MATA22 (Linear Algebra) Course Review

Author: Andrew Hong

April 11, 2020

Contents

1 Introduction

In general, most of the material covered in this section should be material covered in an equivalent (or greater) course to MHF4U/MCV4U (Advanced functions and introduction to calculus and vectors). We will begin with material that should be familiar, but if it is not, take the time to review it before continuing.

2 Vectors, lines and planes

What are vectors? Vectors are geometric objects that have a magnitude and a direction. There are a lot of algebraic operations that can be executed within this field, ranging from vector addition, scalar multiplication, etc.

In terms of notation, we denote vectors as such: $\vec{v} = [v_1, v_2, ..., v_n]$ where v_i are its components and noting that n represents the dimension of the vector. We denote the dimension that a vector is in with $\mathbb{R}^n$ or $\mathbb{C}^n$, depending on whether or not its components are made up of real or complex values.

2.1 Operations with vectors

You are expected to be able to execute basic vector algebra; in other words, please review your notes if you do not remember how to add/subtract vectors or scale them.

2.2 Magnitude/Norm

The magnitude of the vector geometrically is its length and is also a scalar. Recall: A scalar is a quantity that is described by only a magnitude and lacks a direction.

Let $\vec{a} \in \mathbb{R}^n$. In other words, let $\vec{a}$ be a vector of dimension n.
Then, the magnitude of a vector is denoted as:
$$||\vec{a}|| = \sqrt{a_1^2 + a_2^2 + ... + a_n^2}$$

If you do not recognize this, consider the magnitude of a vector in $\mathbb{R}^2$:
Let $\vec{b} \in \mathbb{R}^2$.
$$||\vec{b}|| = \sqrt{b_1^2 + b_2^2} \text{ if } \vec{b} = [b_1, b_2]$$

2.3 Dot product

The dot product is an operation that takes input $\mathbb{R}^n$ and produces a result in $\mathbb{R}$. Geometrically, the dot product represents the product of the magnitudes of some two vectors and the cosine of the angle between them. More formally, it can be seen as:

Let $\vec{a}$, $\vec{b} \in \mathbb{R}^n$. In other words, let $\vec{a}$ and $\vec{b}$ be vectors of the same dimension.
Then, the dot product between vectors **a** and **b** is as follows:
$$\vec{a} \cdot \vec{b} = [a_1, a_2, ..., a_n] \cdot [b_1, b_2, ..., b_n] = a_1 * b_1 + a_2 * b_2 + ... + a_n * b_n$$
More concisely: $= \sum_{i=1}^{n} a_i * b_i$

In addition, one property of the dot product is that it is linear. Note that linearity requires that a function satisfies $f(x + ry) = f(x) + rf(y)$ – later on, we will discuss operations and functions that do not satisfy these conditions. For now, we will provide a simple proof of the dot product's linearity:

Let:
$$\phi_{\vec{v}}(\vec{u}) = \vec{u} \cdot \vec{v}$$
$$\phi : \mathbb{R}^n \to \mathbb{R}$$

To show that the dot product is a real, linear function, we must prove that $\phi_u(x + ry) = u \cdot \vec{x} + r * \vec{u} \cdot \vec{y}$. This is a simple computation, which you will see below:

Proof. Show that $\phi_u(x + ry) = u \cdot \vec{x} + r * \vec{u} \cdot \vec{y}$

$\phi_{\vec{u}}(\vec{x} + r * \vec{y})$
$= \vec{u} \cdot (\vec{x} + r * \vec{y})$
$= \vec{u} \cdot \vec{x} + \vec{u} \cdot (r * \vec{y})$ (Dot product distributivity)
$= \vec{u} \cdot \vec{x} + r * \vec{u} \cdot \vec{y}$ (Dot product scalar multiplication)
$= \phi_{\vec{u}}(\vec{x}) + r * \phi_{\vec{u}}(\vec{y})$ (By definition) ∎

Within the proof, we made use of properties we have not discussed yet. We will discuss them now:

Let $\vec{a}, \vec{b}, \vec{c}$ be vectors in $\mathbb{R}^n$ and r_1, r_2 be arbitrary scalars in $\mathbb{R}$:
- $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$ (Commutative)
- $\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$ (Distributive over vector addition)
- $(r_1 * \vec{a}) \cdot (r_2 * \vec{b}) = (r_1 * r_2)(\vec{a} \cdot \vec{b})$ (Scalar multiplication)
- $\vec{a} \cdot (\vec{b} \cdot \vec{c}) \neq (\vec{a} \cdot \vec{b}) \cdot \vec{c}$ (Lack of associativity) - This is because the dot product of a vector and a scalar does not make sense

2.4 p-norm

The p-norm is much like the magnitude, but is applied for any exponent p instead of just p = 2. It is denoted as follows:

Let:
$$\phi_p(\vec{v}) = (\textstyle\sum_{i=1}^{n} v_i^p)^{1/p}$$
$$\phi : \mathbb{R}^n \to \mathbb{R}$$

For the purpose of this course, we will not use this apart from proving its linearity or a small subset of proofs. This will be taught more in-depth in later topology courses.

2.5 Orthogonality/perpendicular vectors

We can check if a vector is perpendicular to another using a useful property of the dot product.

Let $\vec{a}$, $\vec{b}$ be two orthogonal/perpendicular vectors in $\mathbb{R}^n$. Then:
$$\vec{a} \cdot \vec{b} = 0$$

Why does this make sense?

Recall that the dot product is the product of the magnitudes of some two vectors and the cosine of the angle between the two vectors; it makes sense then that the dot product of two perpendicular vectors results in 0 because the cosine value at 90 degrees (or $\frac{\pi}{2}$) is 0.

