A Quick Glimpse on Vector and Matrix for Octave Starters

Wai-Yiu Keung

Division of Engineering & Technology, HKUSPACE Community College.

September 7, 2022

This document serves as an additional material for CCIT4076: Engineering and Information Sciences as offered in Fall 2022. We aim at providing necessary insights for course participants to start their journey in using Octave as a computational tool for engineering purposes. For a proper understanding in matrix algebra readers are referred to their course instructors in CCMA4002: Linear Algebra.

1 Notations and Definitions

A **vector**: $\mathbf{x} \in \mathbb{R}^n$ means $\mathbf{x}$ is a length n vector:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = (x_1, x_2, \dots, x_n)$$

with all entries being real-valued. A vector is a column vector unless otherwise specified. We may use the round bracket to denote a column vector to save space. It is analogous to a 1-dimensional array where the i-th cell is storing a real number. A **matrix** $\mathbf{A} \in \mathbb{R}^{m \times n}$ means $\mathbf{A}$ is a m-by-n matrix, i.e.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix}$$

where $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ are n real-valued vectors in m-dimensional space. $\mathbf{A}$ is said to be square if $m = n$. It is analogous to a 2-dimensional array where the m, n-th element is a real number.

Let us post the following examples to demonstrate what is a matrix and a vector. For instance, we have

$$\begin{bmatrix} 1 & 3 & 5 \\ 7 & 9 & 11 \\ 13 & 15 & 17 \\ 19 & 21 & 23 \end{bmatrix} \in \mathbb{R}^{4 \times 3}$$

being a matrix; and consequently,

$$\begin{bmatrix} 1 \\ 7 \\ 13 \\ 19 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ 9 \\ 15 \\ 21 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 5 \\ 11 \\ 17 \\ 23 \end{bmatrix}$$

are vectors of $\mathbb{R}^4$. We stress the fact that *data* are very often stored as these format. We are hence interested to look into this topic.

Some special types of matrices are introduced as follows. We refer $\mathbf{0}$ as a **zero matrix/vector** with all entries $= 0$; similarly, $\mathbf{1}$ is **all-one matrix/vector** if all its entries $= 1$, for instance,

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

are both denoted by $\mathbf{0}$ and

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

are both meant by $\mathbf{1}$.

Convention: We use $\mathbf{x} \geq \mathbf{0}$ to indicate that all the components of $\mathbf{x}$ are non–negative, and $\mathbf{x} \geq \mathbf{y}$ to mean that $\mathbf{x} - \mathbf{y} \geq \mathbf{0}$. The notations $\mathbf{x} > \mathbf{0}$, $\mathbf{x} \leq \mathbf{0}$, $\mathbf{x} < \mathbf{0}$, $\mathbf{x} > \mathbf{y}$, $\mathbf{x} \leq \mathbf{y}$, and $\mathbf{x} < \mathbf{y}$ are to be interpreted accordingly. For instance, given vector $\mathbf{a}$ as

$$\mathbf{a} = \begin{bmatrix} 1 \\ 7 \\ 13 \\ 19 \end{bmatrix}$$

and the condition $\mathbf{a} \leq \mathbf{b}$. It is known that the vector $\mathbf{b}$ must have entries satisfying

$$\begin{aligned} b_1 - 1 &\geq 0 \\ b_2 - 7 &\geq 0 \\ b_3 - 13 &\geq 0 \\ b_4 - 19 &\geq 0 \end{aligned}$$

and imaginably, there are a lot of solutions. Some of those vectors satisfying the constraints are:

$$\begin{bmatrix} 1 \\ 7 \\ 13 \\ 19 \end{bmatrix} \quad \begin{bmatrix} 4 \\ 8 \\ 14 \\ 20 \end{bmatrix} \quad \begin{bmatrix} 1000 \\ 2000 \\ 3000 \\ 4000 \end{bmatrix}$$

Diagonal operator: We use the operator $\mathrm{diag}(\mathbf{x})$ to denote a square matrix with all zero entries except the diagonal elements are those of $\mathbf{x}$, and we can use it to extract a vector containing all diagonal elements from a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$:

$$\operatorname{diag}(\mathbf{x}) = \begin{bmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x_n \end{bmatrix} \quad \text{and} \quad \operatorname{diag}(\mathbf{A}) = \begin{bmatrix} a_{11} \\ a_{22} \\ \vdots \\ a_{nn} \end{bmatrix}$$

