Lecture 1, Sep 13, 2021

Introduction to Vectors

- Vectors: direction, magnitude, units; scalars: magnitude, sign
- Core of linear algebra involves 2 operations: Adding vectors and scaling vectors
- Course notation: vector from p (tail) to q (head) is expressed as $\vec{v} = \overrightarrow{PQ} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ (arrowhead is complete)
  - Row vectors are not equal to column vectors: $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \neq \begin{bmatrix} v_1 & v_2 \end{bmatrix}$
- Standard position for vectors is tail at the origin

Vector Operations

- Adding vectors: $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \end{bmatrix}$
  - Geometrically this puts the tail of one vector at the head of another
- Scaling vectors: $c \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} cv_1 \\ cv_2 \end{bmatrix}$
  - This makes the vector longer or shorter/flips it
- When one vector is a scalar multiple of another, they're parallel $\vec{v} \parallel \vec{w}$
- The zero vector is $\vec{v} - \vec{v} = \vec{0}$
- A linear combination of two vectors $\vec{v}$ and $\vec{w}$ is $c\vec{v} + d\vec{w}$ for any c and d

Vector Properties

1. $\vec{v} + \vec{w} = \vec{w} + \vec{v}$ (commutative)

Lecture 2, Sep 15, 2021

Vector Properties

- Suppose you have 3 vectors $\vec{v}, \vec{w}, \vec{z}$ in $\mathbb{R}^3$, for scalars c, d, e
  - The picture of all linear combinations $c\vec{v}$ is a line, $c\vec{v} + d\vec{w}$ is a plane, $c\vec{v} + d\vec{w} + e\vec{z}$ is all of $\mathbb{R}^3$ assuming independence
- Vectors are linearly independent when they cannot be expressed as linear combinations of others
- Vector magnitude $\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots}$
  - Properties of magnitudes (and all norms in general):
    1. $\|c\vec{v}\| = |c|\|\vec{v}\|$
    2. $\|\vec{v}\| \iff \vec{v} = \vec{0}$
- A unit vector $\hat{v}$ is any vector with length 1
  - In $\mathbb{R}^3$ the well known unit vectors are $\hat{\imath}, \hat{\jmath}, \hat{k}$ but there are an infinite number of them
  - Vectors can be made into unit vectors (normalized) by dividing by their magnitude: $\hat{v} = \dfrac{\vec{v}}{\|\vec{v}\|}$

Distance Between Points

- To calculate the distance between two points, we first need to establish a coordinate system, with an origin and axis directions (basis vectors?)
- Make vectors $\overrightarrow{OP_1}$ and $\overrightarrow{OP_2}$ where O is the origin, and $\overrightarrow{P_1P_2}$, which we want the magnitude of
- $\overrightarrow{OP_1} + \overrightarrow{P_1P_2} = \overrightarrow{OP_2} \implies \overrightarrow{P_1P_2} = \overrightarrow{OP_2} - \overrightarrow{OP_1}$ then find $\left\|\overrightarrow{P_1P_2}\right\|$

Dot Product

- Definition: The *dot product* $\vec{v} \cdot \vec{w} = \sum_{i=1}^{n} v_i w_i$ (sum of products of each corresponding pair); note the dot product of vectors is a scalar, so it is sometimes referred to as the *scalar product*

Lecture 3, Sep 20, 2021

Properties of the Dot Product

- Dot products are **distributive**, **commutative**, and **associative** (with scalars: $c(\vec{v} \cdot \vec{w}) = c\vec{v} \cdot \vec{w} = \vec{v} \cdot c\vec{w}$)
- The dot product of a vector with itself is the square of its length: $\vec{v} \cdot \vec{v} = v_1^2 + v_2^2 + v_3^2 = \|\vec{v}\|^2$

Geometric Meaning of the Dot Product

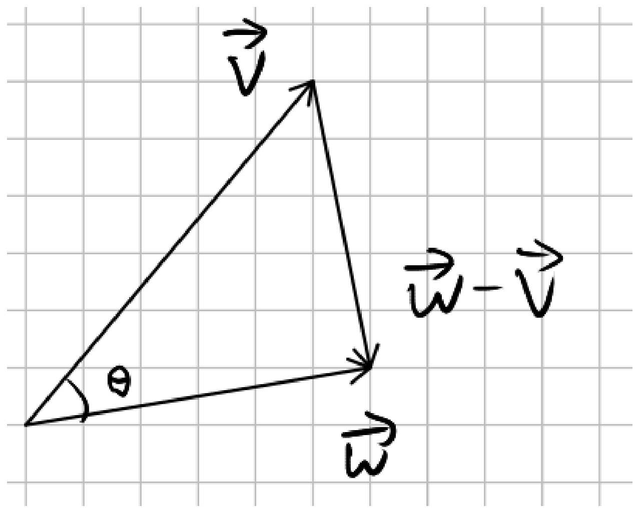


Figure 1: img

-
- $\|\vec{w} - \vec{v}\|^2 = (\vec{w} - \vec{v}) \cdot (\vec{w} - \vec{v}) = \vec{w} \cdot \vec{w} - 2\vec{w} \cdot \vec{v} + \vec{v} \cdot \vec{v} = \|\vec{w}\|^2 + \|\vec{v}\|^2 - 2\vec{w} \cdot \vec{v}$
  - Apply the cosine law: $\|\vec{w} - \vec{v}\|^2 = \|\vec{w}\|^2 + \|\vec{v}\|^2 - 2\|\vec{w}\|\|\vec{v}\| \cos\theta$
  - Therefore: $\vec{w} \cdot \vec{v} = \|\vec{w}\|\|\vec{v}\| \cos\theta \implies \cos\theta = \dfrac{\vec{w} \cdot \vec{v}}{\|\vec{w}\|\|\vec{v}\|}$
- Since cosine has range $[-1, 1]$, this means the dot product can be negative; when the dot product is negative $\theta < \dfrac{\pi}{2}$ so the two vectors are in the same direction
- When $\vec{v} \cdot \vec{w} = 0$, then $\cos\theta = 0 \implies \theta = \dfrac{\pi}{2}$ so $\vec{v}$ and $\vec{w}$ are orthogonal