2.6 Projection

Geometrically, the projection of vector u on vector v refers to the perpendicular/orthogonal projection of u to v.

The projection formula is as follows:

Let $\vec{u}, \vec{v}$ be vectors in $\mathbb{R}^n$.
$$proj_{\vec{v}}(\vec{u}) = \frac{\vec{u} \cdot \vec{v}}{||\vec{v}||^2} * \vec{v} = \frac{\vec{u} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} * \vec{v}$$

2.7 Angle between two vectors

This formula is related to the projection and dot product, and allows us to determine the angle between two vectors. Its formula is shown below:

Let $\vec{u}, \vec{v}$ be vectors in $\mathbb{R}^n$.
Then, $\theta = \arccos(\frac{\vec{u} \cdot \vec{v}}{||\vec{u}||||\vec{v}||})$.
Alternatively: $\cos(\theta) = \frac{\vec{u} \cdot \vec{v}}{||\vec{u}||||\vec{v}||}$

2.8 Lines

We will define a line as:

$r(t) = \vec{p} + t\vec{d}$ where $\vec{p}$ is a point-vector in $\mathbb{R}^n$, $\vec{d}$ is the directional vector in $\mathbb{R}^n$ and t is some real number.

However, it is important to note that the line equation given above does not guarantee a line that is parallel to the original. The point-vector comes from some point on the line and the directional vector is a vector that touches the new line. When we plug in some value for t, we will obtain the sum of two vectors that result in the final line, but is not necessarily parallel to $\vec{p}$ nor $\vec{d}$.

2.9 Planes

Planes are geometrical representations of 2-dimensional flat surfaces that extend infinitely far. We represent these with the following formula:

Let $\vec{n} = [n_1, n_2, n_3]$ be a vector in $\mathbb{R}^3$ that is perpendicular to $\vec{p}$, $\vec{q} \in \mathbb{R}^n$ and let point P $(p_1, p2_p3)$ be a point on the plane.

Then, the equation below holds:

$$(n_1)x + (n_2)y + (n_3)z = D, \text{ where D is a number in } \mathbb{R}.$$

We may obtain D by plugging in point P's x, y and z components:

$$(n_1)(p_1) + (n_2)(p_2) + (n_3)(p_3) = D$$

And plug D back into the original formula:

$$(n_1)x + (n_2)y + (n_3)z = D$$

2.10 Triangle inequality and the Cauchy-Schwartz inequality

Although we haven't quite looked over complex numbers yet, we will also take a look at the complex version. The triangle inequality works for real vectors as well as real and complex numbers.

Real numbers : $|a + b| \leq |a| + |b|$ where $a, b \in \mathbb{R}$
Real vectors : $||\vec{a}| + |\vec{b}|| \leq ||\vec{a}|| + ||\vec{b}||$ where $\vec{a}, \vec{b} \in \mathbb{R}^n$
Complex numbers : $|z_1 + z_2| \leq |z_1| + |z_2|$ where $z_1, z_2 \in \mathbb{C}$

The Cauchy-Schwartz inequality for Vectors ($\mathbb{R}^n$) states that:

$$|\vec{a} \cdot \vec{b}| \leq |\vec{a}| * |\vec{b}|$$

We may use it to make certain calculations, like finding the maximum or minimum given two equations, etc.

3 Complex numbers

In high school, you have covered basic concepts with complex numbers. Here, we will introduce the polar and Euler's form, solving complex equations and algebra in the complex numbers.

3.1 Definition

Complex numbers are numbers that can be expressed in the form $a + ib$, where $a, b \in \mathbb{R}$. In general, the variable z is used to express this; in other words: $z = a + ib$. Note that i is pre-defined as $\sqrt{-1}$. Interestingly, it is because i cannot be defined as a real number that it is called an "imaginary" number.

3.2 Polar form and Euler's form

The polar form and Euler's form of complex numbers is another way of defining a complex number. The following formula has 2 equivalent forms:

$$z = a + ib$$
$$\text{Let } a = \cos(\theta),\ b = \sin(\theta). \text{ Then, the formula } z = a + ib \text{ is equal to:}$$
$$re^{i\theta} = r(\cos(\theta) + i\sin(\theta)).$$

Note that $r = |a + ib| = \sqrt{a^2 + b^2}$ and that $re^{i\theta}$ is Euler's form, the latter $r(\cos(\theta) + i\sin(\theta))$ being the polar form. You may also encounter the formula above as $r * cis(\theta) = re^{i\theta} = r(\cos(\theta) + i\sin(\theta)))$

3.3 Algebra and properties of complex numbers

While many aspects of the algebra you may use in the real numbers remain the same in complex numbers, there are certain properties that you should keep in mind for complex numbers:

Let $z_1 = a + ib$, $z_2 = \alpha + i\beta$; $a,\ b \in \mathbb{R}$. Then:

1. $z_1 + z_2 = (\alpha + a) + i(\beta + b)$
2. $z_1 * z_2 = (a + ib)(\alpha + i\beta) = (a\alpha - b\beta) + i(a\beta + b\alpha)$
3. Let $z_1 = r_1 e^{i\theta_1}$, $z_2 = r_2 e^{i\theta_2}$, $z_1 * z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}$
4. $\overline{z_1} = \overline{a + ib} = a - ib$ (Conjugation of complex numbers)
5. $\overline{\overline{z}} = z$
6. $|z|^2 = z * \bar{z}$

