Compact course notes

Math 245, Fall 2010

Linear Algebra 2

Professor: S. New
transcribed by: J. Lazovskis
University of Waterloo
December 17, 2010

Contents

1 Overview of Linear Algebra 1

1.1 Lines and planes

A line in 3 dimensions is best described parametrically. Given a point p and a vector u, all points on the line are described by $x = p + tu$ for $t \in \mathbb{R}$.

A plane in 3 dimensions is the same; given two points and two vectors $p + u$ and $q + v$, the points on the plane are described by $x = p + ru + st$ for $s, r \in \mathbb{R}$. This can be generalized by $a_1x_1 + a_2x_2 + a_3x_3 = b$.

Definition 1.1.1. A vector space in $\mathbb{R}^n$ is a set of the form $\{t_1u_1, \ldots, t_ku_k | t_i \in \mathbb{R}\} = \text{span}\{u_1, \ldots, u_k \in \mathbb{R}\}$.A vector space includes the origin.

Definition 1.1.2. An affine space in $\mathbb{R}^n$ is a set of the form $p + V = \{p + v | v \in V\}$, for some point $p \in \mathbb{R}^n$ and some vector space $V \in \mathbb{R}^n$. Here, $p + V$ is the affine space through p parallel to V.

Theorem 1.1.3. If U is a basis for a vector space $V \in \mathbb{R}^n$, then the number of elements in U is at most n. If U and W are bases for the same vector space in $\mathbb{R}^n$, then they have the same number of elements. This number is termed dimension.

Definition 1.1.4. A function A is said to be linear if the two following conditions are satisfied for some scalar t and all $x, y \in \mathbb{R}^n$:
$A(tx) = tA(x)$
$A(x + y) = A(x) + A(y)$

Definition 1.1.5. A linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ is a map of the form $L(x) = Ax$ for some $A \in M_{m\times n}$.

Definition 1.1.6. An affine map $L : \mathbb{R}^n \to \mathbb{R}^m$ is a map of the form $L(x) = Ax + b$ for some $A \in M_{m\times n}$ and $b \in \mathbb{R}^n$.

· Note that the span of columns is the column space of the range.

· The nullspace is perpendicular to the rowspace.

Theorem 1.1.7. Suppose that $L : \mathbb{R}^n \to \mathbb{R}^n$ is linear. Let $x \in \mathbb{R}^n$ be such that $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$, where e_k is the kth standard basis vector. Then $L(x) = Ax$ for $A = (L(e_1)\ \ L(e_2)\ \ \ldots\ \ L(e_n))$.

Theorem 1.1.8. Suppose A reduces to R in reduced row echelon form. Then the non-zero rows of R form a basis for the row space of A. Then to obtain a basis for the nullspace of A, solve $Ax = 0$ using Gauss-Jordan elimination to get $x = t_1v_1 + \cdots + t_kv_k$then $\{u_1, \ldots, u_k\}$ is a basis for null(A).

Definition 1.1.9. Given a function $f : X \to Y$:
1. f is 1:1 or injective when for all $y \in Y$ there exists at most 1 $x \in X$ such that $y = f(x)$.
2. f is onto or surjective when for all $y \in Y$ there exists at least 1 $x \in X$ such that $y = f(x)$.
3. f is invertible or bijective when f is one-to-one and onto.

Theorem 1.1.10. A function is differentiable when it can be suitably approximated by an affine map.

Definition 1.1.11. Let U and V be vector spaces with $\dim(U) = n$ and $\dim(V) = n$. Let $\mathcal{U} = \{u_1, \ldots, u_n\}$ and $\mathcal{V} = \{v_1, \ldots, v_n\}$ be ordered bases for for U and V respectively. Then for $x \in \mathcal{U}$ with $x = t_1u_1 + \cdots + t_nu_n$,

define $[x]_{\mathcal{U}} = t = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \in \mathbb{R}^n$

For a linear map $L : U \to V$, there is a unique matrix described by $[L]^{\mathcal{U}}_{\mathcal{V}}$ such that for all $x \in U$, $[L(x)]_{\mathcal{V}} = [L]^{\mathcal{U}}_{\mathcal{V}}[x]_{\mathcal{U}}$. This matrix is given by $[L]^{\mathcal{U}}_{\mathcal{V}} = \big([L(u_1)]_{\mathcal{V}}\ \ \ldots\ \ [L(u_n)]_{\mathcal{V}}\big) \in M_{m\times n}$

Remark 1.1.12. The matrix $[L]^{\mathcal{U}}_{\mathcal{V}}$ is termed the matrix of L with respect to the bases $\mathcal{U}$ and $\mathcal{V}$.

1.2 Determinants

Theorem 1.2.1. Given matrices $A, B \in M_{n\times n}$ and an equation $AB = 0$,
$(A|e_i) \sim (I|B_i)$
$(A|I) \sim (I|B)$
where e_i is the ith column of I and B_i is the ith column of B.

Definition 1.2.2. For $n \geqslant 2$ and $A \in M_{n\times n}$, given a fixed i, $\det(A) = \sum_{j=1}^{n}(-1)^{i+j}A_{i,j}\det(A^{i,j})$

· where $A^{i,j}$ is the $(n-1)\times(n-1)$ matrix obtained from A by removing the i^{th} row and j^{th} column
· and $A_{i,j}$ is the element in the i^{th} row and j^{th} column of A

Theorem 1.2.3. If $\text{Null}(A) \neq \{0\}$, then A is not invertible and $\det(A) = 0$.

Definition 1.2.4. The matrix defined by $\text{Cofac}(A)$ is termed the cofactor matrix (or classical adjoint)of A.

Theorem 1.2.5. For $A \in M_{m\times n}$, A is invertible if and only if $\det(A) \neq 0$, and in that case $A^{-1} = \frac{1}{\det(A)} \cdot \text{Cofac}(A)$ where $(\text{Cofac}(A))_{k,\ell} = (-1)^{k+\ell}\det(A^{\ell,k})$

Theorem 1.2.6. For all $A \in M_{n\times n}$, $A \cdot \text{Cofac}(A) = \det(A)I$. Also, $\text{Cofac}(A) = (A^{adj})^t$.

Theorem 1.2.7. [Inversion properties]
For any $A, B \in M_{n\times n}$, $\det(AB) = \det(A)\det(B)$
For any $A \in M_{n\times n}$ and for $t \in \mathbb{N}$, $det(A) = \det(A^t)$

2 Operations in vector spaces

Remark 2.0.1. A vector space over a field $\mathbb{F}$ is a set closed over addition and multiplication.

Remark 2.0.2. Let U, V be finite dimensional vector spaces with bases $\mathcal{U}_1, \mathcal{U}_2, \mathcal{V}_1, \mathcal{V}_2$. For any $u \in U$ and linear map $L : U \to V$,

$$[u]_{\mathcal{U}_2} = [I]^{\mathcal{U}_1}_{\mathcal{U}_2}[u]_{\mathcal{U}_1} \qquad [L]^{\mathcal{U}_2}_{\mathcal{V}_2} = [I]^{\mathcal{V}_1}_{\mathcal{V}_2}[L]^{\mathcal{U}_1}_{\mathcal{V}_1}[I]^{\mathcal{U}_2}_{\mathcal{U}_1}$$

Definition 2.0.3. For $A, B \in M_{n\times n}$, A and B are similar when there exists an invertible matrix P such that $B = PAP^{-1}$.

2.1 The dot product in $\mathbb{R}^n$

Definition 2.1.1. For $u, v \in \mathbb{R}^n$, the dot product of u and v is $u \cdot v = \sum_{i=1}^{n} u_1 v_1 = u^t v = v^t u$.

Theorem 2.1.2. [Properties of the dot product]
For $t \in \mathbb{R}$ and $u, v, w \in \mathbb{R}^n$:

1. $u \cdot u \geqslant 0$ with $u \cdot u = 0 \iff u = 0$ — Positive definite
2. $u \cdot v = v \cdot u$ — Symmetric
3. $(tu) \cdot v = t(u \cdot v) = u \cdot (tv)$ — Bilinear
4. $(u + w) \cdot w = u \cdot w + v \cdot w$

Remark 2.1.3. For any $A \in M_{m\times n}$ and $x \in \mathbb{R}^n$, we have $Ax = (c_1 \ \dots \ c_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = c_1x_1 + \cdots + c_nx_n$. Also note that the row space of A is equal to the column space of A.

Definition 2.1.4. For $u \in \mathbb{R}^n$, the length of u is $\sqrt{\sum_{i=1}^{n} u_i^2} = \sqrt{u \cdot u} = |u|$.

Theorem 2.1.5. [Properties of length]
For $u, v \in \mathbb{R}^n$ and $t \in \mathbb{R}$,

1. $|u| \geqslant 0$ with $|u| = 0 \iff u = 0$
2. $|tu| = |t|\,|u|$
3. $u \cdot v = \frac{1}{2}(|u+v|^2 - |u|^2 - |v|^2) = \frac{1}{4}(|u+v|^2 - |u-v|^2)$
4. $|u \cdot v| \leqslant |u||v|$ with $|u \cdot v| = |u||v| \iff \{u, v\}$ is linearly dependent
5. $|u - v| \leqslant |u + v| \leqslant |u| + |v|$

Definition 2.1.6. For $u, v \in \mathbb{R}^n$, the distance between u and v is $\mathrm{d}(u, v) = |u - v| = |v - u|$.

Theorem 2.1.7. [Properties of distance]
For $u, v, w \in \mathbb{R}^n$,

1. $\mathrm{d}(u, v) \geqslant 0$ with $\mathrm{d}(u, v) = 0 \iff u = v$
2. $\mathrm{d}(u, v) = \mathrm{d}(v, u)$
3. $\mathrm{d}(u, v) \leqslant \mathrm{d}(u, w) + \mathrm{d}(w, v)$

Definition 2.1.8. For $0 \neq u, v \in \mathbb{R}^n$, the angle between u and v is $\mathrm{angle}(u, v) = \theta$. This is expressed as

$$\theta = \cos^{-1}\left(\frac{u \cdot v}{|u||v|}\right) = \sin^{-1}\left(\frac{|u \times v|}{|u||v|}\right).$$

Theorem 2.1.9. [Properties of angles]
For $0 \neq u, v \in \mathbb{R}^n$ and $\theta = \mathrm{angle}(u, v)$:

1. Law of cosines: $|v - u|^2 = |u|^2 + |v|^2 - 2|u||v|\cos(\theta)$
2. Pythagorean theorem: If $u \cdot (v - u) = 0$, then $|v|^2 = |u|^2 + |v - u|^2$
3. Trigonometric ratios: If $u \cdot (v - u) = 0$, then $\cos(\theta) = \frac{|u|}{|v|}$ and $\sin(\theta) = \frac{|v-u|}{|v|}$

Theorem 2.1.10. For $t \in \mathbb{R}$ and $u, v \in \mathbb{R}^n$ and $t, u, v \neq 0$:

$$\mathrm{angle}(tu, v) = \begin{cases} \mathrm{angle}(u, v) & \text{if } t > 0 \\ \pi - \mathrm{angle}(u, v) & \text{if } t < 0 \end{cases}$$

Definition 2.1.11. For $a, b, c \in \mathbb{R}^n$ all distinct, define $\angle abc = \mathrm{angle}(a - b, c - b)$.

Theorem 2.1.12. For $a, b, c \in \mathbb{R}^n$ all distinct, $\angle abc + \angle cab + \angle bca = \pi$.

Definition 2.1.13. For $0 \neq u \in \mathbb{R}^n$ and $p \in \mathbb{R}^n$, the hyperspace (or hyperplane) in $\mathbb{R}^n$ through p and perpendicular to u is the set of points $x \in \mathbb{R}^n$ such that $(x - p) \cdot u = 0$.

