18.06

Zack Remscrim

Projection Matrices: Motivation

Consider equation $Ax = b$

Given an $m \times n$ matrix A

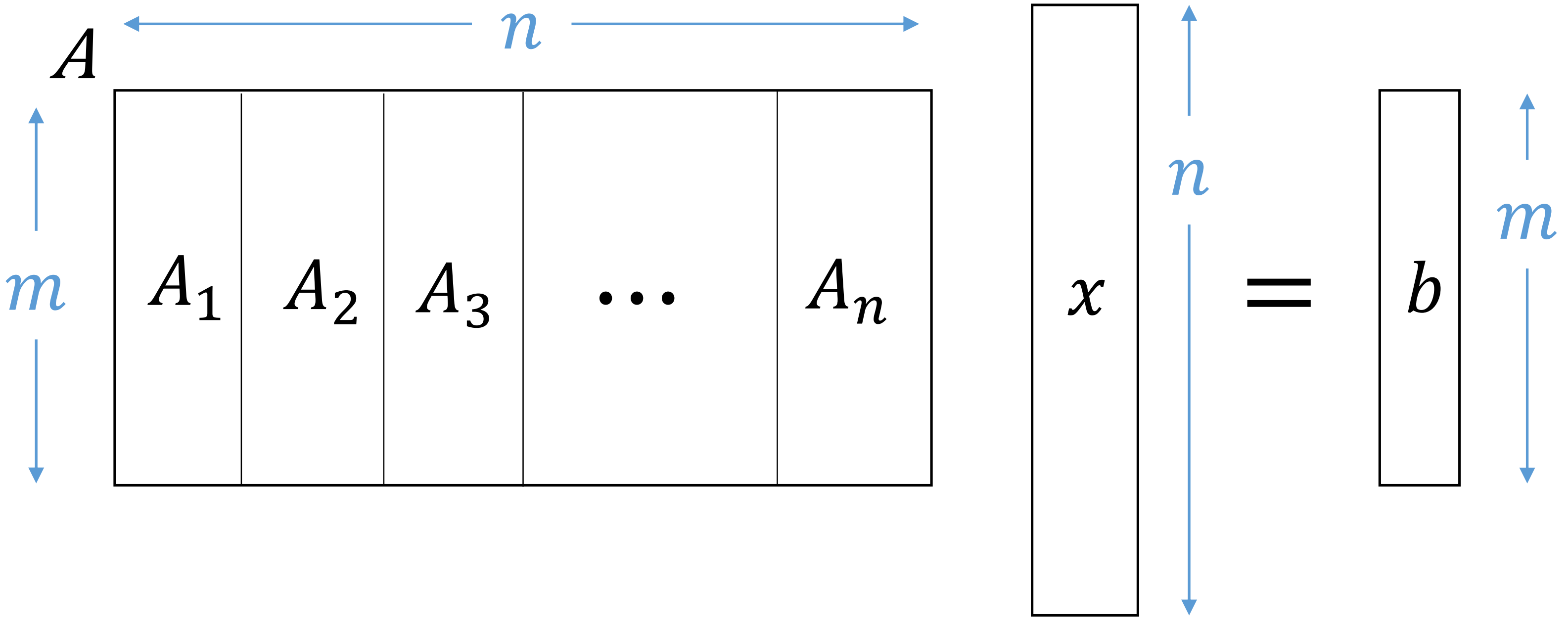
And vector $b \in \mathbb{R}^m$

Does any $x \in \mathbb{R}^n$ satisfy equation?

How many $x \in \mathbb{R}^n$ satisfy equation?

Can we describe set of all $x \in \mathbb{R}^n$ that satisfy equation?