7. $\overline{z_1 * z_2} = \overline{z_1} * \overline{z_2}$
8. $Re(z) = a$ and $Im(z) = b$ (Real/imaginary components)
9. $z^{-1} = \frac{\bar{z}}{|z|^2}$

3.4 Finding distinct, complex roots

We will define a formula that will permit us to calculate distinct roots in the complex numbers. To solve an equation in the format $z^n = a$:

$$z^n = a \Rightarrow (re^{i\theta})^n = r_a(e^{i\theta_a})$$

Since trigonometric functions are periodic, we must account for it now:

$$z = (r_a e^{i\theta})^{1/n} = r_a^{1/n}(\cos(\tfrac{\theta_a+2k\pi}{n}) + i\sin(\tfrac{\theta_a+2k\pi}{n}))$$

Note that n is the degree of z, and $k = \{0, 1, 2, ..., n-1\}$. As an exercise, try calculating the roots of $z^4 = 1 + 0i$. The solution is below:

Find $r = \sqrt{1^2 + 0^2} = 1$
$\implies z^4 = a = (r_a e^{i\theta_a}) = (e^{i(0)}) = 1(\cos(0) + i\sin(0)) = 1$
$\implies z = a^{1/4} = 1^{1/4}$
Case 1: k = 0
$z = 1^{1/4}(\cos(0) + i\sin(0)) = 1 * (1 + 0i) = 1$
Case 2: k = 1
$z = 1^{1/4}(\cos(\frac{0+2\pi(1)}{4}) + i\sin(\frac{0+2\pi(1)}{4})) = 1 * (0 + 1i) = i$
Case 3: k = 2
$z = 1^{1/4}(\cos(\frac{0+2\pi(2)}{4}) + i\sin(\frac{0+2\pi(2)}{4})) = 1 * (-1 + 0i) = -1$
Case 4: k = 3
$z = 1^{1/4}(\cos(\frac{0+2\pi(3)}{4}) + i\sin(\frac{0+2\pi(3)}{4})) = 1 * (0 - 1i) = -i$

Thus, the solutions of z are $z = \pm i$ or $z = \pm 1$.

4 Linear independence and span

These concepts are fundamental to linear algebra, and tie in to concepts you may have learned in high school (e.g solving linear systems, for example).

4.1 Linear combinations

Linear combinations are expressions constructed from a set of terms (for example, vectors or matrices). We'll denote the form for vectors, although it is the same for matrices and other constructs.

Let $S = \{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ and let $r_1, r_2, ..., r_n \in \mathbb{R}$.
The set of linear combinations of S is as follows: $r_1\vec{v}_1 + r_2\vec{v}_2 + ... + r_n\vec{v}_n$

4.2 Linear independence/dependence

If some sequence of vectors $S = \{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ from a vector space is linearly independent, then:

Let $S = \{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ and let $r_1, r_2, ..., r_n \in \mathbb{R}$.
$r_1\vec{v}_1 + r_2\vec{v}_2 + ... + r_n\vec{v}_n = 0$ and $r_1 = r_2 = ... = r_n = 0$

Alternatively, it is said that if $A\vec{x} = \vec{b}$ has an unique solution, then A is linearly independent. This is useful for matrices and other constructs, but is also valid for vectors.

Consider the following example (see section ... for an example of linear independence):
Let $\vec{a} = [1, 2], \vec{b} = [2, 4]$.

Proof. Show that $\vec{a}$ and $\vec{b}$ are linearly independent.

We notice that the vectors appear quite similar; $\vec{b}$ is just $2\vec{a}$.
Now, our motivation is to disprove that it is linearly independent by means of a counterexample.

Recall the definition of linear independence:
$r_1\vec{v}_1 + r_2\vec{v}_2 + ... + r_n\vec{v}_n = 0$ and $r_1 = r_2 = ... = r_n = 0$

In this case, it is in two dimensions, so we want to prove that:
$r_1\vec{a} + r_2\vec{b} = 0$ and $r_1 = r_2 = 0$

Suppose $r_1 = -2, r_2 = 1$. Then:
$-2[1, 2] + 1[2, 4] = 0$ and $r_1 \neq r_2 \neq 0$
This represents a contradiction of the linear independence definition. Thus, it is linearly dependent.

■

4.3 Span

Intuitively, the span of something is the smallest thing that can contains that something. For example, you might have a pencil of a certain length that can "span" the length of a particular direction on a desk (perhaps you need 5 such pencils to reach the horizontal length; we can then call the pencil a "basis" for the horizontal length).

In linear algebra, the span of a set S of vectors in a particular vector space is the smallest "subspace" that contains this particular set. Earlier, we discussed the second dimension (R^2).

By applying some knowledge of vectors, we notice that $span([1,0],[0,1]) = R^2$ given that any vector in R^2 can be represented by a linear combination of the vectors within the span.

Mathematically, we may say that:

1. The subspace spanned by a non-empty subset S of a vector space V is the set of all linear combinations of vectors in S.

2. Every spanning set S of a vector space V must contain at least as many elements as any linearly independent set of vectors of V. That is, the span of a particular space (for instance R^n) must contain a minimum of n linearly independent vectors.

3. Any set of vectors than spans a vector space V can be reduced to a basis for V by discarding un-needed vectors if necessary. Generally, we discard linearly dependent vectors from the span.

For example, if we have $span([1,0],[0,1],[2,3])$, we can discard $[2,3]$ because it can be written as a linear combination of $[1,0]$ and $[0,1]$ (in other words, it is linearly dependent). This leaves us with $span([1,0],[0,1]) = R^2$ again.

5 Matrices and their properties

We won't go into much detail about solving linear systems with matrices as this is something you can easily learn on your own, but we will cover a few operations like the transpose of a matrix, or a matrix in upper triangle form, etc.