2.2 Orthogonal projections

Definition 2.2.1. For $u, v \in \mathbb{R}^n$, we say that u and v are orthogonal (or perpendicular) when $u \cdot v = 0$. If $u, v \neq 0$, then $u \cdot v = 0 \iff \mathrm{angle}(u, v) = \frac{\pi}{2}$.

Definition 2.2.2. For $0 \neq u \in \mathbb{R}^n$ and $x \in \mathbb{R}^n$, the orthogonal projection of x onto u is $\mathrm{proj}_u(x) = \frac{u \cdot x}{|u|}$. If $U = \mathrm{span}\{u\}$, then $\mathrm{proj}_U(x) = \frac{u \cdot x}{|u|^2}u$. Note that $(x - \mathrm{proj}_u(x))$ is orthogonal to u.

With reference to the case above, $[\mathrm{proj}_U x] = \frac{1}{|u|^2} u u^t$.

Definition 2.2.3. For a vector space $U \in \mathbb{R}^n$, the orthogonal complement of U is the vector space $U^\perp = \{x \in \mathbb{R}^n | x \cdot u = 0 \text{ for all } u \in U\} = \mathrm{Null}(U^t)$.

The projection of x onto $U^\perp$ is $x - \mathrm{proj}_U x$.

Theorem 2.2.4. [Properties of the orthogonal complement]
Let U be a vector space in $\mathbb{R}^n$. Then

1. For $A \in M_{m \times n}$ over $\mathbb{R}$, $\mathrm{Null}(A) = \mathrm{Row}(A)^\perp$
2. $U \cap U^\perp = \{0\}$
3. $\dim(U) + \dim(U^\perp) = n$
4. $(U^\perp)^\perp = U$

Theorem 2.2.5. For $A \in M_{m\times n}$, $\mathrm{rank}(A^tA) = \mathrm{rank}(A)$. Also, $\mathrm{Null}(A^tA) = \mathrm{Null}(A)$.

Theorem 2.2.6. Let U be a vector space in $\mathbb{R}^n$ and $x \in \mathbb{R}^n$. Then there exist unique vectors $u, v \in \mathbb{R}^n$ with $u \in U$ and $v \in U^\perp$ such that $u + v = x$.

Corollary 2.2.7. When $\{u_1, \ldots, u_k\}$ is a basis for U and $A = (u_1 \ldots u_k) \in M_{n\times k}$, then
· $\mathrm{Proj}_U(x) = A(A^tA)^{-1}A^tx$
· $\mathrm{Proj}_{U^\perp}(x) = (I - A(A^tA)^{-1}A^t)x$.

Definition 2.2.8. Let U be a subspace of $\mathbb{R}^n$ and let $x \in \mathbb{R}^n$. Let u, v be the unique vectors with $u \in U$ and $v \in U^\perp$ with $u + v = x$. Then u is termed the orthogonal projection of x onto U and we write $u = \mathrm{Proj}_U(x)$. Note that since $(U^\perp)^\perp = U$, we have $v = \mathrm{Proj}_{U^\perp}(x)$.

Theorem 2.2.9. Let U be a subspace of $\mathbb{R}^n$ with $x \in \mathbb{R}^n$. Then the point $u = \mathrm{Proj}_U(x)$ is the unique point on U which is nearest to x.

Theorem 2.2.10. Given a set of data points $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, the polynomial $f \in P_m$ with $f(x) = c_0 + c_1x + c_2x^2 + \cdots + c_mx^m$ that best fits these points has coefficient vector c given by $c = (A^tA)^{-1}A^ty$, where

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix},\ c = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_m \end{pmatrix} \text{ and } f(x) = \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix} = Ac, \text{ with } y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Remark 2.2.11. The above polynomial is termed the least-squares best-fit polynomial for the given data, such that $\displaystyle\sum_{i=1}^{n}(y_i - f(x_i))^2$ is minimized.

Remark 2.2.12. If we have at least $m + 1$ distinct x-coordinates, then A has maximal rank, is invertible, and so $(A^tA)^{-1}$ exists. In general, a best-fit polynomial always exists, but a unique one exists only if the number of distinct x-values is greater than m.

2.3 The cross product in $\mathbb{R}^n$

Theorem 2.3.1. Let $u_1, \ldots, u_{n-1} \in \mathbb{R}^n$. Then the cross product of these vectors is

$$\cdot\ X(u_1, \ldots, u_{n-1}) = \text{formal det}\left(u_1, \ldots, u_{n-1}, \begin{matrix} e_1 \\ \vdots \\ e_n \end{matrix}\right)$$
$$= \sum_{i=1}^{n}(-1)^{i+n}\det(A^i)e_i$$

where $\{e_1, \ldots, e_n\}$ are the standard basis vectors
$A = (u_1 \ldots u_{n-1}) \in M_{n\times(n-1)}$
A^i = the $(n-1)\times(n-1)$ matrix obtained from A by removing the ith row

Theorem 2.3.2. [Properties of the cross product]
For vectors $u, v \in \mathbb{R}^n$:
1. $X(u_1, \ldots, tu_k, \ldots, u_{n-1}) = tX(u_1, \ldots, u_k, \ldots, u_{n-1})$ — $(n-1)$-linear
2. $X(u_1, \ldots, u_k, \ldots, u_\ell, \ldots, u_{n-1}) = -X(u_1, \ldots, u_\ell, \ldots, u_k, \ldots, u_{n-1})$ — skew-symmetric
3. $X(u_1, \ldots, u_{n-1}) \cdot v = \det(u_1\ \ldots\ u_{n-1}\ v)$
4. $X(u_1, \ldots, u_{n-1}) = 0 \iff \{u_1, \ldots, u_{n-1}\}$ is linearly dependent
5. $X(u_1, \ldots, u_{n-1}) \neq 0 \implies \det(u_1\ \ldots\ u_{n-1}\ X(u_1, \ldots, u_{n-1})) > 0$

Theorem 2.3.3. For $u, v, w, x \in \mathbb{R}^3$, $(u \times v) \times w = (u \cdot w)v - (v \cdot w)u$.
Also, $(u \times v) \cdot (w \times x) = (u \cdot w)(v \cdot x) - (u \cdot x)(v \cdot w)$

3 Applications of the cross product

3.1 Geometry

Definition 3.1.1. Let $u_1, \dots, u_k \in \mathbb{R}^n$. The k-parallelotope on these vectors is the set of points x of the form $x = \sum_{i=1}^{k} t_i u_i$ with $0 \leqslant t_i \leqslant 1$ for all i.
- The points $u_1, \dots, u_k$ are termed vertices of the k-parallelotope
- If $\{u_1, \dots, u_k\}$ is linearly dependent, then the k-parallelotope is termed degenerate

Definition 3.1.2. For a k-parallelotope n $u_1, \dots, u_k \in \mathbb{R}^n$, define the k-volume recursively as follows:
$V_1(u_1) = |u_1|$
$V_k(u_1, \dots, u_k) = |u_k| \sin(\theta) V_{k-1}(u_1, \dots, u_{k-1})$ for $k \geqslant 2$
where θ is the angle from u_k (or span$\{u_k\}$) to span$\{u_1, \dots, u_{k-1}\}$, provided that $u_k \neq 0$ and span$\{u_1, \dots, u_{k-1}\} \neq 0$. If $u_k = 0$ or span$\{u_1, \dots, u_{k-1}\} = 0$, then we define $V_k = 0$.

Theorem 3.1.3. Let $u_1, \dots, u_k \in \mathbb{R}^n$. Then $V_k(u_1, \dots, u_k) = \sqrt{\det(A^t A)}$, where $A = (u_1 \ \dots \ u_k) \in M_{n \times k}$. In particular, $V_k(u_1, \dots, u_k) = 0 \iff \{u_1, \dots, u_k\}$ is linearly dependent.

Corollary 3.1.4. $V_k(u_1, \dots, u_i, \dots, u_j, \dots, u_k) = V_k(u_1, \dots, u_j, \dots, u_i, \dots, u_k)$. Or, the k-volume is independent of the order of vectors.

Corollary 3.1.5. $V_k(u_1, \dots, u_k) = |\det(A)|$

Corollary 3.1.6. $|X(u_1, \dots, u_{n-1})| = V_{n-1}(u_1, \dots, u_{n-1})$

Definition 3.1.7. For $a, b \in \mathbb{R}^n$, the perpendicular bisector of $[a, b]$ is the hyperplane through $\frac{a+b}{2}$ perpendicular to $b - a$. It is the set $\{x \in \mathbb{R}^n \,\big|\, \left(x - \frac{a+b}{2}\right) \cdot (b - a) = 0\}$.

3.2 Spherical geometry

Definition 3.2.1. The (standard) unit sphere in $\mathbb{R}^3$ is the set $\mathbb{S}^2 = \{x \in \mathbb{R}^3 \,\big|\, |x| = 1\}$. More generally, $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n \,\big|\, |x| = 1\}$.

Definition 3.2.2. Given $u \in \mathbb{S}^2$, the line in $\mathbb{S}^2$ with poles $\pm u$ is the set $L_u = \{x \in \mathbb{S}^2 \,\big|\, x \cdot u = 0\} = \mathbb{S}^2 \cap P_u$, where $P_u = \{x \in \mathbb{R}^3 \,\big|\, x \cdot u = 0\}$

Axiom 3.2.3. [Axioms of spherical geometry]
For $u, v \in \mathbb{S}^2$ and $u \neq \pm v$:
1. $L_u = L_v \iff u = \pm v$
2. There exists a unique line on S^2 through u and v, given by $L_w = \frac{u \times v}{|u \times v|}$
3. There exists a unique line on S^2 through v and perpendicular to L_u
4. $L_u \cap L_v = \{\pm w\}$ for some $w \in \mathbb{S}^2$

Definition 3.2.4. The (spherical) distance between $u, v \in \mathbb{S}^2$ is given by $\text{dist}_{\mathbb{S}^2}(u, v) = \text{angle}_{\mathbb{R}^3}(u, v)$.

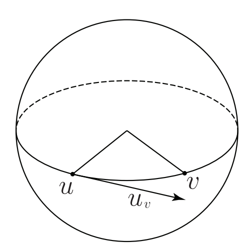
Theorem 3.2.5. For $u, v, w, x \in \mathbb{R}^3$,
$(u \times v) \times w = (u \cdot w)v - (v \cdot w)u$
$(u \times v) \cdot (w \times x) = (u \cdot w)(v \cdot x) - (u \cdot x)(v \cdot w)$

Remark 3.2.6. Properties for spherical distance are identical to properties for distance on the plane.

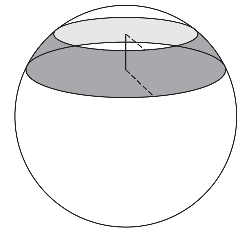
Definition 3.2.7. Given $u \in \mathbb{S}^2$, $r \in (0, \pi)$, the circle on $\mathbb{S}^2$ centered at u of radius r is the set

$$
\begin{aligned}
C(u, r) &= \{x \in \mathbb{S}^2 \,\big|\, \text{dist}(x, u) = r\} \\
&= x \in \mathbb{S}^2 \,\big|\, x \cdot u = \cos(r)\} \\
&= P \cap \mathbb{S}^2 \text{ where } P \text{ is the plane in } R^3 \text{ with equation } x \cdot u = \cos(r)
\end{aligned}
$$

So P is the plane perpendicular to u which goes through the point $\cos(r)u \cdot u$.

