Applied Optimization for Wireless, Machine Learning, Big Data

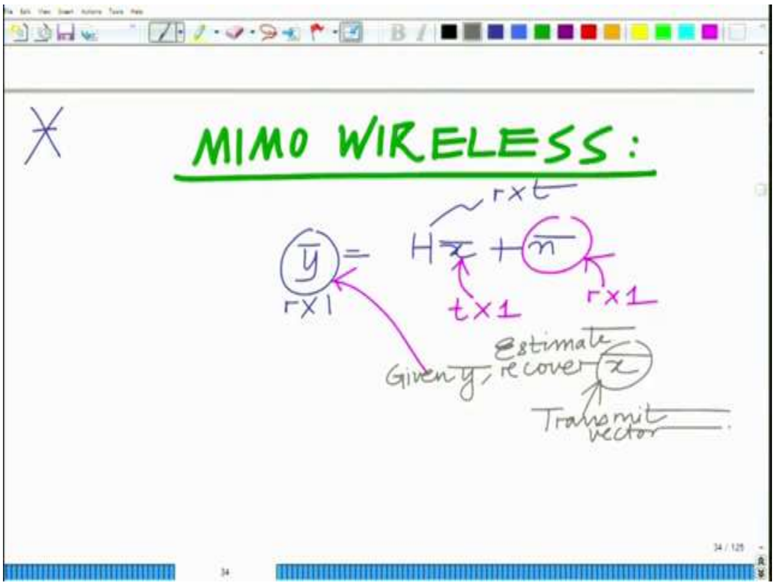
Prof. Aditya K. Jagannatham
Department Of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture –26

Application: MIMO Receiver Design as a Least Squares Problem

Hello, welcome to another module in this massive open online course. So, we are looking at a practical application of a MIMO communication system. Let us continue this discussion.

(Refer Slide Time: 00:30)

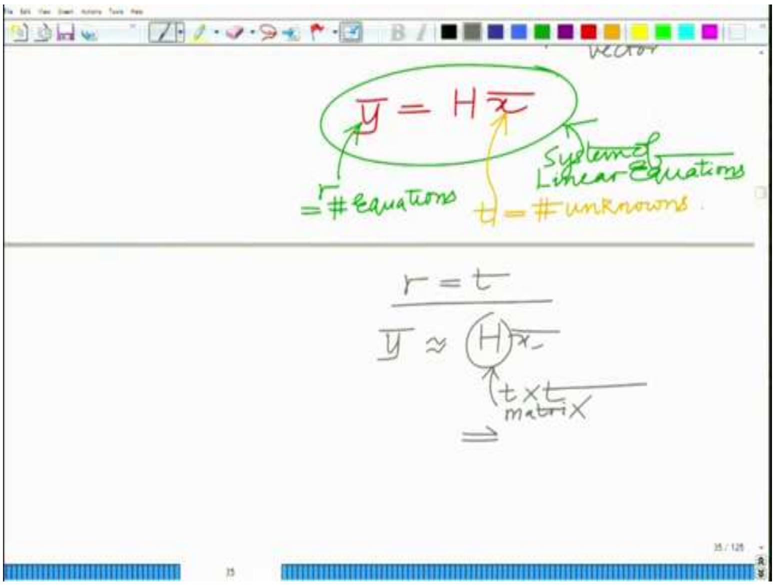


The model for a MIMO system with t number of transmitters and r number of receivers is given as

$$\bar{y} = H\bar{x} + \bar{n}$$

Where $\bar{x}$ is the $t \times 1$ transmit vector, $\bar{y}$ is the $r \times 1$ received vector, H is the $r \times t$ MIMO channel matrix and $\bar{n}$ is the $r \times 1$ additive noise vector.

