Math 235 Reference

Definition 1 (Vector Space) *A vector space over* $\mathbb{F}$ *is a set* V *together with an operation* $+ : V \times V \to V$ *(vector addition) so that for all* $\mathbf{x}, \mathbf{y} \in V$, $\mathbf{x} + \mathbf{y} \in V$ *and an operation* $\cdot : \mathbb{F} \times V \to V$ *(scalar multiplication) so that for all* $s \in \mathbb{F}, \mathbf{x} \in V$, $s \cdot \mathbf{x} \in V$. *The vector space axioms are:*

1. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$
2. $\exists \mathbf{0} \in V$ *such that* $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$ *for all* $\mathbf{x} \in V$
3. $\forall \mathbf{x} \in V$, *there exists* $(-\mathbf{x}) \in V$ *such that* $\mathbf{x} + (-\mathbf{x}) = (-\mathbf{x}) + \mathbf{x} = \mathbf{0}$
4. $\forall \mathbf{x}, \mathbf{y} \in V$, $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
5. $\forall \mathbf{x} \in V$ *and* $s, t \in \mathbb{F}$, $s(t\mathbf{x}) = (st)\mathbf{x}$
6. $\forall \mathbf{x} \in V$ *and* $s, t \in \mathbb{F}$, $(s + t)\mathbf{x} = s\mathbf{x} + t\mathbf{x}$
7. $\forall \mathbf{x}, \mathbf{y} \in V$ *and* $s \in \mathbb{F}$, $s(\mathbf{x} + \mathbf{y}) = s\mathbf{x} + s\mathbf{y}$
8. $1\mathbf{x} = \mathbf{x}$ *for all* $\mathbf{x} \in V$

Theorem 1 *Let* V *be a vector space over* $\mathbb{F}$. *Then,*

1. *The zero vector in* V *is unique*
2. *The additive inverse of any vector in* V *is uniquely determined*

Definition 2 (Subspace) *Let* V *be a vector space over* $\mathbb{F}$ *and* $U \subset V$ *a subset. We call* U *a subspace of* V *if* U, *endowed with the addition and scalar multiplication from* V, *is itself a vector space over* $\mathbb{F}$.

Theorem 2 (Subspace Test) *Let* V *be a vector space over* $\mathbb{F}$ *and let* U *be a subset of* V. *Then* U *is a subspace of* V *if and only if the following three conditions hold:*

1. U *is non-empty*
2. *For all* $\mathbf{u}_1, \mathbf{u}_2 \in U$, $\mathbf{u}_1 + \mathbf{u}_2 \in U$ *(closure under addition)*
3. *For all* $\alpha \in \mathbb{F}$ *and* $\mathbf{u} \in U$, $\alpha\mathbf{u} \in U$ *(closure under scalar multiplication)*

Corollary 2.1 *Let* V *be a vector space over* $\mathbb{F}$ *and suppose that* U *is a subspace of* V. *Then* $\mathbf{0} \in U$.

Definition 3 (Span) *Let* $S = \{\mathbf{v}_1, ..., \mathbf{v}_k\}$ *be a subset of a vector space* V *over* $\mathbb{F}$. *Define the span of* S *by*

$$Span(S) = \{t_1\mathbf{v}_1 + ... + t_k\mathbf{v}_k : t_1, ..., t_k \in \mathbb{F}\}.$$

A vector of the form $t_1\mathbf{v}_1 + ... + t_k\mathbf{v}_k$ *is called a linear combination of vectors* $\mathbf{v}_1, ..., \mathbf{v}_k$. *By convention,* $Span(\emptyset) = \{\mathbf{0}\}$.

Theorem 3 *Let $S = \{\mathbf{v}_1, ..., \mathbf{v}_k\}$ be a subset of vector space V. Then Span(S) is a subspace of V.*

Definition 4 (Spanning Set) *A set of vectors $S = \{\mathbf{v}_1, ..., \mathbf{v}_k\}$ in a vector space V is a spanning set for V if Span$(S) = V$. We also say that S spans V.*

Definition 5 (Linearly Independent) *A set of vectors $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ in a vector space V is linearly independent if the only solution to the equation*

$$t_1\mathbf{v}_1 + ... + t_k\mathbf{v}_k = \mathbf{0}$$

is $t_1 = ... = t_k = 0$. Otherwise, the set is linearly dependent. By convention, the empty set is linearly independent.

Definition 6 (Basis) *A basis for a vector space V is a linearly independent subset that spans V.*

Theorem 4 *Every vector space has a basis.*

Theorem 5 *Let V be a vector space over $\mathbb{F}$ and suppose that $V = Span(\{\mathbf{v}_1, ..., \mathbf{v}_n\})$. If $\{\mathbf{u}_1, ..., \mathbf{u}_k\}$ is a linearly independent set in V, then $k \leq n$.*

Theorem 6 *Let $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ and $\mathcal{C} = \{\mathbf{u}_1, ..., \mathbf{u}_k\}$ both be bases of a vector space V. Then $k = n$.*

Definition 7 (Dimension) *The dimension of a vector space V, $\dim(V)$, is the size of any basis for V.*

Theorem 7 *Let V be an n-dimensional vetor space over $\mathbb{F}$. Then*

1. *A set of more than n vectors in V must be linearly dependent*
2. *A set of fewer than n vectors in V cannot span V*
3. *A set with exactly n vectors in V is a spanning set for V if and only if it is linearly independent*

Theorem 8 *Let V be a finite-dimensional vector space over $\mathbb{F}$ and let W be a subspace of V. Then $\dim(V) \leq \dim(W)$ with equality if and only if $W = V$.*

Theorem 9 *Let V be a vector space and let $S = \{\mathbf{v}_1, ..., \mathbf{v}_k\}$ be a subset of V and let $U = Span(S)$. Then every vector in U can be expressed in a unique way as a linear combination of the vectors in S if and only if S is linearly independent.*

Theorem 10 (Unique Representation Theorem) *Let V be a vector space and let $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ be a basis of V. Then for all $\mathbf{v} \in V$, there exist unique scalars $x_1, ..., x_n \in \mathbb{F}$ such that*

$$\mathbf{v} = x_1\mathbf{v}_1 + ... + x_n\mathbf{v}_n$$

Definition 8 (Ordered Basis) *Let V be a vector space over $\mathbb{F}$. An ordered basis for V is a basis $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ for V together with a fixed ordering.*

Definition 9 (Coordinate Vector) *Let $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ be an ordered basis for vector space V. If $\mathbf{x} \in V$ is written as $\mathbf{x} = x_1\mathbf{v}_1 + ... + x_n\mathbf{v}_n$, then the coordinate vector of $\mathbf{x}$ with respect to $\mathcal{B}$ is*

$$[\mathbf{x}]_{\mathcal{B}} = [x_1, ..., x_n]^T.$$

Theorem 11 *Let V be a vector space over $\mathbb{F}$ with ordered basis $\mathcal{B}$. Then*

1. $[\mathbf{x} + \mathbf{y}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{B}} + [\mathbf{y}]_{\mathcal{B}}$
2. $[t\mathbf{x}]_{\mathcal{B}} = t[\mathbf{x}]_{\mathcal{B}}$

for all $\mathbf{x}, \mathbf{y} \in V$ and $t \in \mathbb{F}$.

