

MATH 235 Linear Algebra 2 for Honours Mathematics

Keven Qiu

Contents

Review of MATH 136 Linear Algebra 1	3
1 Vector Spaces	4
1.1 Vector Spaces	4
1.2 Bases and Dimension	5
1.3 Coordinates	6
2 Fundamental Subspaces	7
2.1 Bases of Fundamental Subspaces	7
3 Linear Mappings	8
3.1 General Linear Mappings	8
3.2 Rank-Nullity Theorem	8
3.3 Matrix of a Linear Mapping	9
3.4 Isomorphisms	9
4 Inner Products	11
4.1 Inner Product Spaces	11
4.2 Orthogonality and Length	11
4.3 The Gram-Schmidt Procedure	13
4.4 General Projections	13
4.5 The Fundamental Theorem	14
4.6 The Method of Least Squares	15
5 Applications of Orthogonal Matrices	16
5.1 Orthogonal Similarity	16
5.2 Orthogonal Diagonalization	16
5.3 Quadratic Forms	16

5.4	Graphing Quadratic Forms (Optional)	17
5.5	Optimizing Quadratic Forms	17
5.6	Singular Value Decomposition	17
6	Complex Vector Spaces	19
6.4	Complex Inner Product	19
6.5	Unitary Diagonalization	20
7	Introduction to Fourier Analysis	21
7.1	Inner Product Space of Continuous Functions	21
7.2	Fourier Series	22

Review of MATH 136 Linear Algebra 1

System-Rank Theorem:

1. $\text{rank } A < \text{rank}(A|\vec{b})$ if and only if the system is inconsistent.
2. If the system is consistent, then the system contains $(n - \text{rank } A)$ free variables.
3. $\text{rank } A = m$ if and only if $[A|\vec{b}]$ is consistent for every \vec{b} .

Solution Theorem (Applies only to consistent systems): Let $[A|\vec{b}]$ be a consistent system of M linear equations in n variables with RREF $[R|\vec{c}]$. If $\text{rank } A = k < n$, then a vector equation of the solution set of $[A|\vec{b}]$ has the form

$$\vec{x} = \vec{d} + t_1\vec{v}_1 + \cdots + t_{n-k}\vec{v}_{n-k}, \quad t_1, \dots, t_{n-k} \in \mathbb{F}$$

where $\vec{d} \in \mathbb{F}^n$ and $\{\vec{v}_1, \dots, \vec{v}_{n-k}\}$ is a linearly independent set in \mathbb{F}^n .

Invertible Matrix Theorem: For any $n \times n$ matrix A , the following are equivalent: 2

1. A is invertible.
2. The RREF of A is I .
3. $\text{rank } A = n$.
4. A^T is invertible.
5. The nullspace of A is $\{\vec{0}\}$.
6. The columns of A form a basis for \mathbb{F}^n .
7. The rows of A form a basis for \mathbb{F}^n .
8. The system of equations $A\vec{x} = \vec{b}$ is consistent with a unique solution for all $\vec{b} \in \mathbb{F}^n$.

Eigenvalues/Diagonalization:

- Diagonalization is finding an invertible P such that $P^{-1}AP = D$.
- Find eigenvalues \rightarrow entries of D .
- Find eigenvectors \rightarrow columns of P .

Cayley-Hamilton Theorem: If $A \in M_{n \times n}(\mathbb{C})$, then A is a root of its characteristic polynomial $C(\lambda)$. i.e. $C(A) = \mathbb{O}$.

1 Vector Spaces

1.1 Vector Spaces

Definition - Vector Space: A set \mathbb{V} with an operation of addition, denoted $\vec{x} + \vec{y}$, and an operation of scalar multiplication, denoted $c\vec{x}$, is called a vector space over \mathbb{F} if for every $\vec{v}, \vec{x}, \vec{y} \in \mathbb{V}$ and $c, d \in \mathbb{F}$ we have:

V1 $\vec{x} + \vec{y} \in \mathbb{V}$.

V2 $(\vec{x} + \vec{y}) + \vec{v} = \vec{x} + (\vec{y} + \vec{v})$.

V3 $\vec{x} + \vec{y} = \vec{y} + \vec{x}$.

V4 There is a vector $\vec{0} \in \mathbb{V}$, called the zero vector, such that $\vec{x} + \vec{0} = \vec{x}, \forall \vec{x} \in \mathbb{V}$.

V5 $\forall \vec{x} \in \mathbb{V}, \exists (-\vec{x}) \in \mathbb{V}$ such that $\vec{x} + (-\vec{x}) = \vec{0}$.

V6 $c\vec{x} \in \mathbb{V}$.

V7 $c(d\vec{x}) = (cd)\vec{x}$.

V8 $(c + d)\vec{x} = c\vec{x} + d\vec{x}$.

V9 $c(\vec{x} + \vec{y}) = c\vec{x} + c\vec{y}$.

V10 $1\vec{x} = \vec{x}$.

Theorem 1: If \mathbb{V} is a vector space, then

1. $0\vec{x} = \vec{0}, \forall \vec{x} \in \mathbb{V}$.
2. $(-\vec{x}) = (-1)\vec{x}, \forall \vec{x} \in \mathbb{V}$.

Definition - Subspace: Let \mathbb{V} be a vector space. If \mathbb{S} is a subset of \mathbb{V} and \mathbb{S} is a vector space under the same operations as \mathbb{V} , then \mathbb{S} is called a subspace of \mathbb{V} .

Theorem 2 - Subspace Test: If \mathbb{S} is a non-empty subset of \mathbb{V} such that $\vec{x} + \vec{y} \in \mathbb{S}$ and $c\vec{x} \in \mathbb{S}, \forall \vec{x}, \vec{y} \in \mathbb{S}, c \in \mathbb{F}$ under the operations of \mathbb{V} , then \mathbb{S} is a subspace of \mathbb{V} .

Definition - Span: Let $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ be a set of vectors in a vector space \mathbb{V} . The span of \mathcal{B} is

$$\text{Span } \mathcal{B} = \{c_1\vec{v}_1 + \dots + c_k\vec{v}_k \mid c_1, \dots, c_k \in \mathbb{F}\}.$$

We say that $\text{Span } \mathcal{B}$ is spanned by \mathcal{B} and \mathcal{B} is a spanning set for $\text{Span } \mathcal{B}$.

Theorem 3: If $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_k\}$ is a set of vectors in vector space \mathbb{V} , then $\text{Span } \mathcal{B}$ is a subspace of \mathbb{V} .

Theorem 4: Let \mathbb{V} be a vector space and $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{V}$. Then $\vec{v}_i \in \text{Span}\{\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \dots, \vec{v}_k\}$ if and only if

$$\text{Span}\{\vec{v}_1, \dots, \vec{v}_k\} = \text{Span}\{\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \dots, \vec{v}_k\}.$$

Definition - Linearly Dependent/Independent: A set of vectors in a vector space \mathbb{V} is said to be linearly dependent if $\exists c_1, \dots, c_k$ not all zero such that

$$\vec{0} = c_1\vec{v}_1 + \dots + c_k\vec{v}_k$$

has a unique solution.

The set is linearly independent if the only solution to

$$\vec{0} = c_1\vec{v}_1 + \dots + c_k\vec{v}_k$$

is the trivial solution $c_1 = \cdots = c_k = 0$.

Theorem 5: A set of vectors $\{\vec{v}_1, \dots, \vec{v}_k\}$ in a vector space \mathbb{V} is linearly dependent if and only if $\vec{v}_i \in \text{Span}\{\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \dots, \vec{v}_k\}$ for some i , $1 \leq i \leq k$.

Theorem 6: Any set of vectors $\{\vec{v}_1, \dots, \vec{v}_k\}$ in a vector space \mathbb{V} which contains the zero vector is linearly dependent.

1.2 Bases and Dimension

Definition - Basis: Let \mathbb{V} be a vector space. A set \mathcal{B} is called a basis for \mathbb{V} if \mathcal{B} is a linearly independent spanning set for \mathbb{V} .

