18.085/18.0851 Computational Science and Engineering I **Summer 2020**

Week 1 (June 8-June 12)

Lecturer: Richard Zhang *Scribes: Richard Zhang*

Welcome to 18.085/18.0851. This series of lectures will be a rapid review of linear algebra.

1.1 Matrices

Matrices are the fundamental building blocks of linear algebra. Briefly put, they are a table of numbers, such as

$$K = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \tag{1.1}$$

This matrix, K, is a 3×3 matrix. The first 3 indicates the number of rows, while the second 3 indicates the number of columns. You will often see the notation $K \in \mathbb{R}^{3\times 3}$, where $\mathbb{R}$ indicates the real numbers.

1.1.1 Basic Operations of Matrices

One can add, subtract, and multiply matrices. Division is a bit complex so we table the discussion for now.

1.1.1.1 Addition/Subtraction

Matrices of the same dimension can be added/subtracted. To add/subtract two matrices, simply add/subtract the numbers in the corresponding entries. For example, suppose

$$K = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, L = \begin{bmatrix} 0.5 & 1 & 2 \\ -0.5 & 3 & 5 \\ 0 & 4 & 8 \end{bmatrix} \tag{1.2}$$

Then

$$K + L = \begin{bmatrix} 2+0.5 & -1+1 & 0+2 \\ -1-0.5 & 2+3 & -1+5 \\ 0+0 & -1+4 & 2+8 \end{bmatrix} = \begin{bmatrix} 2.5 & 0 & 2 \\ -1.5 & 5 & 4 \\ 0 & 3 & 10 \end{bmatrix} \tag{1.3}$$

1.1.1.2 Multiplication

In order to multiply two matrices, the number of columns of the first matrix has to equal the number of rows of the second matrix (otherwise multiplication simply does not make sense). The entry (i, j) of the resulting matrix of the multiplication is equal to the sum of the products between row i and column j. Let's multiply K and L

$$K * L = \begin{bmatrix} 1.5 & -1 & -1 \\ 1.5 & 1 & 0 \\ 0.5 & 5 & 11 \end{bmatrix} \tag{1.4}$$

1.1.1.3 Inverse

To define the inverse of a matrix, we first define a special matrix called the identity matrix. For a given number n, the identity matrix is defined

Definition 1.1 *For a given natural number n, an identity matrix, I_n, is an $n - by - n$ matrix such that all entries are 0 except for those on the diagonal, which equal 1*

For example, a 3-by-3 identity matrix is

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.5}$$

Then for any n-by-n matrix, A, B is the inverse of A if and only if $A * B = I_n$. We will discuss later on how to find the inverse of a matrix.

Note that not every matrix has an inverse. Matrices that have inverses are called invertible matrices.

1.1.1.4 Transpose

The transpose of a matrix is equal to the same matrix with columns and row switched. For a given matrix A, we let A^T be its transpose. As an example, L^T is equal to

$$L^T = \begin{bmatrix} 1.5 & 1 & .5 \\ 0.5 & 1 & 5 \\ 1 & 9 & 8 \end{bmatrix} \tag{1.6}$$

1.1.1.5 Gaussian elimination

Please watch the video by Khan Academy (https://www.youtube.com/watch?v=woqq3Sls1d8) on how to perform the Gaussian elimination on a matrix.

1.1.2 Vectors

Vectors are matrices such that one of the dimensions is 1. A vector that is $n - by - 1$ is called a row vector, while a vector that is $1 - by - n$ is called a column vector.
Another view of vectors is that they are geometric objects in $\mathbb{R}^n$. Both views are equivalent, but have separate implications in numerical linear algebras

1.1.2.1 Dot Product

Let's suppose we have $\vec{u} = (u_1, u_2, ..., u_n)^T$ and $\vec{v} = (v_1, v_2, .., v_n)^T$. Then we define the dot product of $\vec{u}$ and $\vec{v}$ as

