UBC 221 Matrix Algebra

Justin Hua

April 21, 2021

I would like to thank Evan Chen for letting me use his LaTeXpackage.

Contents

§1 March 23rd

Example 1.1 (Fibonacci's Fake Rabbit Problem)

Fibonacci starts with 1 pair of rabbits. Every month, each pair of rabbits produce 1 pair of offspring.

Let r_n be the number of pairs of rabbits after n months. So, $r_0 = 1, r_1 = 2, r_2 = 4$. We find the recursion: $r_{n+1} = 2r_n$. Solving yields: $r_n = 2^n$

Remark 1.2. This is an example of dynamical system.

Example 1.3 (Fibonacci's Real Rabbit Problem)

It takes newborn rabbits a month to grow up. Only after a month, can these rabbits start reproduce. Find a formula for the number of rabbits in the n^{th} month.

Proof. There are 2 types of rabbits: juveniles and adults. Let (j, a) denote the number of pairs of rabbits of each type. The state of the system is described by a vector:

$r = \begin{bmatrix} j \\ a \end{bmatrix}$ We look at the monthly change. Every adult pair of adults produce a pair of newborns. Adults also do not die, so they stay around. All juveniles turn to adults in a month.

$j_{n+1} = a_n$ and $a_{n+1} = j_n + a_n$

$$\begin{bmatrix} j_{n+1} \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} a_n \\ j_n + a_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} j_n \\ a_n \end{bmatrix}$$

So, $r_{n+1} = Ar_n$. We call this matrix A the *transition matrix.*

Let us calculate a few values of r_k $r_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $r_1 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

By induction, we find that $r_n = A^n r_0$

That is, $r_n = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

We assume the vectors of $\begin{bmatrix} j \\ a \end{bmatrix}$ form a line. Then we have that $Ar = \lambda r$ for some λ. We rewrite this as :$Ar - r = 0 \implies Ar - \lambda \mathrm{id}_2 r = 0 \implies (A - \lambda \mathrm{id}_2)r = 0$

So, any vector on this line will be a solution to this homogeneous system of equations. So, we need to solve: $\begin{bmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \cdot \begin{bmatrix} j \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ So, this system has non-trivial solutions. So, the matrix $\begin{bmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix}$ is not invertible. In other words, $det \begin{bmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} = -\lambda(1-\lambda) - 1 = \lambda^2 - \lambda - 1$. Using the quadratic equation, we find the 2 solutions: $\lambda = \frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}$. The latter is negative since we have a positive growth rate.

We can now conclude that as $n \to \infty$, the growth rate of rabbits approaches $\frac{1+\sqrt{5}}{2} \approx 1.6$.

We call $\frac{1+\sqrt{5}}{2}$ an **eigenvalue** of the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$

□

Definition 1.4 (Eigenvalues and Eigenvectors)**.** Let A be a square matrix(input and output vectors have to have the same size). Then a scalar λ is called an *eigenvalue* of A if there exist a nonzero vectors v such that $Av = \lambda \cdot v$. Any such non-zero vector v is called an *eigenvector* of A with corresponding eigenvalue λ

Proposition 1.5 (Fundamental Relation between Eigenvalues and characteristic polynomial)

The following are equivalent:

1. λ is an eigenvalue of A
2. There is a nonzero v such that $Av = \lambda v$
3. The homogeneous system with coefficient matrix has nontrivial solutions, ie: $(A - \lambda \mathrm{id}_n)v = 0$
4. $(A - \lambda \mathrm{id}_n)$ is not invertible.
5. $det(A - \lambda \mathrm{id}_n) = 0$

To drive the point home:

Theorem 1.6 (Eigenvalues of a matrix)

λ is an eigenvalue of a matrix A if and only if $det(A - \lambda \mathrm{id}_n) = 0$

Definition 1.7. The *characteristic polynomial* of A is the function $\lambda \mapsto det(A - \lambda \mathrm{id}_n)$

Example 1.8

The characteristic polynomial of the matrix in Fibonacci's rabbit problem is

$$det \begin{bmatrix} -\lambda & 1 \\ 1 & 1 - \lambda \end{bmatrix} = \lambda^2 - \lambda - 1$$

Proposition 1.9 (Relation of characteristic polynomial and eigenavlues)

λ is an eigenvalue of A iff λ is a root of the characteristic polynomial of A.