Identity matrix: Is an diagonal matrix with all it's diagonal element fixed to 1

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \operatorname{diag}\left(\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right) = \operatorname{diag}(\mathbf{1})$$

2 Basic Operators

Addition/subtraction: Given $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{m \times n}$,

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix} \end{aligned}$$

whereas $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$ is defined in a similar manner. **Scalar multiplication**: Given a scalar term $c \in \mathbb{R}$,

$$c\mathbf{A} = c \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1n} \\ ca_{21} & ca_{22} & \dots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \dots & ca_{mn} \end{bmatrix}$$

Matrix multiplication: The product between 2 matrices is defined more strictly —it may not always exists. Let $\mathbf{C} \in \mathbb{R}^{n \times m}$, then the product

$$\mathbf{AC} = m \text{ rows} \left\{ \underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}}_{n \text{ columns}} \underbrace{\begin{bmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nm} \end{bmatrix}}_{m \text{ columns}} \right\} n \text{ rows}$$

and the (i, j)th element of $\mathbf{AC}$ is defined as

$$[\mathbf{AC}]_{ij} = \sum_{k=1}^{n} a_{ik} c_{kj}$$

Note that even if $\mathbf{A}, \mathbf{C} \in \mathbb{R}^{n\times n}$, $\mathbf{AC} \neq \mathbf{CA}$ in general.

Let us visit a numerical example. Suppose there are three matrices

$$\mathbf{A} = \begin{bmatrix} -8 & -4 & -1 \\ 2 & -6 & 6 \\ -3 & 5 & -1 \\ 0 & 4 & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 7 & -2 \\ 1 & -5 \\ 7 & 2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} -2 & -2 & 2 \\ -1 & -1 & -2 \\ 1 & 1 & -1 \\ 1 & 0 & 0 \end{bmatrix}$$

It is straight forward to see that

$$\mathbf{A} - \mathbf{C} = \begin{bmatrix} -6 & -2 & -3 \\ 3 & -5 & 8 \\ -4 & 4 & 0 \\ -1 & 4 & 3 \end{bmatrix}, \quad \mathbf{A} + \mathbf{C} = \begin{bmatrix} -10 & -6 & 1 \\ 1 & -7 & 4 \\ -2 & 6 & -2 \\ 1 & 4 & 3 \end{bmatrix} \quad 0.3\mathbf{A} + 0.6\mathbf{C} = \begin{bmatrix} -3.6 & -2.4 & 0.9 \\ 0 & -2.4 & 0.6 \\ -0.3 & 2.1 & -0.9 \\ 0.6 & 1.2 & 0.9 \end{bmatrix}$$

Also, it is clear that the sum $\mathbf{A} + \mathbf{B}$ or $\mathbf{C} + \mathbf{B}$ do not exists due to dimensional mismatch. The same goes for all the difference. The matrix products

$$\mathbf{AB} = \begin{bmatrix} -67 & 34 \\ 50 & 38 \\ -23 & -21 \\ 25 & -14 \end{bmatrix}, \quad \mathbf{CB} = \begin{bmatrix} -2 & 18 \\ -22 & 3 \\ 1 & -9 \\ 7 & -2 \end{bmatrix}$$

follows from the definition of matrix multiplication. Other products $\mathbf{AC}$, $\mathbf{CA}$, $\mathbf{BC}$ and $\mathbf{BA}$ do not exists due to, again, mismatch in dimensionality.

Inverse matrix: For $\mathbf{A} \in \mathbb{R}^{n\times n}$, if there exists $\mathbf{B} \in \mathbb{R}^{n\times n}$ such that

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}$$

then $\mathbf{B} = \mathbf{A}^{-1}$ is said to be the inverse matrix of $\mathbf{A}$.

Matrix transposition: We define the transpose operator $(.)^T$ as follow:

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

to put it in words, we treat the ith column of $\mathbf{A}$ as the ith row of $\mathbf{A}^T$. Note some key facts:

1. $(\mathbf{A}^T)^T = \mathbf{A}$
2. $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
3. $(c\mathbf{A})^T = c\mathbf{A}^T$
4. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$

Also, symmetric matrices satisfy $\mathbf{A} = \mathbf{A}^T$; skew-symmetric matrices satisfies $\mathbf{A} = -\mathbf{A}^T$.

