CMO 2: Taylor Series

Eklavya Sharma

Contents

1 Univariate Taylor Series

Let $f : [a, b] \mapsto \mathbb{R}$. Let $x, y \in [a, b]$.

Suppose f is differentiable k times. Then for some $z \in (x, y)$,

$$f(y) = \sum_{i=0}^{k-1} f^{(i)}(x) \frac{(y-x)^i}{i!} + f^{(k)}(z) \frac{(y-x)^k}{k!}$$

C^k is the set of all functions which are k-times differentiable and whose k^{th} derivative is continuous.

When $f^{(k)} \in C^k$,

$$f(y) = \sum_{i=0}^{k} f^{(i)}(x) \frac{(y-x)^i}{i!} + o(1) \frac{(y-x)^k}{k!}$$

Therefore, we can ignore the last term if x is close to y.

2 Multivariate Calculus

Definition 1. *Let $f : \mathbb{R}^m \mapsto \mathbb{R}^n$ be a function and $y = f(x)$. Then the Jacobian of y w.r.t. x is an n by m matrix where*

$$\left(\frac{\partial y}{\partial x}\right)_{i,j} = \frac{\partial y_i}{\partial x_j}$$

Theorem 1 (Chain rule)**.** *Let $y = f(x)$ and $z = g(y)$. Then*

$$\frac{\partial z}{\partial x} = \left(\frac{\partial z}{\partial y}\right)\left(\frac{\partial y}{\partial x}\right)$$

Definition 2. *For $f : \mathbb{R}^d \mapsto \mathbb{R}$, the gradient of f, denoted as ∇_f, is a d-dimensional vector defined as*

$$\nabla_f(x) = \left[\frac{\partial f(x)}{\partial x_i}\right]_{i=1}^{d}$$

For multivariate functions, $f \in C^1$ iff ∇_f exists and all components are continuous. Note that differentiability does not imply C_1.

Definition 3. *For $f : \mathbb{R}^d \mapsto \mathbb{R}$, the hessian of f, denoted as H_f, is a d by d matrix defined as*

$$\mathrm{H}_f(x)_{i,j} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

For multivariate function, $f \in C^2$ iff H_f exists and all its entries are continuous.

Theorem 2 (Proof omitted)**.** *When $f \in C^2$, H_f is symmetric.*

3 Multivariate Taylor Series

Let $g(t) = f(x + tu)$, where $t \in \mathbb{R}$ and $x, u \in \mathbb{R}^d$.

Theorem 3.

$$g'(t) = \nabla_f(x + tu)^T u \qquad \text{(when } f \in C^1\text{, by chain rule)}$$

$$g''(t) = u^T \mathrm{H}_f(x + tu)u \qquad \text{(when } f \in C^2\text{)}$$

Theorem 4. $f \in C^1 \implies g \in C^1$

Theorem 5. *If $f \in C^1$ and y is close to x,*

$$f(y) = f(x) + \nabla_f(x)^T(y - x) + o(\|y - x\|)$$

Proof. Let $g(t) = f(x + tu)$ and let $u = y - x$ be small. By applying univariate Taylor series on g at 0, we get

$$g(1) = g(0) + g'(\alpha), \text{ where } \alpha \in [0, 1]$$
$$\Rightarrow f(x + u) = f(x) + \nabla_f(x + \alpha u)^T u$$
$$\Rightarrow f(x + u) = f(x) + (\nabla_f(x) + o(1))^T u \qquad (\nabla_f \text{ is continuous and } u \text{ is small})$$
$$\Rightarrow f(y) = f(x) + \nabla_f(x)^T(y - x) + o(\|y - x\|)$$

□

Theorem 6. *If $f \in C^2$ and y is close to x,*

$$f(y) = f(x) + \nabla_f(x)^T(y - x) + \frac{1}{2}(y - x)^T \mathrm{H}_f(x)(y - x) + o(\|y - x\|^2)$$

Proof. Similar to previous theorem. □