Definition 1.10 (Eigenspace)**.** Let λ be an eigenvalue of A. Then $E_\lambda = nul((A - \lambda \mathrm{id}_n)$ is called the *eigenspace* of λ. We have that $dim(Nul((A - \lambda \mathrm{id}_n)) \geq 1$ because it contains the eigenvectors and 0.

§2 March 25th

Example 2.1 (Reflections in $\mathbb{R}^2$)
Let $S : \mathbb{R}^2 \mapsto \mathbb{R}^2$ be the matrix transformation which reflects points across a line L. Thus, any non-zero vector v on L is an eigenvector of S with eigenvalue 1.

For a perpendicular line L' to L. Then any nonzero vector on L' is an eigenvector of of S with eigenvalue -1.

There are clearly no other eigenvalues.

There are 2 eigenspaces: E_1 and E_{-1}

Example 2.2 (Orthogonal Projections)
Let $P : \mathbb{R}^2 \mapsto \mathbb{R}^2$ be the projection onto the line L.

1 is an eigenvalue since all points on L are fixed by P.
For any point x on a perpendicular line, L', $P(x) = 0$ so the eigenvalue is 0. The eigenspace of $E_0 = L'$

Proposition 2.3 (Eigenvalue of 0)
0 is an eigenvalue of A if and only if $Nul(A - 0\text{id}_n) \neq \{0\}$. If 0 is an eigenvalue, then $E_0 = Nul(A)$

Example 2.4 (Rotations)
Let $R_\theta : \mathbb{R}^2 \mapsto \mathbb{R}^2$ be the rotation of angle θ about 0, where θ is not a multiple of 180

There are no nonzero vectors that get mapped to a scalar multiple of itself, so there are 0 *real* eigenvalues. However, there may be complex eigenvalues which will be discussed next week.

Example 2.5 (Dilation (say by factor 3))
Let $D : \mathbb{R}^2 \mapsto \mathbb{R}^2$ be the transformation that has the property $D(x) = 3x$ for all vectors x.

We see that all non-zero vectors in $\mathbb{R}^2$ are eigenvectors of D with eigenvalue $\lambda = 3$
The characteristic polynomial of the given matrix is:

$$det \begin{bmatrix} 3-\lambda & 0 \\ 0 & 3-\lambda \end{bmatrix} = (3-\lambda)^2$$

So, we know that 3 is a double root of the characteristic polynomial and that $dim(E_3) = 2$

Example 2.6 (Shear)

Consider the linear transformation given by the matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} : \mathbb{R}^2 \mapsto \mathbb{R}^2$

Notice that e_1 is a fixed point and e_2 maps to $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

For all other vectors, they are mapped to a perpendicular line which is parallel to the x-axis.

In general, a shear is in fact given by any matrix $\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$

Clearly, 1 is an eigenvalue where E_1 is the x-axis and the $dim(E_1) = 1$. There are no other eigenvalues. This holds for all shears for any a that we choose.

Example 2.7 (Back to Fibonnaci's Rabbit Problem)

Recall that we found the 2 eigenvalues $\lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_2 = \frac{1-\sqrt{5}}{2}$.
By graphing, we see that the 2 eigenspaces are perpendicular. However, this is a coincidence since our matrix is symmetric.
We do some stuff to get

$$x_n = c_1 \lambda_1^n v_1 + c_2 \lambda_2^n v_2$$

In this case,

$$x_n = c_1 \left(\frac{1+\sqrt{5}}{2} \right)^n v_1 + c_2 \left(\frac{1-\sqrt{5}}{2} \right)^n v_2$$

As

$$\lim_{n\to\infty}, x_n \approx c_1 \left(\frac{1+\sqrt{5}}{2} \right)^n \begin{bmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{bmatrix}$$

Definition 2.8. A $n \times n$ matrix is *diagonalizable* if there exists a basis of $\mathbb{R}^n$ consisting of eigenvectors for A

Proposition 2.9 (General Solution of Dynamical System)

If A is diagonalizable, then the general solution of the dynamical system is given by

$$x_k = c_1 \lambda_1^n v_1 + c_2 \lambda_2^n v_2 + ... + c_n \lambda_n^n v_n$$

§3 March 30th

Proposition 3.1 (Diagonizability)

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ be a linear transformation and let A be its associated matrix. Then the following are equivalent:

- T or A is diagonalizable
- There exists a basis of $\mathbb{R}^n$ consisting of eigenvectors for T or A
- There exists an invertible $n \times n$ matrix C and a diagonal matrix D such that $A = C \cdot D \cdot C^{-1}$. The columns of C are the eigenvectors of A and D has the eigenvalues $\lambda_1, ..., \lambda_n$ on the diagonal.