Symmetry: We say a matrix $\mathbf{A}$ is symmetric if its elements satisfy $a_{ij} = a_{ji}$. Similarly, $\mathbf{A}$ is said to be skew-symmetric if its elements satisfy $a_{ij} = -a_{ji}$. As an example:

$$\mathbf{S} = \begin{bmatrix} 4 & 7 \\ 7 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 7 \\ 7 & 4 \end{bmatrix}^T = \mathbf{S}^T$$

is symmetric, and

$$\mathbf{H} = \begin{bmatrix} 0 & -5 \\ 5 & 0 \end{bmatrix} = -\begin{bmatrix} 0 & 5 \\ -5 & 0 \end{bmatrix}^T = -\mathbf{H}^T$$

is skew-symmetric. Notably, if a matrix $\mathbf{H}$ is skew-symmetric, it must satisfied that

$$h_{1,1} = h_{2,2} = \cdots = h_{n,n} = 0.$$

Vector inner products: For $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$, their inner product (also referred as the dot product in vector calculus classes) is defined as:

$$\mathbf{x} \bullet \mathbf{y} = \mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_i y_i.$$

– Note that when $\mathbf{y} = \mathbf{x}$, the inner product gives the sum-of-squares of $\mathbf{x}$, which is the squared 2-norm[1]:

$$\|\mathbf{x}\|_2^2 = \sum_{i=1}^{n} x_i^2 = \mathbf{x}^T\mathbf{x}.$$

– Also, the quantity $\sqrt{\mathbf{x}^T\mathbf{y}}$ is also referred as the Euclidean distance between $\mathbf{x}$ and $\mathbf{y}$. It is a real-valued scalar that gives a measurement on the distance between the two vectors.

Let us go through another numerical example. Suppose

$$\mathbf{x} = \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} -2 \\ 4 \\ -8 \end{bmatrix}$$

their vector inner product can be computed as

$$\mathbf{x}^T\mathbf{y} = \begin{bmatrix} 5 & 7 & 9 \end{bmatrix} \begin{bmatrix} -2 \\ 4 \\ -8 \end{bmatrix} = (5)(-2) + (7)(4) + (9)(-8) = -54.$$

Similarly, the 2-norm of theses two vectors are

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T\mathbf{x}} = 12.45; \qquad \|\mathbf{y}\|_2 = \sqrt{\mathbf{y}^T\mathbf{y}} = 9.1652$$

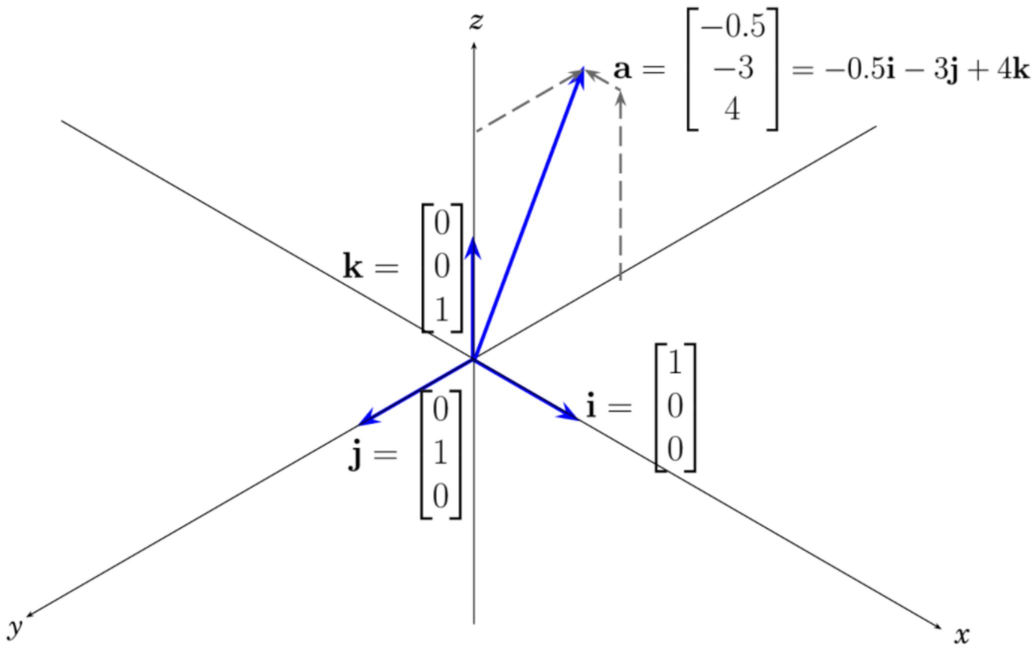
respectively.

