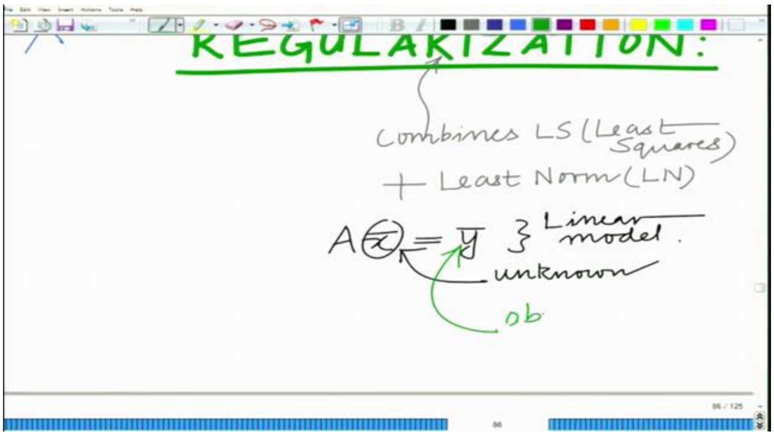
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 46
Regularization: Least Squares + Least Norm

Keywords: *Least Squares, Least Norm*

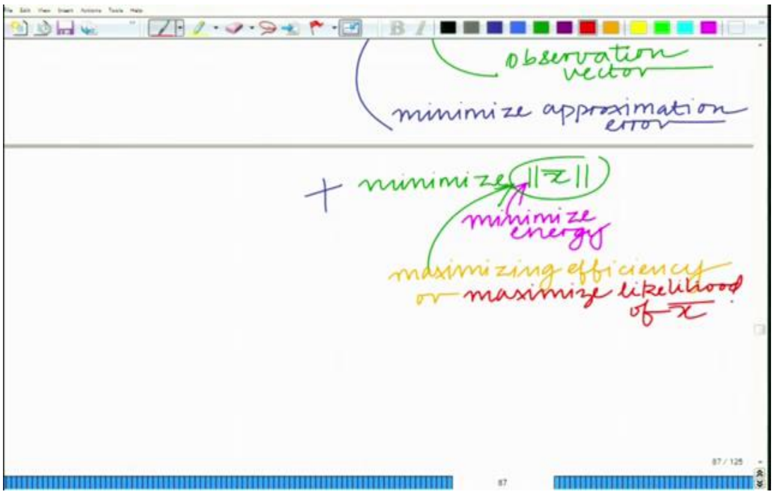
Hello welcome to another module, in this massive open online course. So we are looking at various convex optimisation problems, which are the Least squares and Least norm. Let us look at another problem, which is essentially the combination of both these problems, which is termed as regularization.

(Refer Slide Time: 00:37)



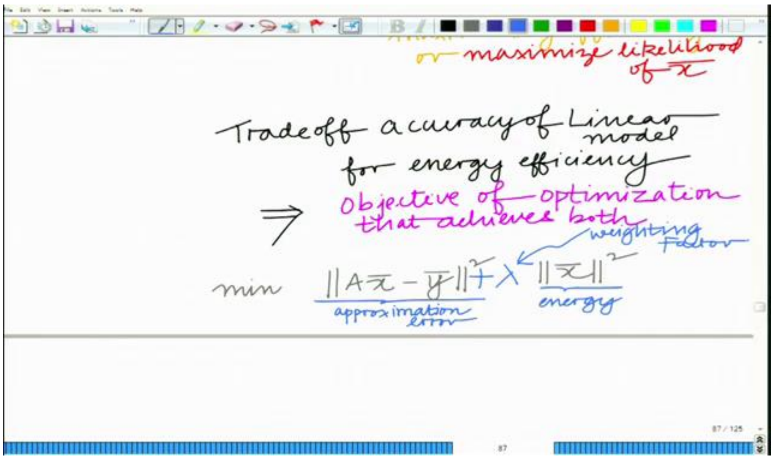
So regularization combines the least squares and least norm frameworks.

