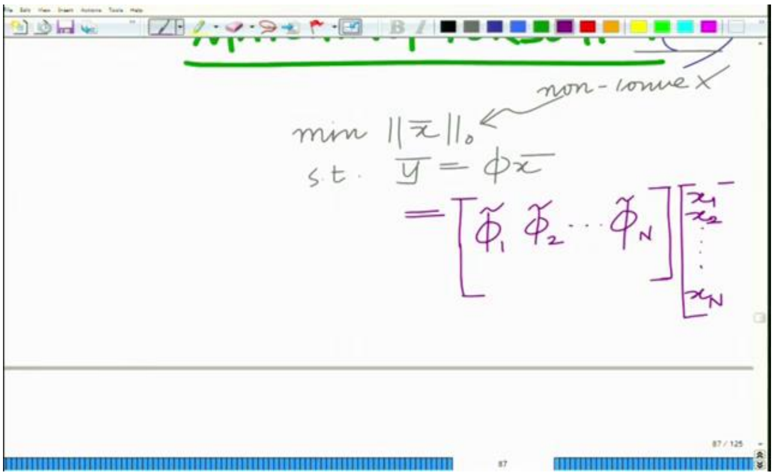
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 57
Practical Application: Orthogonal Matching Pursuit (OMP) algorithm for Compressive Sensing

***Keywords**: Orthogonal Matching Pursuit (OMP)*

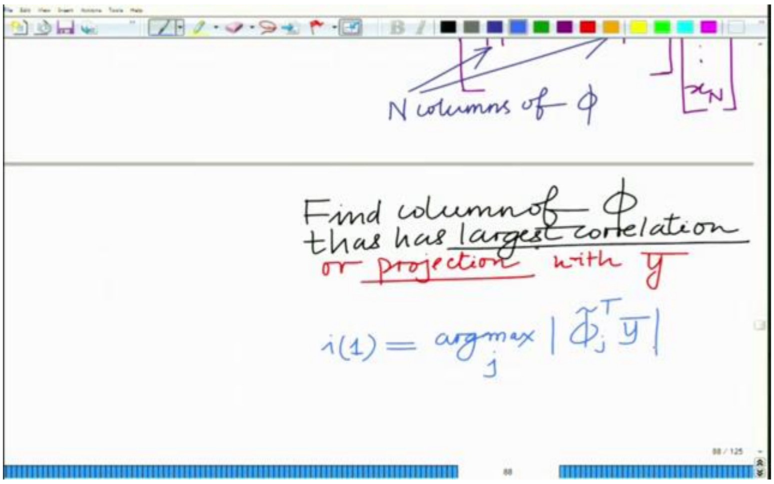
Hello, welcome to another module in this massive open online course. So we are looking at compressive sensing and we have seen that the cost function for the compressive sensing problem is highly non convex and therefore, we have to come up with intelligent techniques to solve this and hence we are going to look at the orthogonal matching pursuit.

(Refer Slide Time: 00:38)



So this is one of the schemes for sparse signal recovery and this is also abbreviated as OMP.

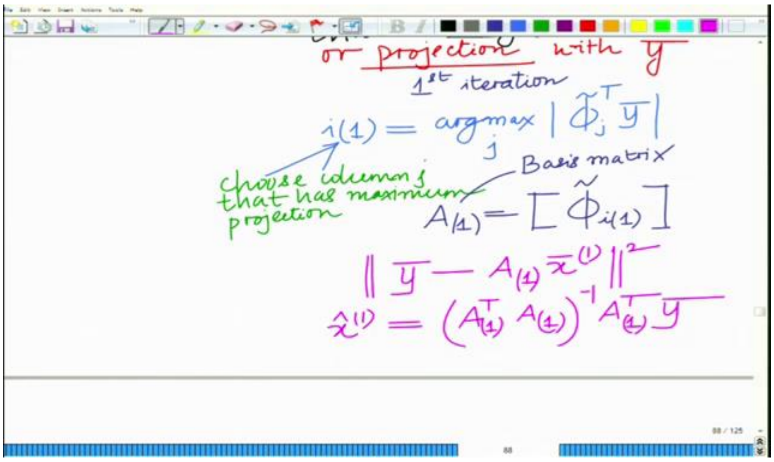
(Refer Slide Time: 02:24)



The name itself implies matching that is we are looking for the column that closely matches the vector $\bar{x}$, which means you have to find the projection of $\bar{x}$ on each of these columns of the matrix ϕ and choose the one that has the maximum value. So we are trying to find the one which has the largest projection on $\bar{y}$ so the way to do that is to find the column of ϕ that will have the largest correlation or basically projection with $\bar{y}$.

