

LECTURE 9 : OUTLINE

- Matrix equations
- The Kronecker product
- Vectorization
- The Khatri-Rao product
- Differentiation



KTH Electrical Engineering

Parts of Chapter 4 in "Topics in Matrix Analysis," by R. A. Horn and C. R. Johnson + additional material, see references on the last slide.

THE KRONECKER PRODUCT

Let $A = [a_{ij}] \in M_{m,n}$ and $B = [b_{ij}] \in M_{p,q}$. The Kronecker product of A and B is defined as

$$A \otimes B \equiv \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix} \in M_{mp,nq}$$



KTH Electrical Engineering

Notice that $A \otimes B \neq B \otimes A$ in general.

The k th Kronecker power is defined as

$$A^{\otimes k} \equiv A \otimes A^{\otimes(k-1)}; \quad A^{\otimes 1} \equiv A$$

MATRIX EQUATIONS

Examples:

$$XA + A^*X = B$$

$$AX = B$$

$$AX = XA$$

$$AXB + CXD = E$$

$$AX + YB = C$$

$$X^2 = A$$

$$X^T AX + B^T X + X^T B = C$$



KTH Electrical Engineering

KRONECKER PRODUCT: SOME PROPERTIES

For matrices A, B, C, D (of suitable dimensions) and scalar α we have:

$$(\alpha A) \otimes B = A \otimes (\alpha B) = \alpha(A \otimes B)$$

$$(A \otimes B)^T = A^T \otimes B^T$$

$$(A \otimes B)^* = A^* \otimes B^*$$

$$(A \otimes B) \otimes C = A \otimes (B \otimes C)$$

$$(A + B) \otimes C = (A \otimes C) + (B \otimes C)$$

$$A \otimes (B + C) = (A \otimes B) + (A \otimes C)$$

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

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

if the inverses exist.



KTH Electrical Engineering

THE VEC OPERATOR

Let $A = [a_{ij}] \in M_{m,n}$. Then the vector $\text{vec}(A) \in \mathbf{C}^{mn}$ is defined as

$$\text{vec}(A) = [a_{11} \ a_{21} \ \dots \ a_{m1} \ a_{12} \ \dots \ a_{m2} \ \dots \ a_{1n} \ \dots \ a_{mn}]^T$$

It is simple to verify that

$$\text{tr}(AB) = \text{vec}^T(A^T) \text{vec}(B) = \text{vec}^T(B^T) \text{vec}(A) = \text{vec}^*(A^*) \text{vec}(B)$$

A very useful result is

$$\text{vec}(ABC) = (C^T \otimes A) \text{vec}(B)$$

for any matrices A, B, C of appropriate dimensions.



KTH Electrical Engineering

KRONECKER PRODUCTS: FURTHER PROPERTIES

Thm: Let $A \in M_m$ and $B \in M_n$. If (λ, x) is an eigenvalue/eigenvector pair of A and similarly (μ, y) an eigenvalue/vector pair of B , then $\lambda\mu$ is an eigenvalue of $A \otimes B$ with the corresponding eigenvector $x \otimes y$. Furthermore, every eigenvalue arises in this way; that is, if $\sigma(A) = \{\lambda_1, \dots, \lambda_m\}$ and $\sigma(B) = \{\mu_1, \dots, \mu_n\}$, then

$$\sigma(A \otimes B) = \{\lambda_i \mu_j : i = 1, 2, \dots, m; j = 1, 2, \dots, n\}$$

Notice also that $\sigma(A \otimes B) = \sigma(B \otimes A)$.



KTH Electrical Engineering

MATRIX EQUATIONS, CONT'D

$$XA + A^*X = B \Leftrightarrow [(A^T \otimes I) + (I \otimes A^*)] \text{vec}(X) = \text{vec}(B)$$

$$AX = B \Leftrightarrow (I \otimes A) \text{vec}(X) = \text{vec}(B)$$

$$AX = XA \Leftrightarrow [(I \otimes A) - (A^T \otimes I)] \text{vec}(X) = 0$$

$$AXB + CXD = E \Leftrightarrow [(B^T \otimes A) + (D^T \otimes C)] \text{vec}(X) = \text{vec}(E)$$



KTH Electrical Engineering

More generally:

Lemma: Let $T : M_{m,n} \rightarrow M_{p,q}$ be a given linear transformation. There exists a unique matrix $K(T) \in M_{pq, mn}$ such that

$$\text{vec}(T(X)) = K(T) \text{vec}(X)$$

for all $X \in M_{m,n}$.

