12.1 Projections

March 27, 2020

Throughout we assume that V is a vector space over a field $\mathbb{F}$.

Recall that $\mathscr{L}(V)$ is the vector space of linear operators on V.

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

$$P^2 = P.$$

Example 12.1.2. If $P \in \mathscr{L}(V)$ is a projection, then

$$I - P \in \mathscr{L}(V)$$

is also a projection, where $I \in \mathscr{L}(V)$ 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 (and we will see why shortly).

Second Reading Quiz Question: An inner product is required to define a projection.

False

Lemma 12.1.3. Suppose $P \in \mathscr{L}(V)$ is a projection. Then

(i) $\mathrm{y} \in \mathscr{R}(P)$ if and only if $P\mathrm{y} = \mathrm{y}$, and

(ii) $\mathscr{N}(P) = \mathscr{R}(I - P)$.

Proof. (i) If $P\mathrm{y} = \mathrm{y}$, then $y \in \mathscr{R}(P)$.
If $y \in \mathscr{R}(P)$, then there exists $\mathrm{x} \in V$ such that $\mathrm{y} = P\mathrm{x}$.
Since $P^2 = P$ we have $P\mathrm{y} = P^2\mathrm{x} = P\mathrm{x} = \mathrm{y}$.

(ii) We have $\mathrm{x} \in \mathscr{N}(P)$ if and only if $P\mathrm{x} = 0$.
We also have $P\mathrm{x} = 0$ if and only if $(I - P)\mathrm{x} = \mathrm{x} - P\mathrm{x} = \mathrm{x}$.
Because $I - P$ is a projection by Example 12.1.2, by part (i) we have $(I - P)\mathrm{x} = \mathrm{x}$ if and only if $\mathrm{x} \in \mathscr{R}(I - P)$.
Thus we have $\mathrm{x} \in \mathscr{N}(P)$ if and only if $\mathrm{x} \in \mathscr{R}(I - P)$. $\square$

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

$$\mathscr{N}(I - P) = \mathscr{R}(P).$$

Theorem 12.1.4. If $P \in \mathscr{L}(V)$ is a projection, then

$$V = \mathscr{R}(P) \oplus \mathscr{N}(P).$$

Corollary 12.1.5. For $\dim(V) < \infty$, if $P \in \mathscr{L}(V)$ is a projection with $S = [\mathrm{s}_1, \ldots, \mathrm{s}_k]$ a basis for $\mathscr{R}(P)$ and $T = [\mathrm{t}_1, \ldots, \mathrm{t}_l]$ a basis for $\mathscr{N}(P)$, 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.

What questions do you have?

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 $P \in \mathscr{L}(V)$ such that

$$\mathscr{R}(P) = W_1 \text{ and } \mathscr{N}(P) = W_2.$$

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

Note. For a projection $P \in \mathscr{L}(V)$, we have by Theorem 12.1.4 that

$$V = \mathscr{R}(P) \oplus \mathscr{N}(P),$$

so that with $W_1 = \mathscr{R}(P)$ and $W_2 = \mathscr{N}(P)$, the projection P is the unique projection onto $\mathscr{R}(P)$ along $\mathscr{N}(P)$.

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

Example. There do exist distinct projections $P, Q \in \mathscr{L}(V)$ with $\mathscr{R}(P) = \mathscr{R}(Q)$ and $\mathscr{N}(P) \neq \mathscr{N}(Q)$.

For example, the projections $P, Q \in \mathscr{L}(\mathbb{C}^2)$ defined by

$$P(e_1) = e_1, P(e_2) = 0, Q(e_1) = e_1, Q(e_1 + e_2) = 0,$$

has the same range but different kernels.

Remark. In a finite dimensional inner product space $(V, \langle \cdot, \cdot \rangle)$, 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 $(V, \langle \cdot, \cdot \rangle)$, a projection P onto W_1 along W_2 is an orthogonal projection only when

- W_1 is a closed subspace, and
- $W_2 = W_1^{\perp}$.

Note. Sometimes nonorthogonal projections are called oblique projections.

What questions do you have?

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 $P : V \to V$ 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 $P \in \mathscr{L}(V)$ is a projection because for all $f \in V$ there holds

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

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

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

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

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

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

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

Is P an orthogonal projection?

