MATH 257 Final CARE Review

Please join the queue!

https://queue.illinois.edu/q/queue/955

Midterm 1 Topics

- Linear systems
  - Solving systems with matrices
- Reduced row echelon form
  - Pivot columns: basic and free variables
  - Row operations
- Vectors and spans
- Matrix operations
  - Addition, subtraction, scalar multiplication, linear combinations
  - Transposition
- Matrix multiplication
  - Properties of matrix multiplication
- Matrix inverses
  - What matrices are invertible?
  - Elementary matrices

Midterm 2 Topic Summary

- LU Decomposition
  - Lower/Upper Triangular Matrix
  - LU for Linear Systems
  - Permutation Matrix
- Vectors and Spans
  - Inner Product
  - Orthogonality
  - Linear Independence
- Subspaces
  - Column Space
  - Null Space
- Basis and Dimension
  - Fundamental Subspaces
  - Orthonormal bases
  - Orthogonal/normal Complements
- Graph and Adjacency Matrices
- Coordinates
  - Coordinate Matrices

Midterm 3 Topic Summary

- Linear Transformation
- Coordinate Matrices
- Determinants
- Eigenvectors and eigenvalues
- Markov Matrices
- Diagonalization
- Matrix powers
    - Matrix exponential
- Linear differential equations
- Matrix projections
- Least squares solutions

Topic Summary – New Content

- Gram-Schmidt Method
- Spectral Theorem
- SVD
- Low Rank SVD
- Spectral Theorem
- Psuedo Inverse
- PCA
- Complex Numbers

^not historically part of the final

Linear Systems

$$a_1x_1 + \ldots + a_nx_n = b$$

and matrices

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Linear systems must have either:

1. One unique solution
2. Infinite solutions
3. No solutions

Equivalent linear systems have the same set of solutions.

You can represent a linear system with matrices…

linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

coefficient matrix

$$\begin{bmatrix} \boxed{\begin{matrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{matrix}} & \begin{matrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{matrix} & \begin{matrix} \cdots \\ \cdots \\ \ddots \\ \cdots \end{matrix} & \begin{matrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{matrix} \end{bmatrix}$$

Column vector

augmented matrix

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \multicolumn{4}{c|}{\boxed{a_{m1} \quad a_{m2} \quad \cdots \quad a_{mn}}} & b_m \end{array}\right]$$

Row vector

We often define a matrix in terms of its columns or its rows:

$\mathbf{a}_n$ are all column vectors

$$A := \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \qquad or \qquad A := \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \vdots \\ \mathbf{R}_m \end{bmatrix}$$

$\mathbf{R}_m$ are all row vectors

Echelon Forms

Pivot

$$\begin{bmatrix} 3 & 1 & 2 & 0 & 5 \\ 0 & 2 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Free

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 5 & -4 \\ 0 & 1 & -2 & 2 & 1 & -3 \end{bmatrix}$$

Row Echelon Form (REF):

1. All nonzero rows above rows of all zeros
2. Leading entry (leftmost nonzero number) is strictly to the right of the leading entry of the row above

Reduced Row Echelon Form (RREF):

1. Leading entries of nonzero rows are all 1
2. Each leading entry is the only nonzero entry in the column

Elementary Row Operations

Elementary operations do not change the solution set of a system.

There are three kinds:

1. Replacement ($R_1 \to R_1 + a*R_2$)
2. Scaling ($R_1 \to a*R_1$)
3. Interchange ($R_1 \to R_2$)

All elementary operations are reversible. Two matrices are **row equivalent** if elementary operations can turn one into the other.

Matrix Operations

a) **The sum of** $A + B$ is

$$\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

b) **The product** cA for a scalar c is

$$\begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}$$

Addition: only defined for matrices with the **same dimensions**

Subtraction: the same as addition

Scalar multiplication: every entry is multiplied by the **scalar**

- Scalar = any real number

Linear combinations: any mixture of scalar multiplication and addition/subtraction of matrices

- **span(a,b)** is a set of ALL the possible linear combinations of **a** and **b**

Matrix Operations (cont.)

Transpose: switch rows and columns

2	4	-1
-10	5	11
18	-7	6

2	-10	18
4	5	-7
-1	11	6

Matrix-vector multiplication: $A\mathbf{x} = x_1\mathbf{a_1} + x_2\mathbf{a_2} + \dots + x_n\mathbf{a_n}$ which means you multiply the **entries** of the vector with the **columns** of the matrix

Matrix multiplication

Only defined for two matrices A and B if

- A has the dimensions *m* x *n* and B has the dimensions *n* x *p*
- A^k (exponent) is only defined for a **square** matrix

Each entry of AB is a linear combination of a **row of A** with a **column of B.**

$$AB = \begin{bmatrix} \mathbf{R}_1\mathbf{C}_1 & \dots & \mathbf{R}_1\mathbf{C}_p \\ \mathbf{R}_2\mathbf{C}_1 & \dots & \mathbf{R}_2\mathbf{C}_p \\ \mathbf{R}_m\mathbf{C}_1 & \dots & \mathbf{R}_m\mathbf{C}_p \end{bmatrix} \quad \text{and} \quad (AB)_{ij} = \mathbf{R}_i\mathbf{C}_j = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

Properties of Matrix Multiplication

(a) $A(BC) = (AB)C$ *(associative law of multiplication)*

(b) $A(B + C) = AB + AC$, $(B + C)A = BA + CA$ *(distributive laws)*

(c) $r(AB) = (rA)B = A(rB)$ *for every scalar* r,

(d) $A(rB + sC) = rAB + sAC$ *for every scalars* r, s *(linearity of matrix multiplication)*

(e) $I_m A = A = A I_n$ *(identity for matrix multiplication)*

Transpose Theorem: $(AB)^T = B^T A^T$

Matrix multiplication is NOT COMMUTATIVE: AB ≠ BA

Elementary Matrices

Identity Matrices

$1 \times 1 \quad [1]$

$2 \times 2 \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$3 \times 3 \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

etc.

Any matrix that can be form from the identity matrix with **one** elementary row operation.

Ex.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{R_2 \to 3R_2}{\rightsquigarrow} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{R_2 \leftrightarrow R_3}{\rightsquigarrow} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Matrix Inverses

Determinants:

$$\frac{1}{ad-bc}$$

For the matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Definition of an inverse:

$$AC = I_n$$

Requirements for a matrix to be invertible:

1. It has to be square
2. The determinant of the matrix cannot be 0 or
3. The RREF of A is the identity matrix or
4. A has as many pivots as columns/rows

Statements 2, 3, and 4 mean the same thing.

Calculating an Inverse

For 2x2:

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

Elementary Matrix strategy:

$$A^{-1} = E_m E_{m-1} \ldots E_1 = E_m E_{m-1} \ldots E_1 I_n.$$

OR: set up an augmented matrix with the identity and reduce to RREF

Properties of Matrix Inverses

(a) A^{-1} *is invertible and* $\left(A^{-1}\right)^{-1} = A$ *(i.e. A is the inverse of A^{-1}).*

(b) AB *is invertible and* $(AB)^{-1} = B^{-1}A^{-1}$.

(c) A^T *is invertible and* $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$.

Inverses are unique! Every invertible matrix only has one inverse.

Multiplying by a matrix inverse is the closest we get to dividing matrices.

