Chapter 6A-C, 7A-D

All vector spaces are finite dimensional over a field $\mathbf{F}$ that is either $\mathbb{R}$ or $\mathbb{C}$. The vector spaces for these chapters all have an inner product, which is defined below, so we call them inner product spaces.

1 Inner product spaces

Over the real numbers the **dot product** of two vectors $u, w \in \mathbb{R}^n$ is $u \cdot w = \sum_{i=1}^n u_i w_i$. We want to have something similar, but over $\mathbb{C}$ as well. Recall that for a complex number $\alpha = a + bi$, we indicate the conjugate by $\bar{\alpha}$, so $\bar{\alpha} = a - bi$. Notice $\alpha\bar{\alpha} = a^2 + b^2$; $\alpha + \bar{\alpha} = 2a$ and $\alpha - \bar{\alpha} = 2b$.

Definition 1.1. A **inner product space** is a vector space V over a field $\mathbf{F}$ that has an inner product. An **inner product** on V is a function from $V \times V$ to $\mathbf{F}$ that is positive definite, left linear, and conjugate symmetric (see below). We write the inner product of u and v as $\langle u, v\rangle$. An inner product must satisfy the following properties.

- **positive definiteness**: $\langle u, u\rangle$ is a nonnegative real number and $\langle u, u\rangle = 0$ only when $u = 0$.
- **left linearity**: The inner product is a linear map (additive and homogenous) on the left hand side, $\langle u + w, v\rangle = \langle u, v\rangle + \langle w, v\rangle$ and $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ for $\lambda \in \mathbf{F}$.
- **conjugate symmetry**: The inner product is not symmetric, but satisfies

$$\langle u, v\rangle = \overline{\langle v, u\rangle}$$

The **norm** of a vector u in an inner product space is $\|u\| = \sqrt{\langle u, u\rangle}$.

There are some additional properties (5.6) for the inner product, and (6.9) for the norm that are easy to show. The one that requires special attention is that inner products are not right homogeneous, you have to conjugate:

$$\langle u, \lambda v\rangle = \bar{\lambda}\langle u, v\rangle$$

The following term is equivalent to the notion of perpendicular for vectors over the reals.

Definition 1.2. Two vectors $u, v \in V$ are **orthogonal** when $\langle u, v\rangle = 0$.

Theorem 1.3 ((6.12), (6.13),(6.21) Equalities for norms)**.** *Let $u, v \in V$.*

(1) (Pythagorus) If u and v are orthogonal, then $\|u\|^2 + \|v\|^2 = \|u + v\|^2$.
(2) (Parallelogram) In any case, $\|u + v\|^2 + \|u - v\|^2 = 2\big(\|u\|^2 + \|v\|^2\big)$.

Theorem 1.4 ((6.14), (6.17) Inequalities for the inner product and norms)**.** *Let $u, v \in V$.*

(1) (Cauchy Schwarz inequality) $|\langle u, v\rangle| \le \|u\|\|v\|$.
(2) (Triangle inequality) $\|u + v\| \le \|u\| + \|v\|$.

Equality holds iff u is a multiple of v, or vice-versa.

2 Orthogonal decomposition

Definition 2.1. A list of vectors $(v_1, \ldots, v_k)$ is called an **orthogonal list** when each pair of distinct vectors is orthogonal. That is, for $i \neq j$, $\langle v_i, v_j \rangle = 0$. The list is called **orthonormal** when it is also true that $\langle v_i, v_i \rangle = 1$.

An **orthonormal basis** is a basis that is also orthonormal. (Similarly for orthogonal basis.)

Proposition 2.2 ((6.24), (6.25), (6.26), (6.28), (6.30))**.** *Let V be an inner product space of dimension n. Let $\mathcal{E} = (e_1, e_2, \ldots, e_k)$ be an orthonormal list.*

(1) The vectors in $\mathcal{E}$ are linearly independent.
(2) If $k = n$ then $\mathcal{E}$ forms a basis for V.
(3) $\|a_1 e_1 + \cdots + a_k e_k\|^2 = |a_1|^2 + \cdots + |a_k|^2$
(4) For any $v \in V$, $|\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_k \rangle|^2 \leq \|v\|^2$ with equality if and only if $v = a_1 e_1 + \cdots + v_k e_k$.
(5) If $k = n$ then

$$v = \langle v, e_1 \rangle e_1 + \langle v, e_2 \rangle e_2 + \cdots + \langle v, e_n \rangle e_n \qquad \textit{and}$$
$$\langle u, v \rangle = \langle u, e_1 \rangle \overline{\langle v, e_1 \rangle} + \langle u, e_2 \rangle \overline{\langle v, e_2 \rangle} + \cdots + \langle u, e_n \rangle \overline{\langle v, e_n \rangle}$$

