MATH 223: Linear Algebra

McGill University

Instructor: Prof. Jeremy Macdonald
Notes by: C. A.

Condensed Notes
July 1, 2025

$$
\begin{tikzcd}
V_\beta \arrow[r, "{[T]_\beta}"] \arrow[d, "Q_\beta^\alpha"'] & V_\beta \\
V_\alpha \arrow[r, "{[T]_\alpha}"] & V_\alpha \arrow[u, "Q_\alpha^\beta"']
\end{tikzcd}
$$

The Similarity Transform (Theorem 68)

Contents

1 Complex Numbers

1.1 Basics of $\mathbb{C}$

Definition 1 (Field of Complex Numbers).
The field of complex numbers, $\mathbb{C}$, consists of all expressions of the form

$$a + b\,i,$$

where $a, b \in \mathbb{R}$ and i is a symbol satisfying $i^2 = -1$.

Definition 2 (Addition and Multiplication).
Let $z = a_1 + b_1\,i$ and $\omega = a_2 + b_2\,i$. Define

$$z+\omega = (a_1+a_2)+(b_1+b_2)\,i, \qquad z\,\omega = (a_1+b_1 i)(a_2+b_2 i) = (a_1a_2-b_1b_2)+(a_1b_2+a_2b_1)\,i.$$

Definition 3 (Conjugate and Modulus).
For $z = a + b\,i$,

$$\overline{z} = a - b\,i \quad \text{(complex conjugate)}, \qquad |z| = \sqrt{a^2+b^2} \quad \text{(absolute value or modulus)}.$$

Proposition 4.
If $z = a + b\,i \neq 0$ (i.e. $z \neq 0 + 0i$), then the number

$$z^{-1} \;=\; \frac{\overline{z}}{|z|^2} \;=\; \frac{a}{a^2+b^2} \;-\; \frac{b}{a^2+b^2}\,i$$

is called the (multiplicative) inverse *of z and satisfies*

$$z\,z^{-1} \;=\; 1 \;=\; z^{-1}\,z.$$

Definition 5 (Division).
For $z, \omega \in \mathbb{C}$ with $\omega \neq 0$, define

$$\frac{z}{\omega} \;=\; z\,\omega^{-1}.$$

Proposition 6.
Let $z, w \in \mathbb{C}$.

1. $\overline{z+w} = \overline{z} + \overline{w}$.
2. $\overline{z\,w} = \overline{z}\,\overline{w}$.
3. $\overline{\overline{z}} = z$.
4. $z\,\overline{z} = |z|^2$.
5. $z \in \mathbb{R} \iff \overline{z} = z$.

1.2 Polar Form and Equation

Definition 7 (Complex Exponential Function).
In the complex plane, $e^{i\theta}$ is the unique number of modulus 1 and argument θ. Equivalently,

$$e^{i\theta} = \cos\theta + i\sin\theta,$$

(Euler's formula).

Theorem 8 (Fundamental Theorem of Algebra).
Let

$$p(z) \;=\; a_n\, z^n \;+\; a_{n-1}\, z^{n-1} \;+\; \cdots +\; a_1\, z \;+\; a_0,$$

where $a_i \in \mathbb{C}$. Then $p(z)$ factors into linear factors:

$$p(z) \;=\; a_n\,(z - r_1)\,(z - r_2)\,\cdots\,(z - r_n),$$

where each $r_i \in \mathbb{C}$ (the r_i may repeat).

2 Vector Spaces

2.1 Vector Space Axioms

Definition 9 (Field).
A field *is a set F equipped with two operations $+$ and $\cdot$ such that:*

- *$(F, +)$ is an abelian group.*
- *$(F \setminus \{0\}, \cdot)$ is an abelian group.*
- *Multiplication distributes over addition: for all $a, b, c \in F$,*

$$a\,(b + c) = ab + ac, \qquad (a + b)\,c = ac + bc.$$

Examples include $\mathbb{Q}, \mathbb{R}, \mathbb{C}$. $\mathbb{Z}$ is not *a field since, e.g., $\frac{2}{3} \notin \mathbb{Z}$.*

Definition 10 (Vector Space).
Let F be a field and V a set. A vector space over F *is a pair $(V, +)$ together with a scalar multiplication $F \times V \to V$ satisfying the following eight axioms for all $u, v, w \in V$ and $a, b \in F$:*

1. *$u + v = v + u$.*
2. *$(u + v) + w = u + (v + w)$.*
3. *There exists $0 \in V$ with $u + 0 = 0 + u = u$.*
4. *For each u there is $-u$ with $u + (-u) = (-u) + u = 0$.*

5. $a\,(u+v) = au + av$.

6. $(a+b)\,u = au + bu$.

7. $a\,(b\,u) = (ab)\,u$.

8. $1\,u = u$, *where* 1 *is the multiplicative identity in* F.

Proposition 11 (Basic Vector Space Properties).
Let V *be a vector space over a field* F.

1. *For all* $u, v, w \in V$, *if* $u + w = v + w$ *then* $u = v$.

2. *The zero vector* $\vec{0}$ *in* V *is unique.*

3. *For each* $u \in V$, *its additive inverse* $-u$ *is unique.*

4. *For all* $u \in V$, $0\,u = \vec{0}$.

5. *For all* $c \in F$, $c\,\vec{0} = \vec{0}$.

6. *For all* $c \in F$ *and* $u \in V$, $(-c)u = c(-u) = -(cu)$.

2.2 Linear Combinations and Subspaces

Definition 12 (Linear Combination).
Given vectors $u_1, u_2, \ldots, u_m \in V$ *and scalars* $c_1, \ldots, c_m \in F$, *any vector of the form*

$$c_1\,u_1 + c_2\,u_2 + \cdots + c_m\,u_m$$

is called a linear combination *of* $u_1, \ldots, u_m$.