Theorem 14. *Let A be an invertible $n \times n$ matrix. Then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.*

Upper/Lower Triangular Matrices

Upper Triangular:

$$\begin{bmatrix} \star & \star & \star & \star & \star \\ 0 & \star & \star & \star & \star \\ 0 & 0 & \star & \star & \star \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & 0 & * \end{bmatrix}$$

Finding this is like doing REF with only row replacement

Lower Triangular:

$$\begin{bmatrix} \star & 0 & 0 & 0 & 0 \\ \star & \star & 0 & 0 & 0 \\ \star & \star & \star & 0 & 0 \\ \star & \star & \star & \ddots & \vdots \\ \star & \star & \star & \star & \star \end{bmatrix}$$

Keep track of your row operations to find L

LU Decomposition:

$A = LU$

- Not all matrices have LU decompositions
- LU decompositions are not unique (unlike inverses)

Finding the LU Decomposition

Determine the LU-decomposition of $\begin{bmatrix} 1 & 2 & 2 \\ 4 & 4 & 4 \\ 4 & 4 & 8 \end{bmatrix}$

$$\begin{bmatrix} 1 & 2 & 2 \\ 4 & 4 & 4 \\ 4 & 4 & 8 \end{bmatrix} \underset{R_2 \to R_2 \boxed{-4} R_1, R_3 \to R_3 \boxed{-4} R_1}{\rightsquigarrow} \begin{bmatrix} 1 & 2 & 2 \\ 0 & -4 & -4 \\ 0 & -4 & 0 \end{bmatrix} \underset{R_3 \to R_3 \boxed{-1} R_2}{\rightsquigarrow} \begin{bmatrix} 1 & 2 & 2 \\ 0 & -4 & -4 \\ 0 & 0 & 4 \end{bmatrix}$$

1) Col 1 Row 2

2) Col 1 Row 3

3) Col 3 Row 3

$$, L := \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 4 & 1 & 1 \end{bmatrix} \qquad U := \begin{bmatrix} 1 & 2 & 2 \\ 0 & -4 & -4 \\ 0 & 0 & 4 \end{bmatrix}$$

LU for Linear Systems

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}}_{U}$$

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}}_{L} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}}_{\mathbf{b}}$$

$$\underbrace{\begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}}_{U} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 5 \\ -12 \\ 2 \end{bmatrix}}_{\mathbf{c}}.$$

Use **LU decomposition** to solve a linear system if:

1. A is nxn matrix
2. A = LU
3. b ∈ R^n

Step-by-step Algorithm

1. Find L and U
2. Solve for c using Lc = b
3. Solve for x using Ux = c

$$A\mathbf{x} = \mathbf{b}$$

$$L\mathbf{c} = \mathbf{b} \longrightarrow U\mathbf{x} = \mathbf{c}$$

$$A\mathbf{x} = (LU)\mathbf{x} = L(U\mathbf{x}) = L\mathbf{c} = \mathbf{b}$$

Permutation Matrices: for matrices that don't have an LU decomposition

Theorem 21. *Let A be $n \times n$ matrix. Then there is a permutation matrix P such that PA has an LU-decomposition.*

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \end{bmatrix} \underset{R_1 \leftrightarrow R_3}{\rightsquigarrow} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = PA$$

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Step-by-step:

- Use the interchange operation done on **A** to an equivalent size identity matrix, this will be your **P** matrix
- Solve the for the **LU** decomposition of **PA**

When we apply the $\mathbf{P}^{-1}$ to **LU** (on the right), we'll be able to get the original value of **A**

$$\boxed{P}A = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ .5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{=:L} \underbrace{\begin{bmatrix} 2 & 1 & 0 \\ 0 & .5 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{=:U}$$

Inner Product, Norm, and Distance

If $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$, then $\mathbf{v} \cdot \mathbf{w}$ is

$$v_1 w_1 + v_2 w_2 + \cdots + v_n w_n$$

The **inner product** of $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ is

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \mathbf{w}$$

AKA the dot product

It is a scalar!

Definition. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$.

The **norm** (or **length**) of $\mathbf{v}$ is

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + \cdots + v_n^2}.$$

The **distance** between $\mathbf{v}$ and $\mathbf{w}$ is

$$\text{dist}(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\|.$$

The norm is also a scalar!

Properties of the Inner Product: similar to scalars

Theorem 22. *Let* $\mathbf{u}, \mathbf{v}$ *and* $\mathbf{w}$ *be vectors in* $\mathbb{R}^n$*, and let* c *be any scalar. Then*

(a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ Commutative!

(b) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$ Distributive!

(c) $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$ Associative!

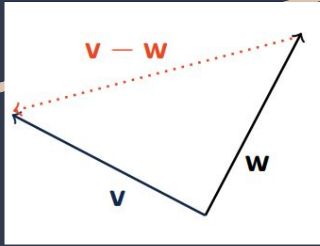
(d) $\mathbf{u} \cdot \mathbf{u} \geq \mathbf{0}$, *and* $\mathbf{u} \cdot \mathbf{u} = \mathbf{0}$ *if and only if* $\mathbf{u} = \mathbf{0}$.

Orthogonality

(fancy word for perpendicular)

Vectors are orthogonal if their **dot product is zero**.

Why? The dot product of two non-zero vectors can only be zero if the angle between them is 90.



Orthonormality

A **unit vector** in $\mathbb{R}^n$ is vector of length 1.

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

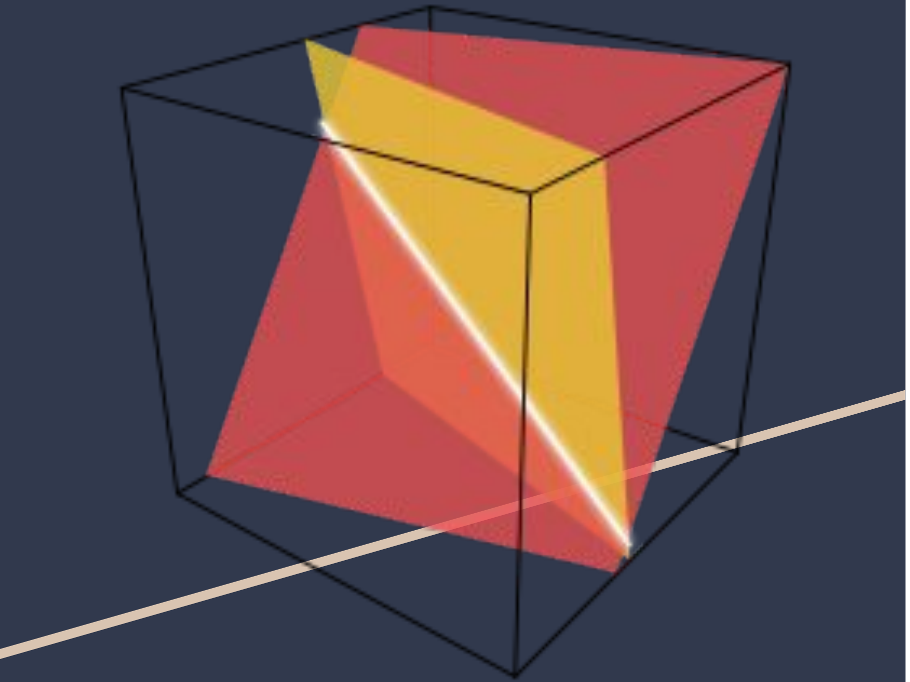
Orthonormal sets are all orthogonal to each other and unit vectors.

Ex.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Subspaces



W is a **subspace** of **V,** if:

- W contains the 0 vector
- Adding any 2 vectors in W together gives a vector also in W
- Multiplying any vector in W by any scalar gives a vector also in W

Theorem 24. *Let* $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_m} \in \mathbb{R}^n$. *Then* $\mathrm{Span}(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_m})$ *is a subspace of* $\mathbb{R}^n$.

