Chapter 2

Projection Matrices

2.1 Definition

Definition 2.1 *Let $\boldsymbol{x} \in E^n = V \oplus W$. Then $\boldsymbol{x}$ can be uniquely decomposed into*

$$\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 \ (\textit{where}\ \boldsymbol{x}_1 \in V \ \textit{and}\ \boldsymbol{x}_2 \in W).$$

The transformation that maps $\boldsymbol{x}$ into $\boldsymbol{x}_1$ is called the projection matrix (or simply projector) onto V along W and is denoted as ϕ. This is a linear transformation; that is,

$$\phi(a_1\boldsymbol{y}_1 + a_2\boldsymbol{y}_2) = a_1\phi(\boldsymbol{y}_1) + a_2\phi(\boldsymbol{y}_2) \tag{2.1}$$

for any $\boldsymbol{y}_1$, $\boldsymbol{y}_2 \in E^n$. This implies that it can be represented by a matrix. This matrix is called a projection matrix and is denoted by $\boldsymbol{P}_{V \cdot W}$. The vector transformed by $\boldsymbol{P}_{V \cdot W}$ (that is, $\boldsymbol{x}_1 = \boldsymbol{P}_{V \cdot W}\boldsymbol{x}$) is called the projection (or the projection vector) of $\boldsymbol{x}$ onto V along W.

Theorem 2.1 *The necessary and sufficient condition for a square matrix $\boldsymbol{P}$ of order n to be the projection matrix onto $V = \mathrm{Sp}(\boldsymbol{P})$ along $W = \mathrm{Ker}(\boldsymbol{P})$ is given by*

$$\boldsymbol{P}^2 = \boldsymbol{P}. \tag{2.2}$$

We need the following lemma to prove the theorem above.

Lemma 2.1 *Let $\boldsymbol{P}$ be a square matrix of order n, and assume that (2.2) holds. Then*

$$E^n = \mathrm{Sp}(\boldsymbol{P}) \oplus \mathrm{Ker}(\boldsymbol{P}) \tag{2.3}$$

H. Yanai et al., *Projection Matrices, Generalized Inverse Matrices, and Singular Value Decomposition*, Statistics for Social and Behavioral Sciences, DOI 10.1007/978-1-4419-9887-3_2,

and

$$\mathrm{Ker}(\boldsymbol{P}) = \mathrm{Sp}(\boldsymbol{I}_n - \boldsymbol{P}). \tag{2.4}$$

Proof of Lemma 2.1. (2.3): Let $\boldsymbol{x} \in \mathrm{Sp}(\boldsymbol{P})$ and $\boldsymbol{y} \in \mathrm{Ker}(\boldsymbol{P})$. From $\boldsymbol{x} = \boldsymbol{Pa}$, we have $\boldsymbol{Px} = \boldsymbol{P}^2\boldsymbol{a} = \boldsymbol{Pa} = \boldsymbol{x}$ and $\boldsymbol{Py} = \boldsymbol{0}$. Hence, from $\boldsymbol{x}+\boldsymbol{y} = \boldsymbol{0} \Rightarrow \boldsymbol{Px}+\boldsymbol{Py} = \boldsymbol{0}$, we obtain $\boldsymbol{Px} = \boldsymbol{x} = \boldsymbol{0} \Rightarrow \boldsymbol{y} = \boldsymbol{0}$. Thus, $\mathrm{Sp}(\boldsymbol{P}) \cap \mathrm{Ker}(\boldsymbol{P}) = \{\boldsymbol{0}\}$. On the other hand, from $\dim(\mathrm{Sp}(\boldsymbol{P})) + \dim(\mathrm{Ker}(\boldsymbol{P})) = \mathrm{rank}(\boldsymbol{P}) + (n - \mathrm{rank}(\boldsymbol{P})) = n$, we have $E^n = \mathrm{Sp}(\boldsymbol{P}) \oplus \mathrm{Ker}(\boldsymbol{P})$.

(2.4): We have $\boldsymbol{Px} = \boldsymbol{0} \Rightarrow \boldsymbol{x} = (\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{x} \Rightarrow \mathrm{Ker}(\boldsymbol{P}) \subset \mathrm{Sp}(\boldsymbol{I}_n - \boldsymbol{P})$ on the one hand and $\boldsymbol{P}(\boldsymbol{I}_n - \boldsymbol{P}) \Rightarrow \mathrm{Sp}(\boldsymbol{I}_n - \boldsymbol{P}) \subset \mathrm{Ker}(\boldsymbol{P})$ on the other. Thus, $\mathrm{Ker}(\boldsymbol{P}) = \mathrm{Sp}(\boldsymbol{I}_n - \boldsymbol{P})$. Q.E.D.

Note When (2.4) holds, $\boldsymbol{P}(\boldsymbol{I}_n - \boldsymbol{P}) = \boldsymbol{O} \Rightarrow \boldsymbol{P}^2 = \boldsymbol{P}$. Thus, (2.2) is the necessary and sufficient condition for (2.4).

Proof of Theorem 2.1. (Necessity) For $\forall \boldsymbol{x} \in E^n$, $\boldsymbol{y} = \boldsymbol{Px} \in V$. Noting that $\boldsymbol{y} = \boldsymbol{y} + \boldsymbol{0}$, we obtain

$$\boldsymbol{P}(\boldsymbol{Px}) = \boldsymbol{Py} = \boldsymbol{y} = \boldsymbol{Px} \Longrightarrow \boldsymbol{P}^2\boldsymbol{x} = \boldsymbol{Px} \Longrightarrow \boldsymbol{P}^2 = \boldsymbol{P}.$$

(Sufficiency) Let $V = \{\boldsymbol{y}|\boldsymbol{y} = \boldsymbol{Px}, \boldsymbol{x} \in E^n\}$ and $W = \{\boldsymbol{y}|\boldsymbol{y} = (\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{x}, \boldsymbol{x} \in E^n\}$. From Lemma 2.1, V and W are disjoint. Then, an arbitrary $\boldsymbol{x} \in E^n$ can be uniquely decomposed into $\boldsymbol{x} = \boldsymbol{Px} + (\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ (where $\boldsymbol{x}_1 \in V$ and $\boldsymbol{x}_2 \in W$). From Definition 2.1, $\boldsymbol{P}$ is the projection matrix onto $V = \mathrm{Sp}(\boldsymbol{P})$ along $W = \mathrm{Ker}(\boldsymbol{P})$. Q.E.D.

Let $E^n = V \oplus W$, and let $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$, where $\boldsymbol{x}_1 \in V$ and $\boldsymbol{x}_2 \in W$. Let $\boldsymbol{P}_{W\cdot V}$ denote the projector that transforms $\boldsymbol{x}$ into $\boldsymbol{x}_2$. Then,

$$\boldsymbol{P}_{V\cdot W}\boldsymbol{x} + \boldsymbol{P}_{W\cdot V}\boldsymbol{x} = (\boldsymbol{P}_{V\cdot W} + \boldsymbol{P}_{W\cdot V})\boldsymbol{x}. \tag{2.5}$$

Because the equation above has to hold for any $\boldsymbol{x} \in E^n$, it must hold that

$$\boldsymbol{I}_n = \boldsymbol{P}_{V\cdot W} + \boldsymbol{P}_{W\cdot V}.$$

Let a square matrix $\boldsymbol{P}$ be the projection matrix onto V along W. Then, $\boldsymbol{Q} = \boldsymbol{I}_n - \boldsymbol{P}$ satisfies $\boldsymbol{Q}^2 = (\boldsymbol{I}_n - \boldsymbol{P})^2 = \boldsymbol{I}_n - 2\boldsymbol{P} + \boldsymbol{P}^2 = \boldsymbol{I}_n - \boldsymbol{P} = \boldsymbol{Q}$, indicating that $\boldsymbol{Q}$ is the projection matrix onto W along V. We also have

$$\boldsymbol{PQ} = \boldsymbol{P}(\boldsymbol{I}_n - \boldsymbol{P}) = \boldsymbol{P} - \boldsymbol{P}^2 = \boldsymbol{O}, \tag{2.6}$$

implying that $\mathrm{Sp}(\boldsymbol{Q})$ constitutes the null space of $\boldsymbol{P}$ (i.e., $\mathrm{Sp}(\boldsymbol{Q}) = \mathrm{Ker}(\boldsymbol{P})$). Similarly, $\boldsymbol{QP} = \boldsymbol{O}$, implying that $\mathrm{Sp}(\boldsymbol{P})$ constitutes the null space of $\boldsymbol{Q}$ (i.e., $\mathrm{Sp}(\boldsymbol{P}) = \mathrm{Ker}(\boldsymbol{Q})$).

Theorem 2.2 *Let $E^n = V \oplus W$. The necessary and sufficient conditions for a square matrix $\boldsymbol{P}$ of order n to be the projection matrix onto V along W are:*

$$\text{(i)}\ \ \boldsymbol{Px} = \boldsymbol{x}\ \text{ for } \forall \boldsymbol{x} \in V, \quad \text{(ii)}\ \ \boldsymbol{Px} = \boldsymbol{0}\ \text{ for } \forall \boldsymbol{x} \in W. \tag{2.7}$$

Proof. (Sufficiency) Let $\boldsymbol{P}_{V \cdot W}$ and $\boldsymbol{P}_{W \cdot V}$ denote the projection matrices onto V along W and onto W along V, respectively. Premultiplying (2.5) by $\boldsymbol{P}$, we obtain $\boldsymbol{P}(\boldsymbol{P}_{V \cdot W}\boldsymbol{x}) = \boldsymbol{P}_{V \cdot W}\boldsymbol{x}$, where $\boldsymbol{P}\boldsymbol{P}_{W \cdot V}\boldsymbol{x} = \boldsymbol{0}$ because of (i) and (ii) above, and $\boldsymbol{P}_{V \cdot W}\boldsymbol{x} \in V$ and $\boldsymbol{P}_{W \cdot V}\boldsymbol{x} \in W$. Since $\boldsymbol{Px} = \boldsymbol{P}_{V \cdot W}\boldsymbol{x}$ holds for any $\boldsymbol{x}$, it must hold that $\boldsymbol{P} = \boldsymbol{P}_{V \cdot W}$.

(Necessity) For any $\boldsymbol{x} \in V$, we have $\boldsymbol{x} = \boldsymbol{x} + \boldsymbol{0}$. Thus, $\boldsymbol{Px} = \boldsymbol{x}$. Similarly, for any $\boldsymbol{y} \in W$, we have $\boldsymbol{y} = \boldsymbol{0} + \boldsymbol{y}$, so that $\boldsymbol{Py} = \boldsymbol{0}$. Q.E.D.

Example 2.1 In Figure 2.1, $\overrightarrow{OA}$ indicates the projection of $\boldsymbol{z}$ onto $\mathrm{Sp}(\boldsymbol{x})$ along $\mathrm{Sp}(\boldsymbol{y})$ (that is, $\overrightarrow{OA} = \boldsymbol{P}_{Sp(\boldsymbol{x}) \cdot Sp(\boldsymbol{y})}\boldsymbol{z}$), where $\boldsymbol{P}_{Sp(\boldsymbol{x}) \cdot Sp(\boldsymbol{y})}$ indicates the projection matrix onto $\mathrm{Sp}(\boldsymbol{x})$ along $\mathrm{Sp}(\boldsymbol{y})$. Clearly, $\overrightarrow{OB} = (\boldsymbol{I}_2 - \boldsymbol{P}_{Sp(\boldsymbol{y}) \cdot Sp(\boldsymbol{x})}) \times \boldsymbol{z}$.

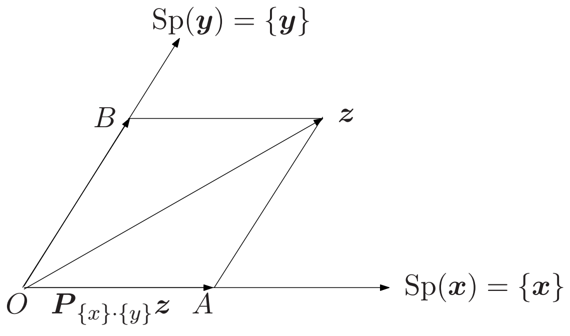

Figure 2.1: Projection onto $\mathrm{Sp}(\boldsymbol{x}) = \{\boldsymbol{x}\}$ along $\mathrm{Sp}(\boldsymbol{y}) = \{\boldsymbol{y}\}$.

Example 2.2 In Figure 2.2, $\overrightarrow{OA}$ indicates the projection of $\boldsymbol{z}$ onto $V = \{\boldsymbol{x} | \boldsymbol{x} = \alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2\}$ along $\mathrm{Sp}(\boldsymbol{y})$ (that is, $\overrightarrow{OA} = \boldsymbol{P}_{V \cdot Sp(\boldsymbol{y})}\boldsymbol{z}$), where $\boldsymbol{P}_{V \cdot Sp(\boldsymbol{y})}$ indicates the projection matrix onto V along $\mathrm{Sp}(\boldsymbol{y})$.

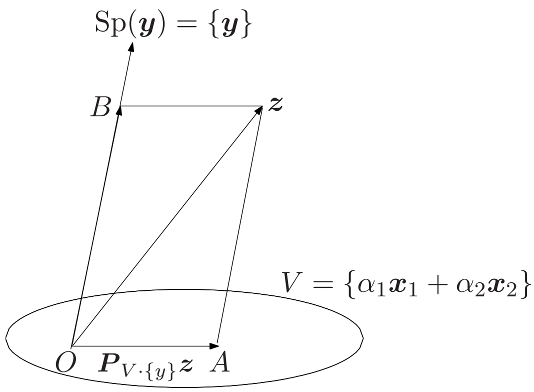

Figure 2.2: Projection onto a two-dimensional space V along $\text{Sp}(\boldsymbol{y}) = \{\boldsymbol{y}\}$.

Theorem 2.3 *The necessary and sufficient condition for a square matrix $\boldsymbol{P}$ of order n to be a projector onto V of dimensionality r ($\dim(V) = r$) is given by*

$$\boldsymbol{P} = \boldsymbol{T}\boldsymbol{\Delta}_r\boldsymbol{T}^{-1}, \tag{2.8}$$

where $\boldsymbol{T}$ is a square nonsingular matrix of order n and

$$\boldsymbol{\Delta}_r = \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

(There are r unities on the leading diagonals, $1 \le r \le n$.)

Proof. (Necessity) Let $E^n = V \oplus W$, and let $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r]$ and $\boldsymbol{B} = [\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots \boldsymbol{b}_{n-r}]$ be matrices of linearly independent basis vectors spanning V and W, respectively. Let $\boldsymbol{T} = [\boldsymbol{A}, \boldsymbol{B}]$. Then $\boldsymbol{T}$ is nonsingular, since $\text{rank}(\boldsymbol{A}) + \text{rank}(\boldsymbol{B}) = \text{rank}(\boldsymbol{T})$. Hence, $\forall \boldsymbol{x} \in V$ and $\forall \boldsymbol{y} \in W$ can be expressed as

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{\alpha} = [\boldsymbol{A}, \boldsymbol{B}]\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{0} \end{pmatrix} = \boldsymbol{T}\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{0} \end{pmatrix},$$

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{\alpha} = [\boldsymbol{A}, \boldsymbol{B}]\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{\beta} \end{pmatrix} = \boldsymbol{T}\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{\beta} \end{pmatrix}.$$

Thus, we obtain

$$\boldsymbol{P}\boldsymbol{x} = \boldsymbol{x} \Longrightarrow \boldsymbol{P}\boldsymbol{T}\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{0} \end{pmatrix} = \boldsymbol{T}\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{0} \end{pmatrix} = \boldsymbol{T}\boldsymbol{\Delta}_r\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{0} \end{pmatrix},$$

$$\boldsymbol{P}\boldsymbol{y} = \boldsymbol{0} \Longrightarrow \boldsymbol{P}\boldsymbol{T}\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix} = \boldsymbol{T}\boldsymbol{\Delta}_r\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{\beta} \end{pmatrix}.$$

Adding the two equations above, we obtain

$$\boldsymbol{P}\boldsymbol{T}\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} = \boldsymbol{T}\boldsymbol{\Delta}_r\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix}.$$

Since $\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix}$ is an arbitrary vector in the n-dimensional space E^n, it follows that

$$\boldsymbol{P}\boldsymbol{T} = \boldsymbol{T}\boldsymbol{\Delta}_r \Longrightarrow \boldsymbol{P} = \boldsymbol{T}\boldsymbol{\Delta}_r\boldsymbol{T}^{-1}.$$

Furthermore, $\boldsymbol{T}$ can be an arbitrary nonsingular matrix since $V = \mathrm{Sp}(\boldsymbol{A})$ and $W = \mathrm{Sp}(\boldsymbol{B})$ such that $E^n = V \oplus W$ can be chosen arbitrarily.