Equation $Ax = b$



Column space $\mathrm{col}(A)$ $= \{$all linear combinations of columns $\}$

$$= \{c_1A_1 + c_2A_2 + \cdots + c_nA_n | c_i \in \mathbb{R}\}$$

$$= \{Ax | x \in \mathbb{R}^n\}$$

$= \{$ all $y \in \mathbb{R}^m$, such that $y = Ax$, for some x $\}$

Equation $Ax = b$ has at least one solution exactly when $b \in \mathrm{col}(A)$

Equation $Ax = b$

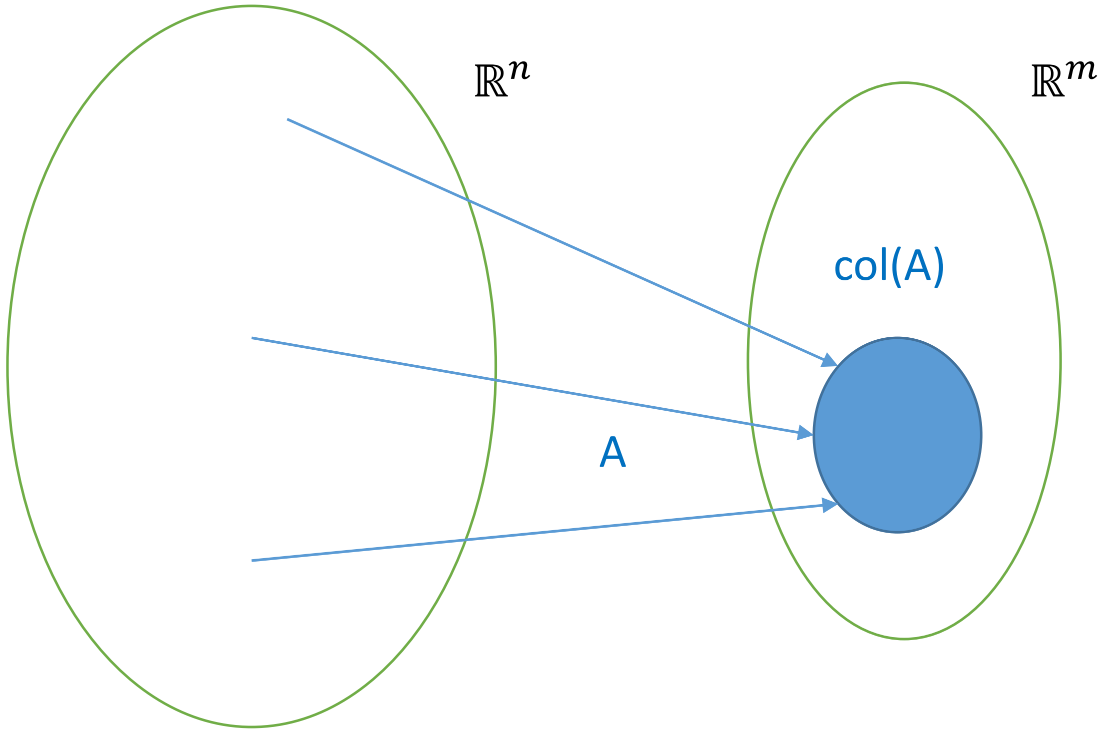
For an $m \times n$ matrix A

A is a function

input: $x \in \mathbb{R}^n$

output: $y \in \mathbb{R}^m$

$y = Ax$



Equation $Ax = b$ has at least one solution exactly when $b \in \mathrm{col}(A)$

Goal: Determine if $b \in \text{col}(A)$

Using $\text{col}(A) = \{c_1A_1 + c_2A_2 + \cdots + c_nA_n | c_i \in \mathbb{R}\}$

Seems hard to determine if $b \in \text{col}(A)$

Instead:

$\text{col}(A)$ is a vector space, which lives inside $\mathbb{R}^m$

Find $\tilde{b} \in \text{col}(A)$ that is closest to b

If $\tilde{b} = b$, $b \in \text{col}(A)$

If $\tilde{b} \neq b$, $b \notin \text{col}(A)$

"Project" b onto $\text{col}(A)$ (find $\tilde{b} \in \text{col}(A)$)

In general, for vector space V with vector space W inside it

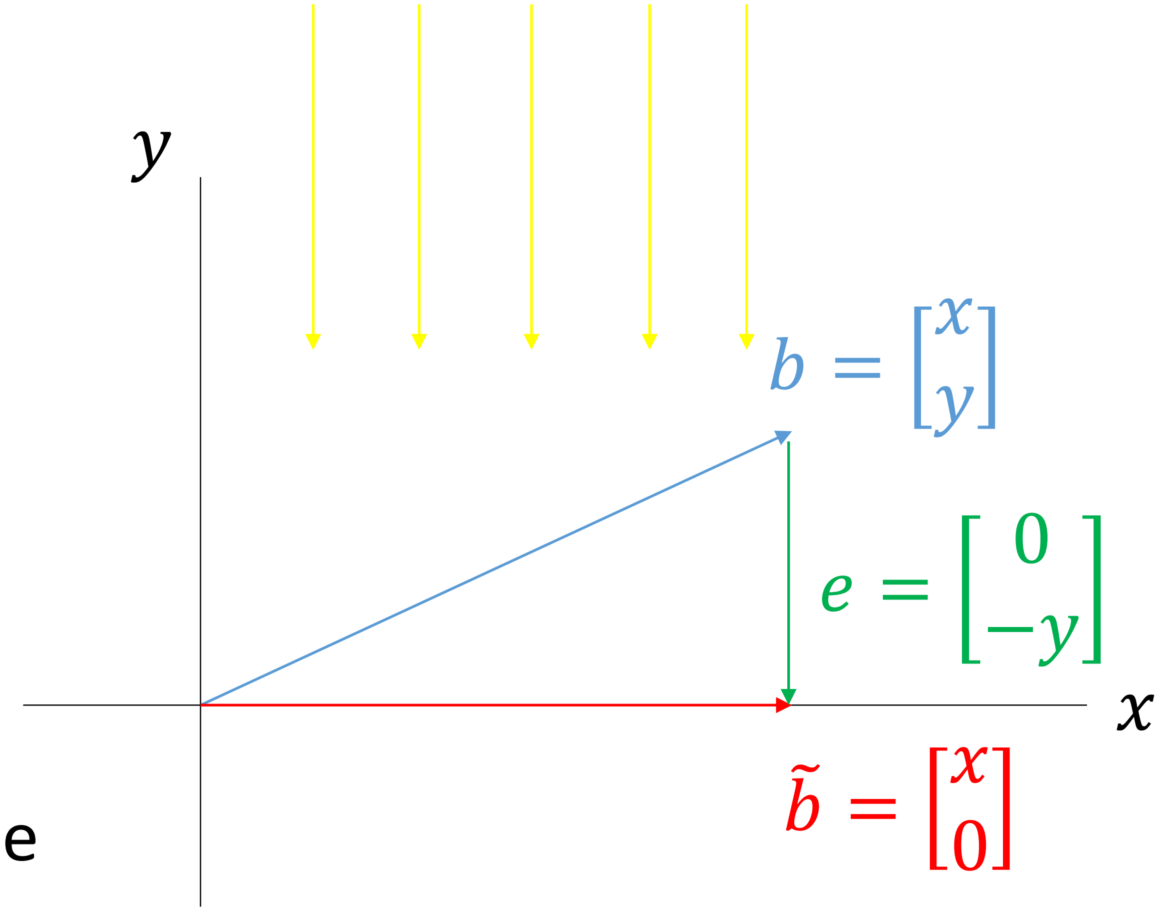
"Project" $b \in V$ onto W (find $\tilde{b} \in W$)