(Refer Slide Time: 01:50)



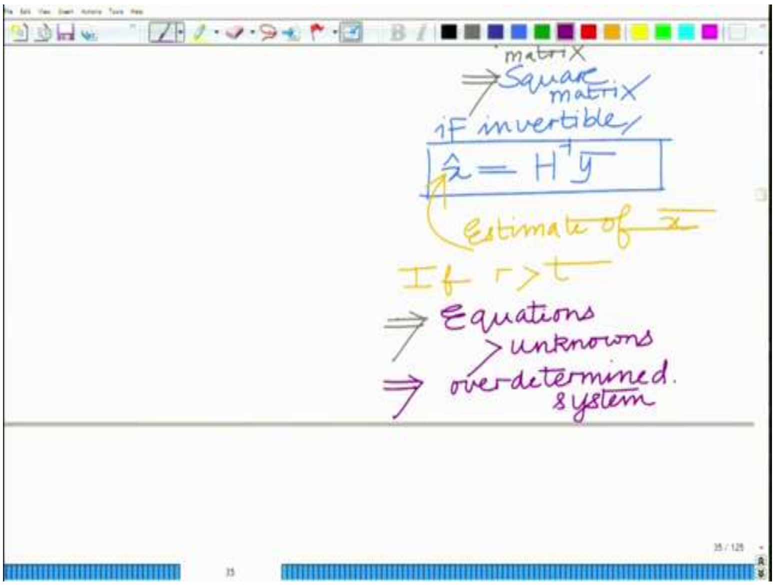
Let us ignore the noise. Then this model reduces to

$$\bar{y} = H\bar{x}$$

So, this is a system of linear equations with r such equations with t unknown variable. The number of unknowns is the transmitted symbols which are meant to be estimated. Consider a simple scenario that number of transmitters is equal to the number of recievers i.e. $r = t$. This means the number of equations is equal to the number of unknowns. Therefore, the matrix H is a square matrix with dimension $t \times t$.

$$\bar{y} \approx H\bar{x}$$

(Refer Slide Time: 03:56)

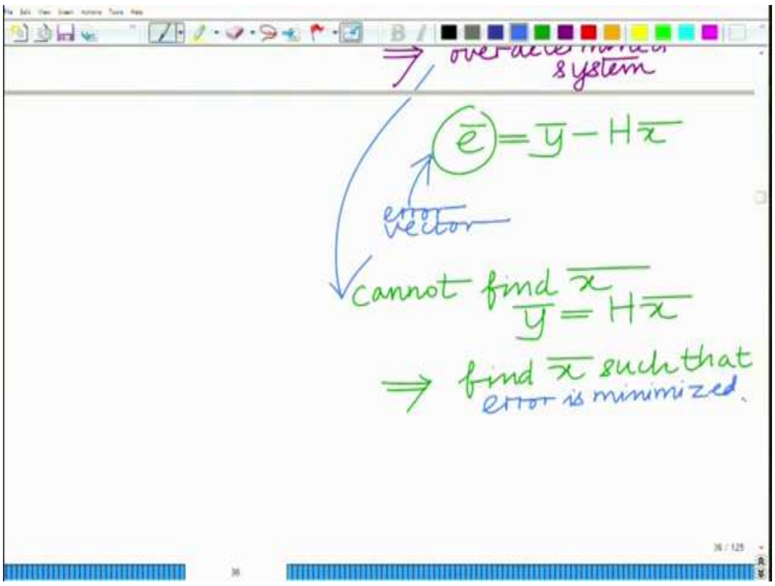


Now if matrix *H* is invertible, then the transmit vector can simply be approximated using received vector. And this estimate of $\bar{x}$ is given as

$$\hat{x} = H^{-1}\bar{y}$$

Now, on the other hand, if *r* is strictly greater than *t*; i.e. $r > t$; then there are more number of equations than the number of unknowns and such systems are known as overdetermined system.

(Refer Slide Time: 05:54)