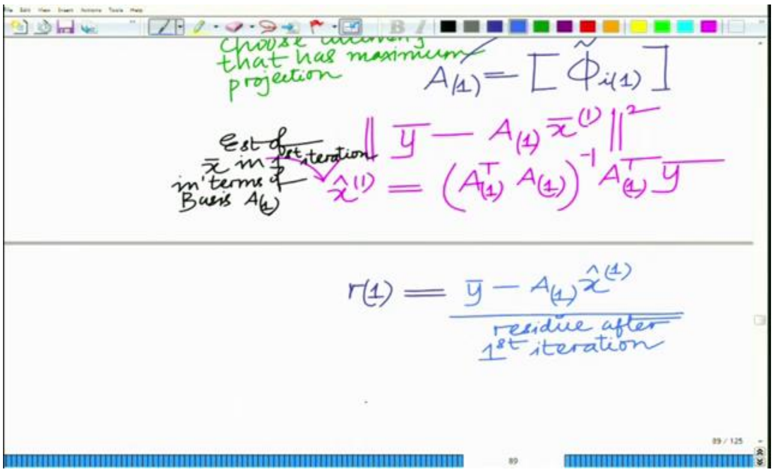
So the way to do that is we have $i(1) = \arg\max_{j} \left|\phi_j^T \bar{y}\right|$.

(Refer Slide Time: 05:02)



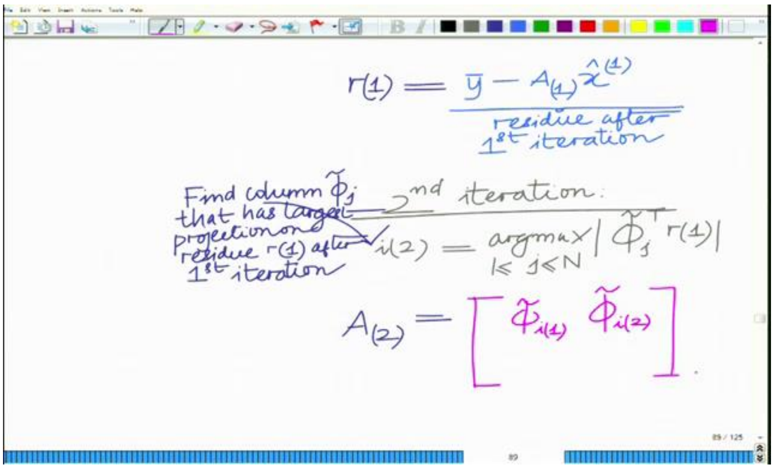
So choose the column j that has the maximum projection. And now we start building the basis matrix that is $A_{(1)} = \left[\phi_{i(1)}\right]$. So at this stage, this is a single column matrix. So this is by the way the first iteration of the algorithm and now we find try to find the best estimate of the vector $\bar{x}$. So $\left\|\bar{y} - A_{(1)}\bar{x}^{(1)}\right\|^2$ that is we are trying to minimize the least squares norm, such that you find the best vector $x^{(1)} = \left(A_{(1)}^T A_{(1)}\right)^{-1} A_{(1)}^T \bar{y}$ in the first iteration that minimizes this error. So what we are doing is we are trying to estimate the columns of ϕ which are present in the linear combination that give rise to only few elements of $\bar{x}$ that are non-zero. So that is what we are trying to find by this orthogonal matching pursuit. So we take the projection of each column on $\bar{y}$, finding the one that has a maximum projection, choosing that column as the basis, then finding the best type for optimization to the $\bar{y}$ based on that basis, that is what we are doing here by solving this least squares problem. Now we find the residue that is left after getting this best possible approximation.

(Refer Slide Time: 08:00)



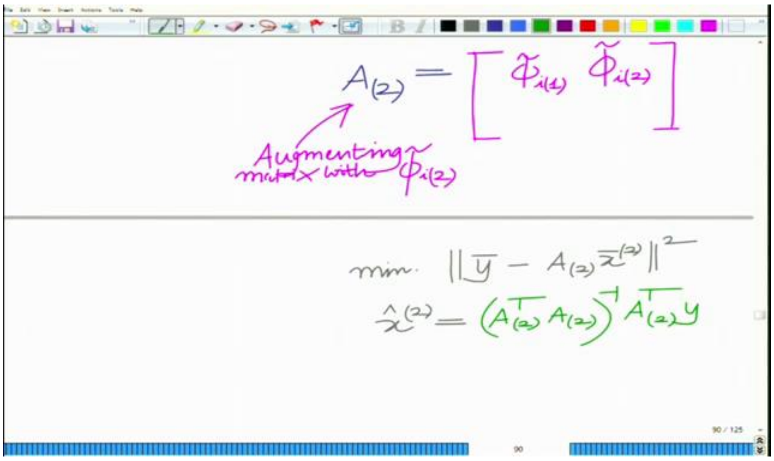
So this is $r(1) = \bar{y} - A_{(1)} x^{(1)}$.

