Part 1:

Introduction/Linear Systems

Linear Equation:

$$A\mathbf{x} = \mathbf{b} \Rightarrow a_1x_1 + a_2x_2 + ... + a_nx_n = b$$

-"a" is a constant -"x" is a variable

Linear System: a set of Linear Equations

Three Possibilities:

- Unique Solution
- Infinitely Many Solutions
- No Solution

Row Reduction

Elementary Row Operations

1. Interchange Rows
2. Multiply a row by a constant
3. Add a multiple of a row to another

Gaussian Elimination: Use elementary row operations on a matrix to reduce to "row echelon form"

Gauss-Jordan Elimination: Same as Gaussian, but reduce to "reduced row echelon form"

Steps:

1. Write System as Augmented Matrix
2. Reduce using Elementary Row Operations
3. Solve Using Back Substitution or in terms of other variables

Example

$$2x_2 + 3x_3 = 8$$
$$2x_1 + 3x_2 + x_3 = 5$$
$$x_1 - x_2 - 2x_3 = -5$$

$$\begin{pmatrix} 0 & 2 & 3 \\ 2 & 3 & 1 \\ 1 & -1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 8 \\ 5 \\ -5 \end{pmatrix}$$

$$\left(\begin{matrix} 0 & 2 & 3 \\ 2 & 3 & 1 \\ 1 & -1 & -2 \end{matrix}\left(\begin{matrix} 8 \\ 5 \\ -5 \end{matrix}\right)\right) \xrightarrow{\text{Interchange R1 and R3}}$$

$$\left(\begin{matrix} 1 & -1 & -2 \\ 2 & 3 & 1 \\ 0 & 2 & 3 \end{matrix}\left(\begin{matrix} -5 \\ 5 \\ 3 \end{matrix}\right)\right) \xrightarrow{\text{Subtract 2R1 from R2}}$$

$$\left(\begin{matrix} 1 & -1 & -2 \\ 0 & 5 & 5 \\ 0 & 2 & 3 \end{matrix}\left(\begin{matrix} -5 \\ 15 \\ 3 \end{matrix}\right)\right) \xrightarrow{\text{Multiply R2 by constant } \frac{1}{5}}$$

$$\left(\begin{matrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 2 & 3 \end{matrix}\left(\begin{matrix} -5 \\ 3 \\ 3 \end{matrix}\right)\right) \xrightarrow{\text{Subtract R2 from R3}}$$

$$\left(\begin{matrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{matrix}\left(\begin{matrix} -5 \\ 3 \\ 2 \end{matrix}\right)\right) \xrightarrow{\text{R2: R2-R3}}$$

This is "row echelon form."

(**Gaussian:** Use back substitution from here)

$$\left(\begin{matrix} 1 & -1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}\left(\begin{matrix} -5 \\ 1 \\ 2 \end{matrix}\right)\right) \xrightarrow{\text{R1: R1+R2+2R3}}$$

$$\left(\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}\left(\begin{matrix} 0 \\ 1 \\ 2 \end{matrix}\right)\right)$$

This is "reduced row echelon form."

(**Gauss-Jordan:** Use back substitution from here)

$$x_1 = 0, x_2 = 1, x_3 = 2 \text{ -Or- } \mathbf{x} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

Vectors, Linear Combinations

Vector: $\mathbf{v} = \langle v_1, v_2, \cdots, v_n \rangle$

Linear Combination: Vector, **v,** is a linear combination of vectors: $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ if

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n$$

-"c" is a scalar and is called a coefficient

Linear Independence: A Set of vectors:

$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is *linearly dependent* if

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n = 0$$

For some set of nontrivial coefficients

$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is *linearly independent* if only the trivial solution exists.

Matrix Operations

Addition/Subtraction:

$$A = [a_{ij}]_{m\times n} \qquad B = [b_{ij}]_{m\times n}$$

$$A \pm B = \left[a_{ij} \pm b_{ij}\right]$$

(A and B must be the same size)

Multiplication:

$$A = [a_{ij}]_{m\times n} \qquad B = [b_{ij}]_{n\times r}$$

$$C = AB = [c_{ij}]_{m\times r}$$

Where,

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$$

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$ (This is useful for proofs)

(A must have k columns, B must have k rows)

Transpose: The transpose of a matrix A is obtained by interchanging its rows and columns