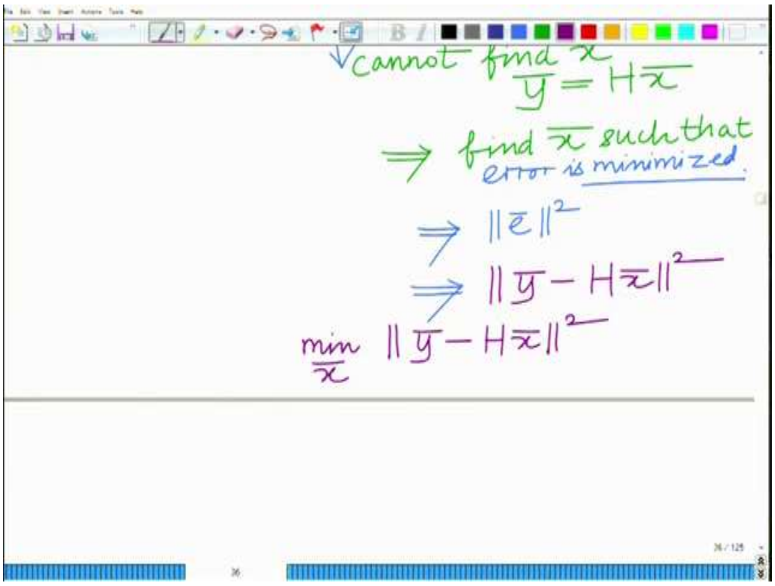


For an over determined system, typically one cannot solve the system of equations which means such systems can only be solves through approximation. This can be done by minimizing the error which is given as

$$\bar{e} = \bar{y} - H\bar{x}$$

Here $\bar{e}$ is the error vector.

(Refer Slide Time: 07:05)

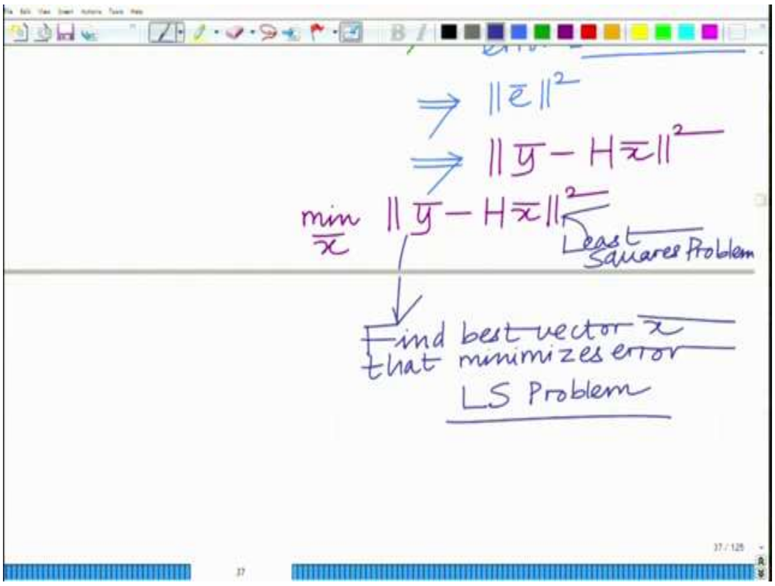


This error is considered as squared norm of this vector, i.e.

$$\|\bar{e}\|^2 = \|\bar{y} - H\bar{x}\|^2$$

Here $\bar{e}$ is the error vector.

(Refer Slide Time: 07:56)

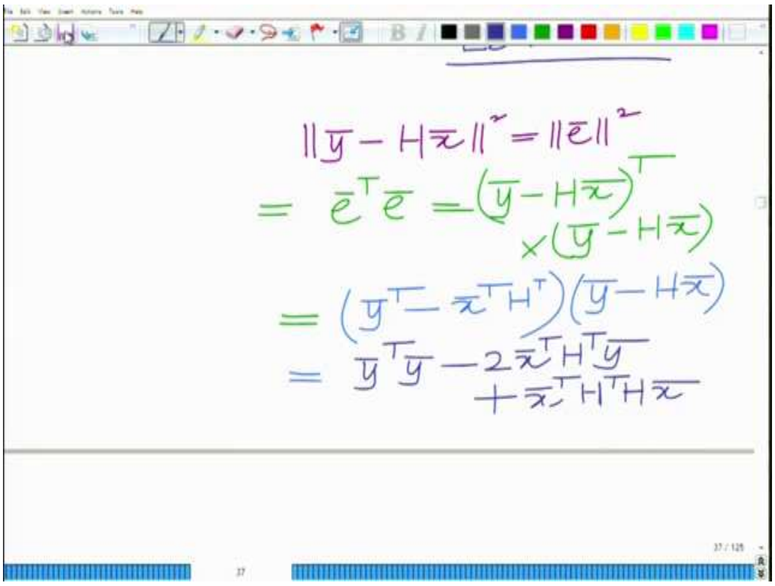


So, to find the best vector $\bar{x}$ that minimizes such error is known as the least square (LS) problem (or in this case the squared norm of the error known as the squared error).