(Sufficiency) $\boldsymbol{P}$ is a projection matrix, since $\boldsymbol{P}^2 = \boldsymbol{P}$, and $\mathrm{rank}(\boldsymbol{P}) = r$ from Theorem 2.1. (Theorem 2.2 can also be used to prove the theorem above.) Q.E.D.

Lemma 2.2 Let $\boldsymbol{P}$ be a projection matrix. Then,

$$\mathrm{rank}(\boldsymbol{P}) = \mathrm{tr}(\boldsymbol{P}). \tag{2.9}$$

Proof. $\mathrm{rank}(\boldsymbol{P}) = \mathrm{rank}(\boldsymbol{T}\boldsymbol{\Delta}_r\boldsymbol{T}^{-1}) = \mathrm{rank}(\boldsymbol{\Delta}_r) = \mathrm{tr}(\boldsymbol{T}\boldsymbol{\Delta}\boldsymbol{T}^{-1}) = \mathrm{tr}(\boldsymbol{P})$. Q.E.D.

The following theorem holds.

Theorem 2.4 *Let $\boldsymbol{P}$ be a square matrix of order n. Then the following three statements are equivalent.*

$$\boldsymbol{P}^2 = \boldsymbol{P}, \tag{2.10}$$

$$\mathrm{rank}(\boldsymbol{P}) + \mathrm{rank}(\boldsymbol{I}_n - \boldsymbol{P}) = n, \tag{2.11}$$

$$E^n = \mathrm{Sp}(\boldsymbol{P}) \oplus \mathrm{Sp}(\boldsymbol{I}_n - \boldsymbol{P}). \tag{2.12}$$

Proof. (2.10) $\rightarrow$ (2.11): It is clear from $\mathrm{rank}(\boldsymbol{P}) = \mathrm{tr}(\boldsymbol{P})$.

(2.11) → (2.12): Let $V = \mathrm{Sp}(\boldsymbol{P})$ and $W = \mathrm{Sp}(\boldsymbol{I}_n - \boldsymbol{P})$. Then, $\dim(V + W) = \dim(V) + \dim(W) - \dim(V \cap W)$. Since $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{x} + (\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{x}$ for an arbitrary n-component vector $\boldsymbol{x}$, we have $E^n = V + W$. Hence, $\dim(V \cap W) = 0 \Longrightarrow V \cap W = \{\boldsymbol{0}\}$, establishing (2.12).

(2.12) → (2.10): Postmultiplying $\boldsymbol{I}_n = \boldsymbol{P} + (\boldsymbol{I}_n - \boldsymbol{P})$ by $\boldsymbol{P}$, we obtain $\boldsymbol{P} = \boldsymbol{P}^2 + (\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{P}$, which implies $\boldsymbol{P}(\boldsymbol{I}_n - \boldsymbol{P}) = (\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{P}$. On the other hand, we have $\boldsymbol{P}(\boldsymbol{I}_n - \boldsymbol{P}) = \boldsymbol{O}$ and $(\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{P} = \boldsymbol{O}$ because $\mathrm{Sp}(\boldsymbol{P}(\boldsymbol{I}_n - \boldsymbol{P})) \subset \mathrm{Sp}(\boldsymbol{P})$ and $\mathrm{Sp}((\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{P}) \subset \mathrm{Sp}(\boldsymbol{I}_n - \boldsymbol{P})$. Q.E.D.

Corollary

$$\boldsymbol{P}^2 = \boldsymbol{P} \Longleftrightarrow \mathrm{Ker}(\boldsymbol{P}) = \mathrm{Sp}(\boldsymbol{I}_n - \boldsymbol{P}). \tag{2.13}$$

Proof. (⇒): It is clear from Lemma 2.1.

(⇐): $\mathrm{Ker}(\boldsymbol{P}) = \mathrm{Sp}(\boldsymbol{I}_n - \boldsymbol{P}) \Leftrightarrow \boldsymbol{P}(\boldsymbol{I}_n - \boldsymbol{P}) = \boldsymbol{O} \Rightarrow \boldsymbol{P}^2 = \boldsymbol{P}$. Q.E.D.

2.2 Orthogonal Projection Matrices

Suppose we specify a subspace V in E^n. There are in general infinitely many ways to choose its complement subspace $V^c = W$. We will discuss some of them in Chapter 4. In this section, we consider the case in which V and W are orthogonal, that is, $W = V^{\perp}$.

Let $\boldsymbol{x}, \boldsymbol{y} \in E^n$, and let $\boldsymbol{x}$ and $\boldsymbol{y}$ be decomposed as $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ and $\boldsymbol{y} = \boldsymbol{y}_1 + \boldsymbol{y}_2$, where $\boldsymbol{x}_1, \boldsymbol{y}_1 \in V$ and $\boldsymbol{x}_2, \boldsymbol{y}_2 \in W$. Let $\boldsymbol{P}$ denote the projection matrix onto V along $V^{\perp}$. Then, $\boldsymbol{x}_1 = \boldsymbol{P}\boldsymbol{x}$ and $\boldsymbol{y}_1 = \boldsymbol{P}\boldsymbol{y}$. Since $(\boldsymbol{x}_2, \boldsymbol{P}\boldsymbol{y}) = (\boldsymbol{y}_2, \boldsymbol{P}\boldsymbol{x}) = 0$, it must hold that

$$\begin{aligned}(\boldsymbol{x}, \boldsymbol{P}\boldsymbol{y}) &= (\boldsymbol{P}\boldsymbol{x} + \boldsymbol{x}_2, \boldsymbol{P}\boldsymbol{y}) = (\boldsymbol{P}\boldsymbol{x}, \boldsymbol{P}\boldsymbol{y}) \\ &= (\boldsymbol{P}\boldsymbol{x}, \boldsymbol{P}\boldsymbol{y} + \boldsymbol{y}_2) = (\boldsymbol{P}\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{P}'\boldsymbol{y})\end{aligned}$$

for any $\boldsymbol{x}$ and $\boldsymbol{y}$, implying

$$\boldsymbol{P}' = \boldsymbol{P}. \tag{2.14}$$

Theorem 2.5 *The necessary and sufficient condition for a square matrix $\boldsymbol{P}$ of order n to be an orthogonal projection matrix (an orthogonal projector) is given by*

$$\text{(i) } \boldsymbol{P}^2 = \boldsymbol{P} \text{ and (ii) } \boldsymbol{P}' = \boldsymbol{P}.$$

Proof. (Necessity) That $\boldsymbol{P}^2 = \boldsymbol{P}$ is clear from the definition of a projection matrix. That $\boldsymbol{P}' = \boldsymbol{P}$ is as shown above.

(Sufficiency) Let $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{\alpha} \in \mathrm{Sp}(\boldsymbol{P})$. Then, $\boldsymbol{P}\boldsymbol{x} = \boldsymbol{P}^2\boldsymbol{\alpha} = \boldsymbol{P}\boldsymbol{\alpha} = \boldsymbol{x}$. Let $\boldsymbol{y} \in \mathrm{Sp}(\boldsymbol{P})^{\perp}$. Then, $\boldsymbol{P}\boldsymbol{y} = \boldsymbol{0}$ since $(\boldsymbol{P}\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}'\boldsymbol{P}'\boldsymbol{y} = \boldsymbol{x}'\boldsymbol{P}\boldsymbol{y} = 0$ must hold for an arbitrary $\boldsymbol{x}$. From Theorem 2.2, $\boldsymbol{P}$ is the projection matrix onto $\mathrm{Sp}(\boldsymbol{P})$ along $\mathrm{Sp}(\boldsymbol{P})^{\perp}$; that is, the orthogonal projection matrix onto $\mathrm{Sp}(\boldsymbol{P})$. Q.E.D.

Definition 2.2 *A projection matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^2 = \boldsymbol{P}$ and $\boldsymbol{P}' = \boldsymbol{P}$ is called an orthogonal projection matrix (projector). Furthermore, the vector $\boldsymbol{Px}$ is called the orthogonal projection of $\boldsymbol{x}$. The orthogonal projector $\boldsymbol{P}$ is in fact the projection matrix onto* $\mathrm{Sp}(\boldsymbol{P})$ *along* $\mathrm{Sp}(\boldsymbol{P})^{\perp}$, *but it is usually referred to as the orthogonal projector onto* $\mathrm{Sp}(\boldsymbol{P})$. See Figure 2.3.

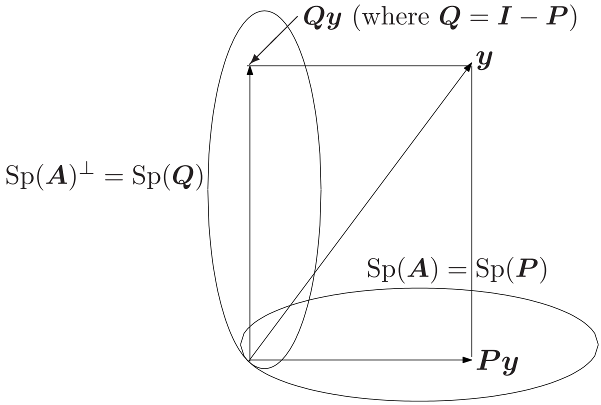

Figure 2.3: Orthogonal projection.

Note A projection matrix that does not satisfy $\boldsymbol{P}' = \boldsymbol{P}$ is called an oblique projector as opposed to an orthogonal projector.

Theorem 2.6 *Let $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m]$, where $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ are linearly independent. Then the orthogonal projector onto $V =$* $\mathrm{Sp}(\boldsymbol{A})$ *spanned by $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ is given by*

$$\boldsymbol{P} = \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'. \tag{2.15}$$

Proof. Let $\boldsymbol{x}_1 \in \mathrm{Sp}(\boldsymbol{A})$. From $\boldsymbol{x}_1 = \boldsymbol{A\alpha}$, we obtain $\boldsymbol{Px}_1 = \boldsymbol{x}_1 = \boldsymbol{A\alpha} = \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{x}_1$. On the other hand, let $\boldsymbol{x}_2 \in \mathrm{Sp}(\boldsymbol{A})^{\perp}$. Then, $\boldsymbol{A}'\boldsymbol{x}_2 = \boldsymbol{0} \Longrightarrow \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{x}_2 = \boldsymbol{0}$. Let $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$. From $\boldsymbol{Px}_2 = \boldsymbol{0}$, we obtain $\boldsymbol{Px} = \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{x}$, and (2.15) follows because $\boldsymbol{x}$ is arbitrary.

Let $\boldsymbol{Q} = \boldsymbol{I}_n - \boldsymbol{P}$. Then $\boldsymbol{Q}$ is the orthogonal projector onto $\text{Sp}(\boldsymbol{A})^\perp$, the ortho-complement subspace of $\text{Sp}(\boldsymbol{A})$.

Example 2.3 Let $\mathbf{1}_n = (1, 1, \cdots, 1)'$ (the vector with n ones). Let $\boldsymbol{P}_M$ denote the orthogonal projector onto $V_M = \text{Sp}(\mathbf{1}_n)$. Then,

$$\boldsymbol{P}_M = \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n' = \begin{bmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \ddots & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix}. \tag{2.16}$$

The orthogonal projector onto $V_M^\perp = \text{Sp}(\mathbf{1}_n)^\perp$, the ortho-complement subspace of $\text{Sp}(\mathbf{1}_n)$, is given by

$$\boldsymbol{I}_n - \boldsymbol{P}_M = \begin{bmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & 1-\frac{1}{n} \end{bmatrix}. \tag{2.17}$$

Let

$$\boldsymbol{Q}_M = \boldsymbol{I}_n - \boldsymbol{P}_M. \tag{2.18}$$

Clearly, $\boldsymbol{P}_M$ and $\boldsymbol{Q}_M$ are both symmetric, and the following relation holds:

$$\boldsymbol{P}_M^2 = \boldsymbol{P}_M,\ \ \boldsymbol{Q}_M^2 = \boldsymbol{Q}_M, \text{ and } \boldsymbol{P}_M\boldsymbol{Q}_M = \boldsymbol{Q}_M\boldsymbol{P}_M = \boldsymbol{O}. \tag{2.19}$$

Note The matrix $\boldsymbol{Q}_M$ in (2.18) is sometimes written as $\boldsymbol{P}_M^\perp$.

Example 2.4 Let

$$\boldsymbol{x}_R = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},\ \ \boldsymbol{x} = \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{pmatrix}, \text{ where } \bar{x} = \frac{1}{n}\sum_{j=1}^n x_j.$$

Then,

$$\boldsymbol{x} = \boldsymbol{Q}_M\boldsymbol{x}_R, \tag{2.20}$$

and so

$$\sum_{j=1}^n (x_j - \bar{x})^2 = ||\boldsymbol{x}||^2 = \boldsymbol{x}'\boldsymbol{x} = \boldsymbol{x}_R'\boldsymbol{Q}_M\boldsymbol{x}_R.$$

The proof is omitted.

2.3 Subspaces and Projection Matrices

In this section, we consider the relationships between subspaces and projectors when the n-dimensional space E^n is decomposed into the sum of several subspaces.

2.3.1 Decomposition into a direct-sum of disjoint subspaces

Lemma 2.3 *When there exist two distinct ways of decomposing E^n,*

$$E^n = V_1 \oplus W_1 = V_2 \oplus W_2, \tag{2.21}$$

and if $V_1 \subset W_2$ or $V_2 \subset W_1$, the following relation holds:

$$E^n = (V_1 \oplus V_2) \oplus (W_1 \cap W_2). \tag{2.22}$$

Proof. When $V_1 \subset W_2$, Theorem 1.5 leads to the following relation:

$$V_1 + (W_1 \cap W_2) = (V_1 + W_1) \cap W_2 = E^n \cap W_2 = W_2.$$

Also from $V_1 \cap (W_1 \cap W_2) = (V_1 \cap W_1) \cap W_2 = \{\mathbf{0}\}$, we have $W_2 = V_1 \oplus (W_1 \cap W_2)$. Hence the following relation holds:

$$E^n = V_2 \oplus W_2 = V_2 \oplus V_1 \oplus (W_1 \cap W_2) = (V_1 \oplus V_2) \oplus (W_1 \cap W_2).$$

When $V_2 \subset W_2$, the same result follows by using $W_1 = V_2 \oplus (W_1 \cap W_2)$. Q.E.D.