5.1 Basic matrix notation and operations (addition, scalar and matrix multiplication)

In general, we represent matrices in two ways:

1. Graphically: $\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,m} \\ \vdots & & & & \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,m} \end{bmatrix}$

2. With notation: Let $A \in M_{n*m}(\mathbb{R}/\mathbb{C})$.

(The two representations above are equivalent)

Note: In the graphical representation of the matrix, do note that some people use the indices "11," "12," and so on to represent "row 1, column 1" and "row 1, column 2". However, for the sake of clarity, this matrix uses the notation "i, j" (row i, column j).

With the second representation, we are defining A to be a matrix of n by m size (in other words, n rows and m columns).

Matrix addition: Define A and B to be matrices such that $A, B \in M_{n*m}(\mathbb{R}/\mathbb{C})$. This does not mean A and B are the same matrix, though!

Then, $A + B = [a_{i,j} + b_{i,j}]$ (i.e. the resulting matrix comes from adding each value at index i and j for each matrix)

Scalar multiplication: Let $A \in M_{nm}(\mathbb{R}/\mathbb{C})$ and r be a scalar. Then, $rA = [r * a_{i,j}]$.

Matrix multiplication: Let $A \in M_{n*m}(\mathbb{R}/\mathbb{C}), B \in M_{j*k}(\mathbb{R}/\mathbb{C})$. It is imperative that you understand that matrix multiplication can only occur if the number of columns on the first matrix is the same as the number of rows on the second matrix.

Matrix multiplication works by taking the "dot product" of the first row of the first matrix and the first column of the second matrix to produce the first value of the matrix and so on. More concisely:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} * \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ \vdots & & & \\ b_{j1} & b_{n2} & \dots & b_{jk} \end{bmatrix} = \begin{bmatrix} (A,\ row\ 1) \cdot (B,\ col\ 1) & \dots & (A,\ row\ 1) \cdot (B,\ col\ j) \\ (A,\ row\ 2) \cdot (B,\ col\ 1) & \dots & (A,\ row\ 1) \cdot (B,\ col\ j) \\ \vdots & & \vdots \\ (A,\ row\ n) \cdot (B,\ col\ 1) & \dots & (A,\ row\ 1) \cdot (B,\ col\ j) \end{bmatrix}$$

5.2 Matrix forms

Definitions:

Upper/lower triangular matrices: Upper/lower triangular matrices are special kinds of square matrices (where the number of rows are the same as the number of columns) where for a upper triangular matrix, all of the entries under the main diagonal are zero while for a lower triangular matrix, all of the entries above the main diagonal are zero.

Row echelon form (REF): In the first non-zero entry of each row (the pivot), anything under that value within the column must be 0.

Reduced row echelon form (RREF): The matrix must be in REF form, every pivot must be 1 and each pivot is the only non-zero entry in its column.

5.3 Matrix invertibility

A matrix of n by n size is invertible iff $\exists A^{-1} = C \in M_{n*n}$ such that $AC = CA = \mathbb{1}_{n*n}$. Note: $\mathbb{1}_{n*n}$ is a n by n sized identity matrix, which is a matrix with the main diagonal filled by 1 and every other value 0)

Singular matrix: If a matrix $A \in M_{n*n}$ is not invertible, then it is called singular.
Theorem: The inverse of a matrix A (noted as A^{-1}) is unique.
Equivalent statements for invertibility:

1. $A\vec{x} = \vec{b}$ has a unique solution iff a matrix with columns $\vec{v}_1, \vec{v}_2$, ..., $\vec{v}_n$ are linearly independent.
2. Matrix $A \in M_{n*n}$ is invertible/has an inverse.
3. Matrix $A \in M_{n*n}$ can be reduced to the identity matrix.
4. For a matrix $A \in M_{n*n}$, A^T is invertible iff A is invertible.
5. $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ are a basis of $V = span\{\vec{v}_1, ..., \vec{v}_n\}$.

5.4 Matrix transpose

The transpose of a matrix $A \in M_{n*n}(\mathbb{R}/\mathbb{R})$ is denoted by A^T. It holds the following properties:

1. For complex matrices: $A^T = (\bar{A})^T$

Let $A \in M_{n*n}(\mathbb{R})$ be a square matrix.

2. A is a symmetric matrix if $A^T = A$ for both real and complex matrices.
3. A is a skew symmetric matrix if $A^T = -A$ for both real and complex matrices.
4. Transpose is a real linear map.
5. $(A^T)^T = A$
6. $(A + B)^T = A^T + B^T$
7. $(rA)^T = rA^T$
8. $(AB)^T = B^T A^T$

5.5 Matrix trace

The trace of a matrix $A \in M_{n*n}(\mathbb{R}/\mathbb{R})$ is defined as $tr(A) = \sum_{i=1}^{n} a_{ii}$. In other words, it is the sum of all entries in the main diagonal.