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + ... + u_n v_n \tag{1.7}$$

Alternatively, we can also show that

$$\vec{v} \cdot \vec{u} = |\vec{v}||\vec{u}| \cos(\theta) \tag{1.8}$$

where $|v|$ indicates the magnitude of $\vec{v} = (v_1^2 + ... + v_n^2)^{1/2}$ and θ indicates the angle in between $\vec{u}$ and $\vec{v}$. The dot product tells us a few things

- If the dot product is zero, then we know that $\cos\theta = 0$, which means that $\theta = \frac{\pi}{2}$. The two vectors therefore are perpendicular to each other. In mathematics, we also call it "orthogonal"
- If the dot product is $|u||v|$, then we know that $\cos\theta = 1$, which means that $\theta = 0$. The two vectors therefore are parallel to each other

1.1.2.2 L^p Distance/Norm of Two Vectors

It is rather trivial to compare the difference between two numbers. On the other hand, comparing two vectors does not seem as straightforward. In particular, we define the L^p distance between two vectors, $\vec{v} = (v_1, .., v_n)$ and $\vec{u} = (u_1, ..., u_n)$ as

$$||v - u||_{L_p} = ((v_1 - u_1)^p + (v_2 - u_2)^p + ... + (v_n - u_n)^p)^{1/p} \tag{1.9}$$

Note that for $p = 2$, this is simply the Euclidean/Pythagorean distance between two vectors. We also define the L_∞ norm/distance as

$$||v - u||_{L_\infty} = \max\{|v_1 - u_1|, |v_2 - u_2|, ..., |v_n - u_n|\} \tag{1.10}$$

Note that there is a technical difference between norm and distance, but do not worry about it for this course.

1.2 Properties of Square Matrices

Here is a list of properties for a given matrix $A \in \mathbb{R}^{n\times n}$.

1.2.1 Invertibility

Definition 1.2 *A is invertible if there exists a matrix B such that $AB = I_n$, where I_n is the n-by-n identity matrix.*

We often denote B as A^{-1}.

There are several ways of finding the inverse of a matrix. This video (https://www.youtube.com/watch?v=HwRRdG$_E$4Y o) describes one general way of finding inverses. For the purpose of this course, you only need to know the inverse of a 2-by-2 matrix.

We let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{1.11}$$

Then

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{1.12}$$

As you can see, if $ad - bc = 0$, A^{-1} stops making sense. We call $ad - bc$ the determinant of A. This will be explained further in the section later. For now, let's introduce a few more properties of A Once $A \in \mathbb{R}^{n\times n}$ is invertible, we can then solve for the the system of linear equations. For $b \in \mathbb{R}^n$, we can solve for $x \in \mathbb{R}^n$

$$Ax = b \tag{1.13}$$

The way we solve it is via the inverse

$$A^{-1}Ax = A^{-1}b \tag{1.14}$$

$$x = A^{-1}b \tag{1.15}$$

Gaussian elimination can also be used to solve linear equations. The motivation for the linear equation is system of linear equations, for example:

$$3x_1 - 2x_2 + 0.5x_3 = 1 \tag{1.16}$$

$$-2x_1 + 1.5x_2 - 3x_3 = -2 \tag{1.17}$$

$$0.1x_1 - 4x_2 + 7x_3 = 3.1 \tag{1.18}$$

This can be recast as $Ax = b$, where

$$A = \begin{bmatrix} 3 & -2 & 0.5 \\ -2 & 1.5 & -3 \\ 0.1 & -4 & 7 \end{bmatrix} \tag{1.19}$$

$$b = \begin{bmatrix} 1 \\ -2 \\ 3.1 \end{bmatrix} \tag{1.20}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{1.21}$$

To solve this equation on MATBLAB, we simply do $A\backslash b$ to get x

1.2.2 Positive Definite

Definition 1.3 *$A \in \mathbb{R}^{n\times n}$ is positive definite if for all $x \in \mathbb{R}^n$, the quadratic form is positive, ie. $x^TAx > 0$*

The relevance of this property will be revealed later.

