Homework 2 due Tuesday.
Extra office hours M 11:30-12:30 EEB 528.
Midterm Thursday Oct 10
Reading: 4.3-4.4

Linear transformations \( \mathbb{R}^n \to \mathbb{R}^m \)
Goal: Understand the geometry of linear maps.

Recall:
Definition: For an \( m \times n \) matrix \( A \),
- The nullspace (or kernel) of \( A \) is \( N(A) = \{ x \in \mathbb{R}^n \mid Ax = \mathbf{0} \} \).
- The range (or column space) of \( A \) is
  \[ R(A) = \{ Ax \mid x \in \mathbb{R}^n \} = \text{Span (columns of } A \text{)} \]
- The row space of \( A \) is
  \( R(A^T) = \text{Span (rows of } A \text{)} \).

Problem: Let \( u = (\frac{1}{2}) \), \( v = (\frac{3}{4}) \), and let
\( A = u v^T = (\frac{1}{2})(3 \ 4) = (\frac{3}{8} \ \frac{4}{8}) \).
What are \( R(A), R(A^T), N(A), N(A^T) \)?

Answer: Yes, we could apply Gaussian elimination.
But we can also just observe that for any vector \( x \),
\( Ax = u v^T x = (v^T x) u \)
\( \Rightarrow R(A) = \text{Span (}u\text{)}, a \text{ line.} \)
Similarly, \( R(A^T) = \text{Span (}v\text{)}, a \text{ line.} \)
Also \( Ax = \mathbf{0} \iff v^T x = \mathbf{0} \)
\[ 3x_1 + 4x_2 \]
\( \iff x = (-\frac{4}{3}, \frac{2}{3}) \)
\( \Rightarrow N(A) = \text{Span (}(-\frac{4}{3}, \frac{2}{3})\text{)}, \text{ the line perpendicular to } u \)
Similarly \( N(A^T) = \text{Span (}(-\frac{2}{3}, 1)\text{)} \).

Definition: A matrix \( A \) is diagonally dominant
if for all rows \( j \),
\[ |a_{jj}| \geq \sum_{i \neq j} |a_{ij}| \]
\[ |a_{ij}| \geq \sum_{i \neq j} |a_{ij}| \]

and is strictly diagonally dominant if the inequality is strict (>) for all rows.

**Theorem:** If \( A \) is strictly diagonally dominant, then

\[ N(A) = \{ \mathbf{0} \} \]

**Proof:** Let \( x \in N(A) \), \( Ax = 0 \).
Let \( j \) be a coordinate so \( |x_j| = \max_i |x_i| \) has maximum magnitude.

\[(Ax)_j = 0 = a_{jj}x_j + \sum_{i \neq j} a_{ij}x_i \]

\[ |a_{jj}x_j| = |\sum_{i \neq j} a_{ij}x_i| \]

\[ \leq \sum_{i \neq j} |a_{ij}||x_i| \]

\[ \leq \left( \sum_{i \neq j} |a_{ij}| \right) |x_j| \]

\[ \Rightarrow x_j = 0 \Rightarrow x = 0 \Rightarrow N(A) = \{ \mathbf{0} \} \quad \Box \]

What about diagonally dominant matrices?

**Example:**

\[
A = \begin{pmatrix}
-2 & 1 & 0 \\
1 & -2 & 1 \\
0 & 1 & -2
\end{pmatrix}
\]

This is proportional to the matrix you get by discretizing the second derivative operator on the circle:

\[ f''(t) \approx \frac{1}{h^2} [f(t+h) - 2f(t) + f(t-h)] \]

with the 1st and last equations wrapping around.

(Note that multiplying by \( \frac{1}{h^2} \) doesn't change the range or nullspace.)