Properties of matrix trace:

1. $tr(A^T) = tr(A)$
2. $tr(rA) = r * tr(A)$
3. $tr(A+B) = tr(A) + tr(B)$
4. $tr(AB) = tr(BA)$ (but $BA \neq AB$)

6 Vector spaces and subspaces

6.1 Requirements and definitions

A vector space on set V satisfies addition and scalar multiplication such that $r \in \mathbb{R}/\mathbb{C}$ for the following properties:

1. $\vec{u} + \vec{v} = \vec{v} + \vec{u}\ \forall \vec{u}, \vec{v} \in V$
2. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})\ \forall \vec{u}, \vec{v}, \vec{w} \in V$
3. $\exists 0 \in V$ such that $0 + \vec{v} = \vec{v} + 0\ \forall \vec{v} \in V$
4. For any $\vec{v} \in V$, there $\exists \vec{w} \in V$ such that $\vec{w} + \vec{v} = 0$
5. $\exists\ e$ such that $e * \vec{v} = \vec{v} * e = \vec{v}$ for $\forall \vec{v} \in V$
6. $r(\vec{u} + \vec{v}) = r\vec{u} + r\vec{v}$ for $\forall \vec{u}, \vec{v} \in V, r \in \mathbb{R}/\mathbb{C}$

And it must also satisfy the following conditions, which is also the requirements for a subspace:

1. Non-empty (contains the 0 vector): $\vec{0} \in V$
2. Closed under addition: For $\forall \vec{u}, \vec{v} \in V$, $\vec{u} + \vec{v} \in V$
3. Closed under scalar multiplication: For $\forall \vec{u} \in V r \in \mathbb{R}/\mathbb{C}$, $r\vec{u} \in V$

(Definition) Subspace: A subspace S of vector space V is a subset of V with the same addition and scalar multiplication as well as S also being a vector space.
(Definition) Basis of a vector space: The basis of a vector space V^n is a collection of non-zero vectors $\{\vec{b}_1, \vec{b}_2, ..., \vec{b}_n\} \in V^n$ (the set of vectors must be linearly independent and they must span V^n by definition).
(Definition) Dimension of V^n The dimension of a vector space V^n denoted by $dim(V)$ is equivalent to the number of bases of V^n.

6.2 Sum of subsets

Suppose $u_1, ..., u_m$ are subsets of a vector space V. The sum of $u_1, ..., u_m$ denoted by $u_1 + u_2 + ... + u_m$ is the set of all possible sums of elements of $u_1, u_2, ..., u_m$.

Mathematically: $u_1 + u_2 + ... + u_m = \{a_1 + a_2 + ... + a_m \mid a_i \in u_i \text{ for } \forall i \in [1, m]\}$

6.3 Direct sums

Suppose $u_1, ..., u_m$ are subspaces of V. The sum of $u_1, ..., u_m$ is called a direct sum if each element of $u_1 + u_2 + ... + u_m$ can be written in only one way as a sum (i.e. each element must be represented as a unique sum) of $a_1 + a_2 + ... + a_m$ where $a_i \in u_i$.

We denote this as: $u_1 \oplus u_2 \oplus ... \oplus u_m$.

6.4 Row space, column space and null-space/kernel

(Definition) Row space: The row space (which also is a subspace) is $span\{\vec{R}_1, \vec{R}_2, ..., \vec{R}_m\}$ where $\vec{R}_i$ represents the vector formed by row i of a matrix.
(Definition) Column space: The column space (which also is a subspace) is $span\{\vec{C}_1, \vec{C}_2, ..., \vec{C}_m\}$ where $\vec{C}_i$ represents the vector formed by column i of a matrix.
(Definition) Null space: The null space or kernel (which also is a subspace) of a matrix A is the set formed by $N(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = 0\}$.
Theorem: If A is a matrix of m by n size such that $m > n$ (the number of rows is greater than the number of columns), then $A\vec{x} = \vec{b}$ has a unique solution $\Leftrightarrow$ A can be reduced to the identity matrix with $m - n$ rows of 0s.

7 Linear transformations

7.1 Definition

A function defined by $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if for $\forall \vec{u}, \vec{v} \in \mathbb{R}^n, r \in \mathbb{R}$ the following properties are satisfied (linearity):

1. $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$
2. $T(r\vec{u}) = rT(\vec{u})$

Theorem (linear transformations must fix 0): Linear transformations must fix 0 (i.e. $T(\vec{0}) = \vec{0}$)

7.2 Properties of linear transformations

Properties:

1. $\mathbb{R}^n$ is called the domain of T
2. $\mathbb{R}^m$ is called the range of T
3. Let $W \subseteq R^n$. The image of W under T is $\{T(\vec{w}) \mid \vec{w} \in W\} = T[W]$
4. The range of T is $T[R^n] = \{T(\vec{v}) \mid \vec{v} \in \mathbb{R}^n\}$
5. If $W' \subseteq R^m$, the inverse image of W' under T is $T^{-1}[W'] = \{\vec{v} \in \mathbb{R}^n \mid T(\vec{v}) = W'\}$
6. The kernel of T is the set $\{\vec{v} \in \mathbb{R}^n \mid T(\vec{v}) = 0\}$.
7. For a set of vectors $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k \in \mathbb{R}^n$, if $T(\vec{v}_1), T(\vec{v}_2), ..., T(\vec{v}_k)$ are linearly independent, then the set of vectors $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_k\}$ are also linearly independent.
8. Let $A \in M_{m*n}$. $T(\vec{x}) = A\vec{x} \forall \vec{x} \in \mathbb{R}^n$ if the matrix A's columns are represented by $T(\vec{e_1}), T(\vec{e_2}), ..., T(\vec{e_n})$ where $\{\vec{e}_1, ..., \vec{e}_n\}$ is a basis for $\mathbb{R}^n$. Note that e_i represents a vector with every value being 0 except for the value at index i being 1.

7.3 Rank and nullity

The rank of a linear transformation, in essence is the dimension of the range of the transformation. Mathematically, this is defined as:

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation with its standard matrix representation $T(\vec{x}) = A\vec{x}$.

Then, the rank of T is: $rank(T) = dim(range(T)) = dim(\text{column space of } A)$

On the other hand, the nullity of a linear transformation is simply the dimension of the null-space or kernel of the linear transformation. Mathematically, this is defined as:

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation with its standard matrix representation $T(\vec{x}) = A\vec{x}$.