For $A = [a_{ij}]_{m\times n} : A^T = [a_{ji}]_{n\times m}$

Example

$A=\begin{pmatrix}2&0\\3&1\end{pmatrix}$ $A^T=\begin{pmatrix}2&3\\0&1\end{pmatrix}$

$B=\begin{pmatrix}1&3\\2&5\end{pmatrix}$

$$A+B=\begin{pmatrix}2+1&0+1\\3+2&1+5\end{pmatrix}=\begin{pmatrix}3&1\\5&6\end{pmatrix}$$

$$AB=\begin{pmatrix}2&0\\3&1\end{pmatrix}\begin{pmatrix}1&3\\2&5\end{pmatrix}=\begin{pmatrix}2\cdot1+0\cdot2&2\cdot3+0\cdot5\\3\cdot1+1\cdot2&3\cdot3+1\cdot5\end{pmatrix}$$

$$AB=\begin{pmatrix}2&6\\5&14\end{pmatrix}$$

Inverse of a Matrix

Inverse: The inverse of matrix A is some matrix: A^{-1} such that:

$$AA^{-1}=I$$

$$A^{-1}A=I$$

Invertible: A matrix is invertible if it has an inverse

Rule for 2×2 matrix:

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

Gauss-Jordan Method for Computing the Inverse:

$$[A\,|\,I]\rightarrow[I\,|\,A^{-1}]$$

Example

$A=\begin{pmatrix}2&2\\2&0\end{pmatrix} over \mathbb{Z}_3$

$A^{-1}:[A\,|\,I]=$

$$\left(\begin{array}{cc|cc}2&2&1&0\\2&0&0&1\end{array}\right)\xrightarrow{\text{2R1}}$$

$$\left(\begin{array}{cc|cc}1&1&2&0\\2&0&0&1\end{array}\right)\xrightarrow{\text{R2:R2+R1}}$$

$$\left(\begin{array}{cc|cc}1&1&2&0\\0&1&2&1\end{array}\right)\xrightarrow{\text{R1:R1+2R2}}$$

$$\left(\begin{array}{cc|cc}1&0&0&2\\0&1&2&1\end{array}\right)=[I\,|\,A^{-1}]$$

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

P^{T}LU Factorization

Matrix A can be written as

$$A=P^TLU$$

where L is a unit lower triangular matrix and
U is a an upper triangular matrix
P is a permutation matrix (for row interchanges)

$$A=P^T\begin{pmatrix}1&0&\cdots&0\\ *&1&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\ *&*&\cdots&1\end{pmatrix}\begin{pmatrix}*&*&\cdots&*\\0&*&\cdots&*\\ \vdots&\vdots&\ddots&\vdots\\0&0&\cdots&*\end{pmatrix}$$

Example

$$A=\begin{pmatrix}0&0&6\\1&2&3\\2&1&4\end{pmatrix} R_1\leftrightarrow R_2$$

$$\begin{pmatrix}1&2&3\\0&0&6\\2&1&4\end{pmatrix} R_3-2R_1$$

$$\begin{pmatrix}1&2&3\\0&0&6\\0&-3&-2\end{pmatrix} R_2\leftrightarrow R_3$$

$$\begin{pmatrix}1&2&3\\0&-3&-2\\0&0&6\end{pmatrix}=U$$

$$P=P_2P_1=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}$$

$$A=P^TLU=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}\begin{pmatrix}1&0&0\\2&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&2&3\\0&-3&-2\\0&0&6\end{pmatrix}$$

Part 2:

Determinants

Determinant: A function that assigns a scalar det(A) to an $n \times n$ matrix

Rules for 2×2 *and* 3×3 *matrices*

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = \begin{vmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{vmatrix} =$$

$$aei + bfg + cdh - bdi - afh - ceg$$

Laplace Expansion Theorem:

Cofactor Expansion:

Along the i-th row:

$$\det A = a_{i1} \det A_{i1} - \cdots + (-1)^{1+n} a_{in} \det A_{in}$$

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{ij} \det A_{ij}$$

Along the j-th column:

$$\det A = a_{1j} \det A_{1j} - \cdots + (-1)^{1+n} a_{nj} \det A_{nj}$$

$$\det A = \sum_{i=1}^{n} (-1)^{1+i} a_{ij} \det A_{ij}$$

Example:

$$\begin{vmatrix} 5 & -3 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \end{vmatrix} = 2\begin{vmatrix} -3 & 2 \\ 0 & 2 \end{vmatrix} - (-1)\begin{vmatrix} 5 & 2 \\ 1 & 2 \end{vmatrix} + 3\begin{vmatrix} 5 & -3 \\ 1 & 0 \end{vmatrix} =$$

