01/12/2022

Week 2 Recitation

CS165B - Machine Learning

UC SANTA BARBARA

Overview

- Matrix calculus (differentials) examples
- Homework 1 Problem 4

Notations

- Vectors (single-column matrices) are denoted by boldfaced lowercase letters like **a**, **b**, **x.**
- Matrices are denoted by boldface uppercase letters **A**, **B**, **X.**

Numerator Layout

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \qquad \frac{\partial \mathbf{y}}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} \\ \frac{\partial y_2}{\partial x} \\ \vdots \\ \frac{\partial y_m}{\partial x} \end{bmatrix}$$

d**y**/dx is a column vector

Numerator Layout

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \frac{\partial y}{\partial \mathbf{x}} = \left[\frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \ldots, \frac{\partial y}{\partial x_n} \right]$$

dy/d**x** is a row vector

Numerator Layout

$$\mathbf{x} \in \mathbb{R}^n, \quad \mathbf{y} \in \mathbb{R}^m$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

Numerator Layout

$$\mathbf{x} \in \mathbb{R}^n, \quad \mathbf{y} \in \mathbb{R}^m, \quad \frac{\partial \mathbf{y}}{\partial \mathbf{x}} \in \mathbb{R}^{m \times n}$$

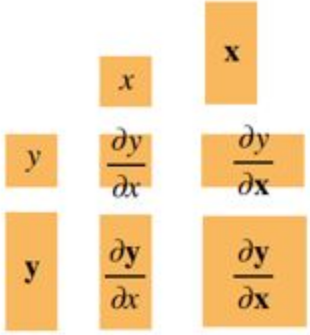
$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial \mathbf{x}} \\ \frac{\partial y_2}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial y_m}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1}, \frac{\partial y_1}{\partial x_2}, \ldots, \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1}, \frac{\partial y_2}{\partial x_2}, \ldots, \frac{\partial y_2}{\partial x_n} \\ \vdots \\ \frac{\partial y_m}{\partial x_1}, \frac{\partial y_m}{\partial x_2}, \ldots, \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

Numerator Layout

$$\mathbf{x} \in \mathbb{R}^n, \quad \mathbf{y} \in \mathbb{R}^m, \quad \frac{\partial \mathbf{y}}{\partial \mathbf{x}} \in \mathbb{R}^{m \times n}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$



$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial \mathbf{x}} \\ \frac{\partial y_2}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial y_m}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1}, \frac{\partial y_1}{\partial x_2}, \ldots, \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1}, \frac{\partial y_2}{\partial x_2}, \ldots, \frac{\partial y_2}{\partial x_n} \\ \vdots \\ \frac{\partial y_m}{\partial x_1}, \frac{\partial y_m}{\partial x_2}, \ldots, \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

Calculus

- Derivative of Sums

$$y = u + v$$
$$\frac{\partial y}{\partial x} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial x}$$

- Product Rule

$$\frac{d(uv)}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}$$

- Chain Rule

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

For more detailed calculus review refer to week 1's recitation slides (6-8pm),
or Matrix Cookbook Chapter 2

Example 1

$$y = x_1^2 + 2x_2^2$$

$$\frac{\partial y}{\partial \mathbf{x}}$$

Example 1

$$y = x_1^2 + 2x_2^2$$

$$\frac{\partial y}{\partial \mathbf{x}} = [2x_1, 4x_2]$$

Example 2

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$$

Example 2

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

$$y_i = \sum_{k=1}^{n} a_{ik} x_k$$

$$\frac{\partial y_i}{\partial x_j} = a_{ij}$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \mathbf{A}$$

Example 3

$$\mathbf{y} = \mathbf{x}^T \mathbf{A}$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$$

Example 3

$$\mathbf{y} = \mathbf{x}^T \mathbf{A}$$

$$y_i = \sum_{k=1}^{n} x_k a_{ki}$$

$$\frac{\partial y_i}{\partial x_j} = a_{ji}$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \mathbf{A}^T$$

Example 4

$$y = \mathbf{u}^T \mathbf{v}$$

$$\frac{\partial y}{\partial \mathbf{x}}$$

Example 4

$$y = \mathbf{u}^T \mathbf{v}$$

$$y = \sum_{i=1}^{n} u_i v_i$$

$$\frac{\partial y}{\partial x_j} = \sum_{i=1}^{n} \left(v_i \frac{\partial u_i}{\partial x_j} + u_i \frac{\partial v_i}{\partial x_j} \right)$$

$$\frac{\partial y}{\partial \mathbf{x}} = \mathbf{v}^T \frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \mathbf{u}^T \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$$