Then, the nullity of T is: $nullity(T) = dim(kernel(T))$

Additional theorems:

1. T is invertible $\Leftrightarrow$ A is invertible.

• If A is invertible, then T is an isomorphism.

2. The rank-nullity equation of linear transformations is defined as: $rank(T) + nullity(T) = dim(domain(T)) \Leftrightarrow dim(range(T)) = dim(domain(T)) - dim(kernel(T))$

7.4 Relationship between linear transformations and subspaces

Linear transformations have certain connections to subspaces. In particular, there are some important theorems that you should remember in the following section.

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

1. If W is a subspace of R^n, then $T(W)$ is a subspace of R^m
2. If W' is a subspace of R^m, then the pre-image of W' is a subspace of $\mathbb{R}^n$.

7.5 Injectivity, surjectivity, bijectivity, isomorphism

Definitions:

1. Injectivity (or one-to-one) is defined as: If $T(\vec{u}) = T(\vec{v})$, then $\vec{u} = \vec{v}$ for $\forall \vec{u}, \vec{v} \in \mathbb{R}^n$. The contrapositive can also be useful, so it is provided here for your convenience: If $\vec{u} \neq \vec{v}$, then $T(\vec{u}) \neq T(\vec{v})$. Equivalently, $kernel(T) = \{\vec{0}\}$.

2. Surjectivity (or onto) is defined as: For $\forall \vec{v} \in \mathbb{R}^m$, there $\exists \vec{u} \in \mathbb{R}^n$ such that $T(\vec{u}) = \vec{v}$. There are two other equivalent forms: $range(T) = \mathbb{R}^m$ and $T(\mathbb{R}^n) = \mathbb{R}^m$

3. Isomorphism: The linear transformation T is injective and surjective (i.e. bijective), which implies that the standard matrix representation of T is invertible. Additionally, if T is isomorphic, then T^{-1} is unique.

3.1. An isomorphism preserves the number of bases through a transformation. If two spaces share the same dimension, they are isomorphic.

4. $\mathcal{L}(V, W)$ means all possible linear transformations from V to W.

8 Linear Maps

8.1 Fundamental Theorem of Linear Maps

Suppose V is a finite-dimensional space for $dim(V) = n$, $T \in \mathcal{L}(V, W)$ and $range(T)$ is finite-dimensional. Then: $dim(V) = null(T) + rank(T)$ (derived from the rank equation).

8.2 Dimension of two vector spaces that are isomorphisms

Suppose there are vector spaces V and W that are isomorphic (i.e. $\exists T : V \to W$ s.t. T is an isomorphism) $\Leftrightarrow dim(V) = dim(W)$. Since the two spaces are isomorphic, their kernels are 0; as a result, $dim(V) = rank(T) = dim(W)$.

8.3 $\mathcal{L}(V, W)$ is a vector space

$\mathcal{L}(V, W) \simeq \{M_{m*n}\}$ (i.e. is isomorphic to a matrix). $\mathcal{L}(V, W)$ is a vector space as it is closed under addition and scalar multiplication.

8.4 Dimension of $\mathcal{L}(V, W)$

The dimension of $\mathcal{L}(V, W)$ is given by the formula $dim(\mathcal{L}(V, W)) = dim(M_{m*n}) = m * n = dim(V) * dim(W)$

9 Determinant

The determinant is a scalar derived from a certain computation performed on a square matrix and encodes certain properties of the linear transformation defined by the matrix. The determinant of a matrix is denoted as $det(A)$. Note that the determinant of any 1 by 1 matrix is just the one value in the matrix.

9.1 Relationship to cross product

You may remember computing the cross product; in effect, the cross product is derived from the determinant of a 3 by 3 matrix. For a matrix defined by:

$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$

The determinant of this matrix is $det(A) = (a_1)(b_2c_3 - b_3c_2) - (a_2)(b_1c_3 - b_3c_1) + (a_3)(b_1c_2 - b_2c_1)$. If we change a_1, a_2, a_3 to i, j, k, the determinant of the matrix is exactly that of the cross product, where the second row and third rows are your vectors.

9.2 Area of parallelogram and volume of parallelepiped

With some work, it is possible to show that the determinant of two linearly independent vectors in a matrix is equivalent to the area of a parallelogram formed by those two vectors:

$det(\begin{bmatrix} \leftarrow \vec{a} \rightarrow \\ \leftarrow \vec{b} \rightarrow \end{bmatrix})$ = (Area of the parallelogram formed by vectors $\vec{a}$, $\vec{b}$) $= ||\vec{b}|| \cdot ||\vec{a}|| sin\theta$

By doing the same thing but with three linearly independent vectors in a matrix, the determinant can also be found to be the volume of a parallelepiped:

$det(\begin{bmatrix} \leftarrow \vec{a} \rightarrow \\ \leftarrow \vec{b} \rightarrow \\ \leftarrow \vec{c} \rightarrow \end{bmatrix})$ = (Volume of the parallelepiped formed by vectors $\vec{a}$, $\vec{b}$, $\vec{c}$)

9.3 Co-factor

The cofactor of a_{ij} in a matrix $A \in M_{n*n}$ is defined by $a'_{ij} = (-1)^{i+j} det(A_{ij})$ where A_{ij} is the minor for a_{ij}. For an n-dimensional matrix, we have that:

$$det(A) = a_{11} * a'_{11} + a_{12} * a'_{12} + ... + a_{1n} * a'_{1n}$$

Properties of cofactors:

If A is a n by n sized matrix and a'_{ji} is the cofactor of a_{ji}, then:

1. $a_{i1}a'_{j1} + a_{i2}a'_{j2} + ... + a_{in}a'_{jn} = \begin{cases} det(A) & i = j \\ 0 & i \neq j \end{cases}$

2. $a_{1i}a'_{1j} + a_{2i}a'_{2j} + ... + a_{ni}a'_{nj} = \begin{cases} det(A) & i = j \\ 0 & i \neq j \end{cases}$

9.4 Adjoined matrix

Adjoined matrices are defined by $adj(A)$ for a matrix A.

