MATH 133: Linear Algebra

Final Revision Guide

DOWN WITH PREP 101!

1 Linear Equations and Matrices

1.1 Matrices

Main idea: matrices are rectangular arrays of numbers. The size of a matrix is determined by the number of rows m and the number of columns n. So, matrices are referred to by these sizes as $m \times n$ matrices.

- We say that two matrices are **equal matrices** if and only if two conditions are met:
  1. The matrices have the same size.
  2. The corresponding entries of the matrices are the same.

1.1.1 Matrix Operations

There are two fundamental matrix operations: matrix addition and matrix multiplication.

- Matrix **addition** is said to be **well-defined** if and only if the matrices are of the same size. For example, if matrix A was given by $p \times q$ and matrix B was given by $s \times t$, then the sum is well-defined if and only if the following equality holds: $p = q = s = t$.
- Matrix **scalar multiplication** is the process of multiplying a matrix A by some scalar α.
  - The resulting matrix would be of the same size.
  - Scalar multiplication is distributive, commutative, and associative.
- The **transpose** of a matrix occurs when an $m \times n$ matrix A is flipped over its main diagonal.
  - The resulting matrix is denoted by A^T.
  - The resulting matrix A^T is of size $n \times m$.
  - The **main diagonal** of a matrix $A = [a_{ij}]$ is given by all entries where $i = j$. That is:
  - The transpose's columns are the rows of A and the transpose's rows are the columns of A.

$$A = \begin{bmatrix} \mathbf{a_{11}} & * & * & * & * \\ * & \mathbf{a_{22}} & * & * & * \\ * & * & \mathbf{a_{33}} & * & * \\ * & * & * & \mathbf{a_{44}} & * \\ * & * & * & * & \mathbf{a_{55}} \end{bmatrix}$$

- Matrix **multiplication** is the process of multiplying two matrices together. We say that the matrix multiplication of a $p \times q$ matrix A and a $s \times t$ matrix B is **well defined** iff
  1. $q = s$.

2. The size of the new matrix, AB, is $p \times t$. If the multiplication were BA, then the size of the new matrix would be $t \times p$.
3. Matrix multiplication is not necessarily commutative.
4. The identity matrix, and any scalar multiple of the identity matrix, commutes with every matrix.

Theorem 1 (Properties of Transposes). *Let A and B denote matrices of the same size, and let α denote a scalar. Then, we obtain:*

1. *A is $m \times n \Rightarrow A^T$ is an $n \times m$ matrix.*
2. $(A^T)^T = A$
3. $(\alpha A)^T = \alpha A^T$
4. $(A + B)^T = A^T + B^T$

Theorem 2. *Properties of matrix multiplication] The important ones:*

1. $IA = A = AI = A$
2. $(AB)^T = B^T A^T$

Types of matrices

- We say that a matrix is **symmetric** if $A^T = A$.
  - Such a matrix is **necessarily** a square matrix.
  - If not a square matrix, the size of the transpose would be different.
- We say that a matrix is **skew-symmetric** if $A^T = -A$. **Remark**: the entries on the diagonal of a skew-symmetric matrix are always zero.
- We say that a matrix is **square** if $m = n$.
- We can obtain the **trace** of a matrix by summing up the entries on the main diagonal of a matrix. This is denoted by $tr(A)$.
- A **diagonal matrix** is one that is obtained by multiplying a scalar α by the identity matrix. As such, it only contains one value, α, on the main diagonal and zeros elsewhere.

1.2 Linear Equations

Definition 1.1. Linear Equation $\rightarrow$ we say that an equation is linear if it is in the form

$$\alpha x_1 + \alpha x_2 + \cdots \alpha_n x_n = b$$

The **variables** of this system are $x_1, x_2, \cdots x_n$, the **constants** of this system are $\alpha_1, \alpha_2, \cdots \alpha_n$, and the **b**

The **solution** of a linear equation is the set of numbers, $s_1, s_2, ..., s_n$ that satisfy that equation. That is:

$$\alpha_1 s_1 + \alpha_2 s_2 + \cdots \alpha_n s_n = b$$

must hold. The solution set is expressed as $X = [s_1 s_2 \cdots s_n]^T$.