Corollary *When $V_1 \subset V_2$ or $W_2 \subset W_1$,*

$$E^n = (V_1 \oplus W_2) \oplus (V_2 \cap W_1). \tag{2.23}$$

Proof. In the proof of Lemma 2.3, exchange the roles of W_2 and V_2. Q.E.D.

Theorem 2.7 *Let $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ denote the projection matrices onto V_1 along W_1 and onto V_2 along W_2, respectively. Then the following three statements are equivalent:*

(i) $\boldsymbol{P}_1 + \boldsymbol{P}_2$ *is the projector onto $V_1 \oplus V_2$ along $W_1 \cap W_2$.*

(ii) $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{O}$.

(iii) $V_1 \subset W_2$ and $V_2 \subset W_1$. *(In this case, V_1 and V_2 are disjoint spaces.)*

Proof. (i) $\rightarrow$ (ii): From $(\boldsymbol{P}_1+\boldsymbol{P}_2)^2 = \boldsymbol{P}_1+\boldsymbol{P}_2$, $\boldsymbol{P}_1^2 = \boldsymbol{P}_1$, and $\boldsymbol{P}_2^2 = \boldsymbol{P}_2$, we have $\boldsymbol{P}_1\boldsymbol{P}_2 = -\boldsymbol{P}_2\boldsymbol{P}_1$. Pre- and postmutiplying both sides by $\boldsymbol{P}_1$, we obtain $\boldsymbol{P}_1\boldsymbol{P}_2 = -\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_1$ and $\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_1 = -\boldsymbol{P}_2\boldsymbol{P}_1$, respectively, which imply $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1$. This and $\boldsymbol{P}_1\boldsymbol{P}_2 = -\boldsymbol{P}_2\boldsymbol{P}_1$ lead to $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{O}$.

(ii) → (iii): For an arbitrary vector $\boldsymbol{x} \in V_1$, $\boldsymbol{P}_1\boldsymbol{x} = \boldsymbol{x}$ because $\boldsymbol{P}_1\boldsymbol{x} \in V_1$. Hence, $\boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{x} = \boldsymbol{P}_2\boldsymbol{x} = \boldsymbol{0}$, which implies $\boldsymbol{x} \in W_2$, and so $V_1 \subset W_2$. On the other hand, when $\boldsymbol{x} \in V_2$, it follows that $\boldsymbol{P}_2\boldsymbol{x} \in V_2$, and so $\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{x} = \boldsymbol{P}_1\boldsymbol{x} = \boldsymbol{0}$, implying $\boldsymbol{x} \in W_2$. We thus have $V_2 \subset W_2$.

(iii) → (ii): For $\boldsymbol{x} \in E^n$, $\boldsymbol{P}_1\boldsymbol{x} \in V_1$, which implies $(\boldsymbol{I}_n - \boldsymbol{P}_2)\boldsymbol{P}_1\boldsymbol{x} = \boldsymbol{P}_1\boldsymbol{x}$, which holds for any $\boldsymbol{x}$. Thus, $(\boldsymbol{I}_n - \boldsymbol{P}_2)\boldsymbol{P}_1 = \boldsymbol{P}_1$, implying $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{O}$. We also have $\boldsymbol{x} \in E^n \Rightarrow \boldsymbol{P}_2\boldsymbol{x} \in V_2 \Rightarrow (\boldsymbol{I}_n - \boldsymbol{P}_1)\boldsymbol{P}_2\boldsymbol{x} = \boldsymbol{P}_2\boldsymbol{x}$, which again holds for any $\boldsymbol{x}$, which implies $(\boldsymbol{I}_n - \boldsymbol{P}_1)\boldsymbol{P}_2 = \boldsymbol{P}_2 \Rightarrow \boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{O}$. Similarly, $\boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{O}$.

(ii) → (i): An arbitrary vector $\boldsymbol{x} \in (V_1 \oplus V_2)$ can be decomposed into $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$, where $\boldsymbol{x}_1 \in V_1$ and $\boldsymbol{x}_2 \in V_2$. From $\boldsymbol{P}_1\boldsymbol{x}_2 = \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{x} = \boldsymbol{0}$ and $\boldsymbol{P}_2\boldsymbol{x}_1 = \boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{x} = \boldsymbol{0}$, we have $(\boldsymbol{P}_1 + \boldsymbol{P}_2)\boldsymbol{x} = (\boldsymbol{P}_1 + \boldsymbol{P}_2)(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \boldsymbol{P}_1\boldsymbol{x}_1 + \boldsymbol{P}_2\boldsymbol{x}_2 = \boldsymbol{x}_1 + \boldsymbol{x}_2 = \boldsymbol{x}$. On the other hand, by noting that $\boldsymbol{P}_1 = \boldsymbol{P}_1(\boldsymbol{I}_n - \boldsymbol{P}_2)$ and $\boldsymbol{P}_2 = \boldsymbol{P}_2(\boldsymbol{I}_n - \boldsymbol{P}_1)$ for any $\boldsymbol{x} \in (W_1 \cap W_2)$, we have $(\boldsymbol{P}_1 + \boldsymbol{P}_2)\boldsymbol{x} = \boldsymbol{P}_1(\boldsymbol{I}_n - \boldsymbol{P}_2)\boldsymbol{x} + \boldsymbol{P}_2(\boldsymbol{I}_n - \boldsymbol{P}_1)\boldsymbol{x} = \boldsymbol{0}$. Since $V_1 \subset W_2$ and $V_2 \subset W_1$, the decomposition on the right-hand side of (2.22) holds. Hence, we know $\boldsymbol{P}_1 + \boldsymbol{P}_2$ is the projector onto $V_1 \oplus V_2$ along $W_1 \cap W_2$ by Theorem 2.2. Q.E.D.

Note In the theorem above, $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{O}$ in (ii) does not imply $\boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{O}$. $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{O}$ corresponds with $V_2 \subset W_1$, and $\boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{O}$ with $V_1 \subset W_2$ in (iii). It should be clear that $V_1 \subset W_2 \Longleftrightarrow V_2 \subset W_1$ does not hold.

Theorem 2.8 *Given the decompositions of E^n in (2.21), the following three statements are equivalent:*

(i) $\boldsymbol{P}_2 - \boldsymbol{P}_1$ *is the projector onto* $V_2 \cap W_1$ *along* $V_1 \oplus W_2$.

(ii) $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{P}_1$.

(iii) $V_1 \subset V_2$ and $W_2 \subset W_1$.

Proof. (i) → (ii): $(\boldsymbol{P}_2 - \boldsymbol{P}_1)^2 = \boldsymbol{P}_2 - \boldsymbol{P}_1$ implies $2\boldsymbol{P}_1 = \boldsymbol{P}_1\boldsymbol{P}_2 + \boldsymbol{P}_2\boldsymbol{P}_1$. Pre- and postmultiplying both sides by $\boldsymbol{P}_2$, we obtain $\boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{P}_2$ and $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{P}_2$, respectively, which imply $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{P}_1$.

(ii) → (iii): For $\forall\boldsymbol{x} \in E^n$, $\boldsymbol{P}_1\boldsymbol{x} \in V_1$, which implies $\boldsymbol{P}_1\boldsymbol{x} = \boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{x} \in V_2$, which in turn implies $V_1 \subset V_2$. Let $\boldsymbol{Q}_j = \boldsymbol{I}_n - \boldsymbol{P}_j$ $(j = 1, 2)$. Then, $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_1$ implies $\boldsymbol{Q}_1\boldsymbol{Q}_2 = \boldsymbol{Q}_2$, and so $\boldsymbol{Q}_2\boldsymbol{x} \in W_2$, which implies $\boldsymbol{Q}_2\boldsymbol{x} = \boldsymbol{Q}_1\boldsymbol{Q}_2\boldsymbol{x} \in W_1$, which in turn implies $W_2 \subset W_1$.

(iii) → (ii): From $V_1 \subset V_2$, for $\forall \boldsymbol{x} \in E^n$, $\boldsymbol{P}_1\boldsymbol{x} \in V_1 \subset V_2 \Rightarrow \boldsymbol{P}_2(\boldsymbol{P}_1\boldsymbol{x}) = \boldsymbol{P}_1\boldsymbol{x} \Rightarrow \boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{P}_1$. On the other hand, from $W_2 \subset W_1$, $\boldsymbol{Q}_2\boldsymbol{x} \in W_2 \subset W_1$ for $\forall \boldsymbol{x} \in E^n \Rightarrow \boldsymbol{Q}_1\boldsymbol{Q}_2\boldsymbol{x} = \boldsymbol{Q}_2\boldsymbol{x} \Rightarrow \boldsymbol{Q}_1\boldsymbol{Q}_2\boldsymbol{Q}_2 \Rightarrow (\boldsymbol{I}_n - \boldsymbol{P}_1)(\boldsymbol{I}_n - \boldsymbol{P}_2) = (\boldsymbol{I}_n - \boldsymbol{P}_2) \Rightarrow \boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_1$.

(ii) → (i): For $\boldsymbol{x} \in (V_2 \cap W_1)$, it holds that $(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{x} = \boldsymbol{Q}_1\boldsymbol{P}_2\boldsymbol{x} = \boldsymbol{Q}_1\boldsymbol{x} = \boldsymbol{x}$. On the other hand, let $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$, where $\boldsymbol{y} \in V_1$ and $\boldsymbol{z} \in W_2$. Then, $(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{x} = (\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y} + (\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{z} = \boldsymbol{P}_2\boldsymbol{Q}_1\boldsymbol{y} + \boldsymbol{Q}_1\boldsymbol{P}_2\boldsymbol{z} = \boldsymbol{0}$. Hence, $\boldsymbol{P}_2 - \boldsymbol{P}_1$ is the projector onto $V_2 \cap W_1$ along $V_1 \oplus W_2$. Q.E.D.

Note As in Theorem 2.7, $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_1$ does not necessarily imply $\boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{P}_1$. Note that $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_1 \Longleftrightarrow W_2 \subset W_1$, and $\boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{P}_1 \Longleftrightarrow V_1 \subset V_2$.

Theorem 2.9 *When the decompositions in (2.21) and (2.22) hold, and if*

$$\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1, \tag{2.24}$$

then $\boldsymbol{P}_1\boldsymbol{P}_2$ *(or* $\boldsymbol{P}_2\boldsymbol{P}_1$*) is the projector onto* $V_1 \cap V_2$ *along* $W_1 + W_2$.

Proof. $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1$ implies $(\boldsymbol{P}_1\boldsymbol{P}_2)^2 = \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_1^2\boldsymbol{P}_2^2 = \boldsymbol{P}_1\boldsymbol{P}_2$, indicating that $\boldsymbol{P}_1\boldsymbol{P}_2$ is a projection matrix. On the other hand, let $\boldsymbol{x} \in V_1 \cap V_2$. Then, $\boldsymbol{P}_1(\boldsymbol{P}_2\boldsymbol{x}) = \boldsymbol{P}_1\boldsymbol{x} = \boldsymbol{x}$. Furthermore, let $\boldsymbol{x} \in W_1 + W_2$ and $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$, where $\boldsymbol{x}_1 \in W_1$ and $\boldsymbol{x}_2 \in W_2$. Then, $\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{x} = \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{x}_1 + \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{x}_2 = \boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{x}_1 + \boldsymbol{0} = \boldsymbol{0}$. Since $E^n = (V_1 \cap V_2) \oplus (W_1 \oplus W_2)$ by the corollary to Lemma 2.3, we know that $\boldsymbol{P}_1\boldsymbol{P}_2$ is the projector onto $V_1 \cap V_2$ along $W_1 \oplus W_2$. Q.E.D.

Note Using the theorem above, (ii) → (i) in Theorem 2.7 can also be proved as follows: From $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{O}$

$$\boldsymbol{Q}_1\boldsymbol{Q}_2 = (\boldsymbol{I}_n - \boldsymbol{P}_1)(\boldsymbol{I}_n - \boldsymbol{P}_2) = \boldsymbol{I}_n - \boldsymbol{P}_1 - \boldsymbol{P}_2 = \boldsymbol{Q}_2\boldsymbol{Q}_1.$$

Hence, $\boldsymbol{Q}_1\boldsymbol{Q}_2$ is the projector onto $W_1 \cap W_2$ along $V_1 \oplus V_2$, and $\boldsymbol{P}_1 + \boldsymbol{P}_2 = \boldsymbol{I}_n - \boldsymbol{Q}_1\boldsymbol{Q}_2$ is the projector onto $V_1 \oplus V_2$ along $W_1 \cap W_2$.

If we take $W_1 = V_1^\perp$ and $W_2 = V_2^\perp$ in the theorem above, $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ become orthogonal projectors.

Theorem 2.10 *Let* $\boldsymbol{P}_1$ *and* $\boldsymbol{P}_2$ *be the orthogonal projectors onto* V_1 *and* V_2*, respectively. Then the following three statements are equivalent:*

(i) $\boldsymbol{P}_1 + \boldsymbol{P}_2$ *is the orthogonal projector onto* $V_1 \dot{\oplus} V_2$.

(ii) $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{O}$.

(iii) V_1 *and* V_2 *are orthogonal.*

Theorem 2.11 *The following three statements are equivalent:*

(i) $\boldsymbol{P}_2 - \boldsymbol{P}_1$ *is the orthogonal projector onto* $V_2 \cap V_1^{\perp}$.

(ii) $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{P}_1$.

(iii) $V_1 \subset V_2$.

The two theorems above can be proved by setting $W_1 = V_1^{\perp}$ and $W_2 = V_2^{\perp}$ in Theorems 2.7 and 2.8.

Theorem 2.12 *The necessary and sufficient condition for* $\boldsymbol{P}_1\boldsymbol{P}_2$ *to be the orthogonal projector onto* $V_1 \cap V_2$ *is (2.24).*

Proof. Sufficiency is clear from Theorem 2.9. Necessity follows from $\boldsymbol{P}_1\boldsymbol{P}_2 = (\boldsymbol{P}_1\boldsymbol{P}_2)'$, which implies $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1$ since $\boldsymbol{P}_1\boldsymbol{P}_2$ is an orthogonal projector. Q.E.D.

We next present a theorem concerning projection matrices when E^n is expressed as a direct-sum of m subspaces, namely

$$E^n = V_1 \oplus V_2 \oplus \cdots \oplus V_m. \tag{2.25}$$

Theorem 2.13 *Let* $\boldsymbol{P}_i$ $(i = 1, \cdots, m)$ *be square matrices that satisfy*

$$\boldsymbol{P}_1 + \boldsymbol{P}_2 + \cdots + \boldsymbol{P}_m = \boldsymbol{I}_n. \tag{2.26}$$

Then the following three statements are equivalent:

$$\boldsymbol{P}_i\boldsymbol{P}_j = \boldsymbol{O} \ \ (i \neq j). \tag{2.27}$$

$$\boldsymbol{P}_i^2 = \boldsymbol{P}_i \ \ (i = 1, \cdots m). \tag{2.28}$$

$$\mathrm{rank}(\boldsymbol{P}_1) + \mathrm{rank}(\boldsymbol{P}_2) + \cdots + \mathrm{rank}(\boldsymbol{P}_m) = n. \tag{2.29}$$

Proof. (i) → (ii): Multiply (2.26) by $\boldsymbol{P}_i$.

