Matrix Algebra

What you will learn in this modul

- Definition and Dimension of a Matrix
- Special Matrices
- Operations with Matrices
- Rank of a Matrix
- Rules of Computation
- Expected Value, Variance and Covariance of Multidimensional Random Variables

Matrix: Definition and Dimension of a Matrix I

A *matrix* **A** is defined as a ordered set of *components* a_{ij} with $i = 1, ..., n$ and $j = 1, ..., m$, that are arranged in n rows and m columns. The components a_{ij} are usually real numbers or sometimes numerical random variables. The numbers n and m are called the *dimensions* of **A**, and **A** is then called a $(n \times m)$-matrix (read: "n by m matrix"), or "*of type* $n \times m$". Here is an example of an (2×3)-matrix:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \quad \text{or} \quad \mathbf{A} = (a_{ij}),$$

where $i = 1, 2$ and $j = 1, 2, 3$.

Matrix: Definition and Dimension of a Matrix II

In general an $(n \times m)$-*matrix* $\mathbf{A}$ may be written as follows:

$$\mathbf{A} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

where $i = 1, ..., n$ and $j = 1, ..., m$.

Special Matrices I

A matrix consisting of a single row is called a *row vector*, while a matrix consisting of one column is called a *column vector*. A (2×3)-matrix, for instance, consists of two row vectors $(a_{11} \;\; a_{12} \;\; a_{13})$ and $(a_{21} \;\; a_{22} \;\; a_{23})$ and of three column vectors:

$$\boldsymbol{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} \quad \text{and} \quad \boldsymbol{a}_3 = \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}.$$

If a matrix $\mathbf{A}$ is of type $n \times n$, then it is called an *square matrix of order n*. In this case the components $a_{11}, a_{22}, \ldots, a_{nn}$ constitute the *main diagonal* of $\mathbf{A}$, and the sum of these diagonal components is called the *trace* of the matrix $\mathbf{A}$.

Special Matrices II

A square matrix **B** is *symmetric*, if for all elements $b_{ij} = b_{ji}$. For example, the following square matrix **B** of order 3 is symmetric:

$$\mathbf{B} = \begin{pmatrix} 3 & 5 & 9 \\ 5 & 7 & 6 \\ 9 & 6 & 8 \end{pmatrix}.$$

Its diagonal components are 3, 7 and 8 and its trace is 3 + 7 + 8 = 18.

Special Matrices III

A *diagonal matrix* is a square matrix, in which all elements off the main diagonal are zero. The following square matrix **D** is a diagonal matrix

$$\mathbf{D} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Special Matrices IV

A *scalar matrix* is a diagonal matrix if all elements of the main diagonal are equal to c, where c is a scalar (a number). The following matrix $\mathbf{S}$ is a scalar matrix:

$$\mathbf{S} = \begin{pmatrix} 6 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix}.$$

An *identity matrix* is a scalar matrix with $c = 1$. It is denoted by $\mathbf{I}$. Example:

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Transpose of a Matrix

The *transpose* $\mathbf{A}'$ of a matrix $\mathbf{A}$ is obtained by interchanging rows and columns, i.e., by writing each component $a_{i,j}$ of $\mathbf{A}$ in the position (j, i) of $\mathbf{A}'$.

For example, if

$$\mathbf{A} = \begin{pmatrix} 5 & 8 & 9 \\ 3 & 0 & 2 \end{pmatrix},$$

then

$$\mathbf{A}' = \begin{pmatrix} 5 & 3 \\ 8 & 0 \\ 9 & 2 \end{pmatrix}.$$

Sometimes $\mathbf{A}^{\mathrm{T}}$ is written instead of $\mathbf{A}'$.

