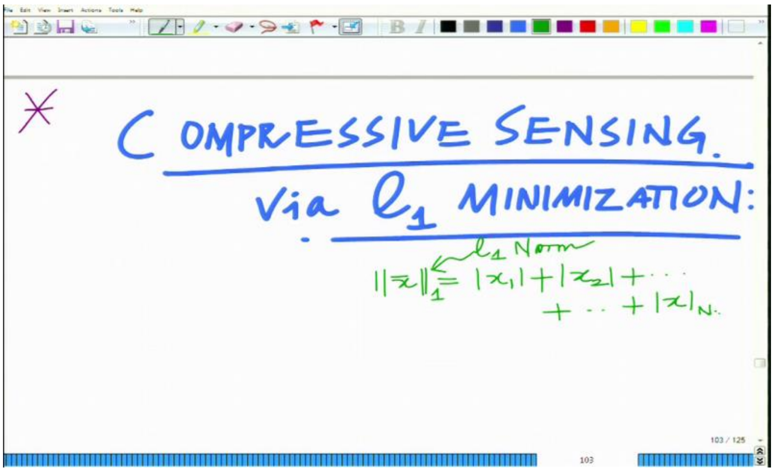
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 59
Practical Application: L_1 norm minimization and Regularization approach for Compressive Sensing Optimization Problem

***Keywords**: L_1 norm minimization, Regularization*

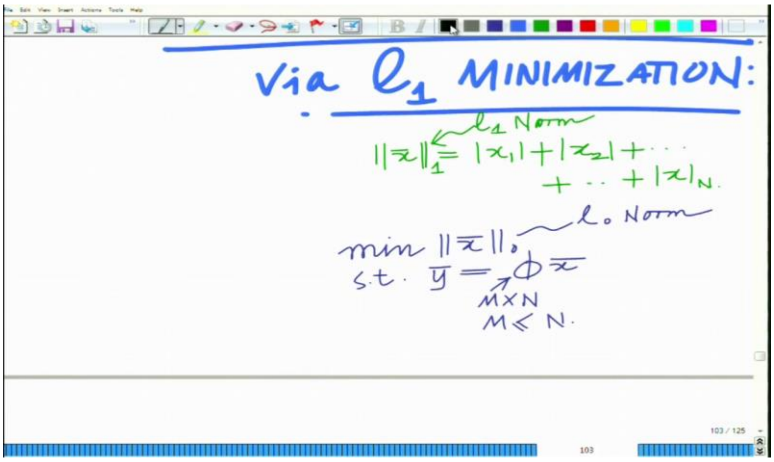
Hello, welcome to another module in this massive open online course. We were looking at compressive sensing and we have discussed the orthogonal matching pursuit. Let us look at another completely different and radical approach to tackle this compressive sensing problem.

(Refer Slide Time: 00:34)



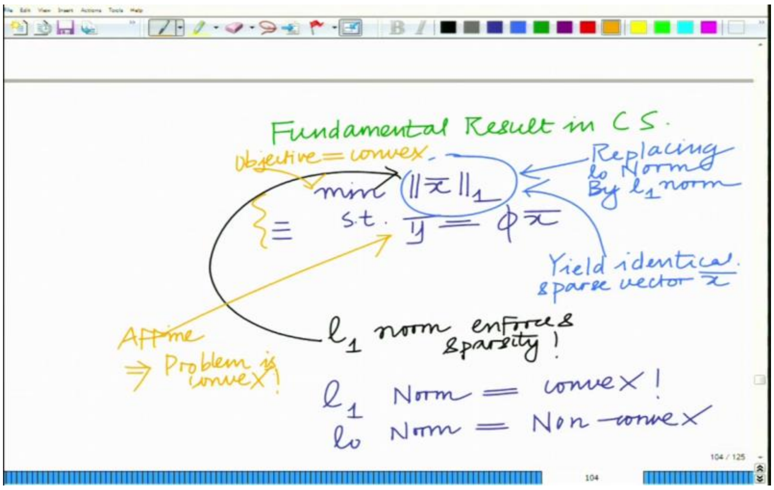
So we want to look at compressive sensing via l_1 norm minimization. We have seen this l_1 norm of a vector that is if you have an N-dimensional vector $\bar{x}$ the l_1 norm is simply the sum of the magnitudes of the components of $\bar{x}$. That is we have $\left\|\bar{x}\right\|_1 = |x_1| + |x_2| + \ldots.. + |x|_N$.