Definition 13 (Span).
Let $S = \{u_1, u_2, \ldots, u_m\} \subseteq V$. *The* span *of* S *is*

$$\operatorname{span}(S) \;=\; \left\{\, c_1 u_1 + c_2 u_2 + \cdots + c_m u_m \;\middle|\; c_i \in F \right\}.$$

If $S = \varnothing$, *we define* $\operatorname{span}(\varnothing) = \{0\}$.

Proposition 14.
If $A, B \in M_{m\times n}(F)$ *and* B *is obtained from* A *by elementary row operations (EROs), then*

$$\operatorname{row}(A) = \operatorname{row}(B).$$

Proposition 15 (Facts About Spans).
Let $S \subseteq V$. *Then:*

1. *For all* $u, w \in \operatorname{span}(S)$, $u + w \in \operatorname{span}(S)$ *(*closure under addition*).*

2. *For all* $u \in \operatorname{span}(S)$ *and* $c \in F$, $cu \in \operatorname{span}(S)$ *(*closure under scalar multiplication*).*

3. $\vec{0} \in \operatorname{span}(S)$.

Definition 16 (Subspace).
Let V *be a vector space over a field* F*, and let* $W \subseteq V$*. We say* W *is a* subspace *of* V *(and write* $W \leq V$*) if:*

1. *For all* $w_1, w_2 \in W$, $w_1 + w_2 \in W$.

2. *For all* $w \in W$ *and all scalars* $c \in F$, $c\,w \in W$.

3. *The zero vector* $0 \in W$.

Theorem 17.
Let $A \in M_{m\times n}(F)$, $b \in F^m$*, and let* $x \in F^n$ *be the vector of variables. Let* S *be the set of all solutions to the linear system* $A\,x = b$*. Then* S *is a subspace of* F^n *if and only if*

$$b = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \vec{0} \quad (\textit{i.e. the system is homogeneous}).$$

Proposition 18.
Subspaces are closed under forming linear combinations. If $W \leq V$*, then for any positive integer* n*, if* $w_1, w_2, \ldots, w_n \in W$ *and* $c_1, \ldots, c_n \in F$*, then*

$$c_1\,w_1 + c_2\,w_2 + \cdots + c_n\,w_n \;\in\; W.$$

Proposition 19 (Spans Are Subspaces).
Let $S \subseteq V$ *(a subset) and* $W \leq V$ *(a subspace). Then:*

1. $S \subseteq \operatorname{span}(S)$.

2. *If* $S \subseteq W$*, then* $\operatorname{span}(S) \subseteq W$.

3. $\operatorname{span}(W) = W$.

2.3 Linear Independence and Dependence

Definition 20 (Linear Dependence).
Let V *be a vector space over* F *and* $S \subseteq V$*. The set* S *is* linearly dependent *if there exist distinct vectors* $u_1, \ldots, u_n \in S$ *and scalars* $c_1, \ldots, c_n \in F$*, not all zero, such that*

$$c_1u_1 + c_2u_2 + \cdots + c_nu_n \;=\; 0.$$

Otherwise, S *is* linearly independent*.*

Definition 21 (Linear Independence).
Let V be a vector space over a field F and let $S \subseteq V$. The set S is said to be linearly independent *if whenever distinct vectors $u_1, \ldots, u_n \in S$ and scalars $c_1, \ldots, c_n \in F$ satisfy*

$$c_1u_1 + c_2u_2 + \cdots + c_nu_n \;=\; 0,$$

it follows that $c_1 = c_2 = \cdots = c_n = 0$.

Proposition 22 (Dependency Special Cases).

1. *The empty set $\varnothing$ is linearly independent.*
2. *Let $S \subseteq V$. If $\vec{0} \in S$, then S is dependent (since $1 \cdot \vec{0} = \vec{0}$ provides a nontrivial dependence).*
3. *Let $u \in V$. Then $\{u\}$ is independent if and only if $u \neq \vec{0}$. Equivalently, $\{u\}$ is dependent if and only if $u = \vec{0}$.*
4. *Let $A \subseteq B \subseteq V$. Then:*
   (a) *If A is dependent, then B is also dependent.*
   (b) *If B is independent, then A is also independent.*

Lemma 23 (Extending a Linearly Independent Set).
Let $S \subseteq V$ be a linearly independent set, and let $w \in V$ with $w \notin S$. Then $S \cup \{w\}$ is independent if and only if $w \notin \operatorname{span}(S)$. (Adding a vector already in the span of S makes the set dependent.)

2.4 Basis and Dimension

Definition 24 (Basis).
Let V be a vector space over F and let $W \leq V$. A subset $\beta \subseteq W$ is called a basis *of W if*

1. $\operatorname{span}(\beta) = W$,
2. *β is linearly independent.*

Theorem 25 ("Bases Exist").
Let $W \leq V$, and suppose $W = \operatorname{span}(S)$ for some finite *set S. Then there exists a subset $\beta \subseteq S$ such that β is a basis of W. (Any finite spanning set can be reduced to a basis.)*

Theorem 26 ("All Bases Have the Same Size").
Let $W = \operatorname{span}(S)$ with S finite. Then W has a finite basis, and all *bases of W have the same cardinality. This common number is called $\dim(W)$.*

Definition 27 (Dimension).
A vector space V (or subspace of one) is called finite–dimensional *if it admits a finite basis. The* dimension *of V, written* $\dim(V)$*, is the number of vectors in any basis of V. If no finite basis exists, then V is* infinite–dimensional.

Theorem 28.
Every vector space (even one without a finite spanning set) has a basis.

Proposition 29.
If we find the general solution to $A\vec{x} = \vec{0}$ as

$$\vec{x} \;=\; t_1\,\vec{v}_1 \;+\; t_2\,\vec{v}_2 \;+\; \cdots +\; t_r\,\vec{v}_r,$$

where $t_1, \ldots, t_r$ are the free variables, then $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r\}$ forms a basis for $\mathrm{null}(A)$.

Theorem 30.
Suppose $\dim(W) = n$ is finite, and let $S \subseteq W$.

