

Advanced 3G and 4G Wireless Communication
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture - 22
MIMO MMSE Receiver and Introduction to SVD

Welcome to another lecture in the course on 3G, 4G wireless Communication Systems. In the last lecture, we had started our discussion on MIMO receiver that is receivers for multiple input multiple output wireless communication systems. And in specifically, we had started looking at linear receivers.

(Refer Slide Time: 00:38)

The image shows a digital whiteboard with handwritten notes. At the top, the word "receiver" is written in blue. Below it, the equation $y = Hx + n$ is written in green. Underneath, the equation $H^{-1}y = x + H^{-1}n$ is written in blue, with a blue circle around H^{-1} and a blue arrow pointing to it. The term $H^{-1}n$ is underlined and labeled "noise" in blue. Below the equation, there are two numbered points in blue: "1. inverse only exists if $r=t$ " and "2. Even for square matrices inverse need not exist." The word "noise" is written in blue below the underlined term.

We said a linear, we said the system can be modeled as y equal to Hx plus n , where H is the channel matrix and we wanted to design receivers for this.

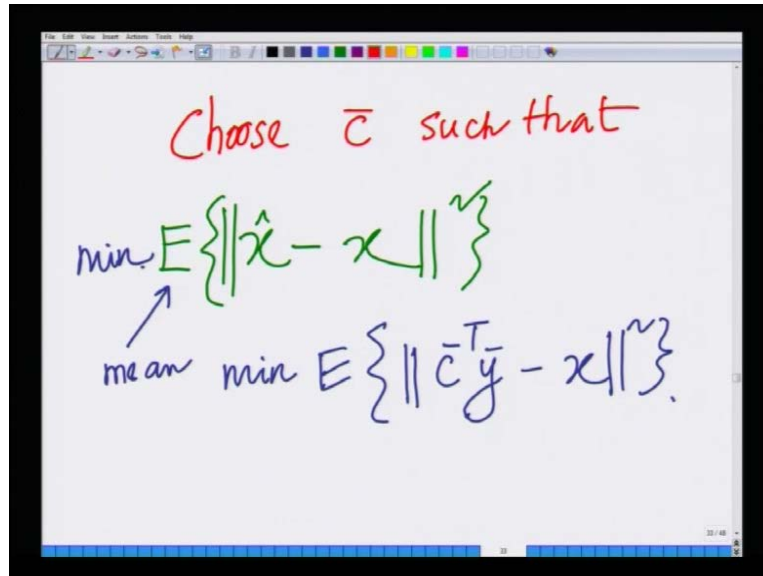
(Refer Slide Time: 00:48)

The image shows a handwritten slide titled "Zero Forcing Receiver" in red. The main equation, enclosed in a green box, is $\hat{x} = (H^T H)^{-1} H^T y$. The variable \hat{x} is circled in purple. A purple arrow points from the text "approximate solution that minimizes the least-squares error" to the equation. Below the equation, the text "Zero-Forcing (ZF) Receiver" is written in blue, with "Zero-Forcing (ZF)" circled in red. A blue arrow points from this circled text to the \hat{x} in the equation.

We said a basic a zero forcing receiver for this system can be denoted as \hat{x} , that is \hat{x} is the estimate of the transmitted symbol vector is given as, \hat{x} equals $H^T H$ that is h transpose H inverse H transpose, that is y . That is I take the received vector y pre-multiply it by H transpose H inverse H transpose and perform my decision on that vector, that is my estimated vector \hat{x} , and this is known as the zero forcing receiver or this is known as the zero forcing receiver technique.

However, we said that the zero forcing receiver has some purpose, most specifically it results in noise enhancement. Hence, we wanted to move to a slightly better receiver and that is basically, we want to consider an MMSE receiver or a Minimum Mean Squared Error receiver for a wireless communication system. And that receiver is simply given as follows, if you look at it.

(Refer Slide Time: 01:56)



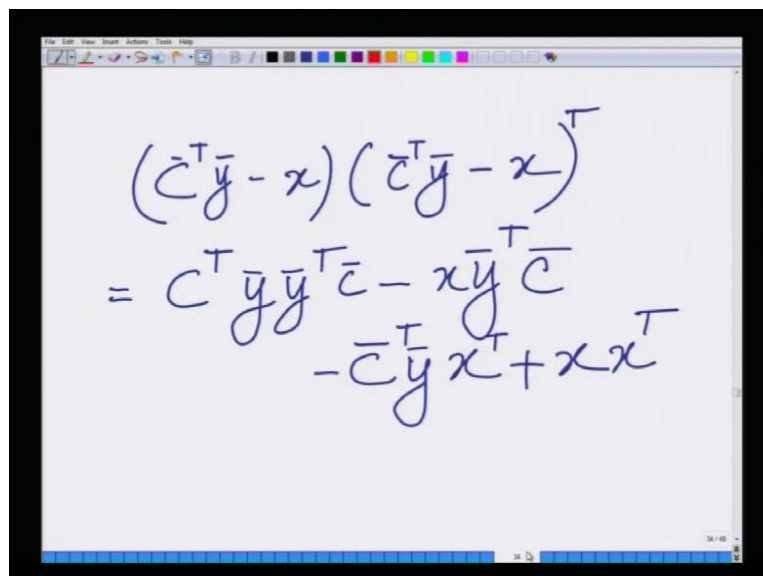
Choose \bar{c} such that

$$\min E \{ \|\hat{x} - x\|^2 \}$$

mean $\min E \{ \|\bar{c}^T \bar{y} - x\|^2 \}$

We want to minimize \hat{x} minus x norms square but look at this we are minimizing this in the average; that is we are minimizing the expected value of norm \hat{x} minus x square. Further we said that \hat{x} is given as, C transpose y that is it is a linear estimate because look at this, this is a linear function, this is a vector c transpose times by y . Hence this is minimum means squared error and also linear, it is also known as, the LMMSE or the Linear Minimum Means Squared Estimator of x .