We say that we have a **system of equations** when we have multiple equations describing the same variables. There are several ways of understanding this. Say that we are given the following system of equations:

$$\alpha_1 x + \alpha_2 y + \alpha_3 z = b_1$$
$$\beta_1 x + \beta_2 y + \beta_3 z = b_2$$
$$\gamma_1 x + \gamma_2 y + \gamma_3 z = b_3$$

How we can understand this:

MATRIX FORM

The matrix form separates a system of equations into a **matrix of coefficients**, denoted as A, a column-matrix of **unknown variables**, denoted as $\vec{x}$, and another matrix of constants called the **matrix of constants**. It is given by:

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$
$$A\vec{x} = b$$

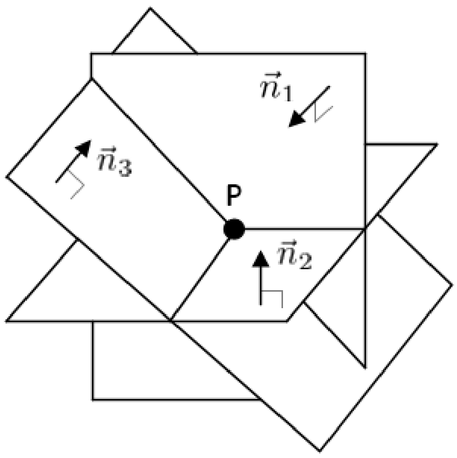
COLUMN PICTURE

This picture involves looking at a linear system of equations as a **linear combination** of the columns of the system of equations. In this case, it would look like this:

$$x \begin{bmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{bmatrix} + y \begin{bmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{bmatrix} + z \begin{bmatrix} \alpha_3 \\ \beta_3 \\ \gamma_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

ROW PICTURE

The row picture involves plotting all the planes (if we are in $\mathbb{R}^3$). If the system of linear equations has a unique solution, then we know that these planes will all intersect at one point $[x, y, z]^T$.



- There are three possible outcomes when we solve a system of equations:
  - A **consistent linear system**:
    1. Consistent linear system with one unique solution.
       (a) Geometric interpretation: this corresponds to an intersection point (of two lines if in $\mathbb{R}^2$ or of three planes if in $\mathbb{R}^3$.
    2. Consistent linear system with infinite solutions.

(a) Geometric interpretation: this corresponds to the two objects being in fact the same exact object. For example, in $\mathbb{R}^2$ this corresponds two the same exact line and in $\mathbb{R}^3$ this corresponds to the objects being the same exact plane.

3. A system is consistent if and only if in its row-echelon form...
   (a) the corresponding right-hand side for every non-zero rows in the coefficient matrix is zero, which is same as
   (b) coefficient matrix rank equals the rank of the augmented matrix.

– An **inconsistent system**, where there exist no solutions.
  1. Geometric interpretation: this corresponds to two objects having no contact points at all in space. In $\mathbb{R}^3$ this corresponds to parallel planes and in $\mathbb{R}^2$ this corresponds to parallel lines.
  2. You can recognize this by looking for contradictions when you row-reduce a matrix. (For example, $0 = 4$ is a contradiction, and indicates an inconsistent system.

- How can we find the solutions to a system of equations?
  - We need to introduce the **augmented matrix**, which combines the matrix of constants and matrix of coefficients. Using **elementary operations**, we reduce this matrix to (reduced) row-echelon form. This entire process is called **Gaussian elimination**.
  - **Elementary operations** are operations performed on a system of equation that yields an equivalent system of linear equation. They are:
    1. Interchanging two equations
    2. Multiplying an equation by a non-zero scalar.
    3. Adding a multiple of one equation to a different equation.
  - We say that a matrix is in **row-echelon form** if the following conditions are satisfied:
    1. All zero rows are at the bottom.
    2. The first non-zero entry from the left of each non-leading row is one (these are called **leading 1's**.
    3. Each leading 1 is to the right of all leading rows above it.
    4. This form is not necessarily unique to a particular matrix.
  - We say that a matrix is in **reduced row-echelon form** if a matrix is in row-echelon form AND all entries above each leading 1 is also zero.
    1. Row and reduced row echelon form is useful because we can obtain the solution set.
    2. This form is necessarily unique to the matrix.
- How does the **Gaussian Algorithm** work?
  1. Using elementary row operations, carry the augmented matrix to reduced-row echelon form.
  2. Assign the non-leading variables as parameters.
  3. Use **back substitution**, the process of solving for leading variables from the bottom up-wards, to determine the values of the leading values in terms of the parameters.

Definition 1.2. Rank $\rightarrow$ the rank is the number of leading 1's in a matrix when it is in row-echelon form. This is denoted by $rank(A)$.

- If we have a system of m equations (so, m rows) in n variables (so, n columns), and if the rank of the matrix is of rank r, then...

- The solution has exactly $(n-r)$ parameters, because by definition the system has $(n-r)$ non-leading variables.