(ii) → (iii): Use $\text{rank}(\boldsymbol{P}_i) = \text{tr}(\boldsymbol{P}_i)$ when $\boldsymbol{P}_i^2 = \boldsymbol{P}_i$. Then,

$$\sum_{i=1}^{m} \text{rank}(\boldsymbol{P}_i) = \sum_{i=1}^{m} \text{tr}(\boldsymbol{P}_i) = \text{tr}\left(\sum_{i=1}^{m} \boldsymbol{P}_i\right) = \text{tr}(\boldsymbol{I}_n) = n.$$

(iii) → (i), (ii): Let $V_i = \text{Sp}(\boldsymbol{P}_i)$. From $\text{rank}(\boldsymbol{P}_i) = \dim(V_i)$, we obtain $\dim(V_1) + \dim(V_2) + \cdots \dim(V_m) = n$; that is, E^n is decomposed into the sum of m disjoint subspaces as in (2.26). By postmultiplying (2.26) by $\boldsymbol{P}_i$, we obtain

$$\boldsymbol{P}_1\boldsymbol{P}_i + \boldsymbol{P}_2\boldsymbol{P}_i + \cdots + \boldsymbol{P}_i(\boldsymbol{P}_i - \boldsymbol{I}_n) + \cdots + \boldsymbol{P}_m\boldsymbol{P}_i = \boldsymbol{O}.$$

Since $\text{Sp}(\boldsymbol{P}_1), \text{Sp}(\boldsymbol{P}_2), \cdots, \text{Sp}(\boldsymbol{P}_m)$ are disjoint, (2.27) and (2.28) hold from Theorem 1.4. Q.E.D.

Note $\boldsymbol{P}_i$ in Theorem 2.13 is a projection matrix. Let $E^n = V_1 \oplus \cdots \oplus V_r$, and let

$$V_{(i)} = V_1 \oplus \cdots \oplus V_{i-1} \oplus V_{i+1} \oplus \cdots \oplus V_r. \tag{2.30}$$

Then, $E^n = V_i \oplus V_{(i)}$. Let $\boldsymbol{P}_{i\cdot(i)}$ denote the projector onto V_i along $V_{(i)}$. This matrix coincides with the $\boldsymbol{P}_i$ that satisfies the four equations given in (2.26) through (2.29).

The following relations hold.

Corollary 1

$$\boldsymbol{P}_{1\cdot(1)} + \boldsymbol{P}_{2\cdot(2)} + \cdots + \boldsymbol{P}_{m\cdot(m)} = \boldsymbol{I}_n, \tag{2.31}$$

$$\boldsymbol{P}_{i\cdot(i)}^2 = \boldsymbol{P}_{i\cdot(i)} \ \ (i = 1, \cdots, m), \tag{2.32}$$

$$\boldsymbol{P}_{i\cdot(i)}\boldsymbol{P}_{j\cdot(j)} = \boldsymbol{O} \ \ (i \neq j). \tag{2.33}$$

Corollary 2 *Let $\boldsymbol{P}_{(i)\cdot i}$ denote the projector onto $V_{(i)}$ along V_i. Then the following relation holds:*

$$\boldsymbol{P}_{(i)\cdot i} = \boldsymbol{P}_{1\cdot(1)} + \cdots + \boldsymbol{P}_{i-1\cdot(i-1)} + \boldsymbol{P}_{i+1\cdot(i+1)} + \cdots + \boldsymbol{P}_{m\cdot(m)}. \tag{2.34}$$

Proof. The proof is straightforward by noting $\boldsymbol{P}_{i\cdot(i)} + \boldsymbol{P}_{(i)\cdot i} = \boldsymbol{I}_n$. Q.E.D.

Note The projection matrix $\boldsymbol{P}_{i\cdot(i)}$ onto V_i along $V_{(i)}$ is uniquely determined. Assume that there are two possible representations, $\boldsymbol{P}_{i\cdot(i)}$ and $\boldsymbol{P}^*_{i\cdot(i)}$. Then,

$$\boldsymbol{P}_{1\cdot(1)} + \boldsymbol{P}_{2\cdot(2)} + \cdots + \boldsymbol{P}_{m\cdot(m)} = \boldsymbol{P}^*_{1\cdot(1)} + \boldsymbol{P}^*_{2\cdot(2)} + \cdots + \boldsymbol{P}^*_{m\cdot(m)},$$

from which

$$(\boldsymbol{P}_{1\cdot(1)} - \boldsymbol{P}^*_{1\cdot(1)}) + (\boldsymbol{P}_{2\cdot(2)} - \boldsymbol{P}^*_{2\cdot(2)}) + \cdots + (\boldsymbol{P}_{m\cdot(m)} - \boldsymbol{P}^*_{m\cdot(m)}) = \boldsymbol{O}.$$

Each term in the equation above belongs to one of the respective subspaces V_1, V_2, $\cdots, V_m$, which are mutually disjoint. Hence, from Theorem 1.4, we obtain $\boldsymbol{P}_{i\cdot(i)} = \boldsymbol{P}^*_{i\cdot(i)}$. This indicates that when a direct-sum of E^n is given, an identity matrix $\boldsymbol{I}_n$ of order n is decomposed accordingly, and the projection matrices that constitute the decomposition are uniquely determined.

The following theorem due to Khatri (1968) generalizes Theorem 2.13.

Theorem 2.14 *Let $\boldsymbol{P}_i$ denote a square matrix of order n such that*

$$\boldsymbol{P} = \boldsymbol{P}_1 + \boldsymbol{P}_2 + \cdots + \boldsymbol{P}_m. \tag{2.35}$$

Consider the following four propositions:

(i) $\boldsymbol{P}_i^2 = \boldsymbol{P}_i \quad (i = 1, \cdots m)$,

(ii) $\boldsymbol{P}_i\boldsymbol{P}_j = \boldsymbol{O} \quad (i \neq j)$, *and* $\mathrm{rank}(\boldsymbol{P}_i^2) = \mathrm{rank}(\boldsymbol{P}_i)$,

(iii) $\boldsymbol{P}^2 = \boldsymbol{P}$,

(iv) $\mathrm{rank}(\boldsymbol{P}) = \mathrm{rank}(\boldsymbol{P}_1) + \cdots + \mathrm{rank}(\boldsymbol{P}_m)$.

All other propositions can be derived from any two of (i), (ii), *and* (iii), *and* (i) *and* (ii) *can be derived from* (iii) *and* (iv).

Proof. That (i) and (ii) imply (iii) is obvious. To show that (ii) and (iii) imply (iv), we may use

$$\boldsymbol{P}^2 = \boldsymbol{P}_1^2 + \boldsymbol{P}_2^2 + \cdots + \boldsymbol{P}_m^2 \text{ and } \boldsymbol{P}^2 = \boldsymbol{P},$$

which follow from (2.35).

(ii), (iii) $\rightarrow$ (i): Postmultiplying (2.35) by $\boldsymbol{P}_i$, we obtain $\boldsymbol{P}\boldsymbol{P}_i = \boldsymbol{P}_i^2$, from which it follows that $\boldsymbol{P}_i^3 = \boldsymbol{P}_i^2$. On the other hand, $\mathrm{rank}(\boldsymbol{P}_i^2) = \mathrm{rank}(\boldsymbol{P}_i)$ implies that there exists $\boldsymbol{W}$ such that $\boldsymbol{P}_i^2\boldsymbol{W}_i = \boldsymbol{P}_i$. Hence, $\boldsymbol{P}_i^3 = \boldsymbol{P}_i^2 \Rightarrow \boldsymbol{P}_i^3\boldsymbol{W}_i = \boldsymbol{P}_i^2\boldsymbol{W}_i \Rightarrow \boldsymbol{P}_i(\boldsymbol{P}_i^2\boldsymbol{W}_i) = \boldsymbol{P}_i^2\boldsymbol{W} \Rightarrow \boldsymbol{P}_i^2 = \boldsymbol{P}_i$.

(iii), (iv) → (i), (ii): We have $\mathrm{Sp}(\boldsymbol{P}) \oplus \mathrm{Sp}(\boldsymbol{I}_n - \boldsymbol{P}) = E^n$ from $\boldsymbol{P}^2 = \boldsymbol{P}$. Hence, by postmultiplying the identity

$$\boldsymbol{P}_1 + \boldsymbol{P}_2 + \cdots + \boldsymbol{P}_m + (\boldsymbol{I}_n - \boldsymbol{P}) = \boldsymbol{I}_n$$

by $\boldsymbol{P}$, we obtain $\boldsymbol{P}_i^2 = \boldsymbol{P}_i$, and $\boldsymbol{P}_i\boldsymbol{P}_j = \boldsymbol{O}$ $(i \neq j)$. Q.E.D.

Next we consider the case in which subspaces have inclusion relationships like the following.

Theorem 2.15 *Let*

$$E^n = V_k \supset V_{k-1} \supset \cdots \supset V_2 \supset V_1 = \{\mathbf{0}\},$$

and let W_i *denote a complement subspace of* V_i. *Let* $\boldsymbol{P}_i$ *be the orthogonal projector onto* V_i *along* W_i, *and let* $\boldsymbol{P}_i^* = \boldsymbol{P}_i - \boldsymbol{P}_{i-1}$, *where* $\boldsymbol{P}_0 = \boldsymbol{O}$ *and* $\boldsymbol{P}_k = \boldsymbol{I}_n$. *Then the following relations hold:*

(i) $\boldsymbol{I}_n = \boldsymbol{P}_1^* + \boldsymbol{P}_2^* + \cdots + \boldsymbol{P}_k^*$.

(ii) $(\boldsymbol{P}_i^*)^2 = \boldsymbol{P}_i^*$.

(iii) $\boldsymbol{P}_i^*\boldsymbol{P}_j^* = \boldsymbol{P}_j^*\boldsymbol{P}_i^* = \boldsymbol{O}$ $(i \neq j)$.

(iv) $\boldsymbol{P}_i$ *is the projector onto* $V_i \cap W_{i-1}$ *along* $V_{i-1} \oplus W_i$.

Proof. (i): Obvious. (ii): Use $\boldsymbol{P}_i\boldsymbol{P}_{i-1} = \boldsymbol{P}_{i-1}\boldsymbol{P}_i = \boldsymbol{P}_{i-1}$. (iii): It follows from $(\boldsymbol{P}_i^*)^2 = \boldsymbol{P}_i^*$ that $\mathrm{rank}(\boldsymbol{P}_i^*) = \mathrm{tr}(\boldsymbol{P}_i^*) = \mathrm{tr}(\boldsymbol{P}_i - \boldsymbol{P}_{i-1}) = \mathrm{tr}(\boldsymbol{P}_i) - \mathrm{tr}(\boldsymbol{P}_{i-1})$. Hence, $\sum_{i=1}^k \mathrm{rank}(\boldsymbol{P}_i^*) = \mathrm{tr}(\boldsymbol{P}_k) - \mathrm{tr}(\boldsymbol{P}_0) = n$, from which $\boldsymbol{P}_i^*\boldsymbol{P}_j^* = \boldsymbol{O}$ follows by Theorem 2.13. (iv): Clear from Theorem 2.8(i). Q.E.D.

Note The theorem above does not presuppose that $\boldsymbol{P}_i$ is an orthogonal projector. However, if $W_i = V_i^\perp$, $\boldsymbol{P}_i$ and $\boldsymbol{P}_i^*$ are orthogonal projectors. The latter, in particular, is the orthogonal projector onto $V_i \cap V_{i-1}^\perp$.

2.3.2 Decomposition into nondisjoint subspaces

In this section, we present several theorems indicating how projectors are decomposed when the corresponding subspaces are not necessarily disjoint. We elucidate their meaning in connection with the commutativity of projectors.

We first consider the case in which there are two direct-sum decompositions of E^n, namely

$$E^n = V_1 \oplus W_1 = V_2 \oplus W_2,$$

as given in (2.21). Let $V_{12} = V_1 \cap V_2$ denote the product space between V_1 and V_2, and let V_3 denote a complement subspace to $V_1 + V_2$ in E^n. Furthermore, let $\boldsymbol{P}_{1+2}$ denote the projection matrix onto $V_{1+2} = V_1 + V_2$ along V_3, and let $\boldsymbol{P}_j$ $(j = 1, 2)$ represent the projection matrix onto V_j $(j = 1, 2)$ along W_j $(j = 1, 2)$. Then the following theorem holds.

Theorem 2.16 (i) *The necessary and sufficient condition for* $\boldsymbol{P}_{1+2} = \boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_2$ *is*

$$(V_{1+2} \cap W_2) \subset (V_1 \oplus V_3). \tag{2.36}$$

(ii) *The necessary and sufficient condition for* $\boldsymbol{P}_{1+2} = \boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_2\boldsymbol{P}_1$ *is*

$$(V_{1+2} \cap W_1) \subset (V_2 \oplus V_3). \tag{2.37}$$

Proof. (i): Since $V_{1+2} \supset V_1$ and $V_{1+2} \supset V_2$, $\boldsymbol{P}_{1+2} - \boldsymbol{P}_1$ is the projector onto $V_{1+2} \cap W_1$ along $V_1 \oplus V_3$ by Theorem 2.8. Hence, $\boldsymbol{P}_{1+2}\boldsymbol{P}_1 = \boldsymbol{P}_1$ and $\boldsymbol{P}_{1+2}\boldsymbol{P}_2 = \boldsymbol{P}_2$. Similarly, $\boldsymbol{P}_{1+2} - \boldsymbol{P}_2$ is the projector onto $V_{1+2} \cap W_2$ along $V_2 \oplus V_3$. Hence, by Theorem 2.8,

$$\boldsymbol{P}_{1+2} - \boldsymbol{P}_1 - \boldsymbol{P}_2 + \boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{O} \Longleftrightarrow (\boldsymbol{P}_{1+2} - \boldsymbol{P}_1)(\boldsymbol{P}_{1+2} - \boldsymbol{P}_2) = \boldsymbol{O}.$$

Furthermore,

$$(\boldsymbol{P}_{1+2} - \boldsymbol{P}_1)(\boldsymbol{P}_{1+2} - \boldsymbol{P}_2) = \boldsymbol{O} \Longleftrightarrow (V_{1+2} \cap W_2) \subset (V_1 \oplus V_3).$$

(ii): Similarly, $\boldsymbol{P}_{1+2} - \boldsymbol{P}_1 - \boldsymbol{P}_2 + \boldsymbol{P}_2\boldsymbol{P}_1 = \boldsymbol{O} \Longleftrightarrow (\boldsymbol{P}_{1+2} - \boldsymbol{P}_2)(\boldsymbol{P}_{1+2} - \boldsymbol{P}_1) = \boldsymbol{O} \Longleftrightarrow (V_{1+2} \cap W_1) \subset (V_2 \oplus V_3)$. Q.E.D.

Corollary *Assume that the decomposition (2.21) holds. The necessary and sufficient condition for* $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1$ *is that both (2.36) and (2.37) hold.*

The following theorem can readily be derived from the theorem above.