Lemma 2.3 ((6.13) Orthogonal decomposition)**.** *Let $u, v \in V$ with $u \neq 0$. We can write v as the sum of a multiple of u and a vector v' that is orthogonal to u.*

$$v = cu + v'$$

where

$$c = \frac{\langle v, u \rangle}{\langle u, u \rangle} \qquad \textit{and} \qquad v' = v - \frac{\langle u, v \rangle}{\langle u, u \rangle} u$$

The space spanned by u, v is the same as the space spanned by u, v'.

Theorem 2.4 ((6.32) Gram-Schmidt procedure for creating an orthonormal list)**.** *Let $v_1, \ldots, v_m$ be linearly independent in V.*

(1) There is a simple process for creating an orthogonal list $f_1, \ldots, f_m$ such that the span of $f_1, \ldots, f_k$ equals the span of $v_1, \ldots, v_k$ for $k = 1, \ldots, m$: Inductively set

$$f_k = v_k - \frac{\langle v_k, f_1 \rangle}{\langle f_1, f_1 \rangle} f_1 - \frac{\langle v_k, f_2 \rangle}{\langle f_2, f_2 \rangle} f_2 \cdots - \frac{\langle v_k, f_{k-1} \rangle}{\langle f_{k-1}, f_{k-1} \rangle} f_{k-1}$$

(2) This can be made into an orthonormal list $e_1, \ldots, e_m$ by setting

$$e_i = \frac{f_i}{\|f_i\|}$$

(3) We also have $\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(f_1, \ldots, f_k) = \operatorname{span}(e_1, \ldots, e_k)$

Corollary 2.5 ((6.35), (6.36), (6.37)). *Suppose V is a finite dimensional inner product space.*

(1) V has an orthonormal basis.
(2) Every orthonormal list can be extended to an orthonormal basis.
(3) Let $T \in \mathcal{L}(V)$ have a minimal polynomial that factors completely. There is an orthonormal basis in which T is upper triangular. (In particular this is true when the field is $\mathbb{C}$.)

The upshot of Corollary 2.5 (1): Every finite dimensional inner product space V over $\mathbf{F}$ is, after choosing an orthonormal, basis "essentially" the same as $\mathbb{R}^n$ (or $\mathbb{C}^n$) by using coordinates with respect to that basis. The mathematically precise way to say this is V is isomorphic to $\mathbf{F}^n$ with the usual inner product.

Corollary 2.6 (7D (7.58)). *Let $v_1, \ldots, v_m$ be linearly independent in $\mathbf{F}^n$ so that the vectors are all $n \times 1$ matrices. Let A be the matrix with columns $v_1, \ldots, v_m$. Apply the Gram-Schmidt process to get an orthonormal basis $e_1, \ldots, e_m$ for the space spanned by the v_i. There is an upper triangular matrix R with positive diagonal entries such that $A = QR$ with Q the matrix with columns $e_1, \ldots, e_m$.*

Exercises 2.7. Apply Gram-Schmidt

(a) Turn u and v into an orthogonal basis for the space they span.

$$u = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \qquad v = \begin{bmatrix} 3 \\ 2 \\ 8 \end{bmatrix}$$

Check the vector v' that you compute is indeed orthogonal to u. Check that u, v and u, v' span the same space. Create an orthonormal basis by dividing by the norm of each vector. Find the matrix factorization of Corollary 2.6

(b) Turn u, v, w into an orthogonal basis for the space they span (use the same process as above).

$$u = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad v = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix} \qquad w = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}$$

(c) Turn u, v into an orthogonal basis for the space they span.

$$u = \begin{bmatrix} 1+i \\ 1 \\ 1 \end{bmatrix} \qquad v = \begin{bmatrix} 1 \\ 3 \\ -3i \end{bmatrix}$$

Exercises 2.8. (6B#8, 6B #4) An inner product space to stretch your imagination. Consider the vector space $\mathcal{P}_2(\mathbb{R})$ consisting of polynomials of degree at most 2.

