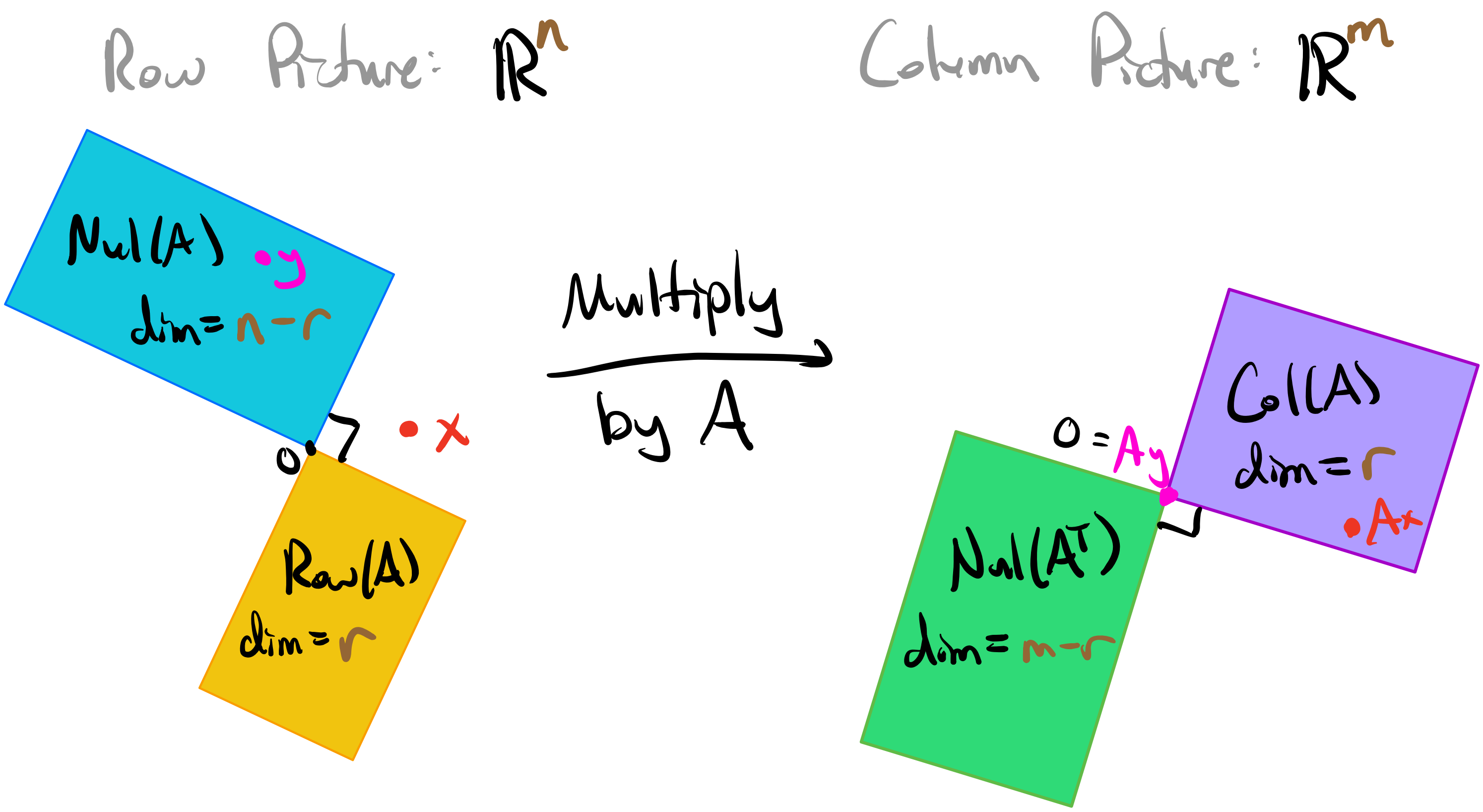
The Big Picture

Last time we discussed orthogonality of the 4 subspaces. Here is a summary:

The Big Picture

for an $m \times n$ matrix A of rank r



NB: The dimensions match up with $\dim V + \dim V^{\perp} = n$:

$$\dim \mathrm{Nul}(A) + \dim \mathrm{Row}(A) = n$$
$$\dim \mathrm{Nul}(A^T) + \dim \mathrm{Col}(A) = m$$

