Miscellaneous Calculus

Andrew Nobel

September, 2021

Taylor’s Theorem in One Dimension

Taylor's Theorem, First Order

Fact 1: Let $g : \mathbb{R} \to \mathbb{R}$ have continuous derivative g'. Then for each x and h

$$g(x+h) = g(x) + hg'(\tilde{x})$$

where $\tilde{x}$ lies between x and $x+h$. As a consequence, if h is small

$$g(x+h) \approx g(x) + hg'(x)$$

Idea: Linear approximation of g in a neighborhood of the point x

Taylor's Theorem, Second Order

Fact 2: If g has two continuous derivatives g' and g'', then for each x and h

$$g(x+h) = g(x) + hg'(x) + h^2 g''(\tilde{x})/2$$

where $\tilde{x}$ lies between x and $x+h$. As a consequence, if h is small

$$g(x+h) \approx g(x) + hg'(x) + h^2 g''(x)/2$$

Idea: Quadratic approximation of g in a neighborhood of the point x

Taylor's Theorem, General Version

Fact: If $g : \mathbb{R} \to \mathbb{R}$ has derivatives $g^{(1)}, \ldots, g^{(k)}$ then for each $x, h \in \mathbb{R}$

$$g(x+h) \; = \; \sum_{j=0}^{k-1} \frac{g^{(j)}(x)}{j!} h^j \; + \; \frac{g^{(k)}(\tilde{x})}{k!} h^k$$

where $g^{(0)} = g$ and $\tilde{x}$ lies between x and $x + h$. If $g^{(k)}$ is continuous then the final "remainder" term is of smaller order than h^k

- Theorem extends to functions g defined on an interval in $\mathbb{R}$
- Result is a consequence of mean value theorem from calculus

Inequalities from Calculus

Increasing and Decreasing Functions

Recall: Let $f, g : [a, b] \to \mathbb{R}$ be functions

- f is *increasing* if $x \leq y$ implies that $f(x) \leq f(y)$
- f is *decreasing* if $x \leq y$ implies that $f(x) \geq f(y)$
- $f \leq g$ if $g(x) - f(x) \geq 0$ for every x

Fact: Suppose that $f : [a, b] \to \mathbb{R}$ is differentiable

- f is increasing if $f'(x) \geq 0$ for every x
- f is decreasing if $f'(x) \leq 0$ for every x
- f is constant if $f'(x) = 0$ for every x

Inequalities Using Derivatives

Common Problem

- Differentiable functions $f, g : I \to \mathbb{R}$ defined on interval $I \subseteq \mathbb{R}$
- Wish to show that $f(x) \geq g(x)$ for each $x \in I$

Method 1: Define difference $h(x) := f(x) - g(x)$ Find $c \in I$, possibly an endpoint of I, such that $h(c) \geq 0$. Then show

$$h'(x) \text{ is } \begin{cases} \geq 0 & \text{if } x \geq c \\ \leq 0 & \text{if } x \leq c \end{cases}$$

Method 2: Expand h in a kth order Taylor series around $c \in I$, usually where $h^{(0)}(c) = \cdots = h^{(k-1)}(c) = 0$, and then examine the remainder term

Inequalities from Calculus: Examples

Ex 1. Show that $h(x) = x/(1+x)$ is increasing for $x \geq 0$

Ex 2. For all x, $1 + x \;\leq\; e^x$

Ex 3. For $0 \leq x < 1$, $\log(1-x) \;\leq\; -x - x^2/2$

Ex 4. For any $a_1, \ldots, a_n \in \mathbb{R}$ we have the inequality

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{a_i \, a_j}{i+j} \right) \;\geq\; 0$$

Example: Birthday Problem

Qu: How likely is it that m unrelated people have different birthdays?

- Let $X_i \in \{1, \dots, 365\}$ be birthday of ith person

Fact: If $X_1, \dots, X_m$ are independent and uniformly distributed then

$$\mathbb{P}(X_1, \dots, X_m \text{ distinct}) \leq \exp\left\{-\frac{(m-1)m}{730}\right\}$$

Cor: If $m \geq 24$ then probability two people have the same birthday $\geq 1/2$

Gradients and Hessians

Gradients and Hessians

Definition: Let $f : \mathbb{R}^d \to \mathbb{R}$ have well defined first and second order partial derivatives. The *gradient* of f is the vector of first partial derivatives

$$\nabla f(\mathbf{x}) \;=\; \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \ldots, \frac{\partial f}{\partial x_d}(\mathbf{x}) \right)^t \in \mathbb{R}^d$$

The *Hessian* of f is the matrix of second partial derivatives

$$H(\mathbf{x}) = \nabla^2 f(\mathbf{x}) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) : 1 \leq i, j, \leq d \right] \in \mathbb{R}^{d \times d}$$

Note: If the second partials $\frac{\partial^2 f}{\partial x_i \partial x_j}$ are continuous then $\nabla^2 f$ is symmetric

Examples

Ex 1: Linear $f(\mathbf{x}) = \langle \mathbf{u}, \mathbf{x} \rangle + b$.

- $\nabla f(\mathbf{x}) = \mathbf{u}$
- $\nabla^2 f(\mathbf{x}) = \mathbf{0}$

Ex 2: Quadratic $f(\mathbf{x}) = \mathbf{x}^t \mathbf{A} \mathbf{x}$ with $\mathbf{A} \in \mathbb{R}^{d \times d}$ symmetric

- $\nabla f(\mathbf{x}) = 2\mathbf{A}\mathbf{x}$
- $\nabla^2 f(\mathbf{x}) = 2\mathbf{A}$

Multivariate Taylor's Theorem

Fact: If $f : \mathbb{R}^d \to \mathbb{R}$ has continuous second partial derivatives $\partial^2 f / \partial x_i \partial x_j$ at each point in $\mathbb{R}^d$ then for every $x, h \in \mathbb{R}^d$

$$f(x+h) \,=\, f(x) \,+\, \langle \nabla f(x), h \rangle \,+\, \frac{1}{2} h^t \nabla^2 f(\tilde{x}) h$$

where $\tilde{x} = x + \alpha h$ for some $\alpha \in [0, 1]$. In particular, we have

$$f(x+h) \,=\, f(x) \,+\, \langle \nabla f(x), h \rangle \,+\, \frac{1}{2} h^t \nabla^2 f(x) h \,+\, o(||h||^2)$$

Stirling’s Approximation

The Gamma Function and Stirling's Approximation

Gamma function: Defined by $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx$ for $\alpha > 0$

(1) $\Gamma(1) = 1$ and $\Gamma(1/2) = \sqrt{\pi}$

(2) $\Gamma(\alpha+1) = \alpha\,\Gamma(\alpha)$

(3) $\Gamma(n) = (n-1)!$

(4) $\log\Gamma(\alpha)$ is convex.

***Stirling's Approximation**

(1) $\Gamma(\alpha+1) \sim \sqrt{2\pi\alpha}\left(\frac{\alpha}{e}\right)^{\alpha}$

(2) $n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^{n} \delta_n$ where $e^{1/(12n+1)} \le \delta_n \le e^{1/(12n)}$