Theorem 2.17 *Let* $E^n = (V_1 + V_2) \oplus V_3$, $V_1 = V_{11} \oplus V_{12}$, *and* $V_2 = V_{22} \oplus V_{12}$, *where* $V_{12} = V_1 \cap V_2$. *Let* $\boldsymbol{P}^*_{1+2}$ *denote the projection matrix onto* $V_1 + V_2$ *along* V_3, *and let* $\boldsymbol{P}^*_1$ *and* $\boldsymbol{P}^*_2$ *denote the projectors onto* V_1 *along* $V_3 \oplus V_{22}$ *and onto* V_2 *along* $V_3 \oplus V_{11}$, *respectively. Then,*

$$\boldsymbol{P}^*_1\boldsymbol{P}^*_2 = \boldsymbol{P}^*_2\boldsymbol{P}^*_1 \tag{2.38}$$

and

$$\boldsymbol{P}^*_{1+2} = \boldsymbol{P}^*_1 + \boldsymbol{P}^*_2 - \boldsymbol{P}^*_1\boldsymbol{P}^*_2. \tag{2.39}$$

Proof. Since $V_{11} \subset V_1$ and $V_{22} \subset V_2$, we obtain

$$V_{1+2} \cap W_2 = V_{11} \subset (V_1 \oplus V_3) \text{ and } V_{1+2} \cap W_1 = V_{22} \subset (V_2 \oplus V_3)$$

by setting $W_1 = V_{22} \oplus V_3$ and $W_2 = V_{11} \oplus V_3$ in Theorem 2.16.
Another proof. Let $\boldsymbol{y} = \boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_{12} + \boldsymbol{y}_3 \in E^n$, where $\boldsymbol{y}_1 \in V_{11}$, $\boldsymbol{y}_2 \in V_{22}$, $\boldsymbol{y}_{12} \in V_{12}$, and $\boldsymbol{y}_3 \in V_3$. Then it suffices to show that $(\boldsymbol{P}^*_1\boldsymbol{P}^*_2)\boldsymbol{y} = (\boldsymbol{P}^*_2\boldsymbol{P}^*_1)\boldsymbol{y}$. Q.E.D.

Let $\boldsymbol{P}_j$ $(j = 1, 2)$ denote the projection matrix onto V_j along W_j. Assume that $E^n = V_1 \oplus W_1 \oplus V_3 = V_2 \oplus W_2 \oplus V_3$ and $V_1 + V_2 = V_{11} \oplus V_{22} \oplus V_{12}$ hold. However, $W_1 = V_{22}$ may not hold, even if $V_1 = V_{11} \oplus V_{12}$. That is, (2.38) and (2.39) hold only when we set $W_1 = V_{22}$ and $W_2 = V_{11}$.

Theorem 2.18 *Let $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ be the orthogonal projectors onto V_1 and V_2, respectively, and let $\boldsymbol{P}_{1+2}$ denote the orthogonal projector onto V_{1+2}. Let $V_{12} = V_1 \cap V_2$. Then the following three statements are equivalent:*

(i) $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1$.

(ii) $\boldsymbol{P}_{1+2} = \boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_2$.

(iii) $V_{11} = V_1 \cap V_{12}^{\perp}$ *and* $V_{22} = V_2 \cap V_{12}^{\perp}$ *are orthogonal.*

Proof. (i) $\to$ (ii): Obvious from Theorem 2.16.

(ii) $\to$ (iii): $\boldsymbol{P}_{1+2} = \boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_2 \Rightarrow (\boldsymbol{P}_{1+2} - \boldsymbol{P}_1)(\boldsymbol{P}_{1+2} - \boldsymbol{P}_2) = (\boldsymbol{P}_{1+2} - \boldsymbol{P}_2)(\boldsymbol{P}_{1+2} - \boldsymbol{P}_1) = \boldsymbol{O} \Rightarrow V_{11}$ and V_{22} are orthogonal.

(iii) $\to$ (i): Set $V_3 = (V_1 + V_2)^{\perp}$ in Theorem 2.17. Since V_{11} and V_{22}, and V_1 and V_{22}, are orthogonal, the result follows. Q.E.D.

When $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, and $\boldsymbol{P}_{1+2}$ are orthogonal projectors, the following corollary holds.

Corollary $\boldsymbol{P}_{1+2} = \boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_2 \Longleftrightarrow \boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1$.

2.3.3 Commutative projectors

In this section, we focus on orthogonal projectors and discuss the meaning of Theorem 2.18 and its corollary. We also generalize the results to the case in which there are three or more subspaces.

Theorem 2.19 *Let $\boldsymbol{P}_j$ denote the orthogonal projector onto V_j. If $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1$, $\boldsymbol{P}_1\boldsymbol{P}_3 = \boldsymbol{P}_3\boldsymbol{P}_1$, and $\boldsymbol{P}_2\boldsymbol{P}_3 = \boldsymbol{P}_3\boldsymbol{P}_2$, the following relations hold:*

$$V_1 + (V_2 \cap V_3) = (V_1 + V_2) \cap (V_1 + V_3), \tag{2.40}$$

$$V_2 + (V_1 \cap V_3) = (V_1 + V_2) \cap (V_2 + V_3), \tag{2.41}$$

$$V_3 + (V_1 \cap V_2) = (V_1 + V_3) \cap (V_2 + V_3). \tag{2.42}$$

Proof. Let $\boldsymbol{P}_{1+(2\cap 3)}$ denote the orthogonal projector onto $V_1 + (V_2 \cap V_3)$. Then the orthogonal projector onto $V_2 \cap V_3$ is given by $\boldsymbol{P}_2\boldsymbol{P}_3$ (or by $\boldsymbol{P}_3\boldsymbol{P}_2$). Since $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1 \Rightarrow \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3 = \boldsymbol{P}_2\boldsymbol{P}_3\boldsymbol{P}_1$, we obtain

$$\boldsymbol{P}_{1+(2\cap 3)} = \boldsymbol{P}_1 + \boldsymbol{P}_2\boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3$$

by Theorem 2.18. On the other hand, from $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1$ and $\boldsymbol{P}_1\boldsymbol{P}_3 = \boldsymbol{P}_3\boldsymbol{P}_1$, the orthogonal projectors onto $V_1 + V_2$ and $V_1 + V_3$ are given by

$$\boldsymbol{P}_{1+2} = \boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_2 \text{ and } \boldsymbol{P}_{1+3} = \boldsymbol{P}_1 + \boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_3,$$

respectively, and so $\boldsymbol{P}_{1+2}\boldsymbol{P}_{1+3} = \boldsymbol{P}_{1+3}\boldsymbol{P}_{1+2}$ holds. Hence, the orthogonal projector onto $(V_1 + V_2) \cap (V_1 + V_3)$ is given by

$$(\boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_2)(\boldsymbol{P}_1 + \boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_3) = \boldsymbol{P}_1 + \boldsymbol{P}_2\boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3,$$

which implies $\boldsymbol{P}_{1+(2\cap 3)} = \boldsymbol{P}_{1+2}\boldsymbol{P}_{1+3}$. Since there is a one-to-one correspondence between projectors and subspaces, (2.40) holds.

Relations (2.41) and (2.42) can be similarly proven by noting that $(\boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_2)(\boldsymbol{P}_2 + \boldsymbol{P}_3 - \boldsymbol{P}_2\boldsymbol{P}_3) = \boldsymbol{P}_2 + \boldsymbol{P}_1\boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3$ and $(\boldsymbol{P}_1 + \boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_3)(\boldsymbol{P}_2 + \boldsymbol{P}_3 - \boldsymbol{P}_2\boldsymbol{P}_3) = \boldsymbol{P}_3 + \boldsymbol{P}_1\boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3$, respectively.

Q.E.D.

The three identities from (2.40) to (2.42) indicate the distributive law of subspaces, which holds only if the commutativity of orthogonal projectors holds.

We now present a theorem on the decomposition of the orthogonal projectors defined on the sum space $V_1 + V_2 + V_3$ of V_1, V_2, and V_3.

Theorem 2.20 *Let $\boldsymbol{P}_{1+2+3}$ denote the orthogonal projector onto $V_1 + V_2 + V_3$, and let $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, and $\boldsymbol{P}_3$ denote the orthogonal projectors onto V_1, V_2, and V_3, respectively. Then a sufficient condition for the decomposition*

$$\boldsymbol{P}_{1+2+3} = \boldsymbol{P}_1 + \boldsymbol{P}_2 + \boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_2 - \boldsymbol{P}_2\boldsymbol{P}_3 - \boldsymbol{P}_3\boldsymbol{P}_1 + \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3 \tag{2.43}$$

to hold is

$$\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1, \;\; \boldsymbol{P}_2\boldsymbol{P}_3 = \boldsymbol{P}_3\boldsymbol{P}_2, \; and \; \boldsymbol{P}_1\boldsymbol{P}_3 = \boldsymbol{P}_3\boldsymbol{P}_1. \tag{2.44}$$

Proof. $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1 \Rightarrow \boldsymbol{P}_{1+2} = \boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_2$ and $\boldsymbol{P}_2\boldsymbol{P}_3 = \boldsymbol{P}_3\boldsymbol{P}_2 \Rightarrow \boldsymbol{P}_{2+3} = \boldsymbol{P}_2 + \boldsymbol{P}_3 - \boldsymbol{P}_2\boldsymbol{P}_3$. We therefore have $\boldsymbol{P}_{1+2}\boldsymbol{P}_{2+3} = \boldsymbol{P}_{2+3}\boldsymbol{P}_{1+2}$. We also have $\boldsymbol{P}_{1+2+3} = \boldsymbol{P}_{(1+2)+(1+3)}$, from which it follows that

$$\begin{aligned}
\boldsymbol{P}_{1+2+3} &= \boldsymbol{P}_{(1+2)+(1+3)} = \boldsymbol{P}_{1+2} + \boldsymbol{P}_{1+3} - \boldsymbol{P}_{1+2}\boldsymbol{P}_{1+3} \\
&= (\boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_2) + (\boldsymbol{P}_1 + \boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_3) \\
&\qquad - (\boldsymbol{P}_2\boldsymbol{P}_3 + \boldsymbol{P}_1 - \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3) \\
&= \boldsymbol{P}_1 + \boldsymbol{P}_2 + \boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_2 - \boldsymbol{P}_2\boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_3 + \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3.
\end{aligned}$$

An alternative proof. From $\boldsymbol{P}_1\boldsymbol{P}_{2+3} = \boldsymbol{P}_{2+3}\boldsymbol{P}_1$, we have $\boldsymbol{P}_{1+2+3} = \boldsymbol{P}_1 + \boldsymbol{P}_{2+3} - \boldsymbol{P}_1\boldsymbol{P}_{2+3}$. If we substitute $\boldsymbol{P}_{2+3} = \boldsymbol{P}_2 + \boldsymbol{P}_3 - \boldsymbol{P}_2\boldsymbol{P}_3$ into this equation, we obtain (2.43). Q.E.D.

Assume that (2.44) holds, and let

$$\boldsymbol{P}_{\tilde{1}} = \boldsymbol{P}_1 - \boldsymbol{P}_1\boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_3 + \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3,$$

$$\boldsymbol{P}_{\tilde{2}} = \boldsymbol{P}_2 - \boldsymbol{P}_2\boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_2 + \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3,$$

$$\boldsymbol{P}_{\tilde{3}} = \boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_3 - \boldsymbol{P}_2\boldsymbol{P}_3 + \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3,$$

$$\boldsymbol{P}_{12(3)} = \boldsymbol{P}_1\boldsymbol{P}_2 - \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3,$$

$$P_{13(2)} = \boldsymbol{P}_1\boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3,$$

$$\boldsymbol{P}_{23(1)} = \boldsymbol{P}_2\boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3,$$

and

$$\boldsymbol{P}_{123} = \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3.$$

Then,

$$\boldsymbol{P}_{1+2+3} = \boldsymbol{P}_{\tilde{1}} + \boldsymbol{P}_{\tilde{2}} + \boldsymbol{P}_{\tilde{3}} + \boldsymbol{P}_{12(3)} + \boldsymbol{P}_{13(2)} + \boldsymbol{P}_{23(1)} + \boldsymbol{P}_{123}. \tag{2.45}$$

Additionally, all matrices on the right-hand side of (2.45) are orthogonal projectors, which are also all mutually orthogonal.

Note Since $\boldsymbol{P}_{\tilde{1}} = \boldsymbol{P}_1(\boldsymbol{I}_n - \boldsymbol{P}_{2+3})$, $\boldsymbol{P}_{\tilde{2}} = \boldsymbol{P}_2(\boldsymbol{I}_n - \boldsymbol{P}_{1+3})$, $\boldsymbol{P}_{\tilde{3}} = \boldsymbol{P}_3(\boldsymbol{I} - \boldsymbol{P}_{1+2})$, $\boldsymbol{P}_{12(3)} = \boldsymbol{P}_1\boldsymbol{P}_2(\boldsymbol{I}_n - \boldsymbol{P}_3)$, $\boldsymbol{P}_{13(2)} = \boldsymbol{P}_1\boldsymbol{P}_3(\boldsymbol{I}_n - \boldsymbol{P}_2)$, and $\boldsymbol{P}_{23(1)} = \boldsymbol{P}_2\boldsymbol{P}_3(\boldsymbol{I}_n - \boldsymbol{P}_1)$, the decomposition of the projector $\boldsymbol{P}_{1\cup 2\cup 3}$ corresponds with the decomposition of the subspace $V_1 + V_2 + V_3$

$$V_1 + V_2 + V_3 = V_{\tilde{1}} \dot{\oplus} V_{\tilde{2}} \dot{\oplus} V_{\tilde{3}} \dot{\oplus} V_{12(3)} \dot{\oplus} V_{13(2)} \dot{\oplus} V_{23(1)} \dot{\oplus} V_{123}, \tag{2.46}$$

where $V_{\tilde{1}} = V_1 \cap (V_2 + V_3)^{\perp}$, $V_{\tilde{2}} = V_2 \cap (V_1 + V_3)^{\perp}$, $V_{\tilde{3}} = V_3 \cap (V_1 + V_2)^{\perp}$, $V_{12(3)} = V_1 \cap V_2 \cap V_3^{\perp}$, $V_{13(2)} = V_1 \cap V_2^{\perp} \cap V_3$, $V_{23(1)} = V_1^{\perp} \cap V_2 \cap V_3$, and $V_{123} = V_1 \cap V_2 \cap V_3$.

Theorem 2.20 can be generalized as follows.

Corollary *Let $V = V_1 + V_2 + \cdots + V_s$ $(s \geq 2)$. Let $\boldsymbol{P}_V$ denote the orthogonal projector onto V, and let $\boldsymbol{P}_j$ denote the orthogonal projector onto V_j. A sufficient condition for*

$$\boldsymbol{P}_V = \sum_{j=1}^{s} \boldsymbol{P}_j - \sum_{i<j} \boldsymbol{P}_i \boldsymbol{P}_j + \sum_{i<j<k} \boldsymbol{P}_i \boldsymbol{P}_j \boldsymbol{P}_k + \cdots + (-1)^{s-1} \boldsymbol{P}_1 \boldsymbol{P}_2 \boldsymbol{P}_3 \cdots \boldsymbol{P}_s \tag{2.47}$$

to hold is

$$\boldsymbol{P}_i \boldsymbol{P}_j = \boldsymbol{P}_j \boldsymbol{P}_i \ \ (i \neq j). \tag{2.48}$$

2.3.4 Noncommutative projectors

We now consider the case in which two subspaces V_1 and V_2 and the corresponding projectors $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are given but $\boldsymbol{P}_1 \boldsymbol{P}_2 = \boldsymbol{P}_2 \boldsymbol{P}_1$ does not necessarily hold. Let $\boldsymbol{Q}_j = \boldsymbol{I}_n - \boldsymbol{P}_j$ $(j = 1, 2)$. Then the following lemma holds.

