CAS CS 365
Lec 21

Vector Calculus (cont.)

1. Gradient of a scalar field
    a. Partial derivative at x = (x_1, …, x_n)
    b. $\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f(x_{1,} \ldots, x_i + h, \ldots, x_n) - f(x_{1,} \ldots, x_i, \ldots, x_n)}{h}, \, i = 1, \ldots, n$
    c. We collect them at the row vector known as the gradient of the function f
    d. $\nabla f(x) = \nabla_x f = \mathrm{grad} f = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \ldots & \frac{\partial f}{\partial x_n} \end{pmatrix} \in \mathbb{R}^{1 \times n}$
    e. Remark: the gradient collects the slopes in the positive x_i direction for all i = 1…n
2. Directional derivative
    a. Instead of computing the slopes in the positive x_i directions for all i=1…n, we can compute the derivative along any direction
    b. Directional derivative
    c. $\nabla_v f(x) = D_v f(x) = \lim_{h \to 0} \frac{f(x + hv) - f(x)}{h} = \nabla f(x) \cdot v$
3. Hessian of a scalar field
    a. If all second partial derivatives of f exist and are continuous over the domain of the function, then the Hessian matrix is a square matrix, usually defined and arranged as follows:

$$\mathbf{H}_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \, \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \, \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \, \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \, \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \, \partial x_1} & \frac{\partial^2 f}{\partial x_n \, \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

4. Example
    a. Compute the Hessian of f(x,y) = xy*(x+y) at (1,1)

$$H_f(x, y) = \begin{pmatrix} 2y & 2(x+y) \\ 2(x+y) & 2x \end{pmatrix}, \, H_f(1, 1) = \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix}$$

    b. The symmetry of H is not coincidence; of (x,y) is a twice continuously differentiable function, then $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$

Lec 21

$\curvearrowright x^T x = x_1^2 + \dots + x_n^2$

$f(x) = \|x\|_2^2 \qquad f: R^n \to R$

Gradient

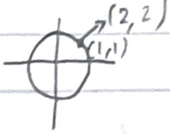
$\nabla f(x)$

↳ what are the dimensions $\to 1 \times n$

$\nabla f(x) = \left(\frac{df}{dx_1}, \dots, \frac{df}{dx_n} \right)$

$= (2x_1, \dots, 2x_n) = 2x^T$

↳ the bigger the x, greater changes



→ Vector that points to the direction of maximum increase of f

$\lim_{h \to 0} \frac{f(\vec{x} + h\vec{u}) - f(\vec{x})}{h}$ → directional derivative → it is a real number, not vector

$\vec{u} \in R^n, \|u\|_2 = 1$

$= \nabla f(x) \cdot v$

↳

ex) $v = (-1/\sqrt{2}, 1/\sqrt{2})$
$(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$] ~~orthogonal to (2,2)~~

$\nabla(x^T x) = 2x^T$

Evaluate at $(1,1)$ $\nabla f(x) \Big|_{(1,1)} = \nabla f(1,1)$
$= (2,2)$

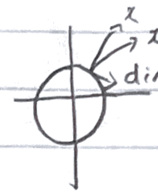
orthogonal vector

$D_{\left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)} f(1,1) = (2,2) \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = 0$

This means that if we move in direction pointed by vector $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$, this means that we don't change much.

$x' \begin{pmatrix} 1 \\ 1 \end{pmatrix} + h \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \qquad \lim_{h \to 0} \frac{f(x') - f(x)}{h}$

$$f(x') = \|x'\|_2^2$$
$$= \left(1+\frac{h}{\sqrt{2}}\right)^2 + \left(1-\frac{h}{\sqrt{2}}\right)^2$$



direction denoted by u
↳ projecting gradient on the direction

$$= 1 + \frac{h^2}{2} + 1 + \frac{h^2}{2}$$
$$= 2 + h^2$$

$$\lim_{h \to 0} \frac{f(x') - f(x)}{h} = \frac{2 + h^2 - 2}{h} = h = 0$$

Move little → orthogonal to gradient
Move maximum → parallel to gradient

* make the derivative unit vector)

ex) $f(x,y) = x^2 y$
gradient of f:
$\nabla f(x,y) = (2xy, x^2)$
gradient at (3,2)
$\nabla f(3,2) = (12, 9)$
The derivative of f in direction of (1,2) at point (3,2)

$$\nabla_{(1/\sqrt{5},\, 2/\sqrt{5})} f(2,3) = (12,9)\begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} = \frac{30}{\sqrt{5}}$$

$f(x) = w^T x \quad \nabla f(x) = w^T$ ← no need to prove
$f(x_1, x_2) = 3x_1 - 10x_2$
$w^T = (3, -10)$

$f(x) = w_1 x_1 + \dots + w_n x_n$

$$\nabla f(x)^{1\times n} = \left(\frac{df}{dx_1}, \dots, \frac{df}{dx_n}\right) = (w_1, \dots, w_n) = w^T$$