Projection Onto x-axis in $\mathbb{R}^2$

Given $b = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$

Goal: find $\tilde{b}$

- $\tilde{b}$ points along x-axis
- $\tilde{b}$ as close to b as possible
- $e = \tilde{b} - b$
- $\|e\| = \|\tilde{b} - b\|$ as small as possible



Say $\tilde{b}$ is projection of b onto x-axis

Projection Onto x-axis in $\mathbb{R}^2$

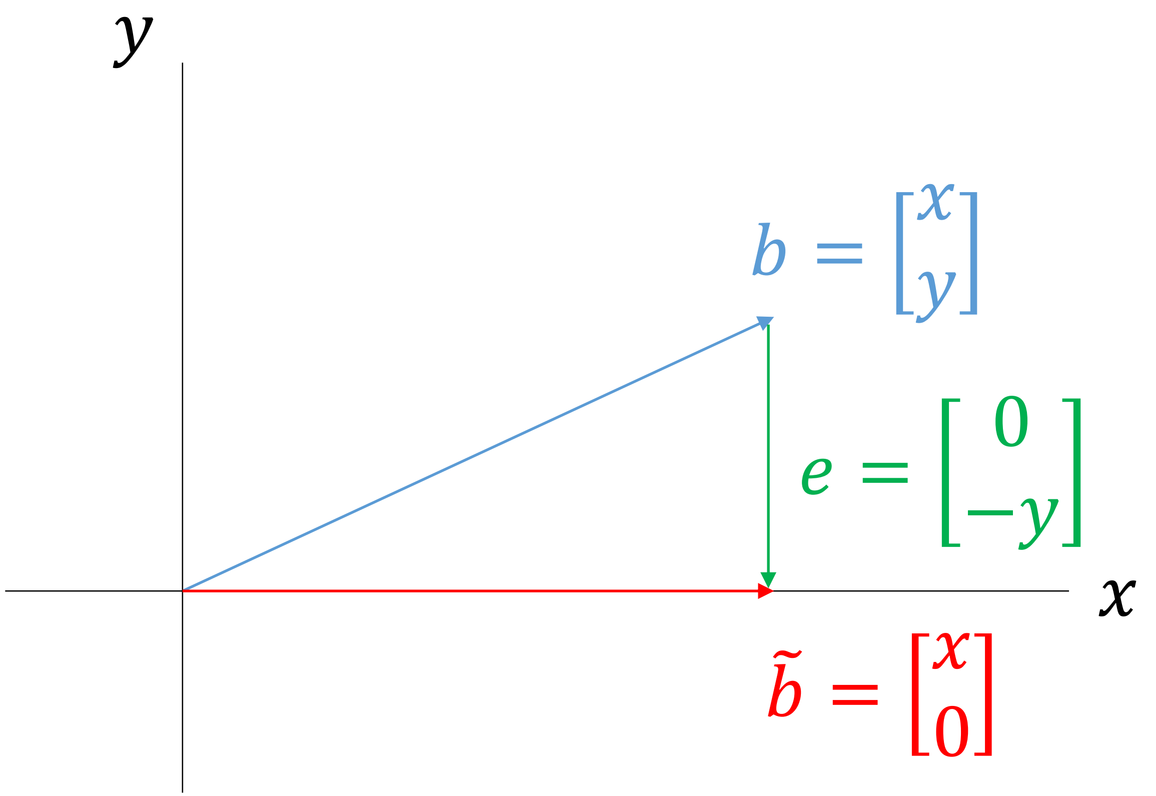
Next: Want Projection matrix P

$$\tilde{b} = Pb$$

$$\begin{bmatrix} x \\ 0 \end{bmatrix} = P \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