(Refer Slide Time: 02:35)


$$\begin{aligned} & (\bar{c}^T \bar{y} - x)(\bar{c}^T \bar{y} - x)^T \\ &= \bar{c}^T \bar{y} \bar{y}^T \bar{c} - x \bar{y}^T \bar{c} \\ &\quad - \bar{c}^T \bar{y} x^T + x x^T \end{aligned}$$

We started our derivation of the Linear Minimum Mean Squared Error we want to we will start looking at this product $C^T y x - C^T y x$ minus $C^T y x$ minus $x^T C$ transpose that is this is nothing but, $C^T y x$ whole square.

(Refer Slide Time: 02:59)

$$E(\bar{y}\bar{y}^T) = R_{yy}$$

covariance matrix of y

$$E(x\bar{y}^T) = R_{xy}$$

cross covariance

$$E(yx^T) = R_{yx}^T = R_{xy}$$

We said this is given as $C^T y y^T C - C^T y x + x x^T C - C^T y x$ plus $x x^T$ transpose. We also saw that this is nothing but, expected $y y^T$ transpose is R_{yy} expected $x y^T$ transpose is R_{xy} expected $y x^T$ transpose is R_{yx} transpose it is also R_{yx} , that is the correlation between vector y and x , which is also R_{xy} transpose.

(Refer Slide Time: 03:25)

$$E\{(C^T \bar{y} - x)(C^T \bar{y} - x)^T\}$$

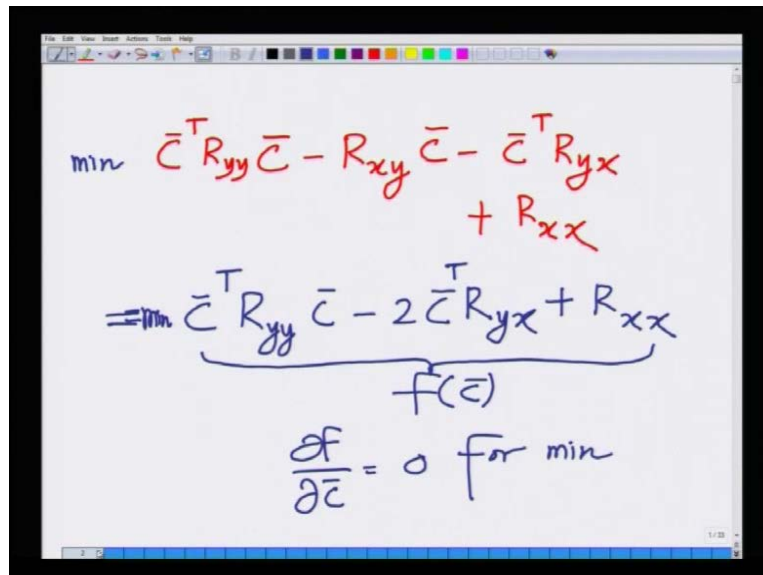
$$= E\{C^T \bar{y} \bar{y}^T C - x \bar{y}^T C - C^T \bar{y} x^T + x x^T\}$$

$$= C^T R_{yy} C - R_{xy} \bar{C} - \bar{C}^T R_{yx} + R_{xx}$$

Hence using that simplification coming back to this, I can write this as $\bar{C}^T R_{yy} \bar{C} - R_{xy} \bar{C} - \bar{C}^T R_{yx} + R_{xx}$; that is when I take the expected value of this, that is the expected value of the error.

I want to minimize the average error, hence I am considering the expected value of this error and expected value of this error reduces to this function. That is $\bar{C}^T R_{yy} \bar{C} - R_{xy} \bar{C} - \bar{C}^T R_{yx} + R_{xx}$ where these are the respective correlation and cross correlation matrices that we introduced earlier.

(Refer Slide Time: 04:47)



$$\begin{aligned} \min \quad & \bar{C}^T R_{yy} \bar{C} - R_{xy} \bar{C} - \bar{C}^T R_{yx} + R_{xx} \\ = \min \quad & \underbrace{\bar{C}^T R_{yy} \bar{C} - 2 \bar{C}^T R_{yx} + R_{xx}}_{f(\bar{C})} \\ & \frac{\partial f}{\partial \bar{C}} = 0 \text{ for min} \end{aligned}$$

