10 The algebra of matrices

(Applied Mathematics — FAPPZ)

Petr Gurka

(Updated on November 28, 2011)

Contents

1 Vector space of matrices of the type (m, n)

Recall that a matrix $\boldsymbol{A}$ of type (m, n) we call a rectangular array of mn real numbers in m rows and n columns,

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

1.1 Transposed matrix

Definition 1. Let $\boldsymbol{A}$ be a matrix of the type (m, n). The matrix $\boldsymbol{A}^T$ of the type (n, m),

$$\boldsymbol{A}^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix},$$

which is obtained from the matrix $\boldsymbol{A}$ by interchanging the rows and the columns is called a *transposed matrix* to the matrix $\boldsymbol{A}$. Elements a_{ii}, $i = 1, 2, \dots, k$, $k = \min\{m, n\}$, are elements of the *main diagonal* of the matrix $\boldsymbol{A}$ (and also the main diagonal of the matrix $\boldsymbol{A}^T$).

1.2 Sum, difference and scalar multiples of matrices

Definition 2. By the symbol $\mathbf{M}_{m,n}$ we denote the vector space of all matrices of the type (m,n) with the following operations.

By the *sum* of two matrices of the type (m,n),

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix},$$

we mean the matrix

$$\boldsymbol{A} + \boldsymbol{B} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix}.$$

If $r \in \mathbb{R}$, then by the r-*multiple of the matrix* $\boldsymbol{A}$ we mean

$$r\boldsymbol{A} = \begin{pmatrix} r\,a_{11} & r\,a_{12} & \dots & r\,a_{1n} \\ r\,a_{21} & r\,a_{22} & \dots & r\,a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r\,a_{m1} & r\,a_{m2} & \dots & r\,a_{mn} \end{pmatrix}.$$

If $\boldsymbol{O}$ is the zero matrix of the type (m,n) then, for any matrix $\boldsymbol{A}$ of the type (m,n), $\boldsymbol{A} + \boldsymbol{O} = \boldsymbol{A}$. Further we define the *difference of matrices* $\boldsymbol{A} - \boldsymbol{B}$ as $\boldsymbol{A} + (-1)\boldsymbol{B}$.

2 Matrix equations

2.1 Multiplication of matrices

Definition 3. Let $\boldsymbol{A} = \big(a_{ij}\big)$ be a matrix of the type (m,p) and $\boldsymbol{B} = \big(b_{ij}\big)$ be a matrix of the type (p,n). By the *product* $\boldsymbol{A}\,\boldsymbol{B}$ of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ we mean the matix $\boldsymbol{C} = \big(c_{ij}\big)$ of the type (m,n) whose each entry c_{ij} is an inner product if i $^{\text{th}}$ *row* of the matrix $\boldsymbol{A}$ and j $^{\text{th}}$ *column* of the matrix $\boldsymbol{B}$, that is,

$$c_{ij} = \mathbf{a}_i \cdot \widetilde{\mathbf{b}}_j = \sum_{k=1}^{p} a_{ik} \cdot b_{kj}, \quad \text{where } i = 1, 2, \dots, m, \;\; j = 1, 2, \dots, n.$$

ATTENTION !!!
Multiplication is not, in general, *commutative*, that is, $\boldsymbol{A}\,\boldsymbol{B}$ need not be the same as $\boldsymbol{B}\,\boldsymbol{A}$.

- $\boldsymbol{A}\,\boldsymbol{B}$ and $\boldsymbol{B}\,\boldsymbol{A}$ may be matrices of different types.
- It may happen that one of the products $\boldsymbol{A}\,\boldsymbol{B}$ and $\boldsymbol{B}\,\boldsymbol{A}$ is defined and the other one is not defined.

- When $\boldsymbol{A}$ and $\boldsymbol{B}$ are matrices of the same type (n, n), then the products $\boldsymbol{A}\,\boldsymbol{B}$ and $\boldsymbol{B}\,\boldsymbol{A}$ are defined and they are both of the type (n, n). However, they may be different.

2.2 Identity matrix

Definition 4.
- A matrix $\boldsymbol{A}$ of the type (n, n) we call a *square matrix of the degree* n.
- A square matrix $\boldsymbol{A} = \big(a_{ij}\big)$ we call a *diagonal matrix*, if $a_{ij} = 0$ for all pairs of indices $i, j \in \{1, 2, \ldots, n\}$, $i \neq j$, that is, a diagonal matrix has nonzero entries only on the main diagonal.
- A diagonal matrix $\boldsymbol{E} = (e_{ij})$ is called the *identity matrix*, if all its diagonal entries are equal to one, that is, if $e_{ij} = 0,\ i \neq j$, and $e_{ii} = 1$ for all $i, j \in \{1, 2, \ldots, n\}$.