1.3 Homogeneous Systems

Definition 1.3. Homogeneous system $\rightarrow$ a system of linear equations where all the constants are equal to zero. The general form of a homogenous system is:

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots \alpha_n x_n = 0$$

Remarks

- A homogeneous system is always consistent.
- If a homogenous system has more variables than equations (that is, n is bigger than m), then the system will have non-trivial solutions. In fact, it will have infinitely many because parameters must be involved.
- A **trivial solution** is one where every variable equals zero, and you obtain that zero equals zero.
- A **non-trivial solution** is a solution where at least one of the variables is a non-zero number.

How do we express solutions?
We can express solutions as linear combinations, basic solutions, and general solutions. For example, say that we obtain the following solution:

$$X = \begin{bmatrix} 2s+t \\ s \\ -2t \\ t \end{bmatrix} = \begin{bmatrix} 2s \\ s \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} t \\ 0 \\ -2t \\ t \end{bmatrix} = s \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

- The **basic solution** corresponds to the coefficient matrices of the parameters. In this example, the basic solution is:

$$X_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} X_2 = \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

- The **general solution** is a linear combination of the basic solutions, with the scalars being the parameters obtained from the Gaussian algorithm.

$$X + sX_1 + tX_2$$

More remarks

- The Gaussian algorithm produces exactly $n-r$, where n is the number of variables and r is the rank of the matrix, basic solution.
- Every solution can be expressed as a linear combination of the basic obtained from the Gaussian algorithm.

1.4 Matrix Multiplication

Motivation: We can combine matrix multiplication, homogeneous systems, and linear equations to obtain something else: the **associated homogeneous system**.

Definition 1.4. Associated homogeneous system → the associated homogeneous system to a system of equations, $A\vec{x} = b$, is the system that sets the matrix of constants equal to zero:

$$A\vec{x} = 0 \tag{1}$$

This is simpler to solve, and it can help us find general solutions. How?

<u>Main idea</u>: **every solution to the system $AX = B$ has the form of $\mathbf{X} = \mathbf{X_0} + \mathbf{X'}$, where $\mathbf{X'}$ is the solution to the associated homogeneous system $\mathbf{AX} = \mathbf{0}$ and X_0 is a solution to the system $\mathbf{AX_0} = \mathbf{B}$**

1.5 Matrix Inverses

Motivation: because matrix inverses "undo" matrices, they can be used to isolate variables, and consequently find solutions to systems of equations.

Definition 1.5. Inverse → the inverse of a matrix, denoted A^{-1}, is the matrix that when multiplied by itself, produces the identity matrix.

Remarks

- The inverse of A is **uniquely determined**.
- In order for a matrix to have an inverse, it must be a **square matrix**.
- The inverse of a 2×2 matrix is given by the following formula:

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{2}$$

A MATRIX A IS INVERTIBLE IF AND ONLY IF (THE FOLLOWING ARE ALL EQUIVALENT...

1. the reduced-row echelon form of A is the identity matrix.
2. The system $A\vec{x} = b$ has a unique solution for any b.
3. A is the product of elementary matrices.
4. the determinant of A is NOT zero.
5. the rows of A are linearly independent.
6. the columns of A are linearly independent.
7. the associated homogeneous system $A\vec{x} = 0$ only contains the trivial solution (a unique solution).
8. 0 is not an eigenvalue of A.

Using these properties of matrix inversion, we obtain the **matrix inversion algorithm**:

- Augment the matrix A with the identity matrix I_n as follows:

$$[\ A \mid I\ ]$$

  The point of augmenting A with I is to have a sort of "clean slate" to keep track of the operations performed to carry A to reduced row-echelon form.
- Carry A to RREF by performing elementary row operations on the augmented matrix.

- Once A is RREF, ten you will obtain the following:

$$[\ I \mid A^{-1}\ ]$$

Remark 1.1. The product of elementary matrices that bring the matrix to reduced row-echelon form is given by $E_n \cdots E_1 E_2$. The product is exactly equal to the inverse A^{-1}.

