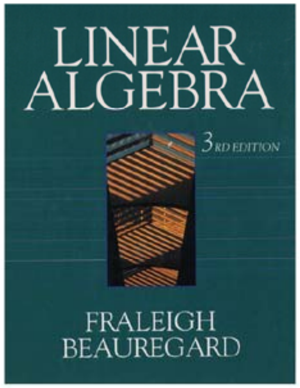
Linear Algebra

Chapter 6: Orthogonality

Section 6.4. The Projection Matrix—Proofs of Theorems



Theorem 6.10

Theorem 6.10. The Rank of $(A^T)A$**.**
Let A be an $m \times n$ matrix of rank r. Then the $n \times n$ symmetric matrix $(A^T)A$ also has rank r.

Proof. We work with nullspaces and the rank-nullity equation (Theorem 2.5). If $\vec{v}$ is in the nullspace of A then $A\vec{v} = \vec{0}$ and so $A^T(A\vec{v}) = A^T\vec{0}$ or $(A^TA)\vec{v} = \vec{0}$. Hence $\vec{v}$ is in the nullspace of A^TA. Conversely, suppose $\vec{w}$ is in the nullspace of A^TA so that $(A^TA)\vec{w} = \vec{0}$. Then $\vec{w}^T(A^TA)\vec{w} = \vec{w}^T\vec{0}$ or $(\vec{w})^T(A\vec{w}) = \vec{w}^T\vec{0} = [0]$ (notice that $\vec{w}$ is $n \times 1$ and $\vec{0}$ is $n \times 1$ so that $\vec{w}^T\vec{0}$ is 1×1). Now $(A\vec{w})^T(A\vec{w}) = [A\vec{w} \cdot A\vec{w}] = [\|A\vec{w}\|^2]$ and so $\|A\vec{w}\| = 0$ or $A\vec{w} = \vec{0}$. That is, $\vec{w}$ is in the nullspace of A^TA and nullity(A) = nullity(A^TA). Since both A and A^TA have n columns then by the rank-nullity equation (Theorem 2.5) rank(A) + nullity(A) = number of columns$(A) = n$ = number of columns(A^TA) = rank(A^TA) + nullity(A^TA) and so rank(A) = rank(A^TA), as claimed. □

Example 6.4.2

Example 6.4.2. Use the definition to find the projection matrix for the subspace in $\mathbb{R}^3$ of the x_2x_2-plane (which is spanned by $\hat{j} = [0,1,0]$ and $\hat{k} = [0,0,1]$).

Solution. We have $A = [\hat{i}\,\hat{j}] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $A^T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Then $A^TA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = (A^TA)^{-1}$, so

$$P = A(A^TA)^{-1}A^T = AA^T = \boxed{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}.$$

Notice that $P\begin{bmatrix} b_1 \\ b_2 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 \\ b_2 \\ b_2 \end{bmatrix}$, as expected. □

Page 368 Number 6

Page 368 Number 6. Find the projection for subspace $3x + 2y + z = 0$ of $\mathbb{R}^3$ and find the projection of $\vec{b} = [4, 2, -1]$ onto the plane $3x + 2y + z = 0$.

Solution. We choose two (nonzero) linearly independent vectors in the plane as a spanning set of the plane. Say, $\vec{a}_1 = [0, 1, -2]$ and $\vec{a}_2 = [1, 0, -3]$. We set $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -2 & -3 \end{bmatrix}$ and so $A^T = \begin{bmatrix} 0 & 1 & -2 \\ 1 & 0 & -3 \end{bmatrix}$.

Hence $A^TA = \begin{bmatrix} 0 & 1 & -2 \\ 1 & 0 & -3 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -2 & -3 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 6 & 10 \end{bmatrix}$. So for $(A^TA)^{-1}$, consider

$$\left[\begin{array}{cc|cc} 5 & 6 & 1 & 0 \\ 6 & 10 & 0 & 1 \end{array}\right] \overset{R_2 \to R_2 - R_1}{\widetilde{\quad}} \left[\begin{array}{cc|cc} 5 & 6 & 1 & 0 \\ 1 & 4 & -1 & 1 \end{array}\right]$$

Page 368 Number 6 (continued 1)

Solution (continued).