We define a basis for $\{\vec{0}_{\mathbb{V}}\}$ to be the empty set.

Finding a Basis Algorithm: Let \mathbb{V} be a vector space that can be spanned by a finite number of vectors. To find a basis for \mathbb{V} :

1. Find a general form of a vector $\vec{v} \in \mathbb{V}$.
2. Write the general form of \vec{x} as a linear combination of vectors $\vec{v}_1, \dots, \vec{v}_k$ in \mathbb{V} .
3. Check if $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_k\}$ is linearly independent. If it is, then stop as \mathcal{B} is a basis.
4. If \mathcal{B} is linearly dependent, pick some vector $\vec{v}_i \in \mathcal{B}$ that can be written as a linear combination of the other vectors in \mathcal{B} and remove it using Theorem 4.1.4. Repeat until linearly independent.

Dimension

Theorem 1: Let $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ be a basis for a vector space \mathbb{V} and let $\mathcal{C} = \{\vec{w}_1, \dots, \vec{w}_k\}$ be a set in \mathbb{V} . If $k > n$, then \mathcal{C} is linearly dependent.

Theorem 2: If $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ and $\mathcal{C} = \{\vec{w}_1, \dots, \vec{w}_k\}$ are bases for a vector space \mathbb{V} , then $k = n$.

Definition - Dimension: If $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for a vector space \mathbb{V} , then we say the dimension of \mathbb{V} is

$$\dim \mathbb{V} = n.$$

If \mathbb{V} is a trivial vector, then $\dim \mathbb{V} = 0$. If \mathbb{V} does not have a basis with a finite number of vectors in it, then \mathbb{V} is said to be infinite dimensional.

Theorem 3: If \mathbb{V} is an n -dimensional vector space with $n > 0$, then

1. A set of more than n vectors in \mathbb{V} must be linearly dependent.
2. A set of fewer than n vectors in \mathbb{V} cannot span \mathbb{V} .
3. A set of n vectors in \mathbb{V} is linearly independent **iff** it spans \mathbb{V} .

Theorem 4: If \mathbb{V} is an n -dimensional vector space and $\{\vec{v}_1, \dots, \vec{v}_k\}$ is a linearly independent set in \mathbb{V} with $k < n$, then there exist vectors $\vec{w}_{k+1}, \dots, \vec{w}_n$ in \mathbb{V} such that $\{\vec{v}_1, \dots, \vec{v}_k, \vec{w}_{k+1}, \dots, \vec{w}_n\}$ is a basis for \mathbb{V} .

Corollary 5: If \mathbb{S} is a subspace of a finite dimensional vector space \mathbb{V} , then $\dim \mathbb{S} \leq \dim \mathbb{V}$.

Important Test of Basis: Let $A = [\vec{v}_1, \dots, \vec{v}_n] \in M_{n \times n}(\mathbb{F})$, $B = \{\vec{v}_1, \dots, \vec{v}_n\}$, $r = \text{rank } A$, then

- If $r = m$, $\text{Span } \mathcal{B} = \mathbb{F}^m$.
- If $r = n$, \mathcal{B} is linearly independent.
- If $r = m = n$, \mathcal{B} is a basis.

1.3 Coordinates

Theorem 1: If $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for a vector space \mathbb{V} , then every $\vec{v} \in \mathbb{V}$ can be written as a unique linear combination of the vectors in \mathcal{B} .

Definition - \mathcal{B} -Coordinates, \mathcal{B} -Coordinate Vector: Let $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ be a basis for a vector space \mathbb{V} . If $\vec{v} = b_1\vec{v}_1 + \dots + b_n\vec{v}_n$, then b_1, \dots, b_n are called the **\mathcal{B} -coordinates** of \vec{v} , and we define the **\mathcal{B} -coordinate vector** by

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Theorem 2: If \mathbb{V} is a vector space with basis $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_k\}$, then for any $\vec{v}, \vec{w} \in \mathbb{V}$ and $s, t \in \mathbb{R}$ we have

$$[s\vec{v} + t\vec{w}]_{\mathcal{B}} = s[\vec{v}]_{\mathcal{B}} + t[\vec{w}]_{\mathcal{B}}.$$

Change of Coordinates

Definition - Change of Coordinates Matrix: Let $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ and \mathcal{C} both be bases for a vector space \mathbb{V} . The change of coordinates matrix from \mathcal{B} -coordinates to \mathcal{C} -coordinates is defined by

$${}_C P_{\mathcal{B}} = \begin{bmatrix} [\vec{v}_1]_{\mathcal{C}} & \dots & [\vec{v}_n]_{\mathcal{C}} \end{bmatrix}$$

and for any $\vec{x} \in \mathbb{V}$ we have

$$[\vec{x}]_{\mathcal{C}} = {}_C P_{\mathcal{B}} [\vec{x}]_{\mathcal{B}}.$$

Theorem 3: If \mathcal{B} and \mathcal{C} are bases for an n -dimensional vector space \mathbb{V} , then the change of coordinate matrices ${}_C P_{\mathcal{B}}$ and ${}_B P_{\mathcal{C}}$ satisfy

$${}_C P_{\mathcal{B}} {}_B P_{\mathcal{C}} = I = {}_B P_{\mathcal{C}} {}_C P_{\mathcal{B}}.$$

2 Fundamental Subspaces

2.1 Bases of Fundamental Subspaces

Definition - Fundamental Subspaces: Let A be an $m \times n$ matrix. The four fundamental subspaces of A are

1. $\text{Col}(A) = \{A\vec{x} \in \mathbb{R}^m \mid \vec{x} \in \mathbb{R}^n\}$, called the **column space**,
2. $\text{Row}(A) = \{A^T\vec{x} \in \mathbb{R}^n \mid \vec{x} \in \mathbb{R}^m\}$, called the **row space**,
3. $\text{Null}(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}$, called the **nullspace**,
4. $\text{Null}(A^T) = \{\vec{x} \in \mathbb{R}^m \mid A^T\vec{x} = \vec{0}\}$, called the **left nullspace**.

Theorem 1: If A is an $m \times n$ matrix, then $\text{Col}(A)$ and $\text{Null}(A^T)$ are subspaces of \mathbb{R}^m , and $\text{Row}(A)$ and $\text{Null}(A)$ are subspaces of \mathbb{R}^n .

Theorem 2: If A is an $m \times n$ matrix, then the columns of A which correspond to leading ones in the reduced row echelon form of A form a basis for $\text{Col}(A)$. Moreover,

$$\dim \text{Col}(A) = \text{rank } A$$

Lemma 3: If R is an $m \times n$ matrix and E is an $n \times n$ invertible matrix, then

$$\{RE\vec{x} \mid \vec{x} \in \mathbb{R}^n\} = \{R\vec{y} \mid \vec{y} \in \mathbb{R}^n\}$$

Theorem 4: If A is an $m \times n$ matrix, then the non-zero rows in the reduced row echelon form of A form a basis for $\text{Row}(A)$. Hence, $\dim \text{Row}(A) = \text{rank } A$.

Corollary 5: For any $m \times n$ matrix A we have $\text{rank } A = \text{rank } A^T$.

Remark: Due to its usefulness, it is very common to define the rank of a matrix as the dimension of the column space (or row space) and then prove that this is equivalent to the number of leading ones in the reduced row echelon form.

Remark: It is not actually necessary to row reduce A^T to find a basis for the left nullspace of A . We can use the fact that $EA = R$ where E is the product of elementary matrices used to bring A to its reduced row echelon form R to find a basis for the left nullspace.