**Observe:** \( A \) is diagonally dominant (but not strictly so).

**What is \( N(A) \)?**

**Claim:** \( N(A) = \text{Span} \{ \mathbf{c}, 1, 1, \ldots, 1 \} \}

- set of all constant vectors

**Intuition:** \( A \) represents double differentiation, so only constant vectors should be sent to \( 0 \).
Proof:
\[
A \left( \vec{i} \right) = \left( \begin{array}{cc}
\text{sum of } a_{ij} \text{ across row } i \\
\text{sum across row } n
\end{array} \right) = 0
\]

\[\Rightarrow (1, 1, ..., 1) \in N(A)\]
\[\Rightarrow \text{Span}(\mathbb{E}(1, 1, ..., 1)) \subseteq N(A).\]

But are there other vectors in \(N(A)\)?
Let \(v \in N(A)\) \Rightarrow for all \(j\),
\[
v_{j-1} - 2v_j + v_{j+1} = 0
\]
where the indices wrap around
or \(v_j = \frac{1}{2}(v_{j-1} + v_{j+1})\) avg. of its neighbors

\[\Rightarrow v_j\text{ can't have any local maxima}\]

\[\Rightarrow v_j \text{ is constant, } v \in \text{Span}((1, 1, ..., 1))\]
\[\Rightarrow N(A) \subseteq \text{Span}((1, 1, ..., 1))\]
\[\Rightarrow N(A^T) = \text{Span}((1, 1, ..., 1)).\checkmark\]

**Question:** What is \(R(A)\)? (Note \(A = A^T\))

**Question:** Can you characterize the nullspaces of diagonally dominant matrices?

**Linear Transformations**

Why matrices??

Why matrix multiplication?

(Why do only square matrices have inverses?)

**Definition:** A function \(f: \mathcal{U} \rightarrow \mathcal{V}\) between two vector spaces \(\mathcal{U}\) and \(\mathcal{V}\) (with the same underlying field) is linear if
\[
f(\alpha \mathbf{u}) = \alpha f(\mathbf{u}) \quad \text{for all vectors } \mathbf{u} \in \mathcal{U}
\]
and \(f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})\) for all \(\mathbf{u}, \mathbf{v} \in \mathcal{U}\).

**Examples:** For \(\mathcal{U} = \mathcal{V} = \mathbb{R}^2\),
- \(f(x, y) = (0, 0)\) \(\checkmark\)
- \(f(\cos \theta, \sin \theta) = (\cos \theta \mathbf{u} - \sin \theta \mathbf{v})\) \(\checkmark\)
Examples: for \( u, v \in \mathbb{R}^2 \),
- \( f(x, y) = (0, 0) \)
- \( f \) a rotation by \( \theta \) : \( f(x, y) = (\cos \theta x - \sin \theta y, \sin \theta x + \cos \theta y) \)
- \( f(x, y) = (x^2, \sin y) \) not linear
- \( f(x, y) = (1 + x, y) \) not linear

One more example: Let \( \mathcal{U} = \mathcal{V} \) = space of polynomials in \( x \) of arbitrary degree, e.g. \( a + bx + cx^2 \).
(\text{This is a vector space; it is closed under addition and under multiplication by real numbers.})

For a polynomial \( p \), let
\[
f(p) = (2 + 3x) \cdot p
\]
This is a linear transformation!

We'll see lots more examples later (e.g., differentiation, ...)

\[\text{LINEAR TRANSFORMATIONS} \quad \uparrow \quad \text{MATRICES}\]

We'll see the correspondence today for \( \mathcal{U} = \mathbb{R}^n, \mathcal{V} = \mathbb{R}^m \), and generalize it to arbitrary vector spaces next week.

**Theorem 1**: Let \( A \in \mathbb{R}^{m \times n} \) be an \( m \times n \) real matrix.
Define \( f : \mathbb{R}^n \to \mathbb{R}^m \) by
\[
f(u) = Au
\]
(left-multiplication by \( A \)).

Then \( f \) is a linear transformation.

**Proof**: \( f(au) = A \cdot (au) = a \cdot (Au) = af(u) \)
\[
f(u + v) = A \cdot (u + v) = Au + Av = f(u) + f(v)
\]
\( \square \)

**Theorem 2**: Let \( f : \mathbb{R}^n \to \mathbb{R}^m \) be any linear transformation.
Then there exists an \( m \times n \) matrix \( A \) such that
\[
f(u) = Au
\]

**Proof**: Let \( e_j = (0, 0, \ldots, 1, 0, \ldots, 0) \in \mathbb{R}^n \).

Let
\[
A = \begin{pmatrix}
f(e_1) & f(e_2) & \cdots & f(e_n)
\end{pmatrix}
\]
its \( j \)-th column is \( f(e_j) \).