$= 2(-6) + 8 + 3(3) = 5$ (expansion along 3rd row)

Properties:

a. $\det(AB) = (\det A)(\det B)$

b. $\det(A^{-1}) = (\dfrac{1}{\det A})$

c. $\det(A^{T}) = \det A$

Determinant of a matrix B:

Due to Elementary Row Operations on matrix A

1. Interchange two rows (or columns) of A;

$$\det B = -\det A$$

2. Multiply a row in A by a scalar, k;

$$\det B = k \det A$$

3. Add a multiple of one row of A to another;

$$\det B = \det A$$

Cramer's Rule: For a given linear system: $A\mathbf{x} = \mathbf{b}$. The unique solution can be given by:

$$x_i = \frac{\det(A_i(\mathbf{b}))}{\det A} \text{ for } i = 1, \ldots, n$$

Where $A_i(\mathbf{b})$ is the i-th row of A replaced by $\mathbf{b}$

Proof of Cramer's Rule:

$$AI_i(\mathbf{x}) = A\left[\mathbf{e}_1 \cdots \mathbf{x} \cdots \mathbf{e}_n\right] = \left[A\mathbf{e}_1 \cdots A\mathbf{x} \cdots A\mathbf{e}_n\right] =$$

$$\left[A\mathbf{e}_1 \cdots \mathbf{b} \cdots A\mathbf{e}_n\right] = \left[\mathbf{a}_1 \cdots \mathbf{b} \cdots \mathbf{a}_n\right] = A_i(\mathbf{b})$$

$$(\det A)(\det I_i(\mathbf{x})) = \det(AI_i(\mathbf{x})) = \det(A_i(\mathbf{b}))$$

$$\det I_i(x) = \begin{vmatrix} 1 & 0 & \cdots & x_1 & \cdots & \cdots & 0 & 0 \\ 0 & 1 & \cdots & x_2 & \cdots & \cdots & 0 & 0 \\ \vdots & & \ddots & \vdots & & & \vdots & \vdots \\ 0 & 0 & \cdots & x_i & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & & \ddots & 0 & 0 \\ 0 & 0 & \cdots & x_{n-1} & \cdots & \cdots & 1 & 0 \\ 0 & 0 & \cdots & x_n & \cdots & \cdots & 0 & 1 \end{vmatrix} = x_i$$

$$(\det A)x_i = \det(A_i(\mathbf{b})) \Rightarrow x_i = \frac{\det(A_i(\mathbf{b}))}{\det A}$$

Eigenvectors

Eigenvector: The eigenvector, **x,** of a matrix A is any vector such that

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

Eigenvalue: λ , it is a constant

$$A\mathbf{x} = \lambda\mathbf{x} = \lambda(I\mathbf{x}) \Rightarrow (A - I\lambda)\mathbf{x} = \mathbf{0}$$

Use $\det(A - I\lambda) = 0$ to create a *characteristic equation* to find the Eigenvalues.

Example:

$$\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \Rightarrow \begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 1 = 0$$

This is the characteristic equation: $\lambda^2 - 6\lambda + 8 = 0$

These are the eigenvalues: $\lambda = 2, 4$

- $$\left[A - 4I \mid \mathbf{0}\right] = \left(\begin{array}{cc|c} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array}\right) \rightarrow \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

  $x_1 - x_2 = 0 \Rightarrow x_1 = x_2$

  $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ This is the corresponding eigenvector

  for $\lambda = 4$

- $$\left[A - 2I \mid \mathbf{0}\right] = \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 1 & 0 \end{array}\right) \rightarrow \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

  $x_1 + x_2 = 0 \Rightarrow x_1 = -x_2$

$\mathbf{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ This is the corresponding eigenvector for $\lambda = 2$

Diagonalization

Similarity: Two square matrices of the same size are said to be similar if for some matrix, *P*, $P^{-1}AP = B$ then $A \sim B$

Note an equivalent equation for similarity is:

$$AP = PB$$

Properties:

For all $A \sim B$,

a. $\det A = \det B$
b. *A* is invertible iff *B* is
c. *A* and *B* have the same rank
d. *A* and *B* have the same characteristic polynomial
e. *A* and *B* have the same eigenvalues

Proof of *a.*

By Definition, $P^{-1}AP = B$.