Proof. Say $v_1, ..., v_n$ is a basis of $\mathbb{R}^n$ of the eigenvectors of T with corresponding eigenvalues $\lambda_1, ..., \lambda_n$

Claim 3.2 — $AC = CD$

To do this, we check that the columns are equal. On the left, the i^{th} column is $A \cdot v_i$. On the right hand side, the i^{th} column is $C \cdot \lambda e_i$. Now, since v_i is an eigenvector, $A \cdot v_i = \lambda_i \cdot v_i = \lambda_i C e_i = C \cdot \lambda_i \cdot e_i$
Multiplying both sides on the right by C^{-1} gives the desired equation. □

Theorem 3.3 (Linear Independence of Eigenvectors of Distinct Eigenvalues)

If $\lambda_1, ..., \lambda_n$ are distinct eigenvalues of A and $v_1, ..., v_n$ are the corresponding eigenvectors, then $v_1, ..., v_n$ are linearly independent

Example 3.4

Consider the matrix $\begin{bmatrix} 2 & 7 & 6 \\ 0 & 3 & 2 \\ 0 & 0 & 5 \end{bmatrix}$

The characteristic polynomial is $(2-\lambda)(3-\lambda)(5-\lambda)$ There are 3 distinct eigenvalues and so the 3 corresponding eigenvectors are linearly independent.

Theorem 3.5 (Linearly independence of bases of eigenspaces for distinct eigenvalues)

Let $\lambda_1, \ldots, \lambda_r$ be distinct eigenvalues of A and the corresponding eigenspaces $E_1, \ldots, E_r$. If we put together all the bases of $E_1, \ldots E_r$, we get a linearly independent set of vectors.

Corollary 3.6 (Diagonalizability Criterion)

If the dimensions of each of the eigenspaces add up to n, then A is diagonizable.

Example 3.7 (Linearly independence of bases of eigenspaces for distinct eigenvalues)

Consider the matrix $\begin{bmatrix} 2 & 0 & ? & ? \\ 0 & 2 & ? & ? \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ There are 2 eigenvalues : $\lambda_1 = 2, \lambda_2 = 3$. Notice that $dim(E_2) = 2$ and $dim(E_3) = 2$. Notice that $E_2 + E_3 = 4$. If we take 2 vectors which form a basis for E_2 and 2 vectors which form a basis for E_3, the set of 4 vectors will be linearly independent.

Theorem 3.8 (Bounds on geometric multiplicity)

For every eigenvalue:

$$1 \leq \text{ geometric multiplicity (dimension of eigenspace)}$$
$$\leq \text{algebraic multiplicity (number of factors in the characteristic polynomial)}$$

Corollary 3.9 (Diagonalizability Criterion)

If the characteristic polynomial factors into linear factors (sum of all algebraic multiplicity is n) and for every eigenvalue, the geometric multiplicity is equal to the algebraic multiplicty, then the matrix is diagonizable. In other words, a matrix is not diagonizable if it satisfies either:

1. Some the eigenvalues may be complex and not all real
2. One of the eigenvalues has geometry multiplicity less than the algebraic multiplicity.

An important fact to notice is that if $Av = \lambda v$, then $A^2 v = \lambda^2 v$

Remark 3.10. If A is a transition matrix, say for some transition which occurs each month, we can look at A^2 and it is the transition matrix for every other month.

More generally, notice that if $Av = \lambda v$, then $A^k v = \lambda^k v$ for all $k \in \mathbb{Z}$

Recall that A is diagonizable if there exists an invertible matrix C and a diagonal matrix D such that $A = CDC^{-1}$. Remember that C is the change of basis matrix from the standard basis to the basis consisting of the eigenvectors of A.

If we have $A = CDC^{-1}$, D describes the effect of A on state vectors expressed in B-coordinates.

Example 3.11

Find the standard matrix of the reflection across the line spanned by span $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Consider $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$. These 2 vectors form an eigenbasis for the eigenvectors with respect to this reflection. So, $A = CDC^{-1} = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}^{-1}$

Definition 3.12 (Similarity)**.** Two square matrices A, B are similar, if there exists an invertible matrix C such that $A = CBC^{-1}$

So, if A is diagonizable, then A is similar to a diagonal matrix.