(Refer Slide Time: 02:23)



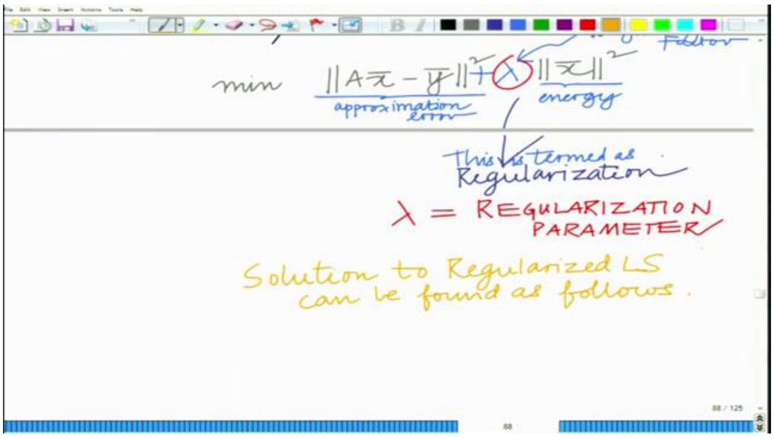
So $A\bar{x} = \bar{y}$ is the linear model and this vector $\bar{x}$ is an unknown vector, vector $\bar{y}$ is the observation matrix and the matrix A is assumed to be known. So we want to minimise the approximation error at the same time, we would also like to minimise the energy. So you are trying to maximize the prior probability of such vectors arising in the problems. So you can also think of this as not probability, but rather the likelihood of $\bar{x}$.

(Refer Slide Time: 05:16)



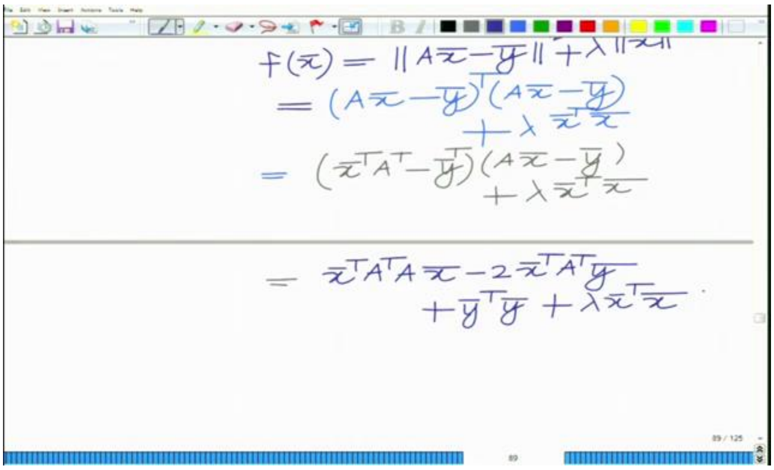
So we would like to have an objective function which achieves both the above objectives and therefore, one can formulate the following optimisation problem by combining the approximation error and the energy of the solution, minimise a linear combination of this as shown. So we have $\min \left\|A\bar{x} - \bar{y}\right\|^2 + \lambda \left\|\bar{x}\right\|^2$ this λ is a weighting factor and not the Lagrange multiplier and therefore, you have a weighted objective. So you have a weighted combination of the approximation error and the energy and this process is known as regularization. So this basically encourages solutions that have lower energy.

(Refer Slide Time: 08:04)



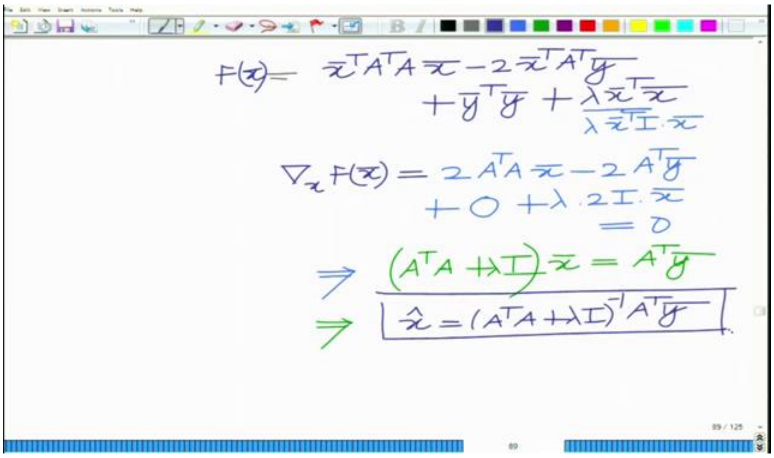
This factor λ is termed as the regularization parameter and basically in a scenario where we would like to achieve a trade-off between the accuracy as well as an energy efficient solution, one can apply this approach and the procedure to solve this is similar to what we have seen before.

