

ALGEBRA

A **matrix** is a collection of numbers ordered by rows and columns. For example, the following table is a matrix:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

This matrix has m rows and n columns, so it is referred to as a “ m by n ” matrix.

The **transpose** of a matrix is denoted by a prime A' . The first row of a matrix becomes the first column of the transpose matrix, the second row of the matrix becomes the second column of the transpose, etc.

$$B = A' \text{ then } b_{ij} = a_{ji}$$

A **square matrix** has as many rows as it has columns.

A **symmetric matrix** is a square matrix in which $a_{ij} = a_{ji}$ for all i and j .

A **diagonal matrix** is a symmetric matrix where all the off diagonal elements are 0. If $i \neq j$ then $a_{ij} = 0$.

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

An **identity matrix** is a diagonal matrix with 1s and only 1s on the diagonal. The identity matrix is denoted as I .

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A **vector** is a special type of matrix that has only one row (**row vector**) or one column (**column vector**).

Row vector:

$$a = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$$

Column vector:

$$a' = [a_1 \quad \dots \quad a_k]$$

A matrix with one row and one column is called a **scalar** (number).

Operations on matrices.

Matrix Addition and Subtraction.

To add matrices, they must have the same size (the same number of rows and the same number of columns). The proper elements of matrices are simply added.

$$A + B = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

Matrix subtraction works in the same way.

Matrix Multiplication.

Multiplication of a matrix by a scalar. Each element in the product matrix is the scalar multiplied by the element in the matrix.

$$cB = c \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} cb_{11} & \cdots & cb_{1n} \\ \vdots & \ddots & \vdots \\ cb_{m1} & \cdots & cb_{mn} \end{bmatrix}$$

Matrix multiplication involving a scalar is commutative ($cB = Bc$).

Multiplication of a row vector by a column vector. To perform this, the row vector must have as many columns as the column vector has rows. The result is a scalar.

$$b'a = [b_1 \quad \cdots \quad b_k] \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} = b_1a_1 + \cdots + b_ka_k$$

Other types of matrix multiplication involve the multiplication of a row vector and a column vector. For matrix multiplication: the first matrix must have as many columns as the second matrix has rows. The resulting matrix has as many rows as the first matrix and as many columns as the second matrix.

$$AB = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \times \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_1'b_1 & \cdots & a_1'b_n \\ \vdots & \ddots & \vdots \\ a_m'b_1 & \cdots & a_m'b_n \end{bmatrix}$$

Matrix multiplication is not commutative ($AB \neq BA$).

$$AI = IA = A$$

Matrix Inverse.

In scalar algebra, the inverse of a number is that number which multiplied by the original number gives 1. In matrix algebra, the inverse of a matrix is that matrix which multiplied by the original matrix gives an identity matrix.

$$AA^{-1} = A^{-1}A = I$$

A matrix must be square to have an inverse, but not all square matrices have an inverse. We can take the inverse of a square matrix which has a non-zero determinant.

Determinant of a Matrix.

The determinant of a matrix is a scalar and is denoted as $|A|$ or $\det(A)$. The determinant has very important mathematical properties.

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Trace of a Matrix.

The trace of a matrix is denoted as $tr(A)$. The trace is computed only for square matrices and equals to the sum of the diagonal elements of the matrix.

$$tr \begin{bmatrix} 7 & 2 & 3 \\ 8 & 1 & 4 \\ 0 & 3 & 5 \end{bmatrix} = 7 + 1 + 5$$

$$tr(cA) = ctr(A)$$

$$tr(A') = tr(A)$$

$$tr(A + B) = tr(A) + tr(B)$$

$$tr(AB) = tr(BA)$$

Properties of operations on matrices:

$$A + B = B + A$$

$$AB \neq BA$$

$$A(BC) = (AB)C = ABC$$

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

$$(A + B)' = A' + B'$$

$$(AB)' = B'A'$$

$$(AB)^{-1} = B^{-1}A^{-1}$$

$$(A^{-1})' = (A')^{-1}$$

$$\det(A') = \det(A)$$

$$\det(AB) = \det(A)\det(B)$$

A **quadratic form** is given by:

$$x'Ax = \sum_{i=1}^m \sum_{j=1}^m a_{ij}x_i x_j \text{ (the result is a scalar, matrix } A \text{ has to be square)}$$

If for any non-zero vector x a quadratic form is larger than zero then matrix is **positive define**, if a quadratic form is smaller than zero then matrix is **negative define**.

The matrix A is given:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

To know whether it is positive, negative, non-positive, non-negative define we need to compute the following determinants:

$$\det(a_{11})$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$\vdots$

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

If all are positive (non-negative)

$$\det(a_{11}) > 0$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0$$

$\vdots$

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} > 0$$

Then the matrix is positive define (non-negative).