Definition 10 (Linear Transformation) *If V and W are vector spaces over $\mathbb{F}$, a function $L : V \to W$ is called a linear transformation if it satisfies*

1. $L(\mathbf{x} + \mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y})$
2. $L(t\mathbf{x}) = tL(\mathbf{x})$

for all $\mathbf{x}, \mathbf{y} \in V$ and $t \in \mathbb{F}$.

Definition 11 (Range and Kernel) *Let $L : V \to W$ be a linear transformation. The range of L is*

$$Range(L) = \{L(\mathbf{x}) \in W : \mathbf{x} \in V\}$$

and the kernel (or nullspace) of L is

$$Ker(L) = \{\mathbf{x} \in V : L(\mathbf{x}) = \mathbf{0}\}$$

Theorem 12 *Let V and W be vector spaces over $\mathbb{F}$ and let $L : V \to W$ be a linear transformation. Then*

1. $L(\mathbf{0}) = \mathbf{0}$
2. *$Range(L)$ is a subspace of W*
3. *$Ker(L)$ is a subspace of V*

Definition 12 (Rank and Nullity) *Let V and W be vector spaces over $\mathbb{F}$. The rank of a linear map $L : V \to W$ is the dimension of the range of L. The nullity of L is the dimension of the kernel of L.*

Theorem 13 (Rank-Nullity Theorem) *Let V and W be vector spaces over $\mathbb{F}$ with $\dim(V) = n$. Let $L : V \to W$ be a linear map. Then*

$$rank(L) + nullity(L) = n.$$

Theorem 14 *Let V be an n-dimensional vector space with ordered basis $\mathcal{B}$. Let W be an m-dimensional vector space with oredered basis $\mathcal{C}$. Then, for every linear map $L : V \to W$, there exists an $m \times n$ matrix A such that $[L(\mathbf{v})]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}}$ for all $\mathbf{v} \in V$. Conversely, every $m \times n$ matrix A defines a linear map $L : V \to W$ by $[L(\mathbf{v})]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}}$.*

Corollary 14.1 *Let V be a vector space with ordered basis $\mathcal{B} = \{\mathbf{b}_1, ..., \mathbf{b}_n\}$. Let W be a vector space with ordered basis $\mathcal{C} = \{\mathbf{c}_1, ..., \mathbf{c}_m\}$. Let $L : V \to W$ be a linear map. Then the $m \times n$ matrix A of the linear map L is given by*

$$ {}_{\mathcal{C}}[L]_{\mathcal{B}} = [[L(\mathbf{b}_1)]_{\mathcal{C}}, ..., [L(\mathbf{b}_n)]_{\mathcal{C}}] $$

If $L : V \to V$, then we write $[L]_{\mathcal{B}} = {}_{\mathcal{B}}[L]_{\mathcal{B}}$.

Theorem 15 *Let V, U, W be vector spaces over $\mathbb{F}$ with bases $\mathcal{B}, \mathcal{C}, \mathcal{D}$ respectively. Let $L : V \to U$ and $M : U \to W$ be linear maps. Then*

$$ {}_{\mathcal{D}}[M \circ L]_{\mathcal{B}} = {}_{\mathcal{D}}[M]_{\mathcal{C}}\,{}_{\mathcal{C}}[L]_{\mathcal{B}}. $$

Definition 13 (Column space and nullspace) *Let $A \in M_{m \times n}(\mathbb{F})$. The column space of A, $Col(A)$, is the span of the columns of A. The rank of A is the dimension of its column space. The nullspace of A, $Null(A)$, is the set of all $\mathbf{v} \in \mathbb{F}^n$ such that $A\mathbf{v} = \mathbf{0}$. The nullity of A is the dimension of its nullspace.*

Theorem 16 *Let $L : V \to W$ be a linear map with bases $\mathcal{B}$ and $\mathcal{C}$ for bases V and W. Let $A = {}_{\mathcal{C}}[L]_{\mathcal{B}}$.*

1. *$\mathbf{v} \in Ker(L)$ if and only if $[\mathbf{v}]_{\mathcal{B}} \in Null(A)$*
2. *$\mathbf{w} \in Range(L)$ if and only if $[\mathbf{w}]_{\mathcal{C}} \in Col(A)$*

Definition 14 (Change of Coordinates Matrix) *Let V be a finite dimensional vector space, and let $\mathcal{B}$ and $\mathcal{C}$ be two bases for V. The change of coordinates matrix ${}_{\mathcal{C}}\mathcal{I}_{\mathcal{B}}$ is the matrix ${}_{\mathcal{C}}[id]_{\mathcal{B}}$, where $id : V \to V$ is the identity map.*

Theorem 17 *Let V be a finite dimensional vector space with bases $\mathcal{B}$ and $\mathcal{C}$. Then ${}_{\mathcal{C}}\mathcal{I}_{\mathcal{B}} = ({}_{\mathcal{B}}\mathcal{I}_{\mathcal{C}})^{-1}$*

Theorem 18 *Let $L : V \to W$ be a linear map between two finite-dimensional vector spaces V and W. Suppose $\mathcal{B}$ and $\mathcal{B}'$ are ordered bases for V and that $\mathcal{C}$ and $\mathcal{C}'$ are ordered bases for W. Then*

$$ {}_{\mathcal{C}'}[L]_{\mathcal{B}'} = {}_{\mathcal{C}'}\mathcal{I}_{\mathcal{C}}\,{}_{\mathcal{C}}[L]_{\mathcal{B}}\,{}_{\mathcal{B}}\mathcal{I}_{\mathcal{B}'} $$

Theorem 19 (Injective and Surjective) *Let $L : V \to W$ be a linear map between vector spaces and let $\mathbf{v}_1, \mathbf{v}_2 \in V$.*

- *We say L is injective if $L(\mathbf{v}_1) = L(\mathbf{v}_2)$ implies $\mathbf{v}_1 = \mathbf{v}_2$*
- *We say L is surjective if $Range(L) = W$*

Theorem 20 *A linear map $L : V \to W$ is injective if and only if $Ker(L) = \{\mathbf{0}\}$.*

Theorem 21 *Let $L : V \to W$ be a linear map between finite-dimensional vector spaces. Then*

1. *L is injective if and only if $nullity(L) = 0$*
2. *L is surjective if and only if $rank(L) = \dim(W)$*

Definition 15 (Isomorphism) *Let $L : V \to W$ be a linear map. If L is injective and surjective, we say L is an isomorphism. We say that two vector spaces V and W are isomorphic and write $V \cong W$, if there is an isomorphism $L : V \to W$.*