Theorem 6 - Dimension Theorem: If A is an $m \times n$ matrix, then

$$\text{rank } A + \dim \text{Null}(A) = n$$

3 Linear Mappings

3.1 General Linear Mappings

Linear Mappings $L : \mathbb{V} \rightarrow \mathbb{W}$

Definition - Linear Mapping: Let \mathbb{V} and \mathbb{W} be vector spaces. A mapping $L : \mathbb{V} \rightarrow \mathbb{W}$ is called linear if

$$L(s\vec{x} + t\vec{y}) = sL(\vec{x}) + tL(\vec{y})$$

for all $\vec{x}, \vec{y} \in \mathbb{V}$.

Two linear mappings $L : \mathbb{V} \rightarrow \mathbb{W}$ and $M : \mathbb{V} \rightarrow \mathbb{W}$ are said to be equal if $L(\vec{v}) = M(\vec{v})$ for all $\vec{v} \in \mathbb{V}$.

Definition - Addition, Scalar Multiplication: Let $L : \mathbb{V} \rightarrow \mathbb{W}$ and $M : \mathbb{V} \rightarrow \mathbb{W}$ be linear mappings. We define $L + M : \mathbb{V} \rightarrow \mathbb{W}$ by

$$(L + M)(\vec{v}) = L(\vec{v}) + M(\vec{v})$$

and for any $t \in \mathbb{R}$ we define $tL : \mathbb{V} \rightarrow \mathbb{W}$ by

$$(tL)(\vec{v}) = tL(\vec{v})$$

Theorem 1: Let \mathbb{V} and \mathbb{W} be vector spaces. The set \mathbb{L} of all linear mappings $L : \mathbb{V} \rightarrow \mathbb{W}$ with standard addition and scalar multiplication of mappings is a vector space.

Definition - Composition: Let $L : \mathbb{V} \rightarrow \mathbb{W}$ and $M : \mathbb{W} \rightarrow \mathbb{U}$ be linear mappings. We define $M \circ L : \mathbb{V} \rightarrow \mathbb{U}$ by

$$(M \circ L)(\vec{v}) = M(L(\vec{v}))$$

for all $\vec{v} \in \mathbb{V}$

Theorem 2: If $L : \mathbb{V} \rightarrow \mathbb{W}$ and $M : \mathbb{W} \rightarrow \mathbb{U}$ are linear mappings, then $M \circ L : \mathbb{V} \rightarrow \mathbb{U}$ is a linear mapping.

Definition - Invertible Mapping: Let $L : \mathbb{V} \rightarrow \mathbb{W}$ and $M : \mathbb{W} \rightarrow \mathbb{V}$ be linear mappings. If $(M \circ L)(\vec{v}) = \vec{v}$ for all $\vec{v} \in \mathbb{V}$ and $(L \circ M)(\vec{w}) = \vec{w}$ for all $\vec{w} \in \mathbb{W}$, then L and M are said to be invertible. We write $M = L^{-1}$ and $L = M^{-1}$.

3.2 Rank-Nullity Theorem

Definition - Range, Kernel: For a linear mapping $L : \mathbb{V} \rightarrow \mathbb{W}$ the kernel of L is defined to be

$$\ker(L) = \{\vec{v} \in \mathbb{V} \mid L(\vec{v}) = \vec{0}_{\mathbb{W}}\}$$

and the range of L is defined to be

$$\text{Range}(L) = \{L(\vec{v}) \in \mathbb{W} \mid \vec{v} \in \mathbb{V}\}$$

Theorem 1: If \mathbb{V} and \mathbb{W} are vector spaces and $L : \mathbb{V} \rightarrow \mathbb{W}$ is a linear mapping, then

$$L(\vec{0}_{\mathbb{V}}) = \vec{0}_{\mathbb{W}}$$

Theorem 2: If $L : \mathbb{V} \rightarrow \mathbb{W}$ is a linear mapping, then $\ker(L)$ is a subspace of \mathbb{V} and $\text{Range}(L)$ is a subspace of \mathbb{W} .

Definition - Rank, Nullity: Let $L : \mathbb{V} \rightarrow \mathbb{W}$ be a linear mapping. We define the rank of L by

$$\text{rank}(L) = \dim(\text{Range}(L))$$

We define the nullity of L to be

$$\text{nullity}(L) = \dim(\ker(L))$$

Theorem 3 - Rank-Nullity Theorem: Let \mathbb{V} be an n -dimensional vector space and let \mathbb{W} be a vector space. If $L : \mathbb{V} \rightarrow \mathbb{W}$ is linear, then

$$\text{rank}(L) + \text{nullity}(L) = n$$

Remark: *The proof of the Rank-Nullity Theorem is essentially identical to that of the Dimension Theorem.*

3.3 Matrix of a Linear Mapping

Definition - Matrix of a Linear Mapping: Suppose $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is any basis for a vector space \mathbb{V} and \mathcal{C} is any basis for a finite dimensional vector space \mathbb{W} . For a linear mapping $L : \mathbb{V} \rightarrow \mathbb{W}$, the matrix of L with respect to bases \mathcal{B} and \mathcal{C} is defined by

$${}_C[L]_{\mathcal{B}} = \begin{bmatrix} [L(\vec{v}_1)]_{\mathcal{C}} \cdots [L(\vec{v}_n)]_{\mathcal{C}} \end{bmatrix}$$

and satisfies

$$[L(\vec{x})]_{\mathcal{C}} = {}_C[L]_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}$$

for all $\vec{x} \in \mathbb{V}$.

Definition - Matrix of a Linear Operator: Suppose $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is any basis for a vector space \mathbb{V} and let $L : \mathbb{V} \rightarrow \mathbb{V}$ be a linear operator. The \mathcal{B} -matrix of L (or the matrix of L with respect to the basis \mathcal{B}) is defined by

$$[L]_{\mathcal{B}} = \begin{bmatrix} [L(\vec{v}_1)]_{\mathcal{B}} \cdots [L(\vec{v}_n)]_{\mathcal{B}} \end{bmatrix}$$

and satisfies

$$[L(\vec{x})]_{\mathcal{B}} = [L]_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}$$

Remark: *The relationship between diagonalization and the matrix of a linear operator is extremely important.*

3.4 Isomorphisms

Definition - One-To-One, Onto: Let $L : \mathbb{V} \rightarrow \mathbb{W}$ be a linear mapping. L is called one-to-one (injective) if for every $\vec{u}, \vec{v} \in \mathbb{V}$ such that $L(\vec{u}) = L(\vec{v})$, we must have $\vec{u} = \vec{v}$. L is called onto (surjective) if for every $\vec{w} \in \mathbb{W}$, there exists $\vec{v} \in \mathbb{V}$ such that $L(\vec{v}) = \vec{w}$.

Lemma 1: Let $L : \mathbb{V} \rightarrow \mathbb{W}$ be a linear mapping. L is one-to-one if and only if $\ker(L) = \{\vec{0}\}$.

Isomorphisms

Definition - Isomorphism, Isomorphic: A vector space \mathbb{V} is said to be isomorphic to a vector space \mathbb{W} if there exists a linear mapping $L : \mathbb{V} \rightarrow \mathbb{W}$ which is one-to-one and onto. L is called an isomorphism from $\mathbb{V} \rightarrow \mathbb{W}$.

Theorem 2: Let \mathbb{V} and \mathbb{W} be finite dimensional vector spaces. \mathbb{V} is isomorphic to \mathbb{W} if and only if $\dim \mathbb{V} = \dim \mathbb{W}$.

Theorem 3: If \mathbb{V} and \mathbb{W} are both n -dimensional vector spaces and $L : \mathbb{V} \rightarrow \mathbb{W}$ is linear, then L is one-to-one if and only if L is onto.

Theorem 4: Let \mathbb{V} and \mathbb{W} be isomorphic vector spaces and let $\{\vec{v}_1, \dots, \vec{v}_n\}$ be a basis for \mathbb{V} . A linear mapping $L : \mathbb{V} \rightarrow \mathbb{W}$ is an isomorphism if and only if $\{L(\vec{v}_1), \dots, L(\vec{v}_n)\}$ is a basis for \mathbb{W} .

