Recall:

1) If

$$\mathbf{u} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

are vectors in $\mathbb{R}^n$ then:

- $\mathbf{u} \cdot \mathbf{v} = a_1 b_1 + \ldots + a_n b_n$
- $||\mathbf{u}|| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$
- $\mathrm{dist}(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}||$

2) Vectors $\mathbf{u}, \mathbf{v}$ are orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$.

3) Pythagorean theorem: $\mathbf{u}, \mathbf{v}$ are orthogonal if and only if

$$||\mathbf{u}||^2 + ||\mathbf{v}||^2 = ||\mathbf{u} + \mathbf{v}||^2$$

4) If $V \subseteq \mathbb{R}^n$ is a subspace then an orthogonal basis of V is a basis which consists of vectors that are orthogonal to one another.

5) If $\mathcal{B} = \{\mathbf{v}_1, \ldots \mathbf{v}_k\}$ is an orthogonal basis of V and $\mathbf{w} \in V$ then

$$[\,\mathbf{w}\,]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}$$

where $c_i = \dfrac{\mathbf{w} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}$.

6) Gram-Schmidt process:

$$\boxed{\begin{array}{c}\text{a basis}\\ \{\mathbf{v}_1, \dots, \mathbf{v}_k\}\\ \text{of } V \subseteq \mathbb{R}^n\end{array}} \xrightarrow{\textit{G-S process}} \boxed{\begin{array}{c}\text{an orthogonal basis}\\ \{\mathbf{w}_1, \dots, \mathbf{w}_k\}\\ \text{of } V\end{array}}$$

$$\mathbf{w}_1 = \mathbf{v}_1$$

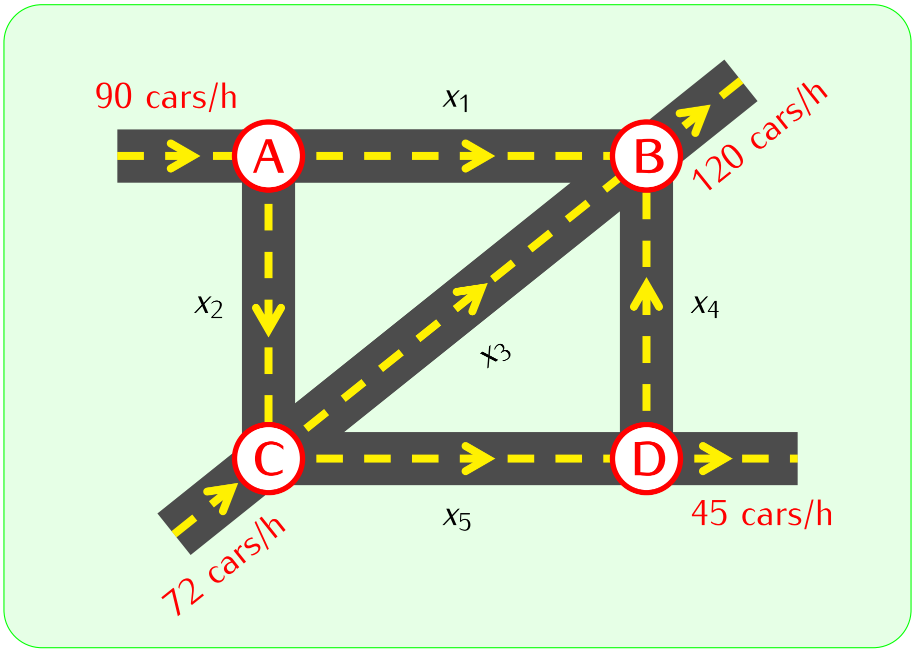
$$\mathbf{w}_2 = \mathbf{v}_2 - \left(\frac{\mathbf{w}_1 \cdot \mathbf{v}_2}{\mathbf{w}_1 \cdot \mathbf{w}_1}\right) \mathbf{w}_1$$