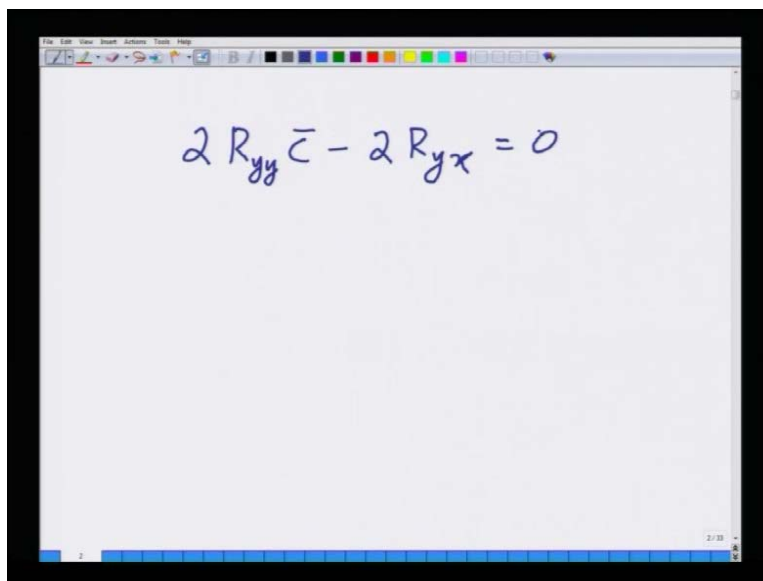
So, let me rewrite this again this can be written as $\bar{C}^T R_{yy} \bar{C} - R_{xy} \bar{C} - \bar{C}^T R_{yx} + R_{xx}$ I want to minimize this, this is the mean squared error I want to minimize this. Also observe that $R_{xy} \bar{C}$ is nothing but, $\bar{C}^T R_{yx}$ hence I will write this as which is the same again as this term hence I will write this as a succinct expression, which is $\bar{C}^T R_{yy} \bar{C} - 2 \bar{C}^T R_{yx} + R_{xx}$ and I want to minimize this, this is my mean squared error.

If you want if you can see, this is clearly a function of \bar{C} , this is clearly a function of the vector \bar{C} that is a vector the linear combiner that you want to use for the estimation of x . This means squared error is a function of that vector \bar{C} , I want to choose that \bar{C} vector that combiner that minimizes this mean squared error. So, what needs to be done is fairly obvious

in this context, I need to differentiate this error as a function of \bar{C} and set it to zero that is what we had done earlier also in the zero forcing case.

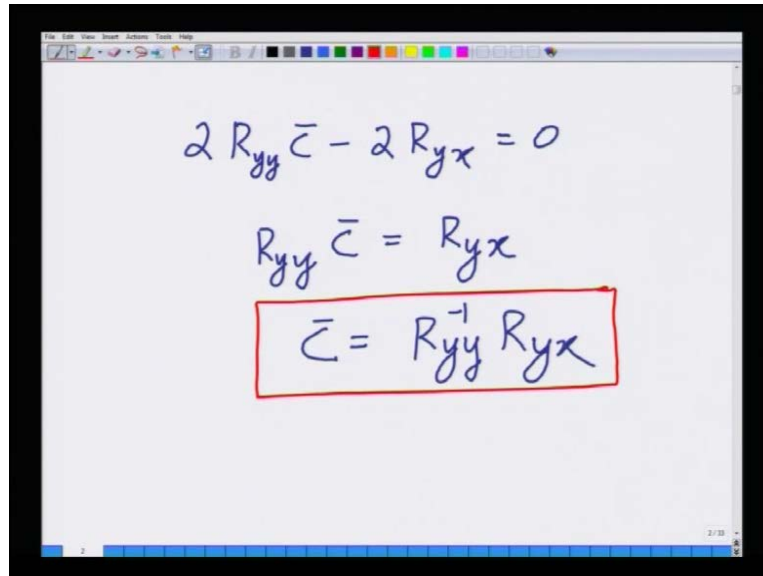
Except now, that we are considering the average error, so I am going to differentiate this with respect to \bar{C} vector \bar{C} and set it 0, so $\frac{df}{d\bar{C}} = 0$ for minimum. Now, the derivative of this we had seen vector differentiation earlier before is nothing but, the differentiation with respect to each component of that vector.

(Refer Slide Time: 07:12)

A screenshot of a digital whiteboard interface. The whiteboard has a toolbar at the top with various drawing tools like pens, eraser, and selection tools. The main area is white and contains the handwritten equation $2 R_{yy} \bar{C} - 2 R_{yx} = 0$ in blue ink. The bottom of the window shows a blue taskbar with a clock and other system icons.

And this can be written as $2 R_{yy} \bar{C} - 2 R_{yx} = 0$. Remember earlier we said derivative of $\bar{C}^T R_{yx}$ is nothing but r_{yx} that is derivative of \bar{C}^T times some vector is nothing but that vector; when it differentiated with respect to \bar{C} .

(Refer Slide Time: 07:45)

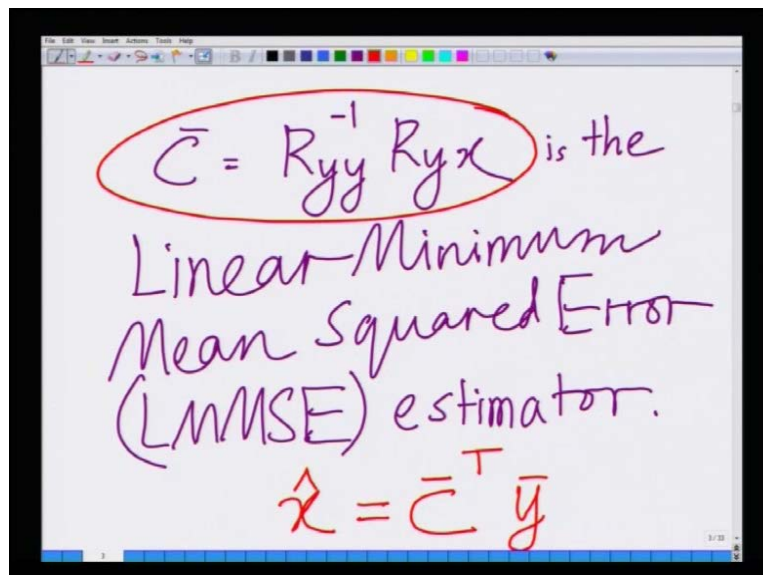


A digital whiteboard interface showing three equations. The first equation is $2R_{yy}\bar{C} - 2R_{yx} = 0$. The second equation is $R_{yy}\bar{C} = R_{yx}$. The third equation, $\bar{C} = R_{yy}^{-1}R_{yx}$, is enclosed in a red rectangular box.

