STA4724: Big Data Analytics Methods

Daniel Palma

January 21, 2025

Contents

1 Definitions of Matrices and Vectors

Matrix

- a matrix is an arrangement of numbers in rectangular form
- a $J \times K$ matrix has J rows and k columns
- a Square matrix is of order $(2, 2)$ as a special case
- Vectors are subcategories of matrices that have either one row or one column

  $(1, k)$ is one rowm, and multiple columns, e.g. $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$

  $(k, 1)$ is one column, and multiple rows, e.g. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$
- a matrix with one row and one column is the same as a scalar. $a = 5 \Leftrightarrow a = \begin{bmatrix} 5 \end{bmatrix}$

2 Addition, Subtraction, Multiplication

- $A + B = C$
- $A + B \Leftrightarrow B + A$
- $(A + B) + C \Leftrightarrow A + (B + C)$

Transposition An order (j, j) matrix is said to be symmetric if $A = A^T$

- $(A^T)^T \Leftrightarrow A$
- $(kA)^T \Leftrightarrow kA^T$ where k is a scalar
- $(A + B)^T \Leftrightarrow A^T + B^T$
- $kA \Rightarrow k \cdot \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \Rightarrow \begin{bmatrix} k \cdot a_1 & k \cdot a_2 & k \cdot a_3 \end{bmatrix}$
- Given matrix A of order (m, n) and matrix B of order (n, r)

  $C = A \cdot B$ is of order $(m, r) = \begin{bmatrix} C_{mr} \end{bmatrix}$ where $C_{mr} = \sum_{i=1}^{n} A_{mi} \cdot B_{ir}$

Example 1. Given the matrices $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix}$, find $C = A \cdot B$

$$C_{11} = 1 \cdot 7 + 2 \cdot 9 + 3 \cdot 11 = 58$$
$$C_{12} = 1 \cdot 8 + 2 \cdot 10 + 3 \cdot 12 = 64$$
$$C_{21} = 4 \cdot 7 + 5 \cdot 9 + 6 \cdot 11 = 139$$
$$C_{22} = 4 \cdot 8 + 5 \cdot 10 + 6 \cdot 12 = 154$$

Therefore, $C = \begin{bmatrix} 58 & 64 \\ 139 & 154 \end{bmatrix}$

Properties

- $AB \neq BA$
- $A(BC) \Leftrightarrow (AB)C$
- $A(B+C) \Leftrightarrow AB + AC$
- $(AB)^T \Leftrightarrow B^T A^T$
- $A^n \Leftrightarrow A_0 \cdot A_1 \cdot ... \cdot A_{n-1}$

3 Diagonal and Identity Matrices

Diagonal matrix A diagonal matrix is a square matrix with zero entries except possible on the main diagonal

Example 2. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is a diagonal matrix. note that they dont need to be 1s, they can be any number, including zero.

In general, a diagonal matrix is given by $D_{mn} = \begin{bmatrix} d_1 & 0 & 0 & ... & 0 \\ 0 & d_2 & 0 & ... & 0 \\ 0 & 0 & d_3 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & d_n \end{bmatrix}$

Echelon Form

1. row echeleon form (ref)

   The first non-zero element in each row is called the leading entry, is always 1

   Each leading entry is in a column to the right of the leading entry in the previous row (if any)

   Rows with all zero elements are below rows with non-zero elements (if any)

2. reduced row echelon form (rref)

   any ref with the leading entry in each row is the only non-zero entry in its column.

Properties of Diagonal Matrices

- A diagonal matrix D is invertible if and only if all the diagonal elements are non zero.

Example 3. given $D^{-1} = \begin{bmatrix} 1/d_1 & 0 & ... & 0 \\ 0 & 1/d_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1/d_n \end{bmatrix}$

so $DD^{-1} = \begin{bmatrix} d_1 \cdot 1/d_1 & 0 & ... & 0 \\ 0 & d_2 \cdot 1/d_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & d_n \cdot 1/d_n \end{bmatrix} \to I$ which is the identity matrix

Identity Matrix The identity matrix is a square matrix, consisting of ones along the diagonal and zeros elsewhere. Typically, I is used to denote the identity matrix.

Example 4.

$$I_{nn} = \begin{bmatrix} 1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 \end{bmatrix}$$

Properties of Identity Matrices

- $AI = IA = A$

Zero Matrix a zero matrix consists of all zero elements.

4 Determinant and Eigenstructure

Determinant Determinants are defined only for square matrices and scalars.

Example 5. let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $\det(A) = ad - cb$

the determinant of a matrix is denoted by $|A|$ or $\det(A)$ and is a number which encodes a lot of information about the matrix.

In general, we need to first define the cofactor $Q_{r,s}$ of each element of A, $a_{r,s}$. The cofactor of $a_{r,s}$ is $Q_{r,s} = -1^{r+s}|A_{r,s}|$ (where $|A_{r,s}|$ is the determinant of the matrix obtained by deleting the r-th row and s-th column of A).

The last step is to define the determinant of the matrix A as

$$|A| = \sum_{j=1}^{n} a_{ij} Q_{ij}$$

or

$$|A| = \sum_{i=1}^{n} a_{ij} Q_{ij}$$

Properties of Determinants

- $|I| = 1$
- if exchanging two rows of a matrix, we only need to reverse the sign of its determinant
- If we multiply one row of a matrix by a scalar k, the determinant is also multiplied by k.
- The determinant behaves like a linear function on the rows of the matrix

Lemma 6. *The geometric multiplicity of an eigenvalue is at most its algebraic multiplicity.*

Characteristic equation $det(A - \lambda I) = 0$

The Geometric Multiplicity of Eigenvalues

- It is the dimension of the linear space of its associated eigenvectors.

Let A be a $k \times k$ matrix, λ_k be one of the eigenvalues of A and denote its associated eigenspace by E_k. Then the dimension of E_k is called the geometric multiplicity of this eigenvalue λ_k

Proof. Suppose that the geometric multiplicity of λ_k is equal to g, so that there are g linearly independent eigenvectors. $x_1, ...x_g$ associated to λ_k. Randomly choose $k - g$ factors $x_{g+1}...x_k$, all having dimension $k \times l$ and such that the k column vectors $x_1, ..., x_k$ are linearly independent.

Define the $k \times k$ matrix

$$x = [x_1, ..., x_k]$$

for any g, denoted by b_g the vector that solves $xb_g = Ax_g = \lambda x_g$

Define the $k \times (k - g)$ matrix

$$B = [b_g + 1, ..., b_k]$$

and denote by C its upper $g \times (k-g)$ block, and denote by D its lower $(k-g) \times (k-g)$ block

$$B = \begin{bmatrix} C \\ D \end{bmatrix}$$

Denote by I the $k \times k$ identity matrix. for any scalar λ, we have that $(A - \lambda I)X$ ($= 0$ to find x for $\lambda_k \times$)

$\square$

5 Inverses and Singularity

6 Systems of Equations

7 Singular Value Decomposition (SVD)

8 Spectral Decomposition

9 Properties and Derivations of Matrix Traces

10 Projection and Isometry

11 Variance-Covariance Matrix

12 Multivariate Normal Distribution