1. *If $W = \mathrm{span}(S)$, then $|S| \geq n$, and there is a subset $\beta \subseteq S$ such that β is a basis of W. (Any spanning set can be "shrunk" to a basis.)*

2. *If S is independent, then $|S| \leq n$, and there is a basis β of W such that $S \subseteq \beta$. (Any independent set can be "extended" to a basis.)*

3. *Let $|S| = n$. If and only if $W = \mathrm{span}(S)$, then S is independent.*

2.5 Subspaces

Proposition 31.
Let $W \leq V$, where $\dim(V) = n$ is finite.

1. $\dim(W) \leq n$.

2. $\dim(W) = n \iff W = V$.

Definition 32 (Column Space).
Let $A \in M_{m\times n}(F)$. The column space *of A, denoted* $\mathrm{col}(A)$*, is the subspace of F^m spanned by the columns of A:*

$$\mathrm{col}(A) = \mathrm{span}\{\textit{columns of } A\} \;\leq F^m.$$

Theorem 33.
Let $A \in M_{m\times n}(F)$, and let R be the row-reduced echelon form of A.

1. *A basis for* $\mathrm{row}(A)$ *is given by the nonzero rows of R.*

2. *The columns of A that correspond to the leading entries (pivots) in R form a basis for* $\mathrm{col}(A)$.

Theorem 34.
Let $U \leq V$ and $W \leq V$ be subspaces of V. Then

$$U \cap W = \{ v \in V : v \in U \text{ and } v \in W \}$$

is a subspace of V.

Definition 35 (Sum of Subspaces).
Let V be a vector space over a field F and let $U, W \leq V$. The sum *of U and W is the subspace*

$$U + W = \{ u + w \mid u \in U,\ w \in W \} \leq V.$$

Proposition 36.
Let U, W be subspaces of V. Then:

1. $U + W = \operatorname{span}(U \cup W)$.
2. $U \leq U + W$ *and* $W \leq U + W$.

Definition 37 (Direct Sum).
Suppose $U, W \leq V$ are subspaces such that every $v \in V$ can be written uniquely as

$$v = u + w \quad \text{with } u \in U,\ w \in W.$$

Then V is called the direct sum *of U and W, and we write*

$$V = U \oplus W.$$

Proposition 38.
Let U, W be subspaces of V. Then

$$V = U \oplus W \iff \big(V = U + W\big) \text{ and } \big(U \cap W = \{\vec{0}\}\big).$$

Theorem 39 (Inclusion-Exclusion Theorem).
Let U, W be finite-dimensional subspaces of V. Then

$$\dim(U + W) = \dim(U) + \dim(W) - \dim\big(U \cap W\big).$$

2.6 Lagrange Interpolation

Definition 40 (Lagrange Polynomials).
Let $a_0, \ldots, a_n \in \mathbb{R}$ be distinct. For each $i = 0, 1, \ldots, n$, the Lagrange polynomial $\ell_i(x)$ *is defined by*

$$\ell_i(x) = \prod_{\substack{0 \leq j \leq n \\ j \neq i}} \frac{x - a_j}{a_i - a_j}.$$

Each $\ell_i(x)$ has degree n and satisfies $\ell_i(a_j) = \delta_{ij}$.

Definition 41 (Kronecker Delta).
The Kronecker delta δ_{ij} *is defined for any indices* i, j *by*

$$\delta_{ij} \;=\; \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Proposition 42.

1. *If* $l_i(a_j) = \delta_{i,j}$*, then* $\delta_{i,j} = \begin{cases} 1 & \text{if } j = i, \\ 0 & \text{if } j \neq i, \end{cases}$ *(Kronecker delta).*

2. *The polynomials* $l_0(x)$*,* $l_1(x), \ldots, l_n(x)$ *form a basis of* $P_n(\mathbb{R})$*.*

3 Linear Transformations

3.1 Definition and Basic Properties

Definition 43 (Linear Transformation).
Let U *and* V *be vector spaces over a field* F*, and let* $T\colon U \to V$ *be a function. If for all* $u_1, u_2 \in U$ *and all* $c \in F$*,*

$$T(u_1 + u_2) = T(u_1) + T(u_2), \qquad T(c\,u) = c\,T(u),$$

then T *is called a* linear transformation*.*

Proposition 44 (Properties of Linear Transformations).
Let $T : U \to V$ *be a linear transformation. Then:*

1. $T(\vec{0}) = \vec{0}$*.*

2. *For all* $u_1, \ldots, u_n \in U$ *and* $c_1, \ldots, c_n \in F$*,*

$$T\Big(\sum_{i=1}^{n} c_i u_i\Big) \;=\; \sum_{i=1}^{n} c_i\,T(u_i).$$

Definition 45 (Matrix-Induced Linear Map).
For any $A \in M_{m\times n}(F)$*, define the map*

$$L_A\colon F^n \;\longrightarrow\; F^m, \qquad L_A(v) = A\,v,$$

for all $v \in F^n$*. One checks that* L_A *is linear.*

Proposition 46.
L_A *is a linear transformation.*

3.2 Kernel and Image

Definition 47 (Kernel and Image).
Let $T\colon U \to V$ be a linear transformation. The kernel *of T is*

$$\ker(T) = \{\, u \in U \mid T(u) = 0\} \subseteq U,$$

and the image *of T is*

$$\operatorname{Im}(T) = \{\, v \in V \mid v = T(u) \text{ for some } u \in U\} \subseteq V.$$

Proposition 48.
Let $T : U \to V$ be a linear transformation. Then:

1. $\ker(T) \leq U$ *(i.e.,* $\ker(T)$ *is a subspace of U).*
2. $\operatorname{Im}(T) \leq V$ *(i.e.,* $\operatorname{Im}(T)$ *is a subspace of V).*

Definition 49 (Rank and Nullity).
Let $T\colon U \to V$ be a linear transformation between finite-dimensional spaces. Then

$$\operatorname{rank}(T) = \dim\bigl(\operatorname{Im}(T)\bigr), \qquad \operatorname{nullity}(T) = \dim\bigl(\ker(T)\bigr).$$

Proposition 50 (Spanning Set of $\operatorname{Im}(T)$).
$T : U \to V$ linear, $U = \operatorname{span}(\alpha)$.
Denote $T(\alpha) = \{\, T(u) \mid u \in \alpha\}$. Then $T(\alpha)$ spans $\operatorname{Im}(T)$.

