

Linear models with applications in R
PUBHLTH 744: Handout 5(Projection Operators)

Instructor: Andrea S. Foulkes

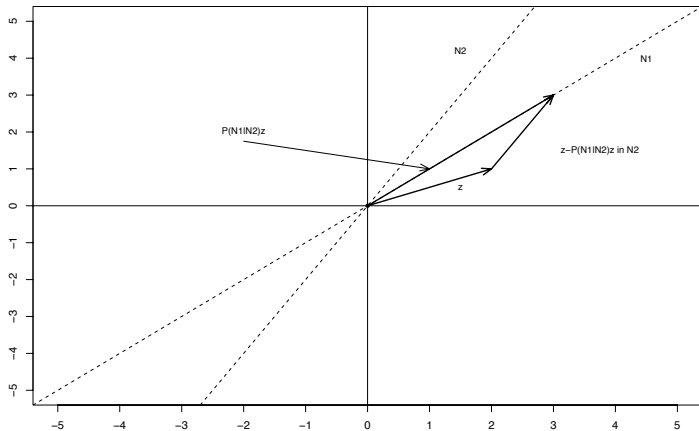
Division of Biostatistics and Epidemiology
UMass School of Public Health and Health Sciences

Fall 2007

Projection operators

Definition: Suppose \mathcal{M} is a vector space and N_1 and N_2 are two subspaces of \mathcal{M} where $N_1 + N_2 = \mathcal{M}$ and $N_1 \cap N_2 = 0$. Consider the unique decomposition of $z = x + y$ where $x \in N_1$ and $y \in N_2$. The linear transformation $P_{N_1|N_2}z = x$ is called the projection of z onto the subspace N_1 along the subspace N_2 . (See Figure)

Projection operators



Projection operators

Definition: Let A be an $n \times n$ matrix. A is called a projection operator onto $C(A)$ along $N(A)$ if for any $v \in C(A)$, $Av = v$.

Example: Let $\mathcal{M} = \mathcal{R}^2$, $N_1 = \mathcal{S} \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ and $N_2 = \mathcal{S} \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$. Since N_1 and N_2 are linearly independent, we have $N_1 + N_2 = \mathcal{R}^2$. To find $P_{N_1|N_2}$. See Figure.

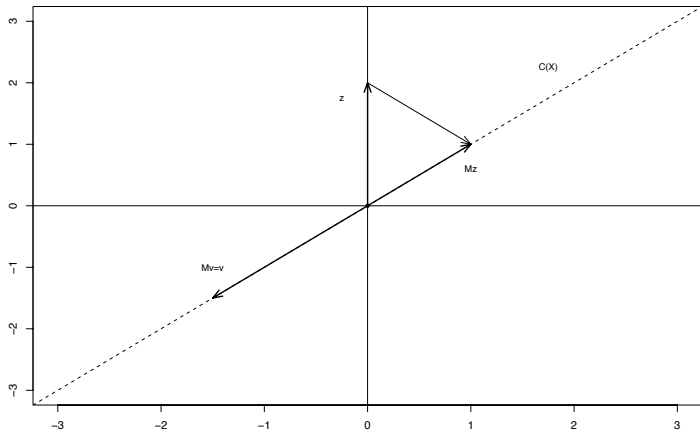
Projection operators

Theorem: $A^2 = A$ (idempotent) if and only if A is a projection operator.

Orthogonal projection operators

Definition: M is an orthogonal projection operator onto $C(X)$ if and only if (i) $v \in C(X) \rightarrow Mv = v$ and (ii) $w \in C(X)^\perp \rightarrow Mw = 0$. (See Figure)

Orthogonal projection operators



Orthogonal projection operators

Example: Let $X = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$. We then have $C(X) = \mathcal{S} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $C(X)^\perp = \mathcal{S} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. For M to be an orthogonal projection operator onto $C(X)$, we need (i) $Mv = v$ where $v = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and (ii) $Mw = 0$ where $w = \beta \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Consider

$$M = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

Orthogonal projection operators

$$(i) Mv = M \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} = \begin{pmatrix} \alpha/2 + \alpha/2 \\ \alpha/2 + \alpha/2 \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} = v$$

$$(ii) Mw = M \begin{pmatrix} \beta \\ -\beta \end{pmatrix} = \begin{pmatrix} \beta/2 - \beta/2 \\ \beta/2 - \beta/2 \end{pmatrix} = 0$$

Orthogonal projection operators

Theorem: If M is an orthogonal projection operator onto $C(X)$, then $C(M) = C(X)$.

