Orthogonality

Definition 1

If a matrix Y in a linear space of matrices is orthogonal to every matrix in a subspace $\mathcal{U}$, Y is said to be *orthogonal* to $\mathcal{U}$.

The statement that Y is orthogonal to $\mathcal{U}$ is sometimes abbreviated to $\mathrm{Y} \perp \mathcal{U}$.

Definition 2

Similarly, to indicate that every matrix in a subspace $\mathcal{U}$ is orthogonal to every matrix in a subspace $\mathcal{W}$, one says that $\mathcal{U}$ is orthogonal to $\mathcal{W}$ or writes $\mathcal{U} \perp \mathcal{W}$.

Orthogonality

Lemma 3

- *Let* Y *represent a matrix in a linear space* $\mathcal{V}$, *let* $\mathcal{U}$ *and* $\mathcal{W}$ *represent subspaces of* $\mathcal{V}$, *and take* $\{\mathrm{X}_1, \ldots, \mathrm{X}_s\}$ *to be a set of matrices that spans* $\mathcal{U}$ *and* $\{\mathrm{Z}_1, \ldots, \mathrm{Z}_t\}$ *to be a set that spans* $\mathcal{W}$. *Then,* $\mathrm{Y} \perp \mathcal{U}$ *if and only if* $\mathrm{Y} \cdot \mathrm{X}_i = 0$ *for* $i = 1, \ldots, s$; *that is,* Y *is orthogonal to* $\mathcal{U}$ *if and only if* Y *is orthogonal to each of the matrices* $\mathrm{X}_1, \ldots, \mathrm{X}_s$.
- *And, similarly,* $\mathcal{U} \perp \mathcal{W}$ *if and only if* $\mathrm{X}_i \cdot \mathrm{Z}_j = 0$ *for* $i = 1, \ldots, s$ *and* $j = 1, \ldots, t$; *that is,* $\mathcal{U}$ *is orthogonal to* $\mathcal{W}$ *if and only if each of the matrices* $\mathrm{X}_1, \ldots, \mathrm{X}_s$ *is orthogonal to each of the matrices* $\mathrm{Z}_1, \ldots, \mathrm{Z}_t$.

Orthogonality

By applying Lemma (3) in the special case where $\mathcal{V} = \mathcal{R}^{m\times 1}$ and $\mathcal{U}$ and $\mathcal{W}$ are the column spaces of two matrices (each of which has m rows), we obtain the following corollary.

Corollary 4

Let $\mathbf{y}$ represent an m–dimensional column vector, and let X represent an $m \times n$ matrix and Z an $m \times p$ matrix. Then, $\mathbf{y}$ is orthogonal to $\mathcal{C}(\mathrm{X})$ if and only if $\mathrm{X}'\mathbf{y} = 0$. Similarly, $\mathcal{C}(\mathrm{X})$ is orthogonal to $\mathcal{C}(\mathrm{Z})$ if and only if $\mathrm{X}'\mathrm{Z} = 0$.

Orthogonality

Theorem 5

Let Y *represent a matrix in a linear space* $\mathcal{V}$*, and let* $\mathcal{U}$ *represent an* r*–dimensional subspace of* $\mathcal{V}$*. Then there exists a unique matrix* Z *in* $\mathcal{U}$ *such that* $(\mathrm{Y}-\mathrm{Z}) \perp \mathcal{U}$*, that is, such that the difference between* Y *and* Z *is orthogonal to every matrix in* $\mathcal{U}$*. If* $r=0$*, then* $\mathrm{Z}=0$*, and, if* $r>0$*,* Z *is expressible as*

$$\mathrm{Z} = c_1\mathrm{X}_1 + \cdots + c_r\mathrm{X}_r, \tag{1}$$

where $\{\mathrm{X}_1, \ldots, \mathrm{X}_r\}$ *is any orthonormal basis for* $\mathcal{U}$ *and* $c_j = \mathrm{Y}\cdot\mathrm{X}_j$ $(j=1,\ldots,r)$*. Moreover,* $\mathrm{Z}=\mathrm{Y}$ *if and only if* $\mathrm{Y}\in\mathcal{U}$*.*