Suppose $A \in M_{n*n}$. Then, the adjoined matrix of A is a matrix with a transpose where each entry is the cofactor of each entry of A. That is, each element at index i, j is a'_{ij}.
Properties (where $A \in M_{n*n}$):
1. $(adj(A))A = A(adj(A)) = det(A) * \mathbb{1}_{n*n}$

9.5 Properties of determinants

There are various properties that may be useful for proofs using determinants. They are as follows:
1. $det(A) = det(A^T)$
2. If A is a triangular matrix, then $det(A)$ is the product of all of its diagonal entries.
3. $Row_i \leftrightarrow Row_j$ where the two rows are rows in A and Row_i is in matrix A, Row_j is in matrix B.
4. If matrix A has two equal rows, then $det(A) = 0$.
5. $A : Row_i \to r * R_i$ in B then $det(B) = r * det(A)$. This also has the effect of producing: $det(r * A) = r^n * det(A)$
6. $A : Row_i \to Row_i + r * R_j$ in B then $det(A) = det(B)$.
7. If A contains proportional rows or columns, then $det(A) = 0$.
8. $det(AB) = det(A) * det(B)$ for n by n sized matrices A and B.
9. If A is invertible, then $det(A^{-1}) = \frac{1}{det(A)}$
10. If $det(A) > 1$, then the resulting shape grows larger. If it is smaller, than the shape shrinks, and if it is 1 exactly, the shape is the same.

9.6 Cramer's Rule

If $A\vec{x} = \vec{b}$ is a system of n linear equations in n unknowns and $det(A) \neq 0$, then the unique solution $\vec{x} = [x_1, x_2, ..., x_n]$ is of the form:

$$x_k = \frac{det(B_k)}{det(A)} \text{ for } k = 1, 2, ..., n$$

where B_k is the matrix A with the k^{th} column replaced by the column vector $\vec{b}$.

10 Eigenvalues and Eigenvectors

10.1 Definition

Let $A \in M_{n*n}$. A scalar λ is an eigenvalue of A is there is a non-zero vector $\vec{v} \in \mathbb{R}^n$ such that $A\vec{v} = \lambda\vec{v}$. In this case, $\vec{v}$ is called an eigenvector of A corresponding to the eigenvalue λ.

Summarized:
- $(\lambda, \vec{v})$ must be a pair
- $A\vec{v} = \lambda\vec{v}$
- $\vec{v} \neq 0$

10.2 Equivalent statements

If λ is an eigenvalue of $A \in M_{n*n}$, then the following statements are equivalent:

- $\exists\vec{v} \neq 0$ such that $A\vec{v} \neq \lambda\vec{v}$

- $\exists\vec{v} \neq 0$ such that $(A - \lambda\mathbb{1})\vec{v} = 0$ where $\mathbb{1}$ is an identity matrix of n by n size. In addition, $(A - \lambda\mathbb{1})\vec{v} = 0$ can be seen as a linear transformation that takes a non-zero vector v and sends it to the kernel (because the image is 0). In addition, $(A - \lambda\mathbb{1})$ is singular (non-invertible and non-isomorphic).

- Note: $A - \lambda\mathbb{1}$ is a matrix of n by n size $\Leftrightarrow T : V^n \to V^n$

10.3 Characteristic polynomial

Let $A \in M_{n*n}$. The characteristic polynomial of A is given by $p(\lambda) = |A - \delta\mathbb{1}|$. If λ is an eigenvalue of A, then the set $E_\lambda = \{\vec{x} \in \mathbb{R}^n | A\vec{x} = \lambda\vec{x}\} \Leftrightarrow E_\lambda = \{\vec{x} \in \mathbb{R}^n | (A - \lambda\mathbb{1})\vec{x} = 0\}$ where $(A - \lambda\mathbb{1})\vec{x} = 0$ is a homogeneous system. Roughly speaking, E_λ represents the null-space of $(A - \lambda\mathbb{1})\vec{x} = 0$.

Computing characteristic polynomials

Start with $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & -1 \\ 1 & 3 & -2 \end{bmatrix}$.

To compute the characteristic polynomial, we must take the determinant:

$det(A - \lambda\mathbb{1}) = det(\begin{bmatrix} 2-\lambda & 0 & 0 \\ 1 & 2-\lambda & -1 \\ 1 & 3 & -2-\lambda \end{bmatrix}) = (2-\lambda)(\lambda-1)(\lambda+1)$

Clearly, $det(A - \lambda\mathbb{1}) = 0$ when $\lambda = 2, \lambda = 1, \lambda = -1$. Computing their eigenvectors:

$E_{\lambda_1=-1} = null(A - (-1)\mathbb{1}) = null(A + \mathbb{1})$

By reducing the matrix from $(A + \mathbb{1})\vec{x} = 0$, we eventually get: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow x_1 = 0, x_2 = \frac{s}{3}, x_3 = s \Rightarrow E_{\lambda_1=-1} = span([0, \frac{1}{3}, 1])$. The corresponding eigenvector for $\lambda_1 = -1$ then is $[0, \frac{1}{3}, 1]$.