(Refer Slide Time: 01:56)



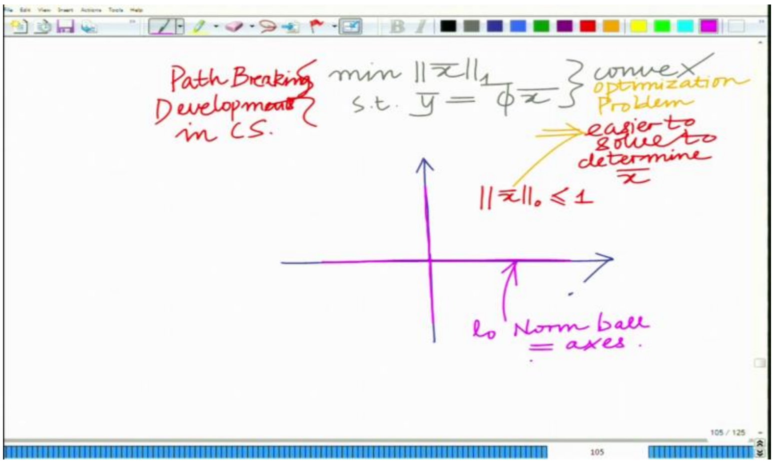
The compressive sensing problem is the following thing that is $\begin{array}{c}\min \left\|\bar{x}\right\|_0 \\ s.t\ \bar{y} = \phi \bar{x}\end{array}$. One of the fundamental results in compressive sensing is that this l_0 norm minimization can be replaced by l_1 norm minimization and still you can recover the sparse vector $\bar{x}$.

(Refer Slide Time: 02:54)



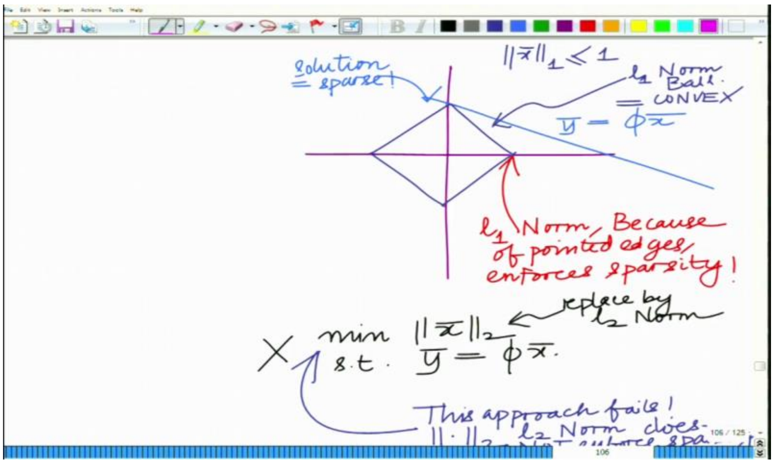
So we have $\begin{array}{c}\min \left\|\bar{x}\right\|_1 \\ s.t\ \bar{y} = \phi \bar{x}\end{array}$ what this says is that the l_1 norm also enforces sparsity. And it can be shown that for a large number of scenarios or with very high probability $\bar{x}$ that is obtained as a solution of both these above optimization problems is the same. The significant advantage is, the l_1 norm is convex in nature and the l_0 norm is highly non-convex. So therefore if you look at this optimization problem, the objective is convex and the constraint is an affine constraint. So we are converting a problem which was previously highly non-convex into something that is convex. So this is much easier to solve and determine the sparse vector $\bar{x}$.

(Refer Slide Time: 06:48)



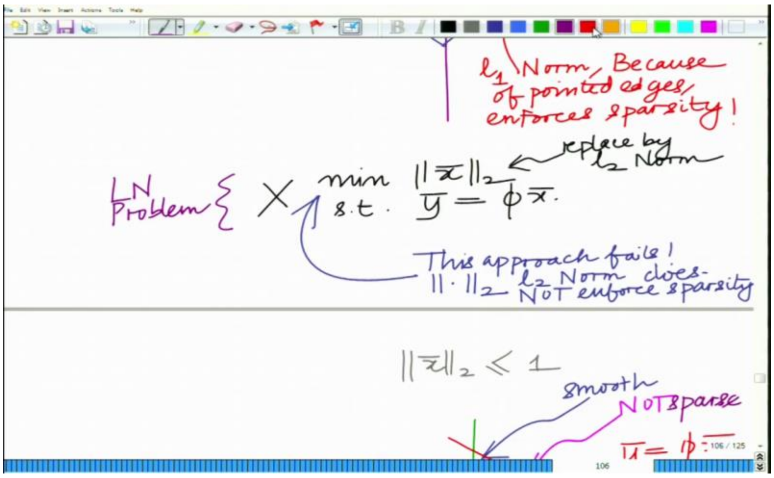
So this is one of the path breaking developments in compressive sensing that is demonstrating that the l_0 norm minimization is equivalent in a large number of scenarios to the l_1 norm minimization. If you look at the l_0 norm ball that is $\left\|\bar{x}\right\|_0 \leq 1$, it is simply along the axis, so this is highly non-convex.