4 Strategies for proving that L is an Isomorphism:

- (1) Prove that L is one-to-one and onto
- (2) $\ker(L) = \{\vec{0}_{\mathbb{V}}\}$ and $\dim \mathbb{V} = \dim \mathbb{W}$
- (3) $\text{Range}(L) = \mathbb{W}$ and $\dim \mathbb{V} = \dim \mathbb{W}$
- (4) $\ker(L) = \{\vec{0}_{\mathbb{V}}\}$ and $\text{Range}(L) = \mathbb{W}$

4 Strategies for proving that L is not an Isomorphism:

- (1) $\ker(L) \neq \{\vec{0}_{\mathbb{V}}\}$
- (2) $\text{Range}(L) \neq \mathbb{W}$
- (3) $\dim \mathbb{V} \neq \dim \mathbb{W}$
- (4) Find bases \mathcal{B} of \mathbb{V} and \mathcal{C} of \mathbb{W} and show that ${}_C[L]_{\mathcal{B}}$ is NOT invertible (RREF is not identity matrix)

4 Inner Products

4.1 Inner Product Spaces

Definition - Inner Product, Inner Product Space: Let \mathbb{V} be a vector space. An inner product on \mathbb{V} is a function $\langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R}$ that has the following properties: for every $\vec{v}, \vec{u}, \vec{w} \in \mathbb{V}$ and $s, t \in \mathbb{R}$ we have

- I1 $\langle \vec{v}, \vec{v} \rangle \geq 0$, and $\langle \vec{v}, \vec{v} \rangle = 0$ if and only if $\vec{v} = 0$ (**Positive Definite**)
- I2 $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$ (**Symmetric**)
- I3 $\langle s\vec{v} + t\vec{u}, \vec{w} \rangle = s\langle \vec{v}, \vec{w} \rangle + t\langle \vec{u}, \vec{w} \rangle$ (**Left linear**)

Remark: Since an inner product is left linear and symmetric, then it is also right linear.

$$\langle \vec{w}, s\vec{v} + t\vec{u} \rangle = s\langle \vec{w}, \vec{v} \rangle + t\langle \vec{w}, \vec{u} \rangle$$

Thus, an inner product is **bilinear**.

The dot product is called the **standard** inner product on \mathbb{R} .

$\langle A, B \rangle = \text{tr}(B^T A)$ is called the standard inner product on $M_{m \times n}(\mathbb{R})$.

Theorem 1: If \mathbb{V} is an inner product space with inner product $\langle \cdot, \cdot \rangle$, then for any $\vec{v} \in \mathbb{V}$ we have

$$\langle \vec{v}, \vec{0} \rangle = 0$$

4.2 Orthogonality and Length

Length

Definition - Length: Let \mathbb{V} be an inner product with inner product $\langle \cdot, \cdot \rangle$. For any $\vec{v} \in \mathbb{V}$ we define the length (or norm) of \vec{v} by

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$$

Theorem 1: Let \mathbb{V} be an inner product space with inner product $\langle \cdot, \cdot \rangle$. For any $\vec{x}, \vec{y} \in \mathbb{V}$ and $t \in \mathbb{R}$ we have

- (1) $\|\vec{x}\| \geq 0$, and $\|\vec{x}\| = 0$ if and only if $\vec{x} = \vec{0}$.
- (2) $\|t\vec{v}\| = |t|\|\vec{v}\|$.
- (3) $\langle \vec{x}, \vec{y} \rangle \leq \|\vec{x}\|\|\vec{y}\|$ (Cauchy-Schwarz-Bunyakovski Inequality)
- (4) $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ (Triangle Inequality)

Definition - Unit Vector: Let \mathbb{V} be an inner product space with inner product $\langle \cdot, \cdot \rangle$. If $\vec{v} \in \mathbb{V}$ is a vector such that $\|\vec{v}\| = 1$, then \vec{v} is called a unit vector.

We can normalize a vector \vec{v} by

$$\hat{v} = \frac{1}{\|\vec{v}\|} \vec{v}$$

Orthogonality

Definition - Orthogonal Vectors: Let \mathbb{V} be an inner product space with inner product \langle, \rangle . If $\vec{x}, \vec{y} \in \mathbb{V}$ such that

$$\langle \vec{x}, \vec{y} \rangle = 0$$

then \vec{x} and \vec{y} are said to be orthogonal.

Definition - Orthogonal Set: If $\mathcal{S} = \{\vec{v}_1, \dots, \vec{v}_k\}$ is a set of vectors in an inner product space \mathbb{V} with inner product \langle, \rangle such that $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ for all $i \neq j$, then \mathcal{S} is called an orthogonal set.

Theorem 2: If $\{\vec{v}_1, \dots, \vec{v}_k\}$ is an orthogonal set in an inner product space \mathbb{V} , then

$$\|\vec{v}_1 + \dots + \vec{v}_k\|^2 = \|\vec{v}_1\|^2 + \dots + \|\vec{v}_k\|^2$$

Theorem 3: If $\mathcal{S} = \{\vec{v}_1, \dots, \vec{v}_k\}$ is an orthogonal set in an inner product space \mathbb{V} with inner product \langle, \rangle such that $\vec{v}_i \neq \vec{0}$ for all $1 \leq i \leq k$, then \mathcal{S} is linearly independent.

Definition - Orthogonal Basis: If \mathcal{B} is an orthogonal set in an inner product space \mathbb{V} that is a basis for \mathbb{V} , then \mathcal{B} is called an orthogonal basis for \mathbb{V} .

Theorem 4: If $\mathcal{S} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthogonal basis for an inner product space \mathbb{V} with inner product \langle, \rangle and $\vec{v} \in \mathbb{V}$, then the coefficient of \vec{v}_i when \vec{v} is written as a linear combination of the vectors in \mathcal{S} is $\frac{\langle \vec{v}, \vec{v}_i \rangle}{\|\vec{v}_i\|^2}$. In particular,

$$\vec{v} = \frac{\langle \vec{v}, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 + \dots + \frac{\langle \vec{v}, \vec{v}_n \rangle}{\|\vec{v}_n\|^2} \vec{v}_n$$

Orthonormal Bases

Definition - Orthonormal Set: If $\mathcal{S} = \{\vec{v}_1, \dots, \vec{v}_k\}$ is an orthogonal set in an inner product space \mathbb{V} such that $\|\vec{v}_i\| = 1$ for $1 \leq i \leq k$, then \mathcal{S} is called an orthonormal set.

Definition - Orthonormal Basis: A basis for an inner product space \mathbb{V} which is an orthonormal set is called an orthonormal basis of \mathbb{V} .

Corollary 5: If \vec{v} is any vector in an inner product space \mathbb{V} with inner product \langle, \rangle and $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthonormal basis for \mathbb{V} , then

$$\vec{v} = \langle \vec{v}, \vec{v}_1 \rangle \vec{v}_1 + \dots + \langle \vec{v}, \vec{v}_n \rangle \vec{v}_n$$

Orthogonal Matrices

Theorem 6: For an $n \times n$ matrix P , the following are equivalent.

- (1) The columns of P form an orthonormal basis for \mathbb{R}^n .
- (2) $P^T = P^{-1}$.
- (3) The rows of P form an orthonormal basis for \mathbb{R}^n .

Definition - Orthogonal Matrix: If the columns of an $n \times n$ matrix P form an orthonormal basis for \mathbb{R}^n , then P is called an orthogonal matrix.