Theorem 22 *A linear map $L : V \to W$ is an isomorphism if and only if there exists a unique linear map $L^{-1} : W \to V$ such that $L \circ L^{-1}(\mathbf{w}) = \mathbf{w}$ for all $\mathbf{w} \in W$ and $L^{-1} \circ L(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in V$. In this case we call L^{-1} the inverse linear map to L.*

Theorem 23 *Let $L : V \to W$ be an isomorphism. Let $\mathcal{B}$ be a basis for V and $\mathcal{C}$ a basis for W. Then ${}_{\mathcal{C}}[L]_{\mathcal{B}}$ is an invertible matrix and $({}_{\mathcal{C}}[L]_{\mathcal{B}})^{-1} = {}_{\mathcal{B}}[L^{-1}]_{\mathcal{C}}$.*

Theorem 24 *Let $L : V \to W$ be a linear map between two finite-dimensional vector spaces. Then*

1. *If $\dim(V) < \dim(W)$, L cannot be surjective*
2. *If $\dim(V) > \dim(W)$, L cannot be injective*
3. *If $\dim(V) = \dim(W)$, L is injective if and only if L is surjective*

Theorem 25 *Suuppose V and W are finite dimensional vector spaces over the same field. Then V and W are isomorphic if and only if $\dim(V) = \dim(W)$.*

Theorem 26 *Let V be an n-dimensional vector space over $\mathbb{F}$. Then $V \cong \mathbb{F}^n$.*

Definition 16 (Linear Operator) *A linear map $T : V \to W$ is a linear operator if $V = W$.*

Definition 17 (Eigenvector and Eigenvalue) *Let $L : V \to V$ be a linear operator. A non-zero vector $\mathbf{v} \in V$ such that $L(\mathbf{v}) = \lambda\mathbf{v}$ for $\lambda \in \mathbb{F}$ is called an eigenvector of L. THe number λ is called an eigenvalue of L.*

Definition 18 (Eigenspace) *Let $L : V \to V$ be a linear operator and let $\lambda \in \mathbb{F}$ be an eigenvalue of L. The eigenspace of L corresponding to λ is*

$$E_\lambda(L) = \{\mathbf{v} \in V : L(\mathbf{v}) = \lambda\mathbf{v}\}.$$

Theorem 27 *Let $L : V \to V$ be linear operator and let $\lambda \in \mathbb{F}$ be an eigenvalue of L. The eigenspace of L corresponding to λ is a subspace of V.*

Theorem 28 *Let $L : V \to V$ be a linear operator on a finite-dimensional vector space V over $\mathbb{F}$, and let $\lambda \in \mathbb{F}$ be an eigenvalue of L. If $\mathcal{B}$ is an ordered basis for V and if $A = [L]_\mathcal{B}$, then the eigenspace of A corresponding to λ is $Null(A - \lambda I)$.*

Definition 19 (Characteristic Polynomial of a Matrix) *Let $A \in M_{n\times n}(\mathbb{F})$. The characteristic polynomial of A is the polynomial in λ is given by $C_A(\lambda) = \det(A - \lambda I)$.*

Theorem 29 *Let $A \in M_{n\times n}(\mathbb{F})$. The eigenvalues of A are the values of $\lambda \in \mathbb{F}$ that are solutions to the equation $\det(A - \lambda I) = 0$. That is, they are the roots of the characteristic polynomial of A that lie in $\mathbb{F}$.*

Theorem 30 *If A is an $n \times n$ matrix with entries in $\mathbb{F}$, then the characteristic polynomial of A is a polynomial of degree n with coefficients in $\mathbb{F}$.*

Corollary 30.1 *Let $A \in M_{n\times n}(\mathbb{C})$. Then A has n (possibly repeated) eigenvalues in $\mathbb{C}$.*

Theorem 31 *Let $A \in M_{n\times n}(\mathbb{C})$. Then*

1. *The determinant of A is the product of eigenvalues of A, where each eigenvalue is repeated according to its multiplicity.*

2. *The trace of A is the sum of eigenvalues of A, where each eigenvalue is repeated according to its multiplicity.*

Definition 20 (Diagonalizable Operator) *Let V be a finite-dimensional vector space over $\mathbb{F}$. A linear operator $L : V \to V$ is diagonalizable if there exists an ordered basis $\mathcal{D}$ for V such that $[L]_\mathcal{D}$ is a diagonal matrix. We say that the basis $\mathcal{D}$ diagonalizes L.*

Theorem 32 *Let $L : V \to V$ be a linear operator and let $\mathcal{B}$ and $\mathcal{C}$ be ordered bases for V. Then*

$$[L]_\mathcal{B} = ({}_\mathcal{C}\mathcal{I}_\mathcal{B})^{-1}[L]_\mathcal{C}\,{}_\mathcal{C}\mathcal{I}_\mathcal{B}$$

Definition 21 (Similar) *If B and C are $n\times n$ matrices such that $B = P^{-1}CP$ for some invertible matrix $P \in M_{n\times n}(\mathbb{F})$, then we say B is similar to C over $\mathbb{F}$.*

Theorem 33 *Let $A, B \in M_{n\times n}(\mathbb{F})$. If A is similar to B over $\mathbb{F}$, then A and B have the same characteristic polynomial, eigenvalues, determinant, trace, rank, and nullity.*

Definition 22 (Diagonalizable Matrix) *A matrix $A \in M_{n\times n}(\mathbb{F})$ is diagonlizable over $\mathbb{F}$ if there exists an invertible matrix $P \in M_{n\times n}(\mathbb{F})$ such that $P^{-1}AP = D$ where D is a diagonal matrix.*

Theorem 34 *A $n \times n$ matrix A is diagonalizable if and only if there exists a basis $\mathcal{D} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ for $\mathbb{F}^n$ such that each $\mathbf{v}_i$ is an eigenvector for A. If such a basis $\mathcal{D}$ exists, and if we let $P = [\mathbf{v}_1, ..., \mathbf{v}_n]$ be the matrix whose columns are the vectors in $\mathcal{D}$, then*

$$P^{-1}AP = D = diag(\lambda_1, ..., \lambda_n)$$

where λ_i is the eigenvalue corresponding to the eigenvector $\mathbf{v}_i$.

