

**Applied Optimization for Wireless, Machine Learning Big Data**  
**Prof. Adithya K. Jagannatham**  
**Department of Electrical Engineering**  
**Indian Institute of Technology Kanpur**

**Lesson - 45**  
**Least Norm Signal Estimation**

**Keywords:** *Least Norm Signal Estimation*

Hello welcome to another module in this massive open online course. So we have looked at the least squares paradigm, let us look at its analogue or a counterpart which is known as the Least Norm Paradigm.

(Refer Slide Time: 00:28)

**LEAST NORM:**

Signal Recovery

Consider the following problem

$$\underline{y} = A \underline{x}$$

$m \times n$   
 $m < n$   
# rows < # columns

Wide matrix

So consider the following problem where we have  $\underline{y} = A \underline{x}$  where  $A$  is an  $m \times n$  matrix. But while previously  $m > n$ , in the least norm framework we will consider  $m < n$ , that is the number of rows is much lower than the number of columns, so this can be called as a wide matrix.

(Refer Slide Time: 02:17)

$y = Ax$   
 $A$  is a Wide matrix with  $m < n$ .  
 # rows = Equations =  $m$   
 # columns = unknowns =  $n$   
 # Equations < # unknowns.  
 Underdetermined.  
 Not enough constraints on  $x$

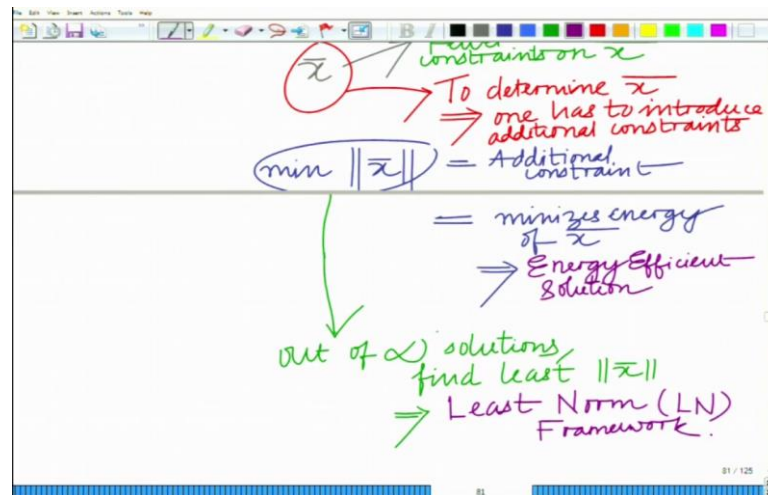
This basically implies that the number of equations is smaller than the number of unknowns which implies that the system is under determined.

(Refer Slide Time: 04:16)

Typically More than one solution or number of solutions.  
 Fewer constraints on  $x$ .  
 To determine  $x$ , one has to introduce additional constraints.  
 $\min \|x\|$  = Additional constraint  
 = minimizes energy of  $x$

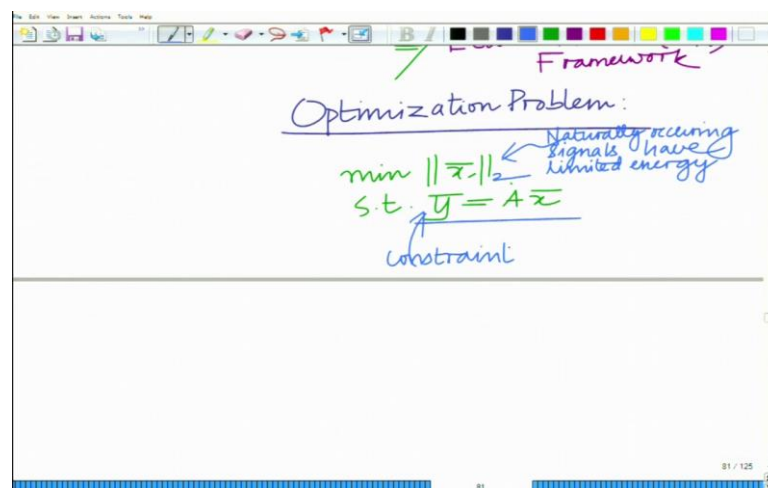
This means that there is an infinite number of solutions and if there are fewer constraints then the only way to determine the possible value of  $\bar{x}$  is to introduce additional constraints. So this can be  $\min \|x\|$ , so this is the additional constraint. It basically minimizes the energy of  $\bar{x}$  which implies that you are trying to find an energy efficient solution. And this is precisely known as the least norm problem.