(Refer Slide Time: 09:05)



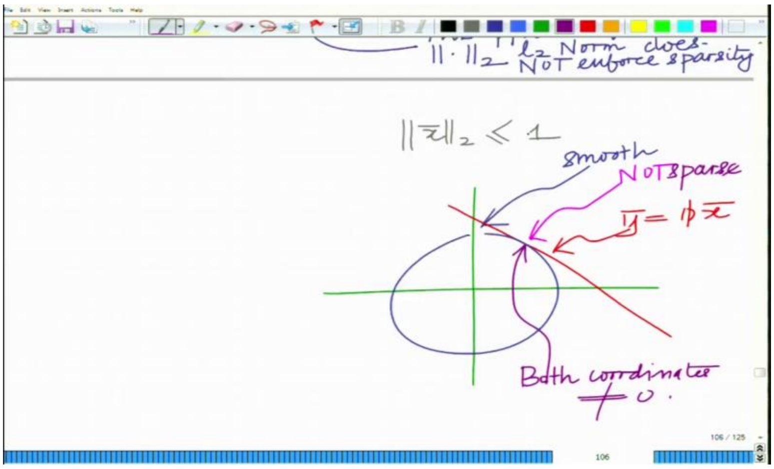
However the l_1 norm ball looks like a diamond shaped object which is a convex shape. And now if you enforce this affine constraint, which is nothing but a line, it intersects at one of these pointed edges, implies the solution is sparse.

(Refer Slide Time: 10:56)



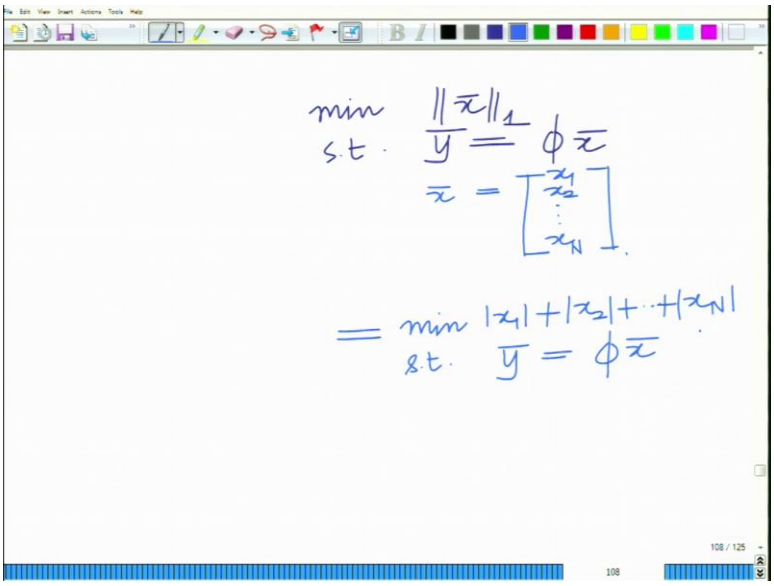
Now, if you replace it by the l_2 norm this approach fails because the l_2 norm does not enforce sparsity.

(Refer Slide Time: 12:12)



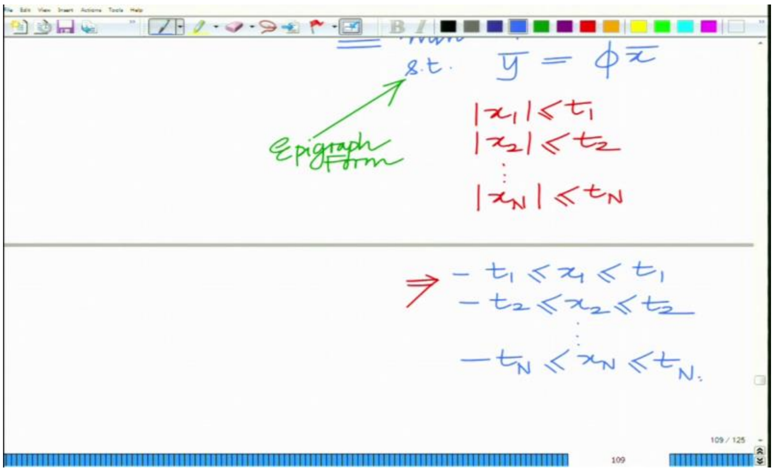
The l_2 norm is smooth with no pointed edges and if you look at this affine constraint it intersects at a point which is not sparse.

