

Math 346 Lecture #31

12.1 Projections

Before, in Chapter 3, we used an inner product on a vector space V to define an orthogonal projection P of V onto a subspace X of V . An algebraic property of an orthogonal projection P is that $P^2 = P$ (called idempotence). This property is actually sufficient to define a projection on a vector space not necessarily equipped with an inner product.

Throughout this note, we assume that V is a vector space over \mathbb{F} .

12.1.1 Projections

Definition 12.1.1. A linear operator $P \in \mathcal{L}(V)$ is called a projection if $P^2 = P$.

Example 12.1.2. If $P \in \mathcal{L}(V)$ is a projection, then so is $I - P \in \mathcal{L}(V)$, where $I \in \mathcal{L}(V)$ is the identity operator defined by $I(v) = v$ for all $v \in V$.

You have it as HW (Exercise 12.1) to show that $I - P$ is a projection.

The linear operator $I - P$ is called the complementary projection of P .

Lemma 12.1.3. If $P \in \mathcal{L}(V)$ is a projection, then

- (i) $y \in \mathcal{R}(P)$ if and only if $Py = y$, and
- (ii) $\mathcal{N}(P) = \mathcal{R}(I - P)$.

Proof. (i) If $Py = y$, then $y \in \mathcal{R}(P)$.

On the other hand, if $y \in \mathcal{R}(P)$, then there exists $x \in V$ such that $y = Px$.

Since $P^2 = P$ we have $Py = P^2x = Px = y$.

(ii) We have $x \in \mathcal{N}(P)$ if and only if $Px = 0$.

We also have $Px = 0$ if and only if $(I - P)x = x - Px = x$.

Because $I - P$ is a projection by Example 12.1.2, we have by part (i) that $(I - P)x = x$ if and only if $x \in \mathcal{R}(I - P)$.

Thus we have that $x \in \mathcal{N}(P)$ if and only if $x \in \mathcal{R}(I - P)$. □

Remark. Because $I - P$ is a projection when P is a projection, we can apply part (ii) of Lemma 12.1.3 to $I - P$ to get $\mathcal{N}(I - P) = \mathcal{R}(P)$.

We show next that a projection P on V decomposes V into the direct sum of two complementary subspaces of V .

Theorem 12.1.4. If $P \in \mathcal{L}(V)$ is a projection, then $V = \mathcal{R}(P) \oplus \mathcal{N}(P)$.

Proof. For every $x \in V$ there holds

$$Px + (I - P)x = Px + Ix - Px = x.$$

By Lemma 12.1.3, part (ii), we have $\mathcal{R}(I - P) = \mathcal{N}(P)$, so that $(I - P)x \in \mathcal{N}(P)$.

Thus each $x \in V$ is the sum of $Px \in \mathcal{R}(P)$ and $(I - P)x \in \mathcal{N}(P)$.

This implies that $V = \mathcal{R}(P) + \mathcal{N}(P)$.

To show that $\mathcal{R}(P) \cap \mathcal{N}(P) = \{0\}$ we take $x \in \mathcal{R}(P) \cap \mathcal{N}(P)$.

Then $x \in \mathcal{R}(P)$, so by Lemma 12.1.3 part (i) we have $Px = x$.

But as $x \in \mathcal{N}(P)$, we have $Px = 0$.

Thus $0 = Px = x$, giving $V = \mathcal{R}(P) \oplus \mathcal{N}(P)$. □

Corollary 12.1.5. For $\dim(V) < \infty$, if $P \in \mathcal{L}(V)$ is a projection with $S = [s_1, \dots, s_k]$ a basis for $\mathcal{R}(P)$ and $T = [t_1, \dots, t_l]$ a basis for $\mathcal{N}(P)$, then $S \cup T$ is a basis for V (i.e., $k + l = \dim(V)$) and the block matrix representation of P in the basis $S \cup T$ is

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

where I is the $k \times k$ identity matrix, and each 0 is a zero matrix of appropriate size.

Proof. By Theorem 12.1.4 the union $S \cup T$ is a basis for V .

