Example: Suppose $Q^T AQ = D$. Solve for A then use the outer product expansion to derive the spectral decomposition.

$$Q^T A Q = D$$

$$\underbrace{\cancel{Q Q^T}}_{I} A Q = Q D$$

$$A Q = Q D$$

$$A \underbrace{\cancel{Q Q^T}}_{I} = Q D Q^T$$

$$A = Q D Q^T$$

$$= \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{q}_1^{\,T} \\ \vec{q}_2^{\,T} \\ \vdots \\ \vec{q}_n^{\,T} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \vec{q}_1 & \lambda_2 \vec{q}_2 & \cdots & \lambda_n \vec{q}_n \end{bmatrix} \begin{bmatrix} \vec{q}_1^{\,T} \\ \vec{q}_2^{\,T} \\ \vdots \\ \vec{q}_n^{\,T} \end{bmatrix}$$

$$= \lambda_1 \vec{q}_1 \vec{q}_1^{\,T} + \lambda_2 \vec{q}_2 \vec{q}_2^{\,T} + \dots + \lambda_n \vec{q}_n \vec{q}_n^{\,T}$$

7.3 Least Squares Approximation

Recall that a system $A\vec{x} = \vec{b}$ may be inconsistent. ← unsolvable

Definition: Given an approximate solution $\vec{s}$, the **error vector** is $\vec{b} - A\vec{s}$ and the **error** is $||\vec{b} - A\vec{s}||$.

Definition: The **least squares solution** $\vec{x}^*$ is the approximate solution with the minimum error.

Comment: Recall that $||\vec{v}|| = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2}$. The terminology **least squares solution** emphasizes that we're making the length of the error vector as small as possible.

Fact: The **least squares solution** to a system $A\vec{x} = \vec{b}$ is $\vec{x}^* = (A^T A)^{-1} A^T \vec{b}$.

Comment: We'll assume that the columns of A are linearly independent so that $(A^T A)^{-1}$ exists.

Example: The system $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}$ is inconsistent.

Find the least squares solution $\vec{x}^*$.

$$\vec{x}^* = (A^TA)^{-1}A^T\vec{b}$$

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

$$(A^TA)^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\vec{x}^* = (A^TA)^{-1}A^T\vec{b}$$
$$= \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}}$$
$$= \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 7 \\ 8 \end{bmatrix}$$
$$= \frac{1}{3}\begin{bmatrix} 6 \\ 9 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Example: Calculate the error for $\vec{x}^*$ above. What can you say about the error for any other vector $\vec{x}$?

$$\text{error} = \|\vec{b} - A\vec{x}^*\|$$

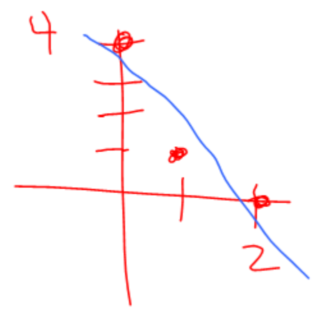
$$\vec{b} - A\vec{x}^* = \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$
$$= \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$$
$$= \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

$$\text{error} = \left\|\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}\right\| = \sqrt{3}$$

For any other $\vec{x}$, error $> \sqrt{3}$

Example: Find the best-fit line $y = a_0 + a_1 x$.
The best-fit line is also called the **least squares regression line**.

x	y
0	4
1	1
2	0



a_0, a_1 are the unknowns

$$a_0 + a_1 x = y$$

$$1(a_0) + x\, a_1 = y$$

$$\begin{array}{cc} 1 & x \\ \end{array}\qquad\qquad y$$

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix}$$

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

$$(A^T A)^{-1} = \frac{1}{6} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix}$$

$$\vec{x}^* = (A^T A)^{-1} A^T \vec{b}$$

$$= \frac{1}{6} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix}}$$

$$= \frac{1}{6} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

$$= \frac{1}{6} \begin{bmatrix} 22 \\ -12 \end{bmatrix} \begin{array}{l} \Leftarrow a_0 = \frac{22}{6} = \frac{11}{3} \\ \Leftarrow a_1 = \frac{-12}{6} = -2 \end{array}$$

$$\boxed{y = \frac{11}{3} - 2x}$$