If the first one is negative (non-positive), the second one positive (non-negative), the third one negative (non-positive), ...

$$\det(a_{11}) < 0$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0$$

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} < 0$$

$\vdots$

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} > 0 \text{ if } n \text{ is even}$$

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} < 0 \text{ if } n \text{ is odd}$$

Then the matrix is negative definite (non-positive).

Idempotent matrix.

$$AA = A$$

Linear independence.

We say that columns of a matrix are linearly independent if $A\alpha = 0$ only when $\alpha = 0$.

$$A\alpha = \begin{bmatrix} a_{11}\alpha_1 & \cdots & a_{1n}\alpha_n \\ \vdots & \ddots & \vdots \\ a_{m1}\alpha_1 & \cdots & a_{mn}\alpha_n \end{bmatrix} = \alpha_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \alpha_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + \alpha_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

If the only solution of above equations is $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ then columns of matrix A are linearly independent. It is not possible to express any column of matrix A as linear combination of the others. We say that matrix A has **full column rank**.

The **column rank** of a matrix A is the maximum number of linearly independent column vectors of A .

EXERCISES.

1.

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 3 \\ 1 & 1 \\ 2 & 3 \end{bmatrix}$$

$$2A = 2 \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 2 \times 1 & 2 \times 2 & 2 \times 1 \\ 2 \times 1 & 2 \times 3 & 2 \times 1 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 2 \\ 2 & 6 & 2 \end{bmatrix}$$

$$B' = \begin{bmatrix} 2 & 1 & 2 \\ 3 & 1 & 3 \end{bmatrix}$$

$$A + B' = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 1 & 2 \\ 3 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 3 \\ 2 & 4 & 4 \end{bmatrix}$$

$$AB = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 1 \end{bmatrix} \times \begin{bmatrix} 2 & 3 \\ 1 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 \times 2 + 2 \times 1 + 1 \times 2 & 1 \times 3 + 2 \times 1 + 1 \times 3 \\ 1 \times 2 + 3 \times 1 + 1 \times 2 & 1 \times 3 + 3 \times 1 + 1 \times 3 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 7 & 9 \end{bmatrix}$$

$$|AB| = \det(AB) = \det \begin{bmatrix} 6 & 8 \\ 7 & 9 \end{bmatrix} = 6 \times 9 - 7 \times 8 = 54 - 56 = -2$$

$$2. A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix}$$

Show that matrix A is symmetric.

Show that $AB \neq BA$

Show that A, AB are singular.

Matrix A is 2 by 2 matrix. We verify whether $\forall_{i,j \in 1, \dots, n} a_{ij} = a_{ji}$. In case of matrix A $a_{12} = 2, a_{21} = 2$ hence matrix is symmetric.

$$AB = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \times \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 4 \\ 18 & 8 \end{bmatrix}$$

$$BA = \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 6 & 12 \end{bmatrix}$$

The determinant of a singular matrix is equal to 0. The column or rows are linearly dependent.

$$\det(AB) = 9 \cdot 8 - 18 \cdot 4 = 0$$

$$\det(A) = 1 \cdot 4 - 2 \cdot 2 = 0$$

4. Show that for any invertible A : $(A^{-1})' = (A')^{-1}$

HINT: $A'B' = (BA)'$

$$(A^{-1})' = (A')^{-1} / \times A'$$

$$(A^{-1})'A' = (A')^{-1}A'$$

Left side:

$$(A^{-1})'A' = (AA^{-1})' = I' = I$$

Right side:

$$(A')^{-1}A' = I$$

11.

$$PP = P$$

$$MM = (I - P)(I - P) = II - IP - PI + PP = I - P - P + P = I - P = M$$

$$MP = (I - P)P = IP - PP = P - P = 0$$

12.

$A'A$ non-singular matrix

Show that $P = A(A'A)^{-1}A'$ is idempotent.

$$PP = A(\textcolor{red}{A'A})^{-1}\textcolor{red}{A'}A(A'A)^{-1}A' = A\textcolor{red}{I}(A'A)^{-1}A' = A(A'A)^{-1}A' = P$$

13.

$$M = I - n^{-1}ll'$$

$$l'M = 0$$

$$l = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$ll' = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} = \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}$$

Lets define $P = n^{-1}ll'$. From exercise 12: if P is idempotent then $M = I - P = I - n^{-1}ll'$ is idempotent.

$$PP = n^{-1}ll'n^{-1}ll' = n^{-2}\textcolor{red}{l'l'}ll' = n^{-2}\textcolor{red}{n}ll' = n^{-1}ll' = P$$

$$l'l = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = n$$

Calculus.

Derivative.

$f(x_1, \dots, x_n): R^n \rightarrow R$ scalar function

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

$$\frac{\partial f}{\partial x'} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \dots & \frac{\partial f}{\partial x_n} \end{bmatrix}$$

$f(x_1, \dots, x_n): R^n \rightarrow R^m$ vector function

$f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$ where f_i are scalar functions.