Two Important Inequalities

1. *Triangle Inequality*: $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$
   -
   - $\|\vec{v} + \vec{w}\|^2 = (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) = \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{w} + 2\vec{v} \cdot \vec{w} = \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\|\vec{v}\|\|\vec{w}\| \cos\theta$
   - Replace $\cos\theta$ with its least upper bound 1: $\|\vec{v} + \vec{w}\|^2 \leq \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\|\vec{v}\|\|\vec{w}\| = (\|\vec{v}\| + \|\vec{w}\|)^2$
   - Take the square root: $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$
2. *Cauchy-Schwartz-Bunyakovsky Inequality*: $|\vec{v} \cdot \vec{w}| \leq \|\vec{v}\|\|\vec{w}\|$

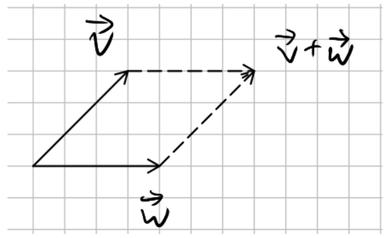
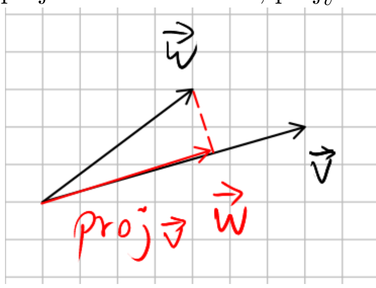


Figure 2: img

Projections

- Key concept: **Orthogonality**: Two vectors are orthogonal iff $\vec{v} \cdot \vec{w} = 0$
- To get the components/coordinates of a vector we project it onto $\hat{\imath}, \hat{\jmath}, \hat{k}$, etc (standard projections)
- The projection of $\vec{w}$ onto $\vec{v}$, $\text{proj}_{\vec{v}}\,\vec{w}$ is a vector parallel to $\vec{v}$ that is as close as $\vec{w}$ as possible:



–

Lecture 4, Sep 22, 2021

Computing Projections

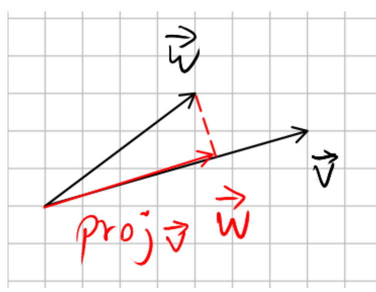


Figure 3: projection

- $\vec{u} = \text{proj}_{\vec{v}}\,\vec{w}$
- $\vec{u}$ has certain properties:
  1. $\vec{u} \parallel \vec{v} \iff \vec{u} = c\vec{v}$
  2. $(\vec{w} - \vec{u}) \perp \vec{v} \iff (\vec{w} - \vec{u}) \cdot \vec{v} = 0$
- Combine the equations to find c: $(\vec{w} - c\vec{v}) \cdot \vec{v} = \vec{w} \cdot \vec{v} - c\|\vec{v}\|^2 = 0 \implies c = \dfrac{\vec{w} \cdot \vec{v}}{\|\vec{v}\|^2}$
- Therefore $\vec{u} = \text{proj}_{\vec{v}}\,\vec{w} = \dfrac{\vec{w} \cdot \vec{v}}{\|\vec{v}\|^2}\vec{v}$ or $\dfrac{\vec{w} \cdot \vec{v}}{\vec{v} \cdot \vec{v}}\vec{v}$
- Alternatively $\dfrac{\vec{w} \cdot \vec{v}}{\|\vec{v}\|^2}\vec{v} = \dfrac{\vec{w} \cdot \vec{v}}{\|\vec{v}\|}\dfrac{\vec{v}}{\|\vec{v}\|} = \dfrac{\vec{w} \cdot \vec{v}}{\|\vec{v}\|}\hat{v}$

Cross Products (aka Vector Products)

- Given $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}, \vec{u} = \vec{v} \times \vec{w} = \begin{bmatrix} v_2w_3 - v_3w_2 \\ -v_1w_3 + v_3w_1 \\ v_1w_2 - v_2w_1 \end{bmatrix}$
- The result of the cross product $\vec{u}$ is orthogonal to both $\vec{v}$ and $\vec{w}$, and has magnitude equal to the area of the parallelogram created by $\vec{v}$ and $\vec{w}$
- Properties of the cross product:
  1. Distributive: $\vec{v} \times (\vec{w} + \vec{z}) = \vec{v} \times \vec{w} + \vec{v} \times \vec{z}$
  2. Anti-commutative: $\vec{v} \times \vec{w} = -(\vec{w} \times \vec{v})$
  3. Associative with scalars: $c\vec{v} \times \vec{w} = \vec{v} \times c\vec{w} = c(\vec{v} \times \vec{w})$
  4. $\vec{v} \times \vec{0} = \vec{0} \times \vec{v} = 0$
  5. **Not** associative with itself: $(\vec{v} \times \vec{w}) \times \vec{z} \neq \vec{v} \times (\vec{w} \times \vec{z})$
- The direction of $\vec{u}$ follows the right-hand rule: index finger in the direction of $\vec{v}$, middle finger in the direction of $\vec{w}$, then the thumb will be pointing in the direction of $\vec{u} = \vec{v} \times \vec{w}$
  - Alternatively start at $\vec{v}$ and curl towards $\vec{w}$
- Lagrange identity says that $\|\vec{v} \times \vec{w}\|^2 = \|\vec{v}\|^2\|\vec{w}\|^2 - (\vec{v} \cdot \vec{w})^2$
  - Use direct proof approach and use the definition of the cross product to brute force the proof
  - Visualization: Suppose θ is the angle between $\vec{w}$ and $\vec{v}$; then $\|\vec{v} \times \vec{w}\|^2 = \|\vec{v}\|^2\|\vec{w}\|^2 - (\|\vec{v}\|\|\vec{w}\| \cos\theta)^2 = \|\vec{v}\|^2\|\vec{w}\|^2(1 - \cos^2\theta) = \|\vec{v}\|^2\|\vec{w}\|^2 \sin^2\theta$
  - Therefore $\|\vec{v} \times \vec{w}\| = \|\vec{v}\|\|\vec{w}\| \sin\theta$ (absolute value not needed because $\sin\theta \geq 0$ for $0 \leq \theta \leq \pi$)
    * Similar to the dot product which is $\vec{v} \cdot \vec{w} = \|\vec{v}\|\|\vec{w}\| \cos\theta$
  - This corresponds to the area of the parallelogram defined by $\vec{v}$ and $\vec{w}$:

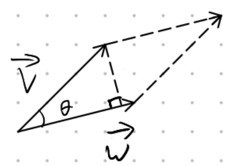


Figure 4: area of paralleogram

    *
    * The base length is $\|\vec{w}\|$, and the height is $\|\vec{v}\| \sin\theta$, so the magnitude of the cross product is the area of this parallelogram
- The dot product is a measure of how much the vectors are parallel; the cross product is a measure of how much the vectors are orthogonal

Lecture 5, Sep 27, 2021

Lines and Planes in 3D

- Lines in $\mathbb{R}^3$:
  - Line in $\mathbb{R}^3$ through the origin: $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = c\vec{d}$ where $\vec{d}$ is a direction vector on the line
  - General line in $\mathbb{R}^3$ not through the origin: $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = p_0 + c\vec{d}$ where p_0 is a known point on the line
  - Both have only one free parameter c, which corresponds with the fact that the line is 1-dimensional
  - To convert $y = mx + b$ to vector form: $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} + c\begin{bmatrix} 1 \\ m \end{bmatrix}$
    * Our known point: $\begin{bmatrix} 0 \\ b \end{bmatrix}$

    * The direction vector: $\begin{bmatrix} 1 \\ m \end{bmatrix}$: as x increases by 1, y increase by m

- Planes in $\mathbb{R}^3$:
  - Plane in $\mathbb{R}^3$ through the origin: $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = c\vec{d_1} + d\vec{d_2}$ where $\vec{d_1}$ and $\vec{d_2}$ are two independent vectors parallel to the plane
  - General plane in $\mathbb{R}^3$ not through the origin: $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = p_0 + c\vec{d_1} + d\vec{d_2}$ where p_0 is a known point on the plane
  - Another way to represent planes is with a normal vector $\vec{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$; while there are many vectors parallel to the plane, there are only two directions that are normal to the plane
    * Since $\vec{n}$ is orthogonal, any vector $\overrightarrow{p_0 p}$ that goes from a known point on the plane p_0 to another point on the plane p_1 is orthogonal to it
    * $0 = \overrightarrow{p_0 p} \cdot \vec{n} = \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} \implies a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \iff ax + by + cz = d$ where $d = ax_0 + by_0 + cz_0$
    * This equation is sometimes referred to as the scalar equation of the plane
  - To go from vector equation to scalar equation, take the cross product to get a normal vector, substitute in a, b, and c and solve for d
  - To go from a scalar equation to a vector equation, find 3 non-collinear points and get two vectors and use any point as the known point

Lecture 6, Sep 29, 2021

Solving Linear Equations

- Solving systems of linear equations is a central problem of linear algebra
- Row picture: Visualize every row as a line or plane in space, and find the intersection of every row
- Column picture: Express the system as a single vector equation, and find the correct linear combination of vectors that satisfy the equation:
  - e.g. $\begin{cases} x - 2y = 1 \\ 3x + 2y = 11 \end{cases} \implies x \begin{bmatrix} 1 \\ 3 \end{bmatrix} + y \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}$
- Matrix form: $A\vec{x} = \vec{b}$
  - The matrix vector product on the left hand side is defined to be the vector given by $A\vec{x} = \begin{bmatrix} | & | & | \\ \vec{a_1} & \vec{a_2} & \cdots \\ | & | & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix} = x_1\vec{a_1} + x_2\vec{a_2} + \cdots$
  - This leads to the dot product rule of calculating a matrix-vector product: Each entry in the resulting vector is the dot product of one row of A and $\vec{x}$

What is a Matrix?

- A rectangular array of numbers: $A = \begin{bmatrix} 4 & 8 & 3 \\ 2 & 1 & 9 \end{bmatrix}$, a 2×3 matrix
- Matrix size is rows $\times$ columns
- More general notation: $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$

Matrix Operations

- Matrix addition: Just add all the corresponding entries
  - Requirement: Two matrices have the same size
  - Commutative, associative
- Scalar multiplication: Multiply every entry by the scalar
- Matrix multiplication:
  - The entry at row i, column j is the dot product of row i of A with column j of B
  - $A \begin{bmatrix} | & | & | \\ \vec{b_1} & \vec{b_2} & \cdots \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ A\vec{b_1} & A\vec{b_2} & \cdots \\ | & | & | \end{bmatrix}$
  - Requirement: The first matrix has the same number of columns as rows of the second matrix: $m \times n$ can multiply $n \times k$ to produce a $m \times k$ matrix
  - **Not commutative**, but associative
  - The notation for multiplying matrices is just AB, without any symbol in between
- Properties:
  1. $A + B = B + A$ addition is commutative
  2. $c(A + B) = cA + cB$ scalar multiplication is distributive
  3. $(A + B) + C = A + (B + C)$ addition is associative
  4. $C(A + B) = CA + CB$ (note C is a matrix) matrix multiplication is distributive
  5. $A(BC) = (AB)C$ matrix multiplication is associative
  6. All exponent laws apply to matrices (e.g. $A^p A^q = A^{p+q}$)