✓

Recall: If A has columns $v_1, \dots, v_n$ then

$$A^TA = \begin{pmatrix} - v_1^T - \\ \vdots \\ - v_n^T - \end{pmatrix} \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} = \begin{pmatrix} v_1 \cdot v_1 & v_1 \cdot v_2 & \cdots & v_1 \cdot v_n \\ v_2 \cdot v_1 & v_2 \cdot v_2 & \cdots & v_2 \cdot v_n \\ \vdots & & \ddots & \vdots \\ v_n \cdot v_1 & v_n \cdot v_2 & \cdots & v_n \cdot v_n \end{pmatrix}$$

This is the matrix of column dot products:
the (i,j)-entry is $(\text{col } i) \cdot (\text{col } j)$

With orthogonality of the 4 subspaces, we can prove:

Important Fact that we will use many times:

$$\boxed{\text{Nul}(A^TA) = \text{Nul}(A)}$$

Proof: $\text{Nul}(A^TA)$ contains $\text{Nul}(A)$: (HW 5)

$x \in \text{Nul}(A) \Rightarrow Ax = 0 \Rightarrow A^TAx = 0 \Rightarrow x \in \text{Nul}(A^TA)$

$\text{Nul}(A)$ contains $\text{Nul}(A^TA)$:

$x \in \text{Nul}(A^TA) \Rightarrow A^TAx = 0 \Rightarrow Ax \in \text{Nul}(A^T)$

$\Rightarrow Ax \in \text{Col}(A)$ and $\text{Nul}(A^T)$

$\Rightarrow (Ax) \cdot (Ax) = 0 \Rightarrow Ax = 0 \Rightarrow x \in \text{Nul}(A)$ ✓

Implicit Equations, Revisited

Recall: $\mathrm{Nul}(A) \overset{\text{PVF}}{\rightsquigarrow} \mathrm{Span}\{v_1, \dots, v_{n-r}\}$

takes the implicit equation $Ax = 0$

and generates the parametric form

$$x = a_1 v_1 + \cdots + a_{n-r} v_{n-r}. \qquad a_1, \dots, a_{n-r} = \text{parameters}$$

Orthogonal complements let us go the other way!

$(\cdot)^\perp$ turns implicit into parametric & vice-versa.

$$\mathrm{Nul}(A)^\perp = \mathrm{Row}(A) \qquad \mathrm{Col}(A)^\perp = \mathrm{Nul}(A^T)$$

Recipe: To produce implicit equations for $\mathrm{Col}(A)$:

parametric

(1) Find PVF for $\mathrm{Nul}(A^T)$:

$$\mathrm{Nul}(A^T) \overset{\text{PVF}}{\rightsquigarrow} \mathrm{Span}\{v_1, \dots, v_{m-r}\}$$

(2) $\mathrm{Col}(A) = \mathrm{Nul}(A^T)^\perp$

$$= \mathrm{Span}\{v_1, \dots, v_{m-r}\}^\perp$$

$$= \mathrm{Nul}\begin{pmatrix} — v_1^T — \\ \vdots \\ — v_{m-r}^T — \end{pmatrix} \leftarrow \text{implicit}$$

Null Space: implicit form

$\xrightarrow{\text{PVF}}$

both require elimination

$\xleftarrow{(-)^\perp \text{ then PVF then } (-)^\perp}$

Column Space: parametric form

Like: easy to check if $x \in V$: $Ax = 0$

Like: can produce vectors in V: $x = a_1 v_1 + \cdots + a_n v_n$

Eg: Find an implicit equation for the plane

$$V = \underbrace{\text{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}}_{\text{parametric description}}$$

$$V^\perp = \text{Nul}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \xrightarrow{\text{PVF}} \text{Span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

$$\Rightarrow V = \text{Span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\}^\perp = \text{Nul}\begin{pmatrix} -1 & 1 & 0 \end{pmatrix}$$

$$= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : \overbrace{-x_1 + x_2 = 0}^{\text{implicit equation}} \right\}$$

Now it's easy to check if a vector is in V:

$-x_1 + x_2 = 0$ means $x_1 = x_2$.

$\begin{pmatrix} 3 \\ 3 \\ 7 \end{pmatrix}$ is in V. $\begin{pmatrix} 3 \\ -3 \\ 7 \end{pmatrix}$ is not.