Orthogonality

Proof: Consider first the case where $r = 0$. In this case, the only matrix in $\mathcal{U}$ is the null matrix 0. Clearly, $\mathrm{Y} - 0$ is orthogonal to 0. Thus, there exists a unique matrix Z in $\mathcal{U}$ such that $(\mathrm{Y} - \mathrm{Z}) \perp \mathcal{U}$, namely, $\mathrm{Z} = 0$. Moreover, it is clear that $\mathrm{Y} = \mathrm{Z}$ if and only if $\mathrm{Y} \in \mathcal{U}$.

Orthogonality

Consider now the case where $r > 0$. Take $\{\mathrm{X}_1, \ldots, \mathrm{X}_r\}$ to be any orthonormal basis for $\mathcal{U}$, and define $c_j = \mathrm{Y} \cdot \mathrm{X}_j$ ($j = 1, \ldots, r$). Clearly, $\sum_j c_j \mathrm{X}_j \in \mathcal{U}$, and

$$(\mathrm{Y} - \sum_j c_j \mathrm{X}_j) \cdot \mathrm{X}_i = (\mathrm{Y} \cdot \mathrm{X}_i) - c_i = 0$$

for $i = 1, \ldots, r$, implying (from Lemma 3) that $(\mathrm{Y} - \sum_j c_j \mathrm{X}_j) \perp \mathcal{U}$.

Orthogonality

Moreover, for any matrix X such that $\mathrm{X} \in \mathcal{U}$ and $(\mathrm{Y}-\mathrm{X}) \perp \mathcal{U}$, we find that $(\mathrm{X}-\sum_j c_j \mathrm{X}_j) \in \mathcal{U}$ and hence

$$\begin{aligned}(\mathrm{X}-\sum_j c_j\mathrm{X}_j)\cdot(\mathrm{X}-\sum_j c_j\mathrm{X}_j) &= (\mathrm{Y}-\sum_j c_j\mathrm{X}_j)\cdot(\mathrm{X}-\sum_j c_j\mathrm{X}_j) \\ &\quad -(\mathrm{Y}-\mathrm{X})\cdot(\mathrm{X}-\sum_j c_j\mathrm{X}_j) \\ &= 0-0=0,\end{aligned}$$

so that $(\mathrm{X}-\sum_j c_j\mathrm{X}_j) = 0$ or equivalently $\mathrm{X} = \sum_j c_j\mathrm{X}_j$.

Orthogonality

We conclude that there exists a unique matrix Z in $\mathcal{U}$ such that $(\mathrm{Y} - \mathrm{Z}) \perp \mathcal{U}$, namely, $\mathrm{Z} = \sum_j c_j \mathrm{X}_j$. To complete the proof, observe that, if $\mathrm{Y} = \mathrm{Z}$, then obviously, $\mathrm{Y} \in \mathcal{U}$, and conversely, if $\mathrm{Y} \in \mathcal{U}$, then, since $\mathrm{Y} - \mathrm{Y}$ is orthogonal to $\mathcal{U}$, $\mathrm{Y} = \mathrm{Z}$.

Projection

Definition 6

Suppose that $\mathcal{U}$ is a subspace of a linear space $\mathcal{V}$. Then, it follows from Theorem 5 that, corresponding to each matrix Y in $\mathcal{V}$, there exists a unique matrix Z in $\mathcal{U}$ such that $\mathrm{Y} - \mathrm{Z}$ is orthogonal to $\mathcal{U}$. The matrix Z is called the *orthogonal projection* of Y on $\mathcal{U}$ or simply the *projection* of Y on $\mathcal{U}$.