$$\underset{\sim}{\overset{R_1\leftrightarrow R_2}{}}\left[\begin{array}{cc|cc} 1 & 4 & -1 & 1 \\ 5 & 6 & 1 & 0 \end{array}\right] \underset{\sim}{\overset{R_2\leftrightarrow R_2-5R_1}{}} \left[\begin{array}{cc|cc} 1 & 4 & -1 & 1 \\ 0 & -14 & 6 & -5 \end{array}\right] \underset{\sim}{\overset{R_2\leftrightarrow R_2/(-14)}{}}$$

$$\left[\begin{array}{cc|cc} 1 & 4 & -1 & 1 \\ 0 & 1 & -6/14 & 5/14 \end{array}\right] \underset{\sim}{\overset{R_1\leftrightarrow R_1-4R_2}{}} \left[\begin{array}{cc|cc} 1 & 0 & 10/14 & -6/14 \\ 0 & 1 & -6/14 & 5/14 \end{array}\right]$$

and so $(A^TA)^{-1} = \frac{1}{14}\begin{bmatrix} 10 & -6 \\ -6 & 5 \end{bmatrix}$. So the projection matrix is

$$P = A(A^TA)^{-1}A^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -2 & -3 \end{bmatrix}\left(\frac{1}{14}\begin{bmatrix} 10 & -6 \\ -6 & 5 \end{bmatrix}\right)\begin{bmatrix} 0 & 1 & -2 \\ 1 & 0 & -3 \end{bmatrix}$$

$$= \frac{1}{14}\begin{bmatrix} -6 & 5 \\ 10 & -6 \\ -2 & -3 \end{bmatrix}\begin{bmatrix} 0 & 1 & -2 \\ 1 & 0 & -3 \end{bmatrix} = \boxed{\frac{1}{14}\begin{bmatrix} 5 & -6 & -3 \\ -6 & 10 & -2 \\ -3 & -2 & 13 \end{bmatrix}}.$$

Page 368 Number 6 (continued 2)

Page 368 Number 6. Find the projection for subspace $3x + 2y + z = 0$ of $\mathbb{R}^3$ and find the projection of $\vec{b} = [4, 2, -1]$ onto the plane $3x + 2y + z = 0$.

Solution (continued). So the projection of $\vec{b} = [4, 2, -1]$ onto the plane is

$$\vec{b}_W = P\vec{b}^T = \frac{1}{14}\begin{bmatrix} 5 & -6 & -3 \\ -6 & 10 & -2 \\ -3 & -2 & 13 \end{bmatrix}\begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix} = \boxed{\frac{1}{14}\begin{bmatrix} 11 \\ -2 \\ -29 \end{bmatrix}}.$$

□

Theorem 6.11

Theorem 6.11. Projection Matrix.
Let W be a subspace of $\mathbb{R}^n$. There is a unique $n \times n$ matrix P such that, for each column vector $\vec{b} \in \mathbb{R}^n$, the vector $P\vec{b}$ is the projection of $\vec{b}$ onto W. The projection matrix can be found by selecting any basis $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\}$ for W and computing $P = A(A^TA)^{-1}A^T$, where A is the $n \times k$ matrix having column vectors $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$.

Proof. We know from above that $P = A(A^Ta)^{-1}A^T$ satisfies the requirements of a projection matrix. For any vector $\vec{x} \in \mathbb{R}^n$, the transformation mapping $\vec{x} \mapsto P(\vec{x})$ is a linear transformation by Theorem 1.3.A(7) and (10). Now with the ith standard basis vector of $\mathbb{R}^n$ as $\hat{e}_i$, we have that $P\hat{e}_i$ is the ith column of P; that is, P is the standard matrix representation of the linear transformation. Since the standard matrix representation of a linear transformation is unique (assume P' is another standard matrix representation of the linear transformation and then $P\hat{e}_i = P'\hat{e}_i$ is the ith column of both P and P'). So the matrix P is unique, as claimed. □

Theorem 6.12

Theorem 6.12. Characterization Projection Matrices.
The projection matrix P for a subspace W of $\mathbb{R}^n$ is both idempotent (that is, $P^2 = P$) and symmetric (that is, $P = P^T$). Conversely, every $n \times n$ matrix that is both idempotent and symmetric is a projection matrix (specifically, it is the projection matrix for its column space).