Lecture 7, Oct 4, 2021

Linear Transformations

- Definition: A linear transformation $\mathcal{L}$ is a function that maps a vector in $\mathbb{R}^n$ to a vector in $\mathbb{R}^m$ with the following properties (more formally a linear transformation is $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$, if $\vec{v}, \vec{w} \in \mathbb{R}^n$ then $\mathcal{L}(\vec{v}), \mathcal{L}(\vec{w}) \in \mathbb{R}^m$):
  1. $\mathcal{L}(c\vec{v}) = c\mathcal{L}(\vec{v})$
  2. $\mathcal{L}(\vec{v} + \vec{w}) = \mathcal{L}(\vec{v}) + \mathcal{L}(\vec{w})$
- Examples:
  - A transformation T_1 that adds a constant vector to every vector is **not** a linear transformation (violates both properties)
  - $y(x) = mx + b$ is **not** a linear transformation
  - A transformation T_2 that projects every vector $\vec{w} \in \mathbb{R}^3$ onto a given vector $\vec{v} \in \mathbb{R}^3$ is a linear transformation
    - Proof is too long so it's not copied down
  - A transformation T_3 that takes the length of every vector **not** a linear transformation (fails property 1 because c could be negative, fails property 2 because of the triangle *inequality*)

Lecture 8, Oct 6, 2021

Matrices as Linear Transformations

- Projections can be written as the product of a vector and a projection matrix
- Generally, all linear transformations can be expressed as a matrix; to determine the matrix associated with a linear transformation, find $\mathcal{L}(\hat{i})$, $\mathcal{L}(\hat{j})$, $\mathcal{L}(\hat{k})$ (and so on for all bases) and put those as the columns of the matrix
  - Consider $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \end{bmatrix} = v_1\vec{b}_1 + v_2\vec{b}_2 + \cdots$ where $\vec{b}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\vec{b}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, etc

  - By linearity $\mathcal{L}(\vec{v}) = v_1\mathcal{L}(\vec{b}_1) + v_2\mathcal{L}(\vec{b}_2) + \cdots$
  - Now consider the matrix-vector product $\begin{bmatrix} | & | & | \\ \mathcal{L}(\vec{b}_1) & \mathcal{L}(\vec{b}_2) & \cdots \\ | & | & | \end{bmatrix} \vec{v}$
  - By definition this is equal to $v_1\mathcal{L}(\vec{b}_1) + v_2\mathcal{L}(\vec{b}_2) + \cdots$ which is equal to $\mathcal{L}(\vec{v})$
  - Therefore $\mathcal{L}$ is represented by the matrix $\begin{bmatrix} | & | & | \\ \mathcal{L}(\vec{b}_1) & \mathcal{L}(\vec{b}_2) & \cdots \\ | & | & | \end{bmatrix}$, so every linear transformation has an associated matrix
- Conversely all matrix transformations are also linear
- Some famous transformations:
  - I – The identity transformation, which takes every vector to itself; the identity matrix has 1s down the diagonal and 0s everywhere else
  - O – The zero transformation, which takes every vector to zero; the zero matrix has all zeroes

Why We Multiply Matrices the Way We Do

- Let T_1 and T_2 be two linear transformations
- The transformation of T_1 followed by T_2 is $T_2(T_1(\vec{v}))$; note order is important here
- $T_2(T_1(\vec{v})) = M_2(M_1\vec{v})$ where M_2 is the matrix for transformation T_2 and M_1 is the matrix for transformation T_1
- Therefore the matrix product M_2M_1 represents the result of applying two linear transformations, one after the other
- $A \begin{bmatrix} | & | & | \\ \vec{b_1} & \vec{b_2} & \cdots \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ A\vec{b_1} & A\vec{b_2} & \cdots \\ | & | & | \end{bmatrix}$
- Example: Define a transformation T_θ that takes every vector in $\mathbb{R}^2$ and rotates it counterclockwise by θ; the matrix for this transformation is $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

Lecture 9, Oct 13, 2021

Proving Trig Identities Using Matrices

- Rotation matrix: $T_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$
- Define the composition $T_\theta(T_\theta(\vec{v})) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta - \sin^2\theta & -2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta \end{bmatrix}$
- Since $T_\theta(T_\theta(\vec{v})) = T_{2\theta} = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}$ but $T_\theta(T_\theta(\vec{v})) = \begin{bmatrix} \cos^2\theta - \sin^2\theta & -2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta \end{bmatrix}$, we can conclude $\begin{cases} \sin 2\theta = 2\sin\theta\cos\theta \\ \cos 2\theta = \sin^2\theta - \cos^2\theta \end{cases}$

Eigenvalues & Eigenvectors

- Certain vectors do not change direction (remain parallel) when transformed by a linear transformation; they are the *eigenvectors* of that transformation, and the amount they are scaled by are the *eigenvalues*
- $A\vec{v} = \lambda\vec{v}$, where $\vec{v}$ is an eigenvector and λ is an eigenvalue
- Example: For a transformation that projects every vector onto a fixed vector, every vector parallel to the fixed vector is an eigenvector of the transformation with eigenvalue 1; every vector orthogonal to it is an eigenvector of the transformation with eigenvalue 0
- The number of eigenvalues (note: these can be complex) is always the same as the dimension of the matrix (note: nonsquare matrices don't have eigenvalues)

- To find eigenvectors, note $A\vec{v} = \lambda\vec{v} \implies A\vec{v} - \lambda\vec{v} = 0 \implies A\vec{v} - \lambda I\vec{v} = 0 \implies (A - \lambda I)\vec{v} = 0$, so the eigenvectors are the null space of $A - \lambda I$
  - Note that this null space is only nontrivial if $\det(A - \lambda I) = 0$, so we can find eigenvalues by finding the values of λ that make $\det(A - \lambda I) = 0$

Lecture 10, Oct 18, 2021

Finding Eigenvectors Given Eigenvalues

- Suppose we have λ, how do we solve $(A - \lambda I)\vec{v} = \vec{0}$?
- Multiplying it out gets us a system, but some equations are not independent; we are left with equations relating the components of $\vec{v}$
- This allows us to express all the other components all in terms of one component, and allows us to write the vector $\vec{v}$ as the product of a scalar (one of the components) and a vector, which an eigenvector