(Refer Slide Time: 15:07)

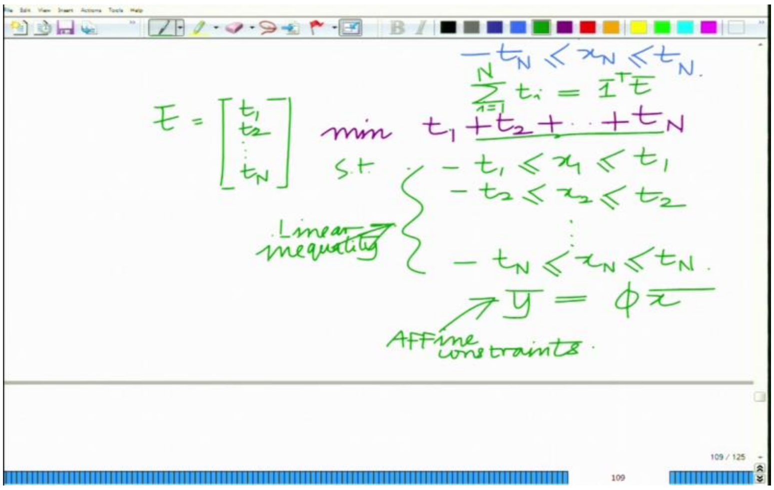


The l_1 norm minimization problem can be further simplified using the epigraph form as shown in the slides below.

(Refer Slide Time: 16:02)



(Refer Slide Time: 17:01)

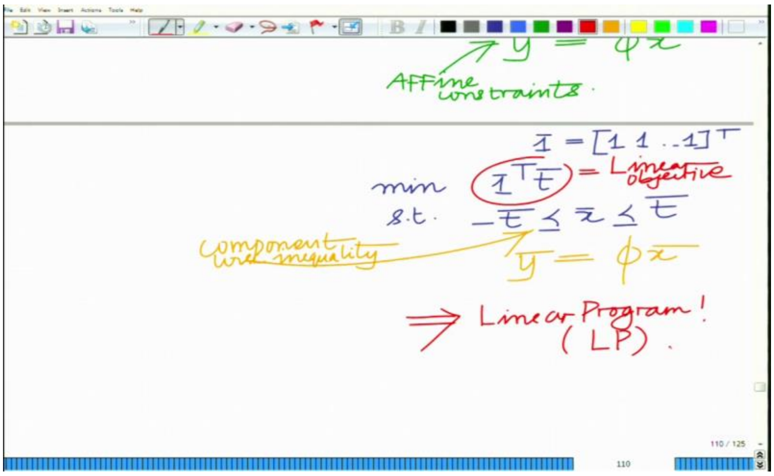


Therefore, this optimization problem can be written as

$$\begin{aligned} &\min t_1 + t_2 + \ldots + t_N \\ &s.t \quad -t_1 \le x_1 \le t_1 \\ &\qquad -t_2 \le x_2 \le t_2 \\ &\qquad \quad . \\ &\qquad \quad . \\ &\qquad \quad . \\ &\qquad -t_N \le x_N \le t_N \\ &\qquad \bar{y} = \phi \bar{x} \end{aligned}$$