$$2R_{yy}\bar{C} - 2R_{yx} = 0$$
$$R_{yy}\bar{C} = R_{yx}$$
$$\bar{C} = R_{yy}^{-1}R_{yx}$$

Hence the \bar{C} the minimum \bar{C} is such that $R_{yy}\bar{C}$ equals R_{yx} and hence the optimal \bar{C} that minimizes the means squared error is \bar{C} equals $R_{yy}^{-1}R_{yx}$. The \bar{C} that minimizes mean squared error is nothing but the solution of this equation. Hence that \bar{C} is $R_{yy}^{-1}R_{yx}$ this is the optimal this is the minimum mean squared linear minimum mean squared error estimator for the quantity x .

(Refer Slide Time: 8:37)



A digital whiteboard interface showing text and equations. The equation $\bar{C} = R_{yy}^{-1}R_{yx}$ is circled in red. Below it, the text 'is the Linear Minimum Mean Squared Error (LMMSE) estimator.' is written. At the bottom, the equation $\hat{x} = \bar{C}^T \bar{y}$ is written in red.

$$\bar{C} = R_{yy}^{-1}R_{yx}$$

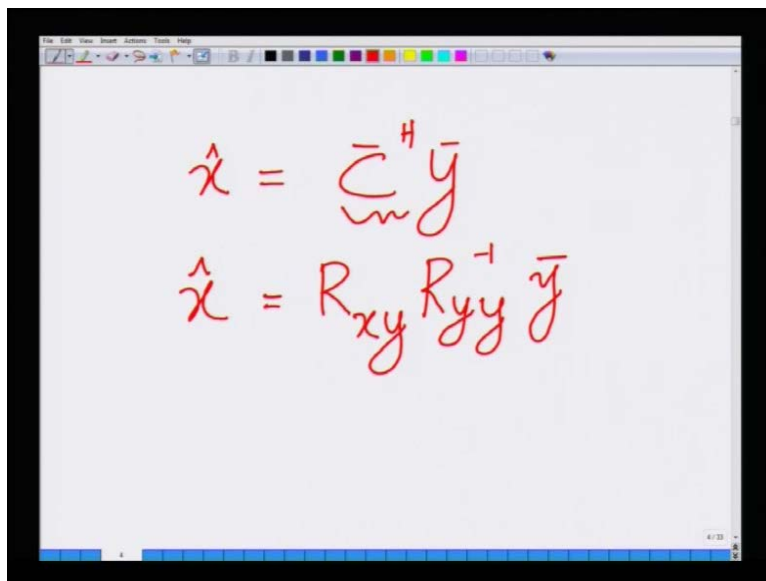
is the Linear Minimum Mean Squared Error (LMMSE) estimator.

$$\hat{x} = \bar{C}^T \bar{y}$$

So, this \bar{C} which is equal to let me write that down here clearly \bar{C} equals $R_{yy}^{-1}R_{yx}$ is the linear minimum mean squared estimator that is the LMM Linear Minimum Mean

Squared Error that is LMMSE estimator. So, this is the expression for the LMMSE estimator. And we know that once we know \bar{C} the estimate is nothing but, \hat{x} equals \bar{C} transpose times \bar{y} that is I compute this vector \bar{C} and whatever vector \bar{y} are you seen I multiply \bar{C} transpose times \bar{y} that is my LMMSE estimate. Now, there are two things first here I have done this example for real vectors, it can be generalized to complex vectors, in a fairly straight forward way; and is except that instead of this transpose use hermitian.

(Refer Slide Time: 10:14)



$$\hat{x} = \bar{C}^H \bar{y}$$

$$\hat{x} = R_{xy} R_{yy}^{-1} \bar{y}$$

So, the LMMSE estimator for complex vectors is nothing but, \hat{x} equals \bar{C} hermitian \bar{y} . Second is slightly a shuttle difference, shuttle extension what I have done it is, I have done this example for a case where \bar{y} is a vector and x is a scalar. However, this expression holds perfectly well, where \bar{y} and \bar{x} are both vectors in which case C is a matrix.

That is our case because remember in a MIMO system \bar{y} the received symbol vector is a vector \bar{x} the transmit symbol vector is a vector. What I am saying is although I have not shown you the proof explicitly the same expression will can be used, that is \hat{x} equals $R_{xy} R_{yy}^{-1} \bar{y}$ I am saying this expression can still be used when \bar{y} and \bar{x} are vectors. So, now we know how to construct the MMSE vector MMSE estimator let us see how to derive these quantities and how to derive the MMSE estimator for our given MIMO system.

(Refer Slide Time: 11:38)

Handwritten notes on a whiteboard:

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

$$E \{ \bar{x} \bar{x}^H \}$$

Annotations:

- Covariance of transmitted symbols (pointing to $E \{ \bar{x} \bar{x}^H \}$)
- Transmit covariance (pointing to $E \{ \bar{x} \bar{x}^H \}$)

$$E \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_t \end{bmatrix} \begin{bmatrix} x_1^* & x_2^* & \dots & x_t^* \end{bmatrix} \right\}$$

So, first let me go back to the system model our system model we know is \bar{y} equals $H \bar{x}$ plus \bar{n} . Hence expected now first let us start with expected $\bar{x} \bar{x}^H$, expected $\bar{x} \bar{x}^H$ this is nothing but, the covariance of the transmitted symbols. I expected $\bar{x} \bar{x}^H$ is nothing but, the covariance of the transmitted symbols this is also termed as the transmit covariance, this is also termed as a transmit covariance. This is nothing but, I can write this as expected let me expand these vectors out \bar{x} is a t dimensional transmit symbol vector. So, this is x_1, x_2 up to x_t , \bar{x}^H is nothing but, x_1^* conjugate that is I take transpose and conjugate vector.