Addition of Matrices I

If two matrices **A** and **B** have the same dimensions, the *sum* (and *the subtraction*) is defined to be the matrix obtained by adding the corresponding elements in **A** and **B**. That is,

$$\mathbf{A} \pm \mathbf{B} = (a_{ij}) \pm (b_{ij}) = (a_{ij} \pm b_{ij}).$$

Example:

$$\begin{pmatrix} 5 & 8 & 9 \\ 3 & 0 & 2 \end{pmatrix} + \begin{pmatrix} -6 & 12 & 0 \\ 0 & 3 & 8 \end{pmatrix} = \begin{pmatrix} 5-6 & 8+12 & 9+0 \\ 3+0 & 0+3 & 2+8 \end{pmatrix} = \begin{pmatrix} -1 & 20 & 9 \\ 3 & 3 & 10 \end{pmatrix}$$

and

$$\begin{pmatrix} 5 & 8 & 9 \\ 3 & 0 & 2 \end{pmatrix} - \begin{pmatrix} -6 & 12 & 0 \\ 0 & 3 & 8 \end{pmatrix} = \begin{pmatrix} 5+6 & 8-12 & 9-0 \\ 3-0 & 0-3 & 2-8 \end{pmatrix} = \begin{pmatrix} 11 & -4 & 9 \\ 3 & -3 & -6 \end{pmatrix}.$$

Addition of Matrices II

The commutative and associative law hold for the addition of matrices, that is,

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

and

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}).$$

Scalar Multiplication

The multiplication of a matrix **A** with a real number c is defined by ,

$$c \cdot \mathbf{A} = c \cdot (a_{ij}) = (c \cdot a_{ij}).$$

Example:

$$3 \cdot \begin{pmatrix} 5 & 8 & 9 \\ 3 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 3\cdot 5 & 3\cdot 8 & 3\cdot 9 \\ 3\cdot 3 & 3\cdot 0 & 3\cdot 2 \end{pmatrix} = \begin{pmatrix} 15 & 24 & 27 \\ 9 & 0 & 6 \end{pmatrix} .$$

Both the associative and the commutative laws hold for the scalar multiplication, i.e.

$$c \cdot \mathbf{A} = \mathbf{A} \cdot c \quad \text{and} \quad c_1 \cdot (c_2 \cdot \mathbf{A}) = (c_1 \cdot c_2) \cdot \mathbf{A}.$$

The *division* of a matrix by a scalar $c \neq 0$ is the multiplication of the matrix by the reciprocal of the scalar.

Example: $$\begin{pmatrix} 5 & 8 & 9 \\ 3 & 0 & 2 \end{pmatrix} / 3 = \frac{1}{3} \cdot \begin{pmatrix} 5 & 8 & 9 \\ 3 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 5/3 & 8/3 & 3 \\ 1 & 0 & 2/3 \end{pmatrix}$$

Multiplication of Matrices

The product **AB** of two matrices is defined if the number of columns of **A** is the same as the number of rows of **B**. The new matrix **AB** has the same number of rows as **A** and the same number of columns as **B**. The component of the product matrix in the position (i, j) is obtained by multiplying the first component of row i of matrix **A** with the first component of column j of matrix **B**, the same with the second components etc. The sum of these m products is then the component in row i and column j of the product matrix **AB**. In formulas:

$$\mathbf{A}\,\mathbf{B} = \Big(\sum_{k=1}^{m} a_{ik} b_{kj} \Big) .$$

Multiplication of Matrices: Example

Example of a multiplication of the (2×3)-matrix **A** multiplied with the (3×2)-matrix **B**:

$$\mathbf{A} = \begin{pmatrix} 5 & 8 & 9 \\ 3 & 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 \\ 4 & 3 \\ 1 & 2 \end{pmatrix} = \mathbf{B}$$

$$\mathbf{A}\ \mathbf{B} = \begin{pmatrix} 5\cdot 2+8\cdot 4+9\cdot 1 & 5\cdot 1+8\cdot 3+9\cdot 2 \\ 3\cdot 2+0\cdot 4+2\cdot 1 & 3\cdot 1+0\cdot 3+2\cdot 2 \end{pmatrix} = \begin{pmatrix} 51 & 47 \\ 8 & 7 \end{pmatrix}.$$

Multiplication of Matrices: Special Case

For the multiplication of matrices the associative law holds but the commutative law does not. This can easily be shown: if you consider the product **AB** of two nonsquare matrices, then the product **BA** is only defined if the number of columns of **B** is the same as the number of rows of **A.**

A special case is the multiplication of a matrix **A** with the identity matrix of corresponding order:

$$\mathbf{A}\,\mathbf{I} = \mathbf{I}\mathbf{A} = \mathbf{A}.$$

If **A** is not square, then **I** has to by of a different type on both sides of the equation.