Proof. First, we show that a projection matrix P is idempotent and symmetric (this is Exercise 6.4.16). For $P = A(A^TA)^{-1}A^T$ we have

$$P^2 = (A(A^TA)^{-1}A^T)(A(A^TA)^{-1}A^T) = A(A^TA)^{-1}(A^TA(A^TA)^{-1}A^T)$$
$$= A(A^TA)^{-1}\mathcal{I}A^T = A(A^TA)^{-1}A^T = P$$

and so P is idempotent, as claimed. Next,

$$\begin{aligned} P^T &= (A(A^TA)^{-1}A^T)^T = (A^T)^T((A^TA)^{-1})^T(A^T) \\ &= A((A^TA)^T)^{-1}A^T \text{ since } (A^T)^T = A \text{ by Note 1.3.A} \\ &\quad \text{and } (A^{-1})^T = (A^T)^{-1} \text{ by Exercise 1.5.24} \\ &= A(A^TA)^{-1}A^T = P \text{ and so } P \text{ is symmetric.} \end{aligned}$$

Theorem 6.12 (continued 1)

Proof (continued). For the converse, let P be an $n \times n$ symmetric and idempotent matrix. Let $\vec{b} \in \mathbb{R}^n$. If we show $P\vec{b} \in W$ and $\vec{b} - P\vec{b}$ is perpendicular to every vector in W where W is some subspace of $\mathbb{R}^n$ (the subspace here will be the column space of matrix P), that is the two conditions of the note after Theorem 6.10 are satisfied, then we know P is a projection matrix (based on the uniqueness of a projection matrix for a given subspace W, as given by Theorem 6.11). Now $P\vec{b}$ is in the column space of P (see Note 1.3.A). Let $p\vec{x}$ be any vector in the column space of P. Then

$$\begin{aligned}
[(\vec{b} - P\vec{b}) \cdot P\vec{x}] &= (\vec{b} - P\vec{b})^T P\vec{x} = ((\mathcal{I} - P)\vec{b})^T P\vec{x} \\
&= \vec{b}^T(\mathcal{I} - P)^T P\vec{x} = \vec{b}^T(\mathcal{I}^T - P^T)P\vec{x} \\
&= \vec{b}^T(\mathcal{I} - P)P\vec{x} \text{ since } P \text{ is symmetric} \\
&= \vec{b}^T(P - P^2)\vec{x} = \vec{b}^T(P - P)\vec{x} \text{ since } P \text{ is idempotent} \\
&= \vec{b}^T(0)\vec{x} = [0].
\end{aligned}$$

Theorem 6.12 (continued 2)

Theorem 6.12. Characterization Projection Matrices.
The projection matrix P for a subspace W of $\mathbb{R}^n$ is both idempotent (that is, $P^2 = P$) and symmetric (that is, $P = P^T$). Conversely, every $n \times n$ matrix that is both idempotent and symmetric is a projection matrix (specifically, it is the projection matrix for its column space).

Proof (continued). So $(\vec{b} - P\vec{b}) \cdot P\vec{x} = 0$ and $\vec{b} - P\vec{b}$ is orthogonal to every vector in the column space of P, as claimed. □

Page 369 Number 28

Page 369 Number 28. Find the projection matrix for the subspace W spanned by the orthonormal vectors $\vec{a}_1 = [1/2, 1/2, 1/2, 1/2]$, $\vec{a}_2 = [-1/2, 1/2, -1/2, 1/2]$, and $\vec{a}_3 = [1/2, 1/2, -1/2, -1/2]$.

Solution. We have $A = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{bmatrix}$ and so

$$P = AA^T = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{bmatrix}\left(\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix}\right)\cdots$$

Page 369 Number 28 (continued)

Page 369 Number 28. Find the projection matrix for the subspace W spanned by the orthonormal vectors $\vec{a}_1 = [1/2, 1/2, 1/2, 1/2]$, $\vec{a}_2 = [-1/2, 1/2, -1/2, 1/2]$, and $\vec{a}_3 = [1/2, 1/2, -1/2, -1/2]$.

Solution (continued). ...

$$= \boxed{\frac{1}{4}\begin{bmatrix} 3 & 1 & 1 & -1 \\ 1 & 3 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{bmatrix}}.$$

Page 369 Number 32

Page 369 Number 32. Find the projection of $\vec{b} = [4, -12, -4, 0]$ onto W the subspace of $\mathbb{R}^4$ given in Exercise 6.4.28.

Solution. We use the projection matrix of Exercise 6.4.28:

$$\vec{b}_W = P\vec{b}^T = \frac{1}{4}\begin{bmatrix} 3 & 1 & 1 & -1 \\ 1 & 3 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ -12 \\ -4 \\ 0 \end{bmatrix}$$

$$= \frac{1}{4}\begin{bmatrix} -4 \\ -28 \\ 4 \\ -20 \end{bmatrix} = \boxed{\begin{bmatrix} -1 \\ -7 \\ 1 \\ -5 \end{bmatrix}}.$$

□