$$\mathbf{w}_3 = \mathbf{v}_3 - \left(\frac{\mathbf{w}_1 \cdot \mathbf{v}_3}{\mathbf{w}_1 \cdot \mathbf{w}_1}\right) \mathbf{w}_1 - \left(\frac{\mathbf{w}_2 \cdot \mathbf{v}_3}{\mathbf{w}_2 \cdot \mathbf{w}_2}\right) \mathbf{w}_2$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots$$

$$\mathbf{w}_k = \mathbf{v}_k - \left(\frac{\mathbf{w}_1 \cdot \mathbf{v}_k}{\mathbf{w}_1 \cdot \mathbf{w}_1}\right) \mathbf{w}_1 - \left(\frac{\mathbf{w}_2 \cdot \mathbf{v}_k}{\mathbf{w}_2 \cdot \mathbf{w}_2}\right) \mathbf{w}_2 - \dots - \left(\frac{\mathbf{w}_{k-1} \cdot \mathbf{v}_k}{\mathbf{w}_{k-1} \cdot \mathbf{w}_{k-1}}\right) \mathbf{w}_{k-1}$$

Problem. Find the flow rate of cars on each segment of streets:



Solution:

	FLOW IN	=	FLOW OUT
@ A:	90	=	$x_1 + x_2$
@ B:	$x_1 + x_3 + x_4$	=	120
@ C:	$72 + x_2$	=	$x_3 + x_5$
@ D:	x_5	=	$x_4 + 45$

$$\longrightarrow \begin{cases} x_1 + x_2 = 90 \\ x_1 + x_3 + x_4 = 120 \\ -x_2 + x_3 + x_5 = 72 \\ -x_4 + x_5 = 45 \end{cases}$$

augmented matrix:

$$\begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 90 \\ 1 & 0 & 1 & 1 & 0 & 120 \\ 0 & -1 & 1 & 0 & 1 & 72 \\ 0 & 0 & 0 & -1 & 1 & 45 \end{array}\right] \end{array} \xrightarrow[\text{reduction}]{\text{row}} \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

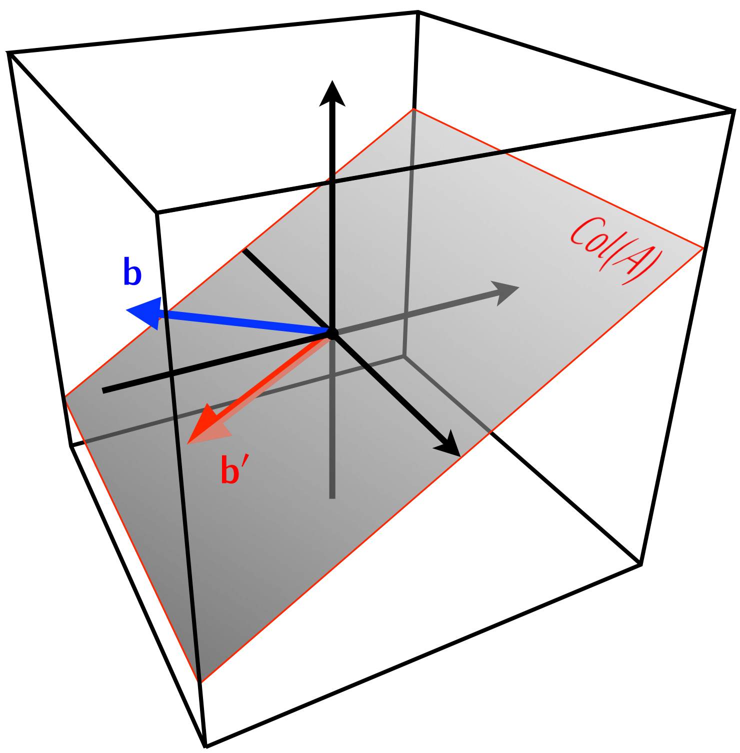
leading one in the last column, so: no solutions

Upshot.

- Recall: a matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b} \in \mathrm{Col}(A)$.

- In practical applications we may obtain a matrix equation that has no solutions, i.e. where $\mathbf{b} \notin \mathrm{Col}(A)$.