If A and B are similar, then they describe the same linear transformation in different coordinate systems.

Proposition 3.13

Two diagonizable matrices are similar to each other if and only if they have the same eigenvalues(with algebraic multiplicty).

§4 April 1st

Recall any matrix of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ describes a rotation and a dilation. For example, the matrix $\begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix}$ describes a rotation with angle θ and dilation by factor r. The characteristic polynomial of the matrix above is $\lambda^2 - 2a\lambda + a^2 + b^2$. Using the quadratic equation, we find that the 2 eigenvalues are $\lambda = a \pm |b|i$

> **Remark 4.1.** Above, we found the characteristic polynomial by using the fact that for a 2×2 matrix A is $\lambda^2 - \text{Tr}(A)\lambda + \det(A)$. Recall that $\text{Tr}(A)$ is the sum of the diagonal entries.

Fact 4.2. Two complex numbers z and z' are equal if and only if $\mathfrak{Re}(z)$ and $\mathfrak{C}(z')$ of both complex numbers are equal

Consider an arbitrary 2×2 matrix. Then we know that the characteristic polynomial $\lambda^2 - \text{Tr}(A) + \det(A)$. Using the quadratic formula, $\lambda = \frac{1}{2}\,\text{Tr}\,A \pm \sqrt{\text{Tr}(A)^2 - 4\det(A)}$. So, if $4\det(A) > \text{Tr}(A)^2$, the eigenvalues of A are complex.

Consider a complex eigenvalue $\lambda + i\mu$ and its corresponding eigenvector $v + iw$. Then by definition,

$$\begin{aligned} A(v+iw) &= (\lambda + i\mu)(v+iw) \\ \implies Av + iAw &= \lambda v + (\lambda w + \mu v)i - \mu w \\ \implies Av + iAw &= (\lambda v - \mu w) + (\lambda w + \mu v)i \end{aligned}$$

By fact 4.2, we get 2 real vector equations:

$$Av = \lambda v - \mu w = \begin{bmatrix} v & w \end{bmatrix} \begin{bmatrix} \lambda \\ -\mu \end{bmatrix}$$

and

$$Aw = \lambda w + \mu v = \begin{bmatrix} v & w \end{bmatrix} \begin{bmatrix} \mu \\ \lambda \end{bmatrix}$$

We can now write $A \begin{bmatrix} v & w \end{bmatrix} = \begin{bmatrix} v & w \end{bmatrix} \begin{bmatrix} \lambda & \mu \\ -\mu & \lambda \end{bmatrix}$

Letting $C = \begin{bmatrix} v & w \end{bmatrix}$, we can rewrite the above as: $AC = C \begin{bmatrix} \lambda & \mu \\ -\mu & \lambda \end{bmatrix}$

Mupltiplying on the left by C^{-1}, we have: $A = C \begin{bmatrix} \lambda & \mu \\ -\mu & \lambda \end{bmatrix} C^{-1}$

> **Theorem 4.3** (Similarity of matrices with complex eigenvalues)
> Every 2×2 matrix with complex eigenvalues is similar to a rotation-dilation matrix.

Example 4.4

Consider the matrix $\begin{bmatrix} 2 & -1 \\ 2 & 0 \end{bmatrix}$.

The characteristic polynomial is $\lambda^2 - 2\lambda + 2$. By the quadratic formula, the eigenvalues are $\vartheta = 1 \pm \frac{1}{2}\sqrt{4-8} = 1 \pm i$. For this example, we only consider the complex eigenvalue $\vartheta = 1 + i$.