By Theorem 4.2.9, the block matrix representation of P in the basis $S \cup T$ is

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

where A_{11} is the $k \times k$ matrix representation of P on $\mathcal{R}(P)$, and A_{22} is the $l \times l$ matrix representation of P on $\mathcal{N}(P)$.

By Lemma 12.1.3 part (i), P is the identity map on $\mathcal{R}(P)$ so that $A_{11} = I$.

Since $Px = 0$ for each $x \in \mathcal{N}(P)$, we have $A_{22} = 0$. □

Theorem 12.1.6. For subspaces W_1 and W_2 of V (not assumed finite dimensional), if $V = W_1 \oplus W_2$, then there exists a unique projection $P \in \mathcal{L}(V)$ such that $\mathcal{R}(P) = W_1$ and $\mathcal{N}(P) = W_2$.

Proof. From $V = W_1 \oplus W_2$, there is for each $x \in V$ unique $x_1 \in W_1$ and unique $x_2 \in W_2$ such that $x = x_1 + x_2$.

Define $P : V \rightarrow V$ by $Px = x_1$.

Since $x_1 = x_1 + 0$ for $x_1 \in W_1$, then $Px_1 = x_1$.

Hence $P^2x_1 = P(Px_1) = Px_1 = x_1$, so that $P^2 = P$ on W_1 .

Since $x_2 = 0 + x_2$ for $x_2 \in W_2$, then $Px_2 = 0$.

Hence $P^2x_2 = P(Px_2) = P0 = 0$ (since $0 \in W_1$).

Thus $P^2 = P$ on $V = W_1 \oplus W_2$.

To show that P is linear, we let $x = x_1 + x_2$ and $y = y_1 + y_2$, where $x_1, y_1 \in W_1$ and $x_2, y_2 \in W_2$.

Then for $a, b \in \mathbb{C}$ we have

$$ax + by = (ax_1 + by_1) + (ax_2 + by_2)$$

where $ax_1 + by_1 \in W_1$ and $ax_2 + by_2 \in W_2$ because W_1 and W_2 are subspaces.

This gives

$$\begin{aligned} P(ax + by) &= P((ax_1 + by_1) + (ax_2 + by_2)) \\ &= ax_1 + by_1 \\ &= aP(x) + bP(y). \end{aligned}$$

Thus P is linear with $\mathcal{R}(P) = W_1$ and $\mathcal{N}(P) = W_2$.

If $Q \in \mathcal{L}(V)$ is another projection with $\mathcal{R}(Q) = W_1$ and $\mathcal{N}(Q) = W_2$, then for all $x = x_1 + x_2 \in W_1 \oplus W_2$ there holds

$$Px = Px_1 + Px_2 = x_1 = Qx_1 = Qx_1 + Qx_2 = Qx.$$

This shows that $P = Q$ on V , so that P is unique. □

Definition. The unique projection $P \in \mathcal{L}(V)$ associated to $V = W_1 \oplus W_2$ in Theorem 12.1.6 is called the projection onto W_1 along W_2 .

For a projection $P \in \mathcal{L}(V)$, we have by Theorem 12.1.4 that $V = \mathcal{R}(P) \oplus \mathcal{N}(P)$, so that with $W_1 = \mathcal{R}(P)$ and $W_2 = \mathcal{N}(P)$, the projection P is the unique projection onto $\mathcal{R}(P)$ along $\mathcal{N}(P)$.

We sometimes say that a projection P is a projection onto $\mathcal{R}(P)$ without reference to along $\mathcal{N}(P)$ because the along part is always given by $\mathcal{N}(P)$.

Keep in mind that there do exist distinct projections $P, Q \in \mathcal{L}(V)$ with $\mathcal{R}(P) = \mathcal{R}(Q)$ but $\mathcal{N}(P) \neq \mathcal{N}(Q)$. For example, the projections $P, Q \in \mathcal{L}(\mathbb{C}^2)$ defined by $P(e_1) = e_1$, $P(e_2) = 0$, $Q(e_1) = e_1$, and $Q(e_1 + e_2) = 0$ has the same range but different kernels.

