Matrix Algebra & Statistics

Contents:

In economic modelling and as well in business, most problems are linear, or approximated by linear models. Such problems are solved by matrix methods. Thus, the syllabus on matrices is essential to these fields.

Variance, Sum of squares and Cross Products

The objective is to account for, or explain, the variation in the data. Variance is the most commonly used measure of dispersion in the data and it is directly proportional to the amount of variation or information in the data. The data below gives two financial ratios, X_1 and X_2, for 12 hypothetical companies.

	Original Data		Mean-Corrected Data		Standardize Data	
Firm	X_1	X_2	x_1	x_2	x_1	x_2
1	13	4	7.92	3.83	1.62	1.11
2	10	6	4.92	5.83	1.01	1.69
3	10	2	4.92	1.83	1.01	0.53
4	8	-2	2.92	-2.17	0.60	-0.63
5	7	4	1.92	3.83	0.39	1.11
6	6	-3	0.92	-3.17	0.19	-0.92
7	5	0	-0.08	-0.17	-0.02	-0.05
8	4	2	-1.08	1.83	-0.22	0.53
9	2	-1	-3.08	-1.17	-0.63	-0.34
10	0	-5	-5.08	-5.17	-1.04	-1.49
11	-1	-1	-6.08	-1.17	-1.24	-0.34
12	-3	-4	-8.08	-4.17	-1.65	-1.20
Mean	5.08	0.17	0	0	0	0
SS			262.92	131.67	11	11
Var	23.90	11.97	23.90	11.97	1	1

The mean of the jth variable is given by:

$$\mu_j = \frac{\sum_{i=1}^{n} X_{ij}}{n},$$

where X_{ij} is the ith observation of the jth variable and n is the number of observations.

The mean-corrected jth variable is denoted by x_j. That is, $x_{ij} = X_{ij} - \mu_j$.

The variance of the jth variable is given by:

$$s_{jj} = \frac{\sum_{i=1}^{n} x_{ij}^2}{n-1} = \frac{\text{SS}}{\text{df}},$$

where SS is the *sum of squares* deviations from the mean and df is the degree of freedom.

The linear relationship, or association, between the two financial ratios can be measured by the covariation between two variables. Covariance, a measure of the covariation between two variables X_i and X_k, is given by:

$$s_{ij} = \frac{\sum_{i=1}^{n} x_{ij}x_{ik}}{n-1} = \frac{\text{SCP}}{\text{df}},$$

where SCP is the *Sum of the cross products* (SCP). That is, the Covariation, is simply the average cross product between two variables for each degree of freedom.

The SS and SCP are usually summarized in a *sum of squares and cross products* (**SSCP**) matrix, while the variance and covariances are usually summarizes in a covariance **S** matrix.

The **SSCP** and **S** of the two financial ratios are given by:

$$\mathbf{SSCP} = \begin{pmatrix} 262.92 & 136.38 \\ 136.38 & 131.67 \end{pmatrix}$$

and

$$\mathbf{S} = \begin{pmatrix} 23.90 & 12.40 \\ 12.40 & 11.97 \end{pmatrix}.$$

Note that the matrices are symmetric. Recall that:

- The variance of a given variable is a measure of its variation in the data. The variances of variables can only be compared if the variables are measured using the same units.
- The Covariance between two variables is a measure of covariation between them. The absolute value of the lower bound covariance is zero implying that the two variables are not linearly associated. However it has no upper bound and this makes it difficult to compare the association between two variables across data sets.

Standardization

Standardized data are obtained by dividing the mean-corrected data by the respective standard deviation (square root of variance). The variances of the standardized variables are always 1. The covariation of standardize variables are always lie between -1 and 1. The value will be:

- 0 (zero) if there is no linear association between the two variables;
- -1 (minus one) if there is perfect inverse linear relationship between the two variables; and
- $+1$ (plus one) if there is a perfect direct linear relationship between the two variables.

The covariance of two standardize variables is called the correlation coefficient. Therefore, the correlation matrix (**R**) is the covariance matrix for standardized data. In the example the correlation matrix is:

$$\mathbf{R} = \begin{pmatrix} 1.00 & 0.733 \\ 0.733 & 1.00 \end{pmatrix}.$$

Matrices

A MATRIX is a two dimensional array of elements. An $m \times n$ (m by n) matrix has m rows and n columns. It is said to have dimension $m \times n$.

Matrix names are usually represented by capital letters and their elements by lower case. An $m \times n$ matrix A containing the real elements a_{ij} is denoted by $A = [a_{ij}] \in \Re^{m \times n}$ and has the general form:

$$A = \begin{array}{cc} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow i\text{th row} \\ \\ \\ \end{array} \end{array}$$

The subscripts of an element a_{ij} indicates that the element is located at the interception of row i and column j, where $1 \leq i \leq m$ and $1 \leq j \leq n$.

A matrix with one row or one column are called row vectors or column vectors, respectively. A row vector R having n real elements is denoted by $R \in \Re^{1 \times n}$ and has the general form $R = (r_1 \ \ldots \ r_n)$.

A column vector C having m real elements is denoted by $C \in \Re^{m\times 1}$ and has the general form

$$C = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}.$$

Generally a C m-elements real vector will be assumed to be a column vector and denoted by $C \in \Re^m$.

SPECIAL TYPES OF MATRICES

- A SQUARE MATRIX is a matrix that has the same number of rows and columns. That is, an $m \times n$ matrix is square if $m = n$. E.g.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 6 & 7 \\ 3 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The variance-covariance and correlation matrices are square.

- An IDENTITY (OR UNIT) MATRIX is a square matrix for which elements along the primary diagonal all equal to 1 and all other elements equal to zero. An $m \times m$ matrix is denoted by I_m. E.g.

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- TRANSPOSED OF A MATRIX.

