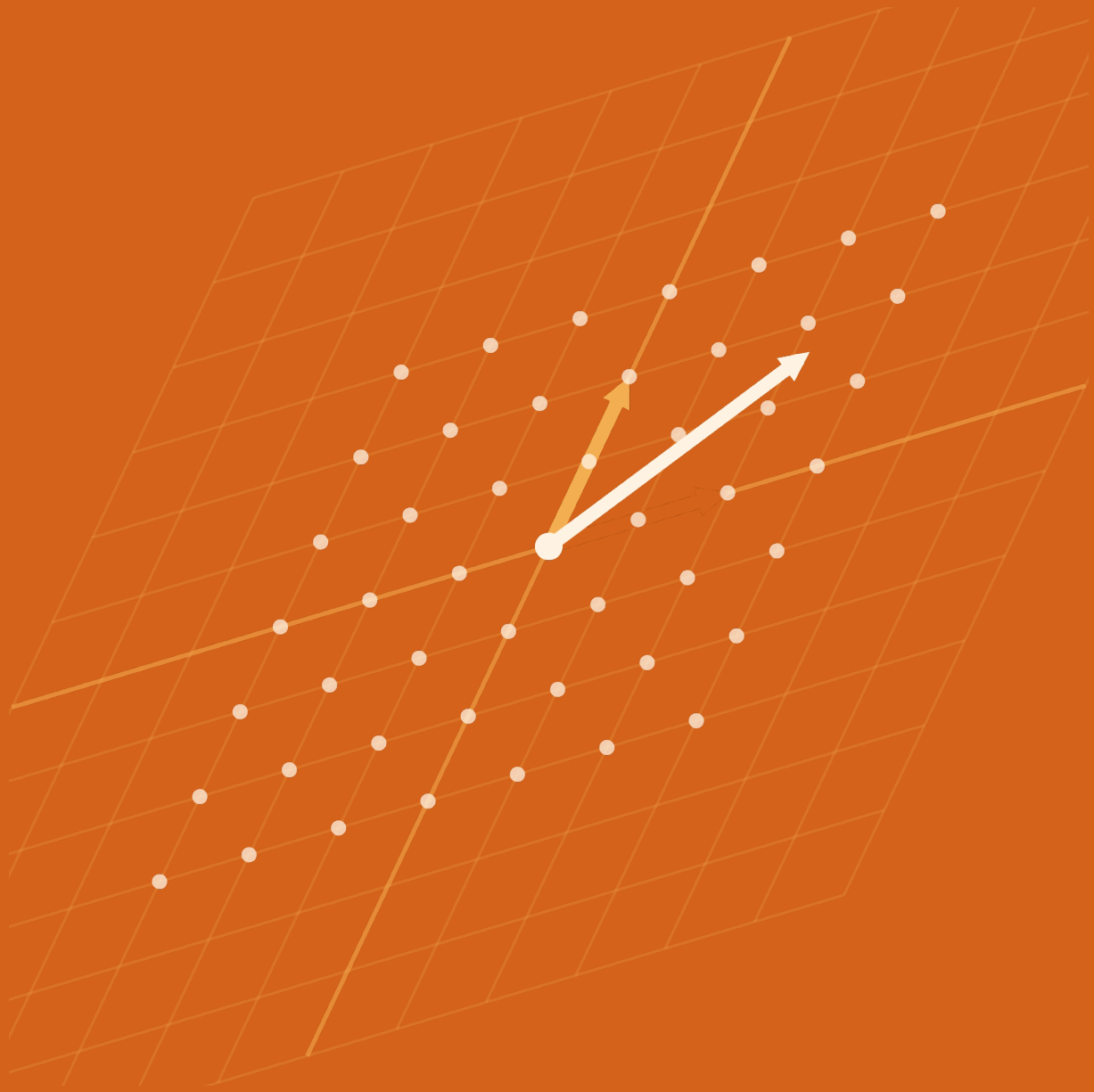
LINEAR ALGEBRA FOR AIML AND 6G

Least Norm Solution



Lecture 22

Prof. Aditya K. Jagannatham
IIT Kanpur

Outline

1. **Least norm problem**
   Underdetermined systems and the minimum-norm solution
2. **Lagrange multiplier solution**
   Solving the constrained problem with multipliers
3. **Geometric intuition**
   Splitting x into row-space and null-space parts
4. **Worked example**
   Least norm solution of a 2 × 4 system