Taking determinants: $\det P^{-1}AP = \det B$

$$\left(\frac{1}{\det P}\right)(\det A)(\det P) = \det B \Rightarrow \det A = \det B$$

Example:

$A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ -2 & -1 \end{pmatrix}$

Then $A \sim B$, since if $P = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$

$AP = PB \Rightarrow$

$$\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -2 & -1 \end{pmatrix}$$

Diagonalizable: If a matrix *A* is similar to a diagonal matrix *D*, then *A* is diagonalizable.

$$P^{-1}AP = D$$

-Let A be an $n \times n$ matrix. A is diagonalizable iff A has n linearly independent eigenvectors.
-The columns of P are the eigenvectors of A.
-The diagonal entries of D are the eigenvalues of A (in the order that their corresponding eigenvectors appear in P)

Example:

For $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ from Example #,

$\lambda_1 = 4$ $\lambda_2 = 2$ *With* $\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\mathbf{x}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$P = [\mathbf{x}_1 \ \mathbf{x}_2] = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ *and* $D = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$

$P^{-1}AP = D \Rightarrow AP = DP$

$$\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 4 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$$

Note: Among other uses, Diagonalization is very useful for computing A^n:

$A^2 = (PDP^{-1})(PDP^{-1}) =$

$PD(P^{-1}P)DP^{-1} = PDIDP^{-1} = PD^2P^{-1}$

Further, $A^n = PD^nP^{-1}$

Example:

$$\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}^5 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}^5\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 4^5 & 0 \\ 0 & 2^5 \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} 4^5 & 2^5 \\ 4^5 & -2^5 \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-1}{2} \end{pmatrix} = \begin{pmatrix} 4^5 \cdot \frac{1}{2} + 2^5 \cdot \frac{1}{2} & 4^5 \cdot \frac{1}{2} - 2^5 \cdot \frac{1}{2} \\ 4^5 \cdot \frac{1}{2} - 2^5 \cdot \frac{1}{2} & 4^5 \cdot \frac{1}{2} + 2^5 \cdot \frac{1}{2} \end{pmatrix} =$$

$$\begin{pmatrix} 2^9 + 2^4 & 2^9 - 2^4 \\ 2^9 - 2^4 & 2^9 + 2^4 \end{pmatrix} = \begin{pmatrix} 528 & 496 \\ 496 & 528 \end{pmatrix}$$

Iterative Estimates for Eigenvalues

Power Method:

1. Let $\mathbf{x}_0 = \mathbf{y}_0$ be any initial vector in $\mathbb{R}^n$ whose largest component is 1.
2. Repeat the following steps for $k = 1, 2, \ldots$:
   (a) Compute $\mathbf{x}_k = A\mathbf{y}_{k-1}$.
   (b) Let m_k be the component of $\mathbf{x}_k$ with the largest absolute value.
   (c) Set $\mathbf{y}_k = (1/m_k)\mathbf{x}_k$.
3. For most choices of $\mathbf{x}_0$,
   m_k will converge to the largest $|\lambda|$
   $\mathbf{y}_0$ will converge to the corresponding vector

Alterations on the Power Method:

- Shifted Power Method:

Use $\mathbf{x}_k = (A - \lambda I)\mathbf{y}_{k-1}$ instead of $\mathbf{x}_k = A\mathbf{y}_{k-1}$ to find the second largest $|\lambda|$

- Inverse Power Method:

Use $\mathbf{x}_k = A^{-1}\mathbf{y}_{k-1}$ instead of $\mathbf{x}_k = A\mathbf{y}_{k-1}$ to find the largest $1/|\lambda|$ or smallest $|\lambda|$.

- Shifted Inverse Power Method:

Use $\mathbf{x}_k = (A - \lambda I)^{-1}\mathbf{y}_{k-1}$ instead of $\mathbf{x}_k = A\mathbf{y}_{k-1}$ to find the second largest $1/|\lambda|$ or second smallest $|\lambda|$.