- In such cases we may look for approximate solutions as follows:
  - replace $\mathbf{b}$ by a vector $\mathbf{b}'$ such that $\mathbf{b}' \in \mathrm{Col}(A)$ and $\mathrm{dist}(\mathbf{b}, \mathbf{b}')$ is a as small as possible.
  - then solve $A\mathbf{x} = \mathbf{b}'$



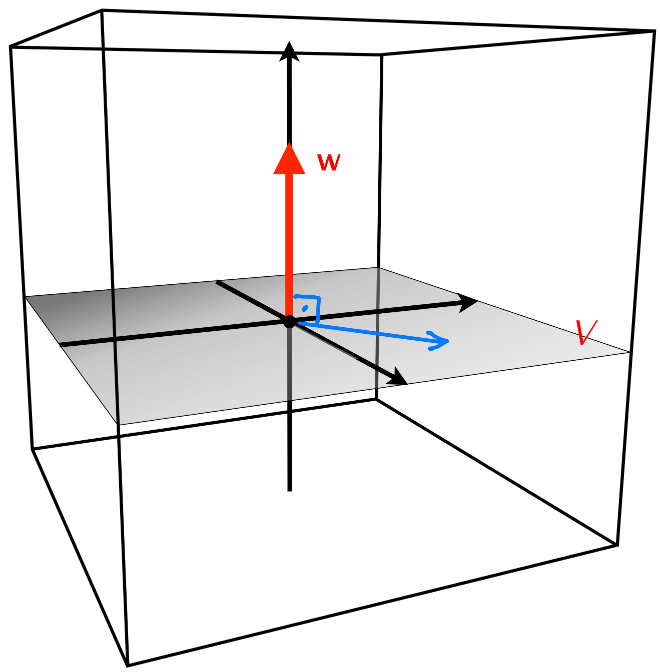
Definition

Given $\mathbf{b}' \in \mathrm{Col}(A)$ as above we will say that a vector $\mathbf{v}$ is a *least square solution* of the equation $A\mathbf{x} = \mathbf{b}$ if $\mathbf{v}$ is a solution of the equation $A\mathbf{x} = \mathbf{b}'$.

Next: How to find the vector $\mathbf{b}'$?

Definition

Let V be a subspace of $\mathbb{R}^n$. A vector $\mathbf{w} \in \mathbb{R}^n$ is *orthogonal to* V if $\mathbf{w} \cdot \mathbf{v} = 0$ for all $\mathbf{v} \in V$.



Proposition

If $V = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ then a vector $\mathbf{w} \in \mathbb{R}^n$ is orthogonal to V if and only if $\mathbf{w} \cdot \mathbf{v}_i = 0$ for $i = 1, \ldots, k$.

Proof: Assume that w is orthogonal to $v_1, \ldots, v_k$.

If $v \in \mathrm{Span}(v_1, \ldots, v_k)$ then $v = c_1 v_1 + \ldots + c_k v_k$

This gives:

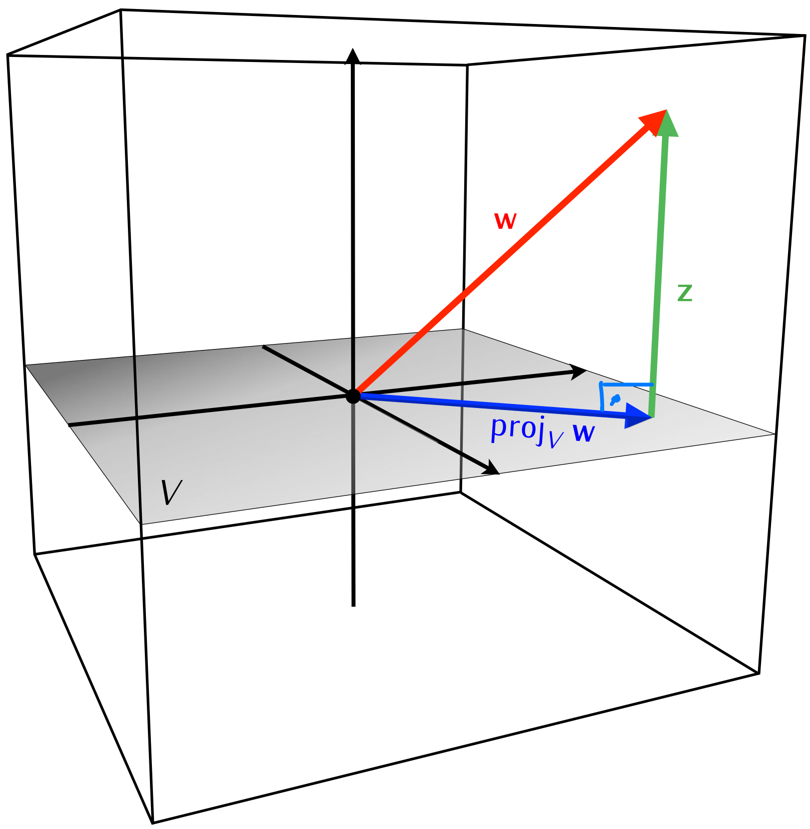
$$\begin{aligned} w \cdot v &= w \cdot (c_1 v_1 + c_2 v_2 + \ldots + c_k v_k) \\ &= c_1 (w \cdot v_1) + c_2 (w \cdot v_2) + \ldots + c_k (w \cdot v_k) = 0 \end{aligned}$$