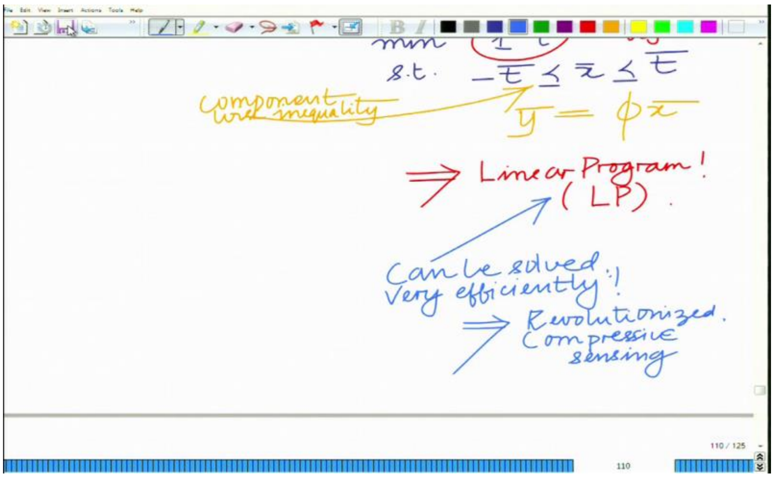
and this set of linear inequalities is also known as box constraint. So these are your linear inequality constraints.

(Refer Slide Time: 18:42)

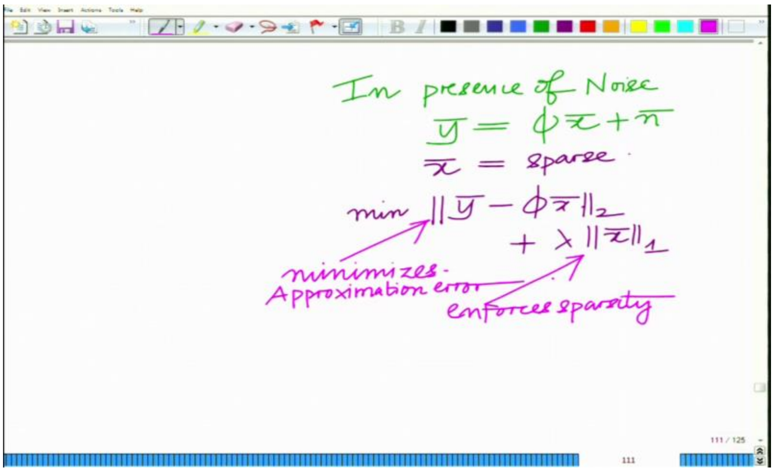


So this can also be written as $\begin{array}{l} \min \bar{1}^T \bar{t} \\ s.t \ -\bar{t} \le \bar{x} \le \bar{t} \\ \quad \bar{y} = \phi \bar{x} \end{array}$. So this is the component wise inequality and you can see these are linear inequalities, affine constraint and linear objective. So the compressive sensing problem to estimate the sparse vector $\bar{x}$ reduces to a linear program for which there are efficient techniques to solve.

(Refer Slide Time: 20:30)

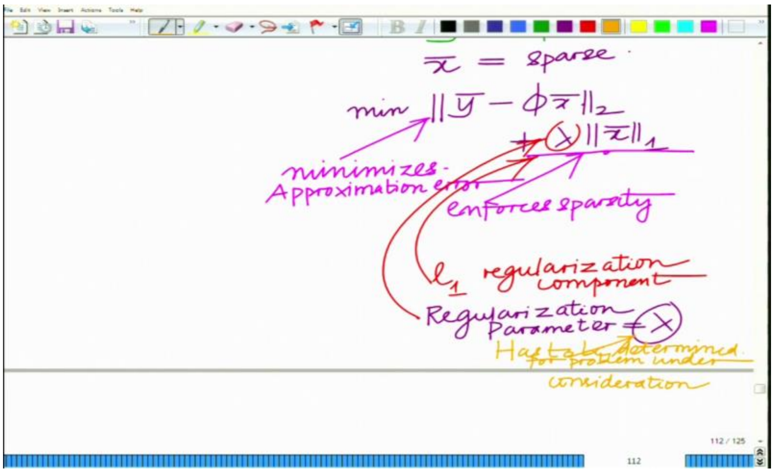


(Refer Slide Time: 21:18)



Now so far we have considered a noiseless observation model. Now in the presence of noise you have the observation model $\bar{y} = \phi \bar{x} + \bar{n}$ and the vector $\bar{x}$ is sparse. Now previously you minimize the least squares. Now here you have $\min \left\| \bar{y} - \phi \bar{x} \right\|_2 + \lambda \left\| \bar{x} \right\|_1$ that is the approximation or fit error or the observation model error and this additional term enforces sparsity.

(Refer Slide Time: 22:59)



So basically you are adding an l_1 regularization component and this λ is termed as the regularization parameter. This has to be determined for the problem under consideration. So here you are recovering the sparse vector as well as at the same time minimizing the approximation error. So basically this is known as the regularized version of the previous problem. So we will stop here. Thank you very much.