Theorem 7: If P and Q are $n \times n$ orthogonal matrices and $\vec{x}, \vec{y} \in \mathbb{R}^n$, then

- (1) $(P\vec{x}) \cdot (P\vec{y}) = \vec{x} \cdot \vec{y}$
- (2) $\|P\vec{x}\| = \|\vec{x}\|$
- (3) $\det P = \pm 1$
- (4) All real eigenvalues of P are 1 or -1
- (5) PQ is an orthogonal matrix

4.3 The Gram-Schmidt Procedure

Theorem 1 - Gram-Schmidt Orthogonalization Theorem: Let $\{\vec{w}_1, \dots, \vec{w}_n\}$ be a basis for an inner product space \mathbb{W} . If we define $\vec{v}_1, \dots, \vec{v}_n$ successively as follows:

$$\begin{aligned}\vec{v}_1 &= \vec{w}_1 \\ \vec{v}_2 &= \vec{w}_2 - \frac{\langle \vec{w}_2, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 \\ \vec{v}_i &= \vec{w}_i - \frac{\langle \vec{w}_i, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 - \frac{\langle \vec{w}_i, \vec{v}_2 \rangle}{\|\vec{v}_2\|^2} \vec{v}_2 - \dots - \frac{\langle \vec{w}_i, \vec{v}_{i-1} \rangle}{\|\vec{v}_{i-1}\|^2} \vec{v}_{i-1}\end{aligned}$$

for $3 \leq k \leq n$, then $\{\vec{v}_1, \dots, \vec{v}_k\}$ is an orthogonal basis for $\text{Span}\{\vec{w}_1, \dots, \vec{w}_k\}$ for $1 \leq k \leq n$.

Theorem 2 - QR-Decomposition: Let $A = [\vec{a}_1, \dots, \vec{a}_n] \in M_{m \times n}(\mathbb{R})$ where $\dim \text{Col}(A) = n$. Let $\vec{q}_1, \dots, \vec{q}_n$ denote the vectors that result from applying the Gram-Schmidt procedure to the columns of A (in order) and then normalizing. If we define

$$Q = [\vec{q}_1 \quad \dots \quad \vec{q}_n]$$

and

$$R = \begin{bmatrix} \vec{a}_1 \cdot \vec{q}_1 & \vec{a}_2 \cdot \vec{q}_1 & \dots & \vec{a}_n \cdot \vec{q}_1 \\ 0 & \vec{a}_2 \cdot \vec{q}_2 & \dots & \vec{a}_n \cdot \vec{q}_2 \\ 0 & 0 & \ddots & \vec{a}_n \cdot \vec{q}_{n-1} \\ 0 & \dots & 0 & \vec{a}_n \cdot \vec{q}_n \end{bmatrix}$$

4.4 General Projections

Definition - Orthogonal Complement: Let \mathbb{W} be a subspace of an inner product space \mathbb{V} . The orthogonal complement \mathbb{W}^\perp of \mathbb{W} in \mathbb{V} is defined by

$$\mathbb{W}^\perp = \{\vec{v} \in \mathbb{V} \mid \langle \vec{w}, \vec{v} \rangle = 0 \text{ for all } \vec{w} \in \mathbb{W}\}$$

Theorem 1: Let $\{\vec{v}_1, \dots, \vec{v}_k\}$ be a spanning set for a subspace \mathbb{W} of an inner product space \mathbb{V} , and let $\vec{x} \in \mathbb{V}$. We have that $\vec{x} \in \mathbb{W}^\perp$ if and only if $\langle \vec{x}, \vec{v}_i \rangle = 0$ for $1 \leq i \leq k$.

Theorem 2: If \mathbb{W} is a subspace of an inner product space \mathbb{V} , then

- (1) \mathbb{W}^\perp is subspace of \mathbb{V}
- (2) If $\dim \mathbb{V} = n$, then $\dim \mathbb{W}^\perp = n - \dim \mathbb{W}$

- (3) If $\dim \mathbb{V} = n$, then $(\mathbb{W}^\perp)^\perp = \mathbb{W}$
- (4) $\mathbb{W} \cap \mathbb{W}^\perp = \{\vec{0}\}$
- (5) If $\dim \mathbb{V} = n$, $\{\vec{v}_1, \dots, \vec{v}_k\}$ is an orthogonal basis for \mathbb{W} , and $\{\vec{v}_{k+1}, \dots, \vec{v}_n\}$ is an orthogonal basis for \mathbb{W}^\perp , then $\{\vec{v}_1, \dots, \vec{v}_k, \vec{v}_{k+1}, \dots, \vec{v}_n\}$ is an orthogonal basis for \mathbb{V}

Remark: Notice that $\mathbb{U} \subset (\mathbb{U}^\perp)^\perp$

Definition - Projection, Perpendicular: Suppose \mathbb{W} is a k -dimensional subspace of an inner product space \mathbb{V} and $\{\vec{v}_1, \dots, \vec{v}_k\}$ is an orthogonal basis for \mathbb{W} . For any $\vec{v} \in \mathbb{V}$ we define the projection of \vec{v} onto \mathbb{W} by

$$\text{proj}_{\mathbb{W}}(\vec{v}) = \frac{\langle \vec{v}, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 + \dots + \frac{\langle \vec{v}, \vec{v}_k \rangle}{\|\vec{v}_k\|^2} \vec{v}_k$$

and the perpendicular of the projection by

$$\text{perp}_{\mathbb{W}}(\vec{v}) = \vec{v} - \text{proj}_{\mathbb{W}}(\vec{v})$$

Theorem 3: If \mathbb{W} is a k -dimensional subspace of an inner product space \mathbb{V} , then for any $\vec{v} \in \mathbb{V}$ we have

$$\text{perp}_{\mathbb{W}}(\vec{v}) = \vec{v} - \text{proj}_{\mathbb{W}}(\vec{v}) \in \mathbb{W}^\perp$$

Theorem 4: If \mathbb{W} is a k -dimensional subspace of an inner product space \mathbb{V} , then $\text{proj}_{\mathbb{W}}$ is a linear operator on \mathbb{V} with kernel \mathbb{W}^\perp .

Theorem 5: If \mathbb{W} is a subspace of a finite dimensional inner product space \mathbb{V} , then for any $\vec{v} \in \mathbb{V}$ we have

$$\text{proj}_{\mathbb{W}^\perp}(\vec{v}) = \text{perp}_{\mathbb{W}}(\vec{v})$$

4.5 The Fundamental Theorem

Definition - Direct Sum: Let \mathbb{V} be a vector space and \mathbb{U} and \mathbb{W} be subspaces of a vector space \mathbb{V} such that $\mathbb{U} \cap \mathbb{W} = \{\vec{0}\}$. The direct sum of \mathbb{U} and \mathbb{W} is

$$\mathbb{U} \oplus \mathbb{W} = \{\vec{u} + \vec{w} \in \mathbb{V} \mid \vec{u} \in \mathbb{U}, \vec{w} \in \mathbb{W}\}$$

Theorem 1: If \mathbb{U} and \mathbb{W} are subspaces of a vector space \mathbb{V} such that $\mathbb{U} \cap \mathbb{W} = \{\vec{0}\}$, then $\mathbb{U} \oplus \mathbb{W}$ is a subspace of \mathbb{V} . Moreover, if $\{\vec{v}_1, \dots, \vec{v}_k\}$ is a basis for \mathbb{U} and $\{\vec{w}_1, \dots, \vec{w}_l\}$ is a basis for \mathbb{W} , then $\{\vec{v}_1, \dots, \vec{v}_k, \vec{w}_1, \dots, \vec{w}_l\}$ is a basis for $\mathbb{U} \oplus \mathbb{W}$.

Corollary 2: If \mathbb{U} and \mathbb{W} are subspaces of a vector space \mathbb{V} such that $\mathbb{U} \cap \mathbb{W} = \{\vec{0}\}$ and $\vec{v} \in \mathbb{U} \oplus \mathbb{W}$, then there exists unique $\vec{u} \in \mathbb{U}$ and $\vec{w} \in \mathbb{W}$ such that $\vec{v} = \vec{u} + \vec{w}$.