  Given the matrix $A = [a_{ij}] \in \Re^{m \times n}$ the matrix B is the transpose of A iff $B = [b_{ij}] \in \Re^{n \times m}$ and $b_{ji} = a_{ij}$. The transpose of a matrix A is denoted by A^T. E.g.

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \quad \rightarrow \quad A^T = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \rightarrow \quad I_3^T = I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$R = \begin{pmatrix} 3 & 4 & 5 \end{pmatrix} \rightarrow \quad R^T = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$$

$$R^T = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} \rightarrow \quad (R^T)^T = R = \begin{pmatrix} 3 & 4 & 5 \end{pmatrix}$$

- A matrix A is SYMMETRIC iff $A = A^T$. E.g.

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = A^T \quad \text{and} \quad A = I_m = A^T.$$

The variance-covariance and correlation matrices are symmetric.

$$S = \begin{pmatrix} 2.7 & 6.4 & -3.2 \\ 6.4 & 5.2 & 1.1 \\ -3.2 & 1.1 & 3.1 \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} 1.0 & 0.7 & -0.6 \\ 0.7 & 1.0 & 0.2 \\ -0.6 & 0.2 & 1.0 \end{pmatrix}.$$

- Upper and lower TRIANGULAR MATRICES. A matrix $A = [a_{ij}] \in \Re^{m\times n}$ is upper triangular iff $\forall i > j$, $a_{ij} = 0$. A matrix $B = [a_{ij}] \in \Re^{m\times n}$ is lower triangular iff $\forall i - j < m - n$, $b_{ij} = 0$. E.g. the matrices U and L below are, respectively, upper and lower triangular:

$$U = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 3 & 6 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ -1 & 2 & 2 \end{pmatrix}.$$

Matrix operations

- Two matrices can be added or subtracted iff they have the same dimension. If $A = [a_{ij}] \in \Re^{m \times n}$, $B = [b_{ij}] \in \Re^{m \times n}$ and $C = [c_{ij}] \in \Re^{m \times n}$, then $C = A + B$ implies that $c_{ij} = a_{ij} + b_{ij}$. That is, the elements of the matrix C are found by adding the corresponding elements of the matrices A and B. Similarly if $C = A - B$, then $c_{ij} = a_{ij} - b_{ij}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. E.g.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}.$$

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1-1 \\ 2-(-1) \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}.$$

- A scalar is a real number. The multiplication of a scalar by a matrix is equivalent into multiplying each element of the matrix by the scalar. E.g.

$$\text{If } k = 2 \text{ and } A = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}, \text{ then}$$

$$kA = \begin{pmatrix} 1 \times k & 4 \times k \\ 2 \times k & 5 \times k \\ 3 \times k & 6 \times k \end{pmatrix} = \begin{pmatrix} 2 & 8 \\ 4 & 10 \\ 6 & 12 \end{pmatrix}.$$

- The INNER PRODUCT is an operation between a row and a column vector (in this order) which contain the same number of elements. The inner product is computed by multiplying corresponding elements in the two vectors and algebraically summing. Let $R = [r_i] \in \Re^{1\times n}$ and $C = [c_i] \in \Re^{n\times 1}$. The inner product between R and C is denoted by RC, and is defined as:

$$RC = \begin{pmatrix} r_1 & r_2 & \dots & r_n \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

$$= r_1c_1 + r_2c_2 + \cdots r_nc_n = \sum_{i=1}^{n} r_ic_i.$$

Example.

$$R = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \quad \text{and} \quad C^T = \begin{pmatrix} 4 & 5 & 6 \end{pmatrix}.$$

$$RC = \sum_{i=1}^{3} r_ic_i = 4 + 10 + 18 = 32.$$

$$C^TC = \sum_{i=1}^{3} c_ic_i = \sum_{i=1}^{3} c_i^2 = 4^2 + 5^2 + 6^2 = 77$$

$$\sqrt{C^TC} = \sqrt{77}.$$

- MATRIX MULTIPLICATION. Given $A \in \Re^{m_a \times n_a}$ and $B \in \Re^{m_b \times n_b}$ the matrix product AB is defined iff $n_a = m_b$. That is, if the number of columns of the matrix A equals the number of rows of matrix B. If $C = AB$, then the matrix C has $m_a \times n_b$ dimension.

  For example if $A \in \Re^{5\times 3}$ and $B \in \Re^{3\times 10}$, then $C = AB$ is a 5×10 matrix, i.e. $C \in \Re^{5\times 10}$.

  The dimension of the matrix C derived from:

$$(5 \times 3) \quad (3 \times 10) \rightarrow (5 \times 10).$$

  Let $A \in \Re^{m\times n}$. The dimension of $A^T A$ is $n \times n$.

$$(n \times m) \quad (m \times n) \rightarrow (n \times n).$$

  Let $A \in \Re^{5\times 3}$, $B \in \Re^{3\times 10}$ and $C \in \Re^{10\times 20}$. The dimension of $D = ABC$ is 5×20, i.e. $D \in \Re^{5\times 20}$.

$$(5 \times 3) \quad (3 \times 10) \quad (10 \times 20) \rightarrow (5 \times 20).$$

Computation

Let $A = [a_{ij}] \in \Re^{m\times n}$, $B = [b_{ij}] \in \Re^{n\times k}$ and $C = AB$, where $C = [c_{ij}] \in \Re^{m\times k}$. The element c_{ij} is defined to be the inner product of row i in matrix A and column j in matrix B. That is, for $i = 1,\ldots,m$ and $j = 1,\ldots,k$, the c_{ij} is computed by:

$$c_{ij} = \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} = \sum_{p=1}^{n} a_{ip} b_{pj}.$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$, $B = \begin{pmatrix} 10 & -1 \\ 0 & 3 \end{pmatrix}$ and $C = AB \in \Re^{3\times 2}$.

