Section 6.5 Least-square Problems

Definition: If A is $m \times n$ and $\mathbf{b}$ is in $\mathbb{R}^m$, a least-squres solution of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$||\mathbf{b} - A\hat{\mathbf{x}}|| \leq ||\mathbf{b} - A\mathbf{x}||$$

↳ arbitrary in $\mathbb{R}^n$

for all $\mathbf{x}$ in $\mathbb{R}^n$.

Remark:

best approx theorem tells us that.

$\|y - \hat{y}\| \leq \|y - v\|$ for all $v \in W$ (is a subspace of $\mathbb{R}^n$)

↳ $\text{Proj}_W y$

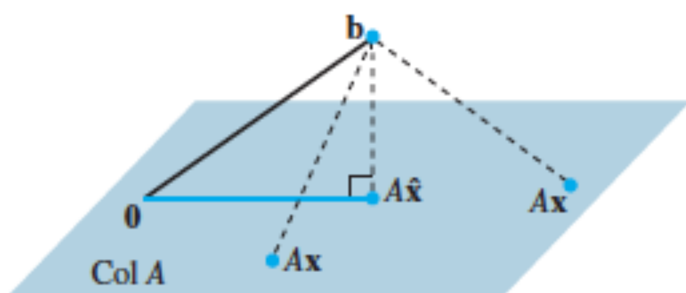


FIGURE 1 The vector $\mathbf{b}$ is closer to $A\hat{\mathbf{x}}$ than to $A\mathbf{x}$ for other $\mathbf{x}$.

Solution of the General Least-Squares Problem:

Given A and $\mathbf{b}$ as above, apply the Best Approximation Theorem in Section 6.3 to the subspace Col A. Let

$$\hat{\mathbf{b}} = \text{proj}_{\text{Col}A}\mathbf{b}$$

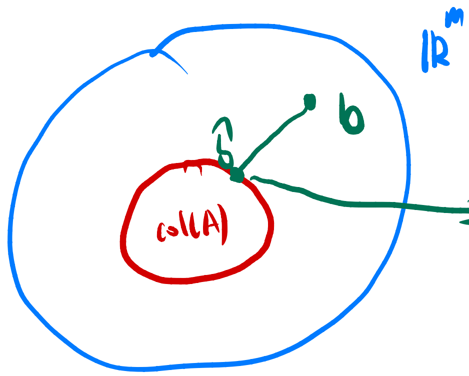
Since $\hat{\mathbf{b}}$ is in the column space of A, then there is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$A\hat{\mathbf{x}} = \hat{\mathbf{b}} \tag{1}$$

Such an $\hat{\mathbf{x}}$ is a least square solution of $A\mathbf{x} = \mathbf{b}$ if and only if $\hat{\mathbf{x}}$ satisfies (1).

$A\hat{x}$ is actual $\text{proj}_{col(A)}\, b$.

why is that?



$\text{proj}_{col(A)}\, b$ (by the best approx theorem)

$$A\hat{x} = \hat{b} = \text{proj}_{col(A)}\, b$$

① can we find $\hat{x}$?

Yes, $\hat{b} \in col(A) \Rightarrow$ there exist $\hat{x}_1 \dots \hat{x}_n \in \mathbb{R}$

such that $\vec{a}_1 \cdot \hat{x}_1 + \vec{a}_2 \cdot \hat{x}_2 + \dots \vec{a}_n \cdot \hat{x}_n = \hat{b}$

$$\hat{x} = \begin{pmatrix} \hat{x}_1 \\ \vdots \\ \hat{x}_n \end{pmatrix}$$

⇒ may not be unique

Theorem: The set of least-squares solutions of $A\mathbf{x} = \mathbf{b}$ coincides with the nonempty set of solutions of the normal equations $A^T A\mathbf{x} = A^T\mathbf{b}$.

↳ normal equation for $Ax = b$.

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

a. The equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution for each $\mathbf{b}$ in $\mathbb{R}^m$.

b. The columns of A are linearly indpendent.

c. The matrix $A^T A$ is invertible.

When these statements are true, the least-squares solution $\hat{\mathbf{x}}$ is given by

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$$

Example 1: Find a least-squares solution of $A\mathbf{x} = \mathbf{b}$ by (a) constructing the normal equations for $\hat{\mathbf{x}}$ and (b) solving for $\hat{\mathbf{x}}$.

$$A = \begin{bmatrix} 2 & 1 \\ -2 & 0 \\ 2 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -5 \\ 8 \\ 1 \end{bmatrix}$$

Step 1 (a)

Normal equation: $A^t A \cdot x = A^t b$

$$A^t A = \begin{pmatrix} 2 & -2 & 2 \\ 1 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -2 & 0 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 12 & 8 \\ 8 & 10 \end{pmatrix}$$

$$A^t b = \begin{pmatrix} 2 & -2 & 2 \\ 1 & 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} -5 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} -24 \\ -2 \end{pmatrix}$$

$$\begin{pmatrix} 12 & 8 \\ 8 & 10 \end{pmatrix} x = \begin{pmatrix} -24 \\ -2 \end{pmatrix}$$

step 2 (b)

$$\hat{x} = (A^t A)^{-1} A^t b = \frac{1}{56} \begin{pmatrix} 10 & -8 \\ -8 & 12 \end{pmatrix} \cdot \begin{pmatrix} -24 \\ -2 \end{pmatrix}$$

$$= \begin{pmatrix} -4 \\ 3 \end{pmatrix}.$$

by the thm, $\hat{x}$ is the least square solution.