Theorem 3: If \mathbb{V} is a finite dimensional inner product space and \mathbb{W} is a subspace of \mathbb{V} , then

$$\mathbb{W} \oplus \mathbb{W}^\perp = \mathbb{V}$$

Theorem 4 - The Fundamental Theorem of Linear Algebra: If A is an $m \times n$ matrix, then $\text{Col}(A)^\perp = \text{Null}(A^T)$ and $\text{Row}(A)^\perp = \text{Null}(A)$. In particular,

$$\mathbb{R}^n = \text{Row}(A) \oplus \text{Null}(A) \text{ and } \mathbb{R}^m = \text{Col}(A) \oplus \text{Null}(A^T)$$

4.6 The Method of Least Squares

Theorem 1 - Approximation Theorem: Let \mathbb{W} be a finite dimensional subspace of an inner product space \mathbb{C} . If $\vec{v} \in \mathbb{V}$, then the vector closest to \vec{v} in \mathbb{W} is $\text{proj}_{\mathbb{W}}(\vec{v})$. That is,

$$\|\vec{v} - \text{proj}_{\mathbb{W}}(\vec{v})\| < \|\vec{v} - \vec{w}\|$$

for all $\vec{w} \in \mathbb{W}$, $\vec{w} \neq \text{proj}_{\mathbb{W}}(\vec{v})$.

Theorem 2: Let m data points $(x_1, y_1), \dots, (x_m, y_m)$ be given and write

$$\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix}$$

If $\vec{a} = \begin{bmatrix} a_0 \\ \vdots \\ a_n \end{bmatrix}$ is any solution to the normal system

$$X^T X \vec{a} = X^T \vec{y}$$

then the polynomial

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

is the best fitting polynomial of degree n for the given data. Moreover, if at least $n + 1$ of the numbers x_1, \dots, x_m are distinct, then the matrix $X^T X$ is invertible and thus \vec{a} is unique with

$$\vec{a} = (X^T X)^{-1} X^T \vec{y}$$

5 Applications of Orthogonal Matrices

5.1 Orthogonal Similarity

Definition - Orthogonally Similar: Two matrices A and B are said to be orthogonally similar if there exists an orthogonal matrix P such that

$$P^T A P = B$$

Theorem 1 - Triangularization Theorem: If A is an $n \times n$ matrix with real eigenvalues, then A is orthogonally similar to an upper triangular matrix T .

5.2 Orthogonal Diagonalization

Definition - Orthogonally Diagonalizable: An $n \times n$ matrix A is said to be orthogonally diagonalizable if there exists an orthogonal matrix P and diagonal matrix D such that

$$P^T A P = D$$

that is, if A is orthogonally similar to a diagonal matrix.

Theorem 1: If A is orthogonally diagonalizable, then $A^T = A$.

Definition - Symmetric Matrix: A matrix A such that $A^T = A$ is said to be symmetric.

Lemma 2: If A is a symmetric matrix with real entries, then all of its eigenvalues are real.

Theorem 3 - Principal Axis Theorem: Every symmetric matrix A is orthogonally diagonalizable.

A matrix is orthogonally diagonalizable if and only if it is symmetric.

Theorem 4: A matrix A is symmetric if and only if $\vec{x} \cdot (A\vec{y}) = (A\vec{x}) \cdot \vec{y}$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$.

Theorem 5: If \vec{v}_1, \vec{v}_2 are eigenvectors of a symmetric matrix A corresponding to distinct eigenvalues λ_1, λ_2 , then \vec{v}_1 is orthogonal to \vec{v}_2 .

5.3 Quadratic Forms

Definition - Quadratic Form: Let A be an $n \times n$ matrix. A function $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ of the form

$$Q(\vec{x}) = \vec{x}^T A \vec{x} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

is called a quadratic form.

Definitions of Quadratic Forms: Let $Q(\vec{x})$ be a quadratic form. We say that:

1. $Q(\vec{x})$ is **positive definite** if $Q(\vec{x}) > 0$ for all $\vec{x} \neq \vec{0}$
2. $Q(\vec{x})$ is **negative definite** if $Q(\vec{x}) < 0$ for all $\vec{x} \neq \vec{0}$
3. $Q(\vec{x})$ is **indefinite** if $Q(\vec{x}) > 0$ for some \vec{x} and $Q(\vec{x}) < 0$ for some \vec{x}

4. $Q(\vec{x})$ is **positive semidefinite** if $Q(\vec{x}) \geq 0$ for all \vec{x}
5. $Q(\vec{x})$ is **negative semidefinite** if $Q(\vec{x}) \leq 0$ for all \vec{x}

Theorem 1: Let A be the symmetric matrix corresponding to a quadratic form $Q(\vec{x}) = \vec{x}^T A \vec{x}$. If P is an orthogonal matrix that diagonalizes A , then $Q(\vec{x})$ can be expressed as

$$\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

where $\vec{y} = P^T \vec{x}$ and where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A corresponding to the columns of P .

Theorem 2: If A is a symmetric matrix, then the quadratic form $Q(\vec{x}) = \vec{x}^T A \vec{x}$ is

- (1) positive definite if and only if the eigenvalues of A are all positive
- (2) negative definite if and only if the eigenvalues of A are all negative
- (3) indefinite if and only if some of the eigenvalues of A are positive and some are negative
- (4) positive semidefinite if and only if all of the eigenvalues are non-negative
- (5) negative semidefinite if and only if all of the eigenvalues are non-positive

5.4 Graphing Quadratic Forms (Optional)

Theorem 1: If $Q(\vec{x}) = a\vec{x}_1^2 + bx_1x_2 + cx_2^2$ with a, b, c not all zero, then there exists an orthogonal matrix P , which corresponds to a rotation in \mathbb{R}^2 , such that the change of variables $\vec{y} = P^T \vec{x}$ brings $Q(\vec{x})$ into diagonal form.

5.5 Optimizing Quadratic Forms

Theorem 1: Let $Q(\vec{x})$ be a quadratic form on \mathbb{R}^n with corresponding symmetric matrix A . The maximum value and minimum value of $Q(\vec{x})$ subject to the constraint $\|\vec{x}\| = 1$ are the greatest and least eigenvalues of A respectively. Moreover, these values occur when \vec{x} is taken to be a unit eigenvector corresponding to the eigenvalue.

5.6 Singular Value Decomposition

Theorem 1: If A is an $m \times n$ matrix and $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $A^T A$ with corresponding unit eigenvectors $\vec{v}_1, \dots, \vec{v}_n$, then $\lambda_1, \dots, \lambda_n$ are all non-negative. In particular,

$$\|A\vec{v}_i\| = \sqrt{\lambda_i}, \quad 1 \leq i \leq n$$

Definition- Singular Values: The singular values $\sigma_1, \dots, \sigma_n$ of an $m \times n$ matrix A are the square roots of the eigenvalues of $A^T A$ arranged so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$.

Lemma 2: If A is an $m \times n$ matrix, then $\text{Null}(A^T A) = \text{Null}(A)$.

Lemma 3: If A is an $m \times n$ matrix, then $\text{rank}(A^T A) = \text{rank}(A)$.

Corollary 4: If A is an $m \times n$ matrix and $\text{rank}(A) = r$, then A has r non-zero singular values.

Definition - Singular Vectors: Let A be an $m \times n$ matrix. If $\vec{v} \in \mathbb{R}^n$ and $\vec{u} \in \mathbb{R}^m$ are unit vectors and $\sigma \neq 0$ is a singular value of A such that

$$A\vec{v} = \sigma\vec{u} \text{ and } A^T\vec{u} = \sigma\vec{v}$$

then we say that \vec{u} is a left singular vector of A corresponding to σ and \vec{v} is a right singular vector of A corresponding to σ . For $\sigma = 0$, if \vec{u} is a unit vector such that $A^T\vec{u} = \vec{0}$, then \vec{u} is a left singular vector of A corresponding to $\sigma = 0$. If \vec{v} is a unit vector such that $A\vec{v} = \vec{0}$, then \vec{v} is a right singular vector of A corresponding to $\sigma = 0$.

