

Fun with matrix algebra

Prove the following statements. All of them can be solved by first expanding out the matrix notation as a combination of their elements, and then use the definitions of trace and matrix derivatives to help finish the proof. For example, the (i, j) element of $Y = AB$ is $Y_{i,j} = \sum_m A_{i,m} B_{m,j}$.

- $\nabla_A \text{tr} AB = B^T$
- $\nabla_{A^T} f(A) = (\nabla_A f(A))^T$
- $\nabla_A \text{tr} ABA^T C = CAB + C^T AB^T$

Hint: Try first solving the easier equation of $\nabla_A \text{tr} BAC = (CB)^T = B^T C^T$

$$\text{גורמים} \quad A_{(2 \times 2)} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B_{(2 \times 2)} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \quad C_{(2 \times 2)} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

$$\textcircled{1} \quad \nabla_A \text{tr} AB = B^T$$

$$AB = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

$$\text{tr}(AB) = a_{11}b_{11} + a_{12}b_{21} + a_{21}b_{12} + a_{22}b_{22}$$

$$\nabla_A \text{tr}(AB) = \begin{bmatrix} \text{tr}(AB)/a_{11} & \text{tr}(AB)/a_{12} \\ \text{tr}(AB)/a_{21} & \text{tr}(AB)/a_{22} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}^T = B^T$$

$$\therefore \nabla_A \text{tr}(AB) = B^T$$

$$\textcircled{2} \quad \nabla_{A^T} f(A) = (\nabla_A f(A))^T$$

$$A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

$$\nabla_{A^T} f(A) = \begin{bmatrix} f(A)/a_{11} & f(A)/a_{21} \\ f(A)/a_{12} & f(A)/a_{22} \end{bmatrix} = \begin{bmatrix} f(A)/a_{11} & f(A)/a_{12} \\ f(A)/a_{21} & f(A)/a_{22} \end{bmatrix}^T = (\nabla_A f(A))^T$$

$$\therefore \nabla_{A^T} f(A) = (\nabla_A f(A))^T$$

$$\textcircled{3} \quad \nabla_A \text{tr}(BAC) = (CB)^T = B^T C^T$$

$$BAC = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{21}b_{12} & a_{12}b_{11} + a_{22}b_{12} \\ a_{11}b_{21} + a_{21}b_{22} & a_{12}b_{21} + a_{22}b_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11}c_{11} + a_{21}b_{12}c_{11} + a_{12}b_{11}c_{21} + a_{22}b_{12}c_{21} & a_{11}b_{11}c_{12} + a_{21}b_{12}c_{12} + a_{12}b_{11}c_{22} + a_{22}b_{12}c_{22} \\ a_{11}b_{21}c_{11} + a_{21}b_{22}c_{11} + a_{12}b_{21}c_{21} + a_{22}b_{22}c_{21} & a_{11}b_{21}c_{12} + a_{21}b_{22}c_{12} + a_{12}b_{21}c_{22} + a_{22}b_{22}c_{22} \end{bmatrix}$$

$$\text{tr}(BAC) = a_{11}(b_{11}c_{11} + b_{21}c_{21}) + a_{12}(b_{11}c_{21} + b_{21}c_{22}) + a_{21}(b_{12}c_{11} + b_{22}c_{21}) + a_{22}(b_{12}c_{21} + b_{22}c_{22})$$

$$\nabla_A \text{tr}(BAC) = \begin{bmatrix} b_{11}c_{11} + b_{21}c_{21} & b_{11}c_{21} + b_{21}c_{22} \\ b_{12}c_{11} + b_{22}c_{21} & b_{12}c_{21} + b_{22}c_{22} \end{bmatrix} = \textcircled{1}$$

$$\text{מה } \textcircled{1} \rightarrow \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{bmatrix} = B^T C^T$$

$$\text{מה } \textcircled{1} \rightarrow \begin{bmatrix} b_{11}c_{11} + b_{21}c_{21} & b_{12}c_{11} + b_{22}c_{21} \\ b_{11}c_{21} + b_{21}c_{22} & b_{12}c_{21} + b_{22}c_{22} \end{bmatrix}^T = \left(\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \right)^T = (CB)^T$$