The answer is no 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

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

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

$$L(W) \subset W.$$

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

$$LP = PLP.$$

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

$$LP = PL.$$

Examples. (i) An invariant space for $L = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ is $W = \mathrm{span}(e_1)$ and $P = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$ is a projection onto W because

$$P^2 = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} = P$$

and $\mathscr{R}(P) = \mathrm{span}(e_1) = W$. We verify Theorem 12.1.8:

$$LP = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 0 & 0 \end{bmatrix}$$

$$PL = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix} \neq LP$$

but

$$PLP = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 0 & 0 \end{bmatrix} = LP.$$

(ii) Complementary invariant subspaces for $L = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ are

$$W_1 = \mathrm{span}(e_1) \text{ and } W_2 = \mathrm{span}(e_2).$$

The linear operator $P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is the projection onto W_1 along W_2 because

$$P^2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = P$$

and $\mathscr{R}(P) = \mathrm{span}(e_1) = W_1$ and $\mathscr{N}(P) = \mathrm{span}(e_2) = W_2$. We verify Theorem 12.1.9:

$$LP = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

and

$$PL = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = LP.$$

What questions do you have?

First Reading Quiz Question: What are the properties of the rank-1 eigenprojections P_i of a simple finite dimensional linear operator A?

Recall for $i, j = 1, \ldots, n$ that δ_{ij} is the $(i, j)^{\text{th}}$ 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 $\lambda_1, \ldots, \lambda_n$. 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 $\ell_1^{\mathrm{T}}, \ldots, \ell_n^{\mathrm{T}}$ be the corresponding left eigenvectors of A, i.e., the rows of S^{-1}. Define the $n \times n$ matrices $P_k = \mathrm{r}_k \ell_k^{\mathrm{T}}$, $k = 1, \ldots, n$. Then

(i) $\ell_i^{\mathrm{T}} \mathrm{r}_j = \delta_{ij}$ for all $i, j = 1, \ldots, n$,

(ii) $P_i P_j = \delta_{ij} P_i$ for all $i, j = 1, \ldots, n$,

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

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

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

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 $P_1, \ldots, P_n$ 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

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

The matrix of right eigenvectors

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

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:

$$\ell_1^{\mathrm{T}} = \frac{1}{4} \begin{bmatrix} 2 & 1 \end{bmatrix},\ \ell_2^{\mathrm{T}} = \frac{1}{4} \begin{bmatrix} 2 & -1 \end{bmatrix}.$$

The eigenprojections are

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

and

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

Each of P_1 and P_2 has rank 1, and we can verify properties (ii)-(v) listed in Proposition 12.1.10.

For property (ii) we have

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

$$P_1^2 = \frac{1}{16}\begin{bmatrix}2 & 1\\4 & 2\end{bmatrix}\begin{bmatrix}2 & 1\\4 & 2\end{bmatrix} = \frac{1}{16}\begin{bmatrix}8 & 4\\16 & 8\end{bmatrix} = \frac{1}{4}\begin{bmatrix}2 & 1\\4 & 2\end{bmatrix} = P_1,$$

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

For property (iii) we have

$$AP_1 = \frac{1}{4}\begin{bmatrix}1 & 1\\ 4 & 1\end{bmatrix}\begin{bmatrix}2 & 1\\ 4 & 2\end{bmatrix} = \frac{1}{4}\begin{bmatrix}6 & 3\\ 12 & 6\end{bmatrix},$$

$$P_1A = \frac{1}{4}\begin{bmatrix}2 & 1\\ 4 & 2\end{bmatrix}\begin{bmatrix}1 & 1\\ 4 & 1\end{bmatrix} = \frac{1}{4}\begin{bmatrix}6 & 3\\ 12 & 6\end{bmatrix} = AP_1 = \frac{3}{4}\begin{bmatrix}2 & 1\\ 4 & 2\end{bmatrix} = \lambda_1 P_1$$

$$AP_2 = \frac{1}{4}\begin{bmatrix}1 & 1\\ 4 & 1\end{bmatrix}\begin{bmatrix}2 & -1\\ -4 & 2\end{bmatrix} = \frac{1}{4}\begin{bmatrix}-2 & 1\\ 4 & -2\end{bmatrix},$$

$$P_2A = \frac{1}{4}\begin{bmatrix}2 & -1\\ -4 & 2\end{bmatrix}\begin{bmatrix}1 & 1\\ 4 & 1\end{bmatrix} = \frac{1}{4}\begin{bmatrix}-2 & 1\\ 4 & -2\end{bmatrix}$$

$$= AP_2 = -\frac{1}{4}\begin{bmatrix}2 & -1\\ -4 & 2\end{bmatrix} = \lambda_2 P_2.$$

For property (iv) we have

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

Finally for property (v), the spectral decomposition, we have

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

What questions do you have?