Lemma 2.4

$$\begin{aligned} V_1 + V_2 &= \mathrm{Sp}(\boldsymbol{P}_1) \oplus \mathrm{Sp}(\boldsymbol{Q}_1 \boldsymbol{P}_2) & (2.49) \\ &= \mathrm{Sp}(\boldsymbol{Q}_2 \boldsymbol{P}_1) \oplus \mathrm{Sp}(\boldsymbol{P}_2). & (2.50) \end{aligned}$$

Proof. $[\boldsymbol{P}_1, \boldsymbol{Q}_1 \boldsymbol{P}_2]$ and $[\boldsymbol{Q}_2 \boldsymbol{P}_1, \boldsymbol{P}_2]$ can be expressed as

$$[\boldsymbol{P}_1, \boldsymbol{Q}_1 \boldsymbol{P}_2] = [\boldsymbol{P}_1, \boldsymbol{P}_2] \begin{bmatrix} \boldsymbol{I}_n & -\boldsymbol{P}_2 \\ \boldsymbol{O} & \boldsymbol{I}_n \end{bmatrix} = [\boldsymbol{P}_1, \boldsymbol{P}_2] \boldsymbol{S}$$

and

$$[\boldsymbol{Q}_2 \boldsymbol{P}_1, \boldsymbol{P}_2] = [\boldsymbol{P}_1, \boldsymbol{P}_2] \begin{bmatrix} \boldsymbol{I}_n & \boldsymbol{O} \\ -\boldsymbol{P}_1 & \boldsymbol{I}_n \end{bmatrix} = [\boldsymbol{P}_1, \boldsymbol{P}_2] \boldsymbol{T}.$$

Since $\boldsymbol{S}$ and $\boldsymbol{T}$ are nonsingular, we have

$$\text{rank}(\boldsymbol{P}_1, \boldsymbol{P}_2) = \text{rank}(\boldsymbol{P}_1, \boldsymbol{Q}_1\boldsymbol{P}_2) = \text{rank}(\boldsymbol{Q}_2\boldsymbol{P}_1, \boldsymbol{P}_1),$$

which implies

$$V_1 + V_2 = \text{Sp}(\boldsymbol{P}_1, \boldsymbol{Q}_1\boldsymbol{P}_2) = \text{Sp}(\boldsymbol{Q}_2\boldsymbol{P}_1, \boldsymbol{P}_2).$$

Furthermore, let $\boldsymbol{P}_1\boldsymbol{x} + \boldsymbol{Q}_1\boldsymbol{P}_2\boldsymbol{y} = \boldsymbol{0}$. Premultiplying both sides by $\boldsymbol{P}_1$, we obtain $\boldsymbol{P}_1\boldsymbol{x} = \boldsymbol{0}$ (since $\boldsymbol{P}_1\boldsymbol{Q}_1 = \boldsymbol{O}$), which implies $\boldsymbol{Q}_1\boldsymbol{P}_2\boldsymbol{y} = \boldsymbol{0}$. Hence, $\text{Sp}(\boldsymbol{P}_1)$ and $\text{Sp}(\boldsymbol{Q}_1\boldsymbol{P}_2)$ give a direct-sum decomposition of $V_1 + V_2$, and so do $\text{Sp}(\boldsymbol{Q}_2\boldsymbol{P}_1)$ and $\text{Sp}(\boldsymbol{P}_2)$. Q.E.D.

The following theorem follows from Lemma 2.4.

Theorem 2.21 *Let $E^n = (V_1 + V_2) \oplus W$. Furthermore, let*

$$V_{2[1]} = \{\boldsymbol{x} | \boldsymbol{x} = \boldsymbol{Q}_1\boldsymbol{y}, \boldsymbol{y} \in V_2\} \tag{2.51}$$

and

$$V_{1[2]} = \{\boldsymbol{x} | \boldsymbol{x} = \boldsymbol{Q}_2\boldsymbol{y}, \boldsymbol{y} \in V_1\}. \tag{2.52}$$

Let $\boldsymbol{Q}_j = \boldsymbol{I}_n - \boldsymbol{P}_j$ $(j = 1, 2)$, where $\boldsymbol{P}_j$ is the orthogonal projector onto V_j, and let $\boldsymbol{P}^*$, $\boldsymbol{P}_1^*$, $\boldsymbol{P}_2^*$, $\boldsymbol{P}_{1[2]}$, and $\boldsymbol{P}_{2[1]}$ denote the projectors onto $V_1 + V_2$ along W, onto V_1 along $V_{2[1]} \oplus W$, onto V_2 along $V_{1[2]} \oplus W$, onto $V_{1[2]}$ along $V_2 \oplus W$, and onto $V_{2[1]}$ along $V_1 \oplus W$, respectively. Then,

$$\boldsymbol{P}^* = \boldsymbol{P}_1^* + \boldsymbol{P}_{2[1]}^* \tag{2.53}$$

or

$$\boldsymbol{P}^* = \boldsymbol{P}_{1[2]}^* + \boldsymbol{P}_2^* \tag{2.54}$$

holds.

Note When $W = (V_1 + V_2)^\perp$, $\boldsymbol{P}_j^*$ is the orthogonal projector onto V_j, while $\boldsymbol{P}_{j[i]}^*$ is the orthogonal projector onto $V_j[i]$.

Corollary *Let $\boldsymbol{P}$ denote the orthogonal projector onto $V = V_1 \oplus V_2$, and let $\boldsymbol{P}_j$ $(j = 1, 2)$ be the orthogonal projectors onto V_j. If V_i and V_j are orthogonal, the following equation holds:*

$$\boldsymbol{P} = \boldsymbol{P}_1 + \boldsymbol{P}_2. \tag{2.55}$$

2.4 Norm of Projection Vectors

We now present theorems concerning the norm of the projection vector $\boldsymbol{Px}$ ($\boldsymbol{x} \in E^n$) obtained by projecting $\boldsymbol{x}$ onto $\mathrm{Sp}(\boldsymbol{P})$ along $\mathrm{Ker}(\boldsymbol{P})$ by $\boldsymbol{P}$.

Lemma 2.5 $\boldsymbol{P}' = \boldsymbol{P}$ *and* $\boldsymbol{P}^2 = \boldsymbol{P} \Longleftrightarrow \boldsymbol{P}'\boldsymbol{P} = \boldsymbol{P}$.

(The proof is trivial and hence omitted.)

Theorem 2.22 *Let* $\boldsymbol{P}$ *denote a projection matrix (i.e.,* $\boldsymbol{P}^2 = \boldsymbol{P}$*). The necessary and sufficient condition to have*

$$||\boldsymbol{Px}|| \leq ||\boldsymbol{x}|| \tag{2.56}$$

for an arbitrary vector $\boldsymbol{x}$ *is*

$$\boldsymbol{P}' = \boldsymbol{P}. \tag{2.57}$$

Proof. (Sufficiency) Let $\boldsymbol{x}$ be decomposed as $\boldsymbol{x} = \boldsymbol{Px} + (\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{x}$. We have $(\boldsymbol{Px})'(\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{x} = \boldsymbol{x}'(\boldsymbol{P}' - \boldsymbol{P}'\boldsymbol{P})\boldsymbol{x} = 0$ because $\boldsymbol{P}' = \boldsymbol{P} \Rightarrow \boldsymbol{P}'\boldsymbol{P} = \boldsymbol{P}'$ from Lemma 2.5. Hence,

$$||\boldsymbol{x}||^2 = ||\boldsymbol{Px}||^2 + ||(\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{x}||^2 \geq ||\boldsymbol{Px}||^2.$$

(Necessity) By assumption, we have $\boldsymbol{x}'(\boldsymbol{I}_n - \boldsymbol{P}'\boldsymbol{P})\boldsymbol{x} \geq 0$, which implies $\boldsymbol{I}_n - \boldsymbol{P}'\boldsymbol{P}$ is *nnd* with all nonnegative eigenvalues. Let $\lambda_1, \lambda_2, \cdots, \lambda_n$ denote the eigenvalues of $\boldsymbol{P}'\boldsymbol{P}$. Then, $1 - \lambda_j \geq 0$ or $0 \geq \lambda_j \geq 1$ $(j = 1, \cdots, n)$. Hence, $\sum_{j=1}^n \lambda_j^2 \leq \sum_{j=1}^n \lambda_j$, which implies $\mathrm{tr}(\boldsymbol{P}'\boldsymbol{P})^2 \leq \mathrm{tr}(\boldsymbol{P}'\boldsymbol{P})$.

On the other hand, we have

$$(\mathrm{tr}(\boldsymbol{P}'\boldsymbol{P}))^2 = (\mathrm{tr}(\boldsymbol{P}\boldsymbol{P}'\boldsymbol{P}))^2 \leq \mathrm{tr}(\boldsymbol{P}'\boldsymbol{P})\mathrm{tr}(\boldsymbol{P}'\boldsymbol{P})^2$$

from the generalized Schwarz inequality (set $\boldsymbol{A}' = \boldsymbol{P}$ and $\boldsymbol{B} = \boldsymbol{P}'\boldsymbol{P}$ in (1.19)) and $\boldsymbol{P}^2 = \boldsymbol{P}$. Hence, $\mathrm{tr}(\boldsymbol{P}'\boldsymbol{P}) \leq \mathrm{tr}(\boldsymbol{P}'\boldsymbol{P})^2 \Rightarrow \mathrm{tr}(\boldsymbol{P}'\boldsymbol{P}) = \mathrm{tr}(\boldsymbol{P}'\boldsymbol{P})^2$, from which it follows that $\mathrm{tr}\{(\boldsymbol{P} - \boldsymbol{P}'\boldsymbol{P})'(\boldsymbol{P} - \boldsymbol{P}'\boldsymbol{P})\} = \mathrm{tr}\{\boldsymbol{P}'\boldsymbol{P} - \boldsymbol{P}'\boldsymbol{P} - \boldsymbol{P}'\boldsymbol{P} + (\boldsymbol{P}'\boldsymbol{P})^2\} = \mathrm{tr}\{\boldsymbol{P}'\boldsymbol{P} - (\boldsymbol{P}'\boldsymbol{P})^2\} = 0$. Thus, $\boldsymbol{P} = \boldsymbol{P}'\boldsymbol{P} \Rightarrow \boldsymbol{P}' = \boldsymbol{P}$. Q.E.D.

Corollary *Let* $\boldsymbol{M}$ *be a symmetric pd matrix, and define the (squared) norm of* $\boldsymbol{x}$ *by*

$$||\boldsymbol{x}||_M^2 = \boldsymbol{x}'\boldsymbol{M}\boldsymbol{x}. \tag{2.58}$$

The necessary and sufficient condition for a projection matrix $\boldsymbol{P}$ (satisfying $\boldsymbol{P}^2 = \boldsymbol{P}$) to satisfy

$$||\boldsymbol{Px}||^2_M \leq ||\boldsymbol{x}||^2_M \tag{2.59}$$

for an arbitrary n-component vector $\boldsymbol{x}$ is given by

$$(\boldsymbol{MP})' = \boldsymbol{MP}. \tag{2.60}$$

Proof. Let $\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Delta}^2\boldsymbol{U}'$ be the spectral decomposition of $\boldsymbol{M}$, and let $\boldsymbol{M}^{1/2} = \boldsymbol{\Delta}\boldsymbol{U}'$. Then, $\boldsymbol{M}^{-1/2} = \boldsymbol{U}\boldsymbol{\Delta}^{-1}$. Define $\boldsymbol{y} = \boldsymbol{M}^{1/2}\boldsymbol{x}$, and let $\tilde{\boldsymbol{P}} = \boldsymbol{M}^{1/2}\boldsymbol{P}\boldsymbol{M}^{-1/2}$. Then, $\tilde{\boldsymbol{P}}^2 = \tilde{\boldsymbol{P}}$, and (2.58) can be rewritten as $||\tilde{\boldsymbol{P}}\boldsymbol{y}||^2 \leq ||\boldsymbol{y}||^2$. By Theorem 2.22, the necessary and sufficient condition for (2.59) to hold is given by

$$\tilde{\boldsymbol{P}}^2 = \tilde{\boldsymbol{P}} \Longrightarrow (\boldsymbol{M}^{1/2}\boldsymbol{P}\boldsymbol{M}^{-1/2})' = \boldsymbol{M}^{1/2}\boldsymbol{P}\boldsymbol{M}^{-1/2}, \tag{2.61}$$

leading to (2.60). Q.E.D.

Note The theorem above implies that with an oblique projector $\boldsymbol{P}$ ($\boldsymbol{P}^2 = \boldsymbol{P}$, but $\boldsymbol{P}' \neq \boldsymbol{P}$) it is possible to have $||\boldsymbol{Px}|| \geq ||\boldsymbol{x}||$. For example, let

$$\boldsymbol{P} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } \boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Then, $||\boldsymbol{Px}|| = 2$ and $||\boldsymbol{x}|| = \sqrt{2}$.

Theorem 2.23 *Let $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ denote the orthogonal projectors onto V_1 and V_2, respectively. Then, for an arbitrary $\boldsymbol{x} \in E^n$, the following relations hold:*

$$||\boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{x}|| \leq ||\boldsymbol{P}_1\boldsymbol{x}|| \leq ||\boldsymbol{x}|| \tag{2.62}$$

and, if $V_2 \subset V_1$,

$$||\boldsymbol{P}_2\boldsymbol{x}|| \leq ||\boldsymbol{P}_1\boldsymbol{x}||. \tag{2.63}$$

Proof. (2.62): Replace $\boldsymbol{x}$ by $\boldsymbol{P}_1\boldsymbol{x}$ in Theorem 2.22.

(2.63): By Theorem 2.11, we have $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2$, from which (2.63) follows immediately.

Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_p$ represent p n-component vectors in E^n, and define $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_p]$. From (1.15) and $\boldsymbol{P} = \boldsymbol{P}'\boldsymbol{P}$, the following identity holds:

$$||\boldsymbol{Px}_1||^2 + ||\boldsymbol{Px}_2||^2 + \cdots + ||\boldsymbol{Px}_p||^2 = \mathrm{tr}(\boldsymbol{X}'\boldsymbol{PX}). \tag{2.64}$$

The above identity and Theorem 2.23 lead to the following corollary.

Corollary

(i) *If* $V_2 \subset V_1$, $\text{tr}(\boldsymbol{X}'\boldsymbol{P}_2\boldsymbol{X}) \leq \text{tr}(\boldsymbol{X}'\boldsymbol{P}_1\boldsymbol{X}) \leq \text{tr}(\boldsymbol{X}'\boldsymbol{X})$.

(ii) *Let* $\boldsymbol{P}$ *denote an orthogonal projector onto an arbitrary subspace in* E^n. *If* $V_1 \supset V_2$,

$$\text{tr}(\boldsymbol{P}_1\boldsymbol{P}) \geq \text{tr}(\boldsymbol{P}_2\boldsymbol{P}).$$

Proof. (i): Obvious from Theorem 2.23. (ii): We have $\text{tr}(\boldsymbol{P}_j\boldsymbol{P}) = \text{tr}(\boldsymbol{P}_j\boldsymbol{P}^2)$ $= \text{tr}(\boldsymbol{P}\boldsymbol{P}_j\boldsymbol{P})$ $(j = 1, 2)$, and $(\boldsymbol{P}_1 - \boldsymbol{P}_2)^2 = \boldsymbol{P}_1 - \boldsymbol{P}_2$, so that

$$\text{tr}(\boldsymbol{P}\boldsymbol{P}_1\boldsymbol{P}) - \text{tr}(\boldsymbol{P}\boldsymbol{P}_2\boldsymbol{P}) = \text{tr}(\boldsymbol{S}\boldsymbol{S}') \geq 0,$$

where $\boldsymbol{S} = (\boldsymbol{P}_1 - \boldsymbol{P}_2)\boldsymbol{P}$. It follows that $\text{tr}(\boldsymbol{P}_1\boldsymbol{P}) \geq \text{tr}(\boldsymbol{P}_2\boldsymbol{P})$.

Q.E.D.

We next present a theorem on the trace of two orthogonal projectors.