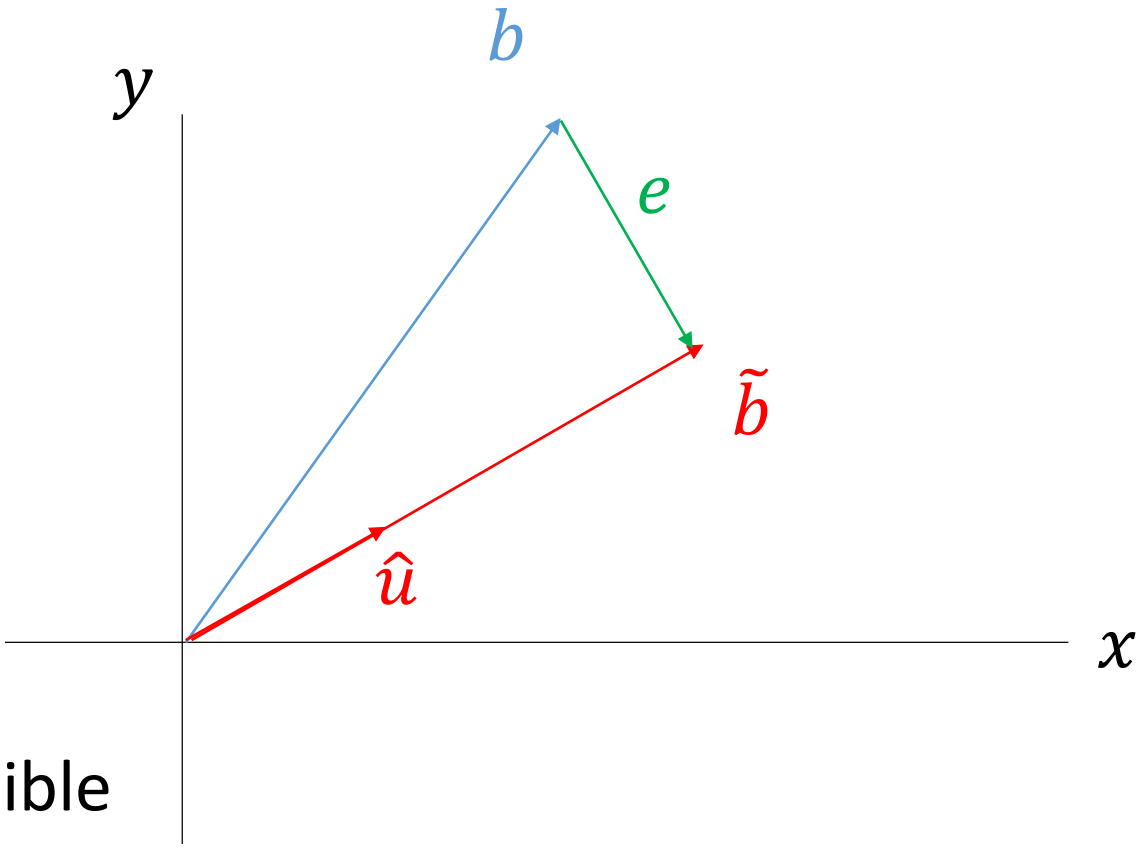


Projection Onto Line Through Origin in $\mathbb{R}^2$

Given $b \in \mathbb{R}^2$ and $\hat{u} \in \mathbb{R}^2$, $\|\hat{u}\| = 1$

Goal: find $\tilde{b}$

- $\tilde{b}$ points along $\hat{u}$
- $\tilde{b}$ as close to b as possible
- $e = \tilde{b} - b$
- $\|e\| = \|\tilde{b} - b\|$ as small as possible