Projection

Theorem 7

Let $\mathrm{Y}_1, \ldots, \mathrm{Y}_p$ represent matrices in a linear space $\mathcal{V}$, let $\mathcal{U}$ represent a subspace of $\mathcal{V}$, and let $\mathrm{Z}_1, \ldots, \mathrm{Z}_p$ represent the projections of $\mathrm{Y}_1, \ldots, \mathrm{Y}_p$, respectively, on $\mathcal{U}$. Then, for any scalars $k_1, \ldots, k_p$, the projection of the linear combination $k_1\mathrm{Y}_1 + \cdots + k_p\mathrm{Y}_p$ (on $\mathcal{U}$) is the corresponding linear combination $k_1\mathrm{Z}_1 + \cdots + k_p\mathrm{Z}_p$ of $\mathrm{Z}_1, \ldots, \mathrm{Z}_p$.

Projection

Proof: By definition, $\mathrm{Z}_i \in \mathcal{U}$ and $(\mathrm{Y}_i - \mathrm{Z}_i) \perp \mathcal{U}$ $(i = 1, \ldots, p)$. Thus, $(k_1\mathrm{Z}_1 + \cdots + k_p\mathrm{Z}_p) \in \mathcal{U}$. Moreover, for every matrix X in $\mathcal{U}$,

$$\begin{aligned}
&[k_1\mathrm{Y}_1 + \cdots + k_p\mathrm{Y}_p - (k_1\mathrm{Z}_1 + \cdots + k_p\mathrm{Z}_p)] \cdot \mathrm{X} \\
&= [k_1(\mathrm{Y}_1 - \mathrm{Z}_1) + \cdots + k_p(\mathrm{Y}_p - \mathrm{Z}_p)] \cdot \mathrm{X} \\
&= k_1[(\mathrm{Y}_1 - \mathrm{Z}_1) \cdot \mathrm{X}] + \cdots + k_p[(\mathrm{Y}_p - \mathrm{Z}_p) \cdot \mathrm{X}] \\
&= k_1 0 + \cdots + k_p 0 = 0.
\end{aligned}$$

Projection

Proof: So that $[k_1 \mathrm{Y}_1 + \cdots + k_p \mathrm{Y}_p - (k_1 \mathrm{Z}_1 + \cdots + k_p \mathrm{Z}_p)] \perp \mathcal{U}$. We conclude that $k_1 \mathrm{Z}_1 + \cdots + k_p \mathrm{Z}_p$ is the projection of $k_1 \mathrm{Y}_1 + \cdots + k_p \mathrm{Y}_p$ on $\mathcal{U}$.

Projection

Theorem 8

Let z represent the projection (with respect to the usual inner product) of an n–dimensional column vector $\mathbf{y}$ *on a subspace* $\mathcal{U}$ *of* $\mathbb{R}^n$, *and let* X *represent any* $n \times p$ *matrix whose columns span* $\mathcal{U}$. *Then,*

$$\mathbf{z} = \mathrm{X}\mathbf{b}^*$$

for any solution $\mathbf{b}^*$ *to the linear system*

$$\mathrm{X}'\mathrm{X}\mathbf{b} = \mathrm{X}'\mathbf{y} \ (\textit{in } \mathbf{b}).$$

Projection

Proof: Suppose that $\mathbf{b}^*$ is a solution to the linear system $\mathrm{X}'\mathrm{X}\mathbf{b} = \mathrm{X}'\mathbf{y}$. Then $\mathrm{X}'(\mathbf{y} - \mathrm{X}\mathbf{b}^*) = 0$, implying that $\mathbf{y} - \mathrm{X}\mathbf{b}^*$ is orthogonal to $\mathcal{C}(\mathrm{X})$ and hence [since $\mathcal{C}(\mathrm{X}) = \mathcal{U}$] to $\mathcal{U}$. Since $\mathrm{X}\mathbf{b}^* \in \mathcal{U}$, we conclude that $\mathrm{X}\mathbf{b}^*$ is the projection of $\mathbf{y}$ on $\mathcal{U}$ and hence, by definition, that $\mathsf{z} = \mathrm{X}\mathbf{b}^*$.