So, $A - \vartheta \mathrm{id}_2 = \begin{bmatrix} 2-(1+i) & -2 \\ 2 & -(1+i) \end{bmatrix} = \begin{bmatrix} 1-i & -2 \\ 2 & -1-i \end{bmatrix}$ A vector in the null space is $\begin{bmatrix} 1 \\ 1-i \end{bmatrix}$ since $\begin{bmatrix} 1-i & -2 \\ 2 & -1-i \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1-i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Our complex eigenvector is $\begin{bmatrix} 1 \\ 1-i \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = i \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

So, $C = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$

Thus, $A = C \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} C^{-1}$

To finish, we write everything out: $\begin{bmatrix} 2 & -1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}^{-1}$

Remark 4.5. Suppose A is a transition matrix of a dynamical system. Then $A^n = CB^nC^{-1}$

Continuing the example above, we write $\begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix}$ for some r and θ. Then by our remark 4.5,

$\begin{bmatrix} 2 & -1 \\ 2 & 0 \end{bmatrix}^n = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix}^n \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}^{-1}$

Using the fact that the rotation dilation matrix does: $e_1 \mapsto \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $e_2 \mapsto \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
Clearly, by highschool trigonometry, we see that $\theta = -\frac{\pi}{4}$ and $r = \sqrt{2}$, by the pythagorean theorem.
We can now conclude that

$$A^n = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}^n \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \sqrt{2}^n \begin{bmatrix} \cos\frac{n\pi}{4} + \sin\frac{n\pi}{4} & -\sin\frac{n\pi}{4} \\ 2\sin\frac{n\pi}{4} & \cos\frac{n\pi}{4} - \sin\frac{n\pi}{4} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}^{-1}$$

Proposition 4.6 (Geometry of complex eigenvalues)

For a complex eigenvalue $\lambda + i\mu$, then $\sqrt{\det A} = \sqrt{\lambda^2 + \mu^2} = \begin{cases} >1, & \text{spirals out} \\ <1, & \text{spirals in} \\ =1, & \text{elliptical rotation} \end{cases}$

Remark 4.7. We found that $\sqrt{\det A} = \sqrt{\lambda^2 + \mu^2}$ by writing $A = C \begin{bmatrix} \lambda & \mu \\ -\mu & \lambda \end{bmatrix} C^{-1}$. Then, we remember the *multiplicative property* of determinants so $\det A = \det C \cdot \det \begin{bmatrix} \lambda & \mu \\ -\mu & \lambda \end{bmatrix} \cdot \det C^{-1}$. The result follows since $\det C^{-1} = (\det C)^{-1}$, by the determinant power rule.

Remark 4.8. We take the square root of $\det A$ because for a rotation dilation matrix $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, the rotation angle, $r = a^2 + b^2$ (by the Pythagorean theorem)

§5 April 6th

Definition 5.1. Dot product Let $x, y \in \mathbb{R}^n$ be vectors. The we define the **dot product** to be $x \cdot y = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix} = x_1y_1 + x_2y_2 + \ldots + x_ny_n$

Proposition 5.2 (Properties of the dot product)

- $x \cdot y = y \cdot x$
- $x \cdot (y + y') = x \cdot y + x \cdot y'$
- $x \cdot (cy) = cx \cdot y$

Definition 5.3. We call $\|x\| = \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$ the **length** or **norm** of the vector $x \in \mathbb{R}^n$

Definition 5.4. Given 2 vectors $x, y \in \mathbb{R}^n$, we define the **distance** between x, y to be $\|y - x\|$

Proposition 5.5 (Properties of the norm)

- $\|cx\| = |c| \cdot \|x\|$
- $\|x\| = 0$ if and only if $x = 0$

Definition 5.6. A vector v is a **unit vector** if $\|v\| = 1$

Notice that the unit vector in the direction of v is $\frac{1}{\|v\|} \cdot v$

Definition 5.7. Two vectors x, y are **orthogonal/perpendicular** if the dot product is zero.

Example 5.8

Using a bit of geometry, we see that $\begin{pmatrix} a \\ b \end{pmatrix}$ is orthogonal to $\begin{pmatrix} -b \\ a \end{pmatrix}$ because the dot product is $-ab + ab = 0$, as desired.

Theorem 5.9

Law of Cosines, High-school version For a triangle with side lengths a, b, c, we have the following relation:

$$c^2 = a^2 + b^2 - 2ab \cos \gamma$$

where γ is the angle between sides with lengths a and b

Theorem 5.10 (Law of Cosines, University Version)
Notice that:

$$c = \text{ dist } (a, b) = \|b - a\| = \sqrt{(b-a)\cdot(b-a)}$$

So, $(b-a)\cdot(b-a) = a^2 + b^2 - 2\,\|a\|\,\|b\|\cos\gamma$
Multiplying stuff out,
$a \cdot b = \|a\|\,\|b\|\cos\gamma \implies \frac{a}{\|a\|}\cdot\frac{b}{\|b\|} = \cos\gamma$

Lemma 5.11 (Theorem 5.10 in English)
The dot product of the two unit vectors is the cosine of the angle between them

To finish, we see that $\gamma = cos^{-1}\left(\frac{a}{\|a\|}\cdot\frac{b}{\|b\|}\right)$ for the 2 vectors a and b.
So, we also see that 2 vectors are orthogonal to each other only when $\cos\gamma = \frac{\pi}{2}$.