Gerschgorin's Disk Theorem: Let r_i be the sum of the absolute values of the off-diagonal terms in the i-th row of $A = [a_{ij}]$. The ***i*-th Gerschgorin disk** is a disk in the complex plane defined by:

$$D_i = \{z \text{ in } \mathbb{C} : |z - a_{ii}| \le r_i\}$$

Center: a_{ii} Radius: r_i

Orthogonality

Orthogonal Set: A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ for which all pairs are orthogonal:

$$\mathbf{v}_i \bullet \mathbf{v}_j = 0 \text{ whenever } i \ne j \text{ for } i, j = 1, 2, \ldots, k$$

Orthogonal Basis: A basis of W that is an *orthogonal* set, where W is a subspace of $\mathbb{R}^n$

Orthonormal Set: A set of orthogonal unit vectors

Orthonormal Set: A basis of W that is an *orthonormal* set, where W is a subspace of $\mathbb{R}^n$

Example:

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

$$\mathbf{v}_1 \bullet \mathbf{v}_2 = 2(0) + 1(1) + (-1)(1) = 0$$

$$\mathbf{v}_2 \bullet \mathbf{v}_3 = 0(1) + 1(-1) + (1)(1) = 0$$

$$\mathbf{v}_1 \bullet \mathbf{v}_3 = 2(1) + 1(-1) + (-1)(1) = 0$$

Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set and forms an orthogonal basis.

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \mathbf{q}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{q}_3 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

Thus $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ is an orthonormal set and forms an orthonormal basis.

Note: An orthonormal set can be summarized by the following

$$\mathbf{q}_i \bullet \mathbf{q}_j = \begin{cases} 0 \text{ if } i \ne j \\ 1 \text{ if } i = j \end{cases}$$

Orthogonal Matrix: A matrix Q whose columns form an orthonormal set.

Properties:

If Q is an orthogonal matrix, then

a. $\|Q\mathbf{x}\| = \mathbf{x}$ For every $\mathbf{x}$ in $\mathbb{R}^n$

b. $Q\mathbf{x} \bullet Q\mathbf{y} = \mathbf{x} \bullet \mathbf{y}$ For every $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$

c. Q^{-1} is orthogonal

d. $\det Q = \pm 1$

e. If is λ an eigenvalue, then $|\lambda| = 1$

f. If Q_1 and Q_2 are orthogonal, then so is $Q_1 Q_2$.

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

Orthogonal Projection: The projection of vector $\mathbf{v}$ with respect to an orthogonal basis:

$$\text{proj}_W(\mathbf{v}) = \left(\frac{\mathbf{u}_1 \bullet \mathbf{v}}{\mathbf{u}_1 \bullet \mathbf{u}_1}\right)\mathbf{u}_1 + \cdots + \left(\frac{\mathbf{u}_k \bullet \mathbf{v}}{\mathbf{u}_k \bullet \mathbf{u}_k}\right)\mathbf{u}_k$$

The ***component of* v *orthogonal to W*** is the vector

$$\text{perp}_W(\mathbf{v}) = \mathbf{v} - \text{proj}_W(\mathbf{v})$$

Gram-Schmidt Process:

Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ be a basis for a subspace W of $\mathbb{R}^n$ and define the following:

$$\mathbf{v}_1 = \mathbf{x}_1, \qquad W_1 = \text{span}(\mathbf{x}_1)$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \left(\frac{\mathbf{v}_1 \bullet \mathbf{x}_2}{\mathbf{v}_1 \bullet \mathbf{v}_1}\right)\mathbf{v}_1, \qquad W_2 = \text{span}(\mathbf{x}_1, \mathbf{x}_2)$$

$$\vdots$$

$$\mathbf{v}_k = \mathbf{x}_k - \left(\frac{\mathbf{v}_1 \bullet \mathbf{x}_k}{\mathbf{v}_1 \bullet \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \bullet \mathbf{x}_k}{\mathbf{v}_2 \bullet \mathbf{v}_2}\right)\mathbf{v}_2 - \cdots - \left(\frac{\mathbf{v}_{k-1} \bullet \mathbf{x}_k}{\mathbf{v}_{k-1} \bullet \mathbf{v}_{k-1}}\right)\mathbf{v}_{k-1}, \quad W_k = \text{span}(\mathbf{x}_1, \cdots, \mathbf{x}_k)$$

Then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is an orthogonal basis for W.

QR factorization: Let A be an $m \times n$ matrix whose columns are linearly independent. Then A can be factored as $A = QR$ where Q is an $m \times n$ matrix with orthonormal columns and R is an upper triangular invertible matrix.

