9. The solution of a least-squares problem

- geometric interpretation
- left inverse of a zero nullspace matrix
- the solution of a least-squares problem
- the normal equations

Geometric interpretation of a LS problem

$$\text{minimize} \quad \|Ax - b\|^2$$

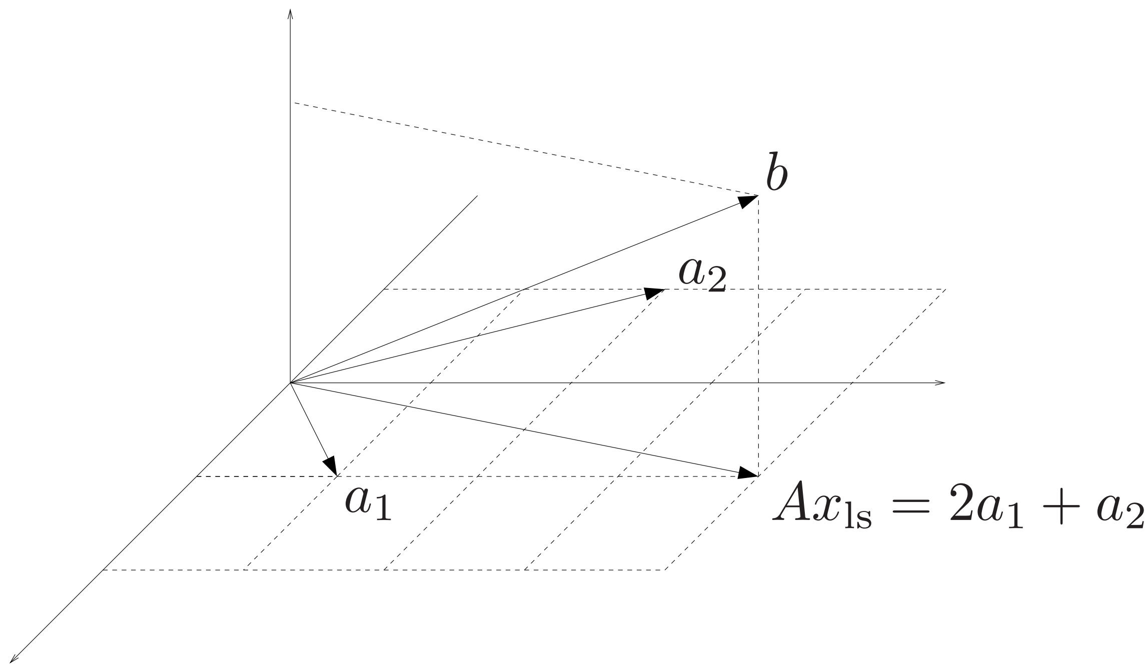
A is $m \times n$ with columns $a_1, \ldots, a_n$

- $\|Ax - b\|$ is the distance of b to the vector

$$Ax = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$$

- solution x_{ls} gives the linear combination of the columns of A closest to b
- Ax_{ls} is the **projection** of b on the range of A

Example: $A = \begin{bmatrix} 1 & -1 \\ 1 & 2 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$



least-squares solution $x_{\rm ls}$

$$Ax_{\rm ls} = \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix}, \qquad x_{\rm ls} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Properties of zero nullspace matrices

suppose A is $m \times n$ with a zero nullspace ($Ax = 0$ only if $x = 0$)

- A^TA is positive definite: if $x \neq 0$, then

$$x^TA^TAx = (Ax)^T(Ax) = \|Ax\|^2 > 0$$

- $X = (A^TA)^{-1}A^T$ is a left inverse of A

- $m \geq n$; to show this, assume $m < n$ and partition $XA = I$ as

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \begin{bmatrix} A_1 & A_2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \qquad \text{(with } X_1 \text{ and } A_1 \ m \times m\text{)}$$

  this is impossible: $X_1A_1 = I$, $X_1A_2 = 0$ imply $X_1 = A_1^{-1}$ and $A_2 = 0$, but then $X_2A_2 = 0 \neq I$

The solution of a least-squares problem

if A has a zero nullspace, then

$$x_{\text{ls}} = (A^T A)^{-1} A^T b$$

is the unique solution of the least-squares problem

$$\text{minimize} \quad \|Ax - b\|^2$$

in other words, if $x \neq x_{\text{ls}}$, then $\|Ax - b\|^2 > \|Ax_{\text{ls}} - b\|^2$

Proof

we have already seen that A^TA is positive definite, hence invertible

we show that $\|Ax - b\|^2 > \|Ax_{\text{ls}} - b\|^2$ for $x \neq x_{\text{ls}}$:

$$\begin{aligned}
\|Ax - b\|^2 &= \|A(x - x_{\text{ls}}) + (Ax_{\text{ls}} - b)\|^2 \\
&= \|A(x - x_{\text{ls}})\|^2 + \|Ax_{\text{ls}} - b\|^2 \\
&> \|Ax_{\text{ls}} - b\|^2
\end{aligned}$$

- 2nd step follows from $A(x - x_{\text{ls}}) \perp (Ax_{\text{ls}} - b)$:

$$(A(x - x_{\text{ls}}))^T(Ax_{\text{ls}} - b) = (x - x_{\text{ls}})^T(A^TAx_{\text{ls}} - A^Tb) = 0$$

- 3rd step follows from $x \neq x_{\text{ls}} \Longrightarrow A(x - x_{\text{ls}}) \neq 0$

The normal equations

$$(A^T A)x = A^T b$$

if A has a zero nullspace:

- least-squares solution can be found by solving the normal equations
- n equations in n variables with a positive definite coefficient matrix
- can be solved using Cholesky factorization