Say $\tilde{b}$ is projection of b onto line through $\hat{u}$

Projection Onto Line Through Origin in $\mathbb{R}^2$

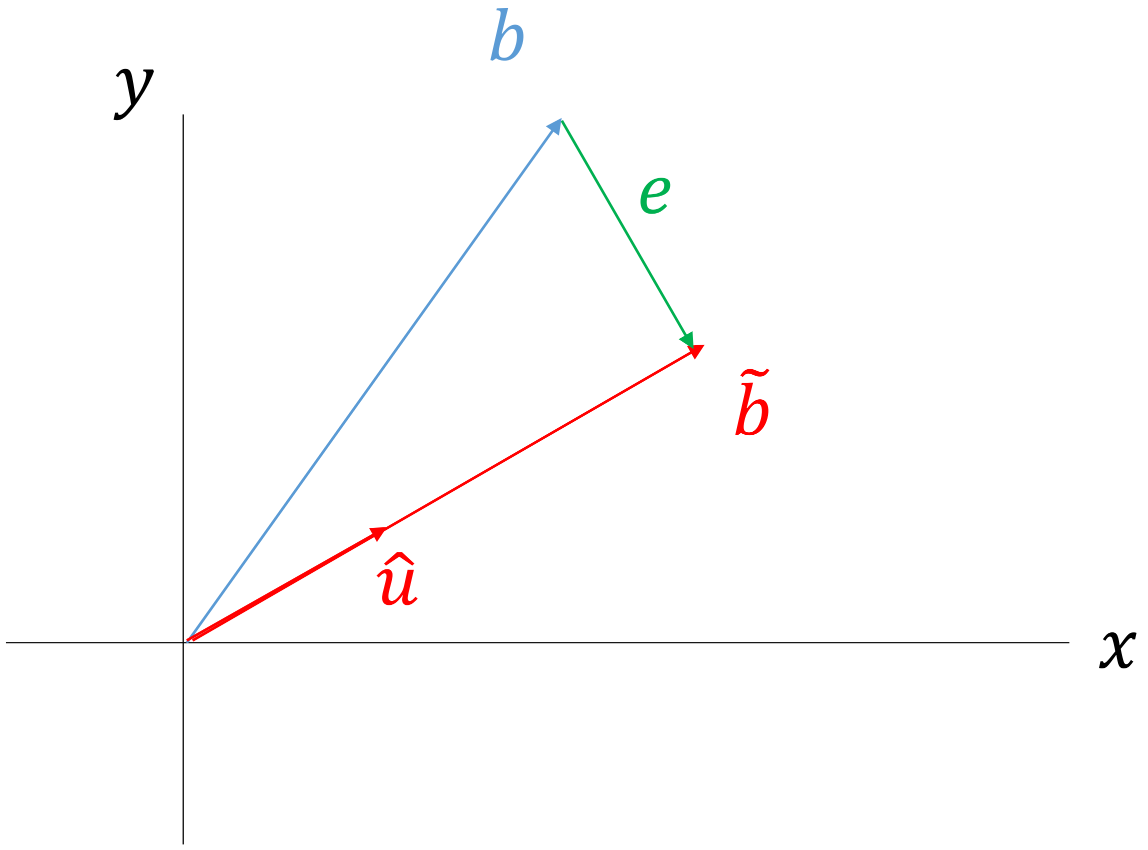
Next: Write expression for $\tilde{b}$

$$\tilde{b} = \alpha\hat{u},\ \alpha \in \mathbb{R}$$

$$\hat{u} \cdot e = 0$$
$$\hat{u} \cdot \left(\tilde{b} - b\right) = 0$$
$$\hat{u} \cdot (\alpha\hat{u} - b) = 0$$
$$\alpha\hat{u} \cdot \hat{u} - \hat{u} \cdot b = 0$$
$$\alpha - \hat{u} \cdot b = 0$$

$$\alpha = \hat{u} \cdot b$$

$$\tilde{b} = (\hat{u} \cdot b\,)\hat{u}$$



Projection Onto Line Through Origin in $\mathbb{R}^2$

Next: Want Projection matrix P

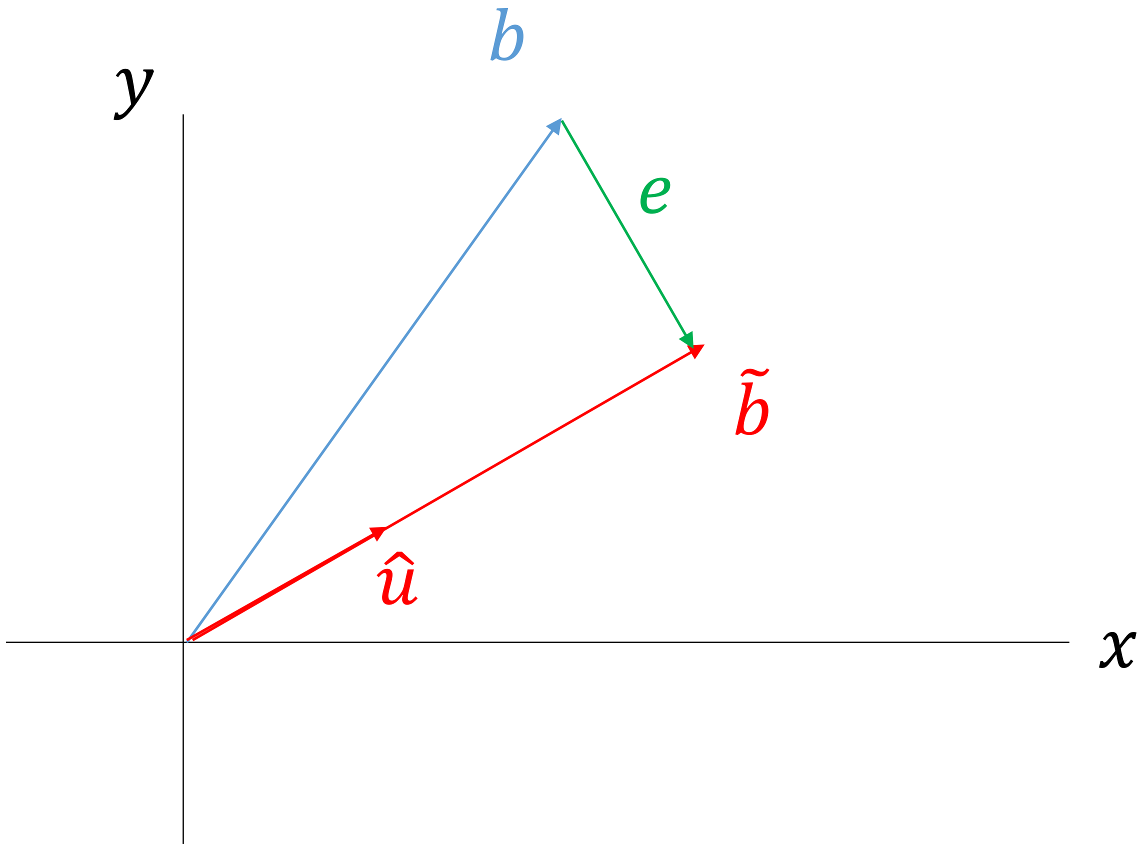
$$\tilde{b} = Pb$$

$$\begin{aligned}\tilde{b} &= (\hat{u} \cdot b)\hat{u} \\ &= \hat{u}(\hat{u} \cdot b) \\ &= \hat{u}(\hat{u}^{\top} b) \\ &= (\hat{u}\hat{u}^{\top})\, b\end{aligned}$$

$$P_{\hat{u}} = \hat{u}\hat{u}^{\top}$$

Check: Projection onto x-axis

$$\hat{u} = \begin{bmatrix}1\\0\end{bmatrix},\ P_{\begin{bmatrix}1\\0\end{bmatrix}} = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$$