Orthogonal Projections

Recall: to find the best approximate solution of $Ax=b$, want to find the closest vector $\hat{b}$ to b in $\mathrm{Col}(A) = \{Ax : x \in \mathbb{R}^n\}$

Want: $b - \hat{b}$ is orthogonal to $\mathrm{Col}(A)$:

$$b - \hat{b} \in \mathrm{Col}(A)^\perp = \mathrm{Nul}(A^T)$$

[demo]

Def: Let V be a subspace of $\mathbb{R}^n$ and $b \in \mathbb{R}^n$. The orthogonal projection of b onto V is the closest vector b_V in V to b. It is characterized by

$$\boxed{b - b_V \in V^\perp.}$$

The orthogonal decomposition of b relative to V is

$$\boxed{b = b_V + b_{V^\perp}}$$

Here $b_{V^\perp} = b - b_V \in V^\perp$. Note that

$$b - b_{V^\perp} = b_V \in V = (V^\perp)^\perp$$

So that $b_{V^\perp}$ is projection onto $V^\perp$.

In other words, the orthogonal decomposition is

$$\boxed{\begin{aligned} b &= \begin{pmatrix}\text{closest vector} & b_V \\ \text{to } b \text{ in } V\end{pmatrix} + \begin{pmatrix}\text{closest vector} & b_{V^\perp} \\ \text{to } b \text{ in } V^\perp\end{pmatrix} \\ b &= \begin{pmatrix}\text{projection of } b \\ \text{onto } V\end{pmatrix} + \begin{pmatrix}\text{projection of } b \\ \text{onto } V^\perp\end{pmatrix}\end{aligned}}$$

[demos]

How to compute b_V?

Step 0: Write V as a column space or a null space.

$V = \mathrm{Col}(A)$: then $V^\perp = \mathrm{Nul}(A^T)$, so

$$b - b_V \in \mathrm{Nul}(A^T) \implies A^T(b - b_V) = 0$$

If $b_V \in \mathrm{Col}(A)$ then $b_V = A\hat{x}$ for $\hat{x} \in \mathbb{R}^n$:

$$A^T(b - A\hat{x}) = 0 \implies A^Tb - A^TA\hat{x} = 0$$

$$\implies A^TA\hat{x} = A^Tb$$

Solve this equation for $\hat{x}$ $\rightsquigarrow$ $b_V = A\hat{x}$

Eg: Let $b=\begin{pmatrix}1\\0\\0\end{pmatrix}$ and $V=\mathrm{Col}\underbrace{\begin{pmatrix}1&1\\1&1\\1&0\end{pmatrix}}_{A}$.

Find $b_V=$ the orthogonal projection of b to V.

We set up the equations $A^TA\hat{x}=A^Tb$:

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

= column dot products

$$A^Tb=\begin{pmatrix}1&1&1\\1&1&0\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$$

In augmented matrix form, $A^TA\hat{x}=A^Tb$ is:

$$\left(\begin{array}{cc|c}3&2&1\\2&2&1\end{array}\right)\xrightarrow{\text{rref}}\left(\begin{array}{cc|c}1&0&0\\0&1&1/2\end{array}\right)$$

So $\hat{x}=\begin{pmatrix}0\\1/2\end{pmatrix}\implies b_V=A\hat{x}=\begin{pmatrix}1&1\\1&1\\1&0\end{pmatrix}\begin{pmatrix}0\\1/2\end{pmatrix}=\begin{pmatrix}1/2\\1/2\\0\end{pmatrix}$

[demo]

Check: $b_{V^\perp}=b-b_V=\begin{pmatrix}1/2\\-1/2\\0\end{pmatrix}$

columns of A

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

$\implies b_{V^\perp}\in\mathrm{Col}(A)^\perp$ ✓

Distance from V: $\|b-b_V\|=\|b_{V^\perp}\|=\left\|\begin{pmatrix}1/2\\-1/2\\0\end{pmatrix}\right\|=\frac{1}{\sqrt{2}}$

Orthogonal Decomposition: $\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}1/2\\1/2\\0\end{pmatrix}+\begin{pmatrix}1/2\\-1/2\\0\end{pmatrix}$

Procedure: To compute the orthogonal projection b_V of b onto $V = \mathrm{Col}(A)$:

(1) Solve the equation $A^TA\hat{x} = A^Tb$

(2) $b_V = A\hat{x}$ for any solution $\hat{x}$.

Then $b_{V^\perp} = b - b_V$, and the orthogonal decomposition of b relative to V is

$$b = b_V + b_{V^\perp}.$$

The distance from b to V is $\|b_{V^\perp}\|$.

Eg: Let $b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $V = \mathrm{Col}\begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & 4 \\ 1 & -1 & -1 \end{pmatrix}$.

Find the orthogonal decomposition of b relative to V.

