

---

---

---

---

---



## Precept 4

### Agenda

- Approx solns of linear eqns w/ multiple RHS
- Varying the RHS in linearly constrained LS
- Nearest vector w/ given average
- Multi-Objective Control } textbook
- Rendez-vous ? coding

} written

### Logistics

- HW1
  - solns on Canvas
  - grades sent out
- HW 3
  - due Friday Feb 18 at 9pm
- Midterm Thursday March 3 (16 days)

• Approximate solution of linear equations w/ multiple RHS's

- Suppose  $A$  is invertible
- Seek  $x$  s.t.  $Ax \approx b_i; i=1,\dots,k$
- Bob proposes a multi-objective least squares approach  

$$x^{Bob} = \arg \min_x \|Ax - b_1\|^2 + \dots + \|Ax - b_k\|^2$$
- Alice proposes to first average the  $b$ 's to get  
 $b = \frac{1}{k}(b_1 + \dots + b_k)$ . Then choose  $x$  by solving one system of linear equations.

$$Ax^{Alice} = b$$

- Oscar says that the 2 approaches sound different, but  $x^{Bob} = x^{Alice}$ . Is Oscar right?

Alice's approach

$$b = \frac{1}{k}(b_1 + \dots + b_k)$$

$$Ax^{Alice} = \frac{1}{k}(b_1 + \dots + b_k)$$

$\xrightarrow{\text{normal eqns}}$   $A^T A x^{Alice} = A^T \left( \frac{1}{k}(b_1 + \dots + b_k) \right)$

$$x^{Alice} = (A^T A)^{-1} A^T \left( \frac{1}{k}(b_1 + \dots + b_k) \right)$$

Bob's approach: multi-objective

- stack the  $A$ s and  $b$ s to create a single-obj. problem

$$x^{\text{Bob}} = \underset{x}{\operatorname{argmin}} \quad \| \tilde{A}x - \tilde{b} \|^2$$

$$\tilde{A} = \begin{bmatrix} A \\ \vdots \\ A \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$$

columns are  
lin. ind.  $\rightarrow$

$$\text{normal eqns} \Rightarrow \tilde{A}^T \tilde{A} x^{\text{Bob}} = \tilde{A}^T \tilde{b}$$

$$\tilde{A}^T \tilde{A} = [A^T | \dots | A^T] \begin{bmatrix} A \\ \vdots \\ A \end{bmatrix} = \sum_{i=1}^k A^T A = k A^T A$$

$$\tilde{A}^T \tilde{b} = [A^T | \dots | A^T] \begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix} = \sum_{i=1}^k A^T b_i$$

$$\Rightarrow k A^T A x^{\text{Bob}} = \sum_{i=1}^k A^T b_i$$

$$A^T A x^{\text{Bob}} = A^T \left( \frac{1}{k} \sum_{i=1}^k b_i \right)$$

$$x^{\text{Bob}} = (A^T A)^{-1} A^T \left( \frac{1}{k} \sum_{i=1}^k b_i \right)$$

$$x^{\text{Alice}} = (A^T A)^{-1} A^T \left( \frac{1}{k} (b_1 + \dots + b_k) \right)$$

Oscar was correct

- Varying the RHS in linearly constrained least squares

- Suppose  $\hat{x} = \underset{x}{\operatorname{argmin}} \|Ax - b\|^2$   
s.t.  $Cx = d$

- Assume KKT matrix is invertible

- Think of  $\hat{x}$  as a fn of  $b$  and  $d$

- A colleague asserts that  $\hat{x}$  is a linear fn of  $b$  and  $d$   
and hence, has the form  $\hat{x} = Fb + Gd$  for some  
appropriate matrices  $F$  and  $G$ .

- Is she right? If not, give a counter-example.  
If so, justify and explain how to find  $F$  and  $G$ .

KKT optimality eqns from lecture

$$\begin{pmatrix} 2A^TA & C^T \\ C & 0 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix} = \begin{pmatrix} 2A^Tb \\ d \end{pmatrix}$$

$$\begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix} = \begin{pmatrix} 2A^TA & C^T \\ C & 0 \end{pmatrix}^{-1} \begin{pmatrix} 2A^Tb \\ d \end{pmatrix}$$

$$\text{Let } H = \begin{pmatrix} 2A^TA & C^T \\ C & 0 \end{pmatrix}^{-1} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$$

$$\begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} 2A^Tb \\ d \end{pmatrix}$$

$$\hat{x} = 2H_{11}A^Tb + H_{12}d$$

Yes it is linear

$$F = 2H_{11}A^T, G = H_{12}$$

• Nearest vector w/ a given average

- $a \in \mathbb{R}^n, \beta \in \mathbb{R}$
- find  $x$  that is closest to  $a$  among all  $n$ -vectors that have an average value of  $\beta$

frame it as a constrained least squares problem

$$\begin{aligned} \min_x \quad & \|x - a\|_2^2 \\ \text{s.t.} \quad & \mathbf{1}^T x = n\beta \end{aligned}$$

can form the KKT system

$$\begin{pmatrix} 2A^TA & C^T \\ C & \mathbf{0} \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix} = \begin{pmatrix} 2A^Tb \\ d \end{pmatrix}$$

here  $A = I$ ,  $b = a$ ,  $C = \mathbf{1}^T$ ,  $d = n\beta$

$$\begin{pmatrix} 2I & \mathbf{1}^T \\ \mathbf{1}^T & 0 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix} = \begin{pmatrix} 2a \\ n\beta \end{pmatrix}$$

$$\begin{cases} 2\hat{x} + \mathbf{1}^T \hat{z} = 2a & \text{now try to eliminate } \hat{z} \\ \mathbf{1}^T \hat{x} = n\beta \end{cases}$$

$$\hat{x} = a - \frac{1}{2} \mathbf{1}^T \hat{z}$$

$$\mathbf{1}^T \hat{x} = \mathbf{1}^T \left( a - \frac{1}{2} \mathbf{1}^T \hat{z} \right) = n\beta$$

$$\mathbf{1}^T a - \frac{1}{2} \mathbf{1}^T \mathbf{1}^T \hat{z} = n\beta$$

$$\mathbf{1}^T a - \frac{n}{2} \hat{z} = n\beta$$

$$\frac{1}{2} \hat{z} = \text{avg}(a) - \beta \quad \hat{z} = 2(\text{avg}(a) - \beta)$$

$$2\hat{x} + \mathbb{1}\hat{z} = 2a$$

$$\hat{z} = 2(\text{avg}(a) - \beta)$$

$$2\hat{x} + 2\mathbb{1}(\text{avg}(a) - \beta) = 2a$$

$$\hat{x} = a - \mathbb{1}(\text{avg}(a) - \beta)$$

$$\hat{x} = (a - \mathbb{1}\text{avg}(a)) + \mathbb{1}\beta$$

check:  $\text{avg}(\hat{x}) = \beta$  ✓

In words: take  $a$ , subtract the mean, then add  $\beta$

example:  $a = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix} \quad \beta = 4$

$$\hat{x} = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix} - \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix} + \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix}$$

$$= \begin{pmatrix} -3 \\ 0 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix}$$

We see that component-wise, the distance between  $\hat{x}$  and  $a$  is the same