Remark 1.2. It follows that if A^{-1} is given by $E_n \cdots E_1 E_2$, then we can invert this to obtain A as a product of elementary matrices: $A = E_1^{-1} \cdots E_n^{-1}$

Properties of Inverses
If the inverse A^{-1} exists, then the following apply:

- $(A^{-1})^{-1} = A$
- $(AB)^{-1} = (B^{-1})(A^{-1})$
- $(A_1 A_2 \cdots A_k)^{-1} = (A_k)^{-1} \cdots (A_2)^{-1}(A_1)^{-1}$
- $(\alpha A)^{-1} = \frac{1}{\alpha}(A)^{-1}$
- $(A^k)^{-1} = (A^{-1})^k$
- $(A^T)^{-1} = (A^{-1})^T$

If A is invertible, then we can obtain the unique solution as:

$$\vec{x} = A^{-1} b \tag{3}$$

1.6 Elementary Matrices

Motivation: how can we tie the Gaussian algorithm into this mess? The answer is with elementary matrices.

Definition 1.6. Elementary matrix → an elementary matrix is a matrix that performs an elementary operation on a matrix. It is obtained by performing an elementary operation on the identity matrix. There are three types:

1. **Type I elementary matrix**: obtained by interchanging two rows.
   - The inverse is given by re-interchanging those two rows.
2. **Type II elementary matrix**: obtained by adding a multiple of one row to another.
   - The inverse is given by subtracting that same multiple from the row.
3. **Type III elementary matrix**: multiplying a row by a non-zero scalar.
   - The inverse is given by dividing that same row by $1/alpha$, where alpha is the scalar.

Performing an elementary row operation on a matrix is the same exact thing as left-multiplying that matrix by its corresponding elementary matrix. Performing an elementary column operation on a matrix is the same exact thing as right-multiplying a matrix by its corresponding elementary matrix.

2 Determinants and Eigenvalues

Motivation: we know how to find a determinant of a 2-by-2 matrix (we have the formula). The purpose of this section is to create a method to determine the determinant of other square matrices.

2.1 Cofactor Expansions

Motivation: cofactors help us compute determinants for matrices larger than 2-by-2.

Definition 2.1. Determinant suppose we already know how to compute the determinant of any $(n-1) \times (n-1)$ matrix (inductive step). Then, the determinant is defined as follows:

$$det(A) = a_{11}C_{11}(A) + a_{12}C_{12}(A) + ... + a_{1n}C_{1n}(A) \tag{4}$$

where...

- the a_{1n} term is the $(1, n)$ entry of the matrix.
- the $C_{1,n}$ entry, the **cofactor**, is obtained by multiplying the determinant of the matrix obtained when deleting the ith row and jth column from the matrix by $(-1)^{i+j}$.
- this is also called a **laplace expansion** of A along row 1. You can actually do this along any row or column.

Remark 2.1. If a square matrix A has a row or columns of zeroes, then $det(A) = 0$. Technique: the more zeroes, the better. So, use elementary row or column operations to make the column nice before actually carrying the expansion out.

Table 1: Effect of ERO on the determinant

ERO on matrix A to produce B	**Effect on det(A)**
Interchanging two rows	$det(B) = -det(A)$
Multiplying a row/column by a scalar α	$det(B) = \alpha det(A)$
Adding multiples of a row/column to a different row/column	$det(B) = det(A)$

Remark 2.2.

- If a square matrix has two identical rows or columns, then the determinant is equal to zero.
- Properties of determinants to know: given a n-square matrix, then we have the following.
  1. $det(\alpha A) = \alpha^n det(A) = (a^n)det(A)$
  2. $det(A^T) = det(A)$
  3. If A is triangular, then its determinant is given by:
  $$det(A) = a_{11} + a_{22} + a_{33} + ... + a_{nn} \tag{5}$$
  4. $det(A_1A_2A_3 \cdots A_n) = det(A_1)det(A_2) \cdots det(A_n)$
  5. $det(A^{-1}) = \frac{1}{det(A)}$

2.2 Determinants and Inverses

Motivation: we can use the determinant to obtain the value of the inverse. The value of the determinant will also tell is if the exists an inverse for a matrix.

Definition 2.2. Adjoint $\rightarrow$ the adjoint of an $n \times n$ matrix A is given by **transposing the matrix of coefficients** of A.

$$adj(A) = [c_{ij}(A)]^T$$

This is nice because putting all of it together, we get a formula to obtain the inverse of any square matrix (if it's invertible).

$$A^{-1} = \frac{1}{det(A)} adj(A) \tag{6}$$

Some other identities/formulas to know cold:

- $A \times adj(A) = (det) \times I = adj(A) \times A$
- $det(adj(A)) = det(A)^{n-1}$

Definition 2.3. Crammer's Rule $\rightarrow$ If we have a linear system of equations $A\vec{x} = b$, where A is an invertible $n \times n$ matrix, x is a vector in $\mathbb{R}^n$, then we can obtain the value of any individual variable x_i by replacing the ith column with b:

$$x_i = \frac{det[A_i(b)]}{det(A)} \tag{7}$$

2.3 Diagonalization and Eigenvalues

Motivation: finding an efficient way to calculate large powers of square matrices. These are easy to do with **diagonal matrices**, matrices containing non-zero numbers only on the main diagonal.