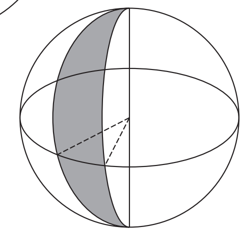


Definition 3.2.8. Given $v \neq \pm u \in \mathbb{S}^2$, define the unit direction vector from u to v to be
$u_v = \frac{(u \times v) \times u}{|(u \times v) \times u|} = \frac{v - \text{Proj}_u(v)}{|v - \text{Proj}_u(v)|} = \frac{v - (v \cdot u)u}{|u \times v|}$



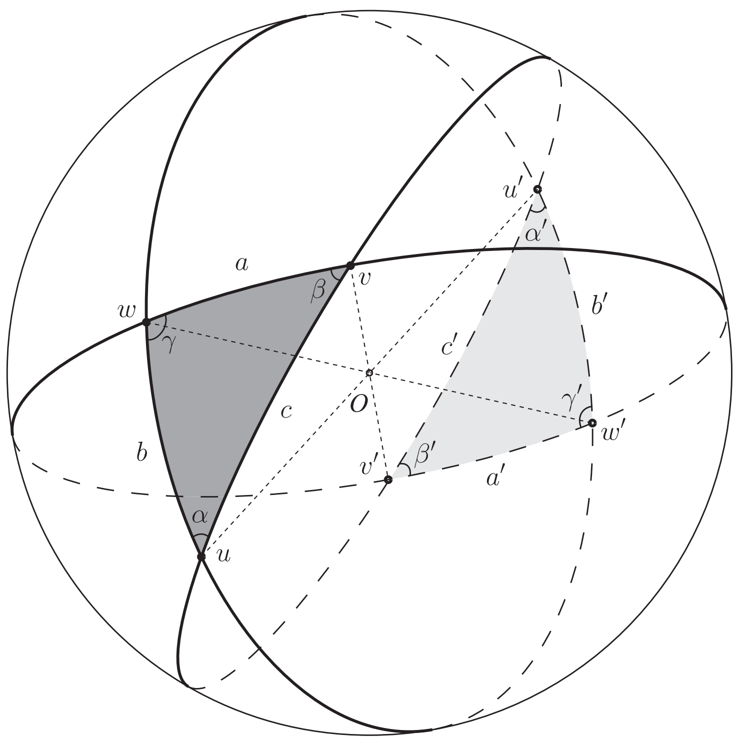
Remark 3.2.9. The set $\{u, u_v\}$ is an orthonormal basis for $\text{span}\{u, v\}$.

Remark 3.2.10. The line segment $[u, v] \in \mathbb{S}^2$ is given parametrically by
$x(t) = cos(t)u + \sin(t)u_v$ with $0 \leqslant t \leqslant \text{dist}(u, v) = \cos^{-1}(u \cdot v)$



Theorem 3.2.11. Two parallel planes a distance $0 \leqslant \ell \leqslant 2r$ apart slicing a sphere of radius r enclose an area of $2\pi r\ell$ on the surface of the sphere.

Theorem 3.2.12. Two planes each bisecting a sphere of radius r with an angle θ to each other enclose an area of $2\theta r^2$ on the surface of the sphere.



With reference to the unit sphere on the left:

$$
\begin{array}{lll}
a = a' & b = b' & c = c' \\
\alpha = \alpha' & \beta = \beta' & \gamma = \gamma' \\
|u| = |u'| & |v| = |v'| & |w| = |w'|
\end{array}
$$

$$
u' = \frac{v \times w}{|v \times w|} \quad v' = \frac{w \times u}{|w \times u|} \quad w' = \frac{u \times v}{|u \times v|}
$$

$[u', v', w']$ is the polar triangle of $[u, v, w]$

Thm 3.1.9. [Spherical law of sines]

$$
\frac{\sin(\alpha)}{\sin(a)} = \frac{\sin(\beta)}{\sin(b)} = \frac{\sin(\gamma)}{\sin(c)}
$$

Thm 3.1.10. [Spherical law of cosines]

$$
\cos(a) = \frac{\cos(\alpha) + \cos(\beta)\cos(\gamma)}{\sin(\beta)\sin(\gamma)}
$$

$$
\cos(\alpha) = \frac{\cos(a) - \cos(b)\cos(c)}{\sin(b)\sin(c)}
$$

3.3 Spherical angles

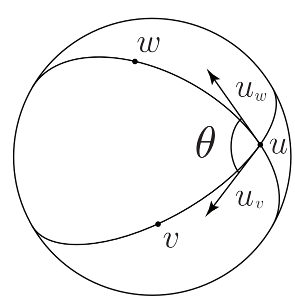
Definition 3.3.1. A non-degenerate triangle on $\mathbb{S}^2$ is determined by 3 non-colinear points $u, v, w \in \mathbb{S}^2$.
Note that u, v, w are colinear $\iff$ u, v, w lie on a plane in $\mathbb{R}^3$ through u
$\iff$ $\{u, v, w\}$ is linearly dependent
$\iff$ $\det(u \;\; v \;\; w) = 0$

Definition 3.3.2. An ordered triangle may be defined as an ordered triple $[u, v, w] = (u, v, w)$ with $u, v, w \in \mathbb{S}^2$ and $\det(u \;\; v \;\; w) \neq 0$. An ordered triangle is positively oriented when $\det(u \;\; v \;\; w) > 0$ and negatively oriented when $\det(u \;\; v \;\; w) < 0$.

Definition 3.3.3. For $u, v, w \in \mathbb{S}^2$ with $v \neq \pm u$ and $w \neq \pm u$, define the oriented angle $\text{angle}(u, v, w)$ to be the angle $\theta \in [0, 2\pi]$ such that

$$
\cos(\theta) = u_v \cdot u_w
$$

$$
\sin(\theta) = (u_v \times u_w) \cdot u = \det(u \;\; u_v \;\; u_w) = \frac{\det(u \;\; v \;\; w)}{|u \times v||u \times w|}
$$



Theorem 3.3.4. Let $[u, v, w]$ be a positively oriented triangle on $\mathbb{S}^2$ with angles α, β, γ. Then the area of $[u, v, w]$ is $A = (\alpha + \beta + \gamma) - \pi$.

4 The inner product

4.1 Fundamental definitions

Definition 4.1.1. Let U be a vector space over $\mathbb{R}$. An inner product on U is a function $\langle\ ,\ \rangle : U^2 \to \mathbb{R}$ such that for all $u, v \in U$ and $c \in \mathbb{R}$

1. $\langle u, u\rangle \geqslant 0$ with $\langle u, u\rangle = 0 \iff u = 0$
2. $\langle u, v\rangle = \langle v, u\rangle$
3. $\langle cu, v\rangle = c\,\langle u, v\rangle = \langle u, cv\rangle$
4. $\langle u + v, w\rangle = \langle u, v\rangle + \langle v, w\rangle$

A vector space closed under an inner product is termed an inner product space.

Definition 4.1.2. Let U be a vector space over $\mathbb{C}$. An inner product on U is a function $\langle\ ,\ \rangle : U^2 \to \mathbb{C}$ such that for all $u, v \in U$ and $c \in \mathbb{C}$

1. $\langle u, u\rangle \in \mathbb{R}$
$\langle u, u\rangle \geqslant 0$ with $\langle u, u\rangle = 0 \iff u = 0$
2. $\langle u, v\rangle = \overline{\langle v, u\rangle}$
3. $\langle cu, v\rangle = c\,\langle u, v\rangle$
$\langle u, cv\rangle = \overline{c}\,\langle u, v\rangle$
4. $\langle u + v, w\rangle = \langle u, v\rangle + \langle v, w\rangle$
$\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$

A vector space over $\mathbb{C}$ closed under an inner product is termed an inner product space over $\mathbb{C}$.

Definition 4.1.3. The vector v^* is termed the conjugate transpose, or the adjoint, or the Hermitian transpose of v, such that $v^* = \overline{v}^t$.

4.2 Standard inner products

Remark 4.2.1. The standard inner product on the following spaces is given by:

on $\mathbb{R}^n$: $\langle u, v\rangle = \displaystyle\sum_{i=1}^{n} u_i v_i = u^t v = v^t u$

on $\mathbb{C}^n$: $\langle u, v\rangle = \displaystyle\sum_{i=1}^{n} u_i \overline{v_i} = u^t \overline{v} = v^* u$

on $M_{m\times n}(\mathbb{R}) : \langle A, B\rangle = \displaystyle\sum_{i,j} A_{ij} B_{ij} = \text{ trace}(A^t B) = \text{ trace}(B^t A)$

on $M_{m\times n}(\mathbb{C}) : \langle A, B\rangle = \displaystyle\sum_{i,j} A_{ij} \overline{B_{ij}} = \text{ trace}(A^t \overline{B}) = \text{ trace}(B^* A)$

on $C[a, b] : \langle f, g\rangle = \displaystyle\int_a^b fg$

Definition 4.2.2. Let U be an inner product space over $\mathbb{F}$. Then for $u \in U$, define the norm or length of u to be $|u| = ||u|| = \sqrt{\langle u, u\rangle}$. Also, a unit vector is a vector of length 1.

Theorem 4.2.3.∗ [Properties of the norm]
Let U be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. Then for $u, v \in U$ and $c \in \mathbb{R}$ or $\mathbb{C}$, we have

1. $|u| \geqslant 0$ with $|u| = 0 \iff u = 0$
2. $|cu| = |c||u|$
3. $|\langle u, v\rangle| \leqslant |u||v|$ with $|\langle u, v\rangle| = |u||v| \iff u, v$ are linearly dependent
4. $|u + v| \leqslant |u| + |v|$

Remark 4.2.4. For a vector space U, a map $|\ | : U \to \mathbb{R}$ which satisfies **1., 2., 3.** above is termed a norm on U.

Theorem 4.2.5. [Polarization identity]
In an inner product space U over $\mathbb{R}$, we have $\langle u, v\rangle = \frac{1}{2}\big(|u + v|^2 - |u - v|^2\big)$.
In an inner product space V over $\mathbb{C}$, we have $\langle u, v\rangle = \frac{1}{4}\big(|u + v|^2 + i|u + iv|^2 - |u - v|^2 - |u - iv|^2\big)$.

Remark 4.2.6. For any non-empty set X a map $d : X \times X \to \mathbb{R}$ which satisfies **1., 2., 3.** above is termed a metric on X.

Definition 4.2.7. Let U be an inner product space over $\mathbb{R}$. For $0 \neq u, v \in U$, define than angle between u and v to be $\text{angle}(u, v) = \cos^{-1}\left(\frac{\langle u,v\rangle}{|u||v|}\right)$

Definition 4.2.8. Let U be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. For $u, v \in U$, we say that u and v are orthogonal if $\langle u, v\rangle = 0$.

Theorem 4.2.9. [Pythagoras]
Let U be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. Let $0 \neq u, v \in U$. Suppose $\langle u, v\rangle = 0$.
Then $|v - u|^2 = |v|^2 + |u|^2$.

4.3 Orthogonal sets / compliments / projections

Definition 4.3.1. Let U be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. A set of vectors $\{u_1, \dots, u_n\}$ in U is termed an orthogonal set when $\langle u_i, u_j\rangle = 0$ for all $i \neq j$, or each pair of vectors is orthogonal. The set is termed orthonormal if $\langle u_i, u_j\rangle = 0$ for all $i \neq j$ and $\langle u_i, u_i\rangle = 1$ for all i.

Remark 4.3.2. Note that $\{u_1, \dots, u_k\} \in \mathbb{R}^n$ is orthogonal $\iff$ A^tA is diagonal for $A = (u_1 \ \dots \ u_k) \in M_{n\times k}$. Similarly, $\{u_1, \dots, u_k\} \in \mathbb{R}^n$ is orthonormal $\iff$ $A^tA = I$ for $A = (u_1 \ \dots \ u_k) \in M_{n\times k}$.

The same may be extended to vectors over $\mathbb{C}^n$, but with conjugate transpose in place of transpose.

Theorem 4.3.3. Let U be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{U} = \{u_1, \dots, u_n\}$ be an orthogonal set of non-zero vectors. Then $\mathcal{U}$ is linearly independent, and also for $x \in \text{span}\{U\}$, $\left([x]_{\mathcal{U}}\right)_k = \frac{\langle x,u_k\rangle}{|u_k|^2}$.

