[MATH-05] Matrices

Miguel-Angel Canela
Associate Professor, IESE Business School

Matrices

A matrix is a two-dimensional (rectangular) array. The words matrix and table are more or less synonymous, but matrix is preferred in Mathematics. The terms of a matrix are arranged in **rows** and **columns**. An (n, m)–matrix is a matrix with n rows and m columns. The pair (n, m) is the **dimension** of the matrix. I only use here matrices whose terms are real numbers, using upper case boldface to distinguish them from numbers and vectors. The position of a term in a matrix of more than one column is indicated by two subindexes. The first subscript indicates the row and the second one the column. Thus, an (n, m)–matrix can be written as

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

The **transpose** of an (n, m)–matrix $\mathbf{A}$ is the (m, n)–matrix obtained by interchanging rows and columns. We denote the transpose by $\mathbf{A}^\mathsf{T}$ (some use $\mathbf{A}'$), so the term in the ith row and the jth column of $\mathbf{A}^\mathsf{T}$ is a_{ji}. n-vectors are usually identified with $(n, 1)$-matrices in documents and the blackboard. Then, the transpose of a vector $\mathbf{x}$ is a one-row matrix,

$$\mathbf{x}^\mathsf{T} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}.$$

Nevertheless, in the computer, vectors and matrices can be objects of different type. For intance, R distinguishes between matrices and column vectors. An (n, n)–matrix is called a **square matrix**. Then the elements a_{ii} are called the **diagonal** terms. The diagonal of $\mathbf{A}$ is the vector of diagonal terms, and the **trace** $\mathrm{tr}(\mathbf{A})$ is the sum of the diagonal terms.

Some special properties of square matrices are:

- $\mathbf{A}$ is **symmetric** when $\mathbf{A} = \mathbf{A}^\mathsf{T}$. Equivalently, when $a_{ij} = a_{ji}$ for any pair i, j.
- $\mathbf{A}$ is **upper triangular** when all the terms below the diagonal are zero, i.e. when $a_{ij} = 0$ for $i > j$. Lower triangular matrices are defined in a similar way.
- $\mathbf{A}$ is **diagonal** when it is both upper and lower triangular, i.e. when $a_{ij} = 0$ for $i \neq j$.

The definitions of the sum and the product by scalars extend those given for vectors. The sum can only be defined for matrices with the same dimension.

The rank of a matrix

The rank of a matrix is the maximum number of linearly independent column vectors. It can be seen that the rank of a matrix coincides with the rank of its transpose, so that the rank can also be defined as the maximum number of linearly independent rows. Therefore, therank cannot exceed the minimum of the number of rows and the number of columns. When the rank is equal to this minimum, we have a **full-rank matrix**.

The rank does not change if one sums to a row (or column) a linear combination of the rest. This fact is the basis of an elementary method for calculating the rank of a matrix, which consists in transforming it into a new matrix with the same rank, but with a triangle of zeros in one of its four corners, which renders trivial to find the rank. Let me show how this works by means of an example.

Example 1. In order to find the rank of the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & -1 & 4 \\ 2 & 4 & 3 & 9 \\ -1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 5 \end{bmatrix},$$

we sum to the second row the first one, multiplied by -2, and to the third row the first one, obtaining

$$\mathbf{A}^{(1)} = \begin{bmatrix} 1 & 3 & -1 & 4 \\ 0 & -2 & 5 & 1 \\ 0 & 4 & 0 & 5 \\ 0 & 2 & 3 & 5 \end{bmatrix}.$$

$\mathbf{A}^{(1)}$ and $\mathbf{A}$ have the same rank. Next, we sum to the third row the second one, multiplied by 2, and to the fourth row the second one, obtaining

$$\mathbf{A}^{(2)} = \begin{bmatrix} 1 & 3 & -1 & 4 \\ 0 & -2 & 5 & 1 \\ 0 & 0 & 10 & 7 \\ 0 & 0 & 8 & 6 \end{bmatrix},$$

again with the same rank. Finally, we sum to the fourth row the third one, multiplied by $-4/5$, and we obtain

$$\mathbf{A}^{(3)} = \begin{bmatrix} 1 & 3 & -1 & 4 \\ 0 & -2 & 5 & 1 \\ 0 & 0 & 10 & 7 \\ 0 & 0 & 0 & 2/5 \end{bmatrix},$$

whose rank, due to the position of the zeros and ones, must be 4. Therefore, $\mathbf{A}$ has rank 4.