(Refer Slide Time: 13:06)

Handwritten notes on a whiteboard:

$$R_{xx} = E \left\{ \begin{bmatrix} |x_1|^2 & x_1 x_2^* & \dots & x_1 x_t^* \\ x_2 x_1^* & |x_2|^2 & \dots & x_2 x_t^* \\ \vdots & \vdots & \ddots & \vdots \\ x_t x_1^* & x_t x_2^* & \dots & |x_t|^2 \end{bmatrix} \right\}$$

$$= \begin{bmatrix} P_d & 0 & 0 & 0 \\ 0 & P_d & & \\ \vdots & & \ddots & \\ 0 & 0 & \dots & P_d \end{bmatrix}$$

$$= P_d I_t$$

Hence this can be written as \mathbf{R}_{xx} the transmit covariance is nothing but, expected $\mathbf{x} \mathbf{x}^H$ conjugate, which is nothing but, norm \mathbf{x} square. So, on the diagonals we will have norm \mathbf{x}_1 square, norm \mathbf{x}_2 square, which is $\mathbf{x}_2 \mathbf{x}_2^H$ conjugate and so on; of diagonal we will have elements such that $\mathbf{x}_2 \mathbf{x}_1^H$ conjugate $\mathbf{x}_1 \mathbf{x}_2^H$ conjugate $\mathbf{x}_1 \mathbf{x}_t^H$ conjugate $\mathbf{x}_t \mathbf{x}_1^H$ conjugate and so on. And now if you see this expression the diagonal terms are nothing but, expected \mathbf{x}_1 square expected \mathbf{x}_2 square expected \mathbf{x}_t square these are nothing but, the transmit powers.

Because if I take $\mathbf{x}_1 \mathbf{x}_2^H$ if I take a $\mathbf{x}_1 \mathbf{x}_1^H$ square norm \mathbf{x}_1 square average it over a large amount of amount of time that is nothing but, the transmit power. Hence the diagonal terms are the transmit power, the off diagonal terms are the correlation between the symbols transmitted on the different transmit antennas. Now, if the symbols are un correlated which is a reasonable assumption, because you want to transmit independent in special multiplexing, you want to transmit independent symbols on the transmit antennas.

Because you want to transmit more information that is you want to transmit independent information symbols which means these different symbols are un correlated. Hence the cross correlation $\mathbf{x}_1 \mathbf{x}_2^H$ conjugate $\mathbf{x}_2 \mathbf{x}_2^H$ conjugate $\mathbf{x}_t \mathbf{x}_1^H$ conjugate all these expectations these expectations are in fact, 0.

Hence this has a very simple structure very simple and the very elegant structure, this is nothing but, P_d power of each data symbols on the diagonal. And all the off diagonal terms are 0, which is the same we have seen this is P_d times the identity matrix of dimension t . So, the transmit covariance \mathbf{R}_{xx} is nothing but, the transmit power times the identity matrix that is the first thing. Now, we will use this to derive the other covariance matrices.

(Refer Slide Time: 15:34)

The image shows a handwritten derivation of the received signal covariance matrix R_{yy} on a whiteboard. The derivation is as follows:

$$\begin{aligned}
 R_{yy} &= E\{\bar{y} \bar{y}^H\} \\
 &= E\{(H\bar{x} + \bar{n})(H\bar{x} + \bar{n})^H\} \\
 &= E\{(H\bar{x}\bar{x}^H H^H + \underbrace{\bar{n}\bar{x}^H H^H}_0 + \underbrace{H\bar{x}\bar{n}^H}_0 + \bar{n}\bar{n}^H)\} \\
 &= H R_{xx} H^H + \sigma_n^2 I
 \end{aligned}$$

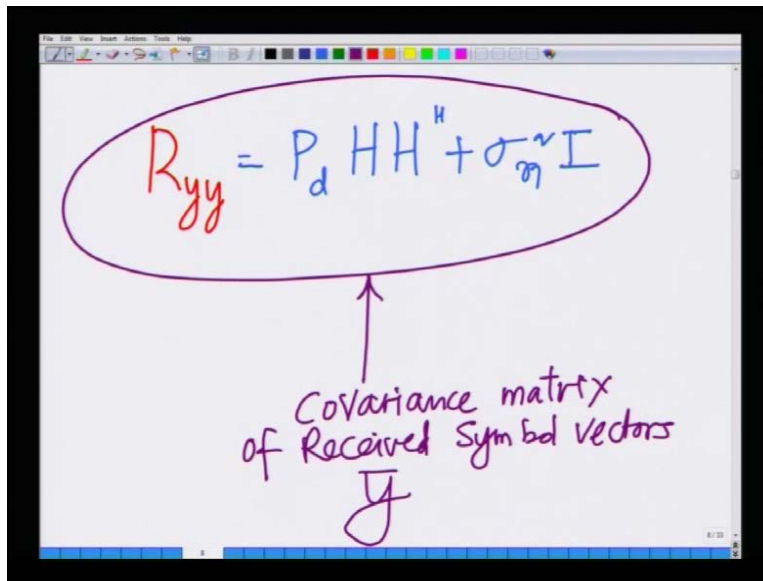
The whiteboard has a standard software interface at the top with various icons and a color palette. The handwriting is in purple and blue ink, with red underlines and annotations indicating that the cross terms are zero.