01

Least norm problem

Underdetermined systems and the minimum-norm solution



Least norm problem

- Consider a matrix $\mathbf{A}$ of size $m \times n$ with $n > m$
  - Such a matrix is termed a **wide matrix**.

$$\text{rank}\,(\mathbf{A}) = m$$

Least norm problem

- Consider the problem

$$\bar{\mathbf{b}} = \mathbf{A}\bar{\mathbf{x}}$$

- This is an **under-determined** system
- **Nullity** $= n - m =$ number of free variables
  - This has **infinitely** many solutions

Least norm problem

- Among these, we find the solution with **minimum norm**
- This can be formulated as the problem below

$$\min \|\bar{\mathbf{x}}\|$$
$$\text{subject to } \bar{\mathbf{b}} = \mathbf{A}\bar{\mathbf{x}}$$

Least norm problem

$$\min \|\bar{\mathbf{x}}\|$$
$$\text{subject to } \bar{\mathbf{b}} = \mathbf{A}\bar{\mathbf{x}}$$

- The above is a **constrained optimization** problem
- There are m constraints
  - One corresponding to each equation in the system $\bar{\mathbf{b}} = \mathbf{A}\bar{\mathbf{x}}$.

02

Lagrange multiplier solution

Solving the constrained problem with multipliers



Least norm solution

- Therefore, we use m **Lagrange multipliers**

$$\bar{\boldsymbol{\lambda}} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix}$$

Least norm solution

- The **Lagrangian** can be formulated as

$$f = \|\bar{\mathbf{x}}\|^2 + \bar{\boldsymbol{\lambda}}^T \left(\bar{\mathbf{b}} - \mathbf{A}\bar{\mathbf{x}}\right)$$

Least norm solution

- The gradient of the **Lagrangian** is

$$\frac{\partial f}{\partial \bar{\mathbf{x}}} = \frac{\partial}{\partial \bar{\mathbf{x}}} \bar{\mathbf{x}}^T \bar{\mathbf{x}} + \frac{\partial}{\partial \bar{\mathbf{x}}} \left(\bar{\boldsymbol{\lambda}}^T \bar{\mathbf{b}} - \bar{\boldsymbol{\lambda}}^T \mathbf{A} \bar{\mathbf{x}} \right)$$

$$= 2\bar{\mathbf{x}} - \left(\bar{\boldsymbol{\lambda}}^T \mathbf{A} \right)^T$$

$$= 2\bar{\mathbf{x}} - \mathbf{A}^T \bar{\boldsymbol{\lambda}} = \mathbf{0}$$

$$\Rightarrow \bar{\mathbf{x}} = \frac{1}{2} \mathbf{A}^T \bar{\boldsymbol{\lambda}}$$

Constraint

- We now use the constraint to determine $\bar{\mathbf{x}}$.

$$\bar{\mathbf{b}} = \mathbf{A}\bar{\mathbf{x}} = \mathbf{A}\frac{1}{2}\mathbf{A}^T\bar{\boldsymbol{\lambda}}$$

$$\Rightarrow \bar{\boldsymbol{\lambda}} = 2\left(\mathbf{A}\mathbf{A}^T\right)^{-1}\bar{\mathbf{b}}$$

$$\Rightarrow \bar{\mathbf{x}} = \frac{1}{2}\mathbf{A}^T\bar{\boldsymbol{\lambda}} = \frac{1}{2}\mathbf{A}^T 2\left(\mathbf{A}\mathbf{A}^T\right)^{-1}\bar{\mathbf{b}}$$

$$= \mathbf{A}^T\left(\mathbf{A}\mathbf{A}^T\right)^{-1}\bar{\mathbf{b}}$$

Least norm solution

- Thus, the **least norm solution** is

$$\bar{\mathbf{x}} = \mathbf{A}^T \left(\mathbf{A}\mathbf{A}^T\right)^{-1} \bar{\mathbf{b}}$$