Theorem 35 *Suppose $\lambda_1, ..., \lambda_k \in \mathbb{F}$ are distinct eigenvalues of a square matrix $A \in M_{n\times n}(\mathbb{F})$ with corresponding eigenvectors $\mathbf{v}_1, ..., \mathbf{v}_k$. Then $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ is linearly independent.*

Theorem 36 *If $A \in M_{n\times n}(\mathbb{F})$ has n distinct eigenvalues, then A is diagonalizable over $\mathbb{F}$.*

Definition 23 (Algebraic and Geometric Multiplicity) *Let $A \in M_{n\times n}(\mathbb{F})$ and let $\lambda \in \mathbb{F}$ be an eigenvalue of A, The algebraic multiplicity of λ is the multiplicity of λ as a root of the characteristic polynomial of A. The geometric multiplicity of λ is defined to be the dimension of the eigenspace $E_\lambda(A) = Null(A - \lambda I)$.*

Theorem 37 *Suppose $\lambda_1, ..., \lambda_k \in \mathbb{F}$ are distinct eigenvalues of a matrix $A \in M_{n\times n}(\mathbb{F})$ and let $\{\mathbf{v}_{i,1}, \mathbf{v}_{i,2}, ..., \mathbf{v}_{i,m_i}\}$ be a basis for the eigenspace corresponding to λ_i. Then the union of these bases is a linearly independent subset of $\mathbb{F}^n$.*

Theorem 38 *Let $A \in M_{n\times n}(\mathbb{F})$ and let $\lambda \in \mathbb{F}$ be an eigenvalue of A. Then*

$$1 \leq \text{geometric multiplicity of } \lambda \leq \text{algebraic multiplicity of } \lambda \leq n.$$

Theorem 39 (Diagonalizability Test) *Let $A \in M_{n\times n}(\mathbb{F})$. Suppose that the characteristic polynomial of A factors over $\mathbb{F}$ as*

$$C_A(\lambda) = (\lambda - \lambda_1)^{a_{\lambda_1}} ... (\lambda - \lambda_k)^{a_{\lambda_k}} h(\lambda)$$

where $\lambda_1, ..., \lambda_k$ are all the distinct eigenvalues in $\mathbb{F}$ with corresponding algebraic multiplicities $a_{\lambda_1}, ..., a_{\lambda_k}$ and $h(\lambda)$ is a polynomial in λ that has no roots in $\mathbb{F}$. Then A is diagonalizable over $\mathbb{F}$ if and only if $h(\lambda)$ is a constant polynomial and for all $i = 1, ..., k$,

$$\text{algebraic multiplicity of } \lambda_i = \text{geometric multiplicity of } \lambda_i$$

Theorem 40 (Diagonalization of an Operator Algorithm) *To diagonalize linear operator $L : V \to V$:*

1. *Pick any basis $\mathcal{B}$ for V and determine matrix $A = [L]_{\mathcal{B}}$*

2. *Compute and factor the characteristic polynomial $C_A(\lambda)$ to find eigenvalues $\lambda_1, ..., \lambda_k$. Let a_i denote the algebraic multiplicity of λ_i. If $a_1 + ... + a_k \neq n$, conclude that A is not diagonalizable over $\mathbb{F}$.*

3. *Otherwise, determine a basis $\mathcal{B}_i$ for each eigenspace $E_{\lambda_i}(A)$ for $i = 1, ..., k$. Let $g_i = \dim E_{\lambda_i}$ denote the geometric multiplicity of λ_i.*

4. *A is diagonalizable if and only if $a_i = g_i$ for $i = 1, ..., k$*

5. *If A is diagonalizable, then $\mathcal{D} = \mathcal{B}_1 \cup ... \cup \mathcal{B}_k$ is a basis for $\mathbb{B}^n$ consisting of eigenvectors of A¿ If P is the matrix whose columns are the vectors in $\mathcal{D}$, then $D = P^{-1}AP$ is a diagonal matrix. The diagonal entries of D are λ_1 (listed a_1 times),..., λ_k (listed a_k times). The order of eigenvalues matches the order in which their corresponding eigenvectors occur as columns in P.*

6. *To determine a basis for V that diagonalizes L, take each of the vectors in $\mathcal{D}$, view it as a coordinate vector in $\mathbb{F}^n$ with respect to the basis $\mathcal{B}$ from step 1, and thereby convert it into a vector in V. The set of all these vectors is then the desired basis for V.*

Definition 24 (Inner Product) *Let V be a vector space over $\mathbb{F}$. An inner product is a function*

$$\langle , \rangle : V \times V \to \mathbb{F}$$

such that for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $\alpha \in \mathbb{F}$,

1. $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$

2. $\alpha \langle \mathbf{v}, \mathbf{w} \rangle = \langle \alpha \mathbf{v}, \mathbf{w} \rangle$

3. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$

4. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$. *If* $\langle \mathbf{v}, \mathbf{v} \rangle = 0$, *then* $\mathbf{v} = \mathbf{0}$

A vector space alongside an inner product is called an inner product space. Property 1 is called conjugate symmetry. Property 2 and 3 are called linearity in the first argument. Property 4 is called positive definiteness.

Theorem 41 *Let V be an inner product space. For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\alpha \in \mathbb{F}$,*

1. $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$

2. $\langle \mathbf{v}, \alpha \mathbf{w} \rangle = \overline{\alpha} \langle \mathbf{v}, \mathbf{w} \rangle$

3. $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$

Theorem 42 *Every finite-dimensional vector space admits an inner product.*

Definition 25 (Norm (Length)) *Let* $\mathbf{v}$ *be a vector in an inner product space* V*. The norm (or length) of* $\mathbf{v}$ *is defined*

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

Definition 26 (Orthogonal) *Let* V *be an inner product space. We say* $\mathbf{v}$ *is orthogonal to* $\mathbf{w}$ *and write* $\mathbf{v} \perp \mathbf{w}$ *if* $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.

Theorem 43 (Pythagorean Theorem) *Let* V *be an inner product space. Suppose* $\mathbf{v} \perp \mathbf{w}$*. Then*

$$\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$$

Theorem 44 *Let* V *be an inner product space and let* $\mathbf{v}, \mathbf{w} \in V$ *such that* $\mathbf{v} \neq \mathbf{w}$*. Then* $\mathbf{v}$ *is orthogonal to*

$$\mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w}$$

Theorem 45 (Cauchy-Schwarz Inequality) *Let* V *be an inner product space. Then*

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \|\mathbf{w}\|$$

for all $\mathbf{v}, \mathbf{w} \in V$ *with equality if and only if* $\mathbf{v}$ *and* $\mathbf{w}$ *are scalar multiples of each other.*

Definition 27 (Angle) *Let* V *be a real inner product space. The angle* θ *between two non-zero vectors* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V *is defined by*

$$\cos(\theta) = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\| \|\mathbf{w}\|},$$

that is, θ *is a unique real number in the interval* $[0, \pi]$.

Theorem 46 *Let* V *be an inner product space. For all* $\mathbf{v}, \mathbf{w} \in V$ *and* $\alpha \in \mathbb{F}$,

1. $\|\alpha \mathbf{v}\| = |\alpha| \|\mathbf{v}\|$
2. $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$
3. $\|\mathbf{v}\| \geq 0$, *and* $\|\mathbf{v}\| = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$

Definition 28 (Distance) *Let* V *be an inner product space and let* $\mathbf{v}, \mathbf{w} \in V$*. The distance between* $\mathbf{v}$ *and* $\mathbf{w}$ *is defined*

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

Definition 29 (Orthogonal Set) *A set* $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ *in an inner product space is called orthogonal if* $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ *whenever* $i \neq j$.

Definition 30 (Unit Vector) *A vector* $\mathbf{v}$ *in an inner product space is a unit vector if* $\|\mathbf{v}\| = 1$.