Theorem 2.24 *Let* $\boldsymbol{P}_1$ *and* $\boldsymbol{P}_2$ *be orthogonal projectors of order* n. *Then the following relations hold:*

$$\text{tr}(\boldsymbol{P}_1\boldsymbol{P}_2) = \text{tr}(\boldsymbol{P}_2\boldsymbol{P}_1) \leq \min(\text{tr}(\boldsymbol{P}_1), \text{tr}(\boldsymbol{P}_2)). \tag{2.65}$$

Proof. We have $\text{tr}(\boldsymbol{P}_1) - \text{tr}(\boldsymbol{P}_1\boldsymbol{P}_2) = \text{tr}(\boldsymbol{P}_1(\boldsymbol{I}_n - \boldsymbol{P}_2)) = \text{tr}(\boldsymbol{P}_1\boldsymbol{Q}_2) = \text{tr}(\boldsymbol{P}_1\boldsymbol{Q}_2\boldsymbol{P}_1) = \text{tr}(\boldsymbol{S}'\boldsymbol{S}) \geq 0$, where $\boldsymbol{S} = \boldsymbol{Q}_2\boldsymbol{P}_1$, establishing $\text{tr}(\boldsymbol{P}_1) \geq \text{tr}(\boldsymbol{P}_1\boldsymbol{P}_2)$. Similarly, (2.65) follows from $\text{tr}(\boldsymbol{P}_2) \geq \text{tr}(\boldsymbol{P}_1\boldsymbol{P}_2) = \text{tr}(\boldsymbol{P}_2\boldsymbol{P}_1)$.

Q.E.D.

Note From (1.19), we obtain

$$\text{tr}(\boldsymbol{P}_1\boldsymbol{P}_2) \leq \sqrt{\text{tr}(\boldsymbol{P}_1)\text{tr}(\boldsymbol{P}_2)}. \tag{2.66}$$

However, (2.65) is more general than (2.66) because $\sqrt{\text{tr}(\boldsymbol{P}_1)\text{tr}(\boldsymbol{P}_2)} \geq \min(\text{tr}(\boldsymbol{P}_1), \text{tr}(\boldsymbol{P}_2))$.

2.5 Matrix Norm and Projection Matrices

Let $\boldsymbol{A} = [a_{ij}]$ be an n by p matrix. We define its Euclidean norm (also called the Frobenius norm) by

$$||\boldsymbol{A}|| = \{\mathrm{tr}(\boldsymbol{A}'\boldsymbol{A})\}^{1/2} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{p} a_{ij}^2}. \tag{2.67}$$

Then the following four relations hold.

Lemma 2.6

$$||\boldsymbol{A}|| \geq 0. \tag{2.68}$$

$$||\boldsymbol{C}\boldsymbol{A}|| \leq ||\boldsymbol{C}|| \cdot ||\boldsymbol{A}||, \tag{2.69}$$

Let both $\boldsymbol{A}$ and $\boldsymbol{B}$ be n by p matrices. Then,

$$||\boldsymbol{A} + \boldsymbol{B}|| \leq ||\boldsymbol{A}|| + ||\boldsymbol{B}||. \tag{2.70}$$

Let $\boldsymbol{U}$ and $\boldsymbol{V}$ be orthogonal matrices of orders n and p, respectively. Then

$$||\boldsymbol{U}\boldsymbol{A}\boldsymbol{V}|| = ||\boldsymbol{A}||. \tag{2.71}$$

Proof. Relations (2.68) and (2.69) are trivial. Relation (2.70) follows immediately from (1.20). Relation (2.71) is obvious from

$$\mathrm{tr}(\boldsymbol{V}'\boldsymbol{A}'\boldsymbol{U}'\boldsymbol{U}\boldsymbol{A}\boldsymbol{V}) = \mathrm{tr}(\boldsymbol{A}'\boldsymbol{A}\boldsymbol{V}\boldsymbol{V}') = \mathrm{tr}(\boldsymbol{A}'\boldsymbol{A}).$$

Q.E.D.

Note Let $\boldsymbol{M}$ be a symmetric *nnd* matrix of order n. Then the norm defined in (2.67) can be generalized as

$$||\boldsymbol{A}||_M = \{\mathrm{tr}(\boldsymbol{A}'\boldsymbol{M}\boldsymbol{A})\}^{1/2}. \tag{2.72}$$

This is called the norm of $\boldsymbol{A}$ with respect to $\boldsymbol{M}$ (sometimes called a metric matrix). Properties analogous to those given in Lemma 2.6 hold for this generalized norm.

There are other possible definitions of the norm of $\boldsymbol{A}$. For example,

(i) $||\boldsymbol{A}||_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$,
(ii) $||\boldsymbol{A}||_2 = \mu_1(\boldsymbol{A})$, where $\mu_1(\boldsymbol{A})$ is the largest singular value of $\boldsymbol{A}$ (see Chapter 5), and
(iii) $||\boldsymbol{A}||_3 = \max_i \sum_{j=1}^{p} |a_{ij}|$.

All of these norms satisfy (2.68), (2.69), and (2.70). (However, only $||\boldsymbol{A}||_2$ satisfies (2.71).)

Lemma 2.7 *Let* $\boldsymbol{P}$ *and* $\tilde{\boldsymbol{P}}$ *denote orthogonal projectors of orders* n *and* p*, respectively. Then,*

$$||\boldsymbol{PA}|| \leq ||\boldsymbol{A}|| \tag{2.73}$$

(the equality holds if and only if $\boldsymbol{PA} = \boldsymbol{A}$*) and*

$$||\boldsymbol{A}\tilde{\boldsymbol{P}}|| \leq ||\boldsymbol{A}|| \tag{2.74}$$

(the equality holds if and only if $\boldsymbol{A}\tilde{\boldsymbol{P}} = \boldsymbol{A}$*).*

Proof. (2.73): Square both sides and subtract the right-hand side from the left. Then,

$$\begin{aligned}\mathrm{tr}(\boldsymbol{A}'\boldsymbol{A}) - \mathrm{tr}(\boldsymbol{A}'\boldsymbol{PA}) &= \mathrm{tr}\{\boldsymbol{A}'(\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{A}\}\\ &= \mathrm{tr}(\boldsymbol{A}'\boldsymbol{QA}) = \mathrm{tr}(\boldsymbol{QA})'(\boldsymbol{QA}) \geq 0 \ \ (\text{where } \ \boldsymbol{Q} = \boldsymbol{I}_n - \boldsymbol{P}).\end{aligned}$$

The equality holds when $\boldsymbol{QA} = \boldsymbol{O} \Longleftrightarrow \boldsymbol{PA} = \boldsymbol{A}$.

(2.74): This can be proven similarly by noting that $||\boldsymbol{A}\tilde{\boldsymbol{P}}||^2 = \mathrm{tr}(\tilde{\boldsymbol{P}}\boldsymbol{A}'\boldsymbol{A}\tilde{\boldsymbol{P}}) = \mathrm{tr}(\boldsymbol{A}\tilde{\boldsymbol{P}}\boldsymbol{A}') = ||\tilde{\boldsymbol{P}}\boldsymbol{A}'||^2$. The equality holds when $\tilde{\boldsymbol{Q}}\boldsymbol{A}' = \boldsymbol{O} \Longleftrightarrow \tilde{\boldsymbol{P}}\boldsymbol{A}' = \boldsymbol{A}' \Longleftrightarrow \boldsymbol{A}\tilde{\boldsymbol{P}} = \boldsymbol{A}$, where $\tilde{\boldsymbol{Q}} = \boldsymbol{I}_n - \tilde{\boldsymbol{P}}$. Q.E.D.

The two lemmas above lead to the following theorem.

Theorem 2.25 *Let* $\boldsymbol{A}$ *be an* n *by* p *matrix,* $\boldsymbol{B}$ *and* $\boldsymbol{Y}$ n *by* r *matrices, and* $\boldsymbol{C}$ *and* $\boldsymbol{X}$ r *by* p *matrices. Then,*

$$||\boldsymbol{A} - \boldsymbol{BX}|| \geq ||(\boldsymbol{I}_n - \boldsymbol{P}_B)\boldsymbol{A}||, \tag{2.75}$$

where $\boldsymbol{P}_B$ *is the orthogonal projector onto* $\mathrm{Sp}(\boldsymbol{B})$*. The equality holds if and only if* $\boldsymbol{BX} = \boldsymbol{P}_B\boldsymbol{A}$*. We also have*

$$||\boldsymbol{A} - \boldsymbol{YC}|| \geq ||\boldsymbol{A}(\boldsymbol{I}_p - \boldsymbol{P}_{C'})||, \tag{2.76}$$

where $\boldsymbol{P}_{C'}$ *is the orthogonal projector onto* $\mathrm{Sp}(\boldsymbol{C}')$*. The equality holds if and only if* $\boldsymbol{YC} = \boldsymbol{AP}_{C'}$*. We also have*

$$||\boldsymbol{A} - \boldsymbol{BX} - \boldsymbol{YC}|| \geq ||(\boldsymbol{I}_n - \boldsymbol{P}_B)\boldsymbol{A}(\boldsymbol{I}_p - \boldsymbol{P}_{C'})||. \tag{2.77}$$

The equality holds if and only if

$$\boldsymbol{P}_B(\boldsymbol{A} - \boldsymbol{YC}) = \boldsymbol{BX} \text{ and } (\boldsymbol{I}_n - \boldsymbol{P}_B)\boldsymbol{AP}_{C'} = (\boldsymbol{I}_n - \boldsymbol{P}_B)\boldsymbol{YC} \tag{2.78}$$

or

$$(\boldsymbol{A} - \boldsymbol{BX})\boldsymbol{P}_{C'} = \boldsymbol{YC} \text{ and } \boldsymbol{P}_B\boldsymbol{A}(\boldsymbol{I}_p - \boldsymbol{P}_{C'}) = \boldsymbol{BX}(\boldsymbol{I}_n - \boldsymbol{P}_{C'}). \quad (2.79)$$

Proof. (2.75): We have $(\boldsymbol{I}_n - \boldsymbol{P}_B)(\boldsymbol{A} - \boldsymbol{BX}) = \boldsymbol{A} - \boldsymbol{BX} - \boldsymbol{P}_B\boldsymbol{A} + \boldsymbol{BX} = (\boldsymbol{I}_n - \boldsymbol{P}_B)\boldsymbol{A}$. Since $\boldsymbol{I}_n - \boldsymbol{P}_B$ is an orthogonal projector, we have $||\boldsymbol{A}-\boldsymbol{BX}|| \geq ||(\boldsymbol{I}_n-\boldsymbol{P}_B)(\boldsymbol{A}-\boldsymbol{BX})|| = ||(\boldsymbol{I}_n-\boldsymbol{P}_B)\boldsymbol{A}||$ by (2.73) in Lemma 2.7. The equality holds when $(\boldsymbol{I}_n - \boldsymbol{P}_B)(\boldsymbol{A} - \boldsymbol{BX}) = \boldsymbol{A} - \boldsymbol{BX}$, namely $\boldsymbol{P}_B\boldsymbol{A} = \boldsymbol{BX}$.

(2.76): It suffices to use $(\boldsymbol{A}-\boldsymbol{YC})(\boldsymbol{I}_p-\boldsymbol{P}_{C'}) = \boldsymbol{A}(\boldsymbol{I}_p-\boldsymbol{P}_{C'})$ and (2.74) in Lemma 2.7. The equality holds when $(\boldsymbol{A} - \boldsymbol{YC})(\boldsymbol{I}_p - \boldsymbol{P}_{C'}) = \boldsymbol{A} - \boldsymbol{YC}$ holds, which implies $\boldsymbol{YC} = \boldsymbol{AP}_{C'}$.

(2.77): $||\boldsymbol{A}-\boldsymbol{BX}-\boldsymbol{YC}|| \geq ||(\boldsymbol{I}_n-\boldsymbol{P}_B)(\boldsymbol{A}-\boldsymbol{YC})|| \geq ||(\boldsymbol{I}_n-\boldsymbol{P}_B)\boldsymbol{A}(\boldsymbol{I}_p-\boldsymbol{P}_{C'})||$ or $||\boldsymbol{A}-\boldsymbol{BX}-\boldsymbol{YC}|| \geq ||(\boldsymbol{A}-\boldsymbol{BX})(\boldsymbol{I}_p-\boldsymbol{P}_{C'})|| \geq ||(\boldsymbol{I}_p-\boldsymbol{P}_B)\boldsymbol{A}(\boldsymbol{I}_p-\boldsymbol{P}_{C'})||$. The first equality condition (2.78) follows from the first relation above, and the second equality condition (2.79) follows from the second relation above. Q.E.D.

Note Relations (2.75), (2.76), and (2.77) can also be shown by the least squares method. Here we show this only for (2.77). We have

$$\begin{aligned}||\boldsymbol{A} - \boldsymbol{BX} - \boldsymbol{YC}||^2 &= \mathrm{tr}\{(\boldsymbol{A} - \boldsymbol{BX} - \boldsymbol{YC})'(\boldsymbol{A} - \boldsymbol{BX} - \boldsymbol{YC})\} \\ &= \mathrm{tr}(\boldsymbol{A} - \boldsymbol{YC})'(\boldsymbol{A} - \boldsymbol{YC}) - 2\mathrm{tr}(\boldsymbol{BX})'(\boldsymbol{A} - \boldsymbol{YC}) + \mathrm{tr}(\boldsymbol{BX})'(\boldsymbol{BX})\end{aligned}$$

to be minimized. Differentiating the criterion above by $\boldsymbol{X}$ and setting the result to zero, we obtain $\boldsymbol{B}'(\boldsymbol{A} - \boldsymbol{YC}) = \boldsymbol{B}'\boldsymbol{BX}$. Premultiplying this equation by $\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^{-1}$, we obtain $\boldsymbol{P}_B(\boldsymbol{A} - \boldsymbol{YC}) = \boldsymbol{BX}$. Furthermore, we may expand the criterion above as

$$\mathrm{tr}(\boldsymbol{A} - \boldsymbol{BX})'(\boldsymbol{A} - \boldsymbol{BX}) - 2\mathrm{tr}(\boldsymbol{YC}(\boldsymbol{A} - \boldsymbol{BX})') + \mathrm{tr}(\boldsymbol{YC})(\boldsymbol{YC})'.$$

Differentiating this criterion with respect to $\boldsymbol{Y}$ and setting the result equal to zero, we obtain $\boldsymbol{C}(\boldsymbol{A} - \boldsymbol{BX}) = \boldsymbol{CC}'\boldsymbol{Y}'$ or $(\boldsymbol{A} - \boldsymbol{BX})\boldsymbol{C}' = \boldsymbol{YCC}'$. Postmultiplying the latter by $(\boldsymbol{CC}')^{-1}\boldsymbol{C}'$, we obtain $(\boldsymbol{A} - \boldsymbol{BX})\boldsymbol{P}_{C'} = \boldsymbol{YC}$. Substituting this into $\boldsymbol{P}_B(\boldsymbol{A} - \boldsymbol{YC}) = \boldsymbol{BX}$, we obtain $\boldsymbol{P}_B\boldsymbol{A}(\boldsymbol{I}_p - \boldsymbol{P}_{C'}) = \boldsymbol{BX}(\boldsymbol{I}_p - \boldsymbol{P}_{C'})$ after some simplification. If, on the other hand, $\boldsymbol{BX} = \boldsymbol{P}_B(\boldsymbol{A} - \boldsymbol{YC})$ is substituted into $(\boldsymbol{A} - \boldsymbol{BX})\boldsymbol{P}_{C'} = \boldsymbol{YC}$, we obtain $(\boldsymbol{I}_n - \boldsymbol{P}_B)\boldsymbol{AP}_{C'} = (\boldsymbol{I}_n - \boldsymbol{P}_B)\boldsymbol{YC}$. (In the derivation above, the regular inverses can be replaced by the respective generalized inverses. See the next chapter.)