$$\min_{\bar{x}} \|\bar{e}\|^2 = \min_{\bar{x}} \|\bar{y} - H\bar{x}\|^2$$

And such vector $\bar{x}$ is known as the least squares estimate. This is a very important problem that arises frequently in both communications as well as single processing.

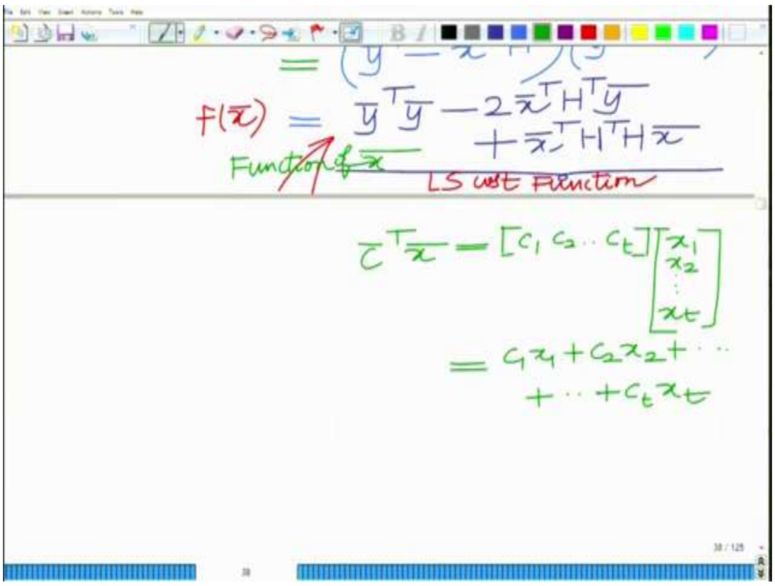
(Refer Slide Time: 09:10)



To simplify this, let us start by simplifying this cost function norm.

$$
\begin{aligned}
\|\bar{e}\|^2 &= \|\bar{y} - H\bar{x}\|^2 \\
&= (\bar{y} - H\bar{x})^T (\bar{y} - H\bar{x}) \\
&= (\bar{y}^T - \bar{x}^T H^T)(\bar{y} - H\bar{x}) \\
&= \bar{y}^T\bar{y} - 2\bar{x}^T H^T \bar{y} + \bar{x}^T H^T \bar{x} H \\
&= f(\bar{x})
\end{aligned}
$$

(Refer Slide Time: 10:45)



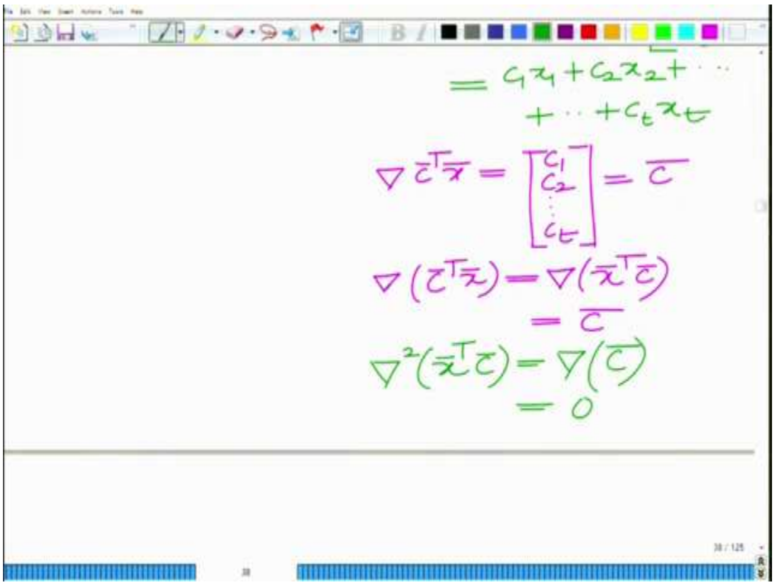
This $f(\bar{x})$ is the cost function and also known as the least squares (LS) cost function.