$$4) \text{Var}(ABA^T C) = CAB + C^T AB^T$$

$$\begin{aligned} \text{RHS : } CAB &= \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}c_{11} + a_{21}c_{12} & a_{12}c_{11} + a_{22}c_{12} \\ a_{11}c_{21} + a_{21}c_{22} & a_{12}c_{21} + a_{22}c_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11}c_{11} + a_{21}b_{11}c_{12} + a_{12}b_{21}c_{11} + a_{22}b_{21}c_{12} & a_{11}b_{12}c_{11} + a_{21}b_{12}c_{12} + a_{12}b_{22}c_{11} + a_{22}b_{22}c_{12} \\ a_{11}b_{21}c_{21} + a_{21}b_{11}c_{22} + a_{12}b_{21}c_{21} + a_{22}b_{21}c_{22} & a_{11}b_{22}c_{21} + a_{21}b_{12}c_{22} + a_{12}b_{22}c_{21} + a_{22}b_{22}c_{22} \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \text{RHS : } C^T AB^T &= \begin{bmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}c_{11} + a_{21}c_{21} & a_{12}c_{11} + a_{22}c_{21} \\ a_{11}c_{12} + a_{21}c_{22} & a_{12}c_{12} + a_{22}c_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11}c_{11} + a_{21}b_{11}c_{21} + a_{12}b_{12}c_{11} + a_{22}b_{12}c_{21} & a_{11}b_{21}c_{11} + a_{21}b_{21}c_{21} + a_{12}b_{22}c_{11} + a_{22}b_{22}c_{21} \\ a_{11}b_{11}c_{12} + a_{21}b_{11}c_{22} + a_{12}b_{12}c_{12} + a_{22}b_{12}c_{22} & a_{11}b_{21}c_{12} + a_{21}b_{21}c_{22} + a_{12}b_{22}c_{12} + a_{22}b_{22}c_{22} \end{bmatrix} \end{aligned}$$

$$\text{RHS : } CAB + C^T AB^T = \begin{bmatrix} a_{11}b_{11}c_{11} + a_{12}b_{12}c_{11} + a_{12}b_{21}c_{11} + a_{21}b_{11}c_{12} + a_{21}b_{11}c_{21} + a_{22}b_{12}c_{21} + a_{22}b_{21}c_{12} & a_{11}b_{12}c_{11} + a_{12}b_{21}c_{11} + 2a_{12}b_{22}c_{11} + a_{21}b_{12}c_{12} + a_{21}b_{21}c_{21} + a_{22}b_{22}c_{12} + a_{22}b_{21}c_{22} \\ a_{11}b_{11}c_{12} + a_{11}b_{11}c_{21} + a_{12}b_{12}c_{12} + a_{12}b_{21}c_{21} + 2a_{21}b_{11}c_{22} + a_{22}b_{12}c_{22} + a_{22}b_{21}c_{22} & a_{11}b_{12}c_{21} + a_{11}b_{21}c_{12} + a_{12}b_{22}c_{12} + a_{12}b_{22}c_{21} + a_{21}b_{12}c_{22} + a_{21}b_{21}c_{22} + 2a_{22}b_{22}c_{22} \end{bmatrix}$$