1.2.3 Symmetric

Definition 1.4 *$A \in \mathbb{R}^{n\times n}$ is symmetric if $A^T = A$.*

1.2.4 Diagonal

Definition 1.5 *$A \in \mathbb{R}^{n\times n}$ is diagonal if the only nonzero entries are along the diagonal of the matrix*

1.2.5 Upper/lower triangular

Definition 1.6 *$A \in \mathbb{R}^{n\times n}$ is lower/upper triangular if all entries above/below the diagonal are zero*

1.2.6 Orthogonal and Orthonormal

Definition 1.7 $x, y \in \mathbb{R}^n$ *are orthonormal vectors if* $x \cdot y = 0$ *and* $||x|| = ||y|| = 1$

Definition 1.8 $A \in \mathbb{R}^{n \times n}$ *is an orthogonal matrix if all columns of* A *are orthonormal to each other*

It's beyond me why A is not called an orthonormal matrix but in any case.

1.3 Spaces of/Formed by Vectors

Vector space is an important concept in linear algebra and mathematics in general. First we introduce a few concepts.

- Linear combination: an expression to construct a new vector using a combination of existing vectors and scalars. For example, given vectors v_1 and v_2 and scalars c_1 and c_2, the new vector $u = c_1v_1 + c_2v_2$ would be called a linear combination of v_1 and v_2
- Span: the span of a set of vectors refers to the collection of all linear combinations of the vectors. If we have vectors $S = \{v_1, ..., v_n\}$, the span of S is often denoted as $Span\{S\} = Span\{v_1, ..., v_n\}$
- Linear (in)dependence: a set of vectors, $\{v_1, ..., v_n\}$, is said to be linearly dependent if there exists scalars $a_1, ..., a_n$, where not all the coefficients are zero, such that $a_1v_1 + ... + a_nv_n = 0$

Now we move on to the important concept of vector space

1.3.1 Vector Space

A vector space V refers to a collection of vectors that observe the following rules:

- The zero vector is in V
- If v_1 and v_2 are in V, then any linear combinations of v_1 and v_2 are also in V

Examples of vector spaces are

- $\mathbb{R}^n$
- $Span\{v_1, v_2, v_3\}$, where $v_1 = (1, 1, 1), v_2 = (1, 2, -1), v_3 = (3, 6, -3)$
- $Span\{v_1...v_n\}$, for any collection of vectors $v_1...v_n$
- The zero vector, ie. $v = (0, ...0)$

1.3.1.1 Basis Vectors

You may notice that in the second bullet above, it is not necessary to have v_3 and v_2 at the same time when constructing the vector space. This is because v_2 and v_3 are linearly dependent. Hence, $V = Span\{v_1, v_2, v_3\} = Span\{v_1, v_2\} = Span\{v_1, v_3\}$. The minimal set of linearly independent vectors to span the vector space is called the basis vectors. In this example, v_1, v_2 or v_1, v_3 are both valid basis vectors of V. Furthermore, the number of basis vectors for a given vector space V is called the dimension of V, often denoted as $dim(V)$

1.3.2 Null Space and Column Space

In this section, we shall discuss the two most important vector spaces in linear algebra, null space and column space. For a given matrix $A \in \mathbb{R}^{m \times n}$

Definition 1.9 *The null space of A, denoted as $nul(A)$, is defined as $\{x \in \mathbb{R}^n | Ax = 0\}$.*

If we think of A as a mapping between $\mathbb{R}^m$ to $\mathbb{R}^n$, then the the null space of A refers to the elements in $\mathbb{R}^m$ that get "killed" (aka. reduced to the zero vector) when being mapped to $\mathbb{R}^n$.

Definition 1.10 *The column space of A, denoted as $Col(A)$, is defined as the span of all column vectors of A*

Finding the basis vectors of the column involves reducing the column vectors of the matrix until you have only linearly independent ones. The technique of doing so can be found in this video (https://www.youtube.com/watch?v=avJDljr s), starting at time-stamp 2:11.