Inverses

- Suppose $\vec{u} = T(\vec{w})$, given $\vec{u}$ and T, can we find $\vec{w}$?
  - If T is linear, $\vec{u} = M_T\vec{w}$
  - To solve the problem we want to find another matrix N such that $NM_T = I$, so that $N\vec{u} = NM_T\vec{w} = I\vec{w} = \vec{w}$
  - N is the inverse of M_T, M_T^{-1}
- For the 2×2 case where $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $M^{-1} = \dfrac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

Determinants

- $ad - bc$ is the *determinant* of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$
- The determinant determines whether the inverse exists; the inverse only exists when $\det A \neq 0$
- $(M - \lambda I)$ is not invertible; proof:
  - Assume it is invertible, this would mean N exists such that $N(M - \lambda I) = I$
  - Therefore, $N(M - \lambda I)\vec{v} = N\vec{0} \implies I\vec{v} = \vec{0} \implies \vec{v} = \vec{0}$
  - Therefore if $M - \lambda I$ is invertible then the only $\vec{v}$ that satisfies the equation is $\vec{v} = \vec{0}$
- Therefore, the condition for finding eigenvalues is to solve for when $\det(M - \lambda I) = 0$

Lecture 11, Oct 20, 2021

Solving Systems of Equations

- Example: Solving $ax_1 + bx_2 + cx_3 = d$
  - Declare x_1 to be the leading variable and x_2, x_3 as the free variables
  - Solve for x_1 in terms of x_2 and x_3: $x_1 = \alpha x_2 + \beta x_3 + \gamma$
  - Use the additional equations $x_2 = x_2$ and $x_3 = x_3$ to write the solution in vector form
  - $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \gamma \\ 0 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} \alpha \\ 1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} \beta \\ 0 \\ 1 \end{bmatrix}$
  - Since x_2 and x_3 are free variables that can be freely assigned, they are arbitrary scalars, which means any solution can be found with $\begin{bmatrix} \gamma \\ 0 \\ 0 \end{bmatrix} + c\begin{bmatrix} \alpha \\ 1 \\ 0 \end{bmatrix} + d\begin{bmatrix} \beta \\ 0 \\ 1 \end{bmatrix}$ for some arbitrary scalars c and d
  - This is basically converting a scalar equation of the plane into a vector equation of the plane
  - In general, express the leading variables in terms of the free variables (number of leading variables equals number of equations), and then you can write the vector solution form

- Solving a system of equations corresponds to solving $A\vec{x} = \vec{b}$
- In the row picture, the solution represents the intersection of m hyperplanes of $n - 1$ dimensions in $\mathbb{R}^n$
- In the column picture, the solution represents all linear combinations of the columns of the matrix that gives $\vec{b}$

Lecture 12, Oct 25, 2021

Reduced Normal Form/Reduced Row Echelon Form

- The leading variable (one per equation) only appearing in one row makes the system easy to solve because the leading variables can be expressed in terms of free variables
- Define the *augmented matrix* $[A|\vec{b}]$ as a shorthand for the system $A\vec{x} = \vec{b}$
- The form of the augmented matrix that makes it easy to solve is called the *reduced normal form* (RNF)
  - Matrices in RNF has the following properties:
    1. The first nonzero entry in each row is 1
    2. The other entries in the columns containing the leading 1s (above and below) are 0
    3. The leading 1s move to the right as we move down the rows
    4. Any and all zero rows are at the bottom
- If the system is in RNF, we can proceed directly to solve it, but if it is not, we can use Gaussian elimination to bring it to RNF $[R|\vec{d}]$, an *equivalent* linear system that has the same solution as the original system $[A|\vec{b}]$

Gaussian Elimination

- By swapping rows, multiplying rows by nonzero scalars, and adding multiples of one row to another, we change the system of equations but don't change the solution
- If we do this algorithmically we can bring the matrix to RNF and make it easy to solve
- For the purposes of just solving the system, we usually don't have to go all the way; stop when the bottom left corner is all zeroes and do the rest using back substitution
- The allowed operations swapping rows, multiplying by nonzero scalars, and adding multiples of one row to another are called *elementary row operations*

Lecture 13, Oct 27, 2021

Gaussian Elimination

- Three outcomes are possible with Gaussian elimination:
  1. Unique solution: When every variable in the RNF is a leading variable
  2. Infinitely many solutions: When there is at least 1 free variable
  3. No solutions: When you get a row of all 0s but the last entry is nonzero
- Cases 1 and 2 are *consistent* systems (at least 1 solution); 3 is an *inconsistent* system

Rank – The True Size of a Matrix

- Suppose $A\vec{x} = \vec{b}$ and A is $m \times n$, define $\text{rank}(A)$ to be the number of leading 1s in the RNF, or equivalently the number of linearly independent rows or columns (row rank equals column rank)
- The RNF is unique even though the Gaussian elimination steps can be done in any order
- The system only has a unique solution when $\text{rank}(A) = m = n$; when $r < m$ or $r < n$ there are no solutions or infinite solutions

Lecture 14, Nov 1, 2021

Nonsquare Matrices

- When $m > n$, there are more equations than unknowns; $\operatorname{rank} A \leq n$; if $\operatorname{rank} A = n$ then there are no solutions or 1 solution
- When $m < n$, there are more unknowns than equations; $\operatorname{rank} A \leq m$; if $\operatorname{rank} A = m$, then there are always infinite solutions
- Full rank square matrices always have only 1 solution (full rank means $\operatorname{rank} A = \min\{m, n\}$)
- Any system that is not full rank can have either zero or infinite solutions