(a) Show that $\langle p, q \rangle = \int_0^1 pq$ defines an inner product on $\mathcal{P}_2(\mathbb{R})$.
(b) Apply Gram-Schmidt to the basis $1, x, x^2$ to get an orthogonal basis.
(c) Divide by the norm to get an orthonormal basis.

3 Orthogonal complements

Let V be an inner product space of dimension n.

Definition 3.1. For U a subset of V, the **orthogonal complement** of U is

$$U^{\perp} = \{v \in V : \langle u, v \rangle = 0 \text{ for all } u \in U\}$$

It is pretty easy to show that each of the following hold:

Proposition 3.2 ((6.48), (6.49), (6.51), (6.52))**.** *Let U be a subspace of V.*

(1) $U^{\perp}$ is a subspace of V.
(2) If $U \subseteq W$ then $U^{\perp} \supseteq W^{\perp}$.
(3) $V = U \oplus U^{\perp}$, consequently $\dim U^{\perp} = \dim V - \dim U$.
(4) $\left(U^{\perp}\right)^{\perp} = U$.

4 Orthogonal Projection

Definition 4.1. Suppose that U is a subspace of V. By Proposition 3.2, each $v \in V$ can be uniquely written as $u + w$ with $u \in U$ and $w \in U^{\perp}$. We call u the **orthogonal projection** of v on U. The operator that takes v to its orthogonal projection on U is denoted p_U.

Proposition 4.2 ((6.57), (6.61))**.** *(1) p_U is a linear map with range U and nullspace $U^{\perp}$.*
(2) $p_U(u) = u$ for $u \in U$ and $p_U^2 = p_U$.
(3) The orthogonal projection of v on U, $p_U(v)$, is the closest point in U to v. That is, for all $u \in U$,

$$\|v - p_U(v)\| \leq \|v - u\|$$

Equality holds if and only if $u = p_U(v)$.

Now we introduce a linear map from one inner product space, V, to another, W. Given some vector $w \in W$ we want to find the closest vector to w that is in the image of T. We do this by projecting onto $\operatorname{range} T$. We would also like to find the shortest vector in V that maps to $p_{\operatorname{range} T}(w)$. We do this by choosing the unique vector in $(\operatorname{null} T)^{\perp}$ that maps to $p_{\operatorname{range} T}(w)$. (How do we know there is a unique vector? We will talk about that.)

Proposition 4.3 ((6.67))**.** *Let $T \in \mathcal{L}(V)$. Let $U = (\operatorname{null} T)^{\perp}$. Then $T|_U$ is a bijection from U to $\operatorname{range} T$.*

Definition 4.4. Let $T \in \mathcal{L}(V)$. The psuedo-inverse of T is the linear map from W to V with range equal to $(\operatorname{null} T)^{\perp}$ defined by

$$T^{\dagger} = T^{-1}_{(\operatorname{null} T)^{\perp}} P_{\operatorname{range}(T)}$$

5 Adjoint map to a linear transformation

Prelude: coordinatizing inner-products

Once you have an orthonormal basis for an inner product space you can think of a vector v as a $n \times 1$ matrix. When working over the real numbers $\langle u, v\rangle$ is just $v^t u$, where t indicates the transpose. When working over the complex numbers $\langle u, v\rangle$ is just $v^* u$, where $*$ indicates the conjugate-transpose.

Definition 5.1. Let A be a $k \times n$ matrix. The **conjugate-transpose** of A is the matrix A^* such that $\left(A^*\right)_{i,j} = \overline{A_{j,i}}$. In words, take the transpose of A and then the complex conjugate of each element.

You can check that $(A^*)^* = A$, $(A+B)^* = A^* + B^*$ and that $(AB)^* = B^* A^*$ (order reversed here!). Using the last one you can check that $\left(A^*\right)^{-1} = \left(A^{-1}\right)^*$.

This comes up later (in 7D) but while we are talking about matrices...

Definition 5.2. An invertible (in particular square) matrix Q over $\mathbb{R}$ is called **orthogonal** when $Q^t = Q^{-1}$. An invertible matrix Q over $\mathbb{C}$ is called **unitary** when $Q^* = Q^{-1}$. An orthogonal matrix is also unitary.

