Algebra

Andrzej Więckowski

August 24, 2020

Contents

1 Vectors

1.1 Matrix notation

We work in linear space where two types of vectors, the vertical $\vec{v}$ and the horizontal ones $\vec{v}^{\mathsf{T}}$, are defined

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \quad \vec{v}^{\mathsf{T}} = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}. \tag{1}$$

We consider them as matrices: $\vec{v}$ as $n \times 1$, and $\vec{v}^{\mathsf{T}}$ as $1 \times n$ matrix, respectively (note: $a \times b$ corresponds to matrix with a rows and b columns and T matrix transposition). Manipulation on a vector can be achieved by applying some matrix A

$$\boxed{A\vec{v} = \vec{w}.} \tag{2}$$

One can say that matrix A *transforms* vector $\vec{v}$ into vector $\vec{w}$. Matrix A can be used for vector $\vec{v}$ rotation, reflection, scaling etc. In general, matrix A does not have to be square

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nm} \end{pmatrix}. \tag{3}$$

Supposing that vector $\vec{v}$ has n rows, then as the result, vector $\vec{w}$ has m rows. However, for our purposes, we will consider transformations in the same space, i.e. square matrices ($n = m$). Note: matrices can be multiplied if, and only if their "inner" dimensions are the same, i.e. if A is $n \times m$ and B is $m \times k$, then $AB = C$, where C is $n \times k$. In particular, matrix A from Equation (2) is $m \times n$, vector $\vec{v}$ is $n \times 1$, then $\vec{w}$ is $m \times 1$.

1.2 Origin of the matrix

How do we multiply matrices? Using $\textbf{math}.\textbf{mul}(A, B)$, however, when we do not have access to computer power resources, matrices multiplication can be calculated using the rule: *row—times—column*

$$(AB)_{ij} = \sum_k A_{ik} B_{kj} \tag{4}$$

Matrices were introduced for describing the system of equations. Consider the following example

$$\begin{cases} ax_1 + bx_2 = c, & \text{(5a)} \\ dx_1 + ex_2 = f. & \text{(5b)} \end{cases}$$

Equations (5a)–(5b) are equivalent to the following matrix equation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} c \\ f \end{pmatrix}. \tag{6}$$

It is worth to mention that matrix product is *non-Abelian* (non-commutative) in general, i.e.

$$\boxed{AB \neq BA.} \tag{7}$$

As an excercice I would recomend checking that

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{8}$$

and

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{9}$$

to verify that matrix multiplication can indeed be non-commutative. However there exist matrices that satisfy $AB = BA$, but this is *not* a general property of matrix algebra.

1.3 Standard operations

First, let us discuss which mathematical operations are standard defined in the Literature:

- vectors, and matrices, *can* be multiplied by some scalar α:

$$\alpha\vec{v} = \alpha \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \alpha v_1 \\ \vdots \\ \alpha v_n \end{pmatrix}; \; [(\alpha\vec{v})_i = \alpha v_i] \tag{10}$$

$$\alpha A = \alpha \begin{pmatrix} A_{11} & \dots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nm} \end{pmatrix} = \begin{pmatrix} \alpha A_{11} & \dots & \alpha A_{1m} \\ \vdots & \ddots & \vdots \\ \alpha A_{n1} & \dots & \alpha A_{nm} \end{pmatrix}. \; [(\alpha A)_{ij} = \alpha A_{ij}] \tag{11}$$

- vectors and matrices *can* be added together, if they have common dimensions (element wise addition):

$$(\vec{v} + \vec{w})_i = v_i + w_i, \tag{12}$$

$$(A + B)_{ij} = A_{ij} + B_{ij}; \tag{13}$$

- vector $\vec{v}$ and scalar α *cannot* add up (unless we treat them as *Quaternions*, see Sec. 4)

$$\cancel{\vec{v} + \alpha}; \tag{14}$$

- square matrix A and scalar α *can* add up (we treat scalar as identity matrix multiplied by the scalar)

$$A + \alpha \equiv A + \alpha \mathbb{1}. \tag{15}$$

1.4 Vector basis

We defined vector $\vec{v}$ in Eq. (1) as a column

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \tag{16}$$

but what do these values v_i mean? We assume that we work in a vector space with some basis defined. The properties of the basis we describe in the next sections. Each vector, which belongs to the space, can be decompose into basis vectors. Each coefficient v_i corresponds to the basis vector $\hat{e}_i$

$$\vec{v} = \sum_i v_i \hat{e}_i \equiv \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}. \tag{17}$$