Vector Spaces ‘V’

$u, v, w \in V$ and for all scalars $c, d \in \mathbb{R}$:

- $\mathbf{u} + \mathbf{v}$ is in V. (V is “closed under addition”.)
- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
- $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
- There is a vector (called the zero vector) $\mathbf{0}_V$ in V such that $\mathbf{u} + \mathbf{0}_V = \mathbf{u}$.
- For each $\mathbf{u}$ in V, there is a vector $-\mathbf{u}$ in V satisfying $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}_V$.
- $c\mathbf{u}$ is in V. (V is “closed under scalar multiplication”.)
- $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
- $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
- $(cd)\mathbf{u} = c(d\mathbf{u})$.
- $1\mathbf{u} = \mathbf{u}$.

Column Spaces

Definition. The **column space**, written as $\text{Col}(A)$, of an $m \times n$ matrix A is the set of all linear combinations of the columns of A. If $A = \begin{bmatrix} \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_n} \end{bmatrix}$, then $\text{Col}(A) = \text{span}(\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n})$.

$$A = \begin{bmatrix} 1 & -10 & -24 & -42 \\ 1 & -8 & -18 & -32 \\ -2 & 20 & 51 & 87 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -10 & -24 & -42 \\ 1 & -8 & -18 & -32 \\ -2 & 20 & 51 & 87 \end{bmatrix} \xrightarrow[R_3 + 2R_1 \rightarrow R_3]{R_2 - R_1 \rightarrow R_2} \begin{bmatrix} \underline{1} & -10 & -24 & -42 \\ 0 & \underline{2} & 6 & 10 \\ 0 & 0 & \underline{3} & 3 \end{bmatrix}$$

$$\text{Col}(A) = \left\{ \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} -10 \\ -8 \\ 20 \end{bmatrix}, \begin{bmatrix} -24 \\ -18 \\ 51 \end{bmatrix} \right\}$$

How to solve for Col(A):

1. Put matrix A into REF
2. Find all the pivots of A
3. Map the pivots to the columns of your original matrix, A

Null Spaces

Definition. The **nullspace** of an $m \times n$ matrix A, written as $\text{Nul}(A)$, is the set of all solutions to the homogeneous equation $A\mathbf{x} = \mathbf{0}$; that is, $\text{Nul}(A) = \{\mathbf{v} \in \mathbb{R}^n \; : \; A\mathbf{v} = \mathbf{0}\}$.

How to solve for Nul(A):

1. Set matrix A into **Augmented Matrix** with zeros on the right (**Ax = 0**)
2. Get A into **RREF**
3. Solve for **x**

Nul(A) = span($\mathbf{x}_1$, $\mathbf{x}_2$,...)

Null Space Example

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix} \xrightarrow{\text{RREF}} A = \begin{bmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- $x_1 - 2x_2 - x_4 + 3x_5 = 0$
- $x_3 + 2x_4 - 2x_5 = 0$

$$x_1 = 2x_2 + x_4 - 3x_5$$
$$x_3 = -2x_4 + 2x_5$$

$$\begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix}$$

$$\text{Nul } A = \text{Span}\left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Linear Independence

Definition. Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are said to be **linearly independent** if the equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution (namely, $x_1 = x_2 = \cdots = x_p = 0$).
We say the vectors are **linearly dependent** if they are not linearly independent.

Theorem 30. *Let A be an $m \times n$ matrix. The following are equivalent:*

- *The columns of A are linearly independent.*
- *$A\mathbf{x} = \mathbf{0}$ has only the solution $\mathbf{x} = \mathbf{0}$.*
- *A has n pivots.*
- *there are no free variables for $A\mathbf{x} = \mathbf{0}$.*

Basis and Dimension

Definition. Let V be a vector space. A sequence of vectors $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ in V is a **basis** of V if

- $V = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_p)$, and
- $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ are linearly independent.

The number of vectors in a basis of V is the **dimension** of V.

Basis and Dimension example

Is $\left(\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \right)$ a basis of $\mathbb{R}^3$?

Our set has 3 vectors and

$$\dim \mathbb{R}^3 = 3$$

General Definition of dimension

$$\dim \mathbb{R}^n = n$$

Next, we check linear independence

$$\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 1 & 3 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 5 \end{bmatrix}$$

*Recall: All pivots for **nxn** matrix means **linear independence**

Theorem 33. *A basis is a minimal spanning set of V; that is the elements of the basis span V but you cannot delete any of these elements and still get all of V.*

Basis and Dim of four subspaces:

Rank [r] : Number of pivots matrix has

Let A be an $m \times n$ matrix with rank r

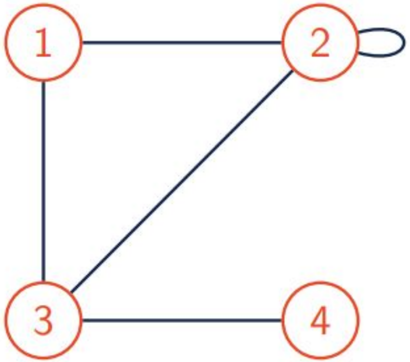
- dim Nul(A) = n - r
- dim Col(A) = r
- dim Nul(A^T) = m - r
- dim Col(A^T) = r

Graphs and Adjacency Matrices

A **graph** is a set of nodes (or: vertices) that are connected through edges.

Definition. Let $\mathcal{G}$ be a graph with n nodes. The **adjacency matrix** of $\mathcal{G}$ is the $n \times n$-matrix $A = (a_{ij})$ such that

$$a_{ij} = \begin{cases} 1 & \text{if there is an edge between node } i \text{ and node } j \\ 0 & \text{otherwise .} \end{cases}$$



$$\begin{array}{cccc} \text{N1} & \text{N2} & \text{N3} & \text{N4} \end{array}$$
$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Node 1: Connected to N2 & N3

Node 2: Connected to N1, N2, & N3

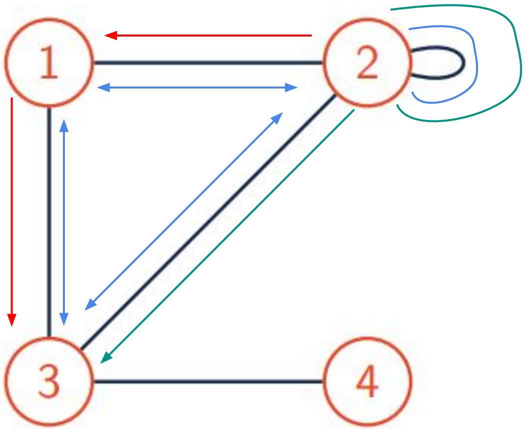
Node 3: Connected to N1, N2 & N4

Node 4: Connected to N3

Walks and Paths

Definition. A **walk of length** k on a graph of is a sequence of $k+1$ vertices and k edges between two nodes (including the start and end) that may repeat. A **path** is walk in which all vertices are distinct.

