Projection Matrices

Recall: If V is a subspace of $\mathbb{R}^m$ and $b \in \mathbb{R}^m$ then b_V is the orthogonal projection of b onto V.

- It is the closest vector to b in V
- It is defined by $b_{V^\perp} = b - b_V \in V^\perp$

The orthogonal decomposition of b relative to V is

$$b = \underbrace{b_V}_{\text{projection onto } V} + \underbrace{b_{V^\perp}}_{\text{projection onto } V^\perp}$$

To compute projections onto $V = \mathrm{Col}(A)$:

- Solve the normal equation $A^TA\hat{x} = A^Tb$
- Then $b_V = A\hat{x}$

Suppose that A has full column rank.

In this case, the $n \times n$ matrix A^TA is invertible:

$\mathrm{Nul}(A) = \{0\}$ because A has FCR

$\mathrm{Nul}(A^TA) = \mathrm{Nul}(A)$ always

$\Rightarrow \mathrm{Nul}(A^TA) = \{0\} \Rightarrow A^TA$ has FCR

$\Rightarrow A^TA$ has n pivots

$\Rightarrow A^TA$ is invertible ✓

(This was HW 6 #16)

Fact: If $V = \mathrm{Col}(A)$ and A has full column rank then for any $b \in \mathbb{R}^m$,

$$\boxed{b_V = A(A^TA)^{-1}A^Tb}$$

This is because $A^TA\hat{x} = A^Tb$ has the unique solution $\hat{x} = (A^TA)^{-1}A^Tb$, so

$$b_V = A\hat{x} = A(A^TA)^{-1}A^Tb$$

Eg: $V = \mathrm{Col}(A)$ $\quad A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$

$$A^TA = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix}$$

$$(A^TA)^{-1} = \frac{1}{6-4}\begin{pmatrix} 2 & -2 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 3/2 \end{pmatrix}$$

$$A(A^TA)^{-1}A^T = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 3/2 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1/2 \\ 0 & 1/2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

So if $b = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ then

$$b_V = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix} \checkmark$$

Observation: $P_V = A(A^TA)^{-1}A^T$ is an $m\times m$ matrix that computes orthogonal projections onto $V = \mathrm{Col}(A)$: $P_Vb = b_V$ for all $b \in \mathbb{R}^m$.

Def: Let V be a subspace of $\mathbb{R}^m$. The projection matrix for V is the $m\times m$ matrix P_V such that $P_V b = b_V$ for all $b \in \mathbb{R}^m$.

NB: If A & B are $m\times n$ matrices and $Ax = Bx$ for all x, then $A = B$. Indeed, $Ae_i = i^{\text{th}}$ col of A.

So "$P_V b = b_V$" determines P_V. In fact, this gives you another way to compute P_V:

$$P_V = \begin{pmatrix} | & & | \\ (e_1)_V & \cdots & (e_m)_V \\ | & & | \end{pmatrix}$$

but this is not efficient.

What if $V = \mathrm{Col}(A)$ but A does not have full column rank?

Eg: $V = \mathrm{Col}(A) \quad A = \begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & 4 \\ 1 & -1 & -1 \end{pmatrix}$

This A does not have full column rank:

$$A \xrightarrow{\text{rref}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

This says that $\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \right\}$ is a basis for V. This means:

(1) $V = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \right\} = \mathrm{Col}\begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 1 & -1 \end{pmatrix}$

(2) $\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \right\}$ is LI

$\rightsquigarrow \begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 1 & -1 \end{pmatrix}$ has full column rank.

So replace A by $B = \begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 1 & -1 \end{pmatrix}$:

$$B^T B = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 3 \end{pmatrix}$$

$$(B^T B)^{-1} = \begin{pmatrix} 1/6 & 0 \\ 0 & 1/3 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

$$P_V = B(B^TB)^{-1}B^T = \frac{1}{6}\begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & -1 \end{pmatrix}$$

$$= \frac{1}{6}\begin{pmatrix} 1 & -2 \\ 2 & 2 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & -1 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 3 & 0 & 3 \\ 0 & 6 & 0 \\ 3 & 0 & 3 \end{pmatrix}$$

$$= \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{pmatrix}$$

Procedure for Computing P_V:

(1) Find a basis $\{v_1, \dots, v_n\}$ of V

for example, if $V = \mathrm{Col}(A)$ then use the pivot columns

(2) $B = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$

(3) $P_V = B(B^TB)^{-1}B^T$

Eg: Suppose $V = \mathrm{Span}\{v\}$ is a line.

$B = v$ (matrix with one column)

$B^TB = v \cdot v$ (a scalar)

$$B(B^TB)^{-1}B^T = v\,(v\cdot v)\,v^T = \frac{\overbrace{vv^T}^{\text{outer product}}}{v \cdot v}$$

Projection Matrix onto a Line

If $V = \text{Span}\{v\}$ then $P_V = \dfrac{vv^T}{v \cdot v}$

Eg: $V = \text{Span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$

$$P_V = \frac{1}{\begin{pmatrix}1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\1\end{pmatrix}}\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1 & 1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} = \begin{pmatrix}1/2 & 1/2\\1/2 & 1/2\end{pmatrix}$$

So if $b = \begin{pmatrix}1\\0\end{pmatrix}$ then $b_V = P_V b = \begin{pmatrix}1/2\\1/2\end{pmatrix}$ ✓

Properties of Projection Matrices:

Let V be a subspace of $\mathbb{R}^m$ and let P_V be its projection matrix.