$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{pmatrix} = \begin{pmatrix} 10 & 5 \\ 30 & 9 \\ 50 & 13 \end{pmatrix}.$$

$$c_{11} = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 10 \\ 0 \end{pmatrix} = 10 \qquad c_{12} = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \end{pmatrix} = 5$$

$$c_{21} = \begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 10 \\ 0 \end{pmatrix} = 30 \qquad c_{22} = \begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \end{pmatrix} = 9$$

$$c_{31} = \begin{pmatrix} 5 & 6 \end{pmatrix} \begin{pmatrix} 10 \\ 0 \end{pmatrix} = 50 \qquad c_{32} = \begin{pmatrix} 5 & 6 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \end{pmatrix} = 13$$

Partitioned matrices

A partitioned matrix contains sub-matrices as elements. The sub-matrices are obtained by partitions of the rows and columns of the original matrix.

E.g. consider the partitioning of $A, B \in \Re^{m\times n}$ as:

$$A = \begin{array}{cc} n_1 \quad\quad n_N & \\ \begin{pmatrix} A_{11} & \dots & A_{1N} \\ \vdots & & \vdots \\ A_{M1} & \dots & A_{MN} \end{pmatrix} & \begin{array}{c} m_1 \\ \\ m_M \end{array} \end{array} \text{ and } B = \begin{array}{cc} n_1 \quad\quad n_N & \\ \begin{pmatrix} B_{11} & \dots & B_{1N} \\ \vdots & & \vdots \\ B_{M1} & \dots & B_{MN} \end{pmatrix} & \begin{array}{c} m_1 \\ \\ m_M \end{array} \end{array},$$

where $n = \sum_{i=1}^{N} n_i$ and $m = \sum_{i=1}^{M} m_i$.

The addition and multiplication of matrices apply directly to the partitioned matrices *provided the sub-matrices are of the appropriate dimension*. E.g. for the addition $A+B$:

$$A+B = \begin{pmatrix} A_{11}+B_{11} & \dots & A_{1N}+B_{MN} \\ \vdots & & \vdots \\ A_{M1}+B_{M1} & \dots & A_{MN}+B_{MN} \end{pmatrix}.$$

Example

Let $A = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ \hline 1 & 1 & 2 & -1 \end{array}\right) \equiv \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$

and

$$B = \left(\begin{array}{cc} -1 & 0 \\ 0 & -1 \\ \hline 3 & 0 \\ 2 & 1 \end{array}\right) \equiv \left(\begin{array}{c} B_{11} \\ \hline B_{21} \end{array}\right).$$

Now,

$$C = AB = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{pmatrix} = \begin{pmatrix} C_{11} \\ C_{21} \end{pmatrix}$$

where

$$C_{11} = A_{11}B_{11} + A_{12}B_{21} = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & -2 \\ -1 & 0 \end{pmatrix} + \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}$$

$$C_{21} = A_{21}B_{11} + A_{22}B_{21} = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 2 & -1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & -1 \end{pmatrix} + \begin{pmatrix} 4 & -1 \end{pmatrix} = \begin{pmatrix} 3 & -2 \end{pmatrix}.$$

Thus, $C = \begin{pmatrix} C_{11} \\ C_{21} \end{pmatrix} = \left(\begin{array}{cc} 2 & -2 \\ 1 & 1 \\ \hline 3 & -2 \end{array}\right).$

Rank of a matrix

- The set $\{a_1, \ldots, a_n\}$ is said to be linearly independent if no a_i can be expressed as a linear combination of the others. This is equivalent to saying that there is no non-null vector $c = (c_1, \ldots, c_n)^T$ such that $\sum_{i=1}^n c_i a_i = 0$. If $\{a_1, \ldots, a_n\}$ are not linear independent, then they are said to be linear dependent.

- The number of independent columns of a matrix A is equal to the number of independent rows. The number of linearly independent columns of a matrix is called *column rank*, hereafter *rank*. It will be denoted by $\text{rank}(A)$.

- The square matrix $A \in \Re^{n\times n}$ is said to be non-singular if the $\text{rank}(A) = n$. Otherwise it is called *singular*.

- Properties

  1. $\text{rank}(A) = \text{rank}(A^T)$.
  2. $\text{rank}(A) = \text{rank}(A^T A) = \text{rank}(AA^T)$.
  3. The rank of A is unchanged by pre- or postmultiplication of A by a non-singular matrix.

Examples

- $\text{rank}(I_n) = n$.
- $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \\ 3 & 1 & 0 \end{pmatrix} \quad \rightarrow \quad \text{rank}(A) = 3.$
- $B = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 3 & 6 & 0 \\ 4 & 8 & 2 \end{pmatrix} \quad \rightarrow \quad \text{rank}(B) = 2.$
- $C = \begin{pmatrix} 1 & 2.001 & 0 \\ 2 & 3.999 & 1 \\ 3 & 6 & 0 \\ 4 & 8 & 2 \end{pmatrix} \quad \rightarrow \quad \text{rank}(C) = 3.$
- $BA = \begin{pmatrix} 5 & 2 & 2 \\ 13 & 5 & 4 \\ 15 & 6 & 6 \\ 26 & 10 & 8 \end{pmatrix} \quad \rightarrow \quad \text{rank}(BA) = 2.$
- $C^T C = \begin{pmatrix} 30.000 & 59.999 & 10.000 \\ 59.999 & 119.996 & 19.999 \\ 10.000 & 19.999 & 5.000 \end{pmatrix} \rightarrow \text{rank}(C^T C) = 3.$
- $CC^T = \begin{pmatrix} 5.004 & 10.002 & 15.006 & 20.008 \\ 10.002 & 20.992 & 29.994 & 41.992 \\ 15.006 & 29.994 & 45.000 & 60.000 \\ 20.008 & 41.992 & 60.000 & 84.000 \end{pmatrix} \rightarrow \text{rank}(CC^T) = 3.$

Trace of a matrix

For a square matrix $A = [a_{ii}] \in \Re^{n\times n}$ the sum of its diagonal elements is called its trace, i.e.

$$\text{trace}(A) = \sum_{i=1}^{n} a_{ii}.$$

Properties

- $\text{trace}(A) = \text{trace}(A^T)$.
- $\text{trace}(AB) = \text{trace}(BA)$.
- $\text{trace}(ABC) = \text{trace}(BCA) = \text{trace}(CAB)$.
- $\text{trace}(A+B) = \text{trace}(B+A) = \text{trace}(A) + \text{trace}(B)$.
- $\text{trace}\left(\sum_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} \text{trace}(A_i)$.
- $\text{trace}(\kappa A) = \kappa\,\text{trace}(A)$.