Definition 31 (Orthonormal Set) *A set* $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ *in an inner product space is an orthonormal set if it is an orthogonal set and if each vector in the set is a unit vector.*

Definition 32 (Normalization) *Let* $\mathbf{v}$ *be a non-zero vector in an inner product space. The normalization of* $\mathbf{v}$ *is the vector*

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

Theorem 47 *Suppose* $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ *is orthogonal and* $\mathbf{v}_i \neq \mathbf{0}$ *for all* i. *Then* $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ *is linearly independent.*

Definition 33 (Orthogonal Basis) *A set* $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ *in an inner product space* V *is an orthogonal basis if it is a basis for* V *and it is an orthogonal set.*

Theorem 48 *Let* V *be an inner product space and let* $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ *be a basis for* V. *If* $\mathbf{x} \in V$ *is given by* $\mathbf{x} = x_1\mathbf{v}_1 + ... + x_n\mathbf{v}_n$, *then*

1. *If* $\mathcal{B}$ *is an orthogonal basis, then*

$$x_i = \frac{\langle \mathbf{x}, \mathbf{v}_i \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \text{ for } 1 \leq i \leq n$$

2. *If* $\mathcal{B}$ *is an orthonormal basis, then*

$$x_i = \langle \mathbf{x}, \mathbf{v}_i \rangle \text{ for } 1 \leq i \leq n$$

Definition 34 (Projection and Perpendicular) *Let* V *be an inner product space, and let* $\mathbf{v}, \mathbf{w} \in V$ *with* $\mathbf{w} \neq \mathbf{0}$. *The projection of* $\mathbf{v}$ *onto* $\mathbf{w}$ *is defined to be the vector*

$$proj_{\mathbf{w}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2}\mathbf{w}.$$

We also define the perpendicular vector of $\mathbf{v}$ *with respect to* $\mathbf{w}$ *by*

$$perp_{\mathbf{w}}(\mathbf{v}) = \mathbf{v} - proj_{\mathbf{w}}(\mathbf{v})$$

Theorem 49 (Gram-Schmidt Orthogonolization Procedure) *Let* V *be an inner product space with basis* $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$. *To obtain an orthogonal basis for* V, *define* $\mathbf{w}_1, ..., \mathbf{w}_n$ *as follows:*

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 \\
\mathbf{w}_2 &= \mathbf{v}_2 - proj_{\mathbf{w}_1}\mathbf{v}_2 \\
\mathbf{w}_3 &= \mathbf{v}_3 - proj_{\mathbf{w}_1}\mathbf{v}_3 - proj_{\mathbf{w}_2}\mathbf{v}_3 \\
&\vdots \\
\mathbf{w}_n &= \mathbf{v}_n - proj_{\mathbf{w}_1}\mathbf{v}_n - ... - proj_{\mathbf{w}_{n-1}}\mathbf{v}_n
\end{aligned}$$

Then $\{\mathbf{w}_1, ..., \mathbf{w}_n\}$ *is an orthogonal basis for* V.

Theorem 50 *Let W be a subspace of a finite-dimensional inner product space V. Then we can find an orthogonal basis $\{\mathbf{w}_1, ..., \mathbf{w}_k, \mathbf{n}_1, ..., \mathbf{n}_l\}$ for V such that $\{\mathbf{w}_1, ..., \mathbf{w}_k\}$ is an orthogonal basis for W.*

Definition 35 (Orthogonal Complement) *Let V be an inner product space and let $W \subset V$ be a subspace. The orthogonal complement of W is the set*

$$W^{\perp} = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{w} \rangle = 0 \text{ for all } \mathbf{w} \in W\}$$

Theorem 51 *Let V be an inner product space and let $W \subset V$ be a subspace.*

1. *$W^{\perp}$ is a subspace of V*
2. *$W \cap W^{\perp} = \{\mathbf{0}\}$*

Theorem 52 *Let V be a finite-dimensional inner product space and let $W \subset V$ be a subspace. Then*

1. *If $\{\mathbf{w}_1, ..., \mathbf{w}_k\}$ is a spanning set for W, then $\mathbf{v} \in W^{\perp}$ if and only if $\langle \mathbf{v}, \mathbf{w}_i \rangle = 0$ for all $1 \leq i \leq k$*
2. *If $\mathcal{B} = \{\mathbf{w}_1, ..., \mathbf{w}_k\}$ is an orthogonal basis for W, then there exists an orthogonal basis $\mathcal{C} = \{\mathbf{n}_1, ..., \mathbf{n}_k\}$ for $W^{\perp}$ such that $\mathcal{B} \cup \mathcal{C}$ is an orthogonal basis for V*
3. $\dim(V) = \dim(W) + \dim(W^{\perp})$
4. $(W^{\perp})^{\perp} = W$

Theorem 53 (Orthogonal Decomposition) *Let W be a subspace of a finite-dimensional inner product space V. Then every $\mathbf{v} \in V$ can be written as $\mathbf{v} = \mathbf{p} + \mathbf{r}$ where $\mathbf{p} \in W$ and $\mathbf{r} \in W^{\perp}$ are uniquely determined by $\mathbf{v}$. Moreover, if $\mathcal{B} = \{\mathbf{w}_1, ..., \mathbf{w}_k\}$ is an orthogonal basis for W, then $\mathbf{p}$ is given by*

$$\mathbf{p} = \sum_{i=1}^{k} proj_{\mathbf{w}_i}(\mathbf{v}).$$

Definition 36 (Projection and Perpendicular onto a Subspace) *Let V be an inner product space and let $W \subset V$ be a subspace. Let $\{\mathbf{w}_1, ..., \mathbf{w}_k\}$ be an orthogonal basis for W. Let $\mathbf{v} \in V$. The projection of $\mathbf{v}$ onto W and the perpendicular vector of $\mathbf{v}$ with respect to W are defined to be*

$$proj_W(\mathbf{v}) = \sum_{i=1}^{k} proj_{\mathbf{w}_i}(\mathbf{v}) \text{ and } perp_W(\mathbf{v}) = \mathbf{v} - proj_W(\mathbf{v}).$$

Theorem 54 *Let V be a finite-dimensional inner product space, W a subspace of V and $\mathbf{v} \in V$. Let $\mathcal{V} = \{\mathbf{w}_1, ..., \mathbf{w}_k\}$ be an orthogonal basis for W and let*

$$\mathbf{p} = \sum_{i=1}^{k} proj_{\mathbf{w}_i}(\mathbf{v}).$$

1. *If $\mathcal{C} = \{\mathbf{u}_1, ..., \mathbf{u}_k\}$ is an orthogonal basis for W, then $\sum_{i=1}^{k} proj_{\mathbf{u}_i}(\mathbf{v}) = \mathbf{p}$*
2. *For all $\mathbf{w} \in W$, $\|\mathbf{v} - \mathbf{p}\| \leq \|\mathbf{v} - \mathbf{w}\|$*
3. *If $\|\mathbf{v} - \mathbf{p}\| = \|\mathbf{v} - \mathbf{w}\|$ for some $\mathbf{w} \in W$, then $\mathbf{w} = \mathbf{p}$*