Steps:

1. Use the Gram-Schmidt Process to find an orthonormal basis for the columns of A.
2. Use these vectors to write $Q = [\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_n]$.
3. Multiply $Q^T A = R$ to find the upper triangular matrix R.

Example:

$A = \begin{pmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ So,

$\mathbf{x}_1 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 2 \end{pmatrix}$

Using the Gram-Schmidt Process,

$$\mathbf{v}_1 = \mathbf{x}_1 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \left(\frac{\mathbf{v}_1 \bullet \mathbf{x}_2}{\mathbf{v}_1 \bullet \mathbf{v}_1} \right) \mathbf{v}_1$$

$$\mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix} - \left(\frac{2}{4} \right) \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 3/2 \\ 1/2 \\ 1/2 \end{pmatrix}, \mathbf{v}_2' = 2\mathbf{v}_2 = \begin{pmatrix} 3 \\ 3 \\ 1 \\ 1 \end{pmatrix}$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \left(\frac{\mathbf{v}_1 \bullet \mathbf{x}_3}{\mathbf{v}_1 \bullet \mathbf{v}_1} \right) \mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \bullet \mathbf{x}_3}{\mathbf{v}_2 \bullet \mathbf{v}_2} \right) \mathbf{v}_2$$

$$\mathbf{v}_3 = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 2 \end{pmatrix} - \left(\frac{1}{4} \right) \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} - \left(\frac{15}{20} \right) \begin{pmatrix} 3 \\ 3 \\ 1 \\ 1 \end{pmatrix}$$

$$\mathbf{v}_3 = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 2 \end{pmatrix} - \left(\frac{1}{4} \right) \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} - \left(\frac{15}{20} \right) \begin{pmatrix} 3 \\ 3 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 0 \\ 1/2 \\ 1 \end{pmatrix}, \mathbf{v}_3' = 2\mathbf{v}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}$$

$\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3'\}$ is an orthogonal basis for the columns of A.

After normalizing each vector, the orthonormal basis is obtained:

$$\mathbf{q}_1 = \begin{pmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{pmatrix}, \mathbf{q}_2 = \begin{pmatrix} 3\sqrt{5}/10 \\ 3\sqrt{5}/10 \\ \sqrt{5}/10 \\ \sqrt{5}/10 \end{pmatrix}, \mathbf{q}_3 = \begin{pmatrix} -\sqrt{6}/6 \\ 0 \\ \sqrt{6}/6 \\ \sqrt{6}/3 \end{pmatrix}$$

So,

$$Q = [\mathbf{q}_1 \ \mathbf{q}_2 \ \mathbf{q}_3] = \begin{pmatrix} 1/2 & 3\sqrt{5}/10 & -\sqrt{6}/6 \\ -1/2 & 3\sqrt{5}/10 & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/3 \end{pmatrix}$$

$$Q^T A = R = \begin{pmatrix} 1/2 & 3\sqrt{5}/10 & -\sqrt{6}/6 \\ -1/2 & 3\sqrt{5}/10 & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/3 \end{pmatrix} =$$

$$\begin{pmatrix} 1/2 & -1/2 & -1/2 & 1/2 \\ 3\sqrt{5}/10 & 3\sqrt{5}/10 & \sqrt{5}/10 & \sqrt{5}/10 \\ -\sqrt{6}/6 & 0 & \sqrt{6}/6 & \sqrt{6}/3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} =$$

$$\begin{pmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{pmatrix} = R$$

Thus,

$$\begin{pmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1/2 & 3\sqrt{5}/10 & -\sqrt{6}/6 \\ -1/2 & 3\sqrt{5}/10 & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/3 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{pmatrix}$$

$A = QR$.

Orthogonally Diagonizable: A square matrix A is **Orthogonally Diagonizable** if there exists an orthogonal matrix Q and a diagonal matrix D such that $Q^T A Q = D$

Spectral Theorem: A matrix is invertible iff it is orthogonally diagonizable.

This leads to the *spectral decomposition* of a matrix $A = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \lambda_n \mathbf{q}_2 \mathbf{q}_2^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T$

<u>Part 3:</u>

<u>Vector Spaces</u>

Vector Space: A vector space is defined by the following properties for vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ and scalars c and d in V.