Examples

Let $A = \begin{pmatrix} 5 & 6 & 7 \\ 3 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \\ 3 & 1 & 0 \end{pmatrix}$ and $A+B = \begin{pmatrix} 6 & 8 & 7 \\ 5 & 2 & 3 \\ 4 & 1 & 1 \end{pmatrix}$.

$\text{trace}(A) = 8, \quad \text{trace}(B) = 1 \quad$ and $\quad \text{trace}(A+B) = 9.$

What is the trace of the variance-covariance and correlation matrix_?

- $AB = \begin{pmatrix} 38 & 17 & 6 \\ 13 & 8 & 2 \\ 4 & 3 & 0 \end{pmatrix} \quad BA = \begin{pmatrix} 11 & 10 & 11 \\ 11 & 12 & 15 \\ 18 & 20 & 23 \end{pmatrix}.$
- Given the vector x, then $\text{trace}(xx^T) = \text{trace}(x^T x) = x^T x$.
- Consider the non-singular matrix Q and $z = Qx$, where $Q^T Q = I$. Then

$$\begin{aligned}\text{trace}(zz^T) &= \text{trace}(Qxx^T Q^T) = \text{trace}(Q^T Qxx^T) \\ &= \text{trace}(Ixx^T) = \text{trace}(x^T x) = x^T x.\end{aligned}$$

Matrix properties

- For any two matrices A and B, it CANNOT be stated that $AB = BA$. E.g.

$$\text{Let} \quad A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 \\ 10 & -1 \end{pmatrix}.$$

$$AB = \begin{pmatrix} 20 & -1 \\ 40 & -1 \end{pmatrix} \qquad \text{and} \qquad BA = \begin{pmatrix} 3 & 4 \\ 7 & 16 \end{pmatrix}.$$

- If A is an $m \times n$ matrix, then $I_m A = A I_m = A$. E.g.

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 4 & 5 \end{pmatrix}.$$

- $\boxed{(AB)^T = B^T A^T.}$ From the latter it follows that:

$$(ABC)^T = ((AB)C)^T = C^T (AB)^T = C^T B^T A^T$$

and

$$(ABCD)^T = D^T C^T B^T A^T.$$

Generally:

$$(A_1 A_2 \cdots A_n)^T = A_n^T \cdots A_2^T A_1^T.$$

- E.g. $\text{trace}(AB) = \text{trace}(BA) = 19$.

Determinants

The determinant is a scalar and derives from the elements of a square matrix. The determinant of the matrix A is denoted by $|A|$.

The determinant of a 1×1 matrix is the value of the single element in the matrix. E.g. If $A = (-3)$, then $|A| = -3$.

The determinant of $A = [a_{ij}] \in \Re^{2\times 2}$ is given by:

$$|A| = a_{11}a_{22} - a_{21}a_{12}.$$

Examples

- $A = \begin{pmatrix} 2 & 4 \\ 3 & 5 \end{pmatrix}$. $\qquad |A| = 2 \times 5 - 3 \times 4 = -2.$

- $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. $\qquad |A| = 1 \times 4 - 2 \times 2 = 0.$

- $A = \begin{pmatrix} -1 & 6 \\ 0 & 3 \end{pmatrix}$. $\qquad |A| = -1 \times 3 + 0 \times 6 = -3.$

Cofactors

In order to find the cofactor of the square matrix A, denoted by A_c, perform the following three steps:

1. Cross off row i and column j of A.
2. Find the determinant of remaining matrix, say d_{ij}.
3. Compute the element of A_c at position (i, j) by $d_{ij}(-1)^{i+j}$.

Example Let $A = \begin{pmatrix} 3 & 1 & 2 \\ -1 & 2 & 4 \\ 3 & -2 & 1 \end{pmatrix}$.

$$d_{11} = \begin{vmatrix} 2 & 4 \\ -2 & 1 \end{vmatrix} = 10; \quad d_{12} = \begin{vmatrix} -1 & 4 \\ 3 & 1 \end{vmatrix} = -13; \quad d_{13} = \begin{vmatrix} -1 & 2 \\ 3 & -2 \end{vmatrix} = -4;$$

$$d_{21} = \begin{vmatrix} 1 & 2 \\ -2 & 1 \end{vmatrix} = 5; \quad d_{22} = \begin{vmatrix} 3 & 2 \\ 3 & 1 \end{vmatrix} = -3; \quad d_{23} = \begin{vmatrix} 3 & 1 \\ 3 & -2 \end{vmatrix} = -9;$$

$$d_{31} = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0; \quad d_{32} = \begin{vmatrix} 3 & 2 \\ -1 & 4 \end{vmatrix} = 14; \quad d_{33} = \begin{vmatrix} 3 & 1 \\ -1 & 2 \end{vmatrix} = 7;$$

Thus, $A_c = \begin{pmatrix} 10 & 13 & -4 \\ -5 & -3 & 9 \\ 0 & -14 & 7 \end{pmatrix}$.

Computing the determinant of matrix

1. Select any row or column of the matrix;

2. Multiply each element in the row (column) by its corresponding cofactors and sum these products.

<u>Example</u>

Let $A = \begin{pmatrix} 3 & 1 & 2 \\ -1 & 2 & 4 \\ 3 & -2 & 1 \end{pmatrix}$ and $A_c = \begin{pmatrix} 10 & 13 & -4 \\ -5 & -3 & 9 \\ 0 & -14 & 7 \end{pmatrix}$.