Theorem 51 (Rank-Nullity Theorem).
$T : U \to V$ linear transformation, with $\dim(U) = n$ *finite. Then*

$$\operatorname{rank}(T) \;+\; \operatorname{nullity}(T) \;=\; \dim(U).$$

3.3 Injective, Surjective

Definition 52 (Injective, Surjective, Bijective).
Let $f : X \to Y$ be a function between sets X and Y.

- *f is* injective *(one-to-one) if*

  $$\forall\, x_1, x_2 \in X,\; x_1 \neq x_2 \implies f(x_1) \neq f(x_2),$$

  equivalently $f(x_1) = f(x_2) \implies x_1 = x_2.$

- *f is* surjective *(onto) if*

  $$\forall\, y \in Y\; \exists\, x \in X \text{ with } f(x) = y,$$

  equivalently $\operatorname{Im}(f) = Y.$

- *f is* bijective *if it is both injective and surjective.*

Proposition 53.
$T : U \to V$ linear, U, V finite dimensional.

1. *T injective $\iff \ker(T) = \{\vec{0}\}$ ($\implies \text{nullity}(T) = 0$).*
2. *T surjective $\iff \text{Im}(T) = V$ ($\implies \text{rank}(T) = \dim(V)$).*
3. *If $\dim(U) = \dim(V)$, then T injective $\iff T$ surjective.*
4. *If $\dim(U) > \dim(V)$, T is not injective. If $\dim(U) < \dim(V)$, T is not surjective.*

3.4 Isomorphism and Coordinates

Definition 54 (Isomorphism of Vector Spaces).
Let $T : U \to V$ be a linear transformation between vector spaces over the same field.

- *If T is bijective, then T is called an* isomorphism.
- *When such an isomorphism T exists, we say U is* isomorphic *to V and write*

$$U \cong V.$$

Proposition 55.
Let $\beta = \{\vec{v}_1, \ldots, \vec{v}_n\}$ be a basis of V. Then every $\vec{v} \in V$ has a unique *expression*

$$\vec{v} \;=\; \sum_{i=1}^{n} c_i \, \vec{v}_i$$

as a linear combination of basis elements. The vector $(c_1, c_2, \ldots, c_n) \in F^n$ of coefficients is called the coordinate vector *of $\vec{v}$ relative to β, denoted $[\vec{v}]_\beta = (c_1, \ldots, c_n)$.*

Theorem 56.
Let V and W be vector spaces, let $\alpha = \{v_1, \ldots, v_n\}$ be a basis of V, and let $w_1, \ldots, w_n \in W$ be arbitrary. Then there is a unique linear transformation $T : V \to W$ such that

$$T(v_i) = w_i \quad (i = 1, \ldots, n).$$

Moreover, if $v = \displaystyle\sum_{i=1}^{n} c_i \, v_i$, then

$$T(v) \;=\; \sum_{i=1}^{n} c_i \, w_i.$$

Corollary 57.
If $T, S : V \to W$ *are linear and* $\alpha = \{v_1, \ldots, v_n\}$ *is a basis of* V *with*

$$T(v_i) = S(v_i) \quad \text{for } i = 1, \ldots, n,$$

then $T = S$.

Theorem 58.
V *vector space,* $\dim V = n$, *finite,* β *basis. The function which computes coordinates*

$$[-]_\beta : V \to F^n$$

is an isomorphism. Hence $V \cong F^n \quad (n = \dim V)$.

Proposition 59 (Composition and Inverses of Lin. Transforms).
Let $S : U \to V$, $T : V \to W$ *be linear.*

1. *The composition* $T \circ S : U \to W$ *is linear.*
2. *If* S, T *both isomorphisms,* $T \circ S$ *also is an isomorphism.*
3. *If* T *is an isomorphism, there is an inverse* $T^{-1} : W \to V$, *and* T^{-1} *is an isomorphism.*

3.5 Matrix of a Linear Transformation

Theorem 60 ("Only Dimensions Matter").
Let U, V *be finite-dimensional vector spaces over* F. *Then*

$$U \cong V \quad (\Longleftrightarrow) \quad \dim(U) = \dim(V).$$

Definition 61 (Matrix of a Linear Transformation).
Let $T : U \to V$ *be a linear transformation, and let* $\alpha = \{u_1, \ldots, u_n\}$ *and* $\beta = \{v_1, \ldots, v_m\}$ *be ordered bases of* U *and* V, *respectively. The* matrix of T relative to α and β *is the* $m \times n$ *matrix*

$$[T]_\alpha^\beta = \big([T(u_1)]_\beta \;\big|\; \ldots \;\big|\; [T(u_n)]_\beta\big),$$

whose i*th column is the coordinate vector of* $T(u_i)$ *in the basis* β.

Theorem 62 ("$[T]_\alpha^\beta$ Computes T in Coordinates").
Let $T : U \to V$ *be linear,* α, β *bases of* U, V. *Then for all* $u \in U$,

$$[T]_\alpha^\beta \, [u]_\alpha \;=\; [\, T(u) \,]_\beta.$$

Proposition 63 (ker and Im in Coordinates).
Let $T : U \to V$, *and let* α, β *be bases of* U *and* V. *Define* $A = [T]_\alpha^\beta$.