So: w is orthogonal to every vector $v \in \mathrm{Span}(v_1, \ldots, v_k)$.

Definition

Let V be a subspace of $\mathbb{R}^n$ and let $\mathbf{w} \in \mathbb{R}^n$ the *orthogonal projection of* $\mathbf{w}$ *onto* V is a vector $\mathrm{proj}_V \mathbf{w}$ such that

1) $\mathrm{proj}_V \mathbf{w} \in V$
2) the vector $\mathbf{z} = \mathbf{w} - \mathrm{proj}_V \mathbf{w}$ is orthogonal to V.



The Best Approximation Theorem

If V is a subspace of $\mathbb{R}^n$ and $\mathbf{w} \in \mathbb{R}^n$ then $\text{proj}_V\mathbf{w}$ is a vector in V which is closest to $\mathbf{w}$:

$$\text{dist}(\mathbf{w}, \text{proj}_V\mathbf{w}) \leq \text{dist}(\mathbf{w}, \mathbf{v})$$

for all $\mathbf{v} \in V$.

Proof: Let $v \in V$. We want to show:

$$\underbrace{\text{dist}(w, \text{proj}_V w)}_{\| w - \text{proj}_V w \|} \leq \underbrace{\text{dist}(w, v)}_{\| w - v \|}$$

Note: 1) $w - \text{proj}_V w$ is a vector orthogonal to V (by the definition of $\text{proj}_V w$)

2) $\text{proj}_V w - v \in V$ (since both $\text{proj}_V w$ and v are vectors in V)

This gives:

$$(w - \text{proj}_V w) \cdot (\text{proj}_V w - v) = 0$$

By the Pythagorean Theorem we obtain:

$$\|w - \text{proj}_V w\|^2 + \underbrace{\|\text{proj}_V w - v\|^2}_{\geq 0} = \underbrace{\|(w - \cancel{\text{proj}_V w}) + (\cancel{\text{proj}_V w} - v)\|^2}_{= \|w - v\|^2}$$