- The "formula" for diagonalizing a matrix A is as such:

$$A = P^{-1}DP^{-1} \tag{8}$$

  where P^{-1} is the inverse of P and D is a diagonal matrix.

Definition 2.4. We say that λ is an **eigenvalue** of the matrix A if $AX = \lambda X$ holds for a certain column X. In this case, X is the **eigenvector** corresponding for λ. Rearranging the equation, we obtain:

$$X(\lambda I - A) = 0 \tag{9}$$

So, the crux of these problems is finding values of λ such that $det(\lambda I - A) = 0$ and the corresponding eigenvectors for the values of λ.

How do we find eigenvalues and their corresponding eigenvectors?

1. Obtain the **characteristic polynomial** of the matrix, A, using the following equation:

$$C_A(x) = det(xI - A) \tag{10}$$

   λ is an eigenvalue of a matrix iff it is a root of the polynomial.
2. Once you obtain all the eigenvalues, solve for the homogenous system (basic solutions) $\vec{X}(\lambda I - A) = \vec{0}$ (once for each eigenvalue) to obtain the eigenvectors corresponding to the eigenvalues.
   (a) λ is an eigenvalue of iff $(\lambda I_n - A)\vec{x} = 0$ has a non-trivial solution $\rightarrow$ any eigenvalue has infinitely many eigenvectors.
3. **Diagonalizing the matrix**. Once we obtain the eigenvalues and their corresponding eigenvectors, we can diagonalize the matrix. The form would be as such:

$$A = [X_1 X_2 \cdots X_n] \begin{bmatrix} \lambda_1 & 0 & \cdots 0_n \\ 0 & \lambda_2 & \cdots 0_n \\ 0 & 0 & \cdots \lambda_n \end{bmatrix} [X_1 X_2 \cdots X_n]^{-1} \tag{11}$$

Definition 2.5. The **geometric multiplicity** of an eigenvalue λ of $A_{n\times n}$ is: (a) the number of parameters needed to describe the solutions of $(\lambda I_n - A)\vec{x} = \vec{0}$) and (b) the difference $n - r$, where r is the rank of $\lambda I_n - A$. It is also the dimension of the eigenvalue's eigenspace.

The **algebraic multiplicity** of an eigenvalue is the number of times it appears in the factorization of the characteristic polynomial; consequently, the sum of all algebraic multiplicities is at most n and strictly less than n if there exist irreducible factors. Consequently, the following conditions are the same:

- λ is an eigenvalue of A.
- The homogeneous system $(\lambda I_n - A)\vec{x} = 0$ has a non-trivial solution (so, infinitely many).
- $\lambda I_n - A$ is a **non-invertible matrix**.
- $det(\lambda I_n - A) = 0$.
- λ is a real root of A's characteristic polynomial.

Remarks/Facts to know:

1. Given an eigenvalue, their algebraic and geometric multiplicities are at least one.
2. An eigenvalue's geometric multiplicity is at most equal to the algebraic multiplicity.
3. A matrix A is diagonalizable iff
    - there exists an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.
    - the characteristic polynomial has no complex roots and the geometric multiplicity of each eigenvalue is exactly equal to the algebraic multiplicity.
    - the sum of the eigenvalues' geometric multiplicity $= n$.

2.3.1 Similar Matrices

Definition 2.6. We say that two matrices A and B are **similar**, denoted by $A \sim B$, if there exists an invertible matrix P such that

$$A = PB^{-1}P \tag{12}$$

Properties of Similar Matrices if $A \sim B$, then we obtain the following...

- $A \sim A$
- $A \sim B \to B \sim A$
- if $A \sim B$ and $B \sim C$ then $A \sim C$
- $A^{-1} \sim B^{-1}$
- $A^T \sim B^T$
- $A^k \sim B^k \; \forall k \in \mathbb{N}$
- $det(A) = det(B)$
- $C_A(x) = C_B(x)$
- A and B have the same eigenvalues (this is a one-way implication, not a bi-conditional)

3 Vector Geometry

Quantities that are completely specified by a single number are called **scalar quantities** and quantities that require both a magnitude and direction to specify it are called **vectors**.