The Inverse of a Matrix

The matrix $\mathbf{A}^{-1}$ is called the *inverse* of a matrix $\mathbf{A}$ if

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I},$$

where $\mathbf{I}$ has the same dimensions as $\mathbf{A}$. Please note that the inverse of a matrix is only defined for square matrices, and that not *all* square matrices have an inverse matrix.

The Inverse of a Matrix: Example

Example: Let $\mathbf{A} = \begin{pmatrix} 3 & 1 & 5 \\ 0 & 1 & 0 \\ 1 & 2 & 2 \end{pmatrix}$ then the inverse is $\mathbf{A}^{-1} = \begin{pmatrix} 2 & 8 & -5 \\ 0 & 1 & 0 \\ -1 & -5 & 3 \end{pmatrix}$.

Even if we do not consider the computation of the inverse of a matrix you can still verify that $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$.

$$\begin{pmatrix} 3 & 1 & 5 \\ 0 & 1 & 0 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} 2 & 8 & -5 \\ 0 & 1 & 0 \\ -1 & -5 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 8 & -5 \\ 0 & 1 & 0 \\ -1 & -5 & 3 \end{pmatrix} \begin{pmatrix} 3 & 1 & 5 \\ 0 & 1 & 0 \\ 1 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The Inverse of a Matrix: Simple Cases I

If the matrix **A** is diagonal,

$$\mathbf{A} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}, \quad \text{then} \quad \mathbf{A}^{-1} = \begin{pmatrix} \frac{1}{a_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{a_{22}} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{a_{nn}} \end{pmatrix}.$$

The Inverse of a Matrix: Simple Cases II

The inverse of a (2×2)-matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

is

$$\mathbf{A}^{-1} = \frac{1}{d} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

This inverse exists only if $d := a_{11} a_{22} - a_{12} a_{21} \neq 0.$

The Rank of a Matrix: Linear Dependency

The vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, all of the same dimension are called *linear independent,* if and only if $\lambda_1 \boldsymbol{a}_1 + \ldots + \lambda_n \boldsymbol{a}_n = \mathbf{0}$ implies $\lambda_1 = \ldots = \lambda_n = 0$, where $\lambda_1, \ldots, \lambda_n$ are real numbers and $\mathbf{0}$ is a vector with the same dimension as the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, and whose components are all zero. Otherwise $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ are called *linearly dependent.*

Example: The two vectors $\boldsymbol{a}_1 = \begin{pmatrix} 84 \\ 91 \\ 119 \\ 161 \end{pmatrix}$ and $\boldsymbol{a}_2 = \begin{pmatrix} 3.6 \\ 3.9 \\ 5.1 \\ 6.9 \end{pmatrix}$

are linearly dependent, because: $3 \cdot \boldsymbol{a}_1 - 70 \cdot \boldsymbol{a}_2 = \mathbf{0}$.

The Rank of a Matrix: Linear Dependency

The *rank* of a matrix **A** in defined by the maximum number of linearly independent column vectors (or row vectors, respectively). The rank of **A** is denoted *rank*(**A**).

A square $(n \times n)$-matrix **A** with *full rank*, i.e. with *rank*(**A**) = n, is called *non-singular.* Otherwise **A** is called *singular.*

The Rank of a Matrix: Linear Dependency II

Theorems

(i) For every non-singular matrix $\mathbf{A}$ there is an inverse matrix $\mathbf{A}^{-1}$. If $\mathbf{A}$ is singular, then there is no inverse of $\mathbf{A}$.

(ii) If $\mathbf{A}$ is an $(n \times m)$-matrix with rank m (*full column rank*), then there is an inverse matrix $(\mathbf{A}'\mathbf{A})^{-1}$ of $\mathbf{A}'\mathbf{A}$. If the rank of $\mathbf{A}$ is less than m, then the inverse $(\mathbf{A}'\mathbf{A})^{-1}$ of $\mathbf{A}'\mathbf{A}$ does not exist..

Rules of Matrix Operations I

(i) $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$ *associative laws*

(ii) $c_1 \cdot (c_2 \cdot \mathbf{A}) = (c_1 \cdot c_2) \cdot \mathbf{A}$

(iii) $(\mathbf{A}\,\mathbf{B})\,\mathbf{C} = \mathbf{A}\,(\mathbf{B}\,\mathbf{C})$

(iv) $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$ *commutative laws*

(v) $c \cdot \mathbf{A} = \mathbf{A} \cdot c$

In general $\mathbf{A}\,\mathbf{B} \neq \mathbf{B}\,\mathbf{A}$, but $trace(\mathbf{A}\,\mathbf{B}) = trace(\mathbf{B}\,\mathbf{A})$, if $\mathbf{A}$ has $n \times m$ dimensions and $\mathbf{B}$ has $m \times n$ dimensions.