1. $\ker(T)$ corresponds, via α-coordinates, to $\text{null}(A)$ (i.e. the solution set of $A\vec{x} = \vec{0}$). In particular,

$$\text{nullity}(T) \;=\; \text{nullity}(A) \;=\; \#\{\text{free variables}\}.$$

2. $\text{Im}(T)$ corresponds, via β-coordinates, to $\text{Im}(L_A) = \text{col}(A)$. Hence

$$\text{rank}(T) \;=\; \text{rank}(A).$$

Definition 64 (Linear Operator).
Let $T\colon V \to V$ be a linear transformation on a vector space V (the domain and codomain are the same vector space). Such a T is called a linear operator. *If* $\dim(V) = n$*, then its matrix relative to any basis α of V is an $n \times n$ matrix, commonly denoted*

$$[T]_\alpha^\alpha = [T]_\alpha,$$

or simply $[T]$ when the basis is understood.

Proposition 65.
Let $T : V \to V$ be a linear operator, and let α be a basis of V with $n = \dim(V)$. Then

$$T \text{ is invertible} \quad\Longleftrightarrow\quad [T]_\alpha \text{ is invertible.}$$

Moreover, if T is invertible, then

$$\left[T^{-1}\right]_\alpha \;=\; \left([T]_\alpha\right)^{-1}.$$

3.6 Change of Basis

Definition 66 (Change-of-Coordinates Matrix).
Let V be an n-dimensional vector space over a field F, and let α and β be two ordered bases of V. The change-of-coordinates matrix from α to β *is the $n \times n$ matrix*

$$Q_\alpha^\beta \;=\; [I]_\alpha^\beta,$$

where $I : V \to V$ is the identity. Equivalently, the jth column of Q_α^β is the coordinate vector of the jth basis vector of α expressed in the basis β.

Proposition 67.

1. *For every $u \in V$,*

$$Q_\alpha^\beta\,[u]_\alpha \;=\; [u]_\beta.$$

2. *Q_α^β is invertible and*

$$\left(Q_\alpha^\beta\right)^{-1} \;=\; Q_\beta^\alpha.$$

Theorem 68 (Similarity Transformation).
Let $T : V \to V$ be a linear operator and let α, β be two bases of V. Then

$$Q_\alpha^\beta \, [T]_\alpha \, Q_\beta^\alpha \;=\; [T]_\beta.$$

Definition 69 (Similar Matrices).
Let $A, B \in M_{n\times n}(F)$. We say A and B are similar *if there exists an invertible matrix $Q \in M_{n\times n}(F)$ such that*

$$Q^{-1} \, A \, Q \;=\; B.$$

Proposition 70.
Let $Q \in M_{n\times n}(F)$ be any invertible matrix and let α be a basis of V (so $n = \dim V$). Then there exists a basis β of V such that

$$Q \;=\; Q_\alpha^\beta.$$

Proposition 71.
Let $T : V \to V$ be a linear operator, let α be a basis of V, and let $B \in M_{n\times n}(F)$ be any matrix. Then

$$[T]_\alpha \text{ is similar to } B \quad \Longleftrightarrow \quad \exists \text{ a basis } \beta \text{ of } V \text{ with } B = [T]_\beta.$$

Theorem 72.
If $\dim V = n$ and $\dim W = m$, then

$$\mathcal{L}(V, W) \;\cong\; M_{m\times n}(F).$$

Moreover, given a basis α of V and a basis β of W, the map

$$\varphi : \mathcal{L}(V, W) \;\longrightarrow\; M_{m\times n}(F), \qquad \varphi(T) = [T]_\alpha^\beta$$

is a vector-space isomorphism.

4 Inner Product Spaces

4.1 Definition and Main Examples

Definition 73 (Inner Product).
Let V be a vector space over $\mathbb{F}$ (where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$). An inner product *on V is a function*

$$\langle \cdot \, , \cdot \rangle : V \times V \;\longrightarrow\; \mathbb{F}$$

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

1. $\langle u, \, v + w\rangle = \langle u, \, v\rangle + \langle u, \, w\rangle, \quad \langle u, \, c\,v\rangle = c\,\langle u, \, v\rangle.$ *(*linearity*).*

2. $\langle u, v\rangle = \overline{\langle v, u\rangle}$ (conjugate symmetry).

3. $\langle u, u\rangle > 0$ *for all* $u \neq 0$ (positive definiteness).

Proposition 74 (Inner Product Properties).
Let V be an inner-product space over $\mathbb{F}$ and let $u, v, w \in V$, $c \in \mathbb{F}$. Then:

1. (Conjugate-linearity in the second slot)

$$\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle, \qquad \langle u, c\,v\rangle = \overline{c}\,\langle u, v\rangle.$$

2. $\langle u, 0\rangle = 0$ *and* $\langle 0, u\rangle = 0$.

3. $\langle u, u\rangle = 0$ *if and only if* $u = 0$.

4. *If* $\langle u, w\rangle = \langle v, w\rangle$ *for all* $w \in V$*, then* $u = v$.

Definition 75 (Conjugate and Adjoint of a Matrix).
Let $A = (a_{ij}) \in M_{m\times n}(\mathbb{F})$.

- *The* conjugate *of A is the entrywise–conjugated matrix*

$$\overline{A} \;=\; \left(\overline{a_{ij}}\right) \;\in\; M_{m\times n}(\mathbb{F}).$$

- *If A is square ($n = m$), the* adjoint *(or Hermitian transpose) of A is*

$$A^* \;=\; \overline{A}^T \;=\; \left(\overline{A}\right)^T \;\in\; M_{n\times n}(\mathbb{F}),$$

  characterized by $\langle A\,u, v\rangle = \langle u, A^* v\rangle$ for all u, v.

4.2 Norm and Angle

Definition 76 (Norm).
Let V be an inner-product space with inner product $\langle\cdot,\cdot\rangle$. The norm *(or length) of a vector $u \in V$ is*

$$\|u\| \;=\; \sqrt{\langle u, u\rangle}\,.$$

Proposition 77.
Let V be an inner-product space, $u \in V$, and $c \in \mathbb{F}$. Then

$$\|c\,u\| \;=\; |c|\,\|u\|.$$

Theorem 78 (Cauchy–Schwarz Inequality).
Let V be an inner-product space and $u, v \in V$. Then

1. $\big|\langle u, v\rangle\big| \leq \|u\|\,\|v\|$.