3.1 Geometric Vectors

Let $\vec{v} = [xyz]^T$ and $\vec{w} = [x_1 y_1 z_1]^T$. Then, we obtain the following:

- The **magnitude** of a vector, given by $||v||$ is equal to $\sqrt{x^2 + y^2 + z^2}$.
- We say that two vectors are **equal** if their components are equal. That is: $v = w \iff x = x_1, y = y_1,$ and $z = z_1$.
- $\vec{v} = \vec{0} \iff ||\vec{v}|| = 0$.
- $||\alpha\vec{v}|| = |\alpha| ||\vec{v}||$

Definition 3.1. We say that a vector is a **unit vector** if it has a length equal to one. If we are given a vector $\vec{x} = [x_1 x_2 x_3]^T$, then a unit vector in the direction of $\vec{x}$ is given by $\frac{\vec{u}}{||\vec{u}||}$.

- The following conditions are equivalent for two non-zero vectors:
  - $\vec{u}$ and $\vec{w}$ are parallel.
  - Each of $\vec{v}$ and $\vec{w}$ are scalar multiples of each other.
  - One of $\vec{v}$ and $\vec{w}$ is a scalar multiple of the other.

3.2 Dot product and Projections

Motivation: The **dot product** is a test of perpendicularity. It allows us to deduce information about the angle between any two non-zero vectors, making it a useful object for optimization problems.

- Given vectors $\vec{v} = [x_1, x_2, x_3]^T$ and $\vec{w} = [w_1, w_2, w_3]^T$, the dot product is given by $\vec{v} \cdot \vec{w} = x_1 w_1 + x_2 w_2 + x_3 w_3$.
- Dot product is a SCALAR quantity.
- $\vec{v} \cdot \vec{v} = ||\vec{v}||^2$
- Dot product commutes and is distributive.
- If θ is the angle between two non-zero vectors $\vec{v}$ and $\vec{w}$, then we obtain

$$\vec{v} \cdot \vec{w} = ||\vec{v}|| ||\vec{w}|| \cos\theta$$

- If the dot product is bigger than zero, we can deduce that the angle is acute. If the dot product is exactly zero, then we can deduce that the vectors are orthogonal to each other. If the dot product is smaller than zero, then the angle is obtuse.

Definition 3.2. The **projection** decomposes a vector into two orthogonal vectors. f we let $\vec{v}$ and $\vec{d}$ be two non-zero vectors. Then, the projection of $\vec{V}$ into $\vec{d}$ is given by

$$proj_{\vec{d}}(\vec{v}) = \left(\frac{\vec{v} \cdot \vec{d}}{||\vec{d}||^2}\right)\vec{d} \tag{13}$$

The projection decomposition takes any vector v and writes it as a **linear combination** of two orthogonal vectors.

3.3 Lines and Planes

The **direction vector** $\vec{d}$ is the non-zero vector that is parallel to the line containing the point we are interested in. For a given line $y = mx + b$, the direction vector is given by

$$\vec{d} = \begin{bmatrix} 1 \\ m \end{bmatrix}$$

The vector equation of a line is given by $\Lambda = \vec{p_0} + t\vec{d}$. To see if a set of points lie on the line, write the equation out in parametric form, and see if the resulting system of equations (with t as the variable) is consistent. Consequently, a line is determined by its direction vector and a point on the line.

A **plane** is defined in terms of it normal vector and any single point on the plane. A non-zero vector $\vec{n}$ is called a **normal vector** to a plane π if $\vec{n}$ is orthogonal to $\vec{v}$ for all $\vec{v}$ on the plane. A plane is an infinitely flat surface.

Scalar Equation of a plane

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

where a, b, and c correspond to the entries of the matrix representing the normal vector, and $x_0, y_0, and z_0$ correspond to the points we choose to define the plane.
A plane in **parametric form** would be described with the following equation: $x = p_0 + tu + sv$.

3.3.1 Optimization Problem Algorithms

Find the shortest distance between...

1. **Point to a point**: just use the distance formula to obtain the distance.
2. **Point to a line**: two sub-cases - (1) if the point lies on the line, then the distance is zero. (2) If the point does not lie on the line, then use the following formula
$$\frac{||P_1P_0 \times u||}{||u||}$$
where u is the direction vector of the line, P_0 is the point we are investigating, and P_1 is the point on the line that the line is defined by.
3. **Point to plane**: given a point P_0 and a plane, the distance is given by
$$\frac{|ax_0 + by_0 + xz_0 + d|}{\sqrt{a^2 + b^2 + c^2}} \tag{14}$$
4. **Line to line**: two sub-cases - (1) if the lines intersect, the distance is zero. (2) if the lines do not intersect, find the points on both the lines such that the dot product of their direction vectors is zero. Then, using the distance formula, find the distance between those two points.