2.6 General Form of Projection Matrices

The projectors we have been discussing so far are based on Definition 2.1, namely square matrices that satisfy $\boldsymbol{P}^2 = \boldsymbol{P}$ (idempotency). In this section, we introduce a generalized form of projection matrices that do not necessarily satisfy $\boldsymbol{P}^2 = \boldsymbol{P}$, based on Rao (1974) and Rao and Yanai (1979).

Definition 2.3 *Let $V \subset E^n$ (but $V \neq E^n$) be decomposed as a direct-sum of m subspaces, namely $V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$. A square matrix $\boldsymbol{P}_j^*$ of order n that maps an arbitrary vector $\boldsymbol{y}$ in V into V_j is called the projection matrix onto V_j along $V_{(j)} = V_1 \oplus \cdots \oplus V_{j-1} \oplus V_{j+1} \oplus \cdots \oplus V_m$ if and only if*

$$\boldsymbol{P}_j^*\boldsymbol{x} = \boldsymbol{x} \quad \forall \boldsymbol{x} \in V_j \ \ (j = 1, \cdots, m) \tag{2.80}$$

and

$$\boldsymbol{P}_j^*\boldsymbol{x} = \boldsymbol{0} \quad \forall \boldsymbol{x} \in V_{(j)} \ \ (j = 1, \cdots, m). \tag{2.81}$$

Let $\boldsymbol{x}_j \in V_j$. Then any $\boldsymbol{x} \in V$ can be expressed as

$$\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_m = (\boldsymbol{P}_1^* + \boldsymbol{P}_2^* + \cdots \boldsymbol{P}_m^*)\boldsymbol{x}.$$

Premultiplying the equation above by $\boldsymbol{P}_j^*$, we obtain

$$\boldsymbol{P}_i^*\boldsymbol{P}_j^*\boldsymbol{x} = \boldsymbol{0} \ \ (i \neq j) \ \text{ and } \ (\boldsymbol{P}_j^*)^2\boldsymbol{x} = \boldsymbol{P}_j^*\boldsymbol{x} \ \ (i = 1, \cdots, m) \tag{2.82}$$

since $\text{Sp}(\boldsymbol{P}_1), \text{Sp}(\boldsymbol{P}_2), \cdots, \text{Sp}(\boldsymbol{P}_m)$ are mutually disjoint. However, $\boldsymbol{V}$ does not cover the entire E^n ($\boldsymbol{x} \in V \neq E^n$), so (2.82) does not imply $(\boldsymbol{P}_j^*)^2 = \boldsymbol{P}_j^*$ or $\boldsymbol{P}_i^*\boldsymbol{P}_j^* = \boldsymbol{O}$ $(i \neq j)$.

Let V_1 and $V_2 \in E^3$ denote the subspaces spanned by $\boldsymbol{e}_1 = (0, 0, 1)'$ and $\boldsymbol{e}_2 = (0, 1, 0)'$, respectively. Suppose

$$\boldsymbol{P}^* = \begin{bmatrix} a & 0 & 0 \\ b & 0 & 0 \\ c & 0 & 1 \end{bmatrix}.$$

Then, $\boldsymbol{P}^*\boldsymbol{e}_1 = \boldsymbol{e}_1$ and $\boldsymbol{P}^*\boldsymbol{e}_2 = \boldsymbol{0}$, so that $\boldsymbol{P}^*$ is the projector onto V_1 along V_2 according to Definition 2.3. However, $(\boldsymbol{P}^*)^2 \neq \boldsymbol{P}^*$ except when $a = b = 0$, or $a = 1$ and $c = 0$. That is, when V does not cover the entire space E^n, the projector $\boldsymbol{P}_j^*$ in the sense of Definition 2.3 is not idempotent. However, by specifying a complement subspace of V, we can construct an idempotent matrix from $\boldsymbol{P}_j^*$ as follows.

Theorem 2.26 *Let $\boldsymbol{P}_j^*$ $(j = 1, \cdots, m)$ denote the projector in the sense of Definition 2.3, and let $\boldsymbol{P}$ denote the projector onto V along V_{m+1}, where $V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$ is a subspace in E^n and where V_{m+1} is a complement subspace to V. Then,*

$$\boldsymbol{P}_j = \boldsymbol{P}_j^* \boldsymbol{P} \ \ (j = 1, \cdots m) \ \ \textit{and} \ \ \boldsymbol{P}_{m+1} = \boldsymbol{I}_n - \boldsymbol{P} \tag{2.83}$$

are projectors (in the sense of Definition 2.1) onto V_j $(j = 1, \cdots, m+1)$ along $V_{(j)}^* = V_1 \oplus \cdots \oplus V_{j-1} \oplus V_{j+1} \oplus \cdots \oplus V_m \oplus V_{m+1}$.

Proof. Let $\boldsymbol{x} \in V$. If $\boldsymbol{x} \in V_j$ $(j = 1, \cdots, m)$, we have $\boldsymbol{P}_j^* \boldsymbol{P} \boldsymbol{x} = \boldsymbol{P}_j^* \boldsymbol{x} = \boldsymbol{x}$. On the other hand, if $\boldsymbol{x} \in V_i$ $(i \neq j,\ i = 1, \cdots, m)$, we have $\boldsymbol{P}_j^* \boldsymbol{P} \boldsymbol{x} = \boldsymbol{P}_j^* \boldsymbol{x} = \boldsymbol{0}$. Furthermore, if $\boldsymbol{x} \in V_{m+1}$, we have $\boldsymbol{P}_j^* \boldsymbol{P} \boldsymbol{x} = \boldsymbol{0}$ $(j = 1, \cdots, m)$. On the other hand, if $\boldsymbol{x} \in V$, we have $\boldsymbol{P}_{m+1} \boldsymbol{x} = (\boldsymbol{I}_n - \boldsymbol{P}) \boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x} = \boldsymbol{0}$, and if $\boldsymbol{x} \in V_{m+1}$, $\boldsymbol{P}_{m+1} \boldsymbol{x} = (\boldsymbol{I}_n - \boldsymbol{P}) \boldsymbol{x} = \boldsymbol{x} - \boldsymbol{0} = \boldsymbol{x}$. Hence, by Theorem 2.2, $\boldsymbol{P}_j$ $(j = 1, \cdots, m+1)$ is the projector onto V_j along $V_{(j)}$. Q.E.D.

2.7 Exercises for Chapter 2

1. Let $\tilde{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_2 \end{bmatrix}$ and $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix}$. Show that $\boldsymbol{P}_{\tilde{A}} \boldsymbol{P}_A = \boldsymbol{P}_A$.

2. Let $\boldsymbol{P}_A$ and $\boldsymbol{P}_B$ denote the orthogonal projectors onto $\text{Sp}(\boldsymbol{A})$ and $\text{Sp}(\boldsymbol{B})$, respectively. Show that the necessary and sufficient condition for $\text{Sp}(\boldsymbol{A}) = \{\text{Sp}(\boldsymbol{A}) \cap \text{Sp}(\boldsymbol{B})\} \dot{\oplus} \{\text{Sp}(\boldsymbol{A}) \cap \text{Sp}(\boldsymbol{B})^\perp\}$ is $\boldsymbol{P}_A \boldsymbol{P}_B = \boldsymbol{P}_B \boldsymbol{P}_A$.

3. Let $\boldsymbol{P}$ be a square matrix of order n such that $\boldsymbol{P}^2 = \boldsymbol{P}$, and suppose

$$||\boldsymbol{P}\boldsymbol{x}|| = ||\boldsymbol{x}||$$

for any n-component vector $\boldsymbol{x}$. Show the following:
(i) When $\boldsymbol{x} \in (\text{Ker}(\boldsymbol{P}))^\perp$, $\boldsymbol{P}\boldsymbol{x} = \boldsymbol{x}$.
(ii) $\boldsymbol{P}' = \boldsymbol{P}$.

4. Let $\text{Sp}(\boldsymbol{A}) = \text{Sp}(\boldsymbol{A}_1) \dot{\oplus} \cdots \dot{\oplus} \text{Sp}(\boldsymbol{A}_m)$, and let $\boldsymbol{P}_j$ $(j = 1, \cdots, m)$ denote the projector onto $\text{Sp}(\boldsymbol{A}_j)$. For $\forall \boldsymbol{x} \in E^n$:
(i) Show that

$$||\boldsymbol{x}||^2 \geq ||\boldsymbol{P}_1 \boldsymbol{x}||^2 + ||\boldsymbol{P}_2 \boldsymbol{x}||^2 + \cdots + ||\boldsymbol{P}_m \boldsymbol{x}||^2. \tag{2.84}$$

(Also, show that the equality holds if and only if $\text{Sp}(\boldsymbol{A}) = E^n$.)
(ii) Show that $\text{Sp}(\boldsymbol{A}_i)$ and $\text{Sp}(\boldsymbol{A}_j)$ $(i \neq j)$ are orthogonal if $\text{Sp}(\boldsymbol{A}) = \text{Sp}(\boldsymbol{A}_1) \oplus \text{Sp}(\boldsymbol{A}_2) \oplus \cdots \oplus \text{Sp}(\boldsymbol{A}_m)$ and the inequality in (i) above holds.
(iii) Let $\boldsymbol{P}_{[j]} = \boldsymbol{P}_1 + \boldsymbol{P}_2 + \cdots + \boldsymbol{P}_j$. Show that

$$||\boldsymbol{P}_{[m]} \boldsymbol{x}|| \geq ||\boldsymbol{P}_{[m-1]} \boldsymbol{x}|| \geq \cdots \geq ||\boldsymbol{P}_{[2]} \boldsymbol{x}|| \geq ||\boldsymbol{P}_{[1]} \boldsymbol{x}||.$$

5. Let $E^n = V_1 \oplus W_1 = V_2 \oplus W_2 = V_3 \oplus W_3$, and let $\boldsymbol{P}_j$ denote the projector onto V_j $(j = 1, 2, 3)$ along W_j. Show the following:
(i) Let $\boldsymbol{P}_i \boldsymbol{P}_j = \boldsymbol{O}$ for $i \neq j$. Then, $\boldsymbol{P}_1 + \boldsymbol{P}_2 + \boldsymbol{P}_3$ is the projector onto $V_1 + V_2 + V_3$ along $W_1 \cap W_2 \cap W_3$.
(ii) Let $\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_2\boldsymbol{P}_1$, $\boldsymbol{P}_1\boldsymbol{P}_3 = \boldsymbol{P}_3\boldsymbol{P}_1$, and $\boldsymbol{P}_2\boldsymbol{P}_3 = \boldsymbol{P}_3\boldsymbol{P}_2$. Then $\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3$ is the projector onto $V_1 \cap V_2 \cap V_3$ along $W_1 + W_2 + W_3$.
(iii) Suppose that the three identities in (ii) hold, and let $\boldsymbol{P}_{1+2+3}$ denote the projection matrix onto $V_1 + V_2 + V_3$ along $W_1 \cap W_2 \cap W_3$. Show that

$$\boldsymbol{P}_{1+2+3} = \boldsymbol{P}_1 + \boldsymbol{P}_2 + \boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_2 - \boldsymbol{P}_2\boldsymbol{P}_3 - \boldsymbol{P}_1\boldsymbol{P}_3 + \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3.$$

6. Show that

$$\boldsymbol{Q}_{[A,B]} = \boldsymbol{Q}_A \boldsymbol{Q}_{Q_A B},$$

where $\boldsymbol{Q}_{[A,B]}$, $\boldsymbol{Q}_A$, and $\boldsymbol{Q}_{Q_A B}$ are the orthogonal projectors onto the null space of $[\boldsymbol{A}, \boldsymbol{B}]$, onto the null space of $\boldsymbol{A}$, and onto the null space of $\boldsymbol{Q}_A\boldsymbol{B}$, respectively.

7. (a) Show that

$$\boldsymbol{P}_X = \boldsymbol{P}_{XA} + \boldsymbol{P}_{X(X'X)^{-1}B},$$

where $\boldsymbol{P}_X$, $\boldsymbol{P}_{XA}$, and $\boldsymbol{P}_{X(X'X)^{-1}B}$ are the orthogonal projectors onto Sp$(\boldsymbol{X})$, Sp$(\boldsymbol{X}\boldsymbol{A})$, and Sp$(\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{B})$, respectively, and $\boldsymbol{A}$ and $\boldsymbol{B}$ are such that Ker$(\boldsymbol{A}')$ $=$ Sp$(\boldsymbol{B})$.
(b) Use the decomposition above to show that

$$\boldsymbol{P}_{[X_1,X_2]} = \boldsymbol{P}_{X_1} + \boldsymbol{P}_{Q_{X_1}X_2},$$

where $\boldsymbol{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2]$, $\boldsymbol{P}_{Q_{X_1}X_2}$ is the orthogonal projector onto Sp$(\boldsymbol{Q}_{X_1}\boldsymbol{X}_2)$, and $\boldsymbol{Q}_{X_1} = \boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'$.

8. Let $E^n = V_1 \oplus W_1 = V_2 \oplus W_2$, and let $\boldsymbol{P}_1 = \boldsymbol{P}_{V_1 \cdot W_1}$ and $\boldsymbol{P}_2 = \boldsymbol{P}_{V_2 \cdot W_2}$ be two projectors (not necessarily orthogonal) of the same size. Show the following:
(a) The necessary and sufficient condition for $\boldsymbol{P}_1\boldsymbol{P}_2$ to be a projector is $V_{12} \subset V_2 \oplus (W_1 \cap W_2)$, where $V_{12} = $ Sp$(\boldsymbol{P}_1\boldsymbol{P}_2)$ (Brown and Page, 1970).
(b) The condition in (a) is equivalent to $V_2 \subset V_1 \oplus (W_1 \cap V_2) \oplus (W_1 \cap W_2)$ (Werner, 1992).

9. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be n by a $(n \geq a)$ and n by b $(n \geq b)$ matrices, respectively. Let $\boldsymbol{P}_A$ and $\boldsymbol{P}_B$ be the orthogonal projectors defined by $\boldsymbol{A}$ and $\boldsymbol{B}$, and let $\boldsymbol{Q}_A$ and $\boldsymbol{Q}_B$ be their orthogonal complements. Show that the following six statements are equivalent: (1) $\boldsymbol{P}_A\boldsymbol{P}_B = \boldsymbol{P}_B\boldsymbol{P}_A$, (2) $\boldsymbol{A}'\boldsymbol{B} = \boldsymbol{A}'\boldsymbol{P}_B\boldsymbol{P}_A\boldsymbol{B}$, (3) $(\boldsymbol{P}_A\boldsymbol{P}_B)^2 = \boldsymbol{P}_A\boldsymbol{P}_B$, (4) $\boldsymbol{P}_{[A,B]} = \boldsymbol{P}_A + \boldsymbol{P}_B - \boldsymbol{P}_A\boldsymbol{P}_B$, (5) $\boldsymbol{A}'\boldsymbol{Q}_B\boldsymbol{Q}_A\boldsymbol{B} = \boldsymbol{O}$, and (6) rank$(\boldsymbol{Q}_A\boldsymbol{B}) = $ rank$(\boldsymbol{B}) - $ rank$(\boldsymbol{A}'\boldsymbol{B})$.

http://www.springer.com/978-1-4419-9886-6

Projection Matrices, Generalized Inverse Matrices, and
Singular Value Decomposition
Yanai, H.; Takeuchi, K.; Takane, Y.
2011, XII, 236 p., Hardcover
ISBN: 978-1-4419-9886-6