2. *Equality holds,* $|\langle u, v\rangle| = \|u\|\|v\|$*, if and only if* $u = c\,v$ *(equivalently* $v = c\,u$*) for some scalar* $c \in \mathbb{F}$*.*

Proposition 79 (Triangle Inequality).
In any inner-product space V*, for all* $u, v \in V$*,*

$$\|u + v\| \;\le\; \|u\| + \|v\|.$$

Definition 80 (Angle).
Let V *be a real inner-product space and let* $u, v \in V$ *be nonzero. By the Cauchy–Schwarz inequality, the quotient*

$$\frac{\langle u, v\rangle}{\|u\|\,\|v\|}$$

lies in $[-1, 1]$*, so there is a unique* $\theta \in [0, \pi]$ *with*

$$\cos\theta \;=\; \frac{\langle u, v\rangle}{\|u\|\,\|v\|}.$$

This θ *is called the* angle *between* u *and* v*.*

4.3 Orthogonal Sets and Complements

Definition 81 (Orthogonal and Orthonormal Sets).
Let V *be an inner-product space.*

1. *Two vectors* $u, v \in V$ *are called* orthogonal *if* $\langle u, v\rangle = 0$*. (In particular, the zero-vector is orthogonal to every vector.)*

2. *A subset* $X \subseteq V$ *is* orthogonal *if every pair of distinct vectors in* X *is orthogonal and* $0 \notin X$*.*

3. *If, in addition,* $\|u\| = 1$ *for all* $u \in X$*, then* X *is* orthonormal.

4. *An orthonormal subset that is also a basis of* V *is called an* orthonormal basis *(ONB).*

Proposition 82.
Let V *be an inner-product space and let* $X \subseteq V$ *be an orthogonal set. Then* X *is linearly independent.*

Theorem 83 (Fourier Coefficients).
Let $\alpha = \{v_1, \ldots, v_n\}$ *be an orthogonal basis of the inner-product space* V*. Then for every* $u \in V$*,*

$$u \;=\; \sum_{i=1}^{n} \frac{\langle u, v_i\rangle}{\langle v_i, v_i\rangle}\, v_i,$$

and the coordinate vector $[u]_\alpha$ is

$$[u]_\alpha = \begin{pmatrix} \dfrac{\langle u, v_1\rangle}{\langle v_1, v_1\rangle} \\ \vdots \\ \dfrac{\langle u, v_n\rangle}{\langle v_n, v_n\rangle} \end{pmatrix}.$$

If α is orthonormal, then this simplifies to

$$u = \sum_{i=1}^{n} \langle u, v_i\rangle\, v_i.$$

Definition 84 (Orthogonal Complement).
Let V be an inner-product space and $X \subseteq V$. The orthogonal complement *of X is*

$$X^\perp \;=\; \{\, v \in V \mid \langle v, x\rangle = 0 \text{ for all } x \in X\},$$

which is a subspace of V.

Proposition 85.
Let $W \leq V$. Then:

1. *$W^\perp$ is a subspace of V.*
2. *If $\alpha = \{w_1, \dots, w_k\}$ is a basis of W, then for any $u \in V$,*

$$u \in W^\perp \iff \langle u, w_i\rangle = 0 \quad (i = 1, \dots, k).$$

3. *$W \cap W^\perp = \{0\}$.*

4.4 Orthogonal Projection and Gram-Schmidt Algorithm

Theorem 86 (Orthogonal Projection).
Let V be an inner-product space and $W \leq V$ a finite-dimensional subspace. Then:

1. *For each $u \in V$ there are unique $w \in W$ and $w' \in W^\perp$ such that $u = w + w'$. Hence $V = W \oplus W^\perp$.*
2. *If $\{u_1, \dots, u_k\}$ is an orthogonal basis of W, then the projection $w \in W$ of any $u \in V$ is*

$$w \;=\; \sum_{i=1}^{k} \frac{\langle u, u_i\rangle}{\langle u_i, u_i\rangle}\, u_i.$$

3. *If V is finite-dimensional, then* $\dim V = \dim W + \dim W^\perp$.

Lemma 87 (Pythagoras' Theorem).
Let V be an inner-product space and let $u, v \in V$ be orthogonal. Then

$$\|u+v\|^2 \;=\; \|u\|^2 + \|v\|^2.$$

Theorem 88 (Projection is Closest Vector in W).
Let V be an inner-product space, $W \le V$ finite-dimensional, and $u \in V$. Write the unique decomposition

$$u = w + w', \quad w \in W,\ w' \in W^{\perp},$$

so that $w = P_W(u)$. Then w is the closest vector in W to u, in the sense that for every $z \in W$,

$$\|u-w\| \;\le\; \|u-z\|.$$

Theorem 89 (Gram–Schmidt Orthogonalization).
With the notation of the Gram–Schmidt process applied to a basis $\{u_1, \ldots, u_n\} \subset V$, one defines vectors $v_1, \ldots, v_n$. Then:

1. *For each $k = 1, \ldots, n$,*

   $$\mathrm{span}\{v_1, \ldots, v_k\} \;=\; \mathrm{span}\{u_1, \ldots, u_k\},$$

   hence $\{v_1, \ldots, v_n\}$ spans the same subspace W.

2. *For each k, the set $\{v_1, \ldots, v_k\}$ is orthogonal in $W_k = \mathrm{span}\{u_1, \ldots, u_k\}$, and in particular $\{v_1, \ldots, v_n\}$ is an orthogonal basis of W.*

4.5 Inner Product Defined by a Matrix

Definition 90 (Positive Definite Matrix).
Let $A \in M_{n\times n}(\mathbb{R})$ be symmetric. A is called positive definite *if*

$$u^T A\, u > 0 \quad \textit{for all nonzero } u \in \mathbb{R}^n.$$

Proposition 91.
Let $A \in M_{n\times n}(\mathbb{R})$ be symmetric and positive definite. Then

$$\langle u, v\rangle_A \;=\; u^T A v$$

defines an inner product on $\mathbb{R}^n$.