It is more important, however, that you familiarize yourself with the method of finding the null space of a matrix. The Khan Academy has a nice explanation in this video (https://www.khanacademy.org/math/linear-algebra/vectors-and-spaces/null-column-space/v/null-space-2-calculating-the-null-space-of-a-matrix)

1.3.3 Rank-Nullity Theorem

One of the most important theorem in linear algebra is stated as follows

Theorem 1.11 *For a given matrix $A \in \mathbb{R}^{m \times n}$, $n = dim(Col(A)) + dim(Nul(A))$. $dim(Col(A))$ is also called the rank of A*

As a result of the rank-nullity theorem, we have some important facts

Theorem 1.12 *For a square matrix $A \in \mathbb{R}^{n \times n}$, the following statements are equivalent*

- *A is invertible*
- *$Rank(A) = n$*
- *All columns of A are linearly independent*
- *$dim(Nul(A)) = 0$*
- *The only member of the null space of A is the zero vectors. This is also called the trivial null space*

1.4 Determinants and Eigenvalues

In this section we introduce the two most important quantities in linear algebra

1.4.1 Determinant

The determinant of a square matrix is a certain quantity that encodes the linear transformation of a matrix. You have already seen the definition of determinat for a $2-by-2$ matrix. Please familiarize yourself with the $n-by-n$ case through this video (https://www.youtube.com/watch?v=HwRRdG$_E$4Yo) by Khan Academy.

Here is a useful add-on to Theorem (1.12):

Theorem 1.13 *$A \in \mathbb{R}^{n\times n}$ is invertible if and only if $det(A) \neq 0$.*

1.4.2 Eigenvalues and Eigenvectors

This is one of the most important concepts in our study. Let $A \in \mathbb{R}^{n\times n}$. Then if a scalar λ and vector v satisfy

$$Av = \lambda v \tag{1.22}$$

Then λ is called an eigenvalue of A and v is called the eigenvector of A.

We first demonstrate how to find eigenvalues. Given the formula above, we write

$$Av - \lambda v = 0 \tag{1.23}$$
$$(A - \lambda I_n)v = 0 \tag{1.24}$$

where I_n is the $n - by - n$ identity matrix. Therefore, v is in the null space of $A - \lambda I_n$. Notice that by Theorem (1.12), either $v \equiv 0$ is the only solution, in which case $A - \lambda I_n$ is invertible, or there exists nontrivial v that satisfies the relation above. By Theorem (1.13), it follows that

$$det\,(A - \lambda I_n) = 0 \tag{1.25}$$

Therefore, to find λ, we just have to compute the determinant and solve the algebraic equation.

Once we obtain λ, we plug it in (1.24) and solve for the null space of $A - \lambda I_n$, which was described in the above section.

As an example, let us compute the eigenvalues and eigenvectors of A, where

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \tag{1.26}$$

First we write the determinant equation

$$det(A - \lambda I_2) = 0 \tag{1.27}$$
$$(2-\lambda)(2-\lambda) - (-1)(-1) = 0 \tag{1.28}$$
$$\lambda^2 - 4\lambda + 3 = 0 \tag{1.29}$$

Hence we obtain that $\lambda_1 = 1$ and $\lambda_2 = 3$. To find the eigenvector, we first compute $A - I\lambda$

$$A - \lambda_1 I_2 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{1.30}$$

$$A - \lambda_2 I_2 = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \tag{1.31}$$

From which we can compute the null spaces as

$$Nul(A - \lambda_1 I_2) = Span\left\{\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}\right\} \tag{1.32}$$

$$Nul(A - \lambda_2 I_2) = Span\left\{\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}\right\} \tag{1.33}$$

It is generally a good habit to normalize the basis vectors (hence the factor of $\sqrt{2}$).

Here is another useful add-on to Theorem (1.12)

Theorem 1.14 *$A \in \mathbb{R}^{n \times n}$ is invertible if and only if all eigenvalues are nonzero.*