Theorem 4.3.4. [Gram-Schmidt procedure]
Let W be an inner product space. Let $\mathcal{U} = \{u_1, \dots, u_n\}$ be a linearly independent set of vectors in W. So $U = \text{span}(\mathcal{U})$ is an n-dimensional subspace of W. Define vectors $v_1, \dots, v_n$ recursively by

$$v_1 = u_1$$

$$v_k = u_k - \sum_{i=1}^{k-1} \frac{\langle u_k, v_i\rangle}{|v_i|^2} v_i$$

Then for each $k = 1, \dots, n$, the set $\{v_1, \dots, v_k\}$ is an orthogonal set of non-zero vectors with $\text{span}\{v_1, \dots, v_{k-1}\} = \text{span}\{u_1, \dots, u_{k-1}\}$.

Corollary 4.3.5. Every finite-dimensional inner product space has an orthonormal basis.

Corollary 4.3.6. Let W be a finite-dimensional inner product space. Let V be a subspace of W. Then every orthonormal basis of U extends to an orthonormal basis of W.

Definition 4.3.7. Let U and V be inner product spaces over $\mathbb{R}$ or $\mathbb{C}$. An isomorphism (of inner product spaces) from U to V is a map $L : U \to V$ such that L is linear, bijective, and preserves inner products ($\langle L(x), L(y)\rangle = \langle x, y\rangle$ for all $x, y \in U$).

It follows as a consequence that the inverse is also linear and also preserves inner products.

The map need only be onto, because the preservation of inner products implies that it is $1:1$.

Definition 4.3.8. Two inner product spaces U, V are said to be isomorphic when there exists an isomorphism $L : U \to V$.

Corollary 4.3.9. Every n-dimensional inner product space over $\mathbb{F}$ for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, is isomorphic to $\mathbb{F}^n$.

Definition 4.3.10. Let W be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. Let U be a subspace of W. Then the orthogonal compliment of U is the vector space $U^\perp = \{x \in W \mid \langle x, u\rangle = 0 \text{ for all } u \in U\}$.

Definition 4.3.11. Let U be a vector space over $\mathbb{F}$. For a set of vectors $\mathcal{U}$, a linear combination of the elements of $\mathcal{U}$ is always a finite sum of the form $\sum_{i=1}^n c_iu_i$ for $c_i \in \mathbb{F}$ and $u_i \in \mathcal{U}$.

Theorem 4.3.12. [Properties of the orthogonal compliment]
Let W be an inner product space over $\mathbb{R}$ or $\mathbb{C}$, and let U be a subspace of W. Then

1. $U \cap U^{\perp} = \{0\}$
2. $U \subset (U^{\perp})^{\perp}$

If W is finite-dimensional, then we also have

3. If $\mathcal{U} = \{u_1, \dots, u_k\}$ is an orthogonal (orthonormal) basis for U, and $\mathcal{W} = \{u_1, \dots, u_k, v_1, \dots, v_\ell\}$ is an orthogonal (orthonormal) basis for W, then $\mathcal{V} = \mathcal{V} \setminus \mathcal{U} = \{v_1, \dots, v_\ell\}$ is an orthogonal (orthonormal) basis for $U^{\perp}$.
4. If $\mathcal{U} = \{u_1, \dots, u_k\}$ is an orthogonal (orthonormal) basis for U, and $\mathcal{V} = \{v_1, \dots, v_\ell\}$ is an orthogonal (orthonormal) basis for $U^{\perp}$, then $\mathcal{W} = \mathcal{U} \cup \mathcal{V}$ is an orthogonal (orthonormal) basis for W.
5. $\dim(U) + \dim(U^{\perp}) = \dim(W)$
6. Given any $x \in W$, there exist unique vectors $u, v \in W$ with $u \in U$ and $v \in U^{\perp}$ such that $u + v = w$.
7. $W = U \oplus U^{\perp}$

Theorem 4.3.13.∗[Orthogonal projections]
Let W be a (possibly infinite-dimensional) inner product space over $\mathbb{R}$ or $\mathbb{C}$ and let U be a finite dimensional subspace of W. Then given $x \in W$, there exist unique vectors $u, v, \in W$ with $u \in U$ and $v \in U^{\perp}$ such that $u + v = x$. In addition, the vector u is the unique vector in U which is nearest to x.

Moreover, if $\mathcal{U} = \{u_1, \dots, u_n\}$ is any orthogonal basis for U, then $u = \displaystyle\sum_{k=1}^{n} \frac{\langle x, u_k \rangle}{|u_k|^2} u_k$.

Definition 4.3.14. Let W be an inner product space over $\mathbb{R}$ or $\mathbb{C}$ and let U be a finite-dimensional subspace. Given $x \in W$, the unique vector u in the above theorem is termed the orthogonal projection of x onto U, and is expressed $u = \text{Proj}_U(x)$.

4.4 Quotient spaces

Definition 4.4.1. Let W be any vector space over $\mathbb{F}$. Let U be a subspace of W. For any $w \in W$, define the coset of U containing w to be

$$\{w\} + U = \{w + u \,\big|\, u \in U\} = w + U$$

Definition 4.4.2. Let W be any vector space over $\mathbb{F}$. Let U be a subspace of W. Then the quotient space, or the collection of all cosets of U, is the vector space

$$W/U = \{p + U \,\big|\, p \in W\}$$

with
$$\begin{aligned} &(p + U) + (q + U) = (p + q) + U \\ &c(p + U) = cp + U \\ &0 = 0 + U = U \end{aligned}$$

Definition 4.4.3. The codimension of U in W is the dimension of W/U.

Definition 4.4.4. A hyperspace of W is a subspace of codimension 1.

Theorem 4.4.5. Let W be a vector space over $\mathbb{F}$. Let U be a subspace of W. If $\mathcal{U}$ is a basis for U and $\mathcal{U}$ extends to a basis $\mathcal{W}$ for W, and if we let $\mathcal{V} = \mathcal{W} \setminus \mathcal{U}$, then $\{v + U \,\big|\, v \in \mathcal{V}\}$ is a basis for W/U, and the dimension of the quotient space is the number of vectors in $\mathcal{V}$, or the cardinality of $\mathcal{V}$, and $\dim(W/U) = |\mathcal{V}|$. Further, if W is finite dimensional, then $\dim(U)+ \dim(W/U) = \dim(W)$.

Theorem 4.4.6. With respect to the above, $W \cong U \oplus W/U$, or $W \cong U \times W/U$.

Definition 4.4.7. If U, V are subspaces of W with $U \cap V = \{0\}$ such that for all $w \in W$, there exist $u \in U, v \in V$ with $u + v = w$, then W is the internal direct sum of U and V, and we write $W = U \oplus V$.

Definition 4.4.8. Given two vector spaces U, V, the external direct sum (or direct product) of U and V is the vector space

$$U \times V = \{(u,v) \,\big|\, u \in U, v \in V\}$$

with $\quad (u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2)$
$\quad c(u+v) = (cu + cv)$

Remark 4.4.9. If U, V are subspaces of W, then $U \oplus V \cong U \times V$. Also, $U \times \{0\} = \{(u,0) \,\big|\, u \in U\} \subset U \times V$.

Definition 4.4.10. Given a set A and vector spaces U_α with $\alpha \in A$, define the direct sum of the spaces to be the vector space

$$\sum_{\alpha \in A} U_\alpha = \{f : A \to \bigcup_{\alpha \in A} U_\alpha \,\big|\, f(\alpha) \in U_\alpha \text{ for all } \alpha \in A \text{ with } f(\alpha) \neq 0 \text{ for only finitely many } \alpha \in A\}$$

and we define the direct product of the vector spaces U_α to be

$$\prod_{\alpha \in A} U_\alpha = \{f : A \to \bigcup_{\alpha \in A} U_\alpha \,\big|\, f(\alpha) \in U_\alpha \text{ for all } \alpha \in A\}$$

When A is a finite, these are equal. When A is infinite, $\sum\limits_{\alpha \in A} U_\alpha \subsetneq \prod\limits_{\alpha \in A} U_\alpha$

Theorem 4.4.11. Suppose $L : W \to V$ is linear. Then $W/\ker(L) \cong \text{Range}(L)$ is an isomorphism given by $\bar{L} : W/\ker(L) \to \text{Range}(L)$ with $\bar{L}(p + \ker(L)) = L(p) \in L$.

4.5 Dual spaces

Definition 4.5.1. Let U be a vector space over $\mathbb{F}$. The dual vector space of U is the vector space

$$U^* = \text{Lin}(U, \mathbb{F}) = \{f : U \to \mathbb{F} \,\big|\, f \text{ is linear}\}$$

Theorem 4.5.2.$*$ Let U be a finite-dimensional vector space over $\mathbb{F}$. Let $\mathcal{U} = \{u_1, \dots, u_n\}$ be a basis for U. For $k = 1, \dots, n$, define $f_k \in U^*$, so $f_k : U \to \mathbb{F}$, to be the unique linear map with $f_k(u_i) = \delta_{ki}$. Then $\mathcal{F} = \{f_1, \dots, f_n\}$ is a basis for U^*.

Definition 4.5.3. The set $\mathcal{F} = \{f_1, \dots, f_n\}$in the above theorem is termed the dual basis of $\mathcal{U}$ for U.

Then $f = \sum\limits_{k=1}^{n} f(u_k) f_k$.

Definition 4.5.4. Let U, V be vector spaces over $\mathbb{F}$. Let $L : U \to V$ be linear. Define the dual (or the transpose) map $L^t : V^* \to U^*$ given by $L^t(g) = g \circ L$ for all $g \in V^*$.

Theorem 4.5.5. Let U, V be finite dimensional vector spaces over $\mathbb{F}$. Let $L : U \to V$ be linear. $\mathcal{U}, \mathcal{V}$ be bases for U, V. Let $\mathcal{F}, \mathcal{G}$ be the dual bases for U^* and V^*. Then $[L^t]_{\mathcal{F}}^{\mathcal{G}} = \left([L]_{\mathcal{V}}^{\mathcal{U}}\right)^t$

Definition 4.5.6. Let U be a vector space over $\mathbb{F}$. The evaluation map $E : U \to U^{**}$ is given by $E(u)(f) = f(u)$ for all $u \in U$ and $f \in U^*$.

Theorem 4.5.7. Let U be a finite dimensional vector space over $\mathbb{F}$. Then the evaluation map $E : U \to U^{**}$ is a (natural) isomorphism.

Remark 4.5.8. Given a basis $\mathcal{U} = \{u_1, \dots, u_n\}$ for a vector space U, we obtain a (non-natural) isomorphism $L_{\mathcal{U}} : U \to U^*$ given by $L_m F(u_i) = f_i$. This is an isomorphism, since $\mathcal{F} = \{f_1, \dots, f_n\}$ is a basis for U^*.

Theorem 4.5.9. Let U be a finite dimensional inner product space over $\mathbb{R}$ or $\mathbb{C}$. Given $f \in U^*$, there exists a unique vector $u \in U$ such that $f(x) = \langle x, u \rangle$ for all $x \in U$.

Corollary 4.5.10. Let U be a finite dimensional inner product space over $\mathbb{R}$. Then the map $L : U \to U^*$ given by $L(u)(x) = \langle x, u \rangle$ is an isomorphism.

Definition 4.5.11. Let W be a vector space over $\mathbb{F}$. Let U be a subspace of W. Then the annihilator of U in W^* is the space $V^\circ = \{f \in W^* \mid f(u) = 0 \text{ for all } u \in U\}$.

Theorem 4.5.12. Let U, V be finite-dimensional inner product spaces over $\mathbb{R}$ or $\mathbb{C}$. Let $L : U \to V$ be a linear map. Then there exists a unique linear map $L^* : V \to U$ such that $\langle L(x), y\rangle = \langle x, L^*(y)\rangle$ for all $x \in U$ and $y \in V$.

Definition 4.5.13. The above map L^* is termed the adjoint of L. In case U and/or V are infinite dimensional, such a map need not exist, but if it does, then it is termed the adjoint of L.