(1) $$A^TA = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & -1 \\ -1 & 4 & -1 \end{pmatrix}\begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & 4 \\ 1 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 6 & 0 & 6 \\ 0 & 3 & 6 \\ 6 & 6 & 18 \end{pmatrix}$$

$$A^Tb = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & -1 \\ -1 & 4 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \\ 2 \end{pmatrix}$$

Solve $A^TA\hat{x} = A^Tb$:

$$\left(\begin{array}{ccc|c} 6 & 0 & 6 & 4 \\ 0 & 3 & 6 & -1 \\ 6 & 6 & 18 & 2 \end{array}\right) \xrightarrow{\text{PVF}} \hat{x} = \begin{pmatrix} 2/3 \\ -1/3 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}$$

(2) $b_V = A\hat{x}$ for any solution. Let's use the particular solution:

$$b_V = \begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & 4 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 2/3 \\ -1/3 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

NB: $b_V = b$: what does that mean?
b was already in V! More on this later.

Def: The normal equation of $Ax = b$ is

$$A^T A\hat{x} = A^T b$$

Fact: $A^T A\hat{x} = A^T b$ is always consistent!
(Otherwise the Procedure wouldn't work.)

Why? I claim $\mathrm{Col}(A^T) = \mathrm{Col}(A^T A)$.

From before: $\mathrm{Nul}(A) = \mathrm{Nul}(A^T A)$

Take $(\cdot)^\perp$: $\mathrm{Nul}(A)^\perp = \mathrm{Nul}(A^T A)^\perp$

$$\mathrm{Nul}(A)^\perp = \mathrm{Row}(A) = \mathrm{Col}(A^T)$$

$$\mathrm{Nul}(A^T A)^\perp = \mathrm{Row}(A^T A) = \mathrm{Col}((A^T A)^T)$$
$$= \mathrm{Col}(A^T A) \checkmark$$

Since $A^T b \in \mathrm{Col}(A^T) = \mathrm{Col}(A^T A)$, the equation $A^T A\hat{x} = A^T b$ is consistent. $\checkmark$

NB: If $\hat{x}$ and $\hat{y}$ both solve

$$A^TA\hat{x} = A^Tx = A^TA\hat{y}$$

then $0 = A^TA\hat{x} - A^TA\hat{y} = A^TA(\hat{x}-\hat{y})$

$$\Rightarrow \hat{x}-\hat{y} \in \mathrm{Nul}(A^TA) \overset{\text{Fact}}{=} \mathrm{Nul}(A) \Rightarrow A(\hat{x}-\hat{y}) = 0$$

$\Rightarrow b_V = A\hat{x} = A\hat{y}$. So any soln of $A^TA\hat{x} = A^Tb$ works.

Now we know how to project onto a column space. What if $V = \mathrm{Nul}(A)$?

Then $V^\perp = \mathrm{Nul}(A)^\perp = \mathrm{Row}(A) = \mathrm{Col}(A^T)$.

So first compute $b_{V^\perp}$ = projection onto a col space, then $b_V = b - b_{V^\perp}$.

Procedure: To compute the orthogonal projection b_V of b onto $V = \mathrm{Nul}(A)$:

(1) Compute $b_{V^\perp}$ = projection onto $V^\perp = \mathrm{Col}(A^T)$

(2) $b_V = b - b_{V^\perp}$

Use the symmetry in the orthogonal decomposition

$$b = b_V + b_{V^\perp}$$

to your advantage!

Eg: Project $b=\begin{pmatrix}1\\0\\0\end{pmatrix}$ onto $V=\mathrm{Nul}\begin{pmatrix}1&1&1\\1&1&0\end{pmatrix}$.

First we project onto $\mathrm{Col}\begin{pmatrix}1&1\\1&1\\1&0\end{pmatrix}$: ← because which is A & which is A^T

$$\begin{pmatrix}1&1&1\\1&1&0\end{pmatrix}\begin{pmatrix}1&1\\1&1\\1&0\end{pmatrix}=\begin{pmatrix}3&2\\2&2\end{pmatrix} \qquad \begin{pmatrix}1&1&1\\1&1&0\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$$

$$\left(\begin{array}{cc|c}3&2&1\\2&2&1\end{array}\right)\xrightarrow{\text{rref}}\left(\begin{array}{cc|c}1&0&0\\0&1&1/2\end{array}\right)$$

So $\hat{x}=\begin{pmatrix}0\\1/2\end{pmatrix} \Longrightarrow b_{V^\perp}=A^T\hat{x}=\begin{pmatrix}1&1\\1&1\\1&0\end{pmatrix}\begin{pmatrix}0\\1/2\end{pmatrix}=\begin{pmatrix}1/2\\1/2\\0\end{pmatrix}$

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

Projection onto a Line:

Suppose $V=\mathrm{Span}\{v\}$.