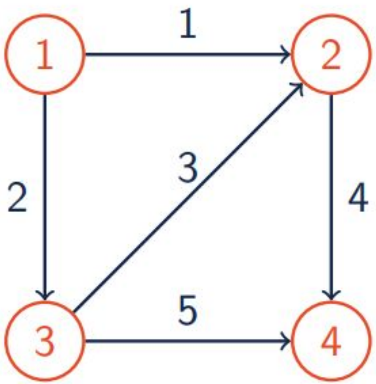
Example. Count the number of walks of length 2 from node 2 to node 3 and the number of walks of length 3 from node 3 back to node 3:



- Node 2 to Node 3: 2 walks of length 2
- Node 3 to Node 3: 3 walks of length 3

Directed Graphs

Definition. A **directed graph** is a set of vertices connected by edges, where the edges have a direction associated with them.



$$\begin{array}{c} \begin{matrix} N1 & N2 & N3 & N4 \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

Node 1: Nothing pointing to N1

Node 2: N1 and N3 pointing to N2

Node 3: N1 points to N3

Node 4: N2 and N3 pointing to N4

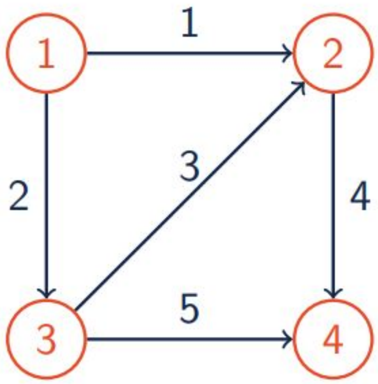
Definition. Let G be a directed graph with m edges and n nodes. The **adjacency matrix** of G is the $n \times n$ matrix $A = (a_{i,j})_{i,j}$ with

$$a_{i,j} = \begin{cases} 1, & \text{if there is a directed edge from node } j \text{ to node } i \\ 0, & \text{otherwise} \end{cases}$$

Edge–Node Incidence

Definition. Let G be a directed graph with m edges and n nodes. The **edge-node incidence matrix** of G is the $m \times n$ matrix $A = (a_{i,j})_{i,j}$ with

$$a_{i,j} = \begin{cases} -1, & \text{if edge } i \text{ leaves node } j \\ +1, & \text{if edge } i \text{ enters node } j \\ 0, & \text{otherwise} \end{cases}$$

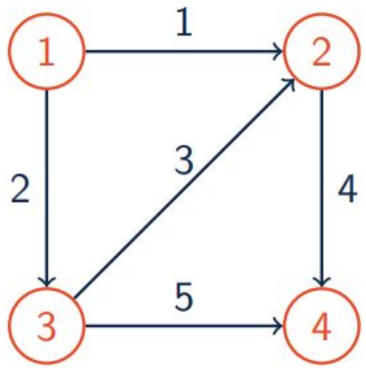


$$\begin{array}{cl} \begin{matrix} \text{N1} & \text{N2} & \text{N3} & \text{N4} \end{matrix} & \\ \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} & \begin{matrix} \text{Edge 1: Leaves N1; Enters N2} \\ \text{Edge 2: Leaves N1; Enters N3} \\ \text{Edge 3: Leaves N3; Enters N2} \\ \text{Edge 4: Leaves N2; Enters N4} \\ \text{Edge 5: Leaves N3; Enters N4} \end{matrix} \end{array}$$

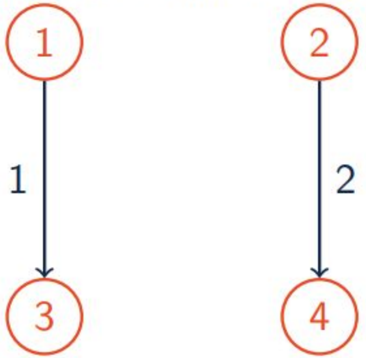
‘Connectedness’

Definition. A **connected component** of an undirected graph is a part in which any two vertices are connected to each other by paths, and which is connected to no additional vertices in the rest of the graph. The connected components of a directed graph are those of its underlying undirected graph. A graph is **connected** if only has one connected component.

A graph with one connected component:



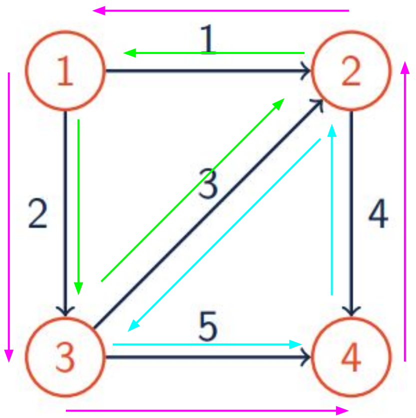
A graph with two connected components:



Theorem 40. *Let $\mathcal{G}$ be a directed graph and let A be its edge-node incidence matrix. Then* $\dim \mathrm{Nul}(A)$ *is equal to the number of connected components of $\mathcal{G}$.*

Cycles

Definition. A **cycle** in an undirected graph is a path in which all edges are distinct and the only repeated vertices are the first and last vertices. By cycles of a directed graph we mean those of its underlying undirected graph.



Cycle 1 $\quad$ Cycle 2 $\quad$ Cycle 3

$$\boxed{\begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}} + \boxed{\begin{bmatrix} 0 \\ 0 \\ -1 \\ -1 \\ 1 \end{bmatrix}} = \boxed{\begin{bmatrix} -1 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}}$$

Theorem 41. *Let $\mathcal{G}$ be a directed graph and let A be its edge-node incidence matrix. Then the cycle space of $\mathcal{G}$ is equal to* $\mathrm{Nul}(A^T)$.

Orthogonal Complements

Definition. Let W be a subspace of $\mathbb{R}^n$. The **orthogonal complement** of W is the subspace $W^\perp$ of all vectors that are orthogonal to W; that is

$$W^\perp := \{\mathbf{v} \in \mathbb{R}^n \ : \ \mathbf{v} \cdot \mathbf{w} = 0 \text{ for all } \mathbf{w} \in W\}.$$

Some helpful theorems:

- $(W^\perp)^\perp = W$
- $\text{Nul}(A) = \text{Col}(A^T)^\perp$
- $\text{Nul}(A)^\perp = \text{Col}(A^T)$
- $\text{Nul}(A^T) = \text{Col}(A)^\perp$

Theorem 43. *Let V be a subspace of $\mathbb{R}^n$. Then $\dim V + \dim V^\perp = n$.*

Coordinates

Standard basis (ε):

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Generally, if $\mathbf{v}_1$, $\mathbf{v}_2$, … $\mathbf{v}_p$ are a basis B of vector space V, the coordinate vector of any vector **w** in V is:

$$\mathbf{w}_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_p \end{bmatrix}, \qquad \text{if } \mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

This coordinate vector is unique!

Coordinate Example

Let $V = \mathbb{R}^2$, and consider the bases

$$\mathcal{B} := \left(\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right)$$

$$\mathcal{E} := \left(\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$$

$\mathbf{w} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$. Determine $\mathbf{w}_{\mathcal{B}}$ and $\mathbf{w}_{\mathcal{E}}$

We want to find ‘**w**’ in terms of **B**’s and **E**’s coordinate planes

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{w}_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

This is what [3,-1] looks like in ‘basis’ **B**

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (-1) \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\mathbf{w}_{\mathcal{E}} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

This is what [3,-1] looks like in ‘basis’ **E**

Change of Basis Matrix

Definition. Let $\mathcal{B}$ and $\mathcal{C}$ be two bases of $\mathbb{R}^n$. The **change of basis matrix** $I_{\mathcal{C},\mathcal{B}}$ is the matrix such that for all $\mathbf{v} \in \mathbb{R}^n$

$$I_{\mathcal{C},\mathcal{B}}\mathbf{v}_{\mathcal{B}} = \mathbf{v}_{\mathcal{C}}$$

Matrix allowing us to go from coordinates **mapped in B** to be **mapped onto C**

Theorem 45. *Let $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b_n})$ be a basis of $\mathbb{R}^n$. Then*

$$I_{\mathcal{E}_n,\mathcal{B}} = \begin{bmatrix} \mathbf{b}_1 & \ldots & \mathbf{b}_n \end{bmatrix}$$

That is, for all $\mathbf{v} \in \mathbb{R}^n$,

$$\mathbf{v} = \begin{bmatrix} \mathbf{b}_1 & \ldots & \mathbf{b}_n \end{bmatrix} \mathbf{v}_{\mathcal{B}}.$$

How do we compute change of basis matrix:

$$I_{\mathcal{B},\mathcal{C}}?$$

What we know:

- $I_{En,B}$ = Matrix that maps **coordinates in B** onto **Standard**
- $I_{B,En}$ = Matrix that maps **coordinates in Standard** onto **B**

- $I_{En,C}$ = Matrix that maps **coordinates in C** onto **Standard**
- $I_{C,En}$ = Matrix that maps **coordinates in Standard** onto **C**

$$I_{\mathcal{B},\mathcal{E}_n} I_{\mathcal{E}_n,\mathcal{C}}$$

From right to left:
We map coordinates **from C into the standard** coordinate plane, **then**, we map the newly acquired **standard coordinates onto B's coordinate plane**

AKA: $I_{\mathcal{B},\mathcal{C}}$

Orthogonal and Orthonormal Bases

Definition. An **orthogonal basis** (an **orthonormal basis**) is an orthogonal set of vectors (an orthonormal set of vectors) that forms a basis.