More About Inverses

- If we have the inverse of A for $A\vec{x} = \vec{b}$, then we can do $A^{-1}A\vec{x} = A^{-1}\vec{b} \implies I\vec{x} = A^{-1}\vec{b} \implies \vec{x} = A^{-1}\vec{b}$
- Definition: The square matrix A is invertible $\iff \exists B$ such that $AB = I$ and $BA = I$
- Inverses are unique: Suppose both B and C are inverses of A, then $BA = I = AC \implies BAC = CAC \implies B = C$, so all inverses are the same
- Properties of inverses:
  1. If A^{-1} exists, then when $[A|\vec{b}]$ is taken into rref, it becomes $[I|\vec{d}]$ where each variable is a leading variable
  2. If $A\vec{x} = \vec{0}$ has nontrivial solutions, then A does not have an inverse; i.e. matrices with a null space of more than the zero vector are not invertible
     - Proof: Assume A^{-1} exists, then $A^{-1}A\vec{x} = A^{-1}\vec{0} \implies I\vec{x} = \vec{0} \implies \vec{x} = \vec{0}$; this means that if A is invertible then $\vec{x} = \vec{0}$ is the only solution to $A\vec{x} = \vec{0}$
  3. If matrices A and B are both invertible and of the same size, then AB is also invertible, and $(AB)^{-1} = B^{-1}A^{-1}$
     - Proof: $B^{-1}A^{-1}AB = B^{-1}IB = B^{-1}B = I$, and $ABB^{-1}A^{-1} = A^{-1}IA = A^{-1}A = I$, so by the definition of the inverse $B^{-1}A^{-1}$ is the inverse of AB
  4. If A is invertible, then A^{-1} is also invertible, and $(A^{-1})^{-1} = A$
  5. Every invertible matrix A can be expressed as a product of elementary matrices

Lecture 15, Nov 3, 2021

Inverses by Gaussian Elimination

- To find the inverse using Gaussian elimination, we introduce elementary matrices for GE operations:
  - Swapping rows: $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$
  - Multiplying a row by a scalar: $\begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ea & eb \\ c & d \end{bmatrix}$
  - Add or subtract multiples of other rows: $\begin{bmatrix} 1 & e \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a + ec & b + ed \\ c & d \end{bmatrix}$
- Notice that the elementary matrices we obtained for these operations can also be obtained by applying the same operation to the identity!
  - Swapping rows, multiplying rows, or adding or subtracting multiples of rows in the identity gets us the elementary matrix we need
  - Suppose the elementary matrix is E and applies the operation we need; then because $EI = E$ and we know what E does to a matrix, we can calculate EI to get E
- Elementary matrices are one operation from the identity, so we can return back to the identity by applying its inverse (elementary matrices are invertible)
- Now apply GE to $[A|I]$, which corresponds to solving $AX = I$
  - As we do this we'll get $E_n E_{n-1} \cdots E_2 E_1 A$
  - Since applying GE to an invertible matrix produces I, we have $E_n E_{n-1} \cdots E_2 E_1 A = I$, therefore $E_n E_{n-1} \cdots E_2 E_1 = A^{-1}$

- Since we're also applying the same operations to the right side, the right side becomes $E_nE_{n-1}\cdots E_2E_1I = E_nE_{n-1}\cdots E_2E_1$, which is what we want
  - Therefore we can apply GE to $[A|I]$, and whatever becomes of I when A is reduced to I is A^{-1}
- Then $A = (E_nE_{n-1}\cdots E_2E_1)^{-1}$ since $(A^{-1})^{-1} = A$; thus $A = E_1^{-1}E_2^{-1}\cdots E_{n-1}^{-1}E_n^{-1}$
  - This leads to a property that any invertible matrix A can be written as a product of elementary matrices

Lecture 16, Nov 15, 2021

Least Squares Projections

- When $A\vec{x} = \vec{b}$ has no solution, the most common cause is because there are too many equations and not enough variables to satisfy all of them at once (tall and thin matrix)
- In such cases we're often still interested in trying to find the best solution possible (try to find $A\vec{x} \approx \vec{b}$)
- An important application is in the context of curve fitting – trying to find the right parameters in an equation so the equation matches the model as closely as possible
  - e.g. we collected 3 data points, and now we want to fit $y = mx + b$ to it, but this has no exact solution since the 3 points may not lie on the same line

The Transpose

- The transpose of a matrix A^T is obtained by making all rows columns
- Properties of the transpose:
  1. $(A^T)^T = A$
  2. $(cA)^T = cA^T$
  3. $(AB)^T = B^TA^T$
  4. If A is invertible, then A^T is also invertible and $(A^T)^{-1} = (A^{-1})^T$
     - Proof: $A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$, and $(A^{-1})^TA^T = (AA^{-1})^T = I^T = I$

Solving the General Least Squares Problem

- Looking at this problem in the column picture, we have $A\vec{x} = x_1\vec{a}_n + x_2\vec{a}_n + \cdots + x_n\vec{a}_n$
- Thinking in the column picture, we're trying to find the x_i such that the linear combination of the columns add up to $\vec{b}$
- All the linear combinations of the columns of A form a subspace, but the vector $\vec{b}$ could lie outside the space, which is why there is no solution
- To find the least squares solution, we need to project $\vec{b}$ into this subspace defined by the columns of A

Lecture 17, Nov 17, 2021

General Least Squares Continued

- Define a vector space that consists of all the linear combinations of the columns of matrix A
- In a least squares problem $\vec{b}$ does not lie in this space
- We want to find $\vec{x}$ such that $\vec{e} = A\vec{x} - \vec{b}$ is minimized
- The smallest such $\vec{e}$ would be orthogonal to the column space of A since shortest distance is the orthogonal distance
  - This means that $\vec{e} \perp \vec{a}_i$ where $\vec{a}_i$ are the vectors in the columns of the matrix
- Since multiplying a matrix by a vector computes the dot product between the rows of the matrix and the vector, we can compute the dot product between the columns of the matrix and the vector by transposing the matrix
- Therefore $A^T\vec{e} = \vec{0}$ since orthogonality means dot product is 0

