10. Least-norm problems

- definition and examples
- the least-norm solution
- computing the least-norm solution

Definition

underdetermined linear equations

$$Ax = b \qquad (A \text{ is } m \times n \text{ with } m < n)$$

- x is underspecified; many choices of x lead to the same b
- if $Ax = b$ is solvable $\left(b \in \mathbf{range}(A)\right)$, there are infinitely many solutions

least-norm solution: solution of $Ax = b$ with the smallest norm $\|x\|$

least-norm problem

$$\begin{array}{ll} \text{minimize} & \|x\| \\ \text{subject to} & Ax = b \end{array}$$

Control example

mass-force example of page 1-10

- unit mass, with zero position and velocity at $t = 0$, subject to force $F(t)$
- $F(t) = x_j$ for $j - 1 \leq t < j$, $j = 1, \ldots, 10$ (x is sequence of forces)
- final position and velocity (at $t = 10$) are given by $y = Ax$ where

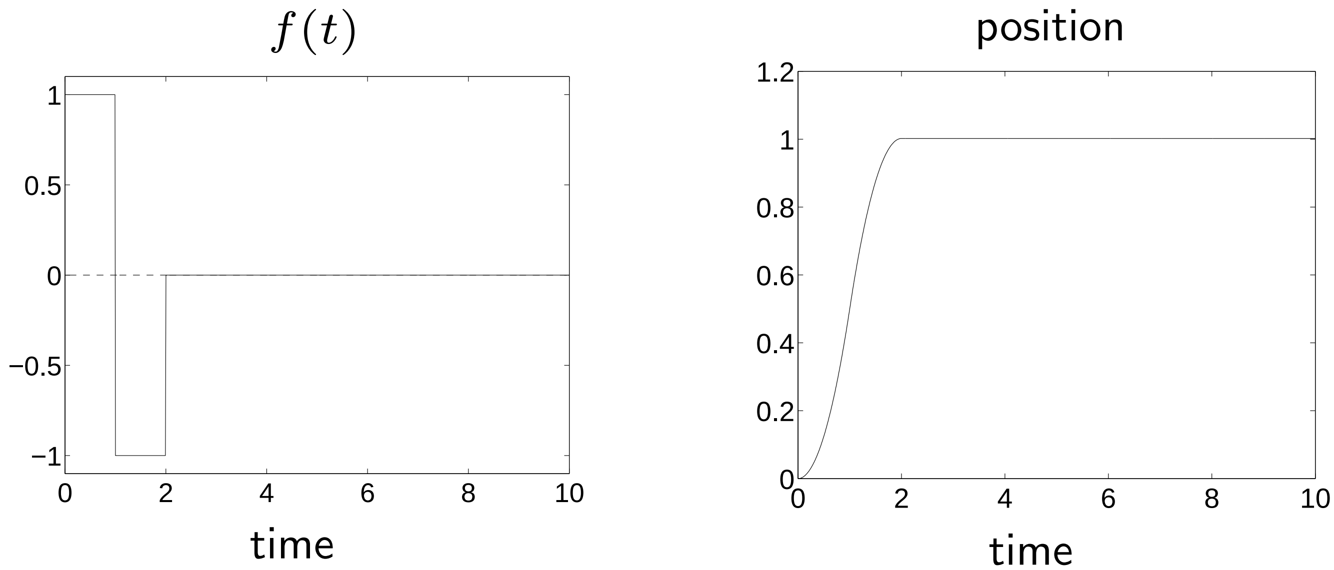
$$A = \begin{bmatrix} 19/2 & 17/2 & 15/2 & \cdots & 1/2 \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}$$