For instance, now I want to derive R_{yy} which is expected $\bar{y} \bar{y}^H$ hermitian, which is the same thing as expected $H \bar{x} \bar{x}^H H^H + \bar{n} \bar{n}^H$ hermitian, which is nothing but, expected. Let me write this down $H \bar{x} \bar{x}^H H^H + \bar{n} \bar{n}^H$ hermitian $H \bar{x} \bar{x}^H H^H + \bar{n} \bar{n}^H$ hermitian $H \bar{x} \bar{x}^H H^H + \bar{n} \bar{n}^H$ hermitian.

Now, if you look at this expression even before simplifying this you can see that the noise and the transmitted symbol they are uncorrelated, because this is the noise at the receiver this is the transmitted symbol there is no correlation. Hence this is 0 by that same for that same by that same token this is also 0. So, both the cross terms are 0, what survives are expected $H \bar{x} \bar{x}^H H^H + \bar{n} \bar{n}^H$ hermitian, which is if I simplify expected $H \bar{x} \bar{x}^H H^H + \bar{n} \bar{n}^H$ hermitian.

I take the expectation operator inside that is $\bar{x} \bar{x}^H$ hermitian that is R_{xx} into $H R_{xx} H^H + \bar{n} \bar{n}^H$ hermitian, this is nothing but, the covariance of the noise. And you know we assume the noise covariance is $\sigma_n^2 I$ that is the different noises are uncorrelated and the power in each noise is σ_n^2 hence this is square listed forward this is $\sigma_n^2 I$.

(Refer Slide Time: 17:41)

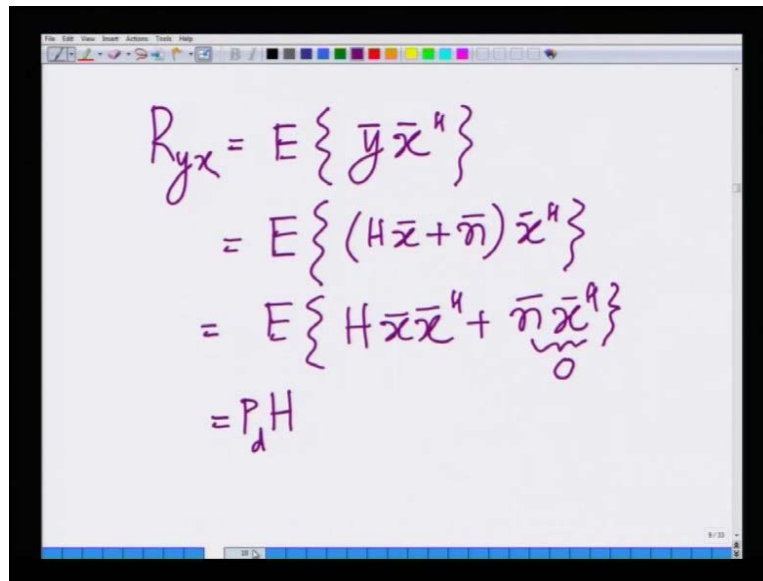


A screenshot of a digital whiteboard showing a handwritten equation $R_{yy} = P_d H H^H + \sigma_n^2 I$. The equation is circled in purple. Below the equation, an arrow points upwards to the expression, with the handwritten text "Covariance matrix of Received Symbol vectors" and the symbol \bar{y} below it.

And I will now substitute for the covariance the transmit covariance R_{xx} which is P_d times I the identity matrix that gives me a simple expression R_{yy} equals P_d because it is a scalar $H H^H$ hermitian plus $\sigma_n^2 I$. And this is nothing but, the covariance matrix of received symbol vectors \bar{y} , this is nothing but, the covariance matrix of the received symbol vectors \bar{y} .

One thing one other thing remains that is the cross covariance that is R_{yx} remember to compute the MMSE estimator we need both (Refer Slide Time: 18:38) R_{xy} which is nothing but, R_{yx} hermitian and R_{yy} . We have computed R_{yy} , so we need to still compute R_{xy} or R_{yx} , so I will compute R_{yx} and one is simply the hermitian of the other and that is also fairly straight forward to compute.

(Refer Slide Time: 18:56)

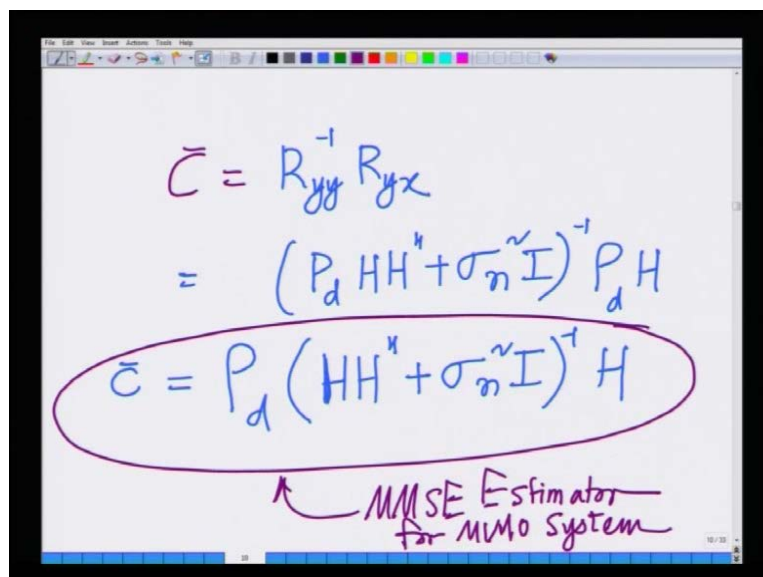


A screenshot of a digital whiteboard showing the derivation of the cross-correlation matrix R_{yx} . The equations are written in purple ink:

$$\begin{aligned} R_{yx} &= E\{\bar{y} \bar{x}^H\} \\ &= E\{(H\bar{x} + \bar{n}) \bar{x}^H\} \\ &= E\{H\bar{x} \bar{x}^H + \underbrace{\bar{n} \bar{x}^H}_0\} \\ &= P_d H \end{aligned}$$

We can see R_{yx} equals expected $\bar{y} \bar{x}^H$, which is nothing but, expected $H \bar{x} \bar{x}^H$ plus $\bar{n} \bar{x}^H$ this is nothing but, expected $H \bar{x} \bar{x}^H$ plus $\bar{n} \bar{x} \bar{x}^H$ as we know those correlation between the noise and \bar{x} this is 0. This is simply H times expected $\bar{x} \bar{x}^H$, which is the transmit covariance which is nothing but, P_d times identity, hence this is P_d times H .

(Refer Slide Time: 19:49)



A screenshot of a digital whiteboard showing the derivation of the MMSE estimator expression. The equations are written in blue ink, with the final result circled in purple:

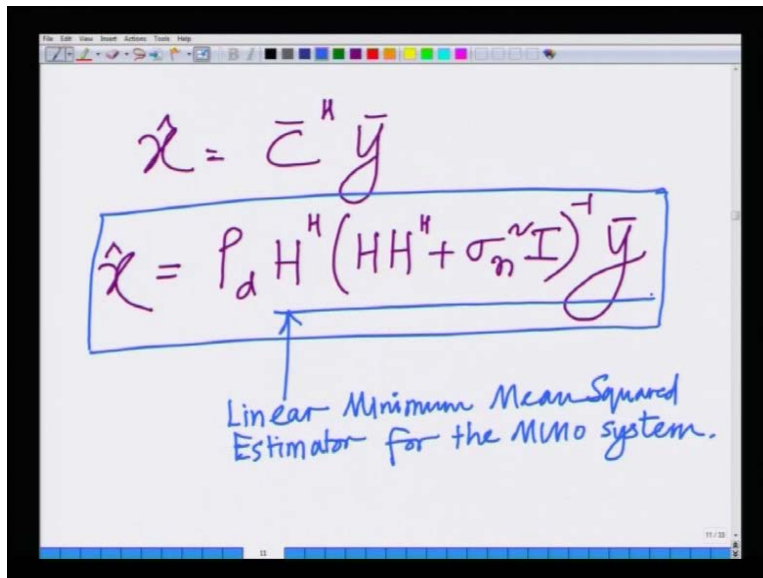
$$\begin{aligned} \bar{c} &= R_{yy}^{-1} R_{yx} \\ &= (P_d H H^H + \sigma_n^2 I)^{-1} P_d H \\ \bar{c} &= P_d (H H^H + \sigma_n^2 I)^{-1} H \end{aligned}$$

Below the circled equation, an arrow points to it with the text: "MMSE Estimator for MIMO System".

Now, we have an expression we know that \bar{c} the MMSE estimator is simply $R_{yy}^{-1} R_{yx}$, which is simply P_d , which is simply $P_d H H^H$ plus σ_n^2

squared I inverse into $R y x$ which is P_d times. Hence the P_d is a scalar, so I can move it to the front of the expression this is $P_d H^H H$ hermitian plus $\sigma_n^2 I$ inverse into $H \bar{C}$ bar, which is this is the MMSE estimator for our MIMO system. This is the MMSE estimator for this is the MMSE estimator for the MIMO system, which is P_d which is \bar{C} equals P_d times $H^H H$ hermitian plus $\sigma_n^2 I$ inverse into H . Where H is the channel matrix σ_n^2 is the noise power at the receiver P_d the transmitted data power.

(Refer Slide Time: 21:22)



The image shows a handwritten derivation on a whiteboard. At the top, the equation $\hat{x} = \bar{C}^H \bar{y}$ is written in purple. Below it, a more complex equation is boxed in blue: $\hat{x} = P_d H^H (H H^H + \sigma_n^2 I)^{-1} \bar{y}$. A blue arrow points from the text "Linear Minimum Mean Squared Estimator for the MIMO system." to the boxed equation.

Now, we also know that the estimate of the symbol is given as \hat{x} equals $\bar{C}^H \bar{y}$, which is essentially means $P_d H^H H$ hermitian plus $\sigma_n^2 I$ inverse into \bar{y} that is \hat{x} and this is nothing but, the MMSE estimate of \hat{x} . So, this is the LMMSE estimator this is the Linear Minimum Mean Squared Estimator for the MIMO system, this is the linear minimum mean squared estimator for the MIMO system alright.