Rules of Matrix Operations II

(vi) $(\mathbf{A}+\mathbf{B})\,\mathbf{C} = \mathbf{A}\,\mathbf{C} + \mathbf{B}\,\mathbf{C}$ *distributive laws*

(vii) $\mathbf{C}\,(\mathbf{A}+\mathbf{B}) = \mathbf{C}\,\mathbf{A} + \mathbf{C}\,\mathbf{B}$

(viii) $(\mathbf{A}+\mathbf{B})' = \mathbf{A}' + \mathbf{B}'$ *transposition laws*

(ix) $(\mathbf{A}\mathbf{B})' = \mathbf{B}'\mathbf{A}'$

(x) $\mathbf{A}\,\mathbf{I} = \mathbf{I}\,\mathbf{A} = \mathbf{A}$ *multiplication with the identity matrix*

Expected Value of Multidimensional Random Variables I

Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_q$ be numerical random variables with finite expected values and variances on a common probability space so that they all have a *joint distribution*. These variables may be gathered in an m- and a q-dimensional row vector $\boldsymbol{x}' = (X_1 \;\ldots\; X_m)$ and $\boldsymbol{y}' = (Y_1 \;\ldots\; Y_q)$, respectively. The expected value of such a vector is defined as the vector of the expected values of its components, i.e.:

$$E(\boldsymbol{x}') := E\big((X_1 \;\ldots\; X_m)\big) := [E(X_1) \;\ldots\; E(X_m)]$$

$$E(\boldsymbol{y}') := E\big((Y_1 \;\ldots\; Y_q)\big) := [E(Y_1) \;\ldots\; E(Y_q)].$$

Expected Value of Multidimensional Random Variables II

For these expected values, the following rules of computation may useful:

(i) $E(\boldsymbol{x}) = \boldsymbol{x}$, if $\boldsymbol{x}$ = const (a vector of real constants).

(ii) $E(\mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{y}) = \mathbf{A}\ E(\boldsymbol{x}) + \mathbf{B}\ E(\boldsymbol{y})$.

Covariance Matrix of Multidimensional Random Variables I

The *covariance matrix* $\mathrm{S}_{\boldsymbol{xy}}$ has the covariances of the variables $X_1, ..., X_m$ with the variables $Y_1, ..., Y_q$ as its entries. It is defined by

$$\mathrm{S}_{\boldsymbol{xy}} := E\big([\boldsymbol{x} - E(\boldsymbol{x})]\,[\boldsymbol{y} - E(\boldsymbol{y})]'\big),$$

and may also be written as

$$\mathrm{S}_{\boldsymbol{xy}} = Cov(\boldsymbol{x}, \boldsymbol{y}) = \begin{pmatrix} \mathrm{s}_{X_1Y_1} & \mathrm{s}_{X_1Y_2} & \cdots & \mathrm{s}_{X_1Y_q} \\ \mathrm{s}_{X_2Y_1} & \mathrm{s}_{X_2Y_2} & \cdots & \mathrm{s}_{X_2Y_q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{s}_{X_mY_1} & \mathrm{s}_{X_mY_2} & \cdots & \mathrm{s}_{X_mY_q} \end{pmatrix},$$

where $\mathrm{s}_{X_iY_j} := Cov(X_i, Y_j)$.