Projection

Corollary 9

Let z represent the projection (with respect to the usual inner product) of an n–dimensional column vector $\mathbf{y}$ *on a subspace* $\mathcal{U}$ *of* $\mathbb{R}^n$*, and let* X *represent any* $n \times p$ *matrix whose columns span* $\mathcal{U}$*. Then,*

$$\mathrm{z} = \mathrm{X}(\mathrm{X}'\mathrm{X})^{-}\mathrm{X}'\mathbf{y}.$$

Projection

Corollary 10

Let $\mathbf{y}$ represent an n–dimensional column vector, X an $n \times p$ matrix, and W any $n \times q$ matrix such that $\mathcal{C}(\mathrm{W}) = \mathcal{C}(\mathrm{X})$. Then,

$$\mathrm{W}\mathbf{a}^* = \mathrm{X}\mathbf{b}^*$$

for any solution $\mathbf{a}^*$ to the linear system $\mathrm{W}'\mathrm{W}\mathbf{a} = \mathrm{W}'\mathbf{y}$ (in $\mathbf{a}$) and any solution $\mathbf{b}^*$ to the linear system $\mathrm{X}'\mathrm{X}\mathbf{b} = \mathrm{X}'\mathbf{y}$ (in $\mathbf{b}$).

Projection

Corollary 11

Let $\mathbf{y}$ *represent an* n*–dimensional column vector, and* X *an* $n \times p$ *matrix. Then,* $\mathrm{X}\mathbf{b}_1 = \mathrm{X}\mathbf{b}_2$ *for any two solutions* $\mathbf{b}_1$ *and* $\mathbf{b}_2$ *to the linear system* $\mathrm{X}'\mathrm{X}\mathbf{b} = \mathrm{X}'\mathbf{y}$ *(in* $\mathbf{b}$*).*

Projection

The following theorem gives a converse of Theorem 8.

Theorem 12

Let z represent the projection (with respect to the usual inner product) of an n–dimensional column vector $\mathbf{y}$ *on a subspace* $\mathcal{U}$ *of* $\mathbb{R}^n$*, and let* X *represent any* $n \times p$ *matrix whose columns span* $\mathcal{U}$*. Then any* $p \times 1$ *vector* $\mathbf{b}^*$ *such that* $\mathrm{z} = \mathrm{X}\mathbf{b}^*$ *is a solution to the linear system* $\mathrm{X}'\mathrm{X}\mathbf{b} = \mathrm{X}\mathbf{y}$ *(in* $\mathbf{b}$*).*

Projection

Proof: In light of Theorem 8, $\mathrm{X}'\mathrm{X}\mathbf{b} = \mathrm{X}'\mathbf{y}$ has a solution, say $\mathbf{a}$, and $\mathrm{z} = \mathrm{X}\mathbf{a}$. Thus,

$$\mathrm{X}'\mathrm{X}\mathbf{b}^* = \mathrm{X}'\mathbf{z} = \mathrm{X}'\mathrm{X}\mathbf{a} = \mathrm{X}'\mathbf{y}. \tag{2}$$

Example

Let us find the projection (with respect to the usual inner product) of an n–dimensional column vector $\mathbf{y}$ on a subspace $\mathcal{U}$ of $\mathbb{R}^n$ in the special case where $n = 2$, $\mathbf{y} = (4, 8)'$, and $\mathcal{U} = \text{sp}\{\mathbf{x}\}$, where $\mathbf{x} = (3, 1)'$.

Upon taking X to be the 2×1 matrix whose only column is $\mathbf{x}$, the linear system $\mathrm{X}'\mathrm{X}\mathbf{b} = \mathrm{X}'\mathbf{y}$ becomes $(10)\mathbf{b} = (20)$, which has the unique solution $\mathbf{b} = (2)$.