(Refer Slide Time: 06:47)



So this means out of infinite solutions, find the one that has least norm.

(Refer Slide Time: 08:32)



The relevant optimization problem for this can be given as  $\min \|\bar{x}\|_2$  and this is the objective function and the constraint is  $\bar{y} = A \bar{x}$ . So this is justified because naturally occurring signals have limited energy.

(Refer Slide Time: 09:59)

$$\begin{aligned} &\equiv \min \| \bar{x} \|^2 = \bar{x}^T \bar{x} \\ &\text{s.t. } \bar{y} = A \bar{x} \\ f(\bar{x}, \lambda) &= \bar{x}^T \bar{x} + \bar{\lambda}^T (A \bar{x} - \bar{y}) \\ \bar{\lambda} &= \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix} \end{aligned}$$

*m constraints*

Therefore, one can form the Lagrangian which is  $f(\bar{x}, \lambda) = \bar{x}^T \bar{x} + \bar{\lambda}^T (A \bar{x} - \bar{y})$  because we have  $m$  constraints, each row is an equation. So there has to be one Lagrange multiplier for each constraint.

(Refer Slide Time: 11:09)

$$\begin{aligned} f &= \bar{x}^T \bar{x} + \bar{\lambda}^T (A \bar{x} - \bar{y}) \\ \nabla F &= 2\bar{x} + A^T \bar{\lambda} - 0 \Rightarrow 0 \end{aligned}$$

*one Lagrange multiplier for each constraint*

*setting gradient = 0*

So we can solve this as shown in slide by taking the gradient with respect to  $\bar{x}$  and setting it to 0.

(Refer Slide Time: 12:29)

The image shows a handwritten derivation on a whiteboard. At the top, the equation  $\nabla F = 2\bar{x} + A^T\lambda = 0$  is written. A green arrow points from this equation to the text "setting gradient = 0". Below this, the equation  $\Rightarrow \bar{x} = -\frac{1}{2} \cdot A^T\lambda$  is written, with a green circle around  $A^T\lambda$ . An orange arrow points from this circle to the text "To determine  $\lambda$  use constraint". Below this, the constraint equation  $A\bar{x} = y$  is written in orange. Finally, the equation  $\Rightarrow A(-\frac{1}{2} \cdot A^T\lambda) = y$  is written in orange.

$$\nabla F = 2\bar{x} + A^T\lambda = 0$$

setting gradient = 0

$$\Rightarrow \bar{x} = -\frac{1}{2} \cdot A^T\lambda$$

To determine  $\lambda$  use constraint

$$A\bar{x} = y$$
$$\Rightarrow A(-\frac{1}{2} \cdot A^T\lambda) = y$$

So on solving, finally you will get  $x = A^T (A A^T)^{-1} y$ .

(Refer Slide Time: 13:29)

The image shows a handwritten derivation on a whiteboard. At the top, the equation  $\Rightarrow \lambda = -2(AA^T)^{-1}y$  is written, with a green circle around the entire expression. Below this, the text "Substitute  $\lambda$  from ② in ①" is written. Then, the equation  $\Rightarrow \bar{x} = -\frac{1}{2} \cdot A^T\lambda$  is written, followed by  $= -\frac{1}{2} \cdot A^T(-2(AA^T)^{-1}y)$ . Finally, the equation  $\hat{x} = A^T(AA^T)^{-1}y$  is written and boxed in purple.

$$\Rightarrow \lambda = -2(AA^T)^{-1}y$$

Substitute  $\lambda$  from ② in ①

$$\Rightarrow \bar{x} = -\frac{1}{2} \cdot A^T\lambda$$
$$= -\frac{1}{2} \cdot A^T(-2(AA^T)^{-1}y)$$
$$\boxed{\hat{x} = A^T(AA^T)^{-1}y}$$

So this is the least norm signal estimate and this also known as the least norm solution.

(Refer Slide Time: 15:10)

The image shows a handwritten slide from a presentation. At the top, there is a toolbar with various drawing tools. Below the toolbar, the text "= Least Norm signal estimate" is written in purple ink. In the center, the formula  $\hat{x} = A^T(AA^T)^{-1}y$  is enclosed in a rectangular box. Below the box, the text "Least Norm(LN) solution" is written in blue ink, with a blue arrow pointing from this text to the  $A^T$  term in the formula. At the bottom right of the slide, the text "85 / 125" is visible.

$$\hat{x} = A^T(AA^T)^{-1}y$$

= Least Norm signal estimate

Least Norm(LN) solution

So this is suitable for scenarios where there are under constrained systems. So we will stop here. Thank you very much.