KRONECKER PRODUCTS: FURTHER PROPERTIES CONT'D

Cor: If $A \in M_m$ and $B \in M_n$ are positive (semi)definite, then $A \otimes B$ is also positive (semi)definite.

Cor: If $A \in M_m$ and $B \in M_n$, then

$$\begin{aligned} \text{tr}(A \otimes B) &= \text{tr}(A) \text{tr}(B) \\ \det(A \otimes B) &= \det(A)^n \det(B)^m \end{aligned}$$



KTH Electrical Engineering

LYAPUNOV EQUATION, CONT'D

Returning to the Lyapunov type equation

$$AX + XB = C \Leftrightarrow [(I_m \otimes A) + (B^T \otimes I_n)] \text{vec}(X) = \text{vec}(C)$$

According to the previous result, this equation has a unique solution \bar{X} if and only if $\lambda_i(A) + \mu_j(B) \neq 0$ for all i, j or equivalently $\sigma(A) \cap \sigma(-B) = \emptyset$ (empty set).

In the complex valued case we had the Lyapunov equation

$$XA + A^* \bar{X} = C$$

It has a unique solution \bar{X} if and only if $\sigma(-A^*) \cap \sigma(A) = \emptyset$ or $\overline{\sigma(-A)} \cap \sigma(A) = \emptyset$. This condition is certainly satisfied when A is stable (positive or negative).



THE HADAMARD PRODUCT

Let $A = [a_{ij}] \in M_{m,n}$ and $B = [b_{ij}] \in M_{m,n}$. The Hadamard or Schur product of A and B is defined as

$$A \circ B \equiv [a_{ij} b_{ij}] \in M_{m,n}$$

More information and properties in Chapter 5 in "Topics in Matrix Analysis."



THE KHATRI-RAO PRODUCT

The Khatri-Rao product of $A \in M_{m,n}$ and $B \in M_{p,n}$ is defined as (the symbol may differ)

$$A \odot B = [a_1 \otimes b_1 \ a_2 \otimes b_2 \ \dots \ a_n \otimes b_n]$$

where a_i, b_i denote the i th column in A and B , respectively.

It is useful, for example, in matrix equations in which diagonal matrices are involved. Let A, B, C be matrices of appropriate dimensions and let B be a diagonal matrix.

Then

$$\text{vec}(ABC) = (C^T \otimes A) \text{vec}(B) = (C^T \odot A) \text{diag}(B)$$

where $\text{diag}(B)$ denotes the column vector of the diagonal elements of B .



DERIVATIVES: SOME DEFINITIONS

The derivative of a matrix $A(t) = [a_{ij}(t)]$ that depends on a (real) scalar t is defined as the matrix

$$\frac{dA(t)}{dt} = \left[\frac{da_{ij}(t)}{dt} \right]$$

We can also define the derivative of a vector $y \in \mathbf{R}^m$ with respect to a vector $x \in \mathbf{R}^n$ as follows

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

(Please notice that other definitions are often used!)



DERIVATIVES, CONT'D

Clearly, if y is a scalar we get

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix}$$

or when x is a scalar

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} & \frac{\partial y_2}{\partial x} & \dots & \frac{\partial y_m}{\partial x} \end{bmatrix}$$



KTH Electrical Engineering

DERIVATIVES; SIMPLE RESULTS

Let $x \in \mathbf{R}^n$ and A a matrix independent of x .

$$\frac{\partial Ax}{\partial x} = A^T$$

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

$$\frac{\partial x^T x}{\partial x} = 2x$$

$$\frac{\partial x^T Ax}{\partial x} = (A + A^T)x \quad (= 2Ax \text{ if } A \text{ is symmetric})$$

$$\frac{\partial}{\partial x} \left[\frac{\partial x^T Ax}{\partial x} \right] = (A + A^T) \quad (= 2A \text{ if } A \text{ is symmetric})$$



KTH Electrical Engineering

CHAIN RULE FOR VECTORS

Assume $z[y(x)]$ where x, y, z are real vectors, then:

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



KTH Electrical Engineering

DERIVATIVE OF SCALARS WITH RESPECT TO A MATRIX

Let $X = [x_{ij}] \in M_{m,n}(\mathbf{R})$ and let $y = f(X)$ be a real valued scalar function of X . Then we define

$$\frac{\partial y}{\partial X} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{12}} & \dots & \frac{\partial y}{\partial x_{1n}} \\ \frac{\partial y}{\partial x_{21}} & \frac{\partial y}{\partial x_{22}} & \dots & \frac{\partial y}{\partial x_{2n}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial y}{\partial x_{m1}} & \frac{\partial y}{\partial x_{m2}} & \dots & \frac{\partial y}{\partial x_{mn}} \end{bmatrix} = \left[\frac{\partial y}{\partial x_{ij}} \right] = \sum_{i,j} E_{ij} \frac{\partial y}{\partial x_{ij}}$$

where $E_{ij} \in M_{m,n}$ is an elementary matrix with a 1 in the i,j th position and zeros elsewhere.



KTH Electrical Engineering

DERIVATIVES, CONT'D

Let Y, X be square matrices:

$$\frac{\partial \operatorname{tr}(X)}{\partial X} = I$$

$$\frac{\partial \det(Y(X))}{\partial X} = ?$$

We do it componentwise: Notice first that

$$\det(Y) = \sum_r y_{rs} c_{rs}$$

where $c_{r,s}$ is the cofactor of $y_{r,s}$.

$$\frac{\partial \det(Y(X))}{\partial x_{ij}} = \sum_{r,s} \frac{\partial \det(Y(X))}{\partial y_{rs}} \frac{\partial y_{rs}}{\partial x_{ij}} = \sum_{r,s} c_{rs} \frac{\partial y_{rs}}{\partial x_{ij}} = \operatorname{tr}(\operatorname{Adj}(Y) \frac{\partial Y}{\partial x_{ij}})$$



KTH Electrical Engineering

DERIVATIVES, CONT'D

Let $Y(X)$ be positive definite and recall $Y^{-1} = \operatorname{Adj}(Y) / \det(Y)$. Then, from the above, we get

$$\frac{\partial \log \det(Y(X))}{\partial x_{ij}} = \frac{1}{\det(Y(X))} \frac{\partial \det(Y(X))}{\partial x_{ij}}$$

$$= \frac{1}{\det(Y)} \operatorname{tr} \left(\operatorname{Adj}(Y) \frac{\partial Y}{\partial x_{ij}} \right)$$

$$= \operatorname{tr} \left(Y^{-1} \frac{\partial Y}{\partial x_{ij}} \right)$$

Other simple results for matrix expressions wrt elements:

$$\frac{\partial AXB}{\partial x_{ij}} = AE_{ij}B$$

$$\frac{\partial AX^{-1}B}{\partial x_{ij}} = -AX^{-1}E_{ij}X^{-1}B$$



KTH Electrical Engineering

DERIVATIVES OF MATRICES WITH RESPECT TO MATRICES

Some different possibilities:

$$\frac{\partial Y}{\partial X} = \text{a partitioned matrix with } ij\text{th block } \frac{\partial y_{ij}}{\partial x_{ij}}$$

$$\frac{\partial Y}{\partial X} = \text{a partitioned matrix with } ij\text{th block } \frac{\partial y_{ij}}{\partial X}$$

$$\frac{\partial Y}{\partial X} = \frac{\partial \operatorname{vec}(Y)}{\partial \operatorname{vec}(X)}$$



KTH Electrical Engineering

SECOND DERIVATIVES

The second derivative of a scalar function of a vector is a matrix called the Hessian and is defined as

$$\frac{\partial^2 f(x)}{\partial x \partial x^T} = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n \partial x_n} \end{bmatrix}$$



KTH Electrical Engineering

SOME FURTHER READING

- Jan R. Magnus & Heinz Neudecker, Matrix differential calculus with applications in statistics and economics, Chichester : Wiley, 1999
- Alexander Graham, Kronecker products and matrix calculus : with applications, Chichester : Horwood, 1981 (vec, tr, Kronecker, differentiation)
- J. Brewer, "Kronecker products and matrix calculus in system theory," Circuits and Systems, IEEE Transactions on, Vol.25, Iss.9, Sep 1978 Pages: 772- 781 (vec, Kronecker, Khatri-Rao)
- D. Brandwood, "A complex gradient operator and its application in adaptive array theory," IEE Proc., vol. 130, no. 1, pp. 11-16, Feb. 1983. (differentiation wrt complex parameters, signal processing/communications)
- Complex-Valued Matrix Derivatives – With Applications in Signal Processing and Communications by Are Hjørungnes, Cambridge U. Press



KTH Electrical Engineering