Example 4

$$y = \mathbf{u}^T \mathbf{v}$$

$$y = \sum_{i=1}^{n} u_i v_i$$

$$\frac{\partial y}{\partial x_j} = \sum_{i=1}^{n} \left(v_i \frac{\partial u_i}{\partial x_j} + u_i \frac{\partial v_i}{\partial x_j} \right)$$

$$\frac{\partial y}{\partial \mathbf{x}} = \mathbf{v}^T \frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \mathbf{u}^T \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$$

$$\frac{\partial y}{\partial \mathbf{x}} = \frac{\partial y}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \frac{\partial y}{\partial \mathbf{v}} \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$$

$$\frac{\partial y}{\partial \mathbf{x}} = \mathbf{v}^T \frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \mathbf{u}^T \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$$

Example 5

$$y = x^T A x$$

$$\frac{\partial y}{\partial x}$$

Example 5

$$y = x^T A x$$

$$b = Ax$$

$$y = x^T b$$

$$\frac{\partial y}{\partial x} = b^T \frac{\partial x}{\partial x} + x^T \frac{\partial b}{\partial x}$$

$$\frac{\partial y}{\partial x} = b^T + x^T A$$

$$\frac{\partial y}{\partial x} = x^T A^T + x^T A$$

$$\frac{\partial y}{\partial x} = x^T (A + A^T)$$

Example 6

Assume $\mathbf{x}, \mathbf{w} \in \mathbb{R}^n, \quad y \in \mathbb{R}$

$$z = \left(\langle \mathbf{x}, \mathbf{w} \rangle - y\right)^2$$

Compute $\dfrac{\partial z}{\partial \mathbf{w}}$

Lecture 2 - pg.35

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \frac{\partial \mathbf{y}}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}}$$

Assume $\mathbf{x}, \mathbf{w} \in \mathbb{R}^n, \quad y \in \mathbb{R}$

$$z = \left(\langle \mathbf{x}, \mathbf{w} \rangle - y\right)^2$$

Compute $\dfrac{\partial z}{\partial \mathbf{w}}$

Decompose

$$a = \langle \mathbf{x}, \mathbf{w} \rangle$$
$$b = a - y$$
$$z = b^2$$

$$\begin{aligned}
\frac{\partial z}{\partial \mathbf{w}} &= \frac{\partial z}{\partial b} \frac{\partial b}{\partial a} \frac{\partial a}{\partial \mathbf{w}} \\
&= \frac{\partial b^2}{\partial b} \frac{\partial a - y}{\partial a} \frac{\partial \langle \mathbf{x}, \mathbf{w} \rangle}{\partial \mathbf{w}} \\
&= 2b \cdot 1 \cdot \mathbf{x}^T \\
&= 2\left(\langle \mathbf{x}, \mathbf{w} \rangle - y\right) \mathbf{x}^T
\end{aligned}$$

Homework 1 problem 4

Problem 4: Vector Calculus (20')

Suppose x is a 3-d vector.

$$f(x) = |e^{A \cdot x + b} - c|_2^2$$

where

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1.5 & -2 \end{bmatrix}, b = \begin{bmatrix} -3 \\ -2 \end{bmatrix}, c = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}$$

$| \cdot |_2$ is 2-norm: $|x|_2 = \sqrt{x_1^2 + x_2^2 + \ldots}$.

What is the differential $\frac{\partial f}{\partial x}$?

Homework 1 problem 4

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (3\times 1)$$

dim of $Ax+b$: 2×1 → dim of u : 2×1

u^T : 1×2

let $u = e^{Ax+b} - C$

$$f(x) = |u|_2^2 = u^T u$$

$$\frac{df}{du} = \frac{d}{du} u^T u = 2u^T$$

$$\frac{df}{dx} = \frac{df}{du}\frac{du}{dx}$$

$$\frac{du}{dx} = \begin{bmatrix} \frac{du_1}{dx_1} & \frac{du_1}{dx_2} & \frac{du_1}{dx_3} \\ \frac{du_2}{dx_1} & \frac{du_2}{dx_2} & \frac{du_2}{dx_3} \end{bmatrix}$$

$= \ldots$ simplify

$$= 2u^T \frac{du}{dx}$$

(1×2) (2×3)

→ (1×3)

UC SANTA BARBARA