Section 6.5

Least Squares Problems

Motivation

Problem

Suppose that $Ax = b$ does not have a solution. What is the *best possible approximate* solution?

Saying $Ax = b$ **has no solution** means that b is not in Col A.

- Using $\widehat{b} = \mathrm{proj}_{\mathrm{Col}\,A}(b)$, then $A\widehat{x} = \widehat{b}$ is a *consistent equation*.
- **Plus:** $\widehat{b}$ is the *closest vector to* b such that $A\widehat{x} = \widehat{b}$ is consistent.

Solution

A solution $\widehat{x}$ to $A\widehat{x} = \widehat{b}$ is a **least squares solution.**

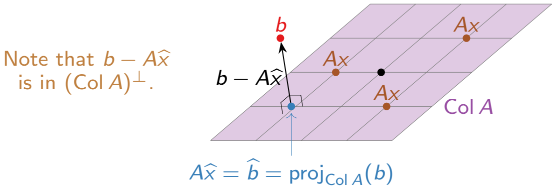
Least Squares Solutions

Definition

Let A be an $m \times n$ matrix. A **least squares solution** to $Ax = b$ is a vector $\widehat{x}$ in $\mathbf{R}^n$ such that

$$A\widehat{x} = \widehat{b} = \mathrm{proj}_{\mathrm{Col}\,A}(b).$$

> A least squares solution $\widehat{x}$ solves $Ax = b$ *as closely as possible*.



In *distance terms*, for all x in $\mathbf{R}^n$:

$$\|b - A\widehat{x}\| \leq \|b - Ax\|$$

Least Squares Solutions: Orthogonal case

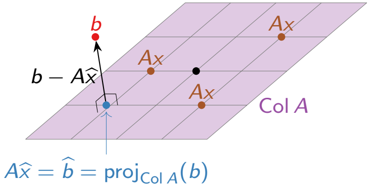
Theorem

Let A be a $m \times n$ matrix with **orthogonal columns** $v_1, v_2, \ldots, v_n$. The least squares solution to $Ax = b$ is the vector

$$\widehat{x} = \left(\frac{b \cdot v_1}{v_1 \cdot v_1}, \ \frac{b \cdot v_2}{v_2 \cdot v_2}, \ \cdots, \ \frac{b \cdot v_n}{v_n \cdot v_n} \right).$$

This is because we have formulas for the **$\mathcal{B}$-coordinates** of orthogonal basis:

$$A\widehat{x} = \sum_{i=1}^{n} \frac{b \cdot v_i}{v_i \cdot v_i} v_i = \mathrm{proj}_{\mathrm{Col}\,A}(b)$$



Least Squares Solutions: General Solution

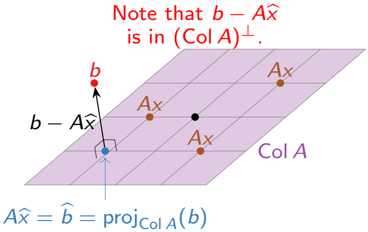
Theorem

Let A be a $m \times n$ **matrix**. Least squares solutions to $Ax = b$ are *any of the solutions to*

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

Now we can solve the problem without computing $\widehat{b}$ first.

This is just another sysmtem of equations,
but now it *is consistent* and uses *square matrix $A^T A$*!



Why is this true?

Recall: $(\text{Col}\, A)^\perp = \text{Nul}(A^T)$.

Now, $b - A\widehat{x}$ is in $(\text{Col}\, A)^\perp$ if and only if

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

In other words, $A^T A\widehat{x} = A^T b$.

Least Squares Solutions

Example 1

Find the least squares solutions to $Ax = b$ where:

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}.$$

First: Compute new matrix and vector

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

and

$$A^T b = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix}.$$

Second: Solve the new system; row reduce:

$$\left(\begin{array}{cc|c} 3 & 3 & 6 \\ 3 & 5 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & -3 \end{array}\right).$$

So the *unique* least squares *solution is* $\widehat{x} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}$.

Least Squares Solutions

Example 2

Find the least squares solutions to $Ax = b$ where:

$$A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{pmatrix} \qquad b = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

First: Compute new matrix and vector

$$A^T A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}$$

and

$$A^T b = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}.$$

Second: Solve the new system; row reduce:

$$\left(\begin{array}{cc|c} 5 & -1 & 2 \\ -1 & 5 & -2 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & 0 & 1/3 \\ 0 & 1 & -1/3 \end{array}\right).$$

So the *unique* least squares *solution is* $\widehat{x} = \begin{pmatrix} 1/3 \\ -1/3 \end{pmatrix}$.

Poll

Poll

What is the *error in the least-squares* solution $\widehat{x}$ for $Ax = b$?

1. $||Ax - b||$
2. $||A\widehat{x} - b||$
3. $||\widehat{b} - b||$
4. $||x - \widehat{x}||$

Solution: Both *b) and c)*; the two expressions are equivalent.

Least Squares Solutions: Uniqueness

When does $Ax = b$ have a *unique* least squares solution $\widehat{x}$?

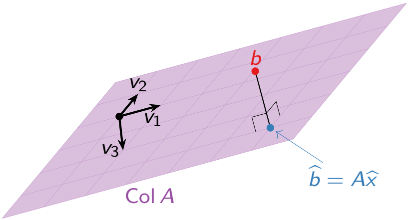
- $A^T A$ is always a square matrix, but it need not be invertible.

Theorem

Let A be an $m \times n$ matrix. The following *are equivalent*:

1. $A^T A$ is invertible.
2. The columns of A are *linearly independent*.
3. $Ax = b$ has a **unique least squares solution** for all b in $\mathbf{R}^n$, which is

$$(A^T A)^{-1}(A^T b).$$



- If the columns of A are *linearly dependent*, then $A\widehat{x} = \widehat{b}$ has many solutions.

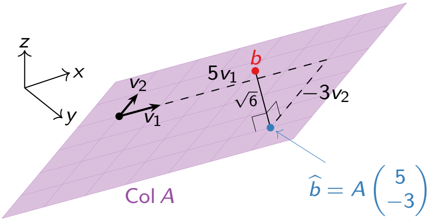
Extra: More details

From Example 1

$$A\widehat{x} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ -3 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix} = \widehat{b}$$

1. The solution $\widehat{x}$ makes the *distance from b to its approximation:*

$$\|b - A\widehat{x}\| = \left\| \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\| = \sqrt{6}.$$



2. If $A^T A$ is invertible: Let v_1, v_2 be the columns of A, and $\mathcal{B} = \{v_1, v_2\}$, then $\widehat{x} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}$ are the $\mathcal{B}$-coordinates of $\widehat{b}$, in Col $A = \text{Span}\{v_1, v_2\}$.