Theorem 5: Let A be an $m \times n$ matrix. If $\vec{v} \in \mathbb{R}^n$ is a unit eigenvector of $A^T A$ corresponding to a non-zero singular value σ of A , then $\vec{u} = \frac{1}{\sigma}A\vec{v}$ is a left singular vector of A corresponding to σ .

Theorem 6: Let A be an $m \times n$ matrix. A vector $\vec{v} \in \mathbb{R}^n$ is a right singular vector of A if and only if \vec{v} is a unit eigenvector of $A^T A$. A vector $\vec{u} \in \mathbb{R}^m$ is a left singular vector of A if and only if \vec{u} is a unit eigenvector of AA^T .

Lemma 7: Let A be an $m \times n$ matrix with $\text{rank}(A) = r$ and singular values $\sigma_1, \dots, \sigma_n$. If $\{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthonormal basis for \mathbb{R}^n consisting of the eigenvectors of $A^T A$ arranged so that the corresponding eigenvalues are arranged from greatest to least, then

$$\left\{ \frac{1}{\sigma_1} A\vec{v}_1, \dots, \frac{1}{\sigma_r} A\vec{v}_r \right\}$$

is an orthonormal basis for $\text{Col } A$.

Theorem 8: If A is an $m \times n$ matrix with rank r and singular values $\sigma_1, \dots, \sigma_n$, then there exists an orthonormal basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ for \mathbb{R}^n of right singular vectors of A and an orthonormal basis $\{\vec{u}_1, \dots, \vec{u}_m\}$ for \mathbb{R}^m of left singular vectors of A such that

$$A\vec{v}_i = \sigma_i \vec{u}_i, \quad 1 \leq i \leq \min(m, n)$$

Definition - Singular Value Decomposition (SVD): Let A be an $m \times n$ matrix with $\text{rank}(A) = r$ and with non-zero singular values $\sigma_1, \dots, \sigma_r$. A singular values decomposition of A is a factorization of the form

$$A = U\Sigma V^T$$

where U is an orthogonal matrix containing left singular vectors of A , V is an orthogonal matrix containing right singular vectors of A , the Σ is the $m \times n$ matrix with $(\Sigma)_{ii} = \sigma_i$ for $1 \leq i \leq r$ and all other entries of Σ are 0.

Algorithm: To find a singular value decomposition of a matrix A with rank r :

- (1) Find the eigenvalues $\lambda_1, \dots, \lambda_n$ of $A^T A$ arranged from greatest to least and a corresponding set of orthonormal eigenvectors $\{\vec{v}_1, \dots, \vec{v}_n\}$. Define $\sigma_i = \sqrt{\lambda_i}$ for $1 \leq i \leq r$.
- (2) Let $V = [\vec{v}_1 \ \dots \ \vec{v}_n]$ and let Σ be the $m \times n$ matrix such that $(\Sigma)_{ii} = \sigma_i$ for $1 \leq i \leq r$ and all other entries of Σ are 0.
- (3) Find left singular vectors of A by computing $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i$ for $1 \leq i \leq r$, and then extend the set $\{\vec{u}_1, \dots, \vec{u}_r\}$ to an orthonormal basis $\{\vec{u}_1, \dots, \vec{u}_m\}$ for \mathbb{R}^m . Take $U = [\vec{u}_1 \ \dots \ \vec{u}_m]$.

Then, $A = U\Sigma V^T$.

Remark: One way to extend $\{\vec{u}_1, \dots, \vec{u}_r\}$ to an orthonormal basis for \mathbb{R}^m is by finding an orthonormal basis for $\text{Null}(A^T)$.

6 Complex Vector Spaces

6.4 Complex Inner Product

Definition - Standard Inner Product on \mathbb{C}^n : The standard inner product \langle, \rangle on \mathbb{C}^n is defined by

$$\langle \vec{z}, \vec{w} \rangle = \vec{z} \cdot \overline{\vec{w}} = z_1 \overline{w_1} + \cdots + z_n \overline{w_n}$$

for any $\vec{z}, \vec{w} \in \mathbb{C}^n$.

Remark: In some other areas like engineering, they use

$$\langle \vec{z}, \vec{w} \rangle = \overline{\vec{z}} \cdot \vec{w}$$

for the definition of the standard inner product for \mathbb{C}^n . Many computer programs also use the engineering definition of the inner product for \mathbb{C}^n

Theorem 1: If $\vec{v}, \vec{z}, \vec{w} \in \mathbb{C}^n$ and $\alpha \in \mathbb{C}$, then

- (1) $\langle \vec{z}, \vec{z} \rangle \in \mathbb{R}, \langle \vec{z}, \vec{z} \rangle \geq 0$, 0 if and only if $\vec{z} = 0$
- (2) $\langle \vec{z}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{z} \rangle}$
- (3) $\langle \vec{v} + \vec{z}, \vec{w} \rangle = \langle \vec{v}, \vec{w} \rangle + \langle \vec{z}, \vec{w} \rangle$
- (4) $\langle \alpha \vec{z}, \vec{w} \rangle = \alpha \langle \vec{z}, \vec{w} \rangle$

Definition - Hermitian Inner Product: Let \mathbb{V} be a complex vector space. A Hermitian inner product on \mathbb{V} is a function $\langle, \rangle : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{C}$ such that for all $\vec{v}, \vec{w}, \vec{z} \in \mathbb{V}$ and $\alpha \in \mathbb{C}$ we have

- (1) $\langle \vec{z}, \vec{z} \rangle \in \mathbb{R}, \langle \vec{z}, \vec{z} \rangle \geq 0$, 0 if and only if $\vec{z} = 0$
- (2) $\langle \vec{z}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{z} \rangle}$
- (3) $\langle \vec{v} + \vec{z}, \vec{w} \rangle = \langle \vec{v}, \vec{w} \rangle + \langle \vec{z}, \vec{w} \rangle$
- (4) $\langle \alpha \vec{z}, \vec{w} \rangle = \alpha \langle \vec{z}, \vec{w} \rangle$

A complex vector space with a Hermitian inner product is called a Hermitian inner product space.

Theorem 2: If \langle, \rangle is a Hermitian inner product on a complex vector space \mathbb{V} , then for all $\vec{v}, \vec{w}, \vec{z} \in \mathbb{V}$ and $\alpha \in \mathbb{C}$ we have

$$\begin{aligned} \langle \vec{z}, \vec{v} + \vec{w} \rangle &= \langle \vec{z}, \vec{v} \rangle + \langle \vec{z}, \vec{w} \rangle \\ \langle \vec{z}, \alpha \vec{w} \rangle &= \alpha \langle \vec{z}, \vec{w} \rangle \end{aligned}$$

Theorem 3: The function $\langle A, B \rangle = \text{tr}(\overline{B^T} A)$ defines a Hermitian inner product on $M_{m \times n}(\mathbb{C})$. It is called the standard inner product for $M_{m \times n}(\mathbb{C})$.

Length and Orthogonality

Definition - Length, Unit Vector: Let \mathbb{V} be a Hermitian inner product space with inner product \langle, \rangle . For any $\vec{z} \in \mathbb{V}$ we define the length of \vec{z} by

$$\|\vec{z}\| = \sqrt{\langle \vec{z}, \vec{z} \rangle}$$

If $\|\vec{z}\| = 1$, then \vec{z} is called a unit vector.

Orthogonality: Let \mathbb{V} be a Hermitian inner product space with inner product \langle, \rangle . For any $\vec{z}, \vec{w} \in \mathbb{V}$ we say that \vec{z} and \vec{w} are orthogonal if $\langle \vec{z}, \vec{w} \rangle = 0$.