- $$\begin{aligned} & A^T\vec{e} = \vec{0} \\ \implies & A^T(A\vec{x} - b) = \vec{0} \\ \implies & A^TA\vec{x} - A^Tb = \vec{0} \\ \implies & A^TA\vec{x} = A^Tb \end{aligned}$$
  - This is called the *normal* system of equations
  - Define $A^* = A^TA$ and $\vec{b}^* = A^T\vec{b}$ then $A^*\vec{x} = \vec{b}^*$
  - This system does have a solution since A^TA is invertible (but only if A has linearly independent columns)
    * Proof: $A^TA\vec{x} = \vec{0} \implies A\vec{x} \in N(A^T)$, but $A\vec{x} \in C(A)$ and $A\vec{x} \in N(A^T) \implies A\vec{x} = \vec{0}$ since $C(A) \perp N(A^T)$; therefore $\vec{x} \in N(A)$, and if A has linearly independent columns the only such $\vec{x} = \vec{0} \implies A\vec{x} = \vec{0} \implies N(A^TA) = \{\vec{0}\} \implies A^TA$ is invertible
      - Note $N(A) \perp C(A^T)$ because $A\vec{x} = \vec{0}$ requires that the dot product of $\vec{x}$ with every row of A is 0, and thus also the dot product of $\vec{x}$ with all the linear combinations of the rows of A is 0 and so the two subspaces are orthogonal
      - So if we transpose the two we get $N(A^T) \perp C(A)$
  - A^* is a square $n \times n$ matrix where n is the number of unknowns
  - Solving directly for the least squares solution gets us $\vec{x} = (A^TA)^{-1}A^T\vec{b}$

Lecture 18, Nov 22, 2021

Initial Value Problems

- An IVP is a differential equation attached with initial conditions that allow for a particular solution
- General problem: Given $\dfrac{\mathrm{d}y}{\mathrm{d}t} = y'(t) = cy(t)$ where c is a scalar and $y(0) = y_0$ solve for $y(t)$ for $t > 0$
- We begin with $y(t) = y_0e^{ct}$, so $\dfrac{\mathrm{d}y}{\mathrm{d}t} = cy_0e^{ct} = cy(t)$ so it satisfies the DE; since $y(0) = y_0e^0 = y_0$ it also satisfies the initial conditions
- Generalizing this we have given $y'(t) = f(t, y(t))$ and $y(0) = y_0$
- In this course we will focus on the numerical approach to solving this equation
- We start with the initial value $y(0) = y_0$, and now we can evaluate $f(t, y(t))$ to obtain a slope and take a step forward by some amount Δt and use the first derivative information to find the y value for this new point, and repeat until the desired t is reached

Reducing the Error

- To reduce the error we can try to get a better slope estimate
- We can take $S = \dfrac{1}{2}(S_L + S_R)$ where $S_L = f(t_n, y_n)$ and $S_R = f(t_{n+1}, y_{n+1})$, with y_{n+1} estimated using the first method, i.e. take slope to be the average of the two points

Lecture 19, Nov 24, 2021

Systems of Linear Differential Equations

- Consider the IVP: $\begin{cases} x'(t) = ax(t) + by(t) \\ y'(t) = cx(t) + dy(t) \end{cases}$
  - This is a coupled system so we can't simply use the method in the last lecture to the two equations separately
  - We can write this using a matrix: Let $Z = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$, then $Z' = \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix}$, so the system can now be represented as $Z' = \begin{bmatrix} a & b \\ c & d \end{bmatrix} Z = AZ$

- To capture the initial conditions, let $Z_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$
  - This is referred to as a *state-space* description of the system and only applies to systems described by linear differential equations with constant coefficients
    - Z is the *state* of the system
  - For these systems there are analytical solutions
- Euler's method can be applied to such a system in the same way
  - Let $t_{n+1} = t_n + \Delta t$
  - $Z_{n+1} = Z_n + \Delta t Z'_n = Z_n + \Delta t A Z_n$
- Improved Euler's method:
  - $Z^*_{n+1} = Z_n + \Delta t A Z_n$ is the Euler update
  - $Z_{n+1} = Z_n + \frac{\Delta t}{2}(AZ_n + AZ^*_{n+1})$ is the improved update, where the average of the rate of change at Z_n and Z^*_{n+1} is averaged

Higher Order Systems

- Higher order systems of linear differential equations can be converted into systems of first order linear ODEs, by making the derivatives a part of the state
- Consider a general second-order differential equation $\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = y''(t) = f(t, y(t), y'(t))$, and for initial values $y(0) = y_0$ and $y'(0) = y'_0$
- Let $Z = \begin{bmatrix} y(t) \\ y'(t) \end{bmatrix} \implies Z' = \begin{bmatrix} y'(t) \\ y''(t) \end{bmatrix}$, and $Z_0 = \begin{bmatrix} y_0 \\ y'_0 \end{bmatrix}$ and if the system is linear, we can express this in the same way using a matrix as a system of first order equations
- Example: $y''(t) = -y(t)$, $y(0) = 1$, $y'(0) = 0$
  - Set up $Z' = AZ$, where $Z = \begin{bmatrix} y(t) \\ y'(t) \end{bmatrix} \implies Z' = \begin{bmatrix} y'(t) \\ y''(t) \end{bmatrix}$
  - $\begin{bmatrix} y'(t) \\ y''(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} y(t) \\ y'(t) \end{bmatrix} \implies A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$
  - Now Euler's method can be used in exactly the same way

Lecture 20, Nov 29, 2021

Boundary Value Problems

- A boundary value problem involves a system described by a differential equation, but unlike an IVP, the known values are no longer at 0
- In a BVP the independent variable is typically not time but some kind of spacial variable (hence boundaries)
- Example: Beam deflection problem; in this case the known points are the supports which have a deflection of 0
- To solve BVPs numerically, begin by partitioning the interval into n evenly spaced subintervals, and set the step size as $\Delta x = \frac{b-a}{n}$
- Approximate derivatives using finite differences (secant line slopes)
  - The forward difference $\Delta_F f(x) = \frac{f(x+\Delta x) - f(x)}{\Delta x}$
  - The backward difference $\Delta_B f(x) = \frac{f(x) - f(x-\Delta x)}{\Delta x}$
  - The central difference $\Delta_C = \frac{f(x+\Delta x) - f(x-\Delta x)}{2\Delta x}$
  - We can use different combinations of these finite differences to estimate higher order derivatives