Remark. In a finite dimensional inner product space V , the projection P onto W_1 along W_2 is an orthogonal projection only when $W_2 = W_1^\perp$. In an infinite dimensional inner product space, a projection P onto W_1 along W_2 is an orthogonal projection only when W_1 is a closed subspace and $W_2 = W_1^\perp$.

Example (in lieu of 12.1.7). Consider the vector space $V = C([0, 1], \mathbb{C})$ equipped with the inner product

$$\langle f, g \rangle = \int_0^1 \overline{f(t)}g(t) dt.$$

Define the operator $P : V \rightarrow V$ by $P(f)$ is the constant function from $[0, 1]$ to \mathbb{C} with value $f(0)$.

The operator P is linear because for $f, g \in V$ and $a, b \in \mathbb{C}$ there holds

$$P(af + bg) = af(0) + bg(0) = aP(f) + bP(g).$$

The operator $P \in \mathcal{L}(V)$ is a projection because for all $f \in V$ there holds

$$P^2(f) = P(f(0)) = f(0) = P(f).$$

The subspace $\mathcal{R}(P)$ consists of the constant functions from $[0, 1]$ to \mathbb{C} .

The subspace $\mathcal{N}(P)$ consists of those continuous functions $f : [0, 1] \rightarrow \mathbb{C}$ such that $f(0) = 0$.

By Theorem 12.1.4 there holds $V = \mathcal{R}(P) \oplus \mathcal{N}(P)$, i.e., each function $f \in V$ can be written uniquely as

$$f(t) = f(0) + (f(t) - f(0))$$

for $f(0) \in \mathcal{R}(P)$ and $f(t) - f(0) \in \mathcal{N}(P)$.

With $W_1 = \mathcal{R}(P)$ and $W_2 = \mathcal{N}(P)$, we have by Theorem 12.1.6 that P is the unique projection onto W_1 along W_2 .

Is P an orthogonal projection? That is, is W_1 closed and is $W_1^\perp = W_2$?

The answer is no for the second condition because there exists $f \in W_1$ and $g \in W_2$ such that $\langle f, g \rangle \neq 0$, i.e., for $f = 1$ and $g(t) = t$ we have

$$\langle f, g \rangle = \int_0^1 t \, dt = 1/2 \neq 0.$$

Note. Sometimes nonorthogonal projections, such as in the previous example, are called oblique projections.

12.1.2 Invariant Subspaces and Their Projections

Recall from Section 4.2 that a subspace W of V is invariant for $L \in \mathcal{L}(V)$ or that W is L -invariant if $L(W) \subset W$.

Theorem 12.1.8. For $L \in \mathcal{L}(V)$, a subspace W of V is L -invariant if and only if for any projection $P \in \mathcal{L}(V)$ onto W there holds

$$LP = PLP.$$

Proof. Suppose W is L -invariant.

Let $P \in \mathcal{L}(L)$ be a projection with $\mathcal{R}(P) = W$.

By the Remark after Lemma 12.1.3 the projection $I - P$ satisfies $\mathcal{N}(I - P) = \mathcal{R}(P) = W$.

For each $w \in W$ the L -invariance of W implies that $Lw \in W$, and hence for all $w \in W$ that $(I - P)Lw = 0$.

Since $Pv \in W$ for all $v \in V$, we obtain $(I - P)LPv = 0$ for all $v \in V$.

This implies that $(I - P)LP = 0$ or rewritten that $LP = PLP$.

Now suppose for a projection P onto W that $LP = PLP$.

Since $Pw = w$ for all $w \in W$ and $LP = PLP$, it follows that

$$Lw = LPw = PLPw = PLw.$$

Since $PLw \in W$, we obtain $Lw = PLw \in W$, whence W is L -invariant. □

Theorem 12.1.9. Suppose W_1, W_2 are subspaces of V for which $V = W_1 \oplus W_2$, and $L \in \mathcal{L}(V)$. Then W_1 and W_2 are both L -invariant if and only if the projection P onto W_1 along W_2 satisfies $LP = PL$.