Definition 37 (Least Squares Solution) *Let $A \in M_{m\times n}(\mathbb{R})$ and $\mathbf{b} \in \mathbb{R}^m$. The vector $\mathbf{s} \in \mathbb{R}^n$ is called a least squares solution to $A\mathbf{x} = \mathbf{b}$ if it is a solution to the system $A\mathbf{x} = \mathbf{p}$ where $\mathbf{p} = proj_{Col(A)}(\mathbf{b})$.*

Theorem 55 *Let $A \in M_{m\times n}(\mathbb{R})$ and $\mathbf{b} \in \mathbb{R}^m$. The vector $\mathbf{s} \in \mathbb{R}^n$ is a least squares solution to $A\mathbf{x} = \mathbf{b}$ if and only if it is a solution to $A^T A\mathbf{x} = A^T\mathbf{b}$.*

Theorem 56 *Let $X \in M_{m\times n}(\mathbb{R})$. Then $X^T X \in M_{n\times n}(\mathbb{R})$ is invertible if and only if the columns of X are linearly independent.*

Definition 38 (Adjoint) *If $A \in M_{n\times n}(\mathbb{F})$, the adjoint of A is the matrix $\overline{A^T} \in M_{n\times n}(\mathbb{F})$. It is denoted A^*.*

Theorem 57 *Let $P \in M_{n\times n}(\mathbb{F})$. Equip $\mathbb{F}^n$ with the standard inner product. Then, the following properties are equivalent:*

- *The columns of P form an orthonormal basis for $\mathbb{F}^n$*
- $P^* = P^{-1}$
- *The rows of P form an orthonormal basis for $\mathbb{F}^n$*

Definition 39 (Unitary Matrix) *A matrix $U \in M_{n\times n}(\mathbb{F})$ is called a unitary matrix if $U^* = U^{-1}$.*

Definition 40 (Orthogonal Matrix) *A matrix $Q \in M_{n\times n}(\mathbb{R})$ is called a (real) orthogonal matrix if $Q^T = Q^{-1}$.*

Theorem 58 *Let $U \in M_{n\times n}(\mathbb{F})$ be a unitary matrix and consider $\mathbb{F}^n$ with the standard inner product. Then:*

1. *$\langle U\mathbf{v}, U\mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w} \in \mathbb{F}^n$*
2. *$\|U\mathbf{v}\| = \|\mathbf{v}\|$ for all $\mathbf{v} \in \mathbb{F}^n$*

Theorem 59 *Let $A, B \in M_{n\times n}(\mathbb{F})$. Then*

1. $(A+B)^* = A^* + B^*$
2. $(AB)^* = B^* A^*$
3. $(A^*)^* = A$
4. $(\alpha A)^* = \bar{\alpha} A^*$ *for all* $\alpha \in \mathbb{F}$

Theorem 60 (The Fundamental Property of the Adjoint of a Matrix) *Let $A \in M_{n\times n}(\mathbb{F})$. Equip $\mathbb{F}^n$ with the standard inner product. Then, for all $\mathbf{v}, \mathbf{w} \in \mathbb{F}^n$,*

$$\langle A\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, A^*\mathbf{w}\rangle$$

Theorem 61 (Schur's Triangulization Theorem) *Let $A \in M_{n\times n}(\mathbb{C})$. There is a unitary matrix $U \in M_{n\times n}(\mathbb{C})$ and an upper-triangular matrix $T \in M_{n\times n}(\mathbb{C})$ such that*

$$U^*AU = T = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n \end{bmatrix}.$$

The diagonal entries of T are the complex eigenvalues of A repeated according to multiplicity. If A has real coefficients and all n eigenvalues of A are in $\mathbb{R}$, then U and T can be chosen to have real coefficients too.

Theorem 62 (Cayley-Hamilton Theorem) *Let $A \in M_{n\times n}(\mathbb{C})$. Then $C_A(A) = \mathbf{0}_{n\times n}$ where $C_A(\lambda)$ is the characteristic polynomial of A.*

Definition 41 (Orthogonally Diagonalizable) *A matrix $A \in M_{n\times n}(\mathbb{R})$ is said to be orthogonally diagonalizable if there is a real orthogonal matrix $Q \in M_{n\times n}(\mathbb{R})$ such that Q^TAQ is diagonal.*

Definition 42 (Unitarily Diagonalizable) *A matrix $A \in M_{n\times n}(\mathbb{C})$ is said to be unitarily diagonalizable if there is a unitary matrix $U \in M_{n\times n}(\mathbb{C})$ such that U^*AU is diagonal.*

Theorem 63 (Criterion for Orthogonal and Unitary Diagonalizability) *Let $A \in M_{n\times n}(\mathbb{R})$.*

- *A is orthogonally diagalizable if and only if there is a basis for $\mathbb{R}^n$ consisting of orthonormal eigenvectors of A*
- *A is unitarily diagaonalizable if and only if there is a basis for $\mathbb{C}^n$ consisting of orthonormal eigenvectors of A*

Definition 43 (Normal, Self-adjoint, Symmetric) *A matrix $A \in M_{n\times n}(\mathbb{F})$ is said to be normal if $AA^* = A^*A$. A matrix $A \in M_{n\times n}(\mathbb{F})$ is said to be self-adjoint if $A = A^*$. Note that if $A \in M_{n\times n}(\mathbb{R})$ and is self-adjoint, then in fact we have $A = A^T$ and we say that A is (real) symmetric.*

Theorem 64 (Spectral Theorem for Self-adjoint Matrices) *A square matrix in $M_{n\times n}(\mathbb{C})$ is self-adjoint if and only if it is unitarily diagaonalizable and if its eigenvalues are all real.*

Theorem 65 (Spectral Theorem for Symmetric Matrices) *A square matrix in $M_{n\times n}(\mathbb{R})$ is symmetric if and only if it is orthogonally diagaonalizable.*

Theorem 66 (Spectral Theorem for Normal Matrices) *A square matrix in $M_{n\times n}(\mathbb{C})$ is normal if and only if it is unitarily diagonalizable.*

Theorem 67 *Let $A \in M_{n\times n}(\mathbb{F})$ be a normal matrix. Equip $\mathbb{F}^n$ with the standard inner product. Then*

1. *For all $\mathbf{x} \in \mathbb{F}^n$, $\|A\mathbf{x}\| = \|A^*\mathbf{x}\|$*
2. *If $\mathbf{x} \in \mathbb{F}^n$ is an eigenvector for A with eigenvalue λ, then $\mathbf{x}$ is an eigenvector for A^* with eigenvalue $\bar{\lambda}$*
3. *If $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{F}^n$ are eigenvectors of A with distinct eigenvalues λ and μ, then $\mathbf{x}$ is orthogonal to $\mathbf{y}$*

Theorem 68 (Algorithm for Unitary Diagonalization of a Normal Matrix) *To unitarily diagonalize a normal matrix $A \in M_{n\times n}(\mathbb{F})$:*