Thus, the projection of $\mathbf{y}$ on $\mathcal{U}$ is

$$z = \begin{pmatrix} 3 \\ 1 \end{pmatrix} (2) = \begin{pmatrix} 6 \\ 2 \end{pmatrix}.$$

Example

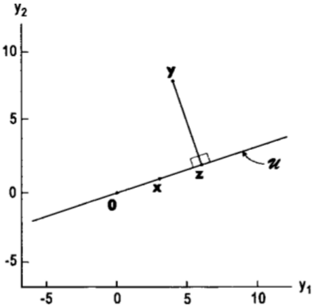

FIGURE 12.1. The projection **z** of a two-dimensional column vector **y** on a one-dimensional subspace $\mathcal{U}$ of $\mathcal{R}^2$.

Example

Let us find the projection of an n–dimensional column vector $\mathbf{y}$ on a subspace $\mathcal{U}$ of $\mathbb{R}^n$ in the special case where $n = 3$, $\mathbf{y} = (3, -38/5, 74/5)'$ and $\mathcal{U} = \text{sp}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, where

$$\mathbf{x}_1 = \begin{pmatrix} 0 \\ 3 \\ 6 \end{pmatrix}, \ \mathbf{x}_2 = \begin{pmatrix} -2 \\ 2 \\ 4 \end{pmatrix}, \ \mathbf{x}_3 = \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}.$$

Clearly, $\mathbf{x}_1$ and $\mathbf{x}_2$ are linearly independent, and $\mathbf{x}_3 = \mathbf{x}_2 - (1/3)\mathbf{x}_3$. Thus $\dim(\mathcal{U}) = 2$.

Example

Upon taking X to be 3×3 matrix whose columns are $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$, respectively, the normal equations $\mathrm{X}'\mathrm{X}\mathbf{b} = \mathrm{X}'\mathbf{y}$ become

$$\begin{pmatrix} 54 & 30 & 15 \\ 30 & 24 & 14 \\ 15 & 14 & 9 \end{pmatrix} \mathbf{b} = \begin{pmatrix} 66 \\ 38 \\ 16 \end{pmatrix}.$$

Example

One solution to these equations is the vector $(32/15, -1/2, -1)'$. Thus, the projection of $\mathbf{y}$ on $\mathcal{U}$ is

$$z = \begin{pmatrix} 0 & -2 & -2 \\ 3 & 2 & 1 \\ 6 & 4 & 2 \end{pmatrix} \begin{pmatrix} 32/15 \\ -1/2 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 22/5 \\ 44/5 \end{pmatrix}.$$

Example

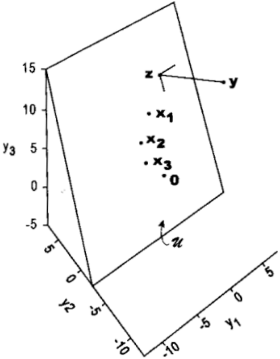

FIGURE 12.2. The projection **z** of a three-dimensional column vector **y** on a two-dimensional subspace $\mathcal{U}$ of $\mathcal{R}^3$.

Projection Matrix

Theorem 13

Let $\mathcal{U}$ represent any subspace of the linear space $\mathbb{R}^n$ of all n–dimensional column vectors. Then, there exists a unique matrix A of dimension $n \times n$ such that $\mathrm{A}\mathbf{y}$ is the projection of $\mathbf{y}$ on $\mathcal{U}$ for every column vector $\mathbf{y}$ in $\mathbb{R}^n$. Moreover, $\mathrm{A} = \mathrm{P}_{\mathrm{X}} = \mathrm{X}(\mathrm{X}'\mathrm{X})^{-}\mathrm{X}$ for any matrix X such that $\mathcal{C}(\mathrm{X}) = \mathcal{U}$.