By repeating the same process for the rest of the matrices:

$E_{\lambda_2=1} = span([0, 1, 1])$
$E_{\lambda_3=-2} = span([1, 1, 1])$
We may obtain the remaining corresponding eigenvectors to their eigenvalues.

10.4 Linear transformation definition of eigenvalues and eigenvectors

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. A scalar λ is an eigenvalue of T if there exists a non-zero vector $\vec{v}$ such that $T(\vec{v}) = \lambda\vec{v}$. The vector $\vec{v}$ then is an eigenvector of T corresponding to λ.

10.5 Properties of eigenvalues and eigenvectors

Let $A \in M_{n*n}$. Let λ be an eigenvalue of A and $\vec{v}$ be an eigenvector of A corresponding to the eigenvalue λ.

1. λ^k is an eigenvalue of A^k and $\vec{v}$ is an eigenvector of A^k corresponding to λ^k where $k \in \mathbb{Z}^{\geq 1}$.
2. A is invertible $\Leftrightarrow \lambda = 0$ is not an eigenvalue of A.
3. If A is invertible, then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} and $\vec{v}$ is an eigenvector of A^{-1} corresponding to $\frac{1}{\lambda}$.
4. If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, then the eigenvalues of T are the eigenvalues of the standard matrix representation of T.
5. A eigenvector cannot correspond to 2 distinct eigenvalues for $A \in M_{n*n}$.
6. Let $\lambda_1, \lambda_2, ..., \lambda_k$ be distinct eigenvalues of $A \in M_{n*n}$. Let β_i be a basis of the eigenspace $E_i = null(A - \lambda_i \mathbb{1})$. Then, $\beta = \bigcup_{i=1}^{k} \beta_i$ is linearly independent.

11 Algebraic and geometric multiplicities

Definition: Let $A \in M_{n*n}$ with a characteristic polynomial $p(\lambda)$ with distinct eigenvalues $\lambda_1, \lambda_2, ..., \lambda_r$. Suppose that $(\lambda - \lambda_i)^{\alpha_i}$ is a factor of $p(\lambda)$, but $(\lambda - \lambda_i)^{\alpha_i+1}$ is not. Then, we say that α_i is the algebraic multiplicity of λ_i. If $dim(E_{\lambda_i}) = \beta_i$, then we say β_i is the geometric multiplicity of λ_i.

Properties and theorems:

1. The geometric multiplicity of an eigenvalue of $A \in M_{n*n} \leq$ its algebraic multiplicity

12 Diagonalizable matrices

12.1 Definition

Let $A, D \in M_{n*n}$.

1. D is a diagonal matrix if all its entries not on the main diagonal are zero.
2. A is diagonalizable if there $\exists P \in M_{n*n}$ that is invertible s.t. $P^{-1}AP = D$ ($P^{-1}AP$ is a diagonal matrix).
3. Equivalently: $PDP^{-1} = A$

12.2 General properties and theorems

Listed are some properties and theorems:

1. Let $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ be eigenvectors corresponding to the distinct eigenvalues (i.e. multiplicities of 1) $\lambda_1, \lambda_2, ..., \lambda_k$ of the square matrix A. Then, the vectors $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ are linearly independent.

2. Let $A \in M_{n*n}$. A is diagonalizable $\Leftrightarrow$ A has n linearly independent eigenvectors.

3. $A^n = PD^nP^{-1}$

4. If $A \in M_{n*n}$ and is invertible, and $B \in M_{n*n}$ then the null-space of AB is the same as the null-space of B and $rank(AB) = rank(B)$.

Proof. Proof of theorem 4.

Suppose $A, B \in M_{n*n}$ and A is invertible.
Notice that A is an invertible matrix. Thus, it can be seen as a isomorphic linear transformation $T_1 : V^n \to V^n$. Since T_1 is an isomorphism, the bases are preserved through the transformation. We don't know if B is isomorphic, but we can define it as a linear transformation $T_2 : V^n \to V^n$.

Looking at $T_1(T_2)$, T_2 takes an input V^n and outputs something in V^n, then T_1 takes the output of T_2 in V^n and maps it to an output in V^n.

Recall that we need to show that $rank(AB) = rank(B)$ and $null(AB) = null(B)$. Suppose T_2 sends k bases to the range, and $n - k$ bases to the kernel. So, $rank(B) = k$, $null(B) = n - k$. Notably, the $n - k$ bases that are sent to the kernel represent a result of $T_2(...) = 0$. By the definition of linear transformations, $T_1(0) = 0$. Thus, the null-space of B is the same as the null-space of A (A is linearly independent by its invertibility). Lastly, the dimension of bases is the same between A and B since T_2 has n bases, and being sent to T_1 also produces n bases as it is linearly independent. Thus, $rank(AB) = rank(B)$.

■

Continued:

5. Let $A \in M_{n*n}$. A is diagonalizable if and only if the characteristic polynomial of A is a product of linear factors (but not necessarily distinct) and the algebraic multiplicity of each eigenvalue is equal to its geometric multiplicity.

12.3 Similar matrices

Definition: If $A, B \in M_{n*n}$ such that $B = P^{-1}AP$ for some invertible matrix P, then we say that A and B are similar matrices.

Properties:

1. $det(A) = det(B)$

2. A is invertible $\Leftrightarrow$ B is invertible
3. $rank(A) = rank(B)$
4. $nullity(A) = nullity(B)$
5. $det(A - \lambda\mathbb{1}) = det(B - \lambda\mathbb{1})$

That's it, folks. Good luck on your exams!