1. *Diagonalize A as usual obtaining $D = diag(\lambda_1, ..., \lambda_n)$ and a basis of eigenvectors for $\mathbb{F}^n$*
2. *Perform the Gram-Schmidt procedure on the bases of each of the eigenspaces E_{λ_i} of A, obtaining orthonormal bases for the eigenspaces*
3. *Take the union of the orthonormal bases to obtain $\mathcal{D} = \{\mathbf{w}_1, ..., \mathbf{w}_n\}$ which is an orthonormal basis of eigenvectors for $\mathbb{F}^n$. Order the baseis so that the orthonormal basis for E_{λ_1} is followed by the orthonormal basis for E_{λ_2}, etc.*
4. *Let $U = [\mathbf{w}_1, ..., \mathbf{w}_n]$. Then U is unitary and $U^*AU = D$ with D as in step 1.*

Theorem 69 (Characterization of Inner Products in Terms of Matrices) *Let V be a vector space over $\mathbb{F}$ with basis $\mathcal{B} = \{\mathbf{g}_1, ..., \mathbf{g}_n\}$, and let $A \in M_{n\times n}(\mathbb{F})$. Then*

$$\langle \mathbf{v}, \mathbf{w} \rangle = [\mathbf{w}]_{\mathcal{B}}^* A [\mathbf{v}]_{\mathcal{B}}$$

where $\mathbf{v}, \mathbf{w} \in V$ defines an inner product on V if and only if $A = A^*$ and all the eigenvalues of A are positive. Furthermore, if $\langle , \rangle$ is an inner product of V, then there is a self-adjoint matrix A such that $\langle \mathbf{v}, \mathbf{w} \rangle = [\mathbf{w}]_{\mathcal{B}}^* A [\mathbf{v}]_{\mathcal{B}}$ for all $\mathbf{v}, \mathbf{w} \in V$. Explicitly, this matrix is given by

$$A = \begin{bmatrix} \langle \mathbf{g}_1, \mathbf{g}_1 \rangle & \cdots & \langle \mathbf{g}_n, \mathbf{g}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{g}_1, \mathbf{g}_n \rangle & \cdots & \langle \mathbf{g}_n, \mathbf{g}_n \rangle \end{bmatrix}$$

Definition 44 (Gram Matrix) *Let V be a finite-dimensional inner product space over $\mathbb{F}$ with inner product $\langle,\rangle$ and basis $\mathcal{B} = \{\mathbf{g}_1, ..., \mathbf{g}_n\}$. The Gram matrix of $\langle,\rangle$ with respect to $\mathcal{B}$ is the matrix*

$$A = \begin{bmatrix} \langle \mathbf{g}_1, \mathbf{g}_1 \rangle & ... & \langle \mathbf{g}_n, \mathbf{g}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{g}_1, \mathbf{g}_n \rangle & ... & \langle \mathbf{g}_n, \mathbf{g}_n \rangle \end{bmatrix}$$

Theorem 70 (Spectral Theorem for Operators) *Let $L : V \to V$ be a linear operator on a finite-dimensional inner product space over $\mathbb{F}$. Let $\mathcal{B}$ be an orthonormal basis for V and let $A = [L]_{\mathcal{B}}$. Then*

1. *If $\mathbb{F} = \mathbb{C}$, there is an orthonormal basis for V consisting of eigenvectors of L if and only if A is normal*

2. *IF $\mathbb{F} = \mathbb{R}$, there is an orthonormal basis for V consisting of eigenvectors of L if and only if A is symmetric*

Definition 45 (Quadratic Form) *A (real) quadratic form in the variables $\mathbf{u} = (u_1, ..., u_n)$ is a polynomial of the form*

$$Q(\mathbf{u}) = \sum_{i,j=1}^{n} a_{ij} u_i u_j$$

where $a_{ij} \in \mathbb{R}$.

Definition 46 (Definiteness) *A quadratic form $Q(\mathbf{u})$ is said to be*

- *Positive definite if $Q(\mathbf{u}) > 0$ for all non-zero $\mathbf{u} \in \mathbb{R}^n$*
- *Positive semi-definite if $Q(\mathbf{u}) \geq 0$ for all $\mathbf{u} \in \mathbb{R}^n$*
- *Negative definite if $Q(\mathbf{u}) \leq 0$ for all $\mathbf{u} \in \mathbb{R}^n$*
- *Negative semi-definite if $Q(\mathbf{u}) \leq 0$ for all $\mathbf{u} \in \mathbb{R}^n$*
- *Indefinite if there exist $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ such that $Q(\mathbf{u}) > 0$ and $Q(\mathbf{v}) < 0$*

Theorem 71 *Let $Q(\mathbf{u}) = \sum_{i,j=1}^{n} a_{ij} u_i u_j$ be a quadratic form. If we let A be the $n \times n$ matrix whose (i, j)th entry is $\frac{a_{ij}+a_{ji}}{2}$, then*

$$Q(\mathbf{u}) = \mathbf{u}^T A \mathbf{u}.$$

Definition 47 (Matrix Associated to a Quadratic Form) *The matrix $A \in M_{n\times n}(\mathbb{R})$ in the previous theorem is called the matrix associated to the quadratic form $Q(\mathbf{u})$.*

Theorem 72 *Let $A \in M_{n\times n}(\mathbb{R})$ be the matrix associated to the quadratic form $Q(\mathbf{u})$. Then A is symmetric.*

Theorem 73 (Classification of Quadratic Forms) *Let $Q(\mathbf{u})$ be a quadratic form with associated matrix A. Let $\lambda_1, ..., \lambda_n$ be the eigenvalues of A. Then*

1. *$Q(\mathbf{u})$ is positive definite if and only if $\lambda_i > 0$ for all $1 \leq i \leq n$*
2. *$Q(\mathbf{u})$ is positive semi-definite if and only if $\lambda_i \geq 0$ for all $1 \leq i \leq n$*
3. *$Q(\mathbf{u})$ is negative definite if and only if $\lambda_i < 0$ for all $1 \leq i \leq n$*
4. *$Q(\mathbf{u})$ is negative semi-definite if and only if $\lambda_i \leq 0$ for all $1 \leq i \leq n$*
5. *$Q(\mathbf{u})$ is indefinite if and only if $\lambda_i > 0$ and $\lambda_j < 0$ for some $1 \leq i, j \leq n$*

Theorem 74 *Let $A \in M_{n\times n}(\mathbb{F})$. Then A^*A is an $n \times n$ self-adjoint matrix and its eigenvalues are non-negative real numbers.*

Definition 48 (Singular Values and Vectors) *Let $A \in M_{n\times n}(\mathbb{F})$. The singular values of A are the non-negative square-roots $\sigma_i = \sqrt{\lambda_i}$ of the eigenvalues λ_i of A^*A. The corresponding eigenvectors of A^*A are called the singular vectors of A.*

Theorem 75 *Let $A \in M_{n\times n}(\mathbb{F})$ and let $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ be singular vectors of A corresponding to the singular values σ_1 and σ_2. If $\sigma_1 \neq \sigma_2$, then $\mathbf{x}$ and $\mathbf{y}$ are orthogonal.*