Theorem 47. *Let $\mathcal{B} := (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ be an orthogonal basis of $\mathbb{R}^n$, and let $\mathbf{v} \in \mathbb{R}^n$. Then*

$$\mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{b}_1}{\mathbf{b}_1 \cdot \mathbf{b}_1}\mathbf{b}_1 + \ldots + \frac{\mathbf{v} \cdot \mathbf{b}_n}{\mathbf{b}_n \cdot \mathbf{b}_n}\mathbf{b}_n.$$

When $\mathcal{B}$ is orthonormal, then $\mathbf{b}_i \cdot \mathbf{b}_i = 1$ for $i = 1, \ldots, n$.

Theorem 48. *Let $\mathcal{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ be an orthonormal basis of $\mathbb{R}^n$. Then*

$$I_{\mathcal{U},\mathcal{E}_n} = \begin{bmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_n \end{bmatrix}^T.$$

Why? An $n \times n$-matrix Q is **orthogonal** if $Q^{-1} = Q^T$

Linear Transformation

Definition. Let V and W be vector spaces. A map $T : V \to W$ is a **linear transformation** if

$$T(a\mathbf{v} + b\mathbf{w}) = aT(\mathbf{v}) + bT(\mathbf{w})$$

for all $\mathbf{v}, \mathbf{w} \in V$ and all $a, b \in \mathbb{R}$.

$T(\mathbf{0}_V) = T(0 \cdot \mathbf{0}_V) = 0 \cdot T(\mathbf{0}_V) = \mathbf{0}_W \rightsquigarrow T(\mathbf{0}_V) = \mathbf{0}_W$

To check linearity for a transformation, we can test with 0, since when we multiply anything by 0, we get 0 back in both spaces

Theorem 50. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there is a $m \times n$ matrix A such that*

- $T(\mathbf{v}) = A\mathbf{v}, \quad$ *for all* $\mathbf{v} \in \mathbb{R}^n$.
- $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \ldots & T(\mathbf{e}_n) \end{bmatrix}$, *where* $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ *is the standard basis of* $\mathbb{R}^n$.

Remark. We call this A the **coordinate matrix of** T with respect to the standard bases - we write $T_{\mathcal{E}_m, \mathcal{E}_n}$.

Coordinate matrices

Theorem 51. *Let V, W be two vector space, let $\mathcal{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ be a basis of V and $\mathcal{C} = (\mathbf{c}_1, \dots, \mathbf{c}_m)$ be a basis of W, and let $T\colon V \to W$ be a linear transformation. Then there is a $m \times n$ matrix $T_{\mathcal{C},\mathcal{B}}$ such that*

- $T(\mathbf{v})_{\mathcal{C}} = T_{\mathcal{C},\mathcal{B}}\mathbf{v}_{\mathcal{B}}$, *for all* $\mathbf{v} \in V$.
- $T_{\mathcal{C},\mathcal{B}} = \begin{bmatrix} T(\mathbf{b}_1)_{\mathcal{C}} & T(\mathbf{b}_2)_{\mathcal{C}} & \dots & T(\mathbf{b}_n)_{\mathcal{C}} \end{bmatrix}$.

$$\begin{tikzcd}
\mathbf{v} : \text{vector in } V \arrow[r, "\text{apply } T"] \arrow[d, "\text{write in coordinates wrt } \mathcal{B}"'] & \text{vector in } W : T(\mathbf{v}) \arrow[d, "\text{write in coordinates wrt } \mathcal{C}"] \\
\mathbf{v}_{\mathcal{B}} : \text{coordinate vector in } \mathbb{R}^n \arrow[r, "\text{multiply by } T_{\mathcal{C},\mathcal{B}}"] & \text{coordinate vector in } \mathbb{R}^m : T_{\mathcal{C},\mathcal{B}}\mathbf{v}
\end{tikzcd}$$

Determinants (how to find them)

2x2: easy formula!

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$$

Triangular: multiply all of the diagonal entries together

Otherwise: cofactor expansion

Note: if the matrix A is not invertible, det(A) = 0 ⟵ this is the definition of a determinant!

Cofactor Expansion Example

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix} = 2\cdot(-1)^{1+2}\cdot\begin{vmatrix} & - & \\ 3 & & 2 \\ 2 & & 1 \end{vmatrix} + (-1)\cdot(-1)^{2+2}\cdot\begin{vmatrix} 1 & & 0 \\ & + & \\ 2 & & 1 \end{vmatrix} + 0\cdot(-1)^{3+2}\cdot\begin{vmatrix} 1 & & 0 \\ 3 & & 2 \\ & - & \end{vmatrix}$$

$$= -2\cdot(-1) + (-1)\cdot 1 - 0 = 1$$

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix} = 0\cdot(-1)^{1+3}\cdot\begin{vmatrix} & & + \\ 3 & -1 & \\ 2 & 0 & \end{vmatrix} + 2\cdot(-1)^{2+3}\cdot\begin{vmatrix} 1 & 2 & \\ & & - \\ 2 & 0 & \end{vmatrix} + 1\cdot(-1)^{3+3}\cdot\begin{vmatrix} 1 & 2 & \\ 3 & -1 & \\ & & + \end{vmatrix}$$

$$= 0 - 2\cdot(-4) + 1\cdot(-7) = 1$$

Properties of determinants

(Replacement) Adding a multiple of one row to another row *does not change* the determinant.
(Interchange) Interchanging two different rows *reverses the sign* of the determinant.
(Scaling) Multiplying all entries in a row by s, *multiplies* the determinant by s.

These three things also apply to the columns of a matrix!

Let A, B be two $n \times n$-matrices. Then $\det(AB) = \det(A)\det(B)$

If A is invertible, then $\det(A^{-1}) = \frac{1}{\det(A)}$

Let A be an $n \times n$-matrix. Then $\det(A^T) = \det(A)$

Eigenvectors and Eigenvalues

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

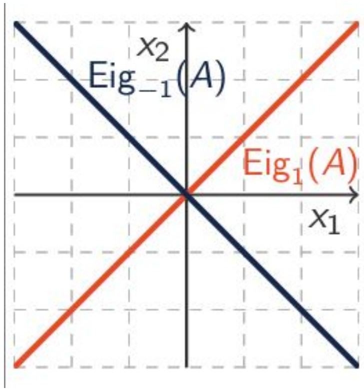
An **eigenvector** of A is a nonzero $\mathbf{v} \in \mathbb{R}^n$

such that

$$A\mathbf{v} = \lambda \mathbf{v}$$

eigenvalue

An eigenspace is all the eigenvectors associated with a specific eigenvalue.



$$A \begin{bmatrix} x \\ x \end{bmatrix} = 1 \cdot \begin{bmatrix} x \\ x \end{bmatrix}$$

$$A \begin{bmatrix} -x \\ x \end{bmatrix} = -1 \cdot \begin{bmatrix} -x \\ x \end{bmatrix}$$

Eigenvectors are always linearly independent!