$$\begin{aligned} \text{LHS : } ABA^T C &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}^2 b_{11} + a_{11}a_{12}b_{21} + a_{11}a_{12}b_{12} + a_{12}^2 b_{22} & a_{11}a_{21}b_{11} + a_{11}a_{21}b_{21} + a_{11}a_{22}b_{12} + a_{12}a_{22}b_{22} \\ a_{11}a_{21}b_{11} + a_{11}a_{22}b_{21} + a_{12}a_{21}b_{12} + a_{12}a_{22}b_{22} & a_{21}^2 b_{11} + a_{21}a_{22}b_{21} + a_{21}a_{22}b_{12} + a_{22}^2 b_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}^2 b_{11}c_{11} + a_{11}a_{12}b_{21}c_{11} + a_{11}a_{12}b_{12}c_{11} + a_{12}^2 b_{22}c_{11} + a_{11}a_{21}b_{11}c_{21} + a_{11}a_{22}b_{12}c_{21} + a_{12}a_{22}b_{22}c_{21} + a_{12}a_{21}b_{21}c_{21} \\ a_{11}^2 b_{11}c_{12} + a_{11}a_{12}b_{21}c_{12} + a_{11}a_{12}b_{12}c_{12} + a_{12}^2 b_{22}c_{12} + a_{11}a_{21}b_{11}c_{22} + a_{11}a_{22}b_{12}c_{22} + a_{12}a_{22}b_{22}c_{22} + a_{12}a_{21}b_{21}c_{22} \\ a_{11}a_{21}b_{11}c_{11} + a_{11}a_{22}b_{21}c_{11} + a_{12}a_{21}b_{12}c_{11} + a_{12}a_{22}b_{22}c_{11} + a_{21}^2 b_{11}c_{21} + a_{21}a_{22}b_{21}c_{21} + a_{21}a_{22}b_{12}c_{21} + a_{22}^2 b_{22}c_{21} \\ a_{11}a_{21}b_{11}c_{12} + a_{11}a_{22}b_{21}c_{12} + a_{12}a_{21}b_{12}c_{12} + a_{12}a_{22}b_{22}c_{12} + a_{21}^2 b_{11}c_{22} + a_{21}a_{22}b_{21}c_{22} + a_{21}a_{22}b_{12}c_{22} + a_{22}^2 b_{22}c_{22} \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \text{tr}(ABA^T C) &= a_{11}^2 b_{11}c_{11} + a_{11}a_{12}b_{21}c_{11} + a_{11}a_{12}b_{12}c_{11} + a_{12}^2 b_{22}c_{11} + a_{11}a_{21}b_{11}c_{21} + a_{11}a_{22}b_{12}c_{21} + a_{12}a_{22}b_{22}c_{21} + a_{12}a_{21}b_{21}c_{21} \\ &\quad + a_{11}a_{21}b_{11}c_{12} + a_{11}a_{22}b_{21}c_{12} + a_{12}a_{21}b_{12}c_{12} + a_{12}a_{22}b_{22}c_{12} + a_{21}^2 b_{11}c_{22} + a_{21}a_{22}b_{21}c_{22} + a_{21}a_{22}b_{12}c_{22} + a_{22}^2 b_{22}c_{22} \end{aligned}$$

$$\text{Var}(ABA^T C) = \begin{bmatrix} a_{11}b_{11}c_{11} + a_{12}b_{12}c_{11} + a_{12}b_{21}c_{11} + a_{21}b_{11}c_{12} + a_{21}b_{11}c_{21} + a_{22}b_{12}c_{21} + a_{22}b_{21}c_{12} & a_{11}b_{12}c_{11} + a_{11}b_{21}c_{11} + a_{12}b_{22}c_{11} + a_{21}b_{12}c_{12} + a_{21}b_{21}c_{21} + a_{22}b_{22}c_{12} + a_{22}b_{21}c_{22} \\ a_{11}b_{11}c_{12} + a_{11}b_{11}c_{21} + a_{12}b_{12}c_{12} + a_{12}b_{21}c_{21} + 2a_{21}b_{11}c_{22} + a_{22}b_{12}c_{22} + a_{22}b_{21}c_{22} & a_{11}b_{12}c_{21} + a_{11}b_{21}c_{12} + a_{12}b_{22}c_{12} + a_{12}b_{22}c_{21} + a_{21}b_{12}c_{22} + a_{21}b_{21}c_{22} + 2a_{22}b_{22}c_{22} \end{bmatrix}$$

$$\therefore \text{LHS} = \text{RHS} \therefore \text{Var}(ABA^T C) = CAB + C^T AB^T$$