(Refer Slide Time: 09:52)



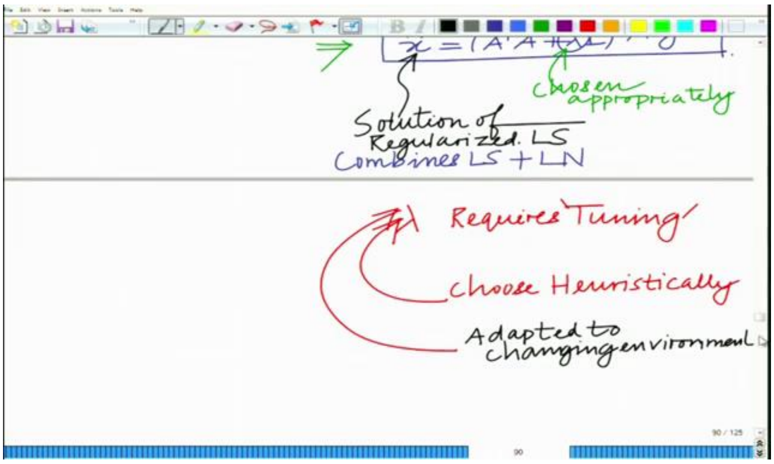
This can also be thought of as the regularized least squares and the solution to the regularized LS can be found as follows where we have the objective function $f(\bar{x}) = \left\|A\bar{x} - \bar{y}\right\|^2 + \lambda\left\|\bar{x}\right\|^2$ and this is solved as shown in slide.

(Refer Slide Time: 11:36)



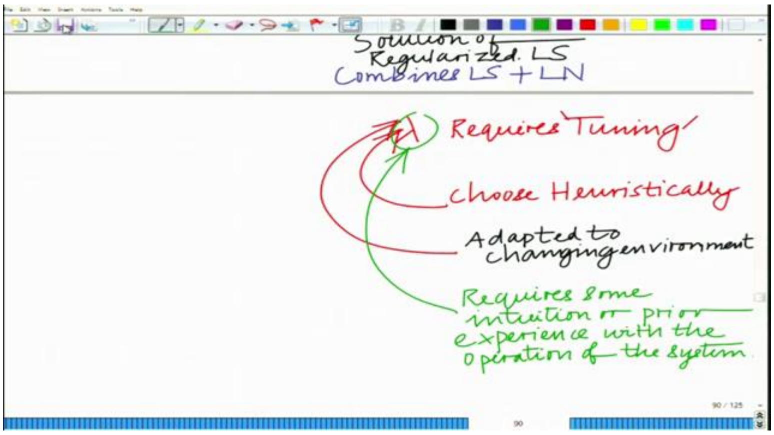
Now we need to take the gradient of this with respect to $\bar{x}$ and we get the solution as shown in slide and finally we get $x = \left(A^T A + \lambda I\right)^{-1} A^T \bar{y}$.

(Refer Slide Time: 13:30)



So this is basically the solution to the regularized that combines both the least squares and the least norm. Now this regularization parameter needs to be chosen appropriately that is it requires a tuning to get the best solution.

(Refer Slide Time: 16:12)



So either choose heuristically or tune it appropriately by adapting it to the changing by changing environment and it requires some prior knowledge with the operation of the system so as to determine the regularizing parameter to find the best solution. So that basically completes our discussion on the regularised least squares which combines both the least squares and the least norm frameworks. Thank you very much.