Calculating eigenvectors and eigenvalues

Theorem 59. *Let A be an $n \times n$ matrix. Then $p_A(t) := \det(A - tI)$ is a polynomial of degree n. Thus A has at most n eigenvalues.*

Definition. We call $p_A(t)$ the **characteristic polynomial** of A.

The roots of the characteristic polynomial are the eigenvalues

Let A be $n \times n$ matrix and let λ be eigenvalue of A. Then

$$\mathrm{Eig}_\lambda(A) = \mathrm{Nul}(A - \lambda I).$$

General algorithm: 1) find det(A-λI) and solve for λ
2) plug each eigenvalue back into A-λI
3) solve for the nullspace

Properties of Eigenvalues and Eigenvectors

For a 2x2 matrix:

$$p(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A)$$

Multiplicity:

- **Algebraic** multiplicity is the multiplicity of λ in the characteristic polynomial
- **Geometric** multiplicity is the dimension of the eigenspace of λ

Trace: the sum of the diagonal entries of a matrix

- Tr(A) = sum of all eigenvalues
- det(A) = product of all eigenvalues

Markov Matrices

$$\begin{bmatrix} 0 & .25 & .4 \\ 1 & .25 & .2 \\ 0 & .5 & .4 \end{bmatrix}$$

Definition: a square matrix with non-negative entries where the sum of terms in each column is 1

A **probability vector** has entries that add up to 1

The λ of a Markov Matrix:

- 1 is always an eigenvalue, and the corresponding eigenvector is called **stationary**
- All other $|\lambda| \leq 1$

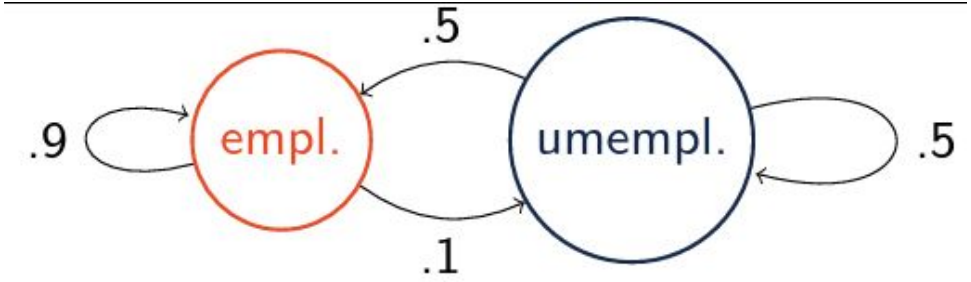
Why is a Markov Matrix useful?

Theorem 65. *Let A be an $n \times n$-Markov matrix with only positive entries and let $\mathbf{z} \in \mathbb{R}^n$ be a probability vector. Then*

$$\mathbf{z}_\infty := \lim_{k\to\infty} A^k\mathbf{z} \text{ exists,}$$

and $\mathbf{z}_\infty$ is a stationary probability vector of A (ie. $A\mathbf{z}_\infty = \mathbf{z}_\infty$).

This basically says you can left multiply A with **z** infinitely and you will get a stationary probability vector (steady state)



x_t: % of population employed at time t

y_t: % of population unemployed at time t

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = \begin{bmatrix} .9x_t + .5y_t \\ .1x_t + .5y_t \end{bmatrix} = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix}$$

Diagonalization

$$P = \begin{bmatrix} \mathbf{v_1} & \dots & \mathbf{v_n} \end{bmatrix}$$

v are eigenvectors

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

For a matrix A to be diagonalizable:

- A must be square
- A must have as many unique eigenvectors as rows/columns (i.e. it has an eigenbasis)
- A = PDP^{-1}

Observe that

$$A = PDP^{-1} = I_{\mathcal{E}_n,\mathcal{B}} D I_{\mathcal{B},\mathcal{E}_n}$$

Where B is the eigenbasis → diagonalizing is a base change to the eigenbasis

Matrix Powers and Matrix Exponential

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

Matrix power: diagonal matrices are easy!

$$A^m = PD^m P^{-1}$$

Where D^m =

$$\begin{bmatrix} (\lambda_1)^m & & \\ & \ddots & \\ & & (\lambda_n)^m \end{bmatrix}$$

Matrix exponential:

$$e^{At} = I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \ldots$$

$$e^{At} = Pe^{Dt} P^{-1}$$

Linear Differential Equations

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u}$$

With initial condition:

$$\mathbf{u}(0) = \mathbf{v}$$

Let A be an $n \times n$ matrix and $\mathbf{v} \in \mathbb{R}^n$ The solution of the differential equation $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ with initial condition $\mathbf{u}(0) = \mathbf{v}$ is $\mathbf{u}(t) = e^{At}\mathbf{v}$

If $v_1, v_2, \ldots v_n$ is an eigenbasis of A:

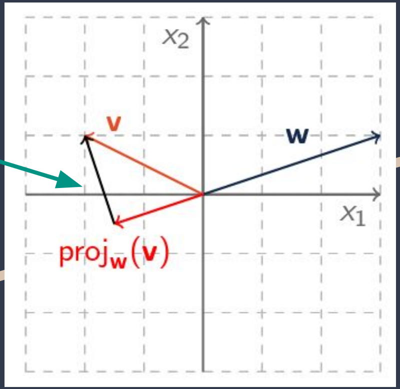
$$e^{At}\mathbf{v} = c_1 e^{\lambda_1 t}\mathbf{v}_1 + \cdots + c_n e^{\lambda_n t}\mathbf{v}_n$$

Vector Projections

Projection of **v** onto **w**

$$\mathrm{proj}_{\mathbf{w}}(\mathbf{v}) := \frac{\mathbf{w} \cdot \mathbf{v}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}$$

Error term



Projecting **v** onto **w** yields the vector in span(**w**) that is closest to **v**.

The **error term** is **v** - $\mathrm{proj}_{\mathbf{w}}(\mathbf{v})$ and is in $\mathrm{span}(\mathbf{w})^{\perp}$

Can also use:

$$\mathrm{proj}_{\mathbf{w}}(\mathbf{v}) = \boxed{\left(\frac{1}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}\mathbf{w}^{T}\right)} \mathbf{v}$$

Where the boxed term is called the orthogonal projection matrix onto span(**w**)

Subspace Projections

Let W be a subspace of $\mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^n$. Then $\mathbf{v}$ can be written uniquely as

$$\mathbf{v} = \underbrace{\hat{\mathbf{v}}}_{\text{in } W} + \underbrace{\mathbf{v}^{\perp}}_{\text{in } W^{\perp}}$$

$\hat{\mathbf{v}}$ is calculated by projecting **v** onto an orthogonal basis of W

P_W is the orthogonal projection matrix for subspace W. Calculate P_W by projecting each column of the identity matrix onto W and join them all in a matrix

$$Q = I - P_W, \text{ where } I \text{ is the identity. Then } P_{W^{\perp}} = Q$$

Least Squares Solutions:

Trying to minimize the distance between Ax and b for an inconsistent system

$$A\hat{\mathbf{x}} = \text{proj}_{\text{Col}(A)}(\mathbf{b})$$

LSQ solution

General algorithm:

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$$

Find A^T and A^TA, then solve the above system with any method you prefer.

For linear regressions:

$$\overset{A}{\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix}}_{\text{design matrix } X}} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}}_{\hat{\mathbf{x}}} = \overset{\mathbf{b}}{\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}}_{\text{observation vector } \mathbf{y}}}$$

The shape of the design matrix depends on the problem!