Selecting row 1:

$$|A| = 3 \times 10 + 1 \times 13 + 2 \times (-4) = 35.$$

Selecting column 2:

$$|A| = 1 \times 13 + 2 \times (-3) - 2 \times (-14) = 35.$$

If only the determinant is required, then compute only the cofactors in the selected row or column. Always select the row or column which simplify your calculations. That is, the row or column with more zeroes or ones.

Example

Compute the determinants of the triangular matrices:

$$U = \begin{pmatrix} 3 & 2 & 1 & 1 \\ 0 & 6 & 0 & 2 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad L = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 6 & 5 & 2 & 4 \end{pmatrix}.$$

$$\begin{aligned}
\det(U) &= 3 \times \det\begin{pmatrix} 6 & 0 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix} + 0 \times \det\begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix} \\
&\quad + 0 \times \det\begin{pmatrix} 2 & 1 & 1 \\ 6 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} + 0 \times \det\begin{pmatrix} 2 & 1 & 1 \\ 6 & 0 & 2 \\ 0 & 2 & 2 \end{pmatrix} \\
&= 3 \times \left(6 \times \det\begin{pmatrix} 2 & 2 \\ 0 & 1 \end{pmatrix} + 0 \times \det\begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix}\right. \\
&\quad \left. + 0 \times \det\begin{pmatrix} 0 & 2 \\ 2 & 2 \end{pmatrix}\right) \\
&= 3 \times 6 \times (2 \times 1 - 0 \times 2) \\
&= 36
\end{aligned}$$

The determinant of a triangular matrix is given by the product of its diagonal elements.

$$\det(L) = 3 \times 2 \times (-1) \times 4 = -24.$$

Properties of the determinant

1. If all elements of any row or column of A equal zero, then $|A| = 0$.

2. If any row (column) is a multiple of another row (column) of A, then $|A| = 0$.

   $\det\begin{pmatrix} 2 & 3 \\ 6 & 9 \end{pmatrix} = 2 \times 9 - 3 \times 6 = 0.$

3. If any two rows or column of A are interchanged, then the sign of the determinant changes.

   $\det\begin{pmatrix} 2 & 3 \\ 1 & 6 \end{pmatrix} = 9. \qquad \det\begin{pmatrix} 1 & 6 \\ 2 & 3 \end{pmatrix} = -9.$

4. If a row, or column, of A is multiplied by the scalar κ, then the determinant of the modified matrix is given by $\kappa\,|A|$.

5. If any multiple of one row (column) is added to another row (column), the value of the determinant is unchanged.

   $\det\begin{pmatrix} 2 & 3 \\ 1 & 6 \end{pmatrix} = 9. \qquad \det\begin{pmatrix} 2 & 11 \\ 1 & 10 \end{pmatrix} = 9.$

Inverse of a matrix

The relationship between a square matrix A and its inverse, denoted by A^{-1} (inverse of A), is that:

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

Note that

- The matrix A must be square.
- The dimensions of A and A^{-1} are the same.
- Only non-singular matrices have an inverse.

Computing the inverse of a a matrix

There are many methods to compute the inverse of a matrix. Consider the use of the cofactor method:

1. Compute the cofactor A_c from A.
2. Compute the *Adjoint matrix* A_j by $A_j = A_c^T$.
3. For $|A| \neq 0$, the inverse of A is given by:

$$A^{-1} = \frac{1}{|A|} A_j.$$

Examples

- Let $A = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$, where
  $\det(A) = 2 \times (-1) - 1 \times 3 = -5$.

  The $A_c = \begin{pmatrix} -1 & -3 \\ -1 & 2 \end{pmatrix}$ and $A_j = A_c^T = \begin{pmatrix} -1 & -1 \\ -3 & 2 \end{pmatrix}$.

  Thus, $A^{-1} = \dfrac{1}{\det(A)} A_j = -\dfrac{1}{5} \begin{pmatrix} -1 & -1 \\ -3 & 2 \end{pmatrix}$.

  Note that $A^{-1}A = AA^{-1} = I_2$.

- Let $A = \begin{pmatrix} 3 & 1 & 2 \\ -1 & 2 & 4 \\ 3 & -2 & 1 \end{pmatrix}$.

  Now, $|A| = 35$, $A_c = \begin{pmatrix} 10 & 13 & -4 \\ -5 & -3 & 9 \\ 0 & -14 & 7 \end{pmatrix}$ and the Adjoint matrix is given by $A_j = A_c^T = \begin{pmatrix} 10 & -5 & 0 \\ 13 & -3 & -14 \\ -4 & 9 & 7 \end{pmatrix}$.

  Thus, $A^{-1} = \dfrac{1}{35} \begin{pmatrix} 10 & -5 & 0 \\ 13 & -3 & -14 \\ -4 & 9 & 7 \end{pmatrix}$.

  Note that $A^{-1}A = AA^{-1} = I_3$.

Gaussian reduction procedure.

The Gaussian reduction procedure can be used to compute the inverse of the matrix. Consider the $m \times m$ matrix A. Construct the augmented matrix $(A \mid I_m)$. Apply the Gaussian elimination method to the whole augmented matrix so that A is transformed to I_m. The resulting matrix will give $(I_m \mid A^{-1})$.

The Gaussian elimination method transforms $(A \mid I_m)$ to $(I_m \mid A^{-1})$ by applying two basic operations:

1. Rows can be multiplied by a non zero constant; and

2. non zero multiples of one row can be added to another row.

Properties of the inverse

- The inverse of a symmetric matrix is also symmetric.

Let $S = \begin{pmatrix} 2.7 & 6.4 & -3.2 \\ 6.4 & 5.2 & 1.1 \\ -3.2 & 1.1 & 3.1 \end{pmatrix}$. $\quad S^{-1} = \begin{pmatrix} -0.08 & 0.13 & -0.13 \\ 0.13 & 0.01 & 0.13 \\ -0.13 & 0.13 & 0.15 \end{pmatrix}$.