forces x that transfer mass a unit distance with zero final velocity satisfy

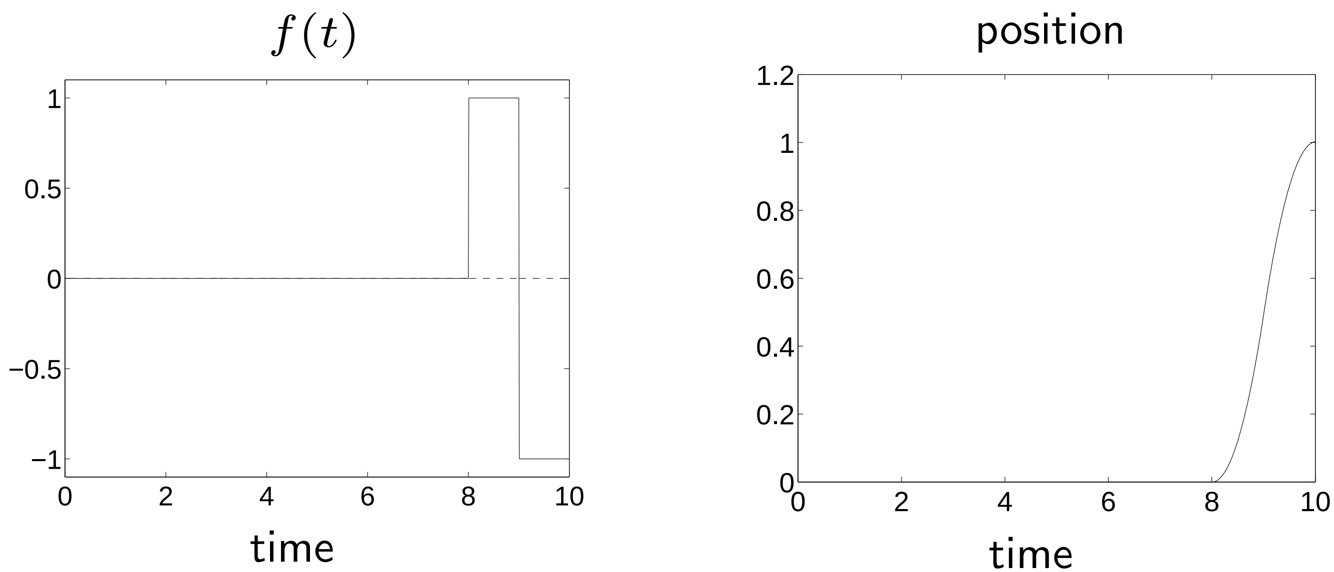
$$\begin{bmatrix} 19/2 & 17/2 & 15/2 & \cdots & 1/2 \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

two equations in ten variables

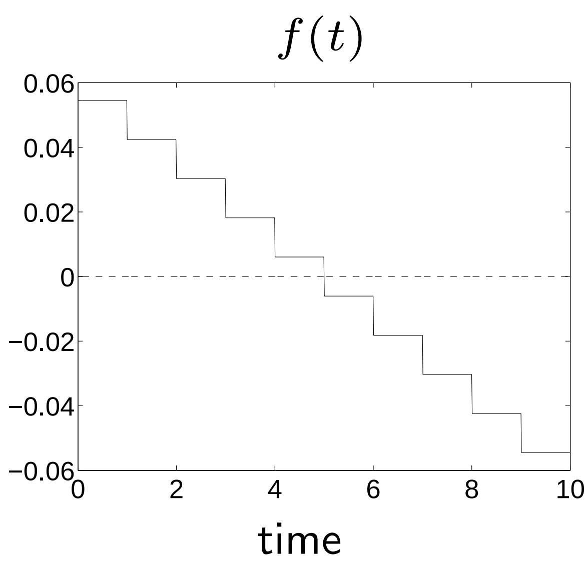
- choose $x_1 = 1$, $x_2 = -1$, $x_3 = \cdots = x_{10} = 0$

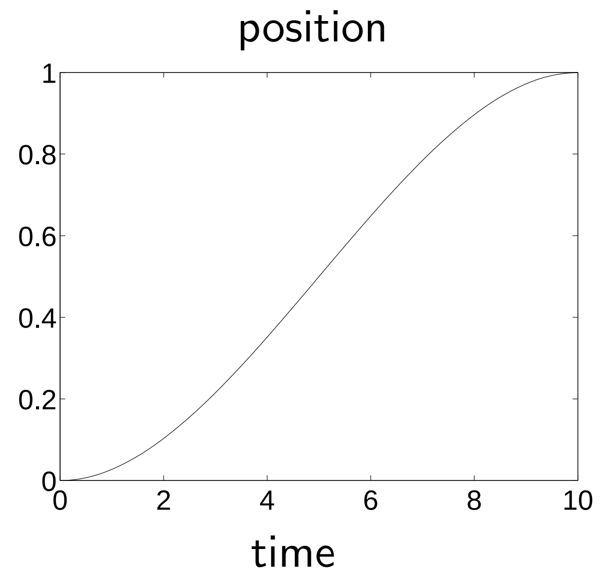


- choose $x_1 = \cdots = x_8 = 0$, $x_9 = 1$, $x_{10} = -1$



• x with smallest $\|x\|^2 = x_1^2 + x_2^2 + \cdots + x_{10}^2$ ('least-norm' solution)