Gram–Schmidt Method

Algorithm. (Gram-Schmidt orthonormalization) Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_m$, produce an orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_m$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_m$.

$$\mathbf{b}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|}$$

$$\mathbf{b}_2 = \mathbf{a}_2 - \underbrace{\operatorname{proj}_{\operatorname{span}(\mathbf{q}_1)}(\mathbf{a}_2)}_{=(\mathbf{a}_2 \cdot \mathbf{q}_1)\mathbf{q}_1}, \qquad \mathbf{q}_2 = \frac{\mathbf{b}_2}{\|\mathbf{b}_2\|}$$

$$\mathbf{b}_3 = \mathbf{a}_3 - \underbrace{\operatorname{proj}_{\operatorname{span}(\mathbf{q}_1, \mathbf{q}_2)}(\mathbf{a}_3)}_{(\mathbf{a}_3 \cdot \mathbf{q}_1)\mathbf{q}_1 + (\mathbf{a}_3 \cdot \mathbf{q}_2)\mathbf{q}_2} \qquad \mathbf{q}_3 = \frac{\mathbf{b}_3}{\|\mathbf{b}_3\|}$$

- QR Decomposition: Let A be an m × n matrix of rank n.
  - There is an m × n-matrix Q with orthonormal columns
  - An upper triangular n × n invertible matrix R such that A = QR.

Spectral Theorem

Theorem 84. *Let A be a symmetric $n \times n$ matrix. Then A has an orthonormal basis of eigenvectors.*

Orthonormal basis: all vectors are orthogonal (perpendicular) to each other and dot product of themselves = 1

Theorem 85. *Let A be a symmetric $n \times n$ matrix. Then there is a diagonal matrix D and a matrix Q with orthonormal columns such that $A = QDQ^T$.*

$$Q^{-1} = Q^T$$

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

$\lambda_1 = 2$: Eigevector $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Normalized, we get $\mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$\lambda_2 = 4$: Eigenvector $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Normalized, we get $\mathbf{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$D = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$ and $Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

Singular Value Decomposition

Definition. Let A be an $m \times n$ matrix. A **singular value decomposition** of A is a decomposition $A = U\Sigma V^T$ where

- U is an $m \times m$ matrix with orthonormal columns,
- Σ is an $m \times n$ rectangular diagonal matrix with non-negative numbers on the diagonal,
- V is an $n \times n$ matrix with orthonormal columns.

Remark. The diagonal entries $\sigma_i = \Sigma_{ii}$ which are positive are called the singular values of A. We usually arrange them in decreasing order, that is $\sigma_1 \geq \sigma_2 \geq \dots$

SVD Algorithm

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

$$A^TA = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Eigenvalues: $\lambda_1 = 3, \lambda_2 = 1, \lambda_3 = 0$

Eigenbasis: $\underbrace{\begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}}_{:=\mathbf{v}_1}, \underbrace{\begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}}_{:=\mathbf{v}_2}, \underbrace{\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}}_{:=\mathbf{v}_3}$

Singular values: $\sigma_1 := \sqrt{3}, \sigma_2 = 1$

$$V = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}, \Sigma = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{u}_1 := \frac{1}{\sigma_1} A\mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} -\frac{3}{\sqrt{6}} \\ \frac{3}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\rightsquigarrow U = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Let A be an mxn matrix with rank r

1. Find orthonormal eigenbasis $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ of A^TA with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = 0 = \cdots = \lambda_n$.
2. Set $\sigma_i = \sqrt{\lambda_i}$ for $i = 1, \ldots, n$.
3. Set $\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1, \ldots, \mathbf{u}_r = \frac{1}{\sigma_r} A\mathbf{v}_r$
4. Find $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m \in \mathbb{R}^m$ such that $(\mathbf{u}_1, \ldots, \mathbf{u}_m)$ is an orthonormal basis of $\mathbb{R}^m$
5. $$U = \begin{bmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_m \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{\min\{m,n\}} \end{bmatrix}, V = \begin{bmatrix} \mathbf{v}_1 & \ldots & \mathbf{v}_n \end{bmatrix}$$

Theorem 86. *Let A be an $m \times n$ matrix with rank r, and let $U = \begin{bmatrix} \mathbf{u}_1 & \ldots \mathbf{u}_m \end{bmatrix}$, $V = \begin{bmatrix} \mathbf{v}_1 & \ldots \mathbf{v}_n \end{bmatrix}$, Σ be such that $A = U\Sigma V^T$ is a SVD of A. Then*

- $(\mathbf{u}_1, \ldots, \mathbf{u}_r)$ *is a basis of* $\mathrm{Col}(A)$.
- $(\mathbf{v}_1, \ldots, \mathbf{v}_r)$ *is a basis of* $\mathrm{Col}(A^T)$.
- $(\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m)$ *is a basis of* $\mathrm{Nul}(A^T)$.
- $(\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n)$ *is a basis of* $\mathrm{Nul}(A)$.

Low rank approximation via SVD

Theorem 87. *Let A be an $m \times n$ matrix with rank r, and let $U = \begin{bmatrix} \mathbf{u}_1 & \dots \mathbf{u}_m \end{bmatrix}$, $V = \begin{bmatrix} \mathbf{v}_1 & \dots \mathbf{v}_n \end{bmatrix}$ be matrices with with orthonormal columns and Σ be a rectangular diagonal $m \times n$ matrix such that $A = U\Sigma V^T$ is an SVD of A. Then*

$$\begin{aligned} A &= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T \\ &= \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_r \end{bmatrix}^T \end{aligned}$$

Compact SVD

For $k \leq r$, define

$$A_k = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \cdots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T$$

If $\sigma_1 \gg \sigma_2 \gg \ldots\ldots,$ Then, A_k is a good approx. since most of the information is inside of first term

Definition. Let A be an $m \times n$ matrix with rank r. A **compact singular value decomposition** of A is a decomposition $A = U_c \Sigma_c V_c^T$ where

- $U_c = \begin{bmatrix} \mathbf{u}_1 & \ldots \mathbf{u}_r \end{bmatrix}$ is an $m \times r$ matrix with orthonormal columns,
- Σ_c is an $r \times r$ diagonal matrix with positive diagonal elements,
- $V_c = \begin{bmatrix} \mathbf{v}_1 & \ldots \mathbf{v}_r \end{bmatrix}$ is an $n \times r$ matrix with orthonormal columns.

rank = 2

$$\text{SVD} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

Pseudo Inverse

Definition. Let A be an $m \times n$ matrix with rank r. Given the compact singular value decomposition $A = U_c \Sigma_c V_c^T$ where

- $U_c = \begin{bmatrix} \mathbf{u}_1 & \dots \mathbf{u}_r \end{bmatrix}$ is an $m \times r$ matrix with orthonormal columns,
- Σ_c is an $r \times r$ diagonal matrix with positive diagonal elements,
- $V_c = \begin{bmatrix} \mathbf{v}_1 & \dots \mathbf{v}_r \end{bmatrix}$ is an $n \times r$ matrix with orthonormal columns,

pseudoinverse A^+ of A as $V_c \Sigma_c^{-1} U_c^T$.