1. $\mathbf{u} + \mathbf{v}$ is in V
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

3. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$
4. There exists an element $\mathbf{0}$ in V such that $\mathbf{u}+\mathbf{0}=\mathbf{u}$.
5. For each $\mathbf{u}$ in V, there is an element $-\mathbf{u}$ in V such that $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$.
6. $c\mathbf{u}$ is in V.
7. $c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$
8. $(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$
9. $c(d\mathbf{u})=(cd)\mathbf{u}$
10. $1\mathbf{u}=\mathbf{u}$

Span: If $S=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k\}$ is a set of vectors in V, then the set of all linear combinations of $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k$ *spans* V.

Basis: A subset $\mathcal{B}$ of a vector space V if:
a. $\mathcal{B}$ spans V
b. $\mathcal{B}$ is linearly independent

Coordinates: Let $\mathcal{B}=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$ be a basis for a vector space V. Let $\mathbf{v}$ be a vector in V, and write $\mathbf{v}=c_1\mathbf{v}_1+c_2\mathbf{v}_2+\ldots+c_n\mathbf{v}_n$. Then $c_1,c_2,\ldots,c_n$ are called the ***coordinates with respect to*** $\mathcal{B}$, and the column vector is called the ***coordinate vector of v with respect to*** $\mathcal{B}$

$[\mathbf{v}]_{\mathcal{B}}=\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}$ is called the *coordinate vector of* $\mathbf{v}$ with respect to $\mathcal{B}$.

Dimension: The number of vectors in a basis for V. Denoted $\dim V$.

Subspaces Associated with Matrix A:
row space: $\mathrm{row}(A)$ the subspace spanned by the <u>rows</u> of A
column space: $\mathrm{col}(A)$ the subspace spanned by the <u>columns</u> of A
null space: $\mathrm{null}(A)$ the subspace spanned by the solutions to $A\mathbf{x}=\mathbf{0}$ See **Kernel**.

Rank: $\mathrm{rank}(A)$ The dimension of the row and column spaces of a matrix A

Subspace: A subset W of a vector space that is itself a vector space with the same scalars, addition, and scalar multiplication as V. W is a subset iff the following properties hold:
a. If $\mathbf{u}$ and $\mathbf{v}$ are in W, then $\mathbf{u}+\mathbf{v}$ is in W.
b. If $\mathbf{u}$ is in W and c is a scalar, then $c\mathbf{u}$ is in W.
(Same as Properties 1. and 6. for vector spaces)
$\dim W\le\dim V$

<u>Change of Basis</u>

Change of Basis Matrix: Let $\mathcal{B}=\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n\}$ and $\mathcal{C}=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$ be bases for a vector space V. An $n\times n$ matrix called the ***change-of-basis matrix*** whose columns are the coordinate vectors of the vectors in $\mathcal{B}$ with respect to $\mathcal{C}$ denoted

$$P_{C\leftarrow\mathcal{B}}=\left[[\mathbf{u}_1]_C\,[\mathbf{u}_2]_C\cdots[\mathbf{u}_n]_C\right]$$

Gauss-Jordan Method for Computing the Change of Basis matrix:

$$[C\,|\,B]\to[I\,|\,P_{C\leftarrow\mathcal{B}}]$$

<u>Linear Transformation</u>

Linear Transformation: A mapping from a vector space vector space V to a vector space W

$T:V\to W$ such that, for all $\mathbf{u}$ and $\mathbf{v}$ in V ..
1. $T(\mathbf{u}+\mathbf{v})=T(\mathbf{u})+T(\mathbf{v})$
2. $T(c\mathbf{u})=cT(\mathbf{u})$

Or Combining these two properties, Linear Combinations are preserved:

$$T(c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_n\mathbf{v}_n)=c_1T(\mathbf{v}_1)+c_2T(\mathbf{v}_2)+\cdots c_kT(\mathbf{v}_k)$$

Inverse of a Linear Transformation: For a linear transformation $T:V\to W$, T is ***invertible*** if there is a linear transformation $T':W\to V$ such that $T'\circ T=I_v$ and $T\circ T'=I_w$ where T' is the ***inverse*** for T

Kernel: The set of all vectors in V that are mapped by $T:V\to W$ to $\mathbf{0}$ in W

$$\ker(T)=\{\mathbf{v}\text{ in }V:T(\mathbf{v})=\mathbf{0}\}$$

The dimension of the kernel of T is **nullity**(T).