Projection Matrix

Proof: Let X represent any matrix such that $\mathcal{C}(\mathrm{X}) = \mathcal{U}$. Then it follows from Corollary 9 that for every $\mathbf{y} \in \mathbb{R}^n$, $\mathrm{P_X}\mathbf{y}$ is the projection of $\mathbf{y}$ on $\mathcal{U}$. Moreover, if a matrix A is such that $\mathrm{A}\mathbf{y}$ is the projection then by Theorem 8 $\mathrm{A}\mathbf{y} = \mathrm{P_X}\mathbf{y}$ for every $\mathbf{y}$, implying that $\mathrm{A} = \mathrm{P_X}$. We conclude that there exists a unique matrix A such that $\mathrm{A}\mathbf{y}$ is the projection of $\mathbf{y}$ on $\mathcal{U}$ for every $\mathbf{y} \in \mathbb{R}^n$ and that $\mathrm{A} = \mathrm{P_X}$.

Projection Matrix

Corollary 14

Let A *represents the projection matrix for a subspace* $\mathcal{U}$ *of a linear space* $\mathbb{R}^n$ *of all n–dimensional column vectors. Then,* $\mathrm{A} = \mathrm{P_X}$ *for any matrix* X *such that* $\mathcal{C}(\mathrm{X}) = \mathcal{U}$*.*

Corollary 15

A matrix A *is a projection matrix if and only if* $\mathrm{A} = \mathrm{P_X}$ *for some matrix* X*.*

Projection Matrix

Theorem 16

Let X *represent any* $n \times p$ *matrix. Then,*

(1) $\mathrm{P_X X} = \mathrm{X}$, *that is,* $\mathrm{X(X'X)^{-}X'X} = \mathrm{X}$, *that is,* $\mathrm{(X'X)^{-}X'}$ *is a generalized inverse of* X*;*

(2) $\mathrm{P_X} = \mathrm{XB^*}$ *for any solution* $\mathrm{B^*}$ *to the (consistent) linear system* $\mathrm{X'XB} = \mathrm{X'}$ *(in* B*);*

(3) $\mathrm{P'_X} = \mathrm{P_X}$*;*

Projection Matrix

Theorem 16

Let X *represent any* $n \times p$ *matrix. Then,*

(4) $\mathrm{X}[(\mathrm{X}'\mathrm{X})^{-}]'\mathrm{X}'\mathrm{X} = \mathrm{X}$*; that is,* $[(\mathrm{X}'\mathrm{X})^{-}]'\mathrm{X}'$ *is a generalized inverse of* X*;*

(5) $\mathrm{X}'\mathrm{P}_{\mathrm{X}} = \mathrm{X}'P'_{\mathrm{X}} = \mathrm{X}'$*; that is,* $\mathrm{X}'\mathrm{X}(\mathrm{X}'\mathrm{X})^{-}\mathrm{X}' = \mathrm{X}'\mathrm{X}[(\mathrm{X}'\mathrm{X})^{-}]'\mathrm{X}' = \mathrm{X}'$*; that is* $\mathrm{X}(\mathrm{X}'\mathrm{X})^{-}$ *and* $\mathrm{X}[(\mathrm{X}'\mathrm{X})^{-}]'$ *are generalized inverses of* X'*.*

(6) $\mathrm{P}_{\mathrm{X}}^2 = \mathrm{P}_{\mathrm{X}}$*, that is* P_{X} *is idempotent;*

Projection Matrix

Theorem 16

Let X *represent any* $n \times p$ *matrix. Then,*

(7) $\mathcal{C}(\mathrm{P_X}) = \mathcal{C}(\mathrm{X})$ *and* $\mathcal{R}(\mathrm{P_X}) = \mathcal{R}(\mathrm{X}')$;

(8) $rank(\mathrm{P_X}) = rank(\mathrm{X})$;

(9) $rank(\mathrm{I} - \mathrm{P_X}) = n - rank(\mathrm{X})$.

Projection matrices

Proof:

(1) By definition nof generalized inverse we have

$$X'X(X'X)^{-}X'X = X'X.$$

Now consider

$$\begin{aligned} & (X(X'X)^{-}X'X - X)'(X(X'X)^{-}X'X - X) \\ = & ((X'X)^{-}X'X - I)'(X'X(X'X)^{-}X'X - X'X) \\ = & 0. \end{aligned}$$

This proves the claim.

