

ECE 532 - lecture 28 - linear matrix problems

①

Linear equations involving matrix variables.

We know how to solve $Ax = b$ or $\min_x \|Ax - b\|^2$

what about something like $AXB = C$ or $\min_x \|AXB - C\|_F^2$?

these are still linear equations, but more annoying to deal with.

Example:
$$\begin{pmatrix} a_1 & a_3 \\ a_2 & a_4 \end{pmatrix} \begin{pmatrix} x_1 & x_3 \\ x_2 & x_4 \end{pmatrix} \begin{pmatrix} b_1 & b_3 \\ b_2 & b_4 \end{pmatrix} = \begin{pmatrix} c_1 & c_3 \\ c_2 & c_4 \end{pmatrix}$$

$$\begin{pmatrix} a_1 x_1 + a_3 x_2 & a_1 x_3 + a_3 x_4 \\ a_2 x_1 + a_4 x_2 & a_2 x_3 + a_4 x_4 \end{pmatrix} \begin{pmatrix} b_1 & b_3 \\ b_2 & b_4 \end{pmatrix} = \begin{pmatrix} c_1 & c_3 \\ c_2 & c_4 \end{pmatrix}$$

$$\begin{pmatrix} a_1 b_1 x_1 + a_3 b_1 x_2 + a_1 b_2 x_3 + a_3 b_2 x_4 & a_1 b_3 x_1 + a_3 b_3 x_2 + a_1 b_4 x_3 + a_3 b_4 x_4 \\ a_2 b_1 x_1 + a_4 b_1 x_2 + a_2 b_2 x_3 + a_4 b_2 x_4 & a_2 b_3 x_1 + a_4 b_3 x_2 + a_2 b_4 x_3 + a_4 b_4 x_4 \end{pmatrix}$$

$$= \begin{pmatrix} a_1 b_1 & a_3 b_1 & a_1 b_2 & a_3 b_2 \\ a_2 b_1 & a_4 b_1 & a_2 b_2 & a_4 b_2 \\ a_1 b_3 & a_3 b_3 & a_1 b_4 & a_3 b_4 \\ a_2 b_3 & a_4 b_3 & a_2 b_4 & a_4 b_4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} \quad \left\{ \begin{array}{l} \text{linear equations} \\ \text{in } x_1, x_2, x_3, x_4 \\ \text{i.e. } X = \begin{pmatrix} x_1 & x_3 \\ x_2 & x_4 \end{pmatrix}. \end{array} \right.$$

$$= \begin{pmatrix} b_1 A & b_2 A \\ b_3 A & b_4 A \end{pmatrix} \begin{matrix} \rightarrow B^T \otimes A \quad (\text{"Kronecker product"}). \\ \text{vec}(x) = \text{vec}(c) \end{matrix}$$

In general, we have:

$$\text{vec}(AXB) = (B^T \otimes A) \text{vec}(X)$$

This fact is useful in transforming matrix norms like the Frobenius norm because it is invariant under reshaping. i.e.

$$\|M\|_F^2 = \sum_i \sum_j m_{ij}^2 = \|\text{vec}(M)\|_2^2$$

Therefore:

$$\min_X \|AXB - C\|_F^2 = \min_x \|(B^T \otimes A)x - \text{vec}(C)\|_2^2$$

where $\text{vec}(X) = x$.

Note: this does not work with the spectral (induced) norm

because $\|M\|_2 \neq \|\text{vec}(M)\|_2$ in general.

* General definition: $A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m1}B & \dots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{mp \times nq}$

works with matrices of any sizes.

* obvious property: $(A \otimes B)^T = A^T \otimes B^T$.

* important fact: $A \otimes B \neq B \otimes A$. (does not commute!)

more obvious properties

$$\left. \begin{aligned} A \otimes (B+C) &= A \otimes B + A \otimes C \\ (A+B) \otimes C &= (A+B) \otimes C \end{aligned} \right\} \text{distributed across addition.}$$

$$(kA) \otimes B = A \otimes (kB) = k(A \otimes B) \quad \left. \vphantom{(kA) \otimes B} \right\} \text{scalar multiplication.}$$

$$(A \otimes B) \otimes C = A \otimes (B \otimes C) \quad \left. \vphantom{(A \otimes B) \otimes C} \right\} \text{associative}$$

Most important property

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

whenever the dimensions are compatible.