Range: The set of all vectors in W that are images of vectors in V under $T:V\to W$.

$$\mathrm{range}(T)=\{T(\mathbf{v}):\mathbf{v}\text{ in }V\}$$
$$=\{\mathbf{w}\text{ in }W:\mathbf{w}=T(\mathbf{v})\text{ for some }\mathbf{v}\text{ in }V\}$$

The dimension of the range of T is **rank**(T).

Rank Theorem: For linear transformation $T:V\to W$,

$$\mathrm{rank}(T)+\mathrm{nullity}(T)=\dim V$$

One-to-One and Onto: A linear transformation $T: V \to W$ is ***one-to-one*** if T maps distinct vectors in V to distinct vectors in W.
If $\text{range}(T) = W$, then T is onto.

Matrix of a Linear Transformation

For vector spaces V and W with respective bases $\mathcal{B}$ and C, where $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ with $T: V \to W$, then matrix

$A = \left[[T(\mathbf{v}_1)]_C \, [T(\mathbf{v}_2)]_C \cdots [T(\mathbf{v}_n)]_C \right]$

Which satisfies $A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_C$

A is called the ***matrix of T with respect to the bases*** $\mathcal{B}$ ***and*** C.

Inverse of a Linear Transformation Matrix:

$\left([T]_{C \leftarrow \mathcal{B}}\right)^{-1} = \left[T^{-1}\right]_{\mathcal{B} \leftarrow \mathcal{C}}$

Change of Basis Matrix: A matrix P such that $[T]_C = P^{-1}[T]_{\mathcal{B}} P$

P is called the ***change-of-basis matrix from*** C ***to*** $\mathcal{B}$.

Inner Product Spaces

Inner Product: An operation on a vector space V that assigns a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ to every pair of vectors in V. The following properties hold:

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
2. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
3. $\langle c\mathbf{u}, \mathbf{v} \rangle = c \langle \mathbf{u}, \mathbf{v} \rangle$
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ iff $\mathbf{u} = \mathbf{0}$

A vector space with an inner product is called an ***inner product space.***

Length, Distance, and Orthogonality:

The ***length*** (or ***norm***) of $\mathbf{v}$ is $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$

The ***distance*** between **u** and **v** is $\mathrm{d}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$

u and **v** are orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Norm: : A mapping on a vector space V that associates with each vector $\mathbf{v}$ a real number $\|\mathbf{v}\|$, called the ***norm*** of $\mathbf{v}$ with the following properties:

1. $\|\mathbf{v}\| \geq 0$ and $\|\mathbf{v}\| = 0$ iff $\mathbf{v} = 0$.
2. $\|c\mathbf{v}\| = |c| \|\mathbf{v}\|$
3. $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

A vector space with a norm is called a ***normed linear space.***

Examples of Norms:

$\|\mathbf{v}\| = \left(|v_1|^p + \cdots + |v_n|^p\right)^{1/p}$

For $p = 1$: Sum norm
For $p = 2$: Euclidean Norm (dot product)
For $p = \infty$: Max norm

Matrix Norm: A mapping on that associates with each $n \times n$ matrix a real number $\|A\|$, called the ***norm*** of A with the following properties:

1. $\|A\| \geq 0$ and $\|A\| = 0$ iff $A = 0$.
2. $\|cA\| = |c| \|A\|$
3. $\|A + B\| \leq \|A\| + \|B\|$
4. $\|AB\| \leq \|A\| \|B\|$

Least Squares Approximation

Best Approximation Theorem: If W is a subspace of a normed linear space V, the ***best approximation to*** $\mathbf{v}$ ***in*** W is the vector $\mathbf{v}$ such that

$\|\mathbf{v} - \overline{\mathbf{v}}\| < \|\mathbf{v} - \overline{\mathbf{w}}\|$

for every vector $\mathbf{w}$ in W different from $\overline{\mathbf{v}}$.
" $\text{proj}_W(\mathbf{v})$ is the best approximation to $\mathbf{v}$ in W."

Least Squares Solution: A vector $\overline{\mathbf{x}}$ such that $\|\mathbf{b} - A\overline{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$ is called the ***least squares solution*** of $A\mathbf{x} = \mathbf{b}$. It is unique and given by

$\overline{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$.

If $A = QR$, then $\overline{\mathbf{x}} = R^{-1} Q^T \mathbf{b}$