- The inverse of A^T is the transpose of A^{-1}. That is,

$$\boxed{(A^T)^{-1} = (A^{-1})^T = A^{-T}.}$$

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \\ 3 & 1 & 0 \end{pmatrix}. \qquad A^{-1} = \frac{1}{5}\begin{pmatrix} -1 & 0 & 2 \\ 3 & 0 & -1 \\ 2 & 5 & -4 \end{pmatrix}.$$

$$A^T = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \qquad (A^T)^{-1} = \frac{1}{5}\begin{pmatrix} -1 & 3 & 2 \\ 0 & 0 & 5 \\ 2 & -1 & -4 \end{pmatrix}.$$

- Let $A_1, \ldots, A_n \in \Re^{n\times n}$. Then,

$$\boxed{(A_1 A_2 \cdots A_n)^{-1} = (A_n^{-1} \cdots A_2^{-1} A_1^{-1}).}$$

$$B = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 2 & 1 \\ 2 & 0 & 1 \end{pmatrix}. \quad AB = \begin{pmatrix} -2 & 4 & 3 \\ -2 & 0 & 3 \\ -6 & 2 & 4 \end{pmatrix}. \quad B^{-1} = \frac{1}{4}\begin{pmatrix} -1 & -0 & 1 \\ -1 & 2 & -1 \\ 2 & 0 & 2 \end{pmatrix}.$$

$$(AB)^{-1} = \frac{1}{20}\begin{pmatrix} 3 & 5 & -6 \\ 5 & -5 & 0 \\ 2 & 10 & -4 \end{pmatrix} = B^{-1}A^{-1}.$$

- If c is a non zero scalar, then

$$\boxed{(cA)^{-1} = \frac{1}{c} A^{-1}.}$$

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \\ 3 & 1 & 0 \end{pmatrix}, \quad c = \frac{1}{5} \quad \text{and} \quad (cA)^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 3 & 0 & -1 \\ 2 & 5 & -4 \end{pmatrix}.$$

- The inverse of a diagonal matrix is a diagonal matrix consisting of the reciprocals of the original elements. E.g.

$$D = \begin{pmatrix} s_{11} & 0 & \dots & 0 \\ 0 & s_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & s_{nn} \end{pmatrix}; \quad D^{-1} = \begin{pmatrix} \frac{1}{s_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{s_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{s_{nn}} \end{pmatrix}$$

$D = \text{diag}(1, 10^4, 10^{15}, 10^{16})$.

$D^{-1} = \text{diag}(1, 10^{-4}, 10^{-15}, 10^{-16})$.

Computed by Octave (Matlab):

$D^{-1} = \text{diag}(1, 0.00010, 0.00000, 0.00000)$.

$\text{rank}(D) = 4 \quad \text{and} \quad \text{rank}(D^{-1}) = 3$.

- The inverse of a triangular matrix is also triangular.

  Let

$$U = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 3 & 6 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ -1 & 2 & 2 \end{pmatrix}.$$

  Then,

$$U^{-1} = \frac{1}{3}\begin{pmatrix} 3 & 2 & 12 \\ 0 & 1 & 6 \\ 0 & 0 & -3 \end{pmatrix} \quad \text{and} \quad L^{-1} = \frac{1}{6}\begin{pmatrix} 6 & 0 & 0 \\ -4 & 2 & 0 \\ 7 & -2 & 3 \end{pmatrix}.$$

System of equations

Consider the $n \times n$ system of equations having the form:

$$\begin{array}{lllll} a_{11}x_1 & +a_{12}x_2 & +\dots & +a_{1n}x_n & = b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\dots & +a_{2n}x_n & = b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & +\dots & +a_{nn}x_n & = b_n \end{array}$$

can be written in a matrix form as

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{1}$$

$$\text{or} \quad Ax = b \tag{2}$$

Assume that the equations are linear independent, that is, A is not singular (it has no inverse). Premultiply both sides of (2) by A^{-1} it gives:

$$A^{-1}Ax = A^{-1}b \quad \text{or} \quad x = A^{-1}b$$

since $A^{-1}Ax = I_n x = x$.

Thus, the solution of (2) is given by $\boxed{x = A^{-1}b}$.

Example

Consider the system of equations:

$$\begin{array}{cccl} x_1 & +x_2 & +x_3 & = 2 \\ 2x_1 & +3x_2 & +4x_3 & = 4 \\ 2x_1 & -x_2 & +x_3 & = 9 \end{array}$$

In matrix form the latter 3×3 system can be written as:

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 2 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 9 \end{pmatrix}$$

or $\quad Ax = b$

Step 1 Compute the inverse of A:

$$A^{-1} = \frac{1}{5} \begin{pmatrix} 7 & -2 & 1 \\ 6 & -1 & -2 \\ -8 & 3 & 1 \end{pmatrix}$$

Step 2 Compute $A^{-1}b$:

$$A^{-1}b = \frac{1}{5} \begin{pmatrix} 7 & -2 & 1 \\ 6 & -1 & -2 \\ -8 & 3 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \\ 9 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}.$$

The solution of the system of equations is given by $x = A^{-1}b = (3 \quad -2 \quad 1)^T$. That is,

$$\boxed{x_1 = 3, \quad x_2 = -2 \quad \text{and} \quad x_3 = 1.}$$

Orthogonal matrices

A square matrix $Q \in \Re^{m\times m}$ is orthogonal iff

$$\boxed{Q^T Q = QQ^T = I_m.}$$

Notice that the inverse of Q is given by Q^T.

Examples of orthogonal matrices:

$$I_m, \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

<u>Property</u>

- It preserves the norm (inner product) of a vector. That is, If $z = Qx$ and Q is orthogonal, then $z^T z = x^T x$.