Now, let us consider a simple function as follows.

$$\bar{C}^T \bar{x} = \begin{bmatrix} c_1 & c_2 & \cdots & c_t \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_t \end{bmatrix}$$

$$= c_1 x_1 + c_2 x_2 + \ldots + c_t x_t$$

(Refer Slide Time: 12:19)



So the first gradient of this matrix with respect to vector $\bar{x}$ is

$$\nabla\left(\bar{C}^T \bar{x}\right) = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_t \end{bmatrix} = \bar{C}$$

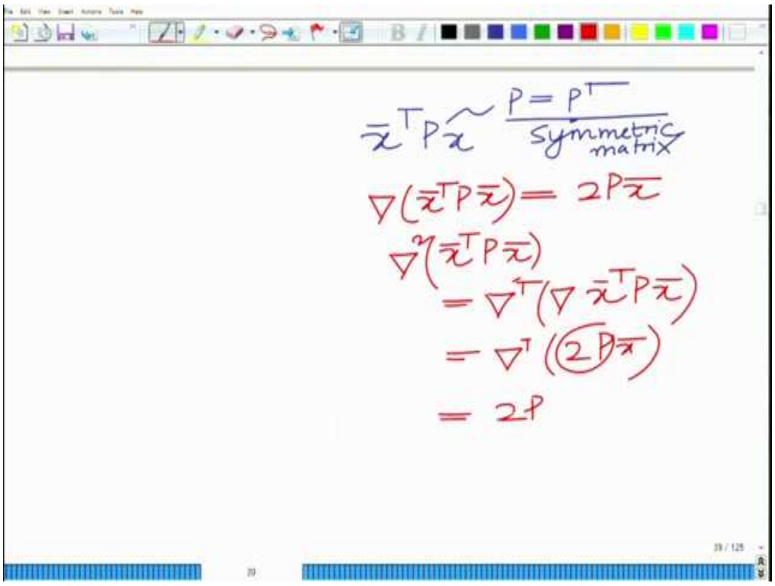
Also

$$\nabla\left(\bar{C}^T \bar{x}\right) = \nabla\left(\bar{x}^T \bar{C}\right) = \bar{C}$$

And as $\bar{C}$ is a constant so Hessian of this function is

$$\nabla^2\left(\bar{x}^T\bar{C}\right)=\nabla\left(\bar{C}\right)=0$$

(Refer Slide Time: 13:46)



On the other hand, let P is the symmetric matrix.

$$P = P^T$$

So the gradient of $\bar{x}^T P\bar{x}$ is

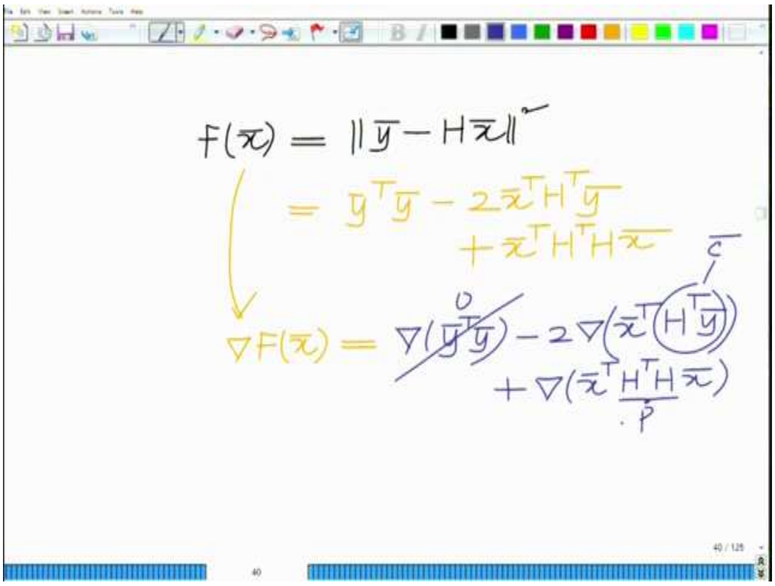
$$\nabla\left(\bar{x}^T P\bar{x}\right)=2P\bar{x}$$