The most common basis is the canonical one

$$\hat{e}_i \in \left\{ \hat{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \hat{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \hat{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}. \tag{18}$$

However, one can rotate basis vectors $\hat{e}_i$ to obtain totally different basis than basis presented in Eq. (18). Except the basis, a proper vector space should have defined dot (scalar) product, which is elaborated in Sec. 1.5. Please note that we focus on *useful* spaces, with application to our field of study. In the notes, we will consider only the *Coordinate space* (mostly $\mathbb{R}^3$ or $\mathbb{R}^2$), and we will skip all abstract infinite spaces or spaces defined with functions and integrals as a scalar product.

1.5 Dot product

In our coordinate space we define dot (scalar) product as follows

$$\boxed{\vec{v} \circ \vec{w} = \vec{v}^{\mathsf{T}} \vec{w} = \sum_i v_i w_i.} \tag{19}$$

Note that the result of such operation $\cdot \circ \cdot$ is symmetric, i.e. dot does not depend on the vectors order

$$\vec{v} \circ \vec{w} = \vec{w} \circ \vec{v}, \tag{20}$$

but be aware of matrix notation! Matrices, and vectors, do not commute in general

$$\vec{v}^{\mathsf{T}} \vec{w} \neq \vec{w} \vec{v}^{\mathsf{T}}, \tag{21}$$

since the result of the left hand side is a scalar and the result of the right hand side of the Eq. (21) is a matrix $n \times n$. Of course the we swap vectors if we transpose them

$$\vec{v}^{\mathsf{T}} \vec{w} = (\vec{v}^{\mathsf{T}} \vec{w})^{\mathsf{T}} = \vec{w}^{\mathsf{T}} \vec{v}. \tag{22}$$

1.6 Basis properties

In the all discussion presented here, we will assume that basis of the vector space is *orthonormal* and *complete*.

1.6.1 Orthonormality

The proper vector space basis should be ortonormal, i.e. basis vectors should satisfy the following condition

$$\hat{e}_i \circ \hat{e}_j = \delta_{ij}, \tag{23}$$

where δ_{ij} is a Kronecker symbol, which is defined as

$$\delta_{ij} = \begin{cases} 0, & \text{for} \quad i \neq j, \\ 1, & \qquad\;\; i = j. \end{cases} \tag{24}$$

Example, canonical basis ($\mathbb{R}^2$):
$\hat{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\hat{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\hat{x} \circ \hat{x} = 1$, $\hat{x} \circ y = 0$, $\hat{y} \circ \hat{x} = 0$, $\hat{y} \circ \hat{y} = 1 \rightarrow$ orthonormal basis!

Presented property in Eq. (23) is sufficient to derive dot product from Eq. (19) between two arbitrary vectors

$$\vec{v} \circ \vec{w} = \left(\sum_i v_i \hat{e}_i \right) \circ \left(\sum_j w_j \hat{e}_j \right) = \sum_{ij} v_i w_j \hat{e}_i \circ \hat{e}_j = \sum_{ij} v_i w_j \delta_{ij} = \sum_i v_i w_i, \tag{25}$$

where in the last step we used the famous delta property

$$\sum_i a_i \delta_{ij} = a_j. \tag{26}$$

1.6.2 Completeness

Other very important property of the basis is *completeness*

$$\boxed{\sum_i \hat{e}_i \hat{e}_i^{\mathsf{T}} = \mathbb{1}.} \tag{27}$$

Remember: we treat vectors as matrices! Certainly, the operation $\vec{v}\,\vec{v}$ does not make any sense, because we cannot multiply matrices with dimensions: $n \times 1$, $n \times 1$, respectively, but $\vec{v}\,\vec{v}^{\mathsf{T}}$ is valid, and it produces $n \times n$ matrix. Such operation name is *outer product* and is commonly used e.g. in quantum mechanics, computer graphics, neural networks, and in every field where matrices plays an important role.

