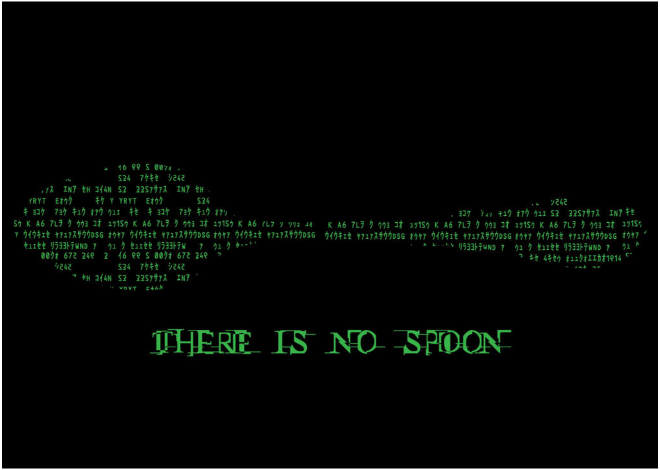

Introduction to Optimization

K. R. Sahasranand

Data Science

sahasranand@iitpkd.ac.in

Recall

Suppose that we are given m equations in n unknowns of the form

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

In matrix form,

$$Ax = b.$$

- The system $Ax = b$ has a solution iff rank A = rank $[A,\ b]$.

Fat matrices

$\approx$ more columns than rows

Consider a system of linear equations

$$Ax = b$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \leq n$, and rank $A = m$.

$$\begin{bmatrix} 0 & 1 & 1 & 2 & 3 & 5 \\ 8 & 13 & 21 & 34 & 55 & 89 \\ 44 & 33 & 77 & 10 & 87 & 97 \end{bmatrix}$$

Fat matrices

$\approx$ more columns than rows

Consider a system of linear equations

$$Ax = b$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \leq n$, and rank $A = m$.

$$\begin{bmatrix} 0 & 1 & 1 & 2 & 3 & 5 \\ 8 & 13 & 21 & 34 & 55 & 89 \\ 44 & 33 & 77 & 10 & 87 & 97 \end{bmatrix}$$

- $b \in \mathrm{Col}(A) = \mathbb{R}^m$

Fat matrices

$\approx$ more columns than rows

Consider a system of linear equations

$$Ax = b$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \leq n$, and rank $A = m$.

$$\begin{bmatrix} 0 & 1 & 1 & 2 & 3 & 5 \\ 8 & 13 & 21 & 34 & 55 & 89 \\ 44 & 33 & 77 & 10 & 87 & 97 \end{bmatrix}$$

- $b \in \mathrm{Col}(A) = \mathbb{R}^m$
- How many solutions?

Minimum norm solution

Suppose we choose x^* such that

$$Ax^* = b$$

and

$$\|x^*\| \leq \|x\|$$

for any x such that $Ax = b$.

Minimum norm solution

Suppose we choose x^* such that

$$Ax^* = b$$

and

$$\|x^*\| \leq \|x\|$$

for any x such that $Ax = b$.

minimize $\|x\|$ subject to $Ax = b$.

Minimum norm solution

The algebraic way

Theorem – The *unique* solution x^* to $Ax = b$ that minimizes the norm $\|x\|$ is given by

$$x^* = A^\top (AA^\top)^{-1} b.$$

Minimum norm solution

The algebraic way

Theorem – The *unique* solution x^* to $Ax = b$ that minimizes the norm $\|x\|$ is given by

$$x^* = A^\top (AA^\top)^{-1} b.$$

Proof – Let $x^* = A^\top (AA^\top)^{-1} b$. Then,

Minimum norm solution

The algebraic way

Theorem – The *unique* solution x^* to $Ax = b$ that minimizes the norm $\|x\|$ is given by

$$x^* = A^\top (AA^\top)^{-1} b.$$

Proof – Let $x^* = A^\top (AA^\top)^{-1} b$. Then,

$$\begin{aligned} \|x\|^2 &= \|(x - x^*) + x^*\|^2 \\ &= \cdots \\ &= \|x - x^*\|^2 + \|x^*\|^2 + 2{x^*}^\top (x - x^*) . \end{aligned}$$

Minimum norm solution

The algebraic way

Theorem – The *unique* solution x^* to $Ax = b$ that minimizes the norm $\|x\|$ is given by

$$x^* = A^\top (AA^\top)^{-1} b.$$

Proof – Let $x^* = A^\top (AA^\top)^{-1} b$. Then,

$$\begin{aligned} \|x\|^2 &= \|(x - x^*) + x^*\|^2 \\ &= \cdots \\ &= \|x - x^*\|^2 + \|x^*\|^2 + 2{x^*}^\top (x - x^*) \, . \end{aligned}$$