And the Hessian of $\bar{x}^T P\bar{x}$ is

$$\begin{aligned}\nabla^2\left(\bar{x}^T P\bar{x}\right)&=\nabla^T\left(\nabla\left(\bar{x}^T P\bar{x}\right)\right)\\&=\nabla^T\left(2P\bar{x}\right)\\&=2P\end{aligned}$$

So, the hessian of this quadratic term is twice of the matrix P.

(Refer Slide Time: 15:51)

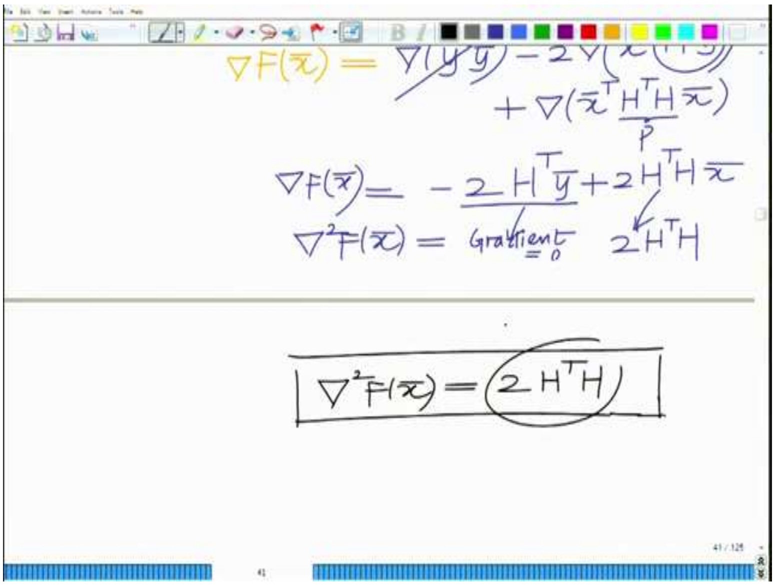


Now coming back to least square cost function $f(\bar{x})$. Let us compute its Hessian.

$$\begin{aligned}
\nabla f(\bar{x}) &= \nabla\left(\bar{y}^T\bar{y} - 2\bar{x}^T H^T \bar{y} + \bar{x}^T H^T \bar{x} H\right) \\
&= \nabla\left(\bar{y}^T\bar{y}\right) - 2\nabla\left(\bar{x}^T H^T \bar{y}\right) + \nabla\left(\bar{x}^T H^T H \bar{x}\right) \\
&= -2\nabla\left(\bar{x}^T \underbrace{H^T \bar{y}}_{\bar{C}}\right) + \nabla\left(\bar{x}^T \underbrace{H^T H}_{P} \bar{x}\right)
\end{aligned}$$

So here $H^T\bar{y}$ is $\bar{C}$ and $H^T H$ is the symmetric matric P.

(Refer Slide Time: 17:27)



It has earlier seen that

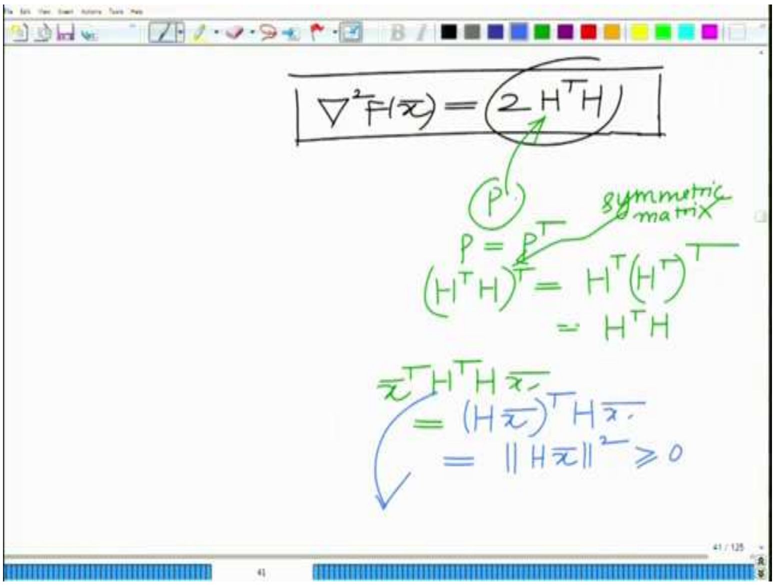
$$\nabla\left(\bar{C}^T\bar{x}\right)=\nabla\left(\bar{x}^T\bar{C}\right)=\bar{C}$$