The Cross Product
Motivation: how can we describe a plane only using two vectors on the plane and a point. And, how can we compute the area of the parallelogram defined by two vectors.

Definition 3.3. The **cross-product** is given by the following formula:

$$\vec{u} \times \vec{v} = det \begin{bmatrix} i & x_1 & x_2 \\ j & y_1 & y_2 \\ k & z_1 & z_2 \end{bmatrix} \tag{15}$$

Key Points

- $\vec{v} \times \vec{w} = 0 \iff \vec{v}$ and $\vec{w}$ are parallel.
- $\vec{u}(\vec{v} \times \vec{w}) = det[\vec{u}\vec{v}\vec{w}]$
- $\vec{u} \times \vec{w}$ is a vector.
-

3.4 Matrix Transformations of R^2

A **matrix transformation** of R^2 is a map T: $R^2 \to R^2$: it maps a 2×1 vector to another 2×1 vector. There are only 3 times of linear transformations:

1. A **rotation** by an angle θ. The inducing matrix is given by

$$R_0 = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \tag{16}$$

   The inverse of a rotation is simply the same matrix, except with $-\theta$.
2. The **projection** into the line $y = mx$ is given by:

$$P_m = \frac{1}{1+m^2} \begin{bmatrix} 1 & m \\ m & m^2 \end{bmatrix} \tag{17}$$

3. The **reflection** in the line $y = mx$ is given by:

$$R_m = \frac{1}{1+m^2} \begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix} \tag{18}$$

   The inverse of the reflection is simply a reflection again over the same line, so the inverse of the reflection is the same reflection itself (that is, $A^{-1} = A$).

A translation is NOT a linear transformation.

What makes a transformation LINEAR
We say that a transformation is linear if it satisfies two properties:

1. Linearity: $T(\vec{u} + \vec{v}) = T(\vec{w}) + T(\vec{v})$
2. Scalar multiplication: $T(\alpha\vec{v}) = \alpha T(\vec{v})$

In 3 blue 1 brown language, this means that gridlines remain evenly spaced and that the origin remains fixed. A linear transformation is linear $\iff$ it can be described with a matrix.

Composition of Transformations

- If S denotes one transformation and T denotes another one, and they are represented by A and B respectively, then the transformation $S \circ T$ (first T, then S) is given by AB.
- An **isometry** is a linear transformation that that preserves the distance between any two vectors. In particular:

$$||T(v) - T(w)|| = ||v - w|| \tag{19}$$

  - In fact, there are only two types of isometries: those with $|det(a)| = 1$. This means that if A is an isometry, then it is either a reflection or rotation.
  - T is a rotation iff $det(A) = 1$
  - T is a reflection iff $det(A) = -1$

The Unit Square
The **unit square** is the parallelogram determined by the coordinate vectors $\vec{i}$ and $\vec{j}$. Given a transformation, the parallelogram determined by $T(\vec{u})$ and $T(\vec{v})$ is called the **image** if the parallelogram determined by $\vec{v}$ and $\vec{w}$.

- The image under T of the unit square, where T is a transformation induced by the matrix A, is det(A).

- "Profs love putting two parts of the course together and watching you get confused" - TA:
  - A projection has two eigenvalues: $\lambda = 0$ (because you lose a dimension) and $\lambda = 1$ (because all vectors on the line that is being projected into remain untouched).
  - A reflection has two eigenvalues: $\lambda = -1$ and $\lambda = 1$.
  - A rotation is an example of an invertible matrix with NO eigenvalues, unless $\theta = k\pi$, where $k \in \mathbb{Z}$.

4 The Vector Space R^n

4.1 Subspaces and Spanning

A **subset** of R^n is a subspace if the following three conditions are met

1. The zero vector is contained within the subspace U.
2. U is closed under addition.
3. U is closed under scalar multiplication.

Key points on this topic

- Lines and planes through the origin are the ONLY proper subspaces
- The **Null Space** is the set of all vectors that get sent to the origin by a transformation induced by A. It is the **solutions to the homogeneous system**.
- The **image** is the set of all columns Y in R^m such that the system $AX = Y$ has a solution. Intuitively, this means that the image is the set of all vectors in R^m (for an $m \times n$ matrix) that can be mapped to by vectors in R^m.
- The **eigenspace** is the set of all X in R^n such that $AX = \lambda X$. It is the space spanned out by the eigenvectors for a given eigenvalue.
- We can use the Gaussian Algorithm to find the null space of a given matrix A.
  - Augment A with the matrix of 0.
  - Row-reduce the matrix until you obtain I.
  - Write the solutions in parametric form. The **general solution** is an integer-linear combination of the basic solutions.
- A **spanning set** describes the vectors that can be used (as linear combinations) to obtain certain areas of a subspace.

