Recall:

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

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

2) If $A\mathbf{x} = \mathbf{b}$ is a consistent equation, then $\mathbf{b} \in \text{Col}(A)$, and $\text{proj}_{\text{Col}(A)}\mathbf{b} = \mathbf{b}$. In such case the least square solutions of $A\mathbf{x} = \mathbf{b}$ are just the ordinary solutions.

3) If $A\mathbf{x} = \mathbf{b}$ is inconsistent, then the least square solutions are the best substitute for the (nonexistent) ordinary solutions.

4) If $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is an orthogonal basis of a subspace V of $\mathbb{R}^n$ then

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

5) If $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is an arbitrary basis of V then we can use the Gram-Schmidt process to obtain an orthogonal basis of V.

How to compute least square solutions of $A\mathbf{x} = \mathbf{b}$

(version 1.0)

1) Compute a basis of $\mathrm{Col}(A)$.

2) Use the Gram-Schmidt process to get an orthogonal basis of $\mathrm{Col}(A)$.

3) Use the orthogonal basis to compute $\mathrm{proj}_{\mathrm{Col}(A)}\mathbf{b}$.

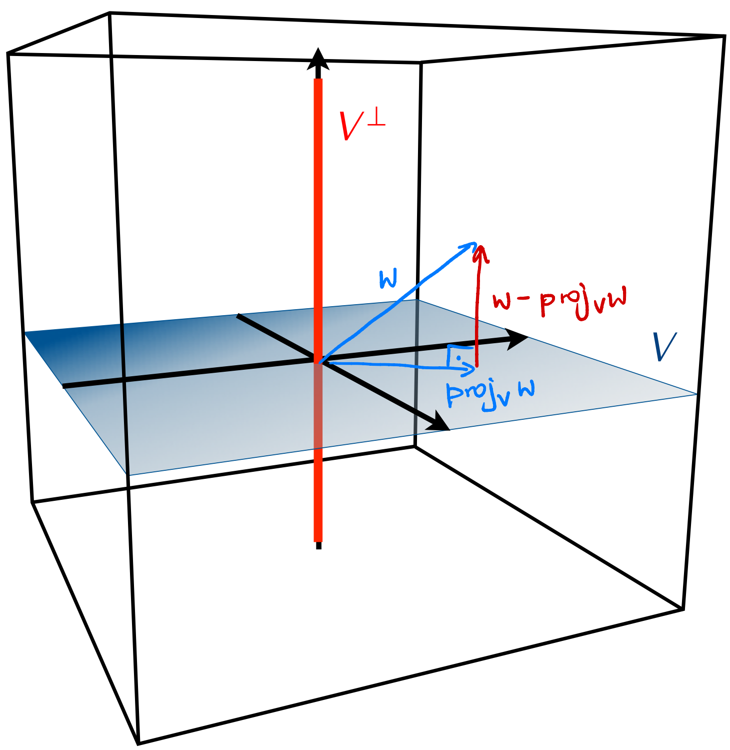
4) Compute solutions of the equation $A\mathbf{x} = \mathrm{proj}_{\mathrm{Col}(A)}\mathbf{b}$.

Next goal: Simplify this.

Definition

If V is a subspace of $\mathbb{R}^n$ then the *orthogonal complement* of V is the set $V^\perp$ of all vectors orthogonal to V:

$$V^\perp = \{\mathbf{w} \in \mathbb{R}^n \mid \mathbf{w} \cdot \mathbf{v} = 0\}$$



<u>Note:</u> If $V \subseteq \mathbb{R}^n$ $w \in \mathbb{R}^n$ then $(w - \text{proj}_V w) \in V^\perp$

Propositon

If V is a subspace of $\mathbb{R}^n$ then:

1) $V^\perp$ is also a subspace of $\mathbb{R}^n$.
2) For each vector $\mathbf{w} \in \mathbb{R}^n$ there exist unique vectors $\mathbf{v} \in V$ and $\mathbf{z} \in V^\perp$ such that $\mathbf{w} = \mathbf{v} + \mathbf{z}$.

<u>Proof of 2)</u> : Take $v = \text{proj}_V w$, $z = w - \text{proj}_V w$.

Definition

If A is an $m \times n$ matrix then the *row space* of A is the subspace $\mathrm{Row}(A)$ of $\mathbb{R}^n$ spanned by rows of A.

Example

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

$$\mathrm{Row}(A) = \mathrm{Span}\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \right)$$

Proposition

If A is a matrix then

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

<u>Note:</u>

If $A = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}$ rows of A

then: $A \cdot v = \begin{bmatrix} r_1 \cdot v \\ r_2 \cdot v \\ \vdots \\ r_m \cdot v \end{bmatrix}$

<u>e.g:</u>

$$\begin{array}{c} r_1 \\ r_2 \end{array} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \overset{v}{\begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}} = \begin{bmatrix} 1\cdot 7 + 2\cdot 8 + 3\cdot 9 \\ 4\cdot 7 \quad 5\cdot 8 + 6\cdot 9 \end{bmatrix} = \begin{bmatrix} r_1 \cdot v \\ r_2 \cdot v \end{bmatrix}$$

<u>Proof of Proposition:</u>

Let $A = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}$. We have:

$$v \in \mathrm{Nul}(A)$$
$$\Updownarrow$$
$$Av = 0$$
$$\Updownarrow$$
$$r_1 \cdot v = 0,\ r_2 \cdot v = 0,\ \ldots,\ r_m \cdot v = 0$$
$$\Updownarrow$$
$$v \in \mathrm{Row}(A)^{\perp}$$

Corollary

If A is a matrix then
$$\mathrm{Col}(A)^{\perp} = \mathrm{Nul}(A^T)$$

Note: $\mathrm{Col}(A) = \mathrm{Row}(A^T)$

e.g.:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

$$\mathrm{Col}(A) = \mathrm{Span}\left(\begin{bmatrix} 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \end{bmatrix}\right) \qquad \mathrm{Row}(A^T) = \mathrm{Span}\left(\begin{bmatrix} 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \end{bmatrix}\right)$$

Proof of Corollary

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

Back to least square solutions

Theorem

A vector $\hat{\mathbf{x}}$ is a least square solution of a matrix equation

$$A\mathbf{x} = \mathbf{b}$$

if and only if $\hat{\mathbf{x}}$ is an ordinary solution of the equation

$$(A^T A)\mathbf{x} = A^T\mathbf{b}$$

Proof: If $\hat{x}$ is a least square solution of $Ax = b$ then:

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

This gives:

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

We obtain:

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

$$A^T b - A^T A\hat{x} = 0$$

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

This shows that $\hat{x}$ is a solution of the equation

$$A^T A x = A^T b$$

Theorem

The equation

$$(A^T A)\mathbf{x} = A^T\mathbf{b}$$

is called the *normal equation* of $A\mathbf{x} = \mathbf{b}$.

How to compute least square solutions of $A\mathbf{x} = \mathbf{b}$

(version 2.0)

1) Compute $A^T A$, $A^T \mathbf{b}$.

2) Solve the normal equation $(A^T A)\mathbf{x} = A^T \mathbf{b}$.

Example. Compute least square solutions of the following equation:

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}}_{A} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}}_{b}$$

Note:

$$\underbrace{\begin{cases} x_1 + x_2 = 1 \\ 2x_2 = 2 \\ 0 = 3 \end{cases}}$$

no solutions!

Solution:

$$A^T = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

The normal equation:

$$A^T A x = A^T b$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

aug. matrix

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 5 & 5 \end{array}\right] \xrightarrow{\text{row red}} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 1 \end{array}\right]$$

least sq. solution:

$$\hat{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$