Theorem 88. *Let* $\mathbf{v} \in \mathrm{Col}(A^T)$ *and* $\mathbf{w} \in \mathrm{Col}(A)$. *Then* $A^+ A\mathbf{v} = \mathbf{v}$ *and* $AA^+\mathbf{w} = \mathbf{w}$

Theorem 89. *Let A be an $m \times n$ matrix and let $\mathbf{b} \in \mathbb{R}^m$. Then $A^+\mathbf{b}$ is the LSQ solution of $A\mathbf{x} = \mathbf{b}$ (with minimum length).*

PCA: Principal Component Analysis

Setup:

- Given **m** objects, measure same **n** variables
- **m** samples, of **n**-dimensional data → mxn matrix → X

Let $X = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}^T$ be an $m \times n$ matrix. We define the **column average** $\mu(X)$ of X as

$$\mu(X) := \frac{1}{m}(\mathbf{a}_1 + \dots + \mathbf{a}_m)$$

We say X is **centered** if $\mu(X) = \mathbf{0}$ **covariance matrix** $\mathrm{cov}(X)$ of X as $\frac{1}{m-1} X^T X$

**:

- ➔ Not centered, replace X by $\begin{bmatrix} \mathbf{a}_1 - \mu(X) & \dots & \mathbf{a}_m - \mu(X) \end{bmatrix}^T$.
- ➔ If the columns of X are orthogonal, then $\mathrm{cov}(X)$ is a diagonal matrix $\rightsquigarrow$ each variable is independent.

PCA Process

- Input: *centered* mxn matrix, X
- Find cov(X)
- Find orthonormal eigenbasis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of $\mathrm{cov}(X)$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$

- Write $\mathrm{cov}(X)$ as a sum of rank 1 matrices:

$$\mathrm{cov}(X) = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T + \cdots + \lambda_n \mathbf{v}_n \mathbf{v}_n^T$$

- The principal component v_i explains part of the variance of the data. The larger λ_i , the more of the variances is explained by v_i

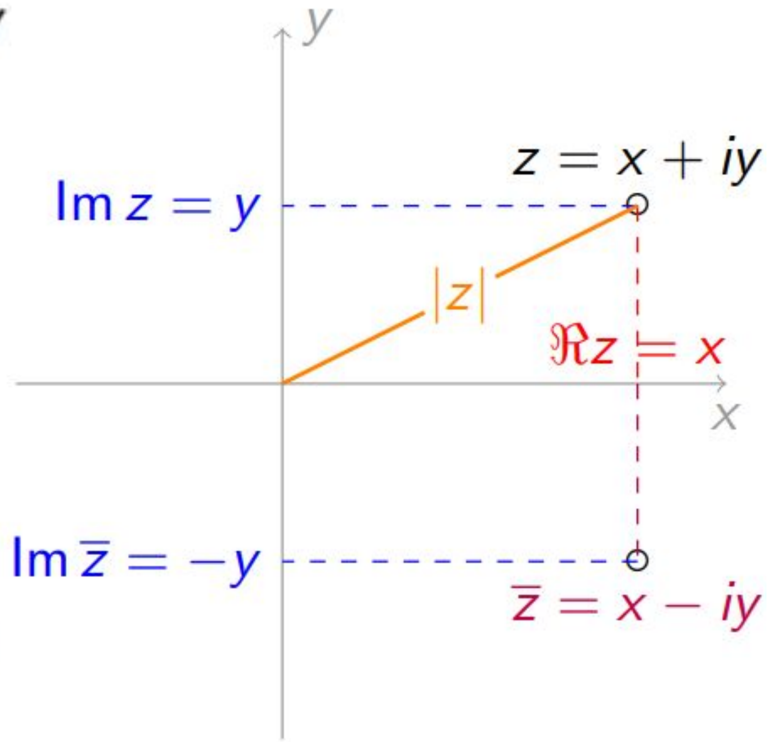
Complex Numbers

$i = \sqrt{-1}$, or $i^2 = -1$

$$z = x + iy$$

- Real Part: **x**
- Imaginary Part: **y**
- Complex Conjugate: $\overline{z}$ **= x - iy**
- Magnitude:

$$|z| = \sqrt{x^2 + y^2}$$



Any point in $\mathbb{R}^2$ can be viewed as a complex number: $\binom{x}{y} \leftrightarrow x + iy$

Complex Numbers cont.

Theorem 90. *Let* $z \in \mathbb{C}$.

- $\bar{\bar{z}} = z$ $\quad (x - iy) \to (x + iy)$
- $|z|^2 = z\bar{z}$ $\quad \sqrt{x^2 + y^2}^2 = (x + iy)(x - iy)$
- $|z| = |\bar{z}|$ $\quad \sqrt{x^2 + y^2} = \sqrt{x^2 + (-y)^2}$

Find:

- Complex eigenvectors/values
- Conjugate Matrix
- Eigen Basis

$$\textit{complex} \text{ column vectors } \mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \det(A - \lambda I) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1$$

Tips on Approaching Conceptual Questions

What topic is the question asking about?

Not always immediately apparent. Think about the relevant theorems: is it a vector or a matrix? Is the matrix invertible? Is the matrix orthonormal?

For true/false questions:

Look for counter examples. Try easy test cases like $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ to prove false

If you think the statement is true, try to connect theorems by breaking down what each part of the statement means

Conceptual Question Example 1

17. (5 points) Let A be an $n \times n$ matrix. Consider the following statements:

(T1) If A is not the zero matrix, then A^2 is also not the zero matrix.

(T2) If A is invertible, then A^2 is also invertible.

Which of the statements are ALWAYS TRUE?

(A) Neither Statement (T1) nor Statement (T2).

(B) Both Statement (T1) and Statement (T2).

(C) Only Statement (T2).

(D) Only Statement (T1).

Topic: matrix multiplication and matrix inverses

T1: FALSE. Look for a counterexample – pick something 2x2 with a lot of zeros, such as $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

T2: TRUE. To be invertible, a matrix has to have a non-zero determinant

- Recall properties of determinants: $\det(A^2) = \det(A)^2$
- If $\det(A)$ is non-zero, so is $\det(A^2)$

Conceptual Question Example 2

18. (5 points) Let A be a diagonalizable 3×3 matrix with only two distinct eigenvalues. Which of the following statements is FALSE?

(A) The matrix $5A$ is diagonalizable.

(B) The matrix A has an eigenbasis.

(C) There are no more than two linearly independent eigenvectors of A.

(D) There is an eigenvalue of A for which the corresponding eigenspace is spanned by two linearly independent eigenvectors.

Topic: diagonalization, eigenvectors

A: TRUE. Look at the definition of an eigenvalue: A**v** = λ**v** – multiplying A by 5 does not change the vector itself so 5A will also be diagonalizable

B: TRUE. This is part of the definition of diagonalizable matrices

C: FALSE. If A has an eigenbasis, it must have **three** linearly independent eigenvectors

D: TRUE. If there are only two distinct eigenvalues but there is an eigenbasis, one of the eigenspaces must have a span of 2 eigenvectors

Professor's Tips (From Spring 2024)

- Realize how important REF & RREF are
  - Help to solve linear systems
  - Number of pivots can help you solve for the dimension of all fundamental subspaces
  - If RREF([A]) = [I], [A] is invertible
  - REF serves as a test for linear independence and solves for the basis of the column space
- Apply your understanding of the fundamental subspaces wherever possible
  - Column space is the possible solution set of a linear system
  - Null space is the 0-eigenspace
  - Rank-nullity is important for conceptual questions
  - All decomposition, span, and transformation questions are fundamental subspace questions
  - Know orthogonal complements

Professor's Tips (From Spring 2024)

- Abstract vector space questions may seem intimidating, but each vector can be written as a column vector and handled normally
  - Matrices, polynomials, etc.
- There are an infinite number of bases for a subspace. How we express a vector is up to us
  - There is nothing special about standard basis except for being convenient
- There is an underlying geometry to all operations in linear algebra
  - Can be used to understand coordinate transforms, determinants, bases, subspaces, decompositions, and more
- Conceptual questions can be solved with applying basic formulas & theorems and inspecting the consequences of them

Questions?

Good luck on your final exam!

No new problems

Join the queue to see the past worksheets and solutions!