4.2 Linear Independence

SIMPLE TEST: put all the vectors as rows in a matrix and row-reduce the mat

- We say that a set of vectors are **linearly independent** iff the only vanishing linear combination of the vectors is the trivial one (all constants are zero).
  - There's no way to obtain any of the vectors as a linear combination of the other vectors.
- How this connects to matrix invertibility. The following conditions are equivalent
  1. A is invertible
  2. Columns are linearly independent in R^n.
  3. Columns span R^n
  4. Rows are linearly independent in R^n
  5. Rows of A span R^n

- If we are given two transformed vectors, then we can obtain A as such:

$$A = [T(\vec{u})T(\vec{v})][\vec{u}\vec{v}]^{-1}$$

A set of vectors is **linearly dependent** if a non-trivial linear combination of the vectors produces the vanishing solution. This means that one of the vectors can be expressed as a combination of the others.

- The vectors are linearly DEPENDENT if they are parallel.
- The vectors are linearly INDEPENDENT if they are not parallel.

4.3 Dimension

A point is zero-dimensional because you have zero degrees of freedom. A line is one-dimensional because you have one degree of freedom. A plane is two-dimensional because you have two degrees of freedom...

- **Fundamental Theorem**. If a subspace U is spanned by m vectors and U contains K linearly independent vectors, then k is at most m. Additionally, the rank is at most the number of columns that are linearly independent.
  - No linearly independent set in R^n can contain more than n vectors.
  - No spanning set for R^n can contain fewer than n vectors.
  - A **basis** of a subspace U of R^n is a set of vectors such that (1) The set is LINEARLY INDEPENDENT and (2) U is spanned by that set of vectors.
- **Invariance Theorem.** any two bases of a subspace must have the same number of elements. Consequence of the Fundamental Theorem.
- The **dimension** of a subspace U of R^n is just the **number of vectors in any basis of U**. The **standard basis** is the set of elementary columns that form a basis for that subspace U.
- **Independent Lemma**. If we have less than n linearly independent vectors in R^n, then if Y is a vector in R^n that is not in the span of the previous vectors, then the larger set is also entirely linearly independent.

4.4 Rank

Recall that the rank is the number of leading one's in the RREF of a matrix. This is also equal to the number of variables, or columns, minus the number of parameters (degrees of freedom) in the general solution. If we have an $m \times n$ matrix, then....

- The **column space (colA)** is the subspace of R^m spanned by the columns of the matrix A. It is a subspace of R^m.
- The **row space (rowA)** is the subspace of R^n spanned by the rows of A. It lives in R^n.

Remarks

- If A is carried to reduced row echelon form, then row(A)=row(RREF). This applies if it's brought there using row operations or column operations.
- The number of leading 1's in RREF is equal to the dimension of the row space of RREF and the dimension of the row space in A.
- **Rank Theorem**. If we let A be an $m \times n$ matrix, then **dim(row(A))** $=$ **dim(col(A))** $=$ **rank(A)**.
  - We obtain the row space by taking the rows containing the leading 1's in RREF.

  - We obtain the column space by taking all columns containing a pivot (leading 1) from the corresponding column in the original matrix A.

The following conditions are equivalent: if A is an m $\times$ n matrix, then the following conditions are the same:

1. the homogeneous system only has the trivial solution.
2. the columns of A are linearly independent
3. the rank of A is n
4. A^TA is an n $\times$ n invertible matrix.

The dimension of the null space is equal to n - r. Quick remarks

- If A is an $m \times n$ matrix, then rank(A) = rank(UAV) for any invertible matrixes U and V.
- For any $m \times n$ matrix, then $dim(imA) + dim(nullA) = n$.

Cramming. A matrix A is invertible iff

- det(A) $\neq$ 0
- 0 is not an eigenvalue for A
- There exists a matrix D such that DA = I
- Rank(A) = n
- Contains exactly n linearly independent columns
- Contains exactly n linearly independent rows
- Row space spans R^n
- Column space spans R^n
- A is the product of elementary matrices
- The homogeneous system AX =0 only contains the trivial solution
- The system AX = B has a unique solution for every b.
- A in RREF = I
- Null(A) = zero vector.