An identity matrix of the degree n we denote by $\boldsymbol{E}$, or $\boldsymbol{E}_n$, that is,

$$\boldsymbol{E}_1 = \begin{pmatrix} 1 \end{pmatrix}, \quad \boldsymbol{E}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{E}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{atd.}$$

Theorem 5. *Let* $\boldsymbol{A}$*,* $\boldsymbol{B}$*,* $\boldsymbol{C}$ *be matrices and let* $r \in \mathbb{R}$*. Then (whenever the following operations are defined):*

1. $r(\boldsymbol{A}\,\boldsymbol{B}) = (r\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}(r\boldsymbol{B})$,
2. $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A}\,\boldsymbol{B} + \boldsymbol{A}\,\boldsymbol{C}$,
3. $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{A}\,\boldsymbol{C} + \boldsymbol{B}\,\boldsymbol{C}$,
4. $(\boldsymbol{A}\,\boldsymbol{B})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{B}\,\boldsymbol{C})$,
5. $0\,\boldsymbol{A} = \boldsymbol{O}$, $\boldsymbol{A}\,0 = \boldsymbol{O}$,
6. $\boldsymbol{E}\,\boldsymbol{A} = \boldsymbol{A}$, $\boldsymbol{A}\,\boldsymbol{E} = \boldsymbol{A}$.

Equations 2 and 3 are called distributive properties, the equation 4 is called an associative property.

2.3 Inverse of a matrix

Definition 6. Let $\boldsymbol{A}$, $\boldsymbol{B}$ be two square matrices of degree n. We say that a matrix $\boldsymbol{B}$ is the *inverse* of a matrix $\boldsymbol{A}$ if $\boldsymbol{A}\,\boldsymbol{B} = \boldsymbol{B}\,\boldsymbol{A} = \boldsymbol{E}$.

Definition 7. A square matrix $\boldsymbol{A}$ of degree n is called a *regular matrix* if its rank $\mathrm{h}(\boldsymbol{A})$ is equal to n. A square matrix which is not regular is called *singular*.

Theorem 8. *Let* $\boldsymbol{A}$ *be a square matrix of degree* n*. Then the following conditions are equivalent.*

1. *A matrix* $\boldsymbol{A}$ *is regular;*

2. *there exists a square matrix* $\boldsymbol{B}$ *of degree* n *such that* $\boldsymbol{A}\,\boldsymbol{B} = \boldsymbol{E}$*;*

3. *there exists a square matrix* $\boldsymbol{C}$ *of degree* n *such that* $\boldsymbol{C}\,\boldsymbol{A} = \boldsymbol{E}$*.*

Clearly, $\boldsymbol{B} = \boldsymbol{C}$ *and it is the inverse of* $\boldsymbol{A}$*.*

The *inverse* of a matrix $\boldsymbol{A}$ we *denote* by $\boldsymbol{A}^{-1}$.

2.4 Finding the inverse (Jordan elimination)

Given regular matrix $\boldsymbol{A}$ of the degree n we associate with the matrix $\big(\boldsymbol{A}|\boldsymbol{E}\big)$ of the type $(n, 2n)$. By suitable elementary row operations we want to obtain on the left (that is, on the place of $\boldsymbol{A}$) the identity matrix of the degree n. On the right (that is, on the place of $\boldsymbol{E}$) we have then the inverse of $\boldsymbol{A}$.

$$\big(\boldsymbol{A}|\boldsymbol{E}\big) \overset{\text{Gaussian elimination on } \boldsymbol{A}}{\sim \cdots \sim} \big(\boldsymbol{B}|\boldsymbol{C}\big) \overset{\text{reverse elimination}}{\sim \cdots \sim} \big(\boldsymbol{E}|\boldsymbol{A}^{-1}\big)$$

At the beginning we obtain from $\big(\boldsymbol{A}|\boldsymbol{E}\big)$ a matrix having on the left a Gaussian matrix $\boldsymbol{B}$.

Reverse elimination: 1st step: Adding suitable multiples of bottom row to the above rows we obtain zeros above the right entry of this row. 2nd step: we repeat the process taking now the second row from the bottom etc. At the end we have a matrix with a diagonal matrix with nonzero diagonal entries on the left (it follows from regularity of $\boldsymbol{A}$). To finish the process we divide each row by the corresponding nonzero diagonal number and we obtain the matrix $\big(\boldsymbol{E}|\boldsymbol{A}^{-1}\big)$.

2.5 Solving matrix equations by the use of the inverse

Assume that $\boldsymbol{B}$ is a regular matrix. For solving matrix equations we use:

$$\boldsymbol{X}\,\boldsymbol{B} = \boldsymbol{A} \qquad \Longrightarrow \qquad \boldsymbol{X} = \boldsymbol{A}\,\boldsymbol{B}^{-1}.$$

$$\boldsymbol{B}\,\boldsymbol{X} = \boldsymbol{A} \qquad \Longrightarrow \qquad \boldsymbol{X} = \boldsymbol{B}^{-1}\,\boldsymbol{A}.$$

Let the matrix $\boldsymbol{B} - \boldsymbol{E}$ be regular. Then

$$\boldsymbol{X}\,\boldsymbol{B} = \boldsymbol{C} + \boldsymbol{X} \quad \Longrightarrow \quad \boldsymbol{X}\,\boldsymbol{B} - \boldsymbol{X}\,\boldsymbol{E} = \boldsymbol{C} \quad \Longrightarrow \quad \boldsymbol{X} = \boldsymbol{C}\,(\boldsymbol{B} - \boldsymbol{E})^{-1}.$$

$$\boldsymbol{B}\,\boldsymbol{X} = \boldsymbol{C} + \boldsymbol{X} \quad \Longrightarrow \quad \boldsymbol{B}\,\boldsymbol{X} - \boldsymbol{E}\,\boldsymbol{X} = \boldsymbol{C} \quad \Longrightarrow \quad \boldsymbol{X} = (\boldsymbol{B} - \boldsymbol{E})^{-1}\,\boldsymbol{C}.$$