Determinants

Let $\mathbf{A}$ be an (n, n)–matrix:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

The **determinant** of $\mathbf{A}$, denoted by $\det \mathbf{A}$, is a number that results from summing products of terms of $\mathbf{A}$. It is also denoted by $|\mathbf{A}|$, but I avoid here this notation, leaving the bars for the absolute value of a real number and the modulus of a complex number. Nevertheless, when replacing in a matrix the brackets by bars, I mean the determinant of this matrix, i.e.

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

The determinant of $\mathbf{A}$ can be defined as the sum of all the possible products of n terms of $\mathbf{A}$, taking one term from each row and one term from each column (there are $n!$ such products), with the same or the opposite sign, according to the following rule: the sign of a product $a_{1m_1} \cdots a_{nm_n}$ is preserved when the number of inversions in the sequence $\big(m_1, m_2, \ldots, m_n\big)$ is even, and changed when the number of inversions is odd.

For n small, the determinant can be calculated easily. For $n = 2$, the rule is

$$\det \mathbf{A} = a_{11}a_{22} - a_{12}a_{21},$$

whereas, for $n = 3$,

$$\det \mathbf{A} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

¶ You probably learned to pick the factors of these terms using a starlike path. Schoolteachers called this trick the rule of Sarrus.

Example 2. The following calculation illustrates these rules.

$$\begin{vmatrix} 1 & 3 & 0 \\ -1 & 1 & 1 \\ 0 & 5 & 2 \end{vmatrix} = 1\times1\times2+(-1)\times5\times0+3\times1\times0-0\times1\times0-(-1)\times3\times2-1\times5\times1 = 2+6-5 = 3.$$

Properties of the determinant

For $n > 3$, the following properties of the determinant are useful:

- The determinant of a triangular matrix is the product of the diagonal terms.
- A matrix and its transpose have the same determinant: $\det \mathbf{A} = \det(\mathbf{A}^\intercal)$.
- Interchanging two rows (columns) changes the sign of the determinant.
- If two rows (columns) are equal, the determinant is null.
- If one sums to a row (column) a linear combination of the others, the determinant does not change.
- If a row (column) is a linear combination of the others, the determinant is null, and conversely.
- For an (n, n)–matrix $\mathbf{A}$, $\operatorname{rank} \mathbf{A} = n$ if and only if $\det \mathbf{A} \neq 0$.
- The rank of a matrix coincides with the dimension of the biggest squared sub-matrix whose determinant is different of zero.

Example 3. Due to these properties (see Example 1),

$$\begin{vmatrix} 1 & 3 & -1 & 4 \\ 2 & 4 & 3 & 9 \\ -1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 3 & -1 & 4 \\ 0 & -2 & 5 & 1 \\ 0 & 0 & 10 & 7 \\ 0 & 0 & 0 & 2/5 \end{vmatrix} = -8.$$

Homework

A. Find the rank of the matrix

$$\mathbf{A} = \begin{bmatrix} 5 & 2 & 2 & 1 \\ 4 & 3 & 3 & 2 \\ 6 & 1 & 2 & 1 \\ 5 & 2 & 3 & 2 \\ 3 & 4 & 4 & 3 \end{bmatrix}.$$

B. Find the rank and the determinant of

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 2 & 0 & 2 \\ 3 & 0 & 1 & 2 \\ 0 & 4 & 4 & 3 \end{bmatrix}.$$

C. Calculate

$$\begin{vmatrix} 1 & 1 & 0 & 1 \\ 2 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \\ 4 & 4 & 2 & 2 \end{vmatrix}.$$