Definition 5.12. Let $W \in \mathbb{R}^n$ be a subspace.
Then let us define $W^\perp = \{v \in \mathbb{R}^n | v \perp w \text{ for all } w \text{ in } W\ \}$. We call $W^\perp$ the **orthogonal complement** of W.

Remark 5.13. In general, $W^\perp$ is always a subsapce of $\mathbb{R}^n$ and $\dim(W^\perp) + \dim(W) = n$. The proof is nontrivial but easy to follow

Claim 5.14 — $(W^\perp)^\perp = W$.

Proof. This is trivial. □

Remark 5.15. All vectors are perpendicular to the zero vector.

Suppose W is spanned some vectors which are the the columns of the matrix A. Then if $x \in W^\perp$, we must have : $x\perp a_1, x\perp a_2, \ldots, x\perp a_n$, where $a_1, a_2, \ldots, a_n$ are the columns of the matrix A. We can write the above relation with matrix products.

$$\begin{pmatrix} a_1^\top \\ a_2^\top \\ . \\ . \\ a_n^\top \end{pmatrix} \cdot x = 0$$

So, $x \in nul(A^\top)$
The important thing to remember is that $col(A)$

Definition 5.16. We define the **row space** of a matrix A as all linear combinations of the rows of A, or simply the columns of $A^\perp$

Theorem 5.17 (Rank Theorem)

Suppose A is matrix with n columns. Then we have the following relation:

$$\dim(Col(A)) + \dim(Nul(A)) = n$$

Since we also have $\dim(W^{\perp}) + \dim(W) = n$, we can conclude that $\dim(Col(A)) = \dim(Nul(A)^{\perp}) = \dim(Row(A))$

This leads to another theorem:

Theorem 5.18

For a matrix A, the column rank of A is equal to the row rank of A.

> **Remark 5.19.** According to the esteemed Prof. Behrend, the row rank-column rank theorem is one of the hardest theorems to prove and marks the culmination of this course on matrix algebra.

§6 April 8th,2021

Let $W \subset \mathbb{R}^n$ be a subspace and $b \in \mathbb{R}^n$ be a vector.
Then there exists a unique vector $b_W \in W$ such that $(b - b_W) \perp W$. This b_W is the **orthogonal projection** of b onto W and minimizes $\|b - w\|$ for all $w \in W$

Now, suppose $W = span(a_1, \dots a_k)$, so $A = (a_1, \dots a_k)$ are the columns of A. Then we can write: $W = Col(A)$. So, for all $b \in \mathbb{R}^n$, there is a unique vector of the form Ac such that $(b - Ac) \perp Col(A)$. This also mean that $(b - Ac) \in Col(A)^\perp = Nul(A^T)$
In other words, we have:

$$A^T(b - Ac) = 0 \implies A^T Ac = A^T b$$

We call the above the **normal** equation of the vector equation $Ax = b$ and it is always consistent.

To emphasize, we have the following:

- Given a matrix A with n rows and k columns, and a vector $b \in \mathbb{R}^n$, there exists $c \in R^k$, such that $A^T Ac = A^T b$
- If we have $A^T Ac = A^T b$, then $Ac = b_{ColA}$ is the projection of b onto $Col(A)$. So, Ac is unique but c is not necessarily.

Theorem 6.1 (Least Squares Solution)
Given a linear system of equations $Ax = b$, then the associated normal equation $A^T Ax = A^T b$ is always consistent and any solution $\hat{x}$ satisfies $A\hat{x} = b_{ColA}$ is called a **least squares solution** of $Ax = b$. It minimizes $\|b - Ax\|^2$: the sum of the squares of the components of Ax.

§7 April 13th

Recall the normal equation:

$$A^T Ax = A^T b$$

Proposition 7.1 (Unique Least Square Solution)

If the columns of A are linearly independent, then $A^T A$ is invertible and there is a unique least squares solution $x = (A^T A)^{-1} A^T b$.