Projection matrices

Proof:

(2) In order to prove this result we need the following result:

Theorem 17

A matrix X^* *is a solution to a consistent linear system* $\mathrm{AX} = \mathrm{B}$ *(in* X*) if and only if*

$$\mathrm{X}^* = \mathrm{A}^-\mathrm{B} + (\mathrm{I} - \mathrm{A}^-\mathrm{A})\mathrm{Y}$$

for some matrix Y.

Projection matrices

Proof:

(2) If B^* is a solution to the linear system $\mathrm{X}'\mathrm{X}\mathrm{B} = \mathrm{X}'$ then by Theorem ??,

$$\mathrm{B}^* = (\mathrm{X}'\mathrm{X})^-\mathrm{X}' + [\mathrm{I} - (\mathrm{X}'\mathrm{X})^-\mathrm{X}'\mathrm{X}]\mathrm{Y}$$

for some matrix Y. Now from Part (1) we get

$$\mathrm{X}\mathrm{B}^* = \mathrm{P}_\mathrm{X} + (\mathrm{X} - \mathrm{P}_\mathrm{X}\mathrm{X})\mathrm{Y} = \mathrm{P}_\mathrm{X}.$$

Projection matrices

Proof:

(3) First we shall show that $[(X'X)^-]'$ is a g–inverse of $X'X$. Observe that

$$\begin{aligned} X'X[(X'X)^-]'X'X &= [X'X(X'X)^-X'X]' \\ &= [X'X]' \\ &= X'X. \end{aligned}$$

Now, since $[(X'X)^-]'$ is g–inverse of $X'X$, therefore $[(X'X)^-]'X'$ is a solution to the linear system $X'XB = X'$. Thus applying Part (2) we find that

$$X(X'X)^-X' = P_X = X[(X'X)^-]'X'$$

Projection matrices

Proof:

(4) It follows from Parts (1) and (3) that

$$\mathrm{X}[(\mathrm{X}'\mathrm{X})^{-}]'\mathrm{X}'\mathrm{X} = \mathrm{P}_{\mathrm{X}}' = \mathrm{P}_{\mathrm{X}} = \mathrm{X}$$

Projection matrices

Proof:

(5) Making use of Parts (3) and (1) we get

$$\mathrm{X}'\mathrm{P}_{\mathrm{X}} = \mathrm{X}'\mathrm{P}_{\mathrm{X}}' = [\mathrm{P}_{\mathrm{X}}\mathrm{X}]' = \mathrm{X}'.$$

Projection matrices

Proof:

(6) Making use of Part (1), we find that

$$\mathrm{P}_{\mathrm{X}}^2 = \mathrm{P}_{\mathrm{X}}\mathrm{X}(\mathrm{X}'\mathrm{X})^{-}\mathrm{X}' = \mathrm{X}(\mathrm{X}'\mathrm{X})^{-}\mathrm{X}' = \mathrm{P}_{\mathrm{X}}.$$

Projection matrices

Proof:

(7) To prove this result we need the following Theorem:

Theorem 17

For any $m \times n$ matrix A and $n \times p$ matrix F, $\mathcal{C}(\mathrm{AF}) \subset \mathcal{C}(\mathrm{A})$. Similarly, for any $m \times n$ matrix A and $q \times m$ matrix L, $\mathcal{R}(\mathrm{LA}) \subset \mathcal{R}(\mathrm{A})$.

In light of above Theorem and by definition of $\mathrm{P_X}$ we have $\mathcal{C}(\mathrm{P_X}) \subset \mathcal{C}(\mathrm{X})$. From Part (1), we have $\mathcal{C}(\mathrm{X}) \subset \mathcal{C}(\mathrm{P_X})$. Thus, $\mathcal{C}(\mathrm{P_X}) = \mathcal{C}(\mathrm{X})$. Next, note that $\mathcal{C}(\mathrm{P_X}) = \mathcal{C}(\mathrm{X})$ if and only if $\mathcal{R}(\mathrm{P'_X}) = \mathcal{R}(\mathrm{X'})$. Thus, $\mathcal{R}(\mathrm{P_X}) = \mathcal{R}(\mathrm{X'})$.