Corollary 4.5.14. Let U, V be finite dimensional inner product spaces. Let $\mathcal{U}, \mathcal{V}$ be orthonormal bases for U, V. Let $L : U \to V$ be linear. Then $[L^*]_{\mathcal{U}}^{\mathcal{V}} = \left([L]_{\mathcal{V}}^{\mathcal{U}}\right)^*$

4.6 Normal linear maps, etc

Theorem 4.6.1. Let U, V be finite dimensional vector spaces over $\mathbb{F}$. Let $L : U \to V$ be linear with $\operatorname{rank}(L) = r$. Then there exist bases $\mathcal{U}, \mathcal{V}$ for U and V such that $[L]_{\mathcal{V}}^{\mathcal{U}}$ is of the form $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Lemma 4.6.2.∗ For every $A \in M_{n\times n}(\mathbb{F})$ whose characteristic polynomial splits, there exists a unitary matrix P (and so $P^{-1} = P^*$) such that $T = P^*AP$ is upper triangular. Further, the diagonal values of T are the eigenvalues of A, repeated by their algebraic multiplicity.

Theorem 4.6.3.∗ [Schur]
Let U be a finite dimensional inner product space over $\mathbb{R}$ or $\mathbb{C}$. Let $L : U \to U$ be linear. Suppose the characteristic polynomial f_L splits over $\mathbb{F}$ (always occurs for $\mathbb{C}$, for $\mathbb{R}$ only when eigenvalues (roots) are real). Then there exists an orthonormal basis $\mathcal{U}$ such that $T = [L]_{\mathcal{U}}$ is upper triangular. Moreover, the diagonal values of T are the eigenvalues of L, repeated according to their algebraic multiplicity.

Remark 4.6.4. The following statements are equivalent:
- · The linear map L is diagonalizable.
- · There exists a basis of eigenvectors of L for L.
- · $\dim(E_{\lambda_i}) = m_i$ for all i

where E_{λ_i} is the eigenspace of the eigenvalue λ_i of L, and m_i is the algebraic multiplicity of eigenvalue λ_i.

Definition 4.6.5. Let U be a finite-dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $L : U \to V$ be linear. The map L is unitarily triangularziable if there exists an orthonormal basis $\mathcal{U}$ for U such that $[L]_{\mathcal{U}}$ is upper triangular. Similarly, L is unitarily diagonalizable if there exists an orthonormal basis $\mathcal{U}$ for U such that $[L]_{\mathcal{U}}$ is diagonal.

Corollary 4.6.6. [from Schur, for $\mathbb{F} = \mathbb{C}$]
Let U be a finite dimensional inner product space over $\mathbb{C}$. Let $L : U \to U$ be linear.

1. $L^*L = LL^* \iff L$ is unitarily diagonalizable
2. $L^* = L \iff L$ is unitarily diagonalizable and the eigenvalues of L are real.
$L^* = -L \iff L$ is unitarily diagonalizable and the eigenvalues of L are imaginary.
4. $L^*L = I \iff L$ is unitarily diagonalizable and the eigenvalues of L have unit norm.

Corollary 4.6.7. [from Schur, for $\mathbb{F} = \mathbb{R}$]
Let U be a finite dimensional inner product space over $\mathbb{R}$. Let $L : U \to U$ be linear.

1. $L^*L = LL^* \iff L$ is orthogonally diagonalizable
2. $L^* = L$ and $L^*L = I \iff L$ is orthogonally diagonalizable and every eigenvalue of L is ± 1

Definition 4.6.8. Let U be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. Let $L : U \to U$ be linear.
when $L^*L = LL^*$, then L is normal
when $L^* = L$, then L is self-adjoint or Hermitian
when $L^* = L$, then L is skew-Hermitian
when $L^*L = I$, then L is unitary

Remark 4.6.9. For any field $\mathbb{F}$, we have the following matrix groups:

$GL(n,\mathbb{F}) = \{A \in M_{n\times n}(\mathbb{F}) \mid \det(A) \neq 0\}$	general linear group
$SL(n,\mathbb{F}) = \{A \in M_{n\times n}(\mathbb{F}) \mid \det(A) = 1\}$	special linear group - *preserves orientation*
$O(n,\mathbb{F}) = \{A \in M_{n\times n}(\mathbb{F}) \mid A^tA = I\}$	orthogonal group - *preserves distance*
$SO(n,\mathbb{F}) = \{A \in M_{n\times n}(\mathbb{F}) \mid A^tA = I, det(A) = 1\}$	special orthogonal group
$U(n,\mathbb{F}) = \{A \in M_{n\times n}(\mathbb{C}) \mid A^*A = I\}$	unitary group
$SU(n,\mathbb{F}) = \{A \in M_{n\times n}(\mathbb{C}) \mid A^*A = I, \det(A) = 1\}$	special unitary group

Corollary 4.6.10. Let U be a finite-dimensional inner product space over $\mathbb{R}$. Let $L : U \to U$ be linear. Then $L^*L = LL^*$ if and only if there is an orthonormal basis $\mathcal{U}$ for U such that $[L]_{\mathcal{U}}$ is in the block diagonal form

$$[L]_{\mathcal{U}} = \begin{pmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_k & & & & & \\ & & & a_1 & b_1 & & & \\ & & & -b_1 & a_1 & & & \\ & & & & & \ddots & & \\ & & & & & & a_\ell & b_\ell \\ & & & & & & b_\ell & a_\ell \end{pmatrix}$$

where each λ_j is a real eigenvalue, and
each $\mu_j = a_j \pm ib_j$ is a pair of complex eigenvalues

for $k \geqslant 0, \ell \geqslant 0, k + 2\ell = n$

Corollary 4.6.11. For the same conditions as above, if $L^*L = I$, then there exists an orthonormal basis $\mathcal{U}$ for U such that $[L]_{\mathcal{U}}$ has the above form, except each real eigenvalue is ±1, and each block matrix of complex eigenvalues has become the block rotation matrix.

Corollary 4.6.12. If L is orthogonally diagonalizable and $\lambda = \pm1$ for all eigenvalues, the map L represents a reflection in the space spanned by the columns in L with $\lambda = 1$.

Corollary 4.6.13. L is a reflection matrix if and only if $L^* = L$ and $L^*L = I$.

Corollary 4.6.14. L is an orthogonal projection if and only if $L^* = L$ and $L^2 = L$.

Definition 4.6.15. For $\mathcal{U} = \{u_1, \dots, u_n\}$ an orthonormal basis for U a subspace of an inner product space W

the scaling map by λ_k in the direction of u_k is represented by the matrix $[\text{scale}_{\lambda_k,u_k}]_{\mathcal{U}} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \lambda_k & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$

the orthogonal projection map onto span$\{u_k\}$ is given by the matrix $[\text{Proj}_{u_k}]_{\mathcal{U}} = \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & 1 & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix}$

Theorem 4.6.16. [Cayley-Hamilton theorem]
Let U be a finite-dimensional vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and $L : U \to U$ linear. If f_L is the characteristic polynomial of L, then $f_L(L) = 0$.

5 Bilinear and quadratic forms

5.1 Bilinear forms

Definition 5.1.1. Let U be a vector space over $\mathbb{F}$. A bilinear form on U is a map $S : U \times U \to \mathbb{F}$ such that for all $x, y, z \in U$ and $c \in \mathbb{F}$

1. $S(x, y + z) = S(x + y) + S(y + z)$
2. $S(x + y, z) = S(x + z) + S(y + z)$

3. $S(cx, y) = c \cdot S(x, y) = S(x, cy)$
A bilinear form S is symmetric if $S(x, y) = S(y, x)$
A bilinear form is skew-symmetric or alternating if $S(x, y) = -S(y, x)$
A bilinear form is non-degenerate if $S(u, x) = 0$ for all $x \in U \iff u = 0$ for all $u \in U$

Remark 5.1.2. If $\mathcal{U}$ is a basis for U, then a bilinear form S on U is determined completely by the values $S(u, v)$ for $u, v \in \mathcal{U}$. Indeed, if we have $x = \sum_{i=1}^n t_i u_i$ and $y = \sum_{j=1}^n r_j u_j$ for $u_i, u_j \in \mathcal{U}$, then

$$S(x, y) = S\left(\sum_{i=1}^n t_i u_i, \sum_{j=1}^n r_i u_i\right) = \sum_{i,j} t_i r_j S(u_i, u_j)$$

Note that this argument also holds for the infinite-dimensional case, since linear combinations are still finite.

Remark 5.1.3. $\mathrm{Bilin}(U \times U) \cong \prod_{(u,v) \in U \times U} \mathbb{F}$

Definition 5.1.4. Let U be a finite dimensional vector space over $\mathbb{F}$. Let $S : U \times U \to \mathbb{F}$ be a bilinear form. Let $\mathcal{U}$ be a basis for U. Then the matrix of S with respect to the basis $\mathcal{U}$ is defined to be the matrix $[S]^{\mathcal{U}}$ such that $S(u, v) = [u]_{\mathcal{U}}^t [S]^{\mathcal{U}} [v]_{\mathcal{U}}$. Furthermore, the (i, j) entry of $[S]^{\mathcal{U}}$ is $S(u_i, u_j)$.

Remark 5.1.5. Let U be a finite-dimensional vector space over $\mathbb{F}$. Let $S : U \times U$ be a bilinear form. Let $\mathcal{U}, \mathcal{V}$ be bases for U. Then $[S]^{\mathcal{V}} = [I]_{\mathcal{U}}^{\mathcal{V}\,t} [S]^{\mathcal{U}} [I]_{\mathcal{U}}^{\mathcal{V}}$.

Definition 5.1.6. For $A, B \in M_{n \times n}(\mathbb{F})$, we say that A and B are congruent if there exists an invertible matrix Q such that $B = Q^t A Q$. Note that congruent matrices have the same rank.

Definition 5.1.7. The rank of a bilinear form S on a finite dimensional vector space U is the rank of $[S]^{\mathcal{U}}$ for any basis $\mathcal{U}$ of U.

Remark 5.1.8. A bilinear form S on a finite-dimensional vector space U is symmetric $\iff$ the matrix $[S]^{\mathcal{U}}$ is symmetric for any basis $\mathcal{U}$ of U.

Theorem 5.1.9. Let U be a finite-dimensional vector space over $\mathbb{F}$. Let S be a symmetric bilinear form.
1. If $\mathrm{char}(\mathbb{F}) \neq 2$, (that is, $1 + 1 \neq 0$), then there exists a basis $\mathcal{U}$ for U such that $[S]^{\mathcal{U}}$ is diagonal.
2. If $\mathbb{F} = \mathbb{C}$, then there exists a basis $\mathcal{U}$ such that $[S]^{\mathcal{U}} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ for $r = \mathrm{rank}(S)$.
3. If $\mathbb{F} = \mathbb{R}$, then there exists a basis $\mathcal{U}$ for U such that $[S]^{\mathcal{U}} = \begin{pmatrix} I_k & & \\ & -I_{r-k} & \\ & & 0 \end{pmatrix}$ for some k.
4. If $\mathbb{F} = \mathbb{R}$, then there exists an orthonormal basis U for U such that $[S]^{\mathcal{U}} = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_k & \\ & & & 0 \end{pmatrix}$ for non-zero eigenvalues $\lambda_1, \ldots, \lambda_k$ of $[S]^{\mathcal{U}}$.
5. If $\mathbb{F} = \mathbb{R}$ and $D = [S]^{\mathcal{U}}$ is diagonal for $\mathcal{U}$ a basis for U, then the number of positive entries of D does not depend on $\mathcal{U}$.

Theorem 5.1.10. [Sylvester]
Let U be a finite-dimensional vector space over U. Let $S : U \times U \to \mathbb{R}$ be a symmetric bilinear form. Let $\mathcal{U}$ and $\mathcal{V}$ be two bases for U such that $[S]^{\mathcal{U}}$ and $[S]^{\mathcal{V}}$ are both diagonal. Then the number of positive entries in $[S]^{\mathcal{U}}$ is the number of positive entries in $[S]^{\mathcal{V}}$.