Proof. Suppose both W_1 and W_2 are L -invariant.

Since $V = W_1 \oplus W_2$, each $v \in V$ can be written uniquely as $v = w_1 + w_2$ for $w_1 \in W_1$ and $w_2 \in W_2$.

We have $Pw_1 = w_1$ and $Pw_2 = 0$, and so $LPw_2 = 0$.

By the L -invariance of W_1 and W_2 we have $Lw_1 \in W_1$ and $Lw_2 \in W_2$.

Since P is the projection onto W_1 along W_2 there holds $PLw_1 = Lw_1$ and $PLw_2 = 0$.

Thus

$$PLv = PLw_1 + PLw_2 = PLw_1 = Lw_1 = LPw_1 = LPw_1 + LPw_2 = LPv.$$

This holds for all $v \in V$ so that $LP = PL$.

Now suppose that $LP = PL$ for the projection P onto W_1 along W_2 .

Then $\mathcal{R}(P) = W_1$ and $\mathcal{N}(P) = W_2$.

For $w_1 \in W_1$ we have

$$Lw_1 = LPw_1 = PLw_1 \in W_1.$$

This shows that W_1 is L -invariant.

Because $LP = PL$ there holds

$$L(I - P) = L - LP = L - PL = (I - P)L.$$

By Lemma 12.1.3 part (ii), we have $\mathcal{R}(I - P) = \mathcal{N}(P) = W_2$.

For $w_2 \in W_2$ we have $Pw_2 = 0$ so that

$$Lw_2 = L(I - P)w_2 = (I - P)Lw_2 \in W_2.$$

This shows that W_2 is L -invariant. □

12.1.3 Eigenprojections for Simple Operators

We apply the theory of projections to a simple operator on a finite dimensional vector space, where the range of the projections are the eigenspaces of that simple operator. This will give a decomposition of a simple operator on a finite dimensional vector space into a sum of scalar multiples of the projections, where the scalars are the corresponding eigenvalues.

Recall for $i, j = 1, \dots, n$ that δ_{ij} is the (i, j) th entry of the $n \times n$ identity matrix I .

Proposition 12.1.10. Suppose $A \in M_n(\mathbb{C})$ is a simple operator whose distinct (complex) eigenvalues are $\lambda_1, \dots, \lambda_n$. Let $S \in M_n(\mathbb{C})$ be the matrix whose columns are the corresponding right eigenvectors of A , and denote the i th column of S by r_i . Let $\ell_1^T, \dots, \ell_n^T$ be the corresponding left eigenvectors of A , i.e., the rows of S^{-1} . Define the $n \times n$ matrices $P_k = r_k \ell_k^T$, $k = 1, \dots, n$. Then

- (i) $\ell_i^T r_j = \delta_{ij}$ for all $i, j = 1, \dots, n$,
- (ii) $P_i P_j = \delta_{ij} P_i$ for all $i, j = 1, \dots, n$,

(iii) $P_i A = A P_i = \lambda_i P_i$ for all $i = 1, \dots, n$,

(iv) $\sum_{i=1}^n P_i = I$, and

(v) $A = \sum_{i=1}^n \lambda_i P_i$.

Proof. (i) Since r_1, \dots, r_n are columns of S , since $\ell_1^T, \dots, \ell_n^T$ are the rows of S^{-1} , and since $S^{-1}S = I$ we have $\ell_i^T r_j = \delta_{ij}$.

(ii) Computing we have

$$P_i P_j = r_i \ell_i^T r_j \ell_j^T = r_i \delta_{ij} \ell_j^T = \delta_{ij} r_i \ell_j^T = \begin{cases} r_i \ell_i^T & \text{if } j = i, \\ 0 & \text{if } i \neq j. \end{cases}$$

Since $r_i \ell_i^T = P_i$, we obtain $P_i P_j = \delta_{ij} P_i$.