Projection matrices

Proof:

(8) Directly follows from Part (7).

Projection matrices

Proof:

(9) To prove this result, following Theorem is needed:

Theorem 17

For any square matrix A *such that* $\mathrm{A}^2 = k\mathrm{A}$ *for some scalar* k*,* $trace(\mathrm{A}) = k\,rank(\mathrm{A})$*.*

Now,

$$\begin{aligned} \text{rank}(\mathrm{I} - \mathrm{P_X}) &= \text{trace}(\mathrm{I} - \mathrm{P_X}) = n - \text{trace}(\mathrm{P_X}) \\ &= n - \text{rank}(\mathrm{P_X}). \end{aligned}$$

Least Squares

Theorem 18

Let Y *represent a matrix in a linear space* $\mathcal{V}$*, and let* $\mathcal{U}$ *represent a subspace of* $\mathcal{V}$*. Then, for* $\mathrm{W} \in \mathcal{U}$*, the distance* $||\mathrm{Y} - \mathrm{W}||$ *between* Y *and* W *is minimized uniquely by taking* W *to be the projection* Z *of* Y *on* $\mathcal{U}$*. Moreover,*

$$||\mathrm{Y} - \mathrm{Z}||^2 = \mathrm{Y} \cdot (\mathrm{Y} - \mathrm{Z}).$$

Least Squares

Proof: For any matrix W in $\mathcal{U}$,

$$\begin{aligned} ||\mathrm{Y}-\mathrm{W}||^2 &= ||(\mathrm{Y}-\mathrm{Z})-(\mathrm{W}-\mathrm{Z})||^2 \\ &= (\mathrm{Y}-\mathrm{Z})\cdot(\mathrm{Y}-\mathrm{Z})-2(\mathrm{Y}-\mathrm{Z})\cdot(\mathrm{W}-\mathrm{Z}) \\ &\quad +(\mathrm{W}-\mathrm{Z})\cdot(\mathrm{W}-\mathrm{Z}). \end{aligned}$$

Least Squares

Proof: Further, $\mathrm{W} - \mathrm{Z}$ is in $\mathcal{U}$ and, by definition, $\mathrm{Y} - \mathrm{Z}$ is orthogonal to every matrix in $\mathcal{U}$. Thus, $(\mathrm{Y} - \mathrm{Z}) \cdot (\mathrm{W} - \mathrm{Z}) = 0$, and hence

$$\begin{aligned} ||\mathrm{Y} - \mathrm{W}||^2 &= (\mathrm{Y} - \mathrm{Z}) \cdot (\mathrm{Y} - \mathrm{Z}) + (\mathrm{W} - \mathrm{Z}) \cdot (\mathrm{W} - \mathrm{Z}) \\ &= ||\mathrm{Y} - \mathrm{Z}||^2 + ||\mathrm{W} - \mathrm{Z}||^2 \\ &\geq ||\mathrm{Y} - \mathrm{Z}||^2, \end{aligned}$$

with equality holding if and only if $\mathrm{W} = \mathrm{Z}$. It follows that, for $\mathrm{W} \in \mathcal{U}$, $||\mathrm{Y} - \mathrm{W}||^2$ and, consequently, $||\mathrm{Y} - \mathrm{W}||$ are minimized uniquely by taking $\mathrm{W} = \mathrm{Z}$.

Least Squares

Proof: That $||\mathrm{Y}-\mathrm{Z}||^2 = \mathrm{Y}\cdot(\mathrm{Y}-\mathrm{Z})$ is clear upon observing that $\mathrm{Z}\in\mathcal{U}$ and hence that $\mathrm{Z}\cdot(\mathrm{Y}-\mathrm{Z}) = 0$.