Remark 5.1.11. We write $\mathrm{Bilin}(U) = \mathrm{Bilin}(U \times U, \mathbb{F})$ for the space of bilinear forms $S : U \times U \to \mathbb{F}$. Given a basis $\mathcal{U}$ of n-dimensional U, the map $\psi_n : \mathrm{Bilin}(U) \to M_{n \times n}(\mathbb{F})$ is a vector space isomorphic map.

Remark 5.1.12. An inner product in a real inner product space is a positive definite symmetric bilinear form. Also, a bilinear form $S : U \times U \to \mathbb{R}$ is non-degenerate when $S(u, x) = 0$ for all $x \in U \iff u = 0$ for all $u \in U$.

5.2 Quadratic forms

Definition 5.2.1. A polynomial $f \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ is of the form

$$f(x_1, x_2, \ldots, x_n) = \sum_{0 \leqslant i_1, \ldots, i_n} a_{i_1, \ldots, i_n} x_1^{i_1} \cdots x_n^{i_n}$$

with only finitely many of the $a_{i_1, \ldots, i_n} = 0$.

Definition 5.2.2. A polynomial homogeneous of degree d may be expressed as

$$K(x) = \sum_{d=0}^{m} \left(\sum_{\substack{0 \leqslant i_1, \ldots, i_n \\ i_1 + \cdots + i_n = d}} a_{i_1, \ldots, i_n} x_1^{i_1} \cdots x_n^{i_n} \right)$$

Definition 5.2.3. Let U be a vector space over $\mathbb{F}$. A quadratic form on U is a map $K : U \to \mathbb{F}$ of the form $K(u) = S(u, u)$ for some symmetric bilinear form S. If $\mathrm{char}(\mathbb{F}) \neq 2$, then
$K(u + v) = S(u + v, u + v) = S(u, u) + 2S(u, v) + S(v, v) = K(u) + 2S(u, v) + K(v)$

Theorem 5.2.4. A quadratic form may be diagonalized if $\mathrm{char}(\mathbb{F}) \neq 2$.

Theorem 5.2.5. Let U be an n-dimensional vector space over $\mathbb{R}$. Let $K : U \to \mathbb{R}$ be a quadratic form on U, and let $S : U \times U \to \mathbb{R}$ be the corresponding symmetric bilinear form. Then the following are equivalent:

1. K (or S) is positive definite
2. the eigenvalues of $[K]^{\mathcal{U}} = [S]^{\mathcal{U}}$ are all positive for some (hence any) basis $\mathcal{U}$ for U
3. for $A = [K]^{\mathcal{U}} = [S]^{\mathcal{U}}$ we have $\det(A^{k \times k}) > 0$ with $1 \leqslant k \leqslant n$

Remark 5.2.6. For $A \in M_{n \times m}(\mathbb{F})$ the notation $A^{k \times \ell}$ denotes the $k \times \ell$ upper left submatrix of A such that $1 \leqslant k \leqslant n$ and $1 \leqslant \ell \leqslant m$.

5.3 Characterization and extreme values

Recall that $K : U \to \mathbb{R}$ or $S : U \times U \to \mathbb{R}$ is positive definite or symmetric bilinear when $K(u) = S(u, u) > 0$ for $u \neq 0$.

Theorem 5.3.1.∗[Characterization of positive definite forms]
Let U be an n-dimensional inner product space over $\mathbb{R}$. Let $K : U \to \mathbb{R}$ be a quadratic form on U, and let $S : U \times U \to \mathbb{R}$ be the corresponding symmetric bilinear form. Then the following statements are equivalent:

1. K (or S) is positive definite
2. the eigenvalues of $[K]^{\mathcal{U}} = [S]^{\mathcal{U}}$ are all positive for some (hence any) basis $\mathcal{U}$ for U
3. for $A = [K]^{\mathcal{U}} = [S]^{\mathcal{U}}$ we have $\det(A^{k \times k}) > 0$

where $A^{k \times k}$ represents the $k \times k$ upper-left submatrix of A.

Theorem 5.3.2. Let $A \in M_{n \times n}(\mathbb{F})$. Suppose $A^* = A$. Recall that the eigenvalues of A are real. Let $\lambda_1, \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ be the eigenvalues of A listed according to algebraic multiplicity in increasing order. Then $\max\limits_{|x|=1}\{x^* A x\} = \lambda_n$ and $\min\limits_{|x|=1}\{x^* A x\} = \lambda_1$.

Corollary 5.3.3. Let U be an n-dimensional inner product space over $\mathbb{R}$. Let $S : U \times U \to \mathbb{R}$ be a symmetric bilinear form and let $K : U \to \mathbb{R}$ be the corresponding quadratic form on U. Let $\lambda_1, \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ be the eigenvalues listed according to algebraic multiplicity in increasing order, of $[K]^{\mathcal{U}} = [S]^{\mathcal{U}}$ for some (hence any) orthonormal basis $\mathcal{U}$ for U. Then $\max\limits_{|u|=1}\{K(u)\} = \lambda_n$ and $\min\limits_{|u|=1}\{K(u)\} = \lambda_1$.

Definition 5.3.4. Let U, V be inner product spaces over $\mathbb{F}$. If a map $L : U \to V$ has an adjoint, then define the singular values of L to be the square roots of the eigenvalues of $L^* L$.

Let U, V be finite dimensional inner product spaces over $\mathbb{F}$. Let $L : U \to V$ be linear. Let $0 \leqslant \sigma_1 \leqslant \sigma_2 \leqslant \cdots \leqslant \sigma_n$ be the singular values of L listed in increasing order, repeated according to algebraic multiplicity. Then $\max\limits_{|u|=1}\{|L(u)|\} = \sigma_n$ and $\min\limits_{|u|=1}\{|L(u)|\} = \sigma_1$.

Definition 5.3.5. The spectrum of a linear map $L : U \to U$ over an inner product space U is the set of eigenvalues of L.

Theorem 5.3.6.$*$ Let U, V be inner product spaces over $\mathbb{F}$. Let $L : U \to V$ be linear. Then there exist orthonormal bases $\mathcal{U}, \mathcal{V}$ for U, V such that $[L]_{\mathcal{V}}^{\mathcal{U}}$ is in the form

$$[L]_{\mathcal{V}}^{\mathcal{U}} = \Sigma = \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right)$$

Corollary 5.3.7. For $A \in M_{m\times n}(\mathbb{F})$, there exist $P \in M_{m\times n}(\mathbb{F})$ and $Q \in M_{n\times n}(\mathbb{F})$ with $P^*P = I_m$ and $Q^*Q = I_n$ such that

$$P^*AQ = \Sigma = \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right)$$

This is termed the singular value decomposition of $A = [L]^{\mathcal{U}}$ with the singular values as described above.

6 Jordan normal form

6.1 Block form

Definition 6.1.1. The $m \times m$ Jordan block for the eigenvalue $\lambda \in \mathbb{F}$ over $\mathbb{F}$ is the $m \times m$ matrix

$$J_\lambda^m = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

Definition 6.1.2. A matrix $B \in M_{n\times n}(\mathbb{F})$ is in Jordan form when it is in the block diagonal form

$$B = \begin{pmatrix} J_{\lambda_1}^{m_1} & & & \\ & J_{\lambda_2}^{m_2} & & \\ & & \ddots & \\ & & & J_{\lambda_\ell}^{m_\ell} \end{pmatrix}$$

6.2 Canonical form

Theorem 6.2.1. Let U be a finite-dimensional vector space over $\mathbb{F}$. Let $L : U \to U$ be linear. Suppose that the characteristic polynomial $f_L(t)$ of L splits over $\mathbb{F}$. Then there exists a basis $\mathcal{U}$ for U such that $[L]_{\mathcal{U}} = B$ is in Jordan form. The matrix B is uniquely determined by L up to the order of the Jordan blocks.

Definition 6.2.2. A generalized eigenvector of a map $L : U \to U$ for an eigenvalue λ of L is a non-zero vector $u \in U$ such that $(L - \lambda I)^P u = 0$ for some $p \geqslant 0$.

Definition 6.2.3. A cycle of generalized eigenvectors of length m for the eigenvalue λ is an ordered set of vectors $C = \{u_1, \ldots, u_m\}$ such that

$$\begin{aligned} u_{m-1} &= (L - \lambda I)u_m \\ u_{m-2} &= (L - \lambda I)^2 u_m \\ &\vdots \\ u_1 &= (L - \lambda I)^{m-1} u_m \\ 0 &= (L - \lambda I)^m u_m \end{aligned}$$

Definition 6.2.4. The generalized eigenspace for λ is $K_\lambda = \{u \in U \mid (L - \lambda I)^p u = 0 \text{ for some } p \geqslant 0\}$

Theorem 6.2.5.$*$ Let U be a finite-dimensional vector space over $\mathbb{F}$ with $L : U \to U$ linear. Then for every eigenvalue λ of L, there exists a basis of cycles corresponding to λ for K_λ.

Definition 6.2.6. Let $L : U \to U$ for U a finite-dimensional vector space over $\mathbb{F}$ be linear. The minimal polynomial of L is the unique monic polynomial $f_L(x)$ of minimum possible degree such that $f_L(L) = 0$.

Note that the minimal polynomial is always a factor of the characteristic polynomial, and the roots of the minimal polynomial are the same as the roots of the characteristic polynomial.

7 Selected proofs

Theorem 4.2.3. [Properties of the norm]
Let U be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. Then for $u, v \in U$ and $c \in \mathbb{R}$ or $\mathbb{C}$, we have

1. $|u| \geqslant 0$ with $|u| = 0 \iff u = 0$
2. $|cu| = |c||u|$
3. $|\langle u, v\rangle| \leqslant |u||v|$ with $|\langle u, v\rangle| = |u||v| \iff u, v$ are linearly dependent
4. $|u + v| \leqslant |u| + |v|$

Proof: For **2.**:

$$|cu|^2 = \langle cu, cu\rangle = c\,\langle u, cu\rangle = c\bar{c}\,\langle u, u\rangle = |c|^2\,|u|^2 \quad \implies \quad |cu| = |c||u|$$

For **3.**: Suppose $\{u, v\}$ is linearly dependent, say $u = cv$ for $c \in \mathbb{C}$.

$$|\langle u, v\rangle| = |c\,\langle v, v\rangle| = |c||v|^2 = |cv||v| = |u||v|$$

Suppose $\{u, v\}$ is linearly independent.

$$\langle v - \mathrm{Proj}_u v, u\rangle = \left\langle v - \frac{\langle v, u\rangle}{\langle u, u\rangle}u, u\right\rangle = \langle v, u\rangle - \left\langle \frac{\langle v, u\rangle}{\langle u, u\rangle}u, u\right\rangle = \langle v, u\rangle - \frac{\langle v, u\rangle}{\langle u, u\rangle}\,\langle u, u\rangle = 0$$

Since $\{u, v\}$ is linearly independent, $v - \frac{\langle v,u\rangle}{\langle u,u\rangle}u \neq 0$, so

$$0 < \left|v - \frac{\langle v, u\rangle}{\langle u, u\rangle}u\right|^2 = \left\langle v - \frac{\langle v, u\rangle}{\langle u, u\rangle}u, v - \frac{\langle v, u\rangle}{\langle u, u\rangle}u\right\rangle = \left\langle v - \frac{\langle v, u\rangle}{\langle u, u\rangle}u, v\right\rangle - 0 = \langle v, v\rangle - \frac{\langle v, u\rangle}{\langle u, u\rangle}\,\langle u, v\rangle$$

$$\begin{aligned}
\langle v, u\rangle\,\langle u, v\rangle &< \langle u, u\rangle\,\langle v, v\rangle \\
\langle v, u\rangle\,\overline{\langle v, u\rangle} &< |u|^2\,|v|^2 \\
|\langle u, v\rangle|^2 &< |u|^2\,|v|^2 \\
|\langle u, v\rangle| &< |u|\,|v|
\end{aligned}$$

For **4.**:

$$\begin{aligned}
|u + v|^2 &= \langle u + v, u + v\rangle \\
&= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\
&= |u|^2 + \langle u, v\rangle + \overline{\langle u, v\rangle} + |v|^2 \\
&= |u|^2 + 2\mathrm{Re}(\langle u, v\rangle) + |v|^2 \\
&\leqslant |u|^2 + 2\,|\mathrm{Re}(\langle u, v\rangle)| + |v|^2 \\
&\leqslant |u|^2 + 2\,|\langle u, v\rangle| + |v|^2 \\
&\leqslant |u|^2 + 2\,|u||v| + |v|^2 \\
&= (|u| + |v|)^2 \\
|u + v| &\leqslant |u| + |v|
\end{aligned}$$

Theorem 4.2.2. Let U be a finite-dimensional vector space over $\mathbb{F}$. Let $\mathcal{U} = \{u_1, \ldots, u_n\}$ be a basis for U. For $k = 1, \ldots, n$, define $f_k \in U^*$, so $f_k : U \to \mathbb{F}$, to be the unique linear map with $f_k(u_i) = \delta_{ki}$. Then $\mathcal{F} = \{f_1, \ldots, f_n\}$ is a basis for U^*.