5 Eigenvalues and Diagonalization

5.1 Eigenvalues and Eigenvectors

Definition 92 (Eigenvector and Eigenvalue).
Let $T\colon V \to V$ be a linear operator on a vector space V. A nonzero vector $u \in V$ is an eigenvector *of T if there exists a scalar $\lambda \in F$ such that*

$$T(u) \;=\; \lambda\, u.$$

The scalar λ is called the corresponding eigenvalue*.*

Definition 93 (Eigenspace).
If λ is an eigenvalue of $T\colon V \to V$, the eigenspace *of λ is the subspace*

$$E_\lambda = \{\, u \in V \mid T(u) = \lambda\, u\}.$$

Definition 94 (Characteristic Polynomial).
Let $T\colon V \to V$ be a linear operator on an n-dimensional vector space V. The characteristic polynomial *of T is*

$$C_T(x) = \det\bigl([T]_\alpha - x\, I\bigr),$$

where α is any basis of V and $[T]_\alpha$ is the matrix of T relative to α.

Proposition 95 (Characteristic Polynomial).
Let $T : V \to V$ be a linear operator on an n-dimensional space V. Its characteristic polynomial $C_T(x)$ satisfies:

1. *$C_T(x)$ is independent of the choice of basis.*
2. *$\deg C_T = n$.*
3. *A scalar λ is an eigenvalue of T if and only if $C_T(\lambda) = 0$.*
4. *The eigenspace $E_\lambda = \ker(T - \lambda I)$ is a subspace of V.*

5.2 Diagonalization

Definition 96 (Diagonalizable Operator).
Let $T\colon V \to V$ be a linear operator on a finite-dimensional vector space V. We say T is diagonalizable *if there exists a basis α of V consisting entirely of eigenvectors of T; equivalently, in that basis the matrix of T is diagonal.*

Proposition 97 (Diagonalizability Criterion).
Let $T : V \to V$ be a linear operator on a finite-dimensional space V. Then T is diagonalizable if and only if there exists a basis of V consisting of eigenvectors of T; equivalently, for any basis β, the matrix $[T]_\beta$ is similar to a diagonal matrix.

Definition 98 (Diagonalizable Matrix).
Let $A \in M_{n\times n}(F)$. We say A is diagonalizable *if there exist an invertible matrix $Q \in M_{n\times n}(F)$ and a diagonal matrix $D \in M_{n\times n}(F)$ such that*

$$Q^{-1}\,A\,Q \;=\; D.$$

5.3 Diagonalizability

Proposition 99.
Let V be a finite-dimensional real vector space and $T : V \to V$ linear. If the characteristic polynomial $C_T(x)$ has a nonreal complex root, then T is not diagonalizable over $\mathbb{R}$.

Lemma 100.
Let $T : V \to V$ be linear and let $\lambda_1 \neq \lambda_2$ be two distinct eigenvalues. Then

$$E_{\lambda_1} \;\cap\; E_{\lambda_2} \;=\; \{0\}.$$

Proposition 101.
Let $T : V \to V$ have distinct eigenvalues $\lambda_1, \ldots, \lambda_m$. For each i let β_i be a basis of the eigenspace E_{λ_i}, and set $\beta = \beta_1 \cup \cdots \cup \beta_m$. Then:

1. $|\beta| = \sum_i |\beta_i|$.
2. *β is linearly independent.*

Corollary 102.
If $T : V \to V$ is linear on an n-dimensional space and has n distinct eigenvalues, then T is diagonalizable.

Definition 103 (Geometric and Algebraic Multiplicity).
Let $T\colon V \to V$ be a linear operator and λ an eigenvalue of T.

- *The* geometric multiplicity *of λ is $\dim E_\lambda$, where $E_\lambda = \ker(T - \lambda I)$.*
- *The* algebraic multiplicity *of λ is the exponent of the factor $(x - \lambda)$ in the characteristic polynomial $C_T(x)$.*

Theorem 104.
Let $T : V \to V$ be a linear operator on a finite-dimensional space V, and let λ be any eigenvalue of T. Then

$$1 \;\leq\; \big(\textit{geometric multiplicity of } \lambda\big) \;\leq\; \big(\textit{algebraic multiplicity of } \lambda\big).$$

Proposition 105.
Let $T : V \to V$ be a linear operator on an n-dimensional vector space over a field F, with characteristic polynomial $C_T(x)$.

1. *If $F = \mathbb{R}$ and $C_T(x)$ has a nonreal root, then T is not diagonalizable over $\mathbb{R}$.*

2. *If $F = \mathbb{R}$ and all roots of $C_T(x)$ lie in $\mathbb{R}$, or if $F = \mathbb{C}$, then*

$$T \text{ is diagonalizable} \iff \forall\,\lambda,\ \big(\text{geom. mult. }\lambda\big) = \big(\text{alg. mult. }\lambda\big).$$

3. *When T is diagonalizable, let $\lambda_1, \ldots, \lambda_k$ be its distinct eigenvalues and β_i a basis of the eigenspace E_{λ_i}. Then*

$$\beta = \beta_1 \,\cup\, \beta_2 \,\cup \cdots \cup\, \beta_k$$

is a basis of V consisting entirely of eigenvectors of T.

5.4 Orthogonal Diagonalization

Definition 106 (Symmetric, Orthogonal, Self-Adjoint, Unitary Matrices).
Let $A \in M_{n\times n}(F)$.

- *If $F = \mathbb{R}$, A is* symmetric *when $A^T = A$.*
- *If $F = \mathbb{R}$, A is* orthogonal *when its columns form an orthonormal basis of $\mathbb{R}^n$ (equivalently $A^T = A^{-1}$).*
- *If $F = \mathbb{C}$, A is* self-adjoint *(Hermitian) when $A^* = A$.*
- *If $F = \mathbb{C}$, A is* unitary *when its columns form an orthonormal basis of $\mathbb{C}^n$ (equivalently $A^* = A^{-1}$).*

Proposition 107 (Orthogonal and Unitary Matrices).