Example, canonical basis ($\mathbb{R}^2$):

$\hat{x} = \begin{pmatrix}1\\0\end{pmatrix}, \hat{y} = \begin{pmatrix}0\\1\end{pmatrix},$

$\hat{x}\hat{x}^\mathsf{T} + \hat{y}\hat{y}^\mathsf{T} = \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} + \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} = \mathbb{1}$

1.7 Basis transform

Assume that we have two different orthonormal and complete basis $\{\hat{e}_i\}$ and $\{\widetilde{e}_i\}$. We know vector $\vec{v}$ coefficients v_i in basis $\{\hat{e}_i\}$, but how find to vector coefficient $\widetilde{v}_i$ in the basis $\{\widetilde{e}_i\}$? It is quite simply, one just has to multiply the basis vector by one

$$\hat{e}_i = \mathbb{1}\hat{e}_i = \sum_j \widetilde{e}_i\widetilde{e}_i^\mathsf{T}\ \hat{e}_i \sum_j \widetilde{e}_i\ \widetilde{e}_j \circ \hat{e}_i = \sum_j \widetilde{e}_i \circ \hat{e}_i\ \widetilde{e}_j. \tag{28}$$

Finally, we ended with a nice recipe

$$\boxed{\hat{e}_i = \sum_j (\widetilde{e}_j \circ \hat{e}_i)\ \widetilde{e}_j.} \tag{29}$$

Using Eq. (29), one can find transformation equation between representations $\{\hat{e}_i\}$ and $\{\widetilde{e}_i\}$. First, consider two dimensional case

$$\vec{v} = \begin{pmatrix}v_1\\v_2\end{pmatrix} = v_1\hat{e}_1 + v_2\hat{e}_2, \tag{30}$$

then apply Eq. (29)

$$\vec{v} = v_1\sum_j \widetilde{e}_j \circ \hat{e}_1\widetilde{e}_j + v_2\sum_j \widetilde{e}_j \circ \hat{e}_2\widetilde{e}_j. \tag{31}$$

After a few algebraic steps, one can find the following

$$\vec{v} = (v_1\widetilde{e}_1 \circ \hat{e}_1 + v_2\widetilde{e}_1 \circ \hat{e}_2)\widetilde{e}_1 + (v_1\widetilde{e}_2 \circ \hat{e}_1 + v_2\widetilde{e}_2 \circ \hat{e}_2)\widetilde{e}_2, \tag{32}$$

which is vector $\vec{v}$ representation in basis $\{\widetilde{e}_i\}$. Equation (32) can be written in matrix notation

$$\mathcal{O}\vec{v} = \begin{pmatrix}\widetilde{e}_1 \circ \hat{e}_1 & \widetilde{e}_1 \circ \hat{e}_2\\ \widetilde{e}_2 \circ \hat{e}_1 & \widetilde{e}_2 \circ \hat{e}_2\end{pmatrix}\begin{pmatrix}v_1\\v_2\end{pmatrix} = \begin{pmatrix}\widetilde{v}_1\\\widetilde{v}_2\end{pmatrix}. \tag{33}$$

Matrix $\mathcal{O}$ is the transformation matrix between basis and has the following compact form

$$\boxed{\mathcal{O} = \sum_{ij}\mathcal{O}_{ij}\ \hat{e}_i\hat{e}_j^\mathsf{T} = \sum_{ij}(\widetilde{e}_i \circ \hat{e}_j)\ \hat{e}_i\hat{e}_j^\mathsf{T}} \tag{34}$$

Not only vectors transform when a vector space basis is changed. When we transform the basis, we have to transform matrices basis as well. First, apply ones from left and right-hand side, respectively, of the arbitrary matrix, A and use the Eq. (27)

$$A = \mathbb{1}A\mathbb{1} = \sum_i \hat{e}_i\hat{e}_i^\mathsf{T}A\sum_j \hat{e}_j\hat{e}_j^\mathsf{T} = \sum_{ij}(\hat{e}_i^\mathsf{T}A\hat{e}_j)\ \hat{e}_i\hat{e}_j^\mathsf{T} = \sum_{ij}A_{ij}\hat{e}_i\hat{e}_j^\mathsf{T}, \tag{35}$$

where we formally introduce matrix coefficient

$$\boxed{A_{ij} = \hat{e}_i^\mathsf{T}A\hat{e}_j,} \tag{36}$$

and we explicitly express matrix A in basis $\{\hat{e}_i\}$

$$\boxed{A = \sum_{ij}A_{ij}\ \hat{e}_i\hat{e}_j^\mathsf{T}.} \tag{37}$$

The question is how to find matrix A coefficients $\widetilde{A}_{ij}$ in basis $\{\widetilde{e}_i\}$? At this step, we will show the transformation scheme beginning with the matrix expressed already in $\{\widetilde{e}_i\}$ basis and then show the way back to the original basis. The first step is to again apply trick with ones