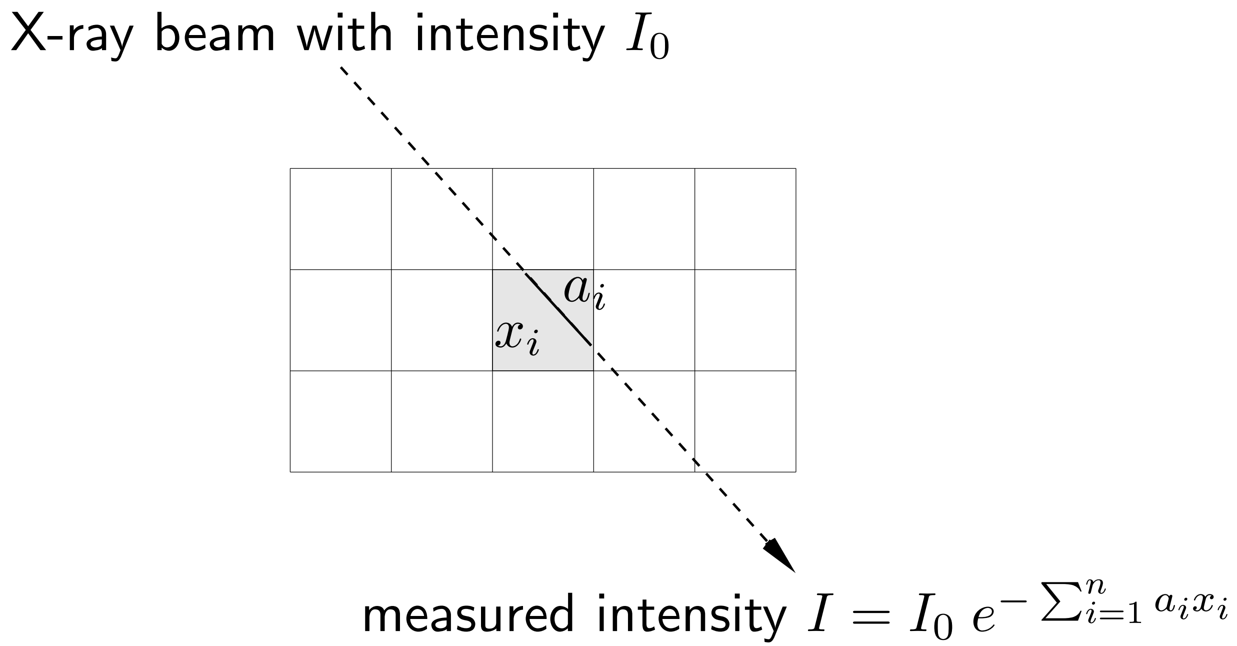


least-norm solution has $\|x\| = 0.110$

Image reconstruction from projections

problem: estimate density of a 3-dimensional object

partition object in n small elements ('pixels' or 'voxels') and assume density is constant on each element



x_i: (unknown) density in element i;
a_i: (known) length of path through element i (up to a physical constant)

after taking logs:

$$\sum_{i=1}^{n} a_i x_i = \log(I_0/I)$$

one linear equation in the variables x_i

repeat with m different source locations and beam directions:

$$Ax = b$$

- m linear equations in n variables
- usually $m \ll n$; many solutions x

a popular method: pick a solution x, consistent with the measurements, and close to a prior guess

$$\begin{array}{ll} \text{minimize} & \|x - \hat{x}\| \\ \text{subject to} & Ax = b \end{array}$$

where $\hat{x}$ is a prior guess of the density ($e.g.$, uniform density)