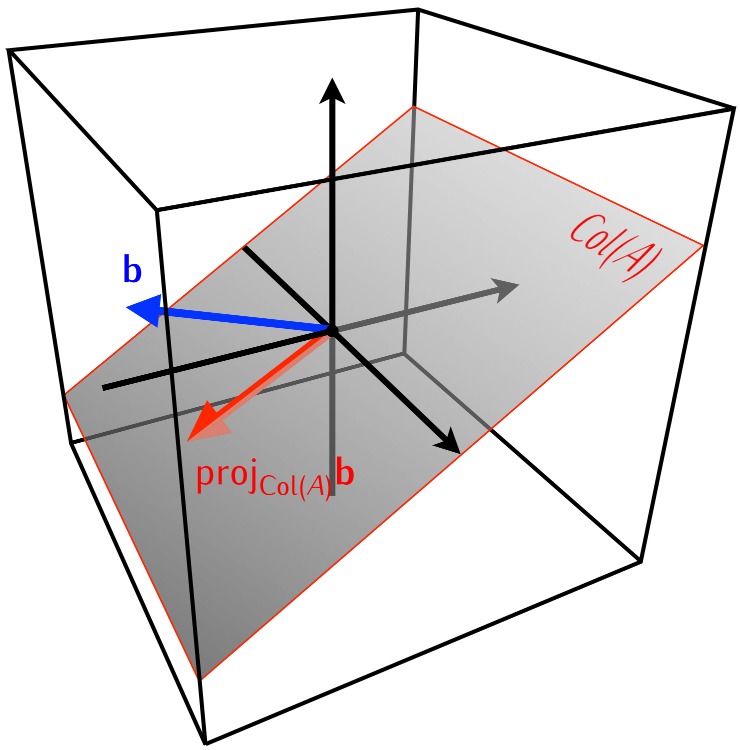
We obtain: $\|w - \text{proj}_V w\|^2 \leq \|w - v\|^2$

so: $\|w - \text{proj}_V w\| \leq \|w - v\|$

Corollary

The least square solutions of a matrix equation $A\mathbf{x} = \mathbf{b}$ are solutions of the equation

$$A\mathbf{x} = \mathrm{proj}_{\mathrm{Col}(A)}\mathbf{b}$$



Next: If V is a subspace of $\mathbb{R}^n$ and $\mathbf{w} \in \mathbb{R}^n$ how to compute $\mathrm{proj}_V\mathbf{w}$?

Theorem

If V is a subspace of $\mathbb{R}^n$ with an orthogonal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ and $\mathbf{w} \in \mathbb{R}^n$ then

$$\text{proj}_V \mathbf{w} = \left(\frac{\mathbf{w} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right) \mathbf{v}_1 + \ldots + \left(\frac{\mathbf{w} \cdot \mathbf{v}_k}{\mathbf{v}_k \cdot \mathbf{v}_k}\right) \mathbf{v}_k$$

Proof: We need to check:

1) $\text{proj}_V w \in V$

2) $w - \text{proj}_V w$ is orthogonal to V.

1) By the formula in the theorem we have

$$\text{proj}_V w \in \text{Span}(v_1, \ldots, v_k) = V$$

since $\{v_1, \ldots, v_k\}$ is a basis of V

2) Since $V = \text{Span}(v_1, \ldots, v_k)$ it suffices to check that $w - \text{proj}_V w$ is orthogonal to $v_1, \ldots, v_k$.

E.g. for v_1: $(w - \text{proj}_V w) \cdot v_1 = w \cdot v_1 - (\text{proj}_V w) \cdot v_1$

$$= w \cdot v_1 - \left[\left(\frac{w \cdot v_1}{\cancel{v_1 \cdot v_1}}\right) \cancel{(v_1 \cdot v_1)} + \left(\frac{w \cdot v_2}{v_2 \cdot v_2}\right) \underbrace{(v_1 \cdot v_2)}_{=0} + \ldots + \left(\frac{w \cdot v_k}{v_k \cdot v_k}\right) \underbrace{(v_1 \cdot v_k)}_{=0} \right]$$

$$= 0$$

Corollary

If V is a subspace of $\mathbb{R}^n$ with an orthonormal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ and $\mathbf{w} \in \mathbb{R}^n$ then

$$\text{proj}_V \mathbf{w} = (\mathbf{w} \cdot \mathbf{v}_1) \mathbf{v}_1 + \ldots + (\mathbf{w} \cdot \mathbf{v}_k) \mathbf{v}_k$$

Example. Let

$$\mathcal{B} = \left\{ \overset{v_1}{\begin{bmatrix} 1 \\ 2 \\ 0 \\ 3 \end{bmatrix}}, \overset{v_2}{\begin{bmatrix} 2 \\ -4 \\ 5 \\ 2 \end{bmatrix}}, \overset{v_3}{\begin{bmatrix} 4 \\ 1 \\ 0 \\ -2 \end{bmatrix}} \right\}, \qquad \mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 1 \end{bmatrix}$$

The set $\mathcal{B}$ is an orthogonal basis of some subspace V of $\mathbb{R}^4$. Compute $\text{proj}_V \mathbf{w}$.