So, $b_{ColA} = Ax = A(A^T A)^{-1} A^T b$ is the projection onto $Col(A)$.

Example 7.2

Find the distance of the point $b = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ from the plane $W = \text{ span } \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right)$

We use the orthogonal decomposition and write $b = b_w + b_{w^\perp}$. First, we find the orthogonal projection.

Using the normal equation, we find that $b_{w^\perp} = \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$. The norm of this vector is

Example 7.3

Consider the projection onto the line spanned by $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$

Let $A = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. By the normal equation,

$$\begin{pmatrix} x \\ y \end{pmatrix}_L = A(A^T A)^{-1} A^T \begin{pmatrix} x \\ y \end{pmatrix}$$

we find that

$$\begin{pmatrix} x \\ y \end{pmatrix}_L = \frac{\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix}}{\begin{pmatrix} 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix}}$$

More generally, the projection of a vector x onto the line spanned by v is

$$\frac{x \cdot v}{\|v\|^2} v$$

Example 7.4

Reflection Consider the reflection across line L spanned by $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

Let R be the projection onto L and R be the reflection across L. By creating a picture, we see that:

$$\mathrm{id}_2 x + Rx = 2Px \implies \mathrm{id}_2 + R = 2P$$

.

What rests is a simple calculation to solve for R.

Definition 7.5 (Orthogonal Sets)**.** Nonzero vectors $v_1, v_2, \ldots, v_k$ form an **orthogonal set** if $v_i \perp v_j$ for all $i \neq j$

Definition 7.6 (Orthonormal Set)**.** An **orthonormal set** is an orthogonal set which has the property that $\|v_i\| = 1$ for all i.

Remark 7.7. The standard basis $e_1, \ldots e_n$ of $\mathbb{R}^n$ is an orthonormal set

Claim 7.8 — If $\{v_i\}$ if an orthogonal set, then $\{\frac{v_i}{\|v_i\|}\}$ is an orthonormal set.

Proof. Trivial. □

Proposition 7.9 (Linearly independence of Orthogonal sets)

Every orthogonal set is linearly independent.

Proposition 7.10 (Formulae for orthogonal projection given orthogonal basis)

For an orthonormal basis $u_1, \ldots, u_k$ of W, we have :

$$x_w = (x \cdot u_1)u_1 + \ldots + (x \cdot u_k)u_k$$

.

For an orthogonal basis $w_1, \ldots, w_k$, simply notice that $u_i = \frac{w_i}{\|w_i\|}$.

Theorem 7.11 (Gram-Schmidt Process)

The algorithm converts a basis of $W \in \mathbb{R}^n$ to an orthogonal basis of the subspace W.

Remark 7.12. The algorithm will not be covered in the matrix algebra 221 course.

Theorem 7.13 (Spectral Theorem)

If A is a symmetric matrix ($A^T = A$), then A is diagonalizable.
Morever, if $\lambda \neq \mu$ are eigenvalues, then $E_\lambda \perp E_\mu$. In particular, if all eigencalues have multiplicity 1, then an eigenbasis is an orthogonal basis.

§7.1 Quantum Mechanics Digression

Suppose the state space if $\mathbb{R}^n$. There are things called **observables**, which are symmetric matrices. Then you can only observe eigenvalues of the matrix.
After you make a measurement, the system will be in an eigenvector.
The big idea is that everything around us is based on symmetric matrices and math is prevalent.

§8 Appendix

Definition 8.1 (Stochastic Matrices)**.** Let A be a $n \times n$ square matrix. Then we call A **stochastic** if the sum of entries in each column is 1.

Theorem 8.2

If A is a stochastic matrix, then A has non-trivial fixed vectors v such that $Av = v$, that is, a eigenvector with corresponding eigenvalue of 1.

Proposition 8.3

The space of fixed vectors is: $Nul(A - \text{id}_n)$

Theorem 8.4 (Perron-Frobenius Theorem)

Let A be a positive stochastic matrix(with entries all greater than 0). Then the following are equivalent:

1. $\dim(Nul(A - \text{id}_n) = 1)$. That is, up to scalar multiplication, there is exactly one fixed vector
2. There exists a unique fixed vector w such that all components of w are greater than 0 and add up to 1. We call w the **steady state vector**
3. If p is any state vector,
$$\lim_{n\to\infty} A^n p = cw$$
where c is equal to the sum of the components of p.