Proof: It is claimed that $\mathcal{F}$ is linearly independent.
Suppose that $\sum_{i=1}^n c_i f_i = 0$
Then $\sum_{i=1}^n c_i f_i(x) = 0$ for all $x \in U$, in particular for all $k = 1, 2, \ldots, n$, so $0 = \sum_{i=1}^n c_i f_i(u_k) = c_k$
It is claimed that $\mathcal{F}$ spans U^*.
Let $g \in U^*$. That is, $g : U \to \mathbb{F}$ is linear.
It is claimed that $g = \sum_{i=1}^n g(u_i) f_i$
Indeed, for each $k = 1, 2, \ldots, n$ we have

$$\left(\sum_{i=1}^n g(u_i) f_i\right)(u_k) = \sum_{i=1}^n g(u_i) f(u_k) = g(u_k)$$

Therefore $g = \sum_{i=1}^n g(u_i) f_i$ as claimed.

Theorem 4.3.13. [Orthogonal projections]
Let W be a (possibly infinite-dimensional) inner product space over $\mathbb{R}$ or $\mathbb{C}$ and let U be a finite dimensional subspace of W. Then given $x \in W$, there exist unique vectors $u, v, \in W$ with $u \in U$ and $v \in U^\perp$ such that $u + v = x$. In addition, the vector u is the unique vector in U which is nearest to x.
Moreover, if $\mathcal{U} = \{u_1, \ldots, u_n\}$ is any orthogonal basis for U, then $u = \displaystyle\sum_{k=1}^n \frac{\langle x, u_k\rangle}{|u_k|^2} u_k$.

Proof: *Uniqueness:* Suppose $u, v, x \in W$ with $u \in U$, $v \in U^\perp$ and $u + v = x$.
Let $\mathcal{U} = \{u_1, \ldots, u_n\}$ be an orthogonal basis for U.
Then $\langle x, u_k\rangle = \langle u + v, u_k\rangle = \langle u, u_k\rangle + \langle v, u_k\rangle = \langle u, u_k\rangle$
Therefore $u = \displaystyle\sum_{k=1}^n \frac{\langle u, u_k\rangle}{|u_k|^2} u_k = \sum_{k=1}^n \frac{\langle x, u_k\rangle}{|u_k|^2} u_k$
And so we have $v = x - u$.
So u and v are uniquely determined in terms of x and $\mathcal{U}$.

Existence: Let x be given.
Let $u = \sum_{k=1}^n \frac{\langle x, u_k\rangle}{|u_k|^2} u_k$ and let $v = x - u$.
Clearly $u \in \text{span}\{U\}$ and $u + v = x$.
To show that $v \in U^\perp$, note that for each k we have

$$\begin{aligned}
\langle v, u_k\rangle &= \langle x - u, u_k\rangle \\
&= \langle x, u_k\rangle - \langle u, u_k\rangle \\
&= \langle x, u_k\rangle - \left\langle \sum_{i=1}^n \frac{\langle x, u_i\rangle}{|u_i|^2} u_i, u_k \right\rangle \\
&= \langle x, u_k\rangle - \sum_{i=1}^n \frac{\langle x, u_i\rangle}{|u_i|^2} \langle u_i, u_k\rangle \\
&= \langle x, u_k\rangle - \frac{\langle x, u_k\rangle}{|u_k|^2} \langle u_k, u_k\rangle \\
&= 0
\end{aligned}$$

Finally, by Pythagoras' theorem, u is the unique point in U nearest to x.

Lemma 4.6.2. For every $A \in M_{n\times n}(\mathbb{F})$ whose characteristic polynomial splits, there exists a unitary matrix P (and so $P^{-1} = P^*$) such that $T = P^*AP$ is upper triangular. Further, the diagonal values of T are the eigenvalues of A, repeated by their algebraic multiplicity.

Proof: This will be done by induction on n.

For $n = 1$, this is clearly true, and we take $P = I = \begin{bmatrix}1\end{bmatrix}$.

Suppose that for every $(n-1)\times(n-1)$ matrix B whose characteristic polynomial splits, there is an $(n-1)\times(n-1)$ unitary matrix Q such that Q^*BQ is upper triangular.

Let $A \in M_{n\times n}(\mathbb{F})$ such that its characteristic polynomial splits.

Let λ_1 be an eigenvalue of A with u_1 the corresponding eigenvector such that $|u_1| = 1$.

Extend $\{u_1\}$ to an orthonormal basis $\{u_1, \ldots, u_n\}$ for $\mathbb{F}^n$.

Let $P = (u_1 \;\; \ldots \;\; u_n)$.

Since $\{u_1, \ldots, u_n\}$ is orthonormal, we have $P^*P = I$, so P is unitary.

Then we have

$$\begin{aligned} P^*AP = \begin{pmatrix} u_1^* \\ \vdots \\ u_n^* \end{pmatrix} A(u_1 \;\; \ldots \;\; u_n) &= \begin{pmatrix} u_1^* \\ \vdots \\ u_n^* \end{pmatrix} \begin{pmatrix}\lambda_1 u_1 & A(u_1 \;\; \ldots \;\; u_n)\end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 u_1^* u_1 & u_1^* A(u_2 \;\; \ldots \;\; u_n) \\ \lambda_1 u_2^* u_1 & \multirow{3}{*}{$\begin{pmatrix} u_2^* \\ \vdots \\ u_n^* \end{pmatrix} A(u_2 \;\; \ldots \;\; u_n)$} \\ \vdots & \\ \lambda_1 u_n^* u_1 & \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & X \\ 0 & B \end{pmatrix} \end{aligned}$$

Since $\left(\begin{smallmatrix} \lambda_1 & X \\ 0 & B \end{smallmatrix}\right)$ is similar to A, they share a common characteristic polynomial, so

$$(\lambda_1 - t)\det(B - tI) = f_A(t) = (-1)^n (t-\lambda_1)^{k_1} \cdots (t-\lambda_\ell)^{k_\ell}$$

Therefore $f_B(t) = (-1)^{n+1}(t-\lambda_1)^{k_1-1}(t-\lambda_2)^{k_2}\cdots(t-\lambda_\ell)^{k_\ell}$, so it splits.

By the induction hypothesis, we can choose $Q \in M_{(n-1)\times(n-1)}$ with $Q^{-1} = Q^*$, so that Q^*BQ is upper triangular.

Then we have

$$\begin{pmatrix} 1 & 0 \\ 0 & Q^* \end{pmatrix} P^*AP \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} = \begin{pmatrix} \lambda_1 & XQ \\ 0 & Q^*BQ \end{pmatrix}$$

Theorem 4.6.3. [Schur]
Let U be a finite dimensional inner product space over $\mathbb{R}$ or $\mathbb{C}$. Let $L : U \to U$ be linear. Suppose the characteristic polynomial f_L splits over $\mathbb{F}$ (always occurs for $\mathbb{C}$, for $\mathbb{R}$ only when eigenvalues (roots) are real). Then there exists an orthonormal basis $\mathcal{U}$ such that $T = [L]_{\mathcal{U}}$ is upper triangular. Moreover, the diagonal values of T are the eigenvalues of L, repeated according to their algebraic multiplicity.

Proof: Let $\mathcal{U}_o$ be an orthonormal basis for U.

Let $A = [L]_{\mathcal{U}_o}$.

Note that $f_A(t) = f_L(t)$.

Choose $P \in M_{n\times n}(\mathbb{F})$ (for $n = \dim(U)$) with $P^*P = I$, so that P^*AP is upper triangular.

Let $\mathcal{U}$ be the basis for U such that $[I]_{\mathcal{U}_o}^{\mathcal{U}} = P$.

Then $\begin{aligned}[t] [L]_{\mathcal{U}} &= [I]_{\mathcal{U}}^{\mathcal{U}_o}[L]_{\mathcal{U}_o}^{\mathcal{U}_o}[I]_{\mathcal{U}_o}^{\mathcal{U}} \\ &= P^{-1}AP \\ &= P^*AP \end{aligned}$

And we have that $\mathcal{U}$ is orthonormal since $P^*P = I$.

Indeed, if $u_k, u_\ell \in U$, then $\begin{aligned}[t] \langle u_k, u_\ell\rangle &= \langle [u_k]_{\mathcal{U}_o}, [u_\ell]_{\mathcal{U}_o}\rangle \\ &= \langle k\text{th column of } P, \ell\text{th column of } P\rangle \\ &= \delta_{k,\ell} \end{aligned}$

Theorem 5.3.6. [Characterization of positive definite forms]
Let U be an n-dimensional inner product space over $\mathbb{R}$. Let $K : U \to \mathbb{R}$ be a quadratic form on U, and let $S : U \times U \to \mathbb{R}$ be the corresponding symmetric bilinear form. Then the following statements are equivalent:

1. K (or S) is positive definite

2. the eigenvalues of $[K]^{\mathcal{U}} = [S]^{\mathcal{U}}$ are all positive for some (hence any) basis $\mathcal{U}$ for U

3. for $A = [K]^{\mathcal{U}} = [S]^{\mathcal{U}}$ we have $\det(A^{k\times k}) > 0$

Proof: 1. $\Longrightarrow$ **2**. Suppose S is positive definite.

Let $\mathcal{U}$ be a basis for U and $A = [S]^{\mathcal{U}}$, so that $S(u, v) = [u]_{\mathcal{U}}^t [S]^{\mathcal{U}} [v]_{\mathcal{U}} = x^t Ay$.

Since $S(u, u) > 0$ for all $0 \neq u \in U$, $x^t Ax > 0$ for all $x \neq 0$.

Let λ be an eigenvalue of A.

Let x be an eigenvector of A so that $Ax = \lambda x$.

Then we have

$$x^t Ax = x^t \lambda x = \lambda x^t x = \lambda |x|^2$$

Therefore $\lambda = \frac{x^t Ax}{|x|^2} > 0$

2. $\Longrightarrow$ **1**. Suppose that the eigenvalues of $[S]^{\mathcal{U}} = A$ are all positive for some basis $\mathcal{U}$ of U.

Since S is symmetric, A is symmetric, and so A is orthogonally diagonalizable.

Suppose $P^*AP = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ for P unitary and $\lambda_i > 0$ for $a \leqslant i \leqslant n$ and $P \in M_{n\times n}(\mathbb{R})$.

So $A = PDP^*$, and

$$x^t Ax = x^t PDP^* x = y^t Dy = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 > 0 \quad \text{for} \quad y \neq 0$$

1. $\Longrightarrow$ **3**. Suppose S is positive definite.

Let $\mathcal{U}$ be a basis for U and let $A = [S]^{\mathcal{U}}$.