$w^T \cdot v = \|w\| \cdot \cos\theta$

$f: \mathbb{R}^n \to \mathbb{R}$

$\nabla^2 f(x) = H_f(x) \leftarrow$ Hessian $\rightarrow \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \partial^2 f / \partial x_i \partial x_j \\ & \ddots \\ & & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$

$\nabla f \in \mathbb{R}^{1 \times n}$

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

Ex) $f: \mathbb{R}^2 \to \mathbb{R}$

$f(x,y) = x^2 y$

$\nabla f(x,y) = (2xy, x^2)$

$$\frac{\partial^2 f}{\partial x^2} = 2y$$

$$\frac{\partial^2 f}{\partial x \partial y} = 2x = \frac{\partial^2 f}{\partial y \partial x} = 2x$$

$$\frac{\partial^2 f}{\partial y^2} = 0$$

$$H_f(x,y) = \begin{bmatrix} 2y & 2x \\ 2x & 0 \end{bmatrix}$$

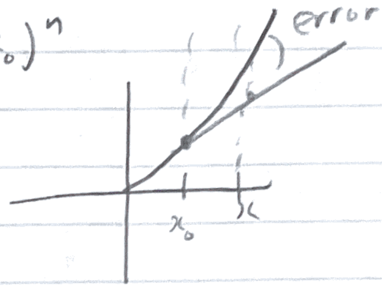
Taylor polynomial of deg n

$f: \mathbb{R} \to \mathbb{R}$ around (x_0)

$$T(x) = f(x_0) + f(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2$$

$$+ \dots + \frac{f^n(x_0)}{n!}(x - x_0)^n$$

$$= \sum_{k=0}^{\infty} \frac{f^k(x_0)}{k!}(x - x_0)^k$$



Taylor Approximation Theorem

Suppose at least $(n+1)$ differentiable in $[a, b]$

$$\forall x \in (a, b)\ T(x) = \sum_{k=0}^{n} \frac{f^k(a)}{k!}(x - a)^k$$

The error $|T(x) - f(x)|$ is at most

$$\max_{a \le x \le b} |f^{(n+1)}(x)|\, |x - a|^{n+1}$$

ex) $f(x)$ around 1, $f(x) = x^4$

$f'(x) = 4x^3 = 4$

$f''(x) = 12x^2 = 12$

$f'''(x) = 24x = 24$

$f''''(x) = 24$

$$T(x) = 1 + 4(x-1) + \frac{12}{2!}(x-1)^2$$

$$+ \frac{24}{3!}(x-1)^3 + \frac{24}{4!}(x-1)^4$$

$$= 1 + 4(x-1) + 6(x-1)^2$$

$$+ 4(x-1)^3 + (x-1)^4$$

ex) $f(x_1, x_2) = x_1^4 + x_2^4 + 3x_1 x_2^3 + e^{x_1 + e^{x_2}}$

$$f(x_1, x_2) = f(1,2) + \nabla f(1,2) \begin{pmatrix} x_1 - 1 \\ x_2 - 2 \end{pmatrix}$$

$$+ (x_1 - 1, x_2 - 2) \, H_f \Big|_{(1,2)} \begin{pmatrix} x_1 - 1 \\ x_2 - 2 \end{pmatrix} + \dots$$

ex) $f(x,y) = x^2 + 2xy + y^3 \quad (1,2)$

$$\nabla f(x,y) = (2x + 2y, 2x + 3y^2)$$

$$H_f = \begin{pmatrix} 2 & 2 \\ 2 & 6y \end{pmatrix}$$

$$\nabla f(1,2) = (6, 14)$$

$$H_f(1,2) = \begin{bmatrix} 2 & 2 \\ 2 & 12 \end{bmatrix}$$

Taylor approximation

$$\hookrightarrow f(1,2) + \nabla f(1,2) \begin{pmatrix} x-1 \\ y-2 \end{pmatrix} + (x-1, y-2) \underbrace{\begin{pmatrix} 2 & 2 \\ 2 & 12 \end{pmatrix}}^{H_f(1,2)} \begin{pmatrix} x-1 \\ y-2 \end{pmatrix}$$

$$= 13 + (6, 14) \begin{pmatrix} x-1 \\ y-2 \end{pmatrix} + (x-1, y-2) \begin{bmatrix} 2 & 2 \\ 2 & 12 \end{bmatrix} \begin{bmatrix} x-1 \\ y-2 \end{bmatrix}$$

$$= 13 + 6(x-1) + 14(y-2) + [x-1 \quad y-2] \begin{bmatrix} 2x - 2 + 2y - 4 \\ 2x - 2 + 12y - 24 \end{bmatrix}$$

$$= 13 + 6(x-1) + 14(y-2) + (2x - 2 + 2y - 4)(x-1)$$
$$+ (y-2)(2x - 2 + 12y - 24)$$

$$T(x_0, y_0) = f(x_0, y_0) + \nabla f(x_0, y_0) \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + (x - x_0 \quad y - y_0) \, H_f(x_0, y_0) \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}$$

or

$$T(x) = f(p) + \nabla f(p)(x - p) + (x - p) \, H_f(p) \, (x - p)$$