

# MATRIX CALCULUS USING "THE TRACE TRICK"

Trace: sum of diagonal Elements

Properties :

- $\text{Tr}[A + B] = \text{Tr}[A] + \text{Tr}[B]$
- $\text{Tr}[cA] = c \text{Tr}[A]$
- $\text{Tr}[AB] = \text{Tr}[BA]$
- $\text{Tr}[A] = \text{Tr}[A^T]$

$$\text{Tr}(A) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

Differential: small change in all elements.

Properties :

- $d(cA) = c dA$
- $d(A + B) = dA + dB$
- $d(AB) = dA B + B dA$

$$dA = \begin{bmatrix} da_{11} & \dots & da_{1n} \\ \vdots & \ddots & \vdots \\ da_{m1} & \dots & da_{mn} \end{bmatrix}$$

Theorem :

If we can write differential of scalar function  $L(x)$  into the form:

$$dL = \text{Tr}(G^T dX)$$

Then gradient of  $L$  w.r.t  $X$  is :

$$\nabla_X L = G$$

Problem 4 : Trace of linear product

Let  $L = \text{Tr}(AX)$ . Find  $\nabla_X L$

Problem 4 : Trace of linear product

Let  $L = \text{Tr}(AX)$ . Find  $\nabla_x L$

Step 1: Take Differential on both sides

$$dL = d(\text{Tr}(AX))$$

Step 2. Since trace is sum and  $d$  is a linear operator, we can swap  $d$  and  $\text{Tr}$ .

$$dL = \text{Tr}(d(AX)) = \text{Tr}(Adx)$$

Step 3. Identify the matrix  $G$ .

$$G^T = A$$

$$\Rightarrow G = A^T$$

Answer:  $\nabla_x L = A^T$

Problem 5 : Trace of Quadratic product

Let  $L = \text{tr}(X^T A X)$ , where  $A$  is constant sq. matrix. Find  $\nabla_X L$ .

## Problem 5: Trace of Quadratic product

Let  $L = \text{Tr}(X^T A X)$ , where  $A$  is constant sq. matrix. Find  $\nabla_X L$ .

Step 1: Take differential on both sides & use product rule of differential.

$$\begin{aligned} dL &= \text{Tr}(d(X^T A X)) \\ &= \text{Tr}(dX^T A X + X^T A dX) \end{aligned}$$

Step 2: Linearity of Trace

$$dL = \text{Tr}(dX^T A X) + (X^T A dX) \quad \textcircled{1}$$

Step 3: Trace of transposed is same.

$$\text{Tr}(dX^T A X) = \text{Tr}(X^T A^T dX) \quad \textcircled{11}$$

using it in  $\textcircled{1}$ :

$$\begin{aligned} dL &= \text{Tr}(X^T A^T dX) + \text{Tr}(X^T A dX) \\ &= \text{Tr}(X^T A^T dX + X^T A dX) \\ &= \text{Tr}(X^T (A^T + A) dX) \end{aligned}$$

Step 4 : Identify  $G$ .

$$G^T = X^T (A^T + A)$$

$$G = (A + A^T)X$$

Answer :  $\nabla_X L = (A + A^T)X$

---

Problem 6 : Linear Regression

Let  $L = \|Ax - b\|_2^2$ . Derive  $\nabla_X L$  and set it = 0.

Step 1. Convert  $L$  to Trace form.

$$\begin{aligned} L &= (Ax - b)^T (Ax - b) \\ &= R^T R \\ &= \text{Tr}(R^T R) \quad \text{where } R = Ax - b \end{aligned}$$

Step 2. Apply differential on both sides

$$\begin{aligned} dL &= \text{Tr}(dR^T R + R^T dR) \\ &= \text{Tr}(2R^T dR) \quad (\text{remember problem 5}) \end{aligned}$$

Step 3. Find  $dR$ .

$$\begin{aligned} dR &= d(Ax - b) \\ &= Adx - \cancel{db} \rightsquigarrow 0 \\ &= Adx \end{aligned}$$

Plug it back to  $dL$ :

$$\begin{aligned} dL &= \text{Tr}(2 Q^T Adx) \\ &= \text{Tr}(2 (Ax - b)^T Adx) \end{aligned}$$

Step 4: Identify  $G$ .

$$G^T = 2 (Ax - b)^T A$$

$$G = 2 A^T (Ax - b)$$

$$\nabla_x L = 2 A^T (Ax - b)$$

Step 5: Set it to zero

$$X = (A^T A)^{-1} A^T b$$