03

Geometric intuition

Splitting x into row-space and null-space parts



Least norm intuition

- The **least norm solution** can be intuitively understood as follows
- Consider

$$\bar{\mathbf{x}} = \bar{\mathbf{x}}_{\|} + \bar{\mathbf{x}}_{\perp}$$

Least norm intuition

- The vector $\bar{\mathbf{x}}_{\|}$ is in the **column space** of $\mathbf{A}^T$
- Therefore,

$$\bar{\mathbf{x}}_{\|} = \mathbf{A}^T \bar{\boldsymbol{\lambda}}$$

Least norm intuition

- The vector $\bar{\mathbf{x}}_\perp$ lies in the space **perpendicular** to the column space of $\mathbf{A}^T$
  - Therefore, it lies in the **null space** of $\mathbf{A}$
- Hence,

$$\mathbf{A}\bar{\mathbf{x}}_\perp = \mathbf{0}$$

Least norm intuition

- Consider now the condition

$$\bar{\mathbf{b}} = \mathbf{A}\bar{\mathbf{x}} = \mathbf{A}\left(\bar{\mathbf{x}}_{\|} + \bar{\mathbf{x}}_{\perp}\right)$$
$$= \mathbf{A}\bar{\mathbf{x}}_{\|} + \mathbf{A}\bar{\mathbf{x}}_{\perp}$$
$$= \mathbf{A}\bar{\mathbf{x}}_{\|} = \mathbf{A}\mathbf{A}^{T}\bar{\boldsymbol{\lambda}}$$
$$\Rightarrow \bar{\boldsymbol{\lambda}} = \left(\mathbf{A}\mathbf{A}^{T}\right)^{-1}\bar{\mathbf{b}}$$
$$\Rightarrow \bar{\mathbf{x}}_{\|} = \mathbf{A}^{T}\left(\mathbf{A}\mathbf{A}^{T}\right)^{-1}\bar{\mathbf{b}}$$

Least norm intuition

- It can be seen that

$$\|\bar{\mathbf{x}}\|^2 = \left\|\bar{\mathbf{x}}_{\|}\right\|^2 + \|\bar{\mathbf{x}}_{\perp}\|^2$$

- Since, $\mathbf{A}\bar{\mathbf{x}}_{\perp} = \mathbf{0}$, any value of $\bar{\mathbf{x}}_{\perp}$ satisfies the constraint
  - Therefore, to **minimize** $\|\bar{\mathbf{x}}\|$, we set $\bar{\mathbf{x}}_{\perp} = \mathbf{0}$

Least norm solution

- The **least norm solution** follows as

$$\bar{\mathbf{x}} = \bar{\mathbf{x}}_{\|} = \mathbf{A}^T \left(\mathbf{A}\mathbf{A}^T\right)^{-1} \bar{\mathbf{b}}$$

04

Worked example

Least norm solution of a 2 × 4 system

Least norm – example

- Consider the following problem

$$\min \|\bar{\mathbf{x}}\|$$

$$\text{subject to } \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \bar{\mathbf{x}}$$

Least norm – example

- We start as follows

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix}$$

Least norm – example

- Next

$$(\mathbf{A}\mathbf{A}^T)^{-1} = \begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix}^{-1} = \frac{1}{20}\begin{bmatrix} 14 & -6 \\ -6 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{7}{10} & -\frac{3}{10} \\ -\frac{3}{10} & \frac{1}{5} \end{bmatrix}$$

Least norm – example

- The solution is

$$\bar{\mathbf{x}} = \mathbf{A}^T\left(\mathbf{A}\mathbf{A}^T\right)^{-1}\bar{\mathbf{b}}$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \frac{1}{20} \begin{bmatrix} 14 & -6 \\ -6 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Least norm – example

- Which can be evaluated as

$$\bar{\mathbf{x}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \frac{1}{20} \begin{bmatrix} 22 \\ -8 \end{bmatrix}$$

$$= \frac{1}{20} \begin{bmatrix} 22 \\ 14 \\ 6 \\ -2 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 11 \\ 7 \\ 3 \\ -1 \end{bmatrix}$$

Summary

- A wide $\mathbf{A}$ $(n > m)$ with rank m gives an **underdetermined** system with infinitely many solutions
- **Least norm**: minimize $\|\bar{\mathbf{x}}\|$ subject to $\bar{\mathbf{b}} = \mathbf{A}\bar{\mathbf{x}}$
- Lagrange multipliers give $\bar{\mathbf{x}} = (1/2)\mathbf{A}^T\bar{\boldsymbol{\lambda}}$
- The solution is $\bar{\mathbf{x}} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\bar{\mathbf{b}}$
- The null-space part of $\bar{\mathbf{x}}$ does not change the constraint, so it is set to zero



Thank You!

Questions?