.

$$\begin{aligned} &f : \text{ convex} \\ &\Leftrightarrow \forall x, y \in S, \forall \lambda : 0 \le \lambda \le 1, f(\lambda x + (1 - \lambda)y) \le \lambda f(x) + (1 - \lambda)f(y) \end{aligned}$$

f : strictly convex

$$\Leftrightarrow \forall x, y \in S, \forall \lambda : 0 < \lambda < 1, f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

Concave function

$$f : \text{ (strictly) concave } \Leftrightarrow -f : \text{ (strictly) convex}$$

Tools to show convexity

(1) f : real-valued on (a, b) with $\exists f''$ on $(a, b), -\infty \le a < b \le +\infty$.

f : convex $\Leftrightarrow f''(x) \geq 0 \quad \forall x \in (a, b)$
f : strictly convex $\Leftrightarrow f''(x) > 0 \quad \forall x \in (a, b)$
(2) f : real-valued on an open convex set $S \subset \mathbb{R}^d$ with $\exists \nabla^2 f$ on S
f : convex $\Leftrightarrow \nabla^2 f$: non-negative definite on S

$$\text{i.e. } a^\top \nabla^2 f(x) a \geq 0, \forall a \forall x \in S$$

f : strictly convex $\Leftrightarrow \nabla^2 f$: positive definite on S

$$\text{i.e. } a^\top \nabla^2 f(x) a > 0, \forall a \neq 0 \forall x \in S$$

(3) f : convex on a convex set S with $f(S)$ being convex.
g : convex and non-decreasing on $f(S)$
$\Rightarrow g \circ f$: convex on S

First and second order characterizations of convex functions

Suppose $f : S \subset \mathbb{R}^d \to \mathbb{R}$ is twice differentiable over an open set int$S$(=the largest open subset of S). Then, the following are equivalent:
(i) f is convex.
(ii) $f(y) \geq f(x) + \nabla f(x)^\top (y - x)$, for all $x, y \in \text{int}S$.
(iii) $\nabla^2 f(x) \succeq 0$, for all $x \in \text{int}S$.

First and second order characterizations of strictly convex functions

Recall that a fuction $f : S \to \mathbb{R}$ is strictly convex if $\forall x, y, x \neq y, \forall \lambda \in (0, 1)$,

$$f(\lambda x + (1 - \lambda) y) < \lambda f(x) + (1 - \lambda) f(y).$$

Like we mentioned before, if f is strictly convex, then f is convex (this is obvious from the definition) but the converse is not true (e.g., $f(x) = x, x \in \mathbb{R}$).

Second order sufficient condition:

$$\nabla^2 f(x) \succ 0, \forall x \in \text{int}S \Rightarrow f \text{ strictly convex on } S$$

The converse is not true though (why?).

First order characterization: A function f is strictly convex on $S \subseteq \mathbb{R}^n$ if and only if

$$f(y) > f(x) + \nabla f(x)^\top (y - x), \forall x, y \in \text{int}S, x \neq y$$

References