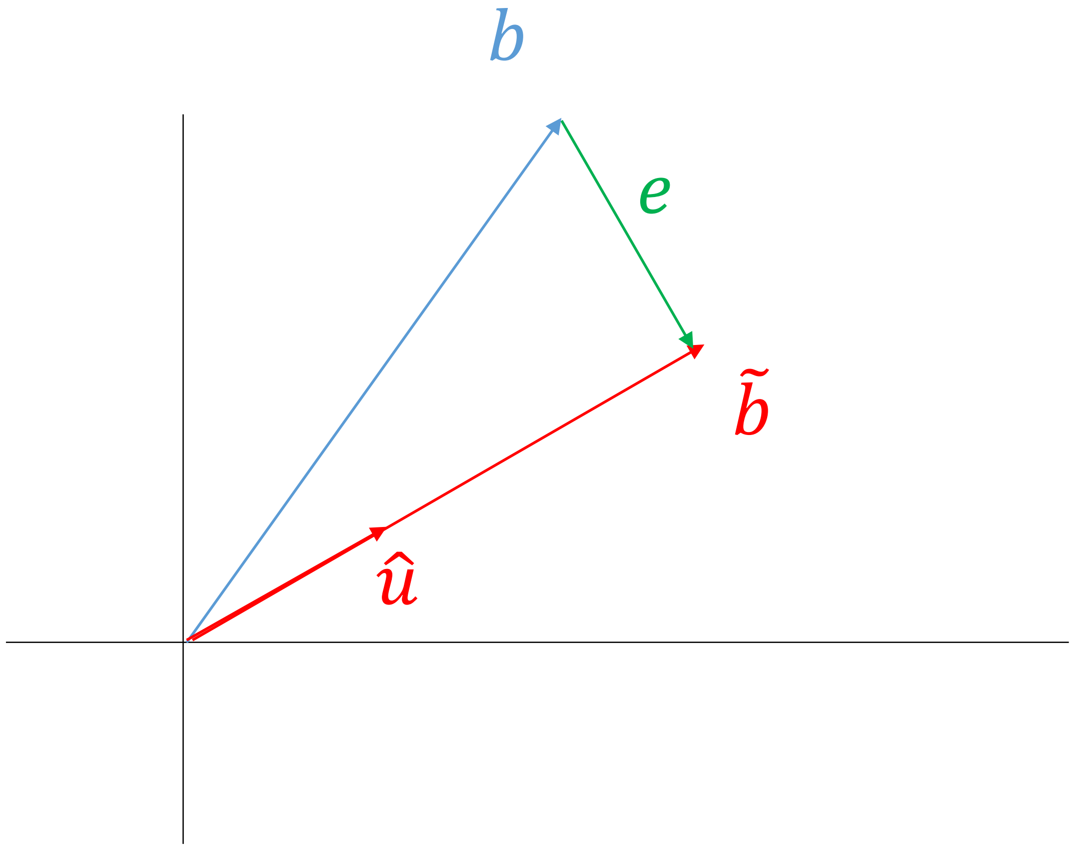
Projection Onto Line Through Origin in $\mathbb{R}^n$

Next: Want Projection matrix P

$$\tilde{b} = Pb$$

$$\begin{aligned}\tilde{b} &= (\hat{u} \cdot b)\hat{u} \\ &= \hat{u}(\hat{u} \cdot b) \\ &= \hat{u}(\hat{u}^{\mathsf{T}} b) \\ &= (\hat{u}\hat{u}^{\mathsf{T}})\, b\end{aligned}$$

$$P_{\hat{u}} = \hat{u}\hat{u}^{\mathsf{T}}$$



Projection Onto Plane Through Origin in $\mathbb{R}^3$

Given $b \in \mathbb{R}^3$ and $\hat{u}_1, \hat{u}_2 \in \mathbb{R}^3$ such that $\hat{u}_1, \hat{u}_2$ are orthonormal