Since S is positive definite, $x^t Ax > 0$ for all $x = [u]_{\mathcal{U}} \neq 0$ and $x \in \mathbb{R}^n$.

For $k = 1, \ldots, n$, $\begin{pmatrix} x \\ 0 \end{pmatrix}^t A \begin{pmatrix} x \\ 0 \end{pmatrix}$ for all $x \in \mathbb{R}^k$.

So the matrix $A^{k\times k}$ is positive definite.

So the eigenvalues of this $k \times k$ submatrix are all positive, so $\det(A^{k\times k}) > 0$, since the determinant of a diagonalizable matrix is the product of the eigenvalues.

3. $\Longrightarrow$ **1**. Suppose $\det(A^{k\times k}) > 0$ for $k = 1, \ldots, n$.

Let $\mathcal{U}$ be a basis for U and let $A = [S]^{\mathcal{U}}$.

Consider the algorithm used to diagonalize a symmetric matrix (or bilinear form) by using row and column operations.

Since $\det(A^{k\times k}) > 0$, we have $A_{11} > 0$ in the form $\left(\begin{array}{c|c} A_{11} & * \\ \hline * & * \end{array}\right)$

Now eliminate $A_{1i} = A_{i1}$ for $i = 2, \ldots n$ by using $C_i \mapsto C_i - \frac{A_{1i}}{A_{11}} C_1$ and $R_i \mapsto R_i - \frac{A_{i1}}{A_{11}} R_1$, so now the matrix is of the form $\left(\begin{array}{c|c} A_{1,1} & 0 \\ \hline 0 & B \end{array}\right)$

So we have $\det(A^{k\times k}) = A_{11} \cdot \det(B^{(k-1)\times(k-1)})$, so $\det(B^{j\times j}) = \dfrac{\det(A^{(j+1)\times(j+1)})}{A_{11}} > 0$

By repeating the procedure, we obtain an invertible matrix Q such that

$$Q^t AQ = D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \text{ with } d_i > 0 \text{ for all } i$$

Then we have

$$x^t Ax = x^t (Q^{-1})^t DQ^{-1} x = y^t Dy = d_1 y_1^2 + \cdots d_n y_n^2 > 0$$

Theorem 5.3.6. Let U, V be inner product spaces over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $L : U \to V$ be linear. Then there exist orthonormal bases $\mathcal{U}, \mathcal{V}$ for U, V such that $[L]_{\mathcal{V}}^{\mathcal{U}}$ is in the form

$$[L]_{\mathcal{V}}^{\mathcal{U}} = \Sigma = \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right)$$

where $\sigma_1, \ldots, \sigma_r$ are the singular values of L.

Proof: <u>*Uniqueness:*</u> Suppose $\mathcal{U}, \mathcal{V}$ are orthonormal bases of U, V such that $[L]_{\mathcal{V}}^{\mathcal{U}}$ is in the form above.
Note that $r = \text{rank}(L)$.
For $\mathcal{U} = \{u_1, \ldots, u_n\}$ and $\mathcal{V} = \{v_1, \ldots, v_m\}$ we have $L(u_i) = \begin{cases} \sigma_i v_i & \text{for } 1 \leqslant i \leqslant r \\ 0 & \text{for } r+1 \leqslant i \leqslant n \end{cases}$
Note that we also have $[L^*]_{\mathcal{U}}^{\mathcal{V}} = \left([L]_{\mathcal{V}}^{\mathcal{U}}\right)^* \in M_{n\times m}(\mathbb{F})$.
Therefore we have $L^*(v_i) = \begin{cases} \sigma_i u_i & \text{for } 1 \leqslant i \leqslant r \\ 0 & \text{for } r+1 \leqslant i \leqslant m \end{cases}$
Therefore $\{v_1, \ldots, v_r\}$ is a basis for range(L) and $\{v_{r+1}, \ldots, v_m\}$ is a basis for range$(L)^\perp$.
Since $L(u_i) = \sigma_i v_i$ and $L^*(v_i) = \sigma_i u_i$ for $1 \leqslant i \leqslant r$

$$L^*(L(u_i)) = L^*(\sigma_i v_i) = \sigma_i \sigma_i u_i = \sigma_i^2 u_i$$

So for $1 \leqslant i \leqslant r$, $\lambda_i = \sigma_i^2$ is an eigenvalue of L^*L and u_i is the corresponding eigenvector.
Note also that $\text{rank}(L^*L) = \text{rank}(L)$, with $\text{null}(L^*L) = \text{null}(L)$.
Therefore for $r+1 \leqslant i \leqslant n$ we take $\lambda_i = 0$ since $\sigma_i = 0$.

<u>*Existence:*</u> Given $L : U \to V$ linear, consider $L^*L : U \to U$.
Since $(L^*L)^* = L^*L$, L^*L has non-negative real eigenvalues.
Let $\lambda_1 \geqslant \cdots \geqslant \lambda_r > 0$ and $\lambda_{r+1} = \cdots = \lambda_n = 0$ be the eigenvalues, so $r = \text{rank}(L^*L) = \text{rank}(L)$.
The map L^*L can then be orthogonally diagonalized with an orthonormal basis of eigenvectors.
Let $\mathcal{U} = \{u_1, \ldots, u_n\}$ for L^*L be such a basis, so

$$[L^*L]_{\mathcal{U}} = \left(\begin{array}{ccc|c} \lambda_1 & & & \\ & \ddots & & 0 \\ & & \lambda_r & \\ \hline & 0 & & 0 \end{array}\right)$$

We want to have $L(u_i) = \sigma_i v_i$ for $1 \leqslant i \leqslant r$.
Choose $v_i = \frac{L(u_i)}{\sigma_i}$ for $1 \leqslant i \leqslant r$.
Note that $\{v_1, \ldots, v_r\}$ is orthonormal, because

$$\langle L(u_i), L(u_j)\rangle = \langle u_i, L^*L(u_j)\rangle = \langle u_i, \lambda_j u_j\rangle = \overline{\lambda_j}\,\langle u_i, u_j\rangle = \lambda_j \delta_{ij} = \sigma_j^2 \delta_{ij}$$

Therefore $\langle v_i, v_j\rangle = \left\langle \frac{L(u_i)}{\sigma_i}, \frac{L(u_j)}{\sigma_j} \right\rangle = \delta_{ij}$.
Extend $\{v_1, \ldots, v_r\}$ to an orthonormal basis $\mathcal{V} = \{v_1, \ldots, v_r, v_{r+1}, \ldots, v_m\}$ for $\mathcal{V}$.
It follows that $[L]_{\mathcal{V}}^{\mathcal{U}} = \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right)$

Theorem 6.2.5. Let U be a finite-dimensional vector space over $\mathbb{F}$ with $L : U \to U$ linear. Then for every eigenvalue λ of L, there exists a basis of cycles corresponding to λ for K_λ.

Proof: Fix an eigenvalue λ of L.

Choose m so that $U = \text{range}(L-\lambda I)^0 \supsetneq \text{range}(L-\lambda I)^1 \supsetneq \cdots \supsetneq \text{range}(L-\lambda I)^m = \text{range}(L-\lambda I)^{m+1} = \cdots$

Previously we saw that $\text{range}(L-\lambda I)^m = \bigoplus_{\mu\neq\lambda} K_\mu$ for an eigenvalue μ of L.

We also have that $\{0\} = \text{null}(L-\lambda I)^0 \subsetneq \text{null}(L-\lambda I)^1 \subsetneq \cdots \subsetneq \text{null}(L-\lambda I)^m = \text{null}(L-\lambda I)^{m+1} = \cdots$

Note that $\text{null}(L-\lambda I) = E_\lambda$ and $\text{null}(L-\lambda I)^m = K_\lambda$.

Now follows the algorithm for finding a basis of cycles for K_λ.

Step 1. Choose a basis $\{u_1^1, \ldots, u_1^{\ell_1}\}$ for $\text{range}(L-\lambda I)^{m-1} \cap K_\lambda = \text{range}(L-\lambda I)^{m-1} \cap E_\lambda$.

Then we obtain cycles $\{u_1^1\}, \{u_1^2\}, \ldots, \{u_1^{\ell_1}\}$.

Step 2. For $1 \leqslant j \leqslant \ell_1$, choose $u_2^j \in \text{range}(L-\lambda I)^{m-2} \cap K_\lambda$ so that $(L-\lambda I)u_2^j = u_1^j$.

Also, extend $\{u_1^1, \ldots, u_1^{\ell_1}\}$ to a basis $\{u_1^1, \ldots, u_1^{\ell_2}\}$ for $\text{range}(L-\lambda I)^{m-2} \cap E_\lambda$.

We obtain the cycles $\{u_1^1, u_2^1\}, \ldots, \{u_1^{\ell_1}, u_2^{\ell_1}\}, \{u_1^{\ell_1+1}\}, \ldots, \{u_1^{\ell_2}\}$

Step k: Suppose we have constructed cycles $B^j = \{u_1^j, \ldots, u_{n_j-1}^j\}$ for $1 \leqslant j \leqslant \ell_{k-1}$ such that $\{u_1^1, \ldots, u_1^{\ell_{k-1}}\}$ is a basis for $\text{range}(L-\lambda I)^{m-(k-1)} \cap E_\lambda$ and such that $\bigcup_{j=1}^{\ell_{k-1}} B^j$ is a basis for $\text{range}(L-\lambda I)^{m-(k-1)} \cap K_\lambda$.

For $1 \leqslant j \leqslant \ell_{k-1}$, choose $u_{n_j}^j \in \text{range}(L-\lambda I)^{m-k} \cap K_\lambda$ so that $(L-\lambda I)u_{n_j}^j = u_{n_j-1}^j$.

Then let $C_j = \{u_1^j, \ldots, u_{n_j}^j\} = B^j \cup \{u_{n_j}^j\}$.

Also, extend $\{u_1^1, \ldots, u_1^{\ell_{k-1}}\}$ to a basis $\{u_1^1, \ldots, u_1^{\ell_k}\}$ for $\text{range}(L-\lambda I)^{m-k} \cap E_\lambda$.

Now it is claimed that $\bigcup_{j=1}^{\ell_k} C^j$ is a basis for $\text{range}(L-\lambda I)^{m-k} \cap K_\lambda$.

To see that $\bigcup C^j$ is linearly independent, let

$V \quad = \text{span} \bigcup C^j \subset \text{range}(L-\lambda I)^{m-k} \cap K_\lambda$

$W \quad = \text{span} \bigcup B^j = \text{range}(L-\lambda I)^{m-(k-1)} \cap K_\lambda$

$M \quad = \text{the restriction } M = (L-\lambda I) : V \to W$

Note that $\text{null}(M) = \text{range}(L-\lambda I)^{m-k} \cap E_\lambda$ and $\text{nullity}(M) = \ell_k$ by definition, so M is onto.

Therefore

$$\dim(V) = \text{rank}(M) + \text{nullity}(M) = \dim(W) + \ell_k = \left|\bigcup B^j\right| + \ell_k = \left|\bigcup C^j\right|$$

Therefore $\bigcup C^j$ is a basis for V.

Therefore $\bigcup C^j$ is linearly independent.

To see that $\bigcup C^j$ spans $\text{range}(L-\lambda I)^{m-k} \cap K_\lambda$, let

$V_2 \quad = \text{range}(L-\lambda I)^{m-k} \cap K_\lambda$

$W \quad = \text{range}(L-\lambda I)^{m-(k-1)} \cap K_\lambda$

$M_2 \quad = \text{the restriction} M_2 = (L-\lambda I) : V_2 \to W$

Now we have that M_2 is onto and $\text{null}(M_2) = \text{range}(L-\lambda I)^{m-k} \cap E_\lambda$ with $\text{nullity}(M_2) = \ell_k$.

So then

$$\dim(V_2) = \dim(W) + \ell_k = \left|\bigcup B^j\right| + \ell_k = \left|\bigcup C^j\right| = \dim(V)$$

Therefore $V_2 = V = \text{span} \bigcup C^j$.