Covariance Matrix of Multidimensional Random Variables II

If you consider only one variable Y_1, then is $\boldsymbol{y} = (Y)$ a vector with only one component, the random variable Y. Then $S_{\boldsymbol{xy}}$ is a single column vector, i.e. a column vector

$$S_{\boldsymbol{xy}} =: s_{\boldsymbol{xy}} = Cov(\boldsymbol{x}, \boldsymbol{y}) = \begin{pmatrix} s_{X_1 Y} \\ s_{X_2 Y} \\ \vdots \\ s_{X_m Y} \end{pmatrix}, \quad \text{with } s_{X_i Y} := Cov(X_i, Y)\ .$$

Variance-Covariance Matrix of a Multidimensional Random Variable

A special case is $\boldsymbol{x} = \boldsymbol{y}$. In this case $S_{\boldsymbol{xx}}$ is called the *variance-covariance matrix*. It is defined $S_{\boldsymbol{xx}} := E\big([\boldsymbol{x} - E(\boldsymbol{x})]\,[\boldsymbol{x} - E(\boldsymbol{x})]'\big)$ and may be written:

$$S_{\boldsymbol{xx}} := Var(\boldsymbol{x}) := Cov(\boldsymbol{x}, \boldsymbol{x}) = \begin{pmatrix} s^2_{X_1} & s_{X_1 X_2} & \cdots & s_{X_1 X_m} \\ s_{X_2 X_1} & s^2_{X_2} & \cdots & s_{X_2 X_m} \\ \vdots & \vdots & \ddots & \vdots \\ s_{X_m X_1} & s_{X_m X_2} & \cdots & s^2_{X_m} \end{pmatrix}.$$

The elements on the main diagonal are the variances of the variables $X_1, \ldots, X_m$, because $s_{X_i X_i} := Cov(X_i, X_i) = Var(X_i) = s^2_{X_i}$.

Rules of Computation I

Let $\boldsymbol{x} = (X_1 \ \dots \ X_m)'$ and $\boldsymbol{y} = (Y_1 \ \dots \ Y_q)'$ be vectors of random variables, and **A** and **B** matrices with dimensions $n \times m$ and $n \times q$, respectively, all containing constants. Then:

(i) $Var(\mathbf{A}\,\boldsymbol{x}) = \mathbf{0}$, if $\boldsymbol{x} = const$

(ii) $Var(\mathbf{A}\,\boldsymbol{x}) := Cov(\mathbf{A}\boldsymbol{x}, \mathbf{A}\boldsymbol{x}) = \mathbf{A}\,Var(\boldsymbol{x})\,\mathbf{A}'$

(iii) $Var(\mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{y}) = \mathbf{A}\,Var(\boldsymbol{x})\;\mathbf{A}'$, if $\boldsymbol{y} = const$

(iv) $Var(\mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{y}) = \mathbf{A}\,Var(\boldsymbol{x})\;\mathbf{A}' + \mathbf{B}\,Var(\boldsymbol{y})\;\mathbf{B}' + \mathbf{A}\,Cov(\boldsymbol{x}, \boldsymbol{y})\,\mathbf{B}' + \mathbf{B}\;Cov(\boldsymbol{y}, \boldsymbol{x})\,\mathbf{A}'$

(v) $Cov(\mathbf{A}\;\boldsymbol{x}, \mathbf{B}\boldsymbol{y}) = \mathbf{0}$, if $\boldsymbol{x} = const.$ or $\boldsymbol{y} = const$

(vi) $Cov(\mathbf{A}\;\boldsymbol{x}, \mathbf{B}\;\boldsymbol{y}) = \mathbf{A}\,Cov(\boldsymbol{x}, \boldsymbol{y})\,\mathbf{B}'$

Rules of Computation II

Let $\boldsymbol{x} = (X_1 \ \dots \ X_m)'$, $\boldsymbol{y} = (Y_1 \ \dots \ Y_q)'$, $\boldsymbol{z} = (Z_1 \ \dots \ Z_r)'$ and $\boldsymbol{w} = (W_1 \ \dots \ W_s)'$ be vectors with random variables, and further **A**, **B**, **C** and **D** matrices with the dimensions $n \times m$, $n \times q$, $m \times r$ and $m \times s$, respectively, all containing constants. Then:

(vii) $Cov(\mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{y}, \mathbf{C}\;\boldsymbol{z} + \mathbf{D}\;\boldsymbol{w}) = \mathbf{A}\;Cov(\boldsymbol{x}, \boldsymbol{z})\,\mathbf{C}' + \mathbf{A}\,Cov(\boldsymbol{x}, \boldsymbol{w})\,\mathbf{D}'$
$+ \mathbf{B}\,Cov(\boldsymbol{y}, \boldsymbol{z})\;\mathbf{C}' + \mathbf{B}\,Cov(\boldsymbol{y}, \boldsymbol{w})\,\mathbf{D}'$