Goal: find $\tilde{b}$

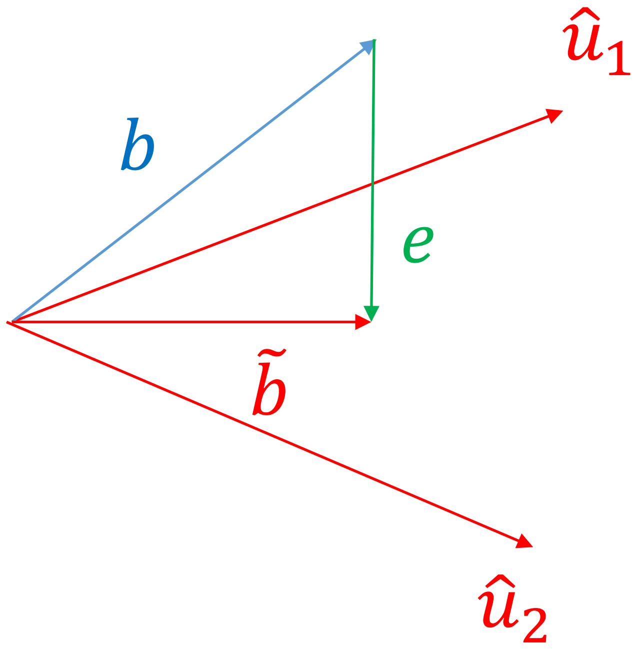
$\tilde{b} \in \text{span}(\hat{u}_1, \hat{u}_2)$

(all linear combinations of $\hat{u}_1, \hat{u}_2$)

$\tilde{b}$ as close to b as possible

$e = \tilde{b} - b$

$\|e\| = \|\tilde{b} - b\|$ as small as possible

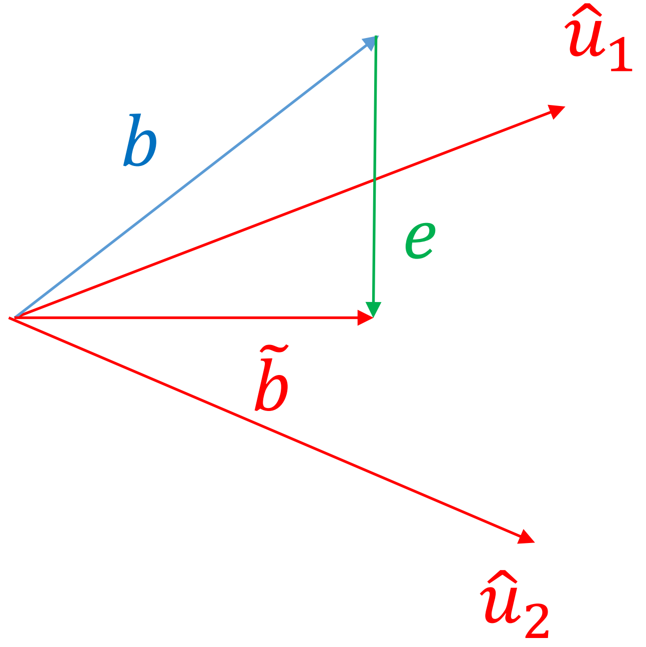


Projection Onto Plane Through Origin in $\mathbb{R}^3$

Next: Write expression for $\tilde{b}$

$$\tilde{b} = \alpha_1 \hat{u}_1 + \alpha_2 \hat{u}_2, \;\; \alpha_1, \alpha_2 \in \mathbb{R}$$

$\hat{u}_1 \cdot e = 0$ and $\hat{u}_2 \cdot e = 0$



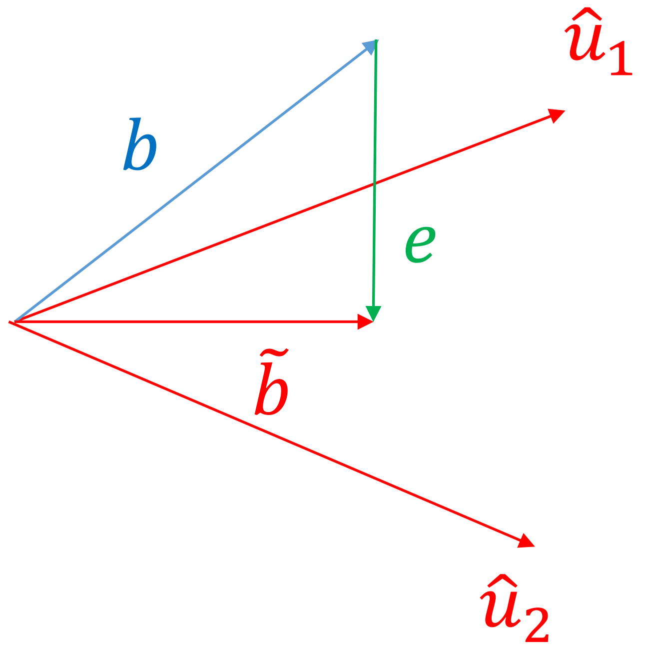
$$\hat{u}_1 \cdot e = 0$$
$$\hat{u}_1 \cdot (\tilde{b} - b) = 0$$
$$\hat{u}_1 \cdot (\alpha_1 \hat{u}_1 + \alpha_2 \hat{u}_2 - b) = 0$$
$$\alpha_1 \hat{u}_1 \cdot \hat{u}_1 + \alpha_2 \hat{u}_1 \cdot \hat{u}_2 - \hat{u}_1 \cdot b = 0$$
$$\alpha_1 - \hat{u}_1 \cdot b = 0$$

$$\left.\begin{aligned} \alpha_1 &= \hat{u}_1 \cdot b \\ \\ \alpha_2 &= \hat{u}_2 \cdot b \end{aligned}\right\} \text{looks familiar...}$$

$$\tilde{b} = (\hat{u}_1 \cdot b)\hat{u}_1 + (\hat{u}_2 \cdot b)\hat{u}_2$$

Projection Onto Plane Through Origin in $\mathbb{R}^3$

$\tilde{b} = (\hat{u}_1 \cdot b)\hat{u}_1 + (\hat{u}_2 \cdot b)\hat{u}_2$



$(\hat{u}_1 \cdot b)\hat{u}_1$ is projection of b onto line through $\hat{u}_1$
$(\hat{u}_2 \cdot b)\hat{u}_2$ is projection of b onto line through $\hat{u}_2$