Example 2: Describe all least-squares solutions of the equation $A\mathbf{x} = \mathbf{b}$.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 8 \\ 2 \end{bmatrix}$$

See the next page !

Theorem: Given an $m \times n$ matrix A with linearly independent columns, let $A = QR$ be a QR factorization of A. Then, for each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution, given by

$$\hat{\mathbf{x}} = (R)^{-1} Q^T \mathbf{b}$$

Example 3: Use the factorization $A = QR$ to find the least-squares solution of $A\mathbf{x} = \mathbf{b}$.

$$A = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2/3 & -1/3 \\ 2/3 & 2/3 \\ 1/3 & -2/3 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 7 \\ 3 \\ 1 \end{bmatrix}$$

the first col of A = the 2^{nd} + the 3^{rd}

$\Rightarrow$ the cols of A are not linear indep.

$$A^t A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

$$A^t b = \begin{pmatrix} 14 \\ 4 \\ 10 \end{pmatrix}$$

Find x such that

$$\underbrace{\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}}_{A^t A \text{ is not invertible.}} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 14 \\ 4 \\ 10 \end{pmatrix}$$

Ref:

$$[A^t A, A^t b] = \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

free

$x_3 = s$

from the 2^{nd} row $\Rightarrow$ $x_2 - s = -3$

$x_2 = s-3$

from the 1^{st} row $\Rightarrow$ $x_1 + s = 5$

$\Rightarrow$ $x_1 = -s + 5$

Finally $x = \begin{pmatrix} -s+5 \\ s-3 \\ s+0 \end{pmatrix} = \begin{pmatrix} -s \\ s \\ s \end{pmatrix} + \begin{pmatrix} 5 \\ -3 \\ 0 \end{pmatrix}$

$= s\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 5 \\ -3 \\ 0 \end{pmatrix}$

all vectors of $s\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 5 \\ -3 \\ 0 \end{pmatrix}$ are the least-squares solution of $Ax = b$. //

Eg 3. Find a least-squares solution of $Ax = b$ for

$$A = \begin{pmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{pmatrix} \quad b = \begin{pmatrix} -1 \\ 2 \\ 1 \\ 6 \end{pmatrix}$$

$$\underbrace{}_{\vec{a}_1} \quad \underbrace{}_{\vec{a}_2}$$

$\vec{a}_1 \cdot \vec{a}_2 = 0 \iff \vec{a}_1$ is orthogonal to $\vec{a}_2$.

what is $A\hat{x}$? $\quad A\hat{x} = \text{Proj}_{col(A)} b$

$\vec{a}_1$ & $\vec{a}_2$ is just an orthogonal basis of $col(A)$
$= \text{span}\{\vec{a}_1, \vec{a}_2\}$

by thm in 6.3.

$$\Rightarrow \quad \hat{b} = \frac{\vec{b} \cdot \vec{a}_1}{\vec{a}_1 \cdot \vec{a}_1} a_1 + \frac{b \cdot \vec{a}_2}{\vec{a}_2 \cdot \vec{a}_2} a_2 = 2\vec{a}_1 + \frac{1}{2}\vec{a}_2$$

$\text{Proj}_{col(A)} b$

$$= \begin{pmatrix} -1 \\ 1 \\ 5/2 \\ 11/2 \end{pmatrix}$$

$A\hat{x} = \hat{b}$, do we really need to solve for $\hat{x}$?

$$\hat{b} = 2\vec{a}_1 + \frac{1}{2}\vec{a}_2 = \underbrace{\begin{pmatrix} \vec{a}_1 & \vec{a}_2 \end{pmatrix}}_{A} \cdot \underbrace{\begin{pmatrix} 2 \\ \frac{1}{2} \end{pmatrix}}_{\hat{x}}$$

NO !

Example 2: Describe all least-squares solutions of the equation $A\mathbf{x} = \mathbf{b}$.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 8 \\ 2 \end{bmatrix}$$

Theorem: Given an $m \times n$ matrix A with linearly independent columns, let $A = QR$ be a QR factorization of A. Then, for each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution, given by

$$\hat{\mathbf{x}} = (R)^{-1} Q^T \mathbf{b} \quad \Leftrightarrow \quad R\hat{x} = Q^{t} b$$

Example 3: Use the factorization $A = QR$ to find the least-squares solution of $A\mathbf{x} = \mathbf{b}$.

$$A = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 2/3 & -1/3 \\ 2/3 & 2/3 \\ 1/3 & -2/3 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 3 & 5 \\ 0 & 1 \end{bmatrix}}_{R}, \mathbf{b} = \begin{bmatrix} 7 \\ 3 \\ 1 \end{bmatrix}$$

$$R = \begin{pmatrix} 3 & 5 \\ 0 & 1 \end{pmatrix}$$

$$Q^t b = \begin{pmatrix} 7 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 5 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ -1 \end{pmatrix}$$

$$x_2 = -1$$

$$3x_1 + 5x_2 = 7$$

$$\Rightarrow x_1 = \frac{1}{3}(7 - 5x_2)$$

$$[R, Q^t b] = \begin{pmatrix} 3 & 5 & 7 \\ 0 & 1 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & -1 \end{pmatrix}$$

$$\Rightarrow \begin{cases} x_2 = -1 \\ x_1 = 4 \end{cases} \quad \Rightarrow \hat{x} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$$