(1) $\text{Col}(P_V) = V$

(2) $\text{Nul}(P_V) = V^\perp$

(3) $P_V^2 = P_V$

(4) $P_V + P_{V^\perp} = I_m$

(5) $P_V = P_V^T$

Def: A (square) matrix S is **symmetric** if $S = S^T$.

Eg: $V = \mathrm{Col}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$ $\qquad P_V = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(1) $\mathrm{Col}(P_V)$ is the plane $\mathrm{Span}\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$.

We have $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \in \mathrm{Span}\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right\} = V$

and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \in \mathrm{Span}\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right\} = V$

so $\mathrm{Col}(P_V) = \mathrm{Col}(A) = V$. ✓

(2) $V^{\perp} = \mathrm{Nul}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \mathrm{Span}\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\right\}$ (W6L1)

$$P_V = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{\text{rref}} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{\text{PVF}} \mathrm{Span}\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\right\}$$ ✓

(3) $$P_V^2 = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ ✓

(4) $V^{\perp} = \mathrm{Span}\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\right\}$ is a line, so

$$P_{v^\perp} = \frac{1}{\begin{pmatrix}-1\\1\\0\end{pmatrix}\cdot\begin{pmatrix}-1\\1\\0\end{pmatrix}} \begin{pmatrix}-1\\1\\0\end{pmatrix}\begin{pmatrix}-1 & 1 & 0\end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix}1 & -1 & 0\\ -1 & 1 & 0\\ 0 & 0 & 0\end{pmatrix} = \begin{pmatrix}1/2 & -1/2 & 0\\ -1/2 & 1/2 & 0\\ 0 & 0 & 0\end{pmatrix}$$

and $P_v + P_{v^\perp}$ is

$$\begin{pmatrix}1/2 & 1/2 & 0\\ 1/2 & 1/2 & 0\\ 0 & 0 & 1\end{pmatrix} + \begin{pmatrix}1/2 & -1/2 & 0\\ -1/2 & 1/2 & 0\\ 0 & 0 & 0\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} \checkmark$$

(5) $$\begin{pmatrix}1/2 & 1/2 & 0\\ 1/2 & 1/2 & 0\\ 0 & 0 & 1\end{pmatrix}^T = \begin{pmatrix}1/2 & 1/2 & 0\\ 1/2 & 1/2 & 0\\ 0 & 0 & 1\end{pmatrix} \checkmark$$

Proofs of the Properties:

This is a translation of properties of projections (WGL2)

(1) $\mathrm{Col}(P_v) = \{P_v b : b \in \mathbb{R}^m\} = \{b_v : b \in \mathbb{R}^m\}$

This equals V

- because $b_v \in V$ for any b,
- and $b_v = b$ for any $b \in V$.

(2) $\mathrm{Nul}(P_v) = \{b \in \mathbb{R}^m : P_v b = 0\} = \{b \in \mathbb{R}^m : b_v = 0\}$

But we know $b_v = 0 \iff b \in V^\perp$.

(3) For any vector b,

$$P_V^2 b = P_V(P_V b) = P_V(b_V) = (b_V)_V$$

This equals b_V because $b_V \in V$ already

$$= b_V = P_V b$$

Since $P_V^2 b = P_V b$ for all vectors b, $P_V^2 = P_V$.

(4) For any vector b,

$$(P_V + P_{V^\perp})b = P_V b + P_{V^\perp} b = b_V + b_{V^\perp}$$

This equals b because $b = b_V + b_{V^\perp}$ is the orthogonal decomposition.

$$= b = I_m b$$

Since $(P_V + P_{V^\perp})b = I_m b$ for all vectors b, $P_V + P_{V^\perp} = I_m$.

(5) Choose a basis for $V \rightsquigarrow P_V = B(B^TB)^{-1}B^T$.

$$P_V^T = (B(B^TB)^{-1}B^T)^T = B^{TT}((B^TB)^{-1})^T B^T$$

$$= B((B^TB)^T)^{-1}B^T = B(B^TB)^{-1}B^T = P_V$$

- For any invertible matrix A,

$$(A^{-1})^T = (A^T)^{-1} \quad \text{because}$$

$$(A^{-1})^T A^T = (AA^{-1})^T = I_n^T = I_n \checkmark$$

Eg: Find P_V if $V = \mathrm{Nul}(A)$.

In this case, $V^\perp = \mathrm{Col}(A^T)$, so we know how to compute $P_{V^\perp}$. Then

$$P_V = I_n - P_{V^\perp}.$$

Eg: $V = \mathrm{Nul}\begin{pmatrix} 1 & 2 & 1 \end{pmatrix} \rightsquigarrow V^\perp = \mathrm{Col}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$

This is a line:

$$P_{V^\perp} = \frac{1}{\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

$$P_V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{6} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 5 & -2 & -1 \\ -2 & 2 & -2 \\ -1 & -2 & 5 \end{pmatrix}$$

This was much easier than finding a basis for V using PVF, then using $P_V = A(A^TA)^{-1}A^T$.

$\rightarrow$ Be intelligent about what you actually have to compute!