Next, we show that (on board)

$${x^*}^\top (x - x^*) = 0.$$

Minimum norm solution

The geometric way

- Any solution to $Ax = b$ that hopes to minimize the norm better be such that
$$x \in \mathrm{Row}(A).$$

Minimum norm solution

The geometric way

- Any solution to $Ax = b$ that hopes to minimize the norm better be such that
$$x \in \mathrm{Row}(A).$$

Why? Decomposition into $\mathrm{Row}(A) + \mathcal{N}(A)$

Minimum norm solution

The geometric way

- Any solution to $Ax = b$ that hopes to minimize the norm better be such that

$$x \in \mathrm{Row}(A).$$

Why? Decomposition into $\mathrm{Row}(A) + \mathcal{N}(A)$

- Then, the solution x^* is of the form

$$x^* = A^\top y \qquad \text{for some } y \in \mathbb{R}^m.$$

Minimum norm solution

The geometric way

- Any solution to $Ax = b$ that hopes to minimize the norm better be such that

$$x \in \mathrm{Row}(A).$$

Why? Decomposition into $\mathrm{Row}(A) + \mathcal{N}(A)$

- Then, the solution x^* is of the form

$$x^* = A^\top y \qquad \text{for some } y \in \mathbb{R}^m.$$

- So, we have

$$AA^\top y = b$$

Minimum norm solution

The geometric way

- Any solution to $Ax = b$ that hopes to minimize the norm better be such that

$$x \in \mathrm{Row}(A).$$

Why? Decomposition into $\mathrm{Row}(A) + \mathcal{N}(A)$

- Then, the solution x^* is of the form

$$x^* = A^\top y \qquad \text{for some } y \in \mathbb{R}^m.$$

- So, we have

$$AA^\top y = b \qquad \Longrightarrow \qquad y = (AA^\top)^{-1} b$$

Minimum norm solution

The geometric way

- Any solution to $Ax = b$ that hopes to minimize the norm better be such that

$$x \in \mathrm{Row}(A).$$

Why? Decomposition into $\mathrm{Row}(A) + \mathcal{N}(A)$

- Then, the solution x^* is of the form

$$x^* = A^\top y \qquad\qquad \text{for some } y \in \mathbb{R}^m.$$

- So, we have

$$AA^\top y = b \qquad \Longrightarrow \qquad y = (AA^\top)^{-1} b$$

whereby

$$x^* = A^\top y = A^\top (AA^\top)^{-1} b.$$

An example

- Consider the space $\mathbb{R}^3$. The intersection of the planes defined by

$$2x_1 + x_2 + 2x_3 = 9$$

and

$$5x_1 + 5x_2 + 7x_3 = 29$$

is a line.

An example

- Consider the space $\mathbb{R}^3$. The intersection of the planes defined by

$$2x_1 + x_2 + 2x_3 = 9$$

and

$$5x_1 + 5x_2 + 7x_3 = 29$$

is a line. Identify the point on this line that is closest to the origin.

An example

- Consider the space $\mathbb{R}^3$. The intersection of the planes defined by

$$2x_1 + x_2 + 2x_3 = 9$$

and

$$5x_1 + 5x_2 + 7x_3 = 29$$

is a line. Identify the point on this line that is closest to the origin.

– We write the system of equations in the form $Ax = b$ as follows:

$$\begin{bmatrix} 2 & 1 & 2 \\ 5 & 5 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 29 \end{bmatrix}.$$

An example

- Consider the space $\mathbb{R}^3$. The intersection of the planes defined by

$$2x_1 + x_2 + 2x_3 = 9$$

and

$$5x_1 + 5x_2 + 7x_3 = 29$$

is a line. Identify the point on this line that is closest to the origin.

– We write the system of equations in the form $Ax = b$ as follows:

$$\begin{bmatrix} 2 & 1 & 2 \\ 5 & 5 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 29 \end{bmatrix}.$$

The solution x^* with the smallest norm is given by

$$x^* = A^T(AA^T)^{-1}b = \cdots = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

More generally . . .

$$Ax = b,\ A \in \mathbb{R}^{m \times n},\ \text{rank } A = r.$$

More generally . . .

$Ax = b,\ A \in \mathbb{R}^{m\times n},$ rank $A = r$.

– Find the vector x^* that minimizes $\|Ax - b\|^2$.

More generally . . .

$Ax = b, \ A \in \mathbb{R}^{m \times n}, \ \text{rank } A = r.$

- Find the vector x^* that minimizes $\|Ax - b\|^2$.
- Among such x^*s, find the one with the minimum norm.

More generally . . .

$Ax = b,\ A \in \mathbb{R}^{m \times n},$ rank $A = r$.

- Find the vector x^* that minimizes $\|Ax - b\|^2$.

- Among such x^*s, find the one with the minimum norm.

$$x^\star = A^\dagger b$$

$A^\dagger \approx$ generalized inverse of A