1. *If $A \in M_n(\mathbb{R})$, then A is orthogonal $\Leftrightarrow A^T = A^{-1}$.*
2. *If $A \in M_n(\mathbb{C})$, then A is unitary $\Leftrightarrow A^* = A^{-1}$.*

Definition 108 (Orthogonal and Unitary Diagonalizability).
Let $A \in M_{n\times n}(F)$.

- *If $F = \mathbb{R}$, A is* orthogonally diagonalizable *if there exists an orthogonal matrix Q and a diagonal D with*

$$Q^T A\, Q = D \quad\iff\quad A = Q\, D\, Q^T.$$

- *If $F = \mathbb{C}$, A is* unitarily diagonalizable *if there exists a unitary matrix Q and a diagonal D with*

$$Q^* A\, Q = D \quad\iff\quad A = Q\, D\, Q^*.$$

Proposition 109.

1. *If $A \in M_n(\mathbb{R})$ is orthogonally diagonalizable, then A is symmetric ($A^T = A$).*

2. If $A \in M_n(\mathbb{C})$ is unitarily diagonalizable, then A is normal ($A^*A = AA^*$).

Lemma 110.

1. If $A \in M_n(\mathbb{R})$ and $u, v \in \mathbb{R}^n$, then

$$\langle Au, v\rangle = \langle u, A^T v\rangle, \quad \langle u, Av\rangle = \langle A^T u, v\rangle.$$

2. If $A \in M_n(\mathbb{C})$ and $u, v \in \mathbb{C}^n$, then

$$\langle Au, v\rangle = \langle u, A^* v\rangle, \quad \langle u, Av\rangle = \langle A^* u, v\rangle.$$

Theorem 111 (Reality of Eigenvalues).
If $A \in M_n(\mathbb{R})$ is symmetric or $A \in M_n(\mathbb{C})$ is self-adjoint, then all roots of its characteristic polynomial lie in $\mathbb{R}$; equivalently, every eigenvalue of A is real.

Theorem 112 (Orthogonality of Eigenspaces).
Let $A \in M_n(\mathbb{R})$ be symmetric or $A \in M_n(\mathbb{C})$ be normal. Then its eigenspaces are mutually orthogonal: if $u \in E_{\lambda_1}$ and $w \in E_{\lambda_2}$ with $\lambda_1 \neq \lambda_2$, then $\langle u, w\rangle = 0$.

Theorem 113 (Spectral Theorem).

1. For $A \in M_n(\mathbb{R})$:

$$A \text{ orthogonally diagonalizable} \iff A \text{ symmetric } (A^T = A).$$

2. For $A \in M_n(\mathbb{C})$:

$$A \text{ unitarily diagonalizable} \iff A \text{ normal } (A^*A = AA^*).$$

Definition 114 (Definiteness of a Symmetric Matrix).
Let $A \in M_{n\times n}(\mathbb{R})$ be symmetric. We say:

- *A is* positive definite *if $u^T A\, u > 0$ for all nonzero $u \in \mathbb{R}^n$.*
- *A is* positive semidefinite *if $u^T A\, u \geq 0$ for all $u \in \mathbb{R}^n$.*
- *A is* negative definite *if $u^T A\, u < 0$ for all nonzero u.*
- *A is* negative semidefinite *if $u^T A\, u \leq 0$ for all u.*
- *Otherwise, A is called* indefinite.

Theorem 115 (Positive Definiteness).
Let $A \in M_n(\mathbb{R})$ be symmetric. Then

$$A \text{ is positive definite} \iff \text{all eigenvalues of } A \text{ are positive.}$$

Lemma 116.
Let $A \in M_{m\times n}(\mathbb{R})$. Then:

1. $A^T A \in M_{n\times n}(\mathbb{R})$ and $AA^T \in M_{m\times m}(\mathbb{R})$ are symmetric and positive semi-definite.

2. $\operatorname{rank}(A^T A) = \operatorname{rank}(A) = \operatorname{rank}(AA^T)$.

5.5 Singular Value Decomposition

Theorem 117 (Singular Value Decomposition).
Let $A \in M_{m\times n}(\mathbb{R})$, and let $L_A : \mathbb{R}^n \to \mathbb{R}^m$ be the associated linear map. Set $r = \operatorname{rank}(A)$. Then there exist orthonormal bases

$$\beta = \{v_1, \ldots, v_n\} \subset \mathbb{R}^n, \qquad \gamma = \{w_1, \ldots, w_m\} \subset \mathbb{R}^m,$$

and positive scalars $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, such that

$$L_A(v_i) = \begin{cases} \sigma_i\, w_i, & i = 1, \ldots, r, \\ 0, & i = r+1, \ldots, n. \end{cases}$$

Consequently, the matrix of L_A relative to β, γ is

$$[L_A]_\beta^\gamma = \begin{pmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \end{pmatrix}_{m\times n}.$$

Definition 118 (Singular Values).
Let $A \in M_{m\times n}(\mathbb{R})$ have rank r. In its singular value decomposition one obtains positive scalars

$$\sigma_1 \;\geq\; \sigma_2 \;\geq\; \cdots \;\geq\; \sigma_r > 0.$$

These $\sigma_1, \ldots, \sigma_r$ are called the singular values *of A.*

Corollary 119.
In the SVD of Theorem 117:

1. *$\{v_1, \ldots, v_n\}$ is an orthonormal basis of eigenvectors of A^TA in $\mathbb{R}^n$.*
2. *$\{w_1, \ldots, w_m\}$ is an orthonormal basis of eigenvectors of AA^T in $\mathbb{R}^m$.*
3. *A^TA and AA^T have the same nonzero eigenvalues, and the singular values σ_i are the positive square roots of these eigenvalues.*