Proof:

$$\left[\begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix} \begin{pmatrix} c_{11}D & \dots & c_{1q}D \\ \vdots & \ddots & \vdots \\ c_{n1}D & \dots & c_{nq}D \end{pmatrix} \right]_{ij}$$

$$= \begin{bmatrix} a_{11}B & \dots & a_{1n}B \end{bmatrix} \begin{bmatrix} c_{ij}D \\ \vdots \\ c_{nj}D \end{bmatrix}$$

$$= (a_{11}c_{ij} + \dots + a_{1n}c_{nj})(BD)$$

$$= (AC)_{ij}(BD)$$

$$= [(AC) \otimes (BD)]_{ij}$$

this is important because it mixes the Kronecker product and the regular product

what about the SVD?

$$\text{let } A = U_A \Sigma_A V_A^T$$

$$\text{and } B = U_B \Sigma_B V_B^T$$

$$\begin{aligned} \text{then } A \otimes B &= (U_A \Sigma_A V_A^T) \otimes (U_B \Sigma_B V_B^T) \\ &= (U_A \otimes U_B) (\Sigma_A \otimes \Sigma_B) (V_A \otimes V_B)^T \end{aligned} \quad (\star)$$

Note: $U_A \otimes U_B$ is orthogonal: (similarly for $V_A \otimes V_B$).

$$\begin{aligned} (U_A \otimes U_B)^T (U_A \otimes U_B) &= (U_A^T \otimes U_B^T) (U_A \otimes U_B) \\ &= (U_A^T U_A) \otimes (U_B^T U_B) \\ &= I \otimes I \\ &= I. \end{aligned}$$

therefore, (\star) is a valid SVD for (\star) , but we will probably have to rearrange rows/cols so that $\Sigma_A \otimes \Sigma_B$ is sorted.

$$\text{if } \Sigma_A = \text{diag}\{\sigma_1, \dots, \sigma_{r_A}\} \quad \text{and} \quad \Sigma_B = \text{diag}\{\mu_1, \dots, \mu_{r_B}\}$$

then we conclude singular values of $A \otimes B$ are $\{\sigma_i \mu_j\}_{\substack{1 \leq i \leq r_A \\ 1 \leq j \leq r_B}}$.

in other words, $\text{rank}(A \otimes B) = (\text{rank } A)(\text{rank } B)$.

We also have:

$$\begin{aligned}
(A \otimes B)^\dagger &= \left[(U_A \otimes U_B) (\Sigma_A \otimes \Sigma_B) (V_A \otimes V_B)^T \right]^\dagger && \text{(use their SVDs).} \\
&= (V_A \otimes V_B) (\Sigma_A \otimes \Sigma_B)^\dagger (U_A \otimes U_B)^T \\
&= (V_A \Sigma_A^\dagger U_A^T) \otimes (V_B \Sigma_B^\dagger U_B^T) \\
&= A^\dagger \otimes B^\dagger.
\end{aligned}$$

if A, B are square, then $(A \otimes B)$ is invertible iff A, B both invertible
 and $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

★ Least squares

$$\text{minimize}_X \|AXB - C\|_F^2$$

$$\Leftrightarrow \|(B^T \otimes A) \text{vec}(X) - \text{vec}(C)\|_2^2$$

Normal equations: $(B^T \otimes A)^T (B^T \otimes A) \text{vec}(X) = (B^T \otimes A)^T \text{vec}(C)$.

$$\Leftrightarrow [(B B^T) \otimes (A A^T)] \text{vec}(X) = (B \otimes A^T) \text{vec}(C).$$

$$\Leftrightarrow A^T A X B B^T = A^T C B^T \quad \text{in matrix form.}$$

Min-norm solution:
 (minimize $\|X\|_F^2$).

$$\text{vec}(X) = (B^T \otimes A)^\dagger \text{vec}(C) = (B^T \otimes A^\dagger) \text{vec}(C).$$

or, $X_{\text{opt}} = A^\dagger C B^\dagger$

Other examples of matrix equations:

★ "Lyapunov equation" in control theory and dynamical systems analysis.

$$A^T X + X A + Q = 0$$

this is linear in X ; vectorize:

$$\Rightarrow A^T X I + I X A + Q = 0$$

$$\Rightarrow (I \otimes A^T) \text{vec}(X) + (A^T \otimes I) \text{vec}(X) + \text{vec}(Q) = 0$$

$$\Rightarrow [I \otimes A^T + A^T \otimes I] \text{vec}(X) = -\text{vec}(Q).$$

And solve via standard method.