[1] Here it means the squared Euclidean norm, or simply, the squared norm of the vector.

Traditionally, vector has a more profound meaning in the physical world. Since everything we touch in the reality is a three dimensional object, we can actually declare the whole reality as a *linear combination* of the unit vectors $\mathbf{i} = [1, 0, 0]^T$, $\mathbf{j} = [0, 1, 0]^T$ and $\mathbf{k} = [0, 0, 1]^T$, i.e. a vector

$$\mathbf{a} = \alpha_1 \cdot \mathbf{i} + \alpha_2 \cdot \mathbf{j} + \alpha_3 \cdot \mathbf{k}$$

where $\alpha_i \in \mathbb{R}$ for $i = 1, 2, 3$ can be used to characterise a certain *position vector*. Below we visualise $\mathbf{a} = [-1/2, -3, 4]^T$ as mapped to a xyz-coordinate space.



The vector 2-norm $\|\mathbf{a}\|_2 = 5.0249$ here is a scalar measuring *how long* is the vector $\mathbf{a}$ in the 3-D space.

Vector outer products: For $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$, their outer product is defined as:

$$\mathbf{x} \otimes \mathbf{y} = \mathbf{x}\mathbf{y}^T = \begin{bmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_m \\ x_2y_1 & x_2y_2 & \dots & x_2y_m \\ \vdots & \vdots & \ddots & \vdots \\ x_ny_1 & x_ny_2 & \dots & x_ny_m \end{bmatrix}.$$

– $\mathbf{x} \otimes \mathbf{y} \neq \mathbf{y} \otimes \mathbf{x}$ in general. But $\mathbf{x} \otimes \mathbf{y} = (\mathbf{y} \otimes \mathbf{x})^T$ for sure.

– When $\mathbf{x}$ and $\mathbf{y}$ are of the same length, the sum of the diagonal entries of $\mathbf{x} \otimes \mathbf{y}$ or $\mathbf{y} \otimes \mathbf{x}$ will give us the vector inner product, i.e.:

$$\sum_{i=1}^{n} \text{diag}(\mathbf{x} \otimes \mathbf{y}) = \sum_{i=1}^{n} \text{diag}(\mathbf{y} \otimes \mathbf{x}) = \mathbf{x} \bullet \mathbf{y}$$

3 Plotting Functions on Octave

Roughly speaking, a function is simply a *map* between the input variable and the output variable. For the simplest example, consider a linear function

$$y(x) = a \cdot x + b \qquad \text{(Linear Function)}$$

where a, b are fixed coefficients that dictates the *shape* of the graph of x-versus-y. Let me name a few more examples of functions as follows:

$$y(x) = ax^2 + bx + c \qquad \text{(Quadratic function)}$$
$$y(x) = \sin(\pi x) \qquad \text{(Sinusoidal function)}$$
$$y(x) = A \cdot e^{-\ell x} \qquad \text{(Exponential decay)}$$

As a hands-on technical document, we do not delve into details of these functions but instead we describe how do engineers compute the numerical values and have the function visualised via computational software such as Octave.

Consider the domain of interest lies in $x \in [0, 5]$. What engineers do is to create a *closely sampled vector* between the range $[0, 5]$ as a vector, for instance:

$$\mathbf{x} = \begin{bmatrix} 0 \\ 0.0001 \\ 0.0002 \\ 0.0003 \\ \vdots \\ 4.9998 \\ 4.9999 \\ 5.0000 \end{bmatrix}$$

and incur the respective coefficients to them in obtaining another vector

$$\mathbf{y} = a \cdot \begin{bmatrix} 0 \\ 0.0001 \\ 0.0002 \\ 0.0003 \\ \vdots \\ 4.9998 \\ 4.9999 \\ 5.0000 \end{bmatrix} + b = \begin{bmatrix} a(0) + b \\ a(0.0001) + b \\ a(0.0002) + b \\ a(0.0003) + b \\ \vdots \\ a(4.9998) + b \\ a(4.9999) + b \\ a(5.0000) + b \end{bmatrix}$$

and eventually, we can use `plot(x, y)` to visualise the graph of the equation $y = ax + b$. Such technique allows us to plot whatever function of x in a graph.

Exercise. Try it yourself. Suppose we want to plot the function

$$f(x) = \frac{1}{x^2} \cdot \cos(200\pi x)$$

for $x \in [0, 1]$ on Octave. Assume the step size is 0.0001. Limit the y-axis to the range $[-40, 40]$ and the x-axis to the same range of x.