The least-norm solution

if A is right-invertible, then

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

is the unique solution of the least-norm problem

$$\begin{array}{ll} \text{minimize} & \|x\| \\ \text{subject to} & Ax = b \end{array}$$

- in other words if $Ax = b$, $x \neq x_{\text{ln}}$, then $\|x\| > \|x_{\text{ln}}\|$
- recall from page 4-27 that AA^T is positive definite and that

$$A^T(AA^T)^{-1}$$

  is a right inverse of A

proof

1. x_{ln} satisfies the equation

$$Ax_{\mathrm{ln}} = AA^T(AA^T)^{-1}b = b$$

2. suppose $Ax = b$, $x \neq x_{\mathrm{ln}}$

$$\begin{aligned}(x - x_{\mathrm{ln}})^T x_{\mathrm{ln}} &= (x - x_{\mathrm{ln}})^T A^T (AA^T)^{-1} b \\ &= (Ax - Ax_{\mathrm{ln}})^T (AA^T)^{-1} b \\ &= 0\end{aligned}$$

therefore

$$\begin{aligned}\|x\|^2 &= \|x - x_{\mathrm{ln}} + x_{\mathrm{ln}}\|^2 \\ &= \|x - x_{\mathrm{ln}}\|^2 + 2(x - x_{\mathrm{ln}})^T x_{\mathrm{ln}} + \|x_{\mathrm{ln}}\|^2 \\ &= \|x - x_{\mathrm{ln}}\|^2 + \|x_{\mathrm{ln}}\|^2 \\ &> \|x_{\mathrm{ln}}\|^2\end{aligned}$$

Computing the least-norm solution

$$x_{\mathrm{ln}} = A^T(AA^T)^{-1}b$$

method 1 (Cholesky factorization)

1. form $C = AA^T$ (m^2n flops)
2. Cholesky factorization $C = LL^T$ ($(1/3)m^3$ flops)
3. forward substitution: solve $Lw = b$ (m^2 flops)
4. back substitution: solve $L^Tz = w$ (m^2 flops)
5. multiply with A^T: $x = A^Tz$ ($2mn$ flops)

total cost for large m, n

$$m^2n + (1/3)m^3 \text{ flops}$$

method 2 (QR factorization)

QR factorization of A^T: $A^T = QR$ (R upper triangular, $Q^TQ = I$)

$$\begin{aligned} x_{\mathrm{ln}} &= A^T(AA^T)^{-1}b \\ &= QR(R^TQ^TQR)^{-1}b \\ &= QR(R^TR)^{-1}b \\ &= QR^{-T}b \end{aligned}$$

1. QR factorization of $A^T = QR$ ($2m^2n$ flops)
2. forward substitution to solve $R^Tz = b$ (m^2 flops)
3. form $x = Qz$ ($2mn$ flops)

total cost for large m, n:

$$2m^2n \text{ flops}$$

more accurate than method 1; up to twice as expensive if $n \gg m$

Example

$$A = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 0 & 1/2 & 1/2 \end{bmatrix}, \qquad b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Cholesky factorization method: solve $AA^T z = b$ and evaluate $x = A^T z$

- Cholesky factorization $AA^T = LL^T$

$$\begin{bmatrix} 4 & 2 \\ 2 & 3/2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 1/\sqrt{2} \end{bmatrix}$$

- solve $Lw = b$

$$\begin{bmatrix} 2 & 0 \\ 1 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$w_1 = 0$, $w_2 = \sqrt{2}$

- solve $L^T z = w$

$$\begin{bmatrix} 2 & 1 \\ 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}$$

$z_1 = -1$, $z_2 = 2$

- evaluate $x = A^T z$

$$x = \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 1 & 1/2 \\ 1 & 1/2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

QR factorization method

- QR factorization $A^T = QR$

$$\begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 1 & 1/2 \\ 1 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/\sqrt{2} \\ -1/2 & 1/\sqrt{2} \\ 1/2 & 0 \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 1/\sqrt{2} \end{bmatrix}$$

- solve $R^T z = b$

$$\begin{bmatrix} 2 & 0 \\ 1 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

  $z_1 = 0$, $z_2 = \sqrt{2}$

- evaluate $x = Qz = (1, 1, 0, 0)$