Then $V=\mathrm{Col}(A)$ where $A=v$ (one column).

$$A^TA=v^Tv=v\cdot v \quad \text{is a } 1\times 1 \text{ matrix}$$

$$A^Tb=v^Tb=v\cdot b$$

so the normal equation becomes

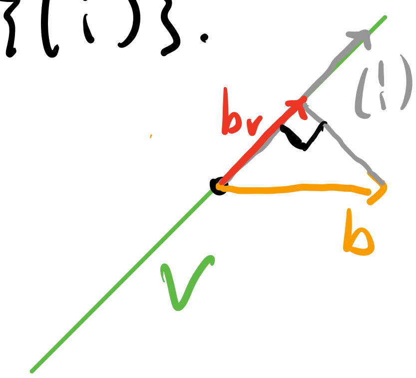
$$A^TA\hat{x}=A^Tb \rightsquigarrow (v\cdot v)\hat{x}=v\cdot b$$

Then $\hat{x}=\dfrac{v\cdot b}{v\cdot v} \rightsquigarrow b_V=A\hat{x}=\dfrac{v\cdot b}{v\cdot v}\,v$

Projection onto the Line Span$\{v\}$

$$b_V = \frac{v\cdot b}{v\cdot v}\, v$$

Eg: Project $b = \begin{pmatrix}1\\0\end{pmatrix}$ onto $V = \text{Span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$.



$$b_V = \frac{\begin{pmatrix}1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix}}{\begin{pmatrix}1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\1\end{pmatrix}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}$$

[demo]

Eg: Compute b_V where

$$V = \text{Span}\left\{\begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix}\right\} \qquad b = \begin{pmatrix}-2\\0\\-1\end{pmatrix}$$

Note V is a plane in $\mathbb{R}^3$ $\rightsquigarrow$ $V^\perp$ is a line.

In fact, $V = \text{Nul}\begin{pmatrix}1 & 1 & 1\end{pmatrix} \Longrightarrow V^\perp = \text{Span}\left\{\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$.

Much easier to compute $b_{V^\perp}$ = proj onto a line.

$$b_{V^\perp} = \frac{b\cdot v}{v\cdot v}\, v = \frac{\begin{pmatrix}-2\\0\\-1\end{pmatrix}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}}{\begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}}\begin{pmatrix}1\\1\\1\end{pmatrix} = \frac{-3}{3}\begin{pmatrix}1\\1\\1\end{pmatrix} = -\begin{pmatrix}1\\1\\1\end{pmatrix}$$

$$\Longrightarrow b_V = b - b_{V^\perp} = \begin{pmatrix}-2\\0\\-1\end{pmatrix} + \begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}-1\\1\\0\end{pmatrix}$$

[demo]

Hint: Ask yourself: is it easier to compute b_V or $b_{V^\perp}$?

NB: You get the same answer if you express V as a column space or a null space! (Or as a Col / Nul space of a different matrix.)

$$b_V = (\text{closest vector to } b \text{ in } V)$$

doesn't care how you describe V!

Eg: Let $b = \begin{pmatrix}1\\0\\0\end{pmatrix}$ and $V = \mathrm{Col}\begin{pmatrix}1 & 1\\1 & 1\\1 & 0\end{pmatrix}$.

$\rightsquigarrow \boxed{b_V = \frac{1}{2}\begin{pmatrix}1\\1\\0\end{pmatrix}}$ (p.7.)

Also $V = \mathrm{Nul}\begin{pmatrix}1 & -1 & 0\end{pmatrix}$ (p.4)

$\Rightarrow V^\perp = \mathrm{Col}\begin{pmatrix}1\\-1\\0\end{pmatrix}$ is a line

$$\Rightarrow b_{V^\perp} = \frac{\begin{pmatrix}1\\-1\\0\end{pmatrix}\cdot\begin{pmatrix}1\\0\\0\end{pmatrix}}{\begin{pmatrix}1\\-1\\0\end{pmatrix}\cdot\begin{pmatrix}1\\-1\\0\end{pmatrix}}\begin{pmatrix}1\\-1\\0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\-1\\0\end{pmatrix}$$

$$\Rightarrow b_V = b - b_{V^\perp} = \boxed{\frac{1}{2}\begin{pmatrix}1\\1\\0\end{pmatrix}}$$ again ✓