Note $z^T z = (Qx)^T (Qx) = x^T Q^T Qx = x^T Ix = x^T x$.

$$x = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 0.5 & 0.866 \\ -0.866 & 0.5 \end{pmatrix}.$$

$$z = Qx = \begin{pmatrix} 2.098 \\ 2.366 \end{pmatrix} \quad \text{and} \quad x^T x = 10 = z^T z.$$

Cholesky Decomposition

The Cholesky decomposition of a symmetric positive definite $n \times n$ matrix A, is given by

$$A = LL^T,$$

where $L \in \Re^{n\times n}$ is lower triangular and non-singular. E.g.

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{12} & A_{22} & A_{23} \\ A_{13} & A_{23} & A_{33} \end{pmatrix} = \begin{pmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{pmatrix} \begin{pmatrix} L_{11} & L_{21} & L_{31} \\ 0 & L_{22} & L_{32} \\ 0 & 0 & L_{33} \end{pmatrix}.$$

Let $\quad A = \begin{pmatrix} 5 & 2 & 3 \\ 2 & 10 & 1 \\ 3 & 1 & 2 \end{pmatrix}$

The Cholesky Decomposition of $A = LL^T$ is given by:

$$\begin{pmatrix} 2.24 & 0 & 0 \\ 0.89 & 3.03 & 0 \\ 1.34 & -0.07 & 0.44 \end{pmatrix} \begin{pmatrix} 2.24 & 0.89 & 1.34 \\ 0 & 3.03 & -0.07 \\ 0 & 0 & 0.44 \end{pmatrix}$$

Solve the matrix problem $Ax = b$, where A is symmetric and has Cholesky decomposition $A = LL^T$.

Notice that $L(L^T x) = b$ is equivalent to $Lz = b$, where $L^T x = z$. That is, the solution of $Ax = b$ comes in three steps:

1. Compute the Cholesky decomposition $A = LL^T$.
2. Solve the lower-triangular system $Lz = b$ for z.
3. Solve the upper-triangular system $L^T x = z$ for x.

<u>Example:</u>

Solve $Ax = b$, where $A = \begin{pmatrix} 5 & 2 & 3 \\ 2 & 10 & 1 \\ 3 & 1 & 2 \end{pmatrix}$ and $b = \begin{pmatrix} 7 \\ -16 \\ 5 \end{pmatrix}$.

1. $A = LL^T$, where $L = \begin{pmatrix} 2.24 & 0 & 0 \\ 0.89 & 3.03 & 0 \\ 1.34 & -0.07 & 0.44 \end{pmatrix}$
2. Solve $Lz = b$ which gives $z = \begin{pmatrix} 3.13 \\ 6.19 \\ 0.88 \end{pmatrix}$
3. Solve $L^T x = z$ which gives $x = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$.

Example

Consider the two financial ratios:

X_1	13	10	10	8	7	6	5	4	2	0	-1	-3
X_2	4	6	2	-2	4	-3	0	2	-1	-5	-1	-4

which have a variance-covariance matrix given by:

$$\mathbf{S} = \begin{pmatrix} 23.90 & 12.40 \\ 12.40 & 11.97 \end{pmatrix}.$$

Thus, if $X = (X_1 \quad X_2)$, then $\text{Var}(X) = S$.

Find the variance-covariance matrix of XL^{-T}, where L is the Cholesky factor of X. I.e.

$$L = \begin{pmatrix} 4.89 & 0 \\ 2.51 & 2.38 \end{pmatrix} \quad \text{and} \quad L^{-1} = \begin{pmatrix} 0.20 & 0 \\ -0.22 & 0.42 \end{pmatrix}.$$

Now, XL^{-T} gives:

X_1	2.7	2.0	2.0	1.6	1.4	1.2	1.0	0.8	0.4	0	-0.2	-0.6
X_2	-1.1	0.4	1.3	2.6	0.2	-2.5	-1.1	-0.0	-0.8	-2.1	-0.2	-1.0

Its variance-covariance matrix is given by:

$$\mathbf{S_T} = \begin{pmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{pmatrix}.$$

Note that

$$\begin{aligned} \text{Var}(XL^{-T}) &= L^{-1}\text{Var}(X)L^{-T} = L^{-1}SL^{-T} \\ &= L^{-1}LL^{T}L^{-T} = I_2. \end{aligned}$$

The Eigenvalue problem

Let A be a square matrix of order $n \times n$, $x \neq 0$ is an n-element column vector and λ is a scalar. The EIGENVALUE PROBLEM is the solution of:

$$Ax = \lambda x.$$

The solution come in pairs: to each λ corresponds an x vector. The λ's are known as eigenvalues (or latent, or characteristic roots) and the x's as eigenvectors (or latent, or, characteristic vectors).

In matrix format the Eigenvalue problem can be written as:

$$(A - \lambda I_n)x = 0$$

In order for $x \neq 0$ it implies that

$$|A - \lambda I_n| = 0.$$

The latter is known as the characteristic equation for A. It gives a polynomial equation in the unknown λ.

Example

Let $A = \begin{pmatrix} 1 & 0 \\ 1 & 3 \end{pmatrix}$ so that $A - \lambda I_2 = \begin{pmatrix} 1-\lambda & 0 \\ 1 & 3-\lambda \end{pmatrix}$.

Now, $|A - \lambda I_2| = (1-\lambda)(3-\lambda)$.
Thus, $\lambda_1 = 1$ and $\lambda_2 = 3$ are the eigenvalues of A.

For the eigenvalue $\lambda_1 = 1$ we have $Ax = \lambda_1 x$:

$$\begin{pmatrix} 1 & 0 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \text{or} \quad \begin{array}{l} x_1 = x_1 \\ x_1 = -2x_2. \end{array}$$

Thus, an eigenvector of A corresponding to the eigenvalue $\lambda_1 = 1$ is given by $x = (-2 \quad 1)^T$.
Normalizing x, i.e. dividing each of its entries by $\sqrt{x^T x}$, it gives the eigenvector

$$\frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

An eigenvector associated with the eigenvalue of $\lambda_2 = 3$ is given by $(0 \quad 1)^T$.