Proposition 5.3. *(TFAE) Let Q be an $n \times n$ matrix. The following are equivalent.*

(1) Q is unitary.
(2) The columns of Q are orthonormal.
(3) The rows of Q are orthonormal.
(4) Q is length preserving: $\|Qv\| = \|v\|$ for all $v \in V$.

Exercises 5.4. Understanding unitary matrices

(a) Show that 2×2 rotation matrices and reflection matrices are orthogonal.
(b) Show that any 2×2 unitary matrix is a rotation or a reflection matrix.
(c) Find examples of unitary 2×2 matrices that are not real.

The adjoint and the conjugate-transpose

Now we consider two inner product spaces, V and W.

Definition 5.5. Let $T \in \mathcal{L}(V, W)$. The **adjoint** of T is the function $T^* : W \longrightarrow V$ such that for every $v \in V$ and $w \in W$,

$$\langle Tv, w\rangle = \langle v, T^* w\rangle$$

This is an abstract and complicated definition. How do we even now that such a function exists? Is it a linear map? What sort of properties does it have? Each of these questions has simple answers involving straightforward computations. But they also require a level of comfort with abstraction.

Let me make it simpler. Choose orthonormal bases $\mathcal{E} = e_1, \ldots, e_n$ for V and $\mathcal{F} = f_1, \ldots, f_k$ for W. Then T corresponds to a $k \times n$ matrix. Let's call it A. We can now think of $v \in V$ as a column vector with n components (with respect to the $\mathcal{E}$ basis). We can think of Tv as a column vector with k components (in the $\mathcal{F}$ basis). You get that column vector for Tv by doing the matrix multiplication Av. Then we can see below that the matrix for T^* is just A^*, the conjugate transpose of A.

$$\begin{aligned}\langle Tv, w\rangle &= w^*(Av)\\ &= (A^*w)^*v \qquad \text{(check this!)}\\ &= \langle v, A^*w\rangle\end{aligned}$$

This needs to be justified by considering arbitrary orthonormal bases; here is Axler's formal statement.

Proposition 5.6 ((7.9))**.** *Let $T \in \mathcal{L}(V, W)$. Let $\mathcal{E}$ be an orthonormal basis for V and let $\mathcal{F}$ be an orthonormal basis for W.*

$$\mathcal{M}(T^*, \mathcal{F}, \mathcal{E}) = \mathcal{M}(T, \mathcal{E}, \mathcal{F})^*$$

Proposition 5.7 ((7.4), (7.6))**.** *Let $T \in \mathcal{L}(V, W)$. Then $T \in \mathcal{L}(W, V)$ and*

(1) $\operatorname{null} T^* = \big(\operatorname{range} T\big)^\perp$
(2) $\operatorname{range} T^* = \big(\operatorname{null} T\big)^\perp$

6 Self-adjoint and normal operators

Now suppose that $T \in \mathcal{L}(V)$, so the matrix A for T is a square matrix.

Definition 6.1. $T \in \mathcal{L}(V)$ is **self-adjoint** when $T^* = T$. Let A be the matrix for T in some orthonormal basis. Then T is self-adjoint when $A = A^*$. Over the reals, we also call such a matrix **symmetric**.

Proposition 6.2 ((7.12))**.** *If T is self adjoint then all its eigenvalues are real. [A nice exercise to prove this.]*

Exercises 6.3. Becoming adroit with the adjoint

- **(a)** For any $m \times n$ matrix A, show that all the diagonal entries of AA^* and of A^*A are nonnegative real numbers.
- **(b)** Let A be a real square matrix. Show Ai is self-adjoint implies A is antisymmetric $A^t = -A$. What can you say about the diagonal.?
- **(c)** If A and B are both self-adjoint show that $A + B$ is self-adjoint. If also A and B commute show that AB is self-adjoint.

Definition 6.4. An operator on an inner product space is **normal** when it commutes with its adjoint: $TT^* = T^*T$ (or for matrices $AA^* = A^*A$).

Proposition 6.5. *Let $T \in \mathcal{L}(V)$. Then T is normal if and only if $\|Tv\| = \|T^*v\|$ for all $v \in V$.*

Proposition 6.6 ((7.21))**.** *Let $T \in \mathcal{L}(V)$ be normal. Let A be its matrix with respect to an orthonormal basis.*

(1) $\operatorname{null} T = \operatorname{null} T^*$
(2) $\operatorname{range} T = \operatorname{range} T^*$
(3) $V = \operatorname{null} T \oplus \operatorname{range} T$.