- $f''(x) = \Delta_B[\Delta_F f(x)]$
$$= \frac{\Delta_F f(x) - \Delta_F f(x-\Delta x)}{\Delta x}$$
$$= \frac{1}{\Delta x}\left[\frac{f(x+\Delta x) - f(x)}{\Delta x} - \frac{f(x) - f(x-\Delta x)}{\Delta x}\right]$$
$$= \frac{f(x+\Delta x) - 2f(x) + f(x-\Delta x)}{(\Delta x)^2}$$
- Example: Given $y'' + 2y' + y = 0$ and $y(0) = 0$, $y(1) = 1$, find y that satisfies this for $x \in [0,1]$
  - Approximate $y'(x) \approx \Delta_C y(x)$ and $y''(x) \approx \Delta_B \Delta_F y(x)$
  - Substitute this back into the equation: $\frac{f(x+\Delta x) - 2f(x) + f(x-\Delta x)}{(\Delta x)^2} + 2\frac{y(x+\Delta x) - y(x-\Delta x)}{2\Delta x} + y(x) = 0$
  - Simplify: $\left(\frac{1}{(\Delta x)^2} + \frac{2}{2\Delta x}\right) y(x+\Delta x) + \left(\frac{-2}{(\Delta x)^2} + 1\right) + \left(\frac{1}{(\Delta x)^2} - \frac{2}{2\Delta x}\right) y(x-\Delta x) = 0$
  - This is now a difference equation, with no more derivatives but instead values of y at discrete points
  - Divide the interval $[0,1]$ into $n = 5 \implies \Delta x = 0.2$ subintervals
  - Now we evaluate the difference equation at each of the interior grid points (open interval $(0,1)$), which will get us an algebraic equation:
    * $$\begin{cases} 30y(0.4) - 49y(0.2) + 20y(0) = 0 \\ 30y(0.6) - 49y(0.4) + 20y(0.2) = 0 \\ 30y(0.8) - 49y(0.6) + 20y(0.4) = 0 \\ 30y(1) - 49y(0.8) + 20y(0.6) = 0 \end{cases}$$
      - Note how the coefficients aren't changing since they only depend on step size
  - Now we can sub in $y(1)$ and $y(0)$ into the equations above
  - The system of y values can now be formulated in the form of $A\vec{x} = \vec{b}$, which allows it to be solved

Lecture 21, Dec 6, 2021

LU Decomposition

- Let A be an invertible square $n \times n$ matrix
- We can write A in a factored form as $A = LU$, where L is a lower-triangular matrix (zeros above the diagonal), and U is an upper-triangular matrix (zeros below the diagonal)
- To solve $A\vec{x} = \vec{b}$ using LU decomposition:
  1. Rewrite $A\vec{x} = \vec{b} \implies LU\vec{x} = \vec{b}$, and let $\vec{y} = U\vec{x}$, then $LU\vec{x} = \vec{b} \implies L\vec{y} = \vec{b}$
  2. Solve for $\vec{y}$
  3. Substitute $\vec{y}$ in $U\vec{x} = \vec{y}$, and solve for $\vec{x}$
- We've separated the decomposition phase, which is computationally heavy, and the part where $\vec{b}$ is brought in, which is computationally light
  - LU decomposition is better than computing the inverse, because if A is sparse computing the inverse might lead to a lot of numerical errors
- Because L and U have special forms, solving $U\vec{x} = \vec{y}$ is much easier
- LU factorizations aren't unique, but to solve the problem we can use any one
- Example: $\begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$
  1. Factor $A = LU \implies \begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$

2. Let $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = U\vec{x}$
3. $L\vec{y} = \vec{b} \implies \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \implies \vec{y} = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix}$
   - By forward substitution $\begin{cases} 2y_1 = 2 & \implies y_1 = 1 \\ -3y_1 + y_2 = 2 & \implies y_2 = 5 \\ 4y_1 - 3y_2 + 7y_3 = 3 & \implies y_3 = 2 \end{cases}$
4. $U\vec{x} = \vec{y} \implies \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} \implies \vec{x} = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$
   - By back substitution $\begin{cases} x_3 = 2 & \implies x_3 = 2 \\ x_2 + 3x_3 = 5 & \implies x_2 = -1 \\ x_1 + 3x_2 + x_3 = 1 & \implies x_1 = 2 \end{cases}$

- To find LU:
  - During elimination we have $E_n \cdots E_1 A = U$, so $L = E_1^{-1} \cdots E_n^{-1}$
  - As long as we don't do any back substitution, $E_1^{-1} \cdots E_n^{-1}$ is going to be lower-triangular

Lecture 22, Dec 8, 2021

Computing LU Decompositions

- There is a simpler procedure for doing LU decompositions:
  1. Reduce matrix A to upper-triangular U by elimination; keep track of the multipliers used to introduce the leading 1s and zeroes below
  2. Along the main diagonal of L, put the reciprocal of the multiplier that introduced the leading 1 in U in that position
  3. Below the main diagonal of L, put the negative of the multiplier used to introduce the 0 in U in that position

Review: The 4 Central Problems of Linear Algebra

1. Solving linear systems $A\vec{x} = \vec{b}$ where $m = n$
   - Arises from mathematical models of static engineering systems
2. Least squares $A\vec{x} = \vec{b}$ where $m \neq n$
   - Arises when fitting mathematical models to real data with noise
3. Eigenvalues $A\vec{x} = \lambda\vec{x}$ for $m = n$
   - Arises when analyzing and designing dynamic systems (systems of differential equations, control systems)
4. The singular value decomposition $A\vec{v} = \sigma\vec{u}$ for $m \neq n$ (analogue of eigenvalues for nonsquare matrices)
   - Arises with data and image compression problems