Theorem 4: If \mathbb{V} is a Hermitian inner product space with inner product \langle, \rangle , then for any $\vec{z}, \vec{w} \in \mathbb{V}$ and $\alpha \in \mathbb{C}$ we have

- (1) $\|\alpha \vec{v}\| = |\alpha| \|\vec{v}\|$.
- (2) $\|\vec{z} + \vec{w}\| \leq \|\vec{z}\| + \|\vec{w}\|$
- (3) $\frac{1}{\|\vec{z}\|} \vec{z}$ is a unit vector in the direction of \vec{z}

Definition Orthogonal, Orthonormal: S is said to be orthogonal if $\langle \vec{z}_l, \vec{z}_j \rangle = 0$ for all $l \neq j$. S is said to be orthonormal if it is orthogonal and $\|\vec{z}_j\| = 1$ for $1 \leq j \leq k$.

Theorem 5: If $\{\vec{z}_1, \dots, \vec{z}_k\}$ is an orthogonal set in a Hermitian inner product space \mathbb{V} , then

$$\|\vec{z}_1 + \dots + \vec{z}_k\|^2 = \|\vec{z}_1\|^2 + \dots + \|\vec{z}_k\|^2$$

Theorem 6: If $S = \{\vec{z}_1, \dots, \vec{z}_k\}$ is an orthogonal set of non-zero vectors in a Hermitian inner product space \mathbb{V} , then S is linearly independent.

Unitary Matrix: If the columns of a matrix U form an orthonormal basis for \mathbb{C}^n , then U is called unitary.

Theorem 7: If $U \in M_{n \times n}(\mathbb{C})$, then the following are equivalent:

- (1) the columns of U form an orthonormal basis for \mathbb{C}^n
- (2) $\overline{U}^T = U^{-1}$
- (3) the rows of U form an orthonormal basis for \mathbb{C}^n

Theorem 8: If U and W are unitary matrices, then UW is a unitary matrix.

Definition - Conjugate Transpose: For $A \in M_{m \times n}(\mathbb{C})$ the conjugate transpose A^* of A is defined to be

$$A^* = \overline{A}^T$$

Remark: Observe that if A is real matrix, then $A^* = A^T$.

Theorem 9: If A, B are matrices of the appropriate size, $\vec{z}, \vec{w} \in \mathbb{C}^n$, and $\alpha \in \mathbb{C}$, then

- (1) $\langle A\vec{z}, \vec{w} \rangle = \langle \vec{z}, A^*\vec{w} \rangle$
- (2) $(A^*)^* = A$
- (3) $(A + B)^* = A^* + B^*$
- (4) $(\alpha A)^* = \overline{\alpha} A^*$
- (5) $(AB)^* = B^* A^*$

6.5 Unitary Diagonalization

Definition - Unitarily Similar: let $A, B \in M_{m \times n}(\mathbb{C})$. If there exists a unitary matrix U such that $U^* A U = B$, then we say that A and B are unitarily similar.

Definition - Unitarily Diagonalizable: A matrix $A \in M_{n \times n}(\mathbb{C})$ is said to be unitarily diagonalizable if it is unitarily similar to a diagonal matrix.

Definition - Hermitian matrix: A matrix $A \in M_{n \times n}(\mathbb{C})$ such that $A^* = A$ is called Hermitian.

Theorem 1 - Schur's Theorem: If $A \in M_{n \times n}(\mathbb{C})$, then A is unitarily similar to an upper triangular matrix T . Moreover, the diagonal entries of T are the eigenvalues of A .

Remark: Schur's Theorem shows that for every $n \times n$ matrix A there exists a unitary matrix U such that $U^*AU = T$ is upper triangular. Since $U^* = U^{-1}$ we can solve this for A to get $A = UTU^*$. This is called a Schur decomposition of A .

Theorem 2 - Spectral Theorem for Hermitian Matrices: If A is Hermitian, then A is unitarily diagonalizable.

Theorem 3: A matrix $A \in M_{n \times n}(\mathbb{C})$ is Hermitian if and only if for all $\vec{z}, \vec{w} \in \mathbb{C}^n$ we have

$$\langle A\vec{z}, \vec{w} \rangle = \langle \vec{z}, A\vec{w} \rangle$$

Theorem 4: If A is a Hermitian matrix, then

- (1) All the eigenvalues of A are real.
- (2) If λ_1 and λ_2 are distinct eigenvalues with corresponding eigenvectors \vec{v}_1 and \vec{v}_2 respectively, then \vec{v}_1 and \vec{v}_2 are orthogonal.

Remark: Since every real symmetric matrix is Hermitian, this result implies Lemma 10.2.2.

Definition - Normal Matrix: A matrix $A \in M_{n \times n}(\mathbb{C})$ is called normal if $AA^* = A^*A$.

Theorem 5 - Spectral Theorem for Normal Matrices: A matrix A is normal if and only if it is unitarily diagonalizable.

Theorem 6: If $A \in M_{n \times n}(\mathbb{C})$ is normal, then

- (1) $\|A\vec{z}\| = \|A^*\vec{z}\|$ for all $\vec{z} \in \mathbb{C}^n$
- (2) $A - \lambda I$ is normal for every $\lambda \in \mathbb{C}$
- (3) If $A\vec{z} = \lambda\vec{z}$, then $A^*\vec{z} = \bar{\lambda}\vec{z}$
- (4) If λ_1 and λ_2 are eigenvalues of A corresponding to distinct eigenvalues λ_1 and λ_2 of A , then \vec{z}_1 and \vec{z}_2 are orthogonal

7 Introduction to Fourier Analysis

7.1 Inner Product Space of Continuous Functions

Theorem 1 - Inner Product Space of Continuous Functions: Let a and b be real numbers such that $a < b$. The vector space of all functions $f : [a, b] \rightarrow \mathbb{R}$ that are continuous on the interval $[a, b]$, denoted by $C([a, b])$, is an inner product space when equipped with the inner product

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

Definition - Norm: Let a and b be real numbers such that $a < b$. The norm of a function $f(x)$ in $C([a, b])$ is defined to be

$$\|f\| = \sqrt{\langle f, f \rangle} = \sqrt{\int_a^b f(x)^2 dx}$$

Proposition 1 - Orthogonal Subsets of $C([0, 2\pi])$: For a positive integer k , the subset

$$\mathcal{B}_k = \{1, \cos(x), \sin(x), \cos(2x), \sin(2x), \dots, \cos(kx), \sin(kx)\}$$

of the inner product space $C([0, 2\pi])$ is orthogonal.

Theorem 2 - Projection of a Function:

For a positive integer k , let $\mathcal{B}_k = \{1, \cos(x), \sin(x), \dots, \cos(kx), \sin(kx)\}$, and let $\mathbb{S}_k = \text{Span}(\mathcal{B}_k)$. Then for any $f(x)$ in $C([0, 2\pi])$ we have

$$\text{proj}_{\mathbb{S}_k} f(x) = \frac{a_0}{2} + \sum_{n=1}^k (a_n \cos(nx) + b_n \sin(nx))$$

where

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(t) dt, a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nt) dt, b_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(nt) dt$$

7.2 Fourier Series

Theorem 3 - Fourier Series: For a positive integer k , let $\mathcal{B}_k = \{1, \cos(x), \sin(x), \dots, \cos(kx), \sin(kx)\}$, and let $\mathbb{S}_k = \text{Span}(\mathcal{B}_k)$. Then for any $f(x)$ in $C([0, 2\pi])$ we have

$$\text{proj}_{\mathbb{S}_k} f(x) = \frac{a_0}{2} + \sum_{n=1}^k (a_n \cos(nx) + b_n \sin(nx))$$

where the coefficients a_0, a_n and b_n , called the Fourier coefficients of $f(x)$, are defined as

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(t) dt, a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nt) dt, b_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(nt) dt$$

Furthermore,

$$\lim_{k \rightarrow \infty} \|f(x) - \text{proj}_{\mathbb{S}_k} f(x)\| = 0$$

Remark: In calculus you may have encountered *infinite series*. Quite often a function $f(x)$ in $C([0, 2\pi])$ is written in the form of an infinite series as

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

where a_0, a_n and b_n are defined as in Theorem 3. Such a representation is called the Fourier Series of $f(x)$.