Properties of eigenvalues and eigenvectors

Given an $m \times m$ SYMMETRIC matrix, e.g. the variance-covariance matrix: $A = \begin{pmatrix} 5 & 2 & 3 \\ 2 & 10 & 1 \\ 3 & 1 & 2 \end{pmatrix}$

- The eigenvalues are real.

  The eigenvalue of A are given by:
  $\lambda_1 = 0.14, \quad \lambda_2 = 5.70 \quad$ and $\quad \lambda_3 = 11.16.$

- Eigenvectors corresponding to distinct eigenvalues are pairwise orthogonal[a]. I.e. if x_1 and x_2 are the eigenvectors corresponding to the eigenvalues λ_1 and λ_2 $(\lambda_1 \neq \lambda_2)$, then $x_1^T x_2 = 0$.

  The eigenvectors of A are given by the columns of $X = (x_1, x_2, x_3)$, where

$$X = \begin{pmatrix} -0.532 & 0.747 & 0.400 \\ 0.022 & -0.459 & 0.888 \\ 0.847 & 0.481 & 0.228 \end{pmatrix}$$

and

$$X^T X = X X^T = I_3.$$

[a] Notice that $Ax_1 = \lambda_1 x_1$, and after premultiplication by x_2^T it gives $x_2^T A x_1 = \lambda_1 x_2^T x_1$. Similarly, $x_1^T A x_2 = \lambda_2 x_1^T x_2$. Since $x_2^T A x_1 = x_1^T A x_2$ it follows that $\lambda_1 x_2^T x_1 = \lambda_2 x_1^T x_2$ and thus, $x_1^T x_2 = 0$.

- The orthogonal matrix of eigenvectors diagonalizes[b] A. That is,

$$\boxed{X^T A X = \Lambda,}$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ and $X = (x_1 \ldots x_m)$.

$$X^T A X = \begin{pmatrix} 0.14 & 0 & 0 \\ 0 & 5.70 & 0 \\ 0 & 0 & 11.16 \end{pmatrix} = \Lambda$$

- The matrix A is singular if one of its eigenvalues is zero.
- The rank of A is equal to the number of non-zero eigenvalues.

- $A^2 = AA = X\Lambda^2 X^T$ and generally $A^n = X\Lambda^n X^T$.

$$A^2 = \begin{pmatrix} 38 & 33 & 23 \\ 33 & 105 & 18 \\ 23 & 18 & 14 \end{pmatrix} \quad \text{and} \quad \Lambda^2 = \begin{pmatrix} 0.02 & 0 & 0 \\ 0 & 32.51 & 0 \\ 0 & 0 & 124.47 \end{pmatrix}.$$

- $A^{-1} = X\Lambda^{-1}X^T$ since $(X\Lambda X^T)^{-1} = X\Lambda^{-1}X^T$.

$$A^{-1} = \frac{1}{9}\begin{pmatrix} 19 & -1 & -28 \\ -1 & 1 & 1 \\ -28 & 1 & 46 \end{pmatrix} \text{ and } \Lambda^{-1} = \begin{pmatrix} 7.07 & 0 & 0 \\ 0 & 0.18 & 0 \\ 0 & 0 & 0.09 \end{pmatrix}.$$

---

[b] The Eigenvalue problem in matrix form is equivalent to $AX = X\Lambda$. Premultiplying by X^T it gives $X^T AX = X^T X\Lambda$ which is equivalent to $X^T AX = \Lambda$ since $X^T X = I_m$.

- Let $A = A^{-1/2}A^{-1/2}$. Then, $A^{-1/2} = X\Lambda^{-1/2}X^T$.

- If A is triangular then its diagonal elements are also its eigenvalues.

  E.g. If $\quad U = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 3 & 6 \\ 0 & 0 & -1 \end{pmatrix}, \quad$ then

  $\lambda_1 = 1, \quad ,\lambda_2 = 3, \quad ,\lambda_3 = -1.$

- If A is orthogonal then its eigenvalues are either 1 or -1.

- If B is non-singular, then BAB^{-1} and A have the same eigenvalues.

Conditioning of $A^T A$

- The eigenvalues of the $n \times n$ symmetric matrix $A^T A$ are given by $\lambda_1, \ldots, \lambda_n$ (in descreasing order).
- The ratio $\kappa(A) = \sqrt{\lambda_1}/\sqrt{\lambda_n}$ is called the condition number of A.

Example

$$A = \begin{pmatrix} -6 & -12 & 8 \\ 2 & 12 & -11 \\ -6 & -17 & 10 \\ 19 & 3 & 6 \\ -9 & 6 & 15 \end{pmatrix}, \qquad \sqrt{\Lambda} = \begin{pmatrix} 31.71 & 0 & 0 \\ 0 & 19.80 & 0 \\ 0 & 0 & 16.99 \end{pmatrix}.$$

The Condition number of A is given by

$$\sqrt{\lambda_1}/\sqrt{\lambda_3} = 31.77/16.99 = 1.87.$$

Consider the matrices:

$$A_0 = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 3 & 6 & 0 \\ 4 & 8 & 2 \end{pmatrix}, A_1 = \begin{pmatrix} 1 & 2.01 & 0 \\ 2 & 3.99 & 1 \\ 3 & 6 & 0 \\ 4 & 8 & 2 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 2.1 & 0 \\ 2 & 3.9 & 1 \\ 3 & 6 & 0 \\ 4 & 8 & 2 \end{pmatrix}, A_3 = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 4 & 1 \\ 3 & 9 & 0 \\ 4 & 16 & 2 \end{pmatrix}.$$

Cond(A_0)=8.82e+16, Cond(A_1)=2124.5, Cond(A_2)=213.02, Cond(A_3)=17.77, and Cond(I_n)=n.