Theorem: M is an orthogonal projection operator onto $C(M)$ if and only if $M = M^2$ (idempotent) and $M = M'$ (symmetric).

In order to prove M is the orthogonal projection operator onto $C(X)$ we need to show (1) $M = M'$ (2) $M = M^2$ and (3) $C(X) = C(M)$.

Orthogonal projection operators

Example: For X defined in Example above, we have $M = M'$ and

$$\begin{aligned} M^2 &= \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \\ &= \begin{pmatrix} 1/4 + 1/4 & 1/4 + 1/4 \\ 1/4 + 1/4 & 1/4 + 1/4 \end{pmatrix} \\ &= M \end{aligned}$$

Orthogonal projection operators

Theorem: Orthogonal projection operators are unique.

Theorem: Suppose M is an $n \times n$ orthogonal projection operator for rank $r \leq n$, then (1) The eigenvalues of M are 0 or 1, (2) $r(M) = \text{tr}(M) = r$ and (3) M is positive semidefinite.

Orthogonal projection operators

Theorem: Suppose X is an $n \times p$ matrix of rank $r \leq \min(n, p)$ and let $A = (a_1, \dots, a_r)$ where a_i is $n \times 1$ and the columns of A form an orthonormal basis for $C(X)$. Then AA' is the unique orthogonal projection operator onto $C(X)$.

Orthogonal projection operators

Proof:

$$\blacktriangleright (AA^T)^T = (A^T)^T A^T = AA^T \quad (2)$$

$$(AA^T)^2 = (AA^T)(AA^T) = A(A^T A)A^T = AIA^T = AA^T.$$

Note this is true since A is orthonormal and so, by definition, the product of each column with itself is 1 and the product of each column with another is 0.

- $\blacktriangleright (\Rightarrow)$: Let $x \in C(AA')$. Then $x = AA't$ for some t . Let $t^* = A't$. Then we have $x = AA't = At^*$ which implies $x \in C(X)$ since $C(X) = \{z : At^* = z\}$. So $C(AA') \subset C(X)$.
- $\blacktriangleright (\Leftarrow)$: Let $x \in C(X)$. We know $x \in C(A)$ since A forms an orthonormal basis for X and so we can write $x = At$ for some $t \in \mathcal{R}^p$. But A is full rank and so any vector $t \in \mathcal{R}^p$ can be written as $t = A'z$ where $z \in \mathcal{R}^n$. Thus, $x = At = AA'z$ which implies $x \in C(AA')$. Therefore $C(X) \subset C(AA')$. \square

Orthogonal projection operators

Theorem: Suppose X is an $n \times p$ matrix of rank p . Define $M = X(X'X)^{-1}X'$. Then M is the orthogonal projection operator onto $C(X)$ (along $C(X)^\perp$).

Proof:

- ▶ $M' = (X(X'X)^{-1}X')' = (X')'(X'X)^{-1'}X' = X(X'X)^{-1}X' = M$.
- ▶ $M^2 = X(X'X)^{-1}X'X(X'X)^{-1}X' = X(X'X)^{-1}X' = M$.
- ▶ Recall, $C(X) = \{z : Xt = z, t \in \mathcal{R}^p\}$.
 $C(M) = \{z^* : X(X'X)^{-1}X't = z^*, z \in \mathcal{R}^n\} = \{z^* : Xt^* = z^*, t^* \in \mathcal{R}^p\} = C(X)$. \square

Orthogonal projection operators

Theorem 5.8: $I - M$ is the unique orthogonal projection operator onto $C(X)^\perp$.

Least squares regression

Suppose $E(Y) = \mu = X\beta$. If X is full rank, then $X'X$ is invertible (also of full rank) and the least squares estimator of β is $\hat{\beta} = (X'X)^{-1}X'Y$. The least squares estimator of μ is $\hat{\mu} = X\hat{\beta} = X(X'X)^{-1}X'Y = MY$ where M is the orthogonal projection operator onto $C(X)$ and MY is the projection of Y onto this space.