Proposition 6.7 ((7.21), (7.22))**.** *Let $T \in \mathcal{L}(V)$ be normal.*

(1) If v is an eigenvector for T with eigenvalue λ, then v is also an eigenvector for T^*, but with eigenvalue λ^*.
(2) The eigenvectors of T associated to distinct eigenvalues are orthogonal.

Exercises 6.8. Getting used to normality

(a) Show that any self-adjoint matrix is normal. (Simple)
(b) Show that these two matrices are normal.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ i & 1 \end{bmatrix}$$

(c) Suppose A is normal. Show that the following are all normal: (i) cA for $c \in \mathbb{C}$, (ii) A^*, D (iii) A^{-1}, (iv) $Q^{-1}AQ$ for Q any unitary matrix.
(d) Suppose that A is upper triangular. Show that if A is normal, then A is actually a diagonal matrix.
(e) Show that a real matrix that is anti-symmetric (that is $A^t = -A$) is normal.

7 The spectral theorems

Theorem 7.1 (7.31))**.** *Let V be a finite dimensional inner product space over $\mathbb{C}$. The following are equivalent (TFAE).*

(1) T is normal.
(2) T has a diagonal matrix with respect to an orthonormal basis.
(3) V has an orthonormal basis consisting of eigenvalues of T.

Theorem 7.2 ((7.29))**.** *Let V be a finite dimensional inner product space over $\mathbb{R}$. The following are equivalent (TFAE).*

(1) T is self adjoint.
(2) T has a diagonal matrix with respect to an orthonormal basis.
(3) V has an orthonormal basis consisting of eigenvalues of T.

Positive matrices

Definition 7.3. An operator $T \in \mathcal{L}(V)$ is called **positive** if T is self-adjoint and for all $v \in V$, $\langle Tv, v\rangle \geq 0$.

Theorem 7.4 ((7.38), (7.39))**.** *The following are equivalent for $T \in \mathcal{L}(V)$. (TFAE)*

(1) T is a positive operator.
(2) There is some orthonormal basis of V for which T is diagonal with nonnegative numbers on the diagonal.
(3) T has a positive (or self-adjoint) square root. (and it is also unique).
(4) $T = R^*R$ for some $R \in \mathcal{L}(V)$.

Exercises 7.5. Creating a normal matrix. Let $T \in \mathcal{L}(V, W)$. Use the definition of the adjoint and properties of the inner product to prove (b) and (c).

(a) Show that T^*T is self adjoint.
(b) Show that T^*T is positive.
(c) Show that $\operatorname{null} T^*T = \operatorname{null} T$
(d) Show that $\operatorname{range} T^*T = \operatorname{range} T^*$

8 Singular value decomposition

Definition 8.1. Let $T \in \mathcal{L}(V, W)$. The **singular values** of T are the nonnegative square roots of the eigenvalues of T^*T, listed in decreasing order, each included as many times as the dimension of the corresponding eigenspace of T^*T.

Proposition 8.2 ((7.68),(7.69))**.** *Let $T \in \mathcal{L}(V, W)$.*

(1) T is injective if and only if 0 is not a singular value.
(2) The number of positive singular values of T equals $\dim(\operatorname{range} T)$.
(3) T is surjective if and only if the number of positive singular values of T equals
(4) T is an isometry if and only if all singular values are 1.

Theorem 8.3 ((7.70))**.** *Let $T \in \mathcal{L}(V, W)$ and let its positive singular values be $s_1, \ldots, s_m$. There exist orthonormal lists $e_1, \ldots, e_m$ in V and $f_1, \ldots, f_m$ in W such that, for every $v \in V$*

$$Tv = s_1\langle v, e_1\rangle + \ldots s_m\langle v, e_m\rangle$$

Corollary 8.4. *Suppose A is a $k \times n$ matrix of rank $r \geq 1$. Then there exists a factorization*

$$A = BDC^*$$

where B is a $k \times r$ matrix with orthonormal columns, C is an $n \times r$ matrix with orthonormal columns and D is an $r \times r$ diagonal matrix with strictly positive numbers on the diagonal.