(Refer Slide Time: 09:08)



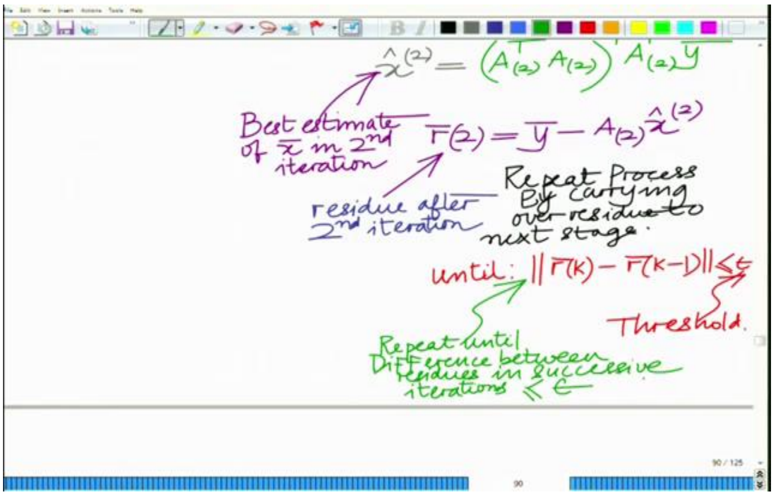
So in the second iteration we take the projection of this on the residue that is we find the column which has the maximum projection on the residue after the first iteration. So we find the projection of each column of ϕ on this residue and choose the column which now has the maximum projection on this residue.

(Refer Slide Time: 11:17)



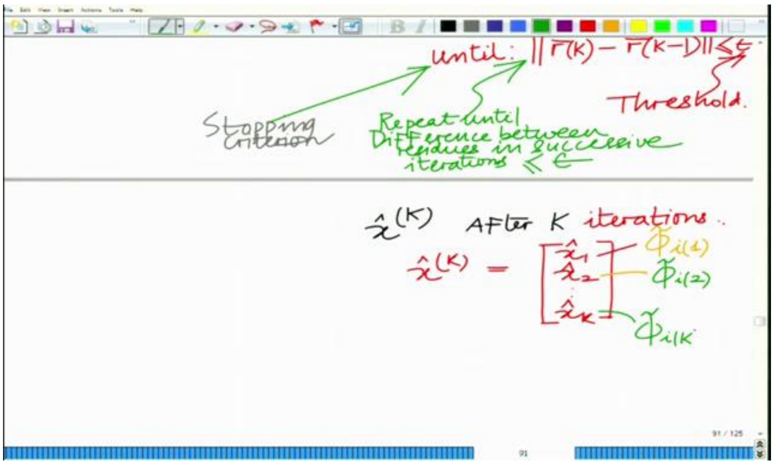
Now, you augment your basis matrix and we get $A_{(2)}$ as shown in slide. Once again you find the best estimate $\bar{x}$ via least squares as shown in slide.

(Refer Slide Time: 12:14)



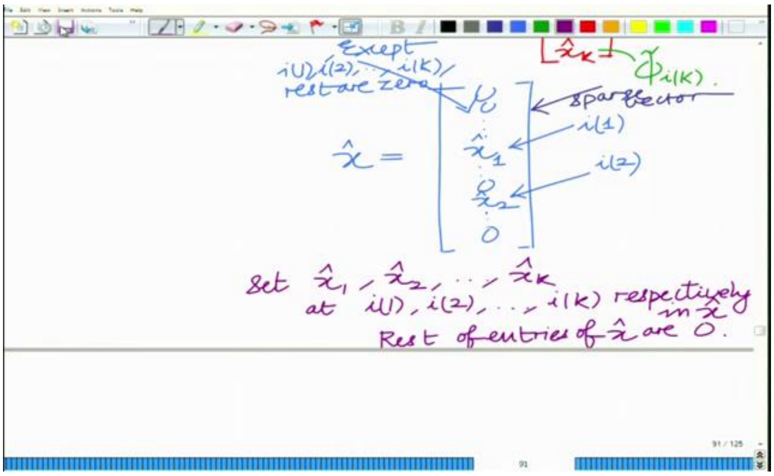
Now we find the residue after the second iteration and then subsequently we do the third iteration and keep repeating this process until the residue stops decreasing, that is you repeat until such a stage that is let us say, you have K iterations, that is $\left\|\bar{r}(K) - \bar{r}(K-2\right\| \le \varepsilon$ which is some threshold.

(Refer Slide Time: 15:54)



So this is termed as the stopping criteria. Let us say, you stop after K iterations and after this you have $x^{(K)}$, which means you basically obtained a fairly good estimate of approximation to $\bar{y}$ and the residue is not decreasing any further.

(Refer Slide Time: 18:17)



And now x will simply be a vector that mostly contains zeros, except corresponding to the location i(1), i(2) and i(K), let us say so. So this is a sparse vector that is estimated using the orthogonal matching pursuit. So we will stop here and we will look at an example in the subsequent module. Thank you very much.