For example:

$f(x_1, x_2, x_3): R^3 \rightarrow R^2$

$f(x_1, x_2, x_3) = (x_1, x_2 + x_3)$

$f_1(x_1, x_2, x_3) = x_1$

$f_2(x_1, x_2, x_3) = x_2 + x_3$

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_n} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

2.

$$\frac{\partial a' \beta}{\partial \beta} = a$$

$$a' = [a_1 \ a_2 \ \dots \ a_k], \quad \beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}$$

We define scalar function:

$$f = a' \beta = a_1 \beta_1 + a_2 \beta_2 + \dots + a_k \beta_k$$

$$\frac{\partial f}{\partial \beta} = \frac{\partial a' \beta}{\partial \beta} = \begin{bmatrix} \frac{\partial a' \beta}{\partial \beta_1} \\ \vdots \\ \frac{\partial a' \beta}{\partial \beta_k} \end{bmatrix} = \begin{bmatrix} \frac{\partial (a_1 \beta_1 + a_2 \beta_2 + \dots + a_k \beta_k)}{\partial \beta_1} \\ \vdots \\ \frac{\partial (a_1 \beta_1 + a_2 \beta_2 + \dots + a_k \beta_k)}{\partial \beta_k} \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} = a$$

$$\frac{\partial a' \beta}{\partial \beta'} = a'$$

$$\frac{\partial a' \beta}{\partial \beta'} = \left[\frac{\partial a' \beta}{\partial \beta_1} \quad \frac{\partial a' \beta}{\partial \beta_2} \quad \dots \quad \frac{\partial a' \beta}{\partial \beta_k} \right] = [a_1 \quad a_2 \quad \dots \quad a_k] = a'$$

3.

$$\frac{\partial A \beta}{\partial \beta'} = A \quad \frac{\partial \beta' A}{\partial \beta} = A$$

$A_{n \times n}$ matrix

β n element column vector.

$$A \beta = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_{1i} \beta_i \\ \vdots \\ \sum_{i=1}^n a_{ni} \beta_i \end{bmatrix}$$

So

$$\frac{\partial A \beta}{\partial \beta'} = \begin{bmatrix} \frac{\sum_{i=1}^n a_{1i} \beta_i}{\partial \beta'} \\ \vdots \\ \frac{\sum_{i=1}^n a_{ni} \beta_i}{\partial \beta'} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} = A$$

$$\beta' A = [\beta_1 \quad \dots \quad \beta_n] \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} = \left[\sum_{i=1}^n a_{i1} \beta_i \quad \dots \quad \sum_{i=1}^n a_{in} \beta_i \right]$$

$$\frac{\partial \beta' A}{\partial \beta} = \left[\frac{\sum_{i=1}^n a_{i1} \beta_i}{\partial \beta} \quad \dots \quad \frac{\sum_{i=1}^n a_{in} \beta_i}{\partial \beta} \right] = A$$

4.

$$\frac{\partial \beta' A \beta}{\partial \beta} = (A + A') \beta$$

$$\beta' A \beta = \sum_{ij} \beta_i \beta_j a_{ij}$$

Then

$$\frac{\partial \beta' A \beta}{\partial \beta} = \begin{bmatrix} \frac{\sum_{ij} \beta_i \beta_j a_{ij}}{\partial \beta_1} \\ \vdots \\ \frac{\sum_{ij} \beta_i \beta_j a_{ij}}{\partial \beta_n} \end{bmatrix}$$

Let's consider l-th element of above matrix:

$$\frac{\sum_{ij} \beta_i \beta_j a_{ij}}{\partial \beta_n} = \frac{\beta_1 \sum_j \beta_j a_{1j} + \beta_2 \sum_j \beta_j a_{2j} + \dots + \beta_l \sum_j \beta_j a_{lj} + \dots + \beta_n \sum_j \beta_j a_{nj}}{\partial \beta_l} =$$

$$\beta_1 a_{1l} + \beta_2 a_{2l} + \dots + \beta_l a_{ll} + \sum_j \beta_j a_{lj} \dots + \beta_n a_{nl} = \sum_j \beta_j a_{lj} + \sum_j \beta_j a_{jl} =$$

$$\sum_j \beta_j (a_{lj} + a_{jl})$$

So:

$$\frac{\partial \beta' A \beta}{\partial \beta} = \begin{bmatrix} \sum_j \beta_j (a_{1j} + a_{j1}) \\ \vdots \\ \sum_j \beta_j (a_{nj} + a_{jn}) \end{bmatrix} = (A + A') \beta$$

If matrix A is symmetric then $A = A'$ and $\frac{\partial \beta' A \beta}{\partial \beta} = 2A\beta$

5. Gradient

$$\frac{\partial f}{\partial \beta} = \begin{bmatrix} \frac{\partial f}{\partial \beta_1} \\ \vdots \\ \frac{\partial f}{\partial \beta_n} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

6. Hessian.

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