And $\nabla^2\left(\bar{x}^T\bar{C}\right)=\nabla\left(\bar{C}\right)=0$

Therefore, using the above solution of hessian of symmetric matrix; the hessian of cost function is

$$\begin{aligned}\nabla^2 f\left(\bar{x}\right)&=\nabla\left(-2\nabla\left(\bar{x}^T\bar{C}\right)+\nabla\left(\bar{x}^TP\bar{x}\right)\right)\\&=2P\\&=2H^TH\end{aligned}$$

(Refer Slide Time: 18:40)



Now as $H^T H$ is symmetric matrix therefore

$$H^T H = \left(H^T H\right)^T$$

And further look at $\overline{x}^T \left(H^T H\right)\overline{x}$.

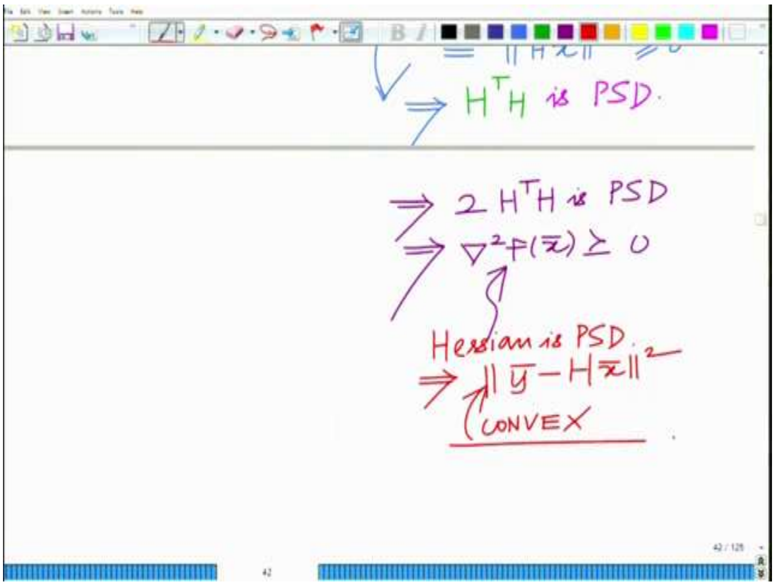
$$\begin{aligned}\overline{x}^T \left(H^T H\right)\overline{x} &= \left(H\overline{x}\right)^T H\overline{x} \\ &= \left\|H\overline{x}\right\|^2 \qquad \geq 0\end{aligned}$$

This means that this matrix $P = H^T H$ is always positive semi definite which implies that $2H^T H$ is also going to be positive semi definite.

Therefore, the Hessian of cost function is positive semi definite and this further implies that the least squares cost function is a convex function.

This is a very important property which helps in designing of the receiver in MIMO system which uses convex optimization techniques.

(Refer Slide Time: 20:20)



This means that this matrix $P = H^T H$ is always positive semi definite which implies that $2H^T H$ is also going to be positive semi definite.

Therefore, the Hessian of cost function is positive semi definite and this further implies that the least squares cost function is a convex function.

This is a very important property which helps in designing of the receiver in MIMO system which uses convex optimization techniques. The convex optimization helps in obtaining an optimized estimate of the transmit vector. And this least squares problem also occurs in several other different scenarios, one of which is to define an efficient receiver for MIMO system.

So our aim is to find the best transmit vector corresponding to a received vector that minimizes the approximation error and closely predicts the transmit vector in MIMO system.