Theorem 76 *Let $A \in M_{n\times n}(\mathbb{F})$. Then $Null(A^*A) = Null(A)$.*

Theorem 77 *Let $A \in M_{n\times n}(\mathbb{F})$. The number of non-zero singular values of A is equal to $rank(A)$ where each repeated singular value is counted according to its multiplicity.*

Theorem 78 (Singular Value Decomposition of Matrices) *Let $A \in M_{n\times n}(\mathbb{F})$ be a matrix of rank r with non-zero singular values $\sigma_1 \geq ... \geq \sigma_r > 0$. Then there exist unitary matrices $U \in M_{m\times m}(\mathbb{F})$ and $V \in M_{n\times n}(\mathbb{F})$ such that*

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

where Σ is the $m \times n$ matrix whose entries are

$$\Sigma_{ij} = \begin{cases} \sigma_i & i = j \leq r \\ 0 & otherwise \end{cases}.$$

If A is real, then U and V can be chosen to be orthogonal matrices.

Definition 49 (Singular Value Decomposition) *A decomposition $A = U\Sigma V^*$ of the type occurring in the previous theorem is called a singular value decomposition of A.*

Theorem 79 (Algorithm for Finding an SVD for a Matrix) *Let $A \in M_{m\times n}(\mathbb{F})$ be a matrix of rank r. To find an SVD for A:*

1. *Find eigenvalues $\lambda_1 \leq ... \leq \lambda_n$ and a corresponding set of orthonormal eigenvectors $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ for A^*A*
2. *Set $\sigma_i = \sqrt{\lambda_i}$ for $i \leq r$*
3. *Set $\mathbf{u}_i = \frac{1}{\sigma_i}A\mathbf{v}_i$ for $i \leq r$. If $r < m$, extend $\{\mathbf{u}_1, ..., \mathbf{u}_r\}$ to an orthonormal basis $\{\mathbf{u}_1, ..., \mathbf{u}_r, \mathbf{u}_{r+1}, ..., \mathbf{u}_m\}$ of $\mathbb{F}^m$*
4. *Set $V = [\mathbf{v}_1, ..., \mathbf{v}_n]$ and $U = [\mathbf{u}_1, ..., \mathbf{u}_m]$ and let Σ be the $m \times n$ matrix defined in Theorem 78.*

Then U and V are unitary square matrices and $A = U\Sigma V^*$.

Theorem 80 *Let $A \in M_{m\times n}(\mathbb{F})$. Then $Col(A)^\perp = Null(A^*)$.*

Theorem 81 (Singular Value Decomposition of Linear Maps) *Let $L : W_1 \to W_2$ be a linear map between finite-dimensional inner product spaces of dimensions n and m respectively. If $r = rank(A)$, there exists orthonormal bases $\mathcal{B}$ and $\mathcal{C}$ for W_1 and W_2 and an $r \times r$ diagonal matrix D such that*

$$_{\mathcal{C}}[L]_{\mathcal{B}} = \begin{bmatrix} D & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix}$$

Definition 50 (Compact SVD) *Let $A \in M_{m\times n}(\mathbb{F}$ be a rank r matrix with singular value decomposition $A = U\Sigma V^*$. Let $r = rank(A)$; Σ will have r non-zero entries along its diagonal. By deleting all 0 rows and columns from Σ, we get Σ_r. Let U_r and V_r be the matrices formed from the first r columns of U and V. Then*

$$A = U_r\Sigma_r V_r^*$$

is called a compact singular value decomposition of A.

Theorem 82 *Let $A \in M_{m\times n}(\mathbb{F}$ have rank r and compact SVD $A = U_r\Sigma_r V_r^*$ where $\Sigma_r = diag(\sigma_1, ..., \sigma_r)$ and $U = [\mathbf{u}_1, ..., \mathbf{u}_r]$ and $V = [\mathbf{v}_1, ..., \mathbf{v}_r]$. Then*

$$A = \sigma_1\mathbf{u}_1\mathbf{v}_1^* + ... + \sigma_r\mathbf{u}_r\mathbf{v}_r^*.$$

Definition 51 (Rank k Truncation) *Let $A \in M_{m\times n}(\mathbb{F})$ be a rank r matrix with singular values $\sigma_1 \geq ... \geq \sigma_r > 0$ and compact SVD $A = U_r\Sigma_r V_r^*$ where $U = [\mathbf{u}_1, ..., \mathbf{u}_r]$ and $V = [\mathbf{v}_1, ..., \mathbf{v}_r]$. Let $k \leq r$ be a positive integer. The rank k truncation of A is*

$$A = \sigma_1\mathbf{u}_1\mathbf{v}_1^* + ... + \sigma_k\mathbf{u}_k\mathbf{v}_k^*.$$

Theorem 83 *Let $A \in M_{m\times n}(\mathbb{F})$ be a matrix of rank r with rank-k truncation*

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^* + ... + \sigma_k \mathbf{u}_k \mathbf{v}_k^*.$$

Then:

1. $A = A_r$
2. $rank(A_k) = k$
3. $\|A - A_k\| \leq \sum_{i=k+1}^{r} \sigma_i$ *where* $\|X\| = \sqrt{tr(X^*X)}$ *is the norm induced by the Frobenius inner product.*

Theorem 84 (Eckart-Young Theorem) *Let $A \in M_{m\times n}(\mathbb{F})$ and let A_k be the rank-k truncation of A. Let $B \in M_{m\times n}(\mathbb{F})$ be an arbitrary rank k matrix. Then*

$$\|A - B\| \geq \|A - A_k\|.$$

Definition 52 (Pseudoinverse) *Let $A \in M_{m\times n}(\mathbb{F})$ be a matrix of rank r with SVD $A = U\Sigma V^*$. The pseudoinverse of A is the $n \times m$ matrix*

$$A^\dagger = V\Sigma^\dagger U^*$$

where $\Sigma^\dagger$ is the $n \times m$ matrix whose (i,j)th entry is $\frac{1}{\sigma_i}$ for $i = j \leq r$ and 0 otherwise.

Theorem 85 (Minimal Norm Solutions) *Consider the system of linear equations $A\mathbf{x} = \mathbf{b}$ where $A \in M_{m\times n}(\mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^m$. Let $\mathbf{x}_0 = A^\dagger \mathbf{b}$. Then*

1. *If $A\mathbf{x} = \mathbf{b}$ is consistent, then $\mathbf{x}_0$ is a solution to the system. Moreover, it is the solution of minimal norm; if $\mathbf{x}$ is any solution to the system, $\|\mathbf{x}\| \geq \|\mathbf{x}_0\|$ with equality if and only if $\mathbf{x} = \mathbf{x}_0$*
2. *If $A\mathbf{x} = \mathbf{b}$ is inconsistent, then $\mathbf{x}_0$ is a least squares solution. Moreover, it is the least squares solution of minimal norm.*