$$A = \sum_{ij} (\widetilde{e}_i^\top A \widetilde{e}_j)\, \widetilde{e}_i \widetilde{e}_j^\top = A = \sum_{ij} (\widetilde{e}_i^\top \mathbb{1} A \mathbb{1} \widetilde{e}_j)\, \widetilde{e}_i \widetilde{e}_j^\top = A = \sum_{ijkl} \widetilde{e}_i^\top \hat{e}_k \hat{e}_k^\top A \hat{e}_l \hat{e}_l^\top \widetilde{e}_j\, \widetilde{e}_i \widetilde{e}_j^\top. \tag{38}$$

Next, recall the transformation matrix elements $\mathscr{O}_{ij} = \widetilde{e}_i \circ \hat{e}_j$ from Eq. (34)

$$A = \sum_{ijkl} \underbrace{\widetilde{e}_i^\top \hat{e}_k}_{\mathscr{O}_{ik}} \underbrace{\hat{e}_k^\top A \hat{e}_l}_{A_{kl}} \underbrace{\hat{e}_l^\top \widetilde{e}_j}_{\mathscr{O}_{lj}^\top}\, \widetilde{e}_i \widetilde{e}_j^\top = \sum_{ijkl} \mathscr{O}_{ik} A_{kl} \mathscr{O}_{lj}^\top\, \widetilde{e}_i \widetilde{e}_j^\top \tag{39}$$

The last step is to use the three-matrix multiplication formula [see Eq. (4)]

$$(ABC)_{ij} = \sum_{kl} A_{ik} B_{kl} C_{lj}, \tag{40}$$

then Eq. (39) takes form

$$A = \sum_{ij} (\mathscr{O} A\, \mathscr{O}^\top)_{ij} \widetilde{e}_i \widetilde{e}_j^\top = \sum_{ij} \widetilde{A}_{ij} \widetilde{e}_i \widetilde{e}_j^\top \tag{41}$$

Finally to obtain matrix element in new basis $\{\widetilde{e}_i\}$ one has to follow the transformation given by

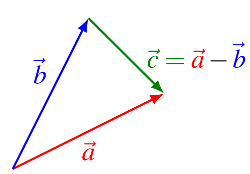
$$\boxed{\widetilde{A}_{ij} = (\mathscr{O} A \mathscr{O}^\top)_{ij}.} \tag{42}$$

1.8 Famous equation

During physics and mathematics classes in high-school, you have to hear about the famous cosine formula

$$\boxed{\vec{a} \circ \vec{b} = \|\vec{a}\| \|\vec{b}\| \cos\theta.} \tag{43}$$

The origin of the Eq. (43) lies in the basic geometry theorem. Let's take two arbitrary vectors $\vec{a}$ and $\vec{b}$, and construct with them the third vector $\vec{c} = \vec{a} - \vec{b}$. The sketch of the described situation can be found in the figure attached below:



All three vectors form a triangle with

$$\|\vec{c}\|^2 = \vec{c} \circ \vec{c} = (\vec{a} - \vec{b}) \circ (\vec{a} - \vec{b}) = \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\vec{a} \circ \vec{b}. \tag{44}$$

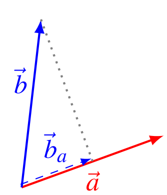
Using cosine theorem

$$c^2 = a^2 + b^2 - 2ab\cos\theta, \tag{45}$$

one can prove the famous Equation (43).

1.9 Projection

The dot product is usually used for calculating vector projection. Consider the following situation



In order to calculate vector $\vec{b}_a$ length one can use dot product. Using cosine definition one gets

$$\|\vec{b}_a\| = \|\vec{b}\| \cos\theta. \tag{46}$$

Using Eq. (43)

$$\boxed{\|\vec{b}_a\| = \vec{b} \circ \hat{a},} \tag{47}$$

where $\hat{a} = \vec{a}/\|\vec{a}\|$ is a unit vector in $\vec{a}$ direction. Finally, vector $\vec{b}_a$ reads

$$\boxed{\vec{b}_a = (\vec{b} \circ \hat{a})\, \hat{a}.} \tag{48}$$

2 Cross product

Future tech talks

3 Rotation

Future tech talks

4 Quaternions

Future tech talks