We claim that \( f(u) = Au \) for any \( u \in \mathbb{R}^n \). Indeed,
\[
u = (u_1, u_2, u_3, \ldots, u_n)
\]
\[ = u_1 e_1 + u_2 e_2 + \cdots + u_n e_n \]
\[ \Rightarrow f(u) = u_1 f(e_1) + u_2 f(e_2) + \cdots + u_n f(e_n) \]
by applying the linearity property repeatedly.

The right-hand side is
\[ A \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = Au. \]

These theorems are why matrix-vector multiplication is defined the way it is.

What about matrix-matrix multiplication?

**MATRIX MULTIPLICATION ⇐ LINEAR FUNCTION COMPOSITION**

\[ \begin{array}{ccc}
\mathbb{R}^m & \overset{f}{\longrightarrow} & \mathbb{R}^n \\
\overset{g}{\longrightarrow} & \mathbb{R}^r
\end{array} \]

\[ f(u) = Au \quad g(y) = By \]
\[ g(f(u)) = B \cdot A \cdot u \]

How to visualize matrices

**I. \(2 \times 2\) matrices as linear transformations**

\[ A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad A^T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \]

Claim: Any \(2 \times 2\) matrix can be expressed as a product of shearing and scaling matrices, e.g.,
\[ \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix} \]
\( (a, b, c, d) \neq (0, 0, 0, 0) \)

Proof: Starting with any \(2 \times 2\) matrix \((a, b, c, d)\), G.E. row operations change it to a matrix of the form either
\[ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & * \\ 0 & 0 \end{pmatrix} \text{ or } \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \]

\[ \begin{pmatrix} & & \\ & & & \end{pmatrix} \begin{pmatrix} & \end{pmatrix} \begin{pmatrix} & \\ & \end{pmatrix} \begin{pmatrix} & \end{pmatrix} \begin{pmatrix} & \\ & \end{pmatrix} \]

\[ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \]
These final matrices are $S(0,0), S(1,1), S(1,0)$. Row operations (scaling a row, adding one row to another) can be implemented by multiplying on the left by $A, A^T$ or $S(a,p)$. Column operations by right-multiplying. Therefore we can get from the end $S(0,0), S(1,1)$ or $S(1,0)$ to the beginning $(a, b)$ by left- or right-multiplying by the generator matrices.

**Exercise:** Generalize this claim to $n \times n$ matrices. Using row and column operations, acting on only two coordinates at a time, any $n \times n$ matrix can be reduced to the form

\[
\begin{pmatrix}
1 & & & \\
 & \ddots & & \\
0 & & \ddots & \\
0 & & & 0
\end{pmatrix}.
\]

**II. The geometry of linear transformations**

Let $A$ be an $m \times n$ real matrix.

Let's refine this picture ...
Observations:

1. If $x, y \in \text{N}(A)$, then $Ax = Ay$. Graphically,

   All points on this line are mapped to the same point.

In our heads, we can thus break $A$ into two steps:

1. First map $y$ to $x$, i.e., take a point and move it parallel to $\text{N}(A)$ to get to $\text{N}(A)^\perp$. This is a projection; it flattens the space to $\text{N}(A)^\perp$.

2. Map $x$ to $Ax$.

2. If $x, z$ are distinct points in $\text{N}(A)^\perp$, then $Ax \neq Az$.

   (Because if $Ax = Az$, then $x - z \in \text{N}(A)$.)

Thus $A$ is a 1-to-1 map $\text{N}(A)^\perp \rightarrow \text{R}(A)$, so restricted to $\text{N}(A)^\perp$ it is actually invertible...

\[ \text{This picture is not yet complete. For example, Homework 2,} \]
Matrix and function inverses

Definition: A function $f : D \rightarrow C$ is invertible if every point in $C$ is the image of exactly one point in $D$.