(iii) Computing we have

$$P_i A = r_i \ell_i^T A = r_i \lambda_i \ell_i^T = \lambda_i r_i \ell_i^T = \lambda_i P_i$$

and

$$A P_i = A r_i \ell_i^T = \lambda_i r_i \ell_i^T = \lambda_i P_i$$

thus giving $P_i A = A P_i = \lambda_i P_i$.

(iv) We notice that

$$\sum_{i=1}^n P_i = \sum_{i=1}^n r_i \ell_i^T$$

is the outer product expansion of SS^{-1} obtained by partitioning S into columns and S^{-1} into rows.

Since $SS^{-1} = I$, we obtain $\sum_{i=1}^n P_i = I$.

(v) Using (iii) and (iv) we compute

$$\sum_{i=1}^n \lambda_i P_i = \sum_{i=1}^n A P_i = A \sum_{i=1}^n P_i = A I = A$$

giving the result. □

Remark. The matrices P_i are projections by part (ii) of Proposition 12.1.10 because $P_i^2 = P_i P_i = \delta_{ii} P_i = P_i$. The rank of each of these projections is one because the columns of P_i are all scalar multiples of the nonzero right eigenvector r_i . Indeed the range of P is the one-dimensional eigenspace of A corresponding to the eigenvalue λ_i .

Definition. For a simple operator $A \in M_n(\mathbb{C})$ the rank-1 projections P_1, \dots, P_n in Proposition 12.1.10 are called the eigenprojections of A .

Example (in lieu of 12.1.11). The eigenvalues and right eigenvectors of the simple

$$A = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} \in M_2(\mathbb{C})$$

are

$$\lambda_1 = 3, \mathbf{r}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \lambda_2 = -1, \mathbf{r}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

The matrix of right eigenvectors

$$S = [\mathbf{r}_1 \ \mathbf{r}_2] = \begin{bmatrix} 1 & 1 \\ 2 & -2 \end{bmatrix}$$

has inverse

$$S^{-1} = -\frac{1}{4} \begin{bmatrix} -2 & -1 \\ -2 & 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 2 & 1 \\ 2 & -1 \end{bmatrix}.$$

The rows of S^{-1} give left eigenvectors of A :

$$\ell_1^T = \frac{1}{4} [2 \ 1], \ell_2^T = \frac{1}{4} [2 \ -1].$$

The eigenprojections are

$$P_1 = \mathbf{r}_1 \ell_1^T = \frac{1}{4} \begin{bmatrix} 1 \\ 2 \end{bmatrix} [2 \ 1] = \frac{1}{4} \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}$$

and

$$P_2 = \mathbf{r}_2 \ell_2^T = \frac{1}{4} \begin{bmatrix} 1 \\ -2 \end{bmatrix} [2 \ -1] = \frac{1}{4} \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix}.$$

Each of P_1 and P_2 has rank 1, and we can verify the properties listed in Proposition 12.1.10.

For property (ii) we have

$$\begin{aligned} P_1 P_2 &= \frac{1}{16} \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0, \\ P_1^2 &= \frac{1}{16} \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} = \frac{1}{16} \begin{bmatrix} 8 & 4 \\ 16 & 8 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} = P_1, \\ P_2^2 &= \frac{1}{16} \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} = \frac{1}{16} \begin{bmatrix} 8 & -4 \\ -16 & 8 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} = P_2. \end{aligned}$$

For property (iii) we have

$$\begin{aligned} AP_1 &= \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 6 & 3 \\ 12 & 6 \end{bmatrix}, \\ P_1 A &= \frac{1}{4} \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 6 & 3 \\ 12 & 6 \end{bmatrix} = \frac{3}{4} \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} = \lambda_1 P_1 \\ AP_2 &= \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix}, \\ P_2 A &= \frac{1}{4} \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix} = -\frac{1}{4} \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} = \lambda_2 P_2. \end{aligned}$$

For property (iv) we have

$$P_1 + P_2 = \frac{1}{4} \left\{ \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} + \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} \right\} = \frac{1}{4} \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = I.$$

Finally for property (v) we have

$$\lambda_1 P_1 + \lambda_2 P_2 = \frac{3}{4} \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} - \frac{1}{4} \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} = A.$$