$\tilde{b}$ (projection of b onto plane spanned by $\hat{u}_1, \hat{u}_2$)
$=$ sum of these two projections

$P_{\hat{u}_1,\hat{u}_2} = \hat{u}_1\hat{u}_1^{\mathsf{T}} + \hat{u}_2\hat{u}_2^{\mathsf{T}}$

Projection Onto col(U)

Given: vector $b \in \mathbb{R}^m$

U an $m \times n$ orthogonal matrix (either square or “tall-skinny”)

Goal: find $\tilde{b}$

$\tilde{b} \in$ col(U)

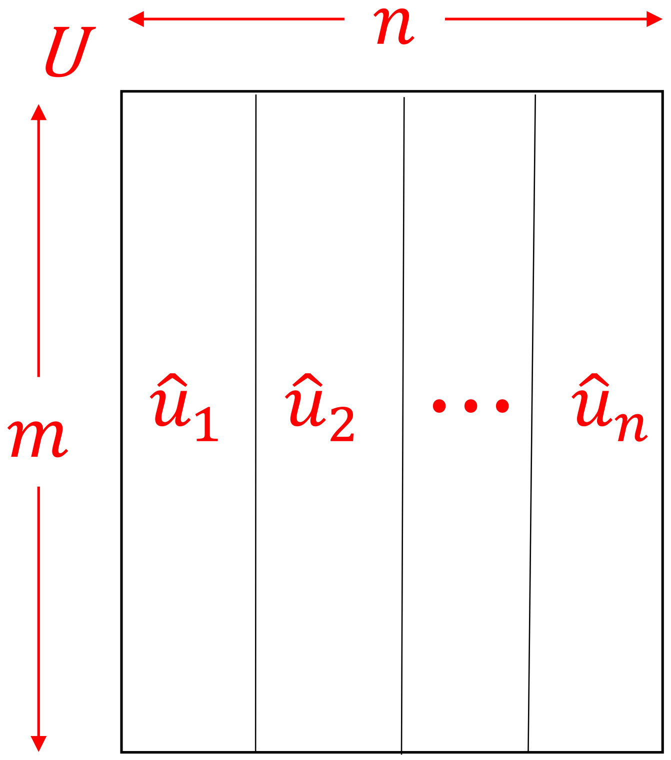
$\tilde{b}$ as close to b as possible

$e = \tilde{b} - b$

$\|e\| = \|\tilde{b} - b\|$ as small as possible

Projection Onto col(U)

By definition, U orthogonal means columns of U orthonormal



$\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_n$ are orthonormal

$\text{col}(U) = \text{span}(\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_n)$

Projection onto $\text{col}(U)$ given by

$$
\begin{aligned}
P_U &= u_1 u_1^\top + u_2 u_2^\top + \cdots + u_n u_n^\top \\
&= U U^\top
\end{aligned}
$$

Equation $Ax = b$

Recall:

$Ax = b$ has at least one solution exactly when $b \in \mathrm{col}(A)$

If A has (compact) SVD $A = U_1 \Sigma_r V_1^\top$

$\mathrm{col}(A) = \mathrm{col}(U_1)$

If $\tilde{b} = U_1 U_1^\top b$, then $\tilde{b}$ is closest element in $\mathrm{col}(U_1)$ to b

$b \in \mathrm{col}(U_1)$ exactly when $b = \tilde{b} = U_1 U_1^\top b$

$Ax = b$ has at least one solution exactly when $b = U_1 U_1^\top b$

Why Use the SVD?

$Ax = b$ has at least one solution exactly when $b \in \text{col}(A)$

Why not immediately project b onto $\text{col}(A)$?

Not clear how

Previous argument gets much messier

Important Vector Space Definitions

Given vector space V over $\mathbb{R}$ (e.g., $\mathbb{R}^3$, col(A))
And vectors $v_1, v_2, \dots, v_k \in V$

Say $v_1, v_2, \dots, v_k$ *span* V if, for each vector $x \in V$

Can write x as lin. combo. of $v_1, v_2, \dots, v_k$

$x = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$, for $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}$

Say $v_1, v_2, \dots, v_k$ are *linearly independent* if

No non-trivial lin. combo. of $v_1, v_2, \dots, v_k$ is zero (zero vector)

$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0$ only when $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$

Say $v_1, v_2, \dots, v_k$ are a *basis* of V if they span V and are linearly independent

Why Use the SVD?

If A has (compact) SVD $A = U_1 \Sigma_r V_1^\top$, then col$(A)$ =col(U_1)

Columns of U_1 span col(A)
Columns of A span col(A)

Columns of U_1 are linearly independent
Columns of A may or may not be linearly independent

Columns of U_1 form a basis of col(A)
Columns of A may or may not form a basis of col(A)