The inverse function $f^{-1} : C \rightarrow D$ takes each point in $C$ to its unique preimage.

Equivalently, $f^{-1}$ satisfies $f^{-1} \circ f = \text{identity on } D$
$f \circ f^{-1} = \text{identity on } C$

Exercise: Prove that the inverse of a linear function, if it exists, is also linear.

Definition: The inverse of a matrix $A$

is a matrix $B$ that satisfies $BA = \text{the identity matrix}$ and $AB = \text{the identity matrix}$.

Observe: Not every matrix is invertible, e.g.,
$A = (0)$ is not invertible.

A matrix that is not invertible is called singular.

If an inverse exists, then it is unique.

Proof: Assume $B$ and $C$ are both inverses of $A$.
Consider $BAC$.

$BAC$
\[
C = (BA)^T C \quad \Rightarrow \quad B(AC) = B \quad \Rightarrow \quad B = C
\]

**Example:** The 2\times2 matrix \( A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \) is invertible/nonsingular if and only if \( ad-bc \neq 0 \). The inverse is given by
\[
A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.
\]
(The proof is an exercise.)

**Question:** How to compute the inverse of a matrix?

Gaussian elimination, of course...

**When is a matrix invertible?**

**Lemma 1:** If \( N(A) \neq \emptyset \), then \( A \) is not invertible.

*Proof:* Take \( y \in N(A), y \neq 0 \).

Then the point \( 0 \) has two preimages, so \( A \) is not invertible.

**Lemma 2:** If \( A \) is an \( m \times n \) matrix with \( m \leq n \), then \( N(A) \neq \emptyset \).

**Corollary:** Only square matrices can be invertible.

*Proof:* Let \( A \) be an \( m \times n \) matrix.

\[
\begin{array}{ll}
\text{if} & m < n : \quad N(A) \neq \emptyset \quad \text{(Lemma 2)} \\
& \Rightarrow A \text{ not invertible (Lemma 1)} \checkmark \\
\text{if} & m \geq n : \quad \text{Assume } A^T \text{ exists, an } n \times m \text{ matrix.} \\
& \Rightarrow N(A^T) \neq \emptyset \quad \text{(Lemma 2)} \\
& \Rightarrow A^T \text{ not invertible (Lemma 1)} \\
& \Rightarrow \text{contradiction, since } (A^T)^T = A. \quad \checkmark
\end{array}
\]

Proof of Lemma 2 (\( m < n \Rightarrow N(A) \neq \emptyset \)).

Applying Gaussian elimination to \( A \) in order to solve for the nullspace results in a matrix like

\[
\begin{pmatrix}
\vdots & \vdots & \vdots & \vdots \\
\alpha & \beta & \gamma & \delta \\
\end{pmatrix}
\]

The point is that there are at most \( m \) pivots. Since \( m < n \), there are necessarily at least \( n-m \geq 1 \) free variables, so the nullspace is infinite. \( \checkmark \)
**Theorem:** An $m \times n$ matrix $A$ is invertible if and only if $m = n$ and $N(A) = \{0\}$.

**Proof:**

$(\Rightarrow)$ If $A$ is invertible, then we have just shown that $m = n$ and $N(A) = \{0\}$.

$(\Leftarrow)$ Follows from:

**Lemma 3:** For an $n \times n$ square matrix $A$,

$$N(A) = \{0\} \iff R(A) = \mathbb{R}^n.$$  

**Proof:** As usual, use Gaussian elimination.

$$A = G \ U$$

$A$ is upper triangular.

If $U$ has $n$ pivots:

$U$ = identity, so $A = G \ U$ is invertible with $N(A) = \{0\}$, $R(A) = \mathbb{R}^n$.

If $U$ has $r < n$ pivots:

$N(U) \neq \{0\}$.

$N(A) = N(U) \neq \{0\}$

If $R(A) = \mathbb{R}^n$, then

$R(U) = G' \ (R(A)) = \mathbb{R}^n$. But that's impossible.

Note: $\square$

**Example:**

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$N(A) = \text{Span}(e_1)$ but $R(A) = \text{everything}$.  

EE 441 Page 10