Solution:

$$\text{proj}_V w = \left(\frac{w \cdot v_1}{v_1 \cdot v_1}\right) v_1 + \left(\frac{w \cdot v_2}{v_2 \cdot v_2}\right) v_2 + \left(\frac{w \cdot v_3}{v_3 \cdot v_3}\right) v_3$$

$$w \cdot v_1 = 1 \cdot 1 + 2 \cdot 2 + 2 \cdot 0 + 1 \cdot 3 = 8$$

$$v_1 \cdot v_1 = 1^2 + 2^2 + 0^2 + 3^2 = 14$$

$$w \cdot v_2 = 1 \cdot 2 + 2 \cdot (-4) + 2 \cdot 5 + 1 \cdot 2 = 6$$

$$v_2 \cdot v_2 = 2^2 + (-4)^2 + 5^2 + 2^2 = 49$$

$$w \cdot v_3 = 1 \cdot 4 + 2 \cdot 1 + 2 \cdot 0 + 1 \cdot (-2) = 4$$

$$v_3 \cdot v_3 = 4^2 + 1^2 + 0^2 + (-2)^2 = 21$$

This gives:

$$\text{proj}_V w = \frac{8}{14} \begin{bmatrix} 1 \\ 2 \\ 0 \\ 3 \end{bmatrix} + \frac{6}{49} \begin{bmatrix} 2 \\ -4 \\ 5 \\ 2 \end{bmatrix} + \frac{4}{21} \begin{bmatrix} 4 \\ 1 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 232/147 \\ 124/147 \\ 90/147 \\ 232/147 \end{bmatrix}$$

Note. In general if V is a subspace of $\mathbb{R}^n$ and $\mathbf{w} \in \mathbb{R}^n$ then in order to find $\text{proj}_V \mathbf{w}$ we need to do the following:

1) find a basis of V.

2) use the Gram-Schmidt process to get an orthogonal basis of V

3) use the orthogonal basis to compute $\text{proj}_V \mathbf{w}$.

Example. Consider the following matrix A and vector $\mathbf{u}$:

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 3 & 4 & 2 \\ 2 & 6 & 3 & -1 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$$

Compute $\text{proj}_{\text{Col}(A)} \mathbf{u}$.

Solution:

1 Find a basis of Col(A):

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 3 & 4 & 2 \\ 2 & 6 & 3 & -1 \end{bmatrix} \xrightarrow{\text{row reduction}} \begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \text{basis of Col}(A) = \left\{ \overset{v_1}{\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}}, \overset{v_2}{\begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}} \right\}$$

2 Use G-S process to get an orthogonal basis of Col(A):

$$w_1 = v_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \qquad w_2 = v_2 - \left(\frac{w_1 \cdot v_2}{w_1 \cdot w_1}\right) w_1 = v_2 - \frac{10}{5} w_1 = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

3 Use the orthogonal basis to compute $\text{proj}_{\text{Col}(A)} u$:

$$\text{proj}_{\text{Col}(A)} u = \left(\frac{u \cdot w_1}{w_1 \cdot w_1}\right) w_1 + \left(\frac{u \cdot w_2}{w_2 \cdot w_2}\right) w_2 = \frac{3}{5} w_1 + \frac{6}{6} w_2 = \frac{3}{5} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 13/5 \\ 1/5 \end{bmatrix}$$

Example. Find least square solutions of the matrix equation $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 90 \\ 120 \\ 72 \\ 45 \end{bmatrix}$$

Solution: (exercise)

1. Find a basis of Col(A).
2. Use the G-S process to get an orthogonal basis of Col(A)
3. Use the orthogonal projection to compute $\text{proj}_{\text{Col}(A)} b$
5. Solve the equation $Ax = \text{proj}_{\text{Col}(A)} b$. Solutions of this equation are the least square solutions of $Ax = b$.

Next: How to simplify this?