(Refer Slide Time 22:58)

$$\begin{aligned}
 & H^H (P_d H H^H + \sigma_n^2 I)^{-1} \\
 &= (P_d H^H H + \sigma_n^2 I)^{-1} H^H \\
 \Leftrightarrow & (P_d H^H H + \sigma_n^2 I) H^H \\
 &= H^H (P_d H H^H + \sigma_n^2 I) \\
 \Leftrightarrow & P_d H^H H H^H + \sigma_n^2 H^H \\
 &= P_d H^H H H^H + \sigma_n^2 H^H
 \end{aligned}$$

So, and now let me do us let me do a slightly technical manipulation to show the equivalence to between two forms of this MIMO MMSE estimator. What I want to show here is that this quantity that we have over here, if you look at this quantity that we have over here I simply want to show that this quantity H hermitian $P_d H H$ hermitian plus sigma n squared I inverse equals $P_d H$ hermitian H plus sigma n squared I inverse into H hermitian.

It might seem a priory that these two things are exactly the same but, look at this here in the brackets I have $H H$ hermitian. Here I have H hermitian H while $H H$ hermitian is r cross r dimensional H hermitian H is t cross t dimensional. And also remember the number of transmit antennas is smaller than the number of receiver antennas that is the case we are considering, which means inversion of this matrix, which is t cross t dimensional is more is simpler than inversion of this matrix, which r cross r dimensional alright.

So, it has a nice structure I just want to show the equivalence and that is fairly simple I take this inverse to that side I take this inverse this side. All I need to show is these two are equal if and only if P_d into if and only if I am taking this matrix to this side that is P_d into H hermitian H plus sigma n squared I into H hermitian equals. Now, I take this to this side that is H hermitian into $P_d H H$ hermitian plus sigma n squared I and now you can clearly receive when I make this simplification I will just multiply this thing out. That is $P_d H$ hermitian $H H$ hermitian plus sigma n square H hermitian equals $P_d H$ hermitian $H H$ hermitian plus sigma

n squared H hermitian. You can see this two quantities are clearly equal they are one and the same thing, hence these are equal, hence these quantities on the top are also equal.

(Refer Slide Time: 25:29)

$$\hat{x} = P_d H^H (P_d H H^H + \sigma_n^2 I)^{-1} \bar{y}$$

$$\hat{x}_{\text{MMSE}} = P_d (P_d H^H H + \sigma_n^2 I)^{-1} H^H \bar{y}$$

Hence we have nothing but, when I look at the MIMO estimator that is \hat{x} let me rewrite this \hat{x} equals $P_d H^H (P_d H H^H + \sigma_n^2 I)^{-1} \bar{y}$ equals from the derivation from the simplification that we did above equals $P_d P_d H^H H + \sigma_n^2 I)^{-1} H^H \bar{y}$. So, this is another expression for the MIMO MMSE estimator more simpler expression. And this is what you will find popularly used in text books if you refer to the text books that we have mentioned as references for the course.

You will find this definition of the MIMO MMSE estimator used more popularly, which is the MIMO, MMSE, LMMSE estimate is nothing but, P_d times P_d into $H^H H$ plus $\sigma_n^2 I$ inverse into H^H into \bar{y} . Now, let me make another slide observation here, so this is the MIMO, so this is the MIMO MMSE estimator again I repeat both these are exactly equal.

Remember this estimator that we derived here and this one that we have they are they are equivalent alright they are exactly result in the same estimate is just convenience. And one form is most commonly used and the reason I said is that inversion of this might be slightly easier. Because, more commonly we have number of transmit antennas much smaller than the number of receive antennas.

(Refer Slide Time: 27:26)

The image shows a handwritten derivation on a whiteboard. At the top, it states $H = h$. Below this, the estimator \hat{x} is given as $\hat{x} = P_d \left(\frac{h^*}{P_d |h|^2 + \sigma_n^2} \right) \bar{y}$. This is then simplified to $= \frac{P_d h^*}{\sigma_n^2} \bar{y}$, where the fraction $\frac{P_d h^*}{\sigma_n^2}$ is circled in purple. To the right of the circled term, green text reads: "MIMO MMSE Estimator does NOT result in Noise enhancement".

Let me consider this to see show the main difference between a MIMO and zero forcing receiver. For instance in a MIMO system consider a single input single output case that is h equals the scalar h then the MIMO receiver reduces to \hat{x} equals P_d times h conjugate times $P_d H H$ hermitian H is a scalar is nothing but, P_d times norm h square P_d times norm h square plus σ_n^2 times \bar{y} .

Now, if the magnitude of h is small that is as h tends to zero this reduces to σ_n^2 this norm of h square is significant compare to σ_n^2 this tends to $P_d h$ conjugate divided by $\sigma_n^2 \bar{y}$. Hence it does not blow up to infinity remember in the zero forcing case as h tends to 0, 1 over 1 over h progressively moves towards infinity there by resulting in noise enhancement.

However, in this case because there is the σ_n^2 , which is also known as the regularization term that prevent this from becoming very large it is bounded by σ_n^2 ; hence this is bounded it does not result in noise enhancement. So, this hence the MIMO MMSE estimator by this simple example for the size of case we can show that MIMO MMSE estimator does not result in noise enhancement. Thus it is superior compare to the zero forcing receivers. Remember earlier we said there is a problem with the zero forcing receiver, the problem is it result in noise enhancement the MIMO MMSE estimator avoids that problem that is what we are saying here.

(Refer Slide Time: 29:40)

$$\begin{aligned}\hat{\mathbf{x}} &= \mathbf{P}_d (\mathbf{P}_d \mathbf{H}^H \mathbf{H} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{H}^H \bar{\mathbf{y}} \\ &\text{at high SNR} \\ &\approx \mathbf{P}_d (\mathbf{P}_d \mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \bar{\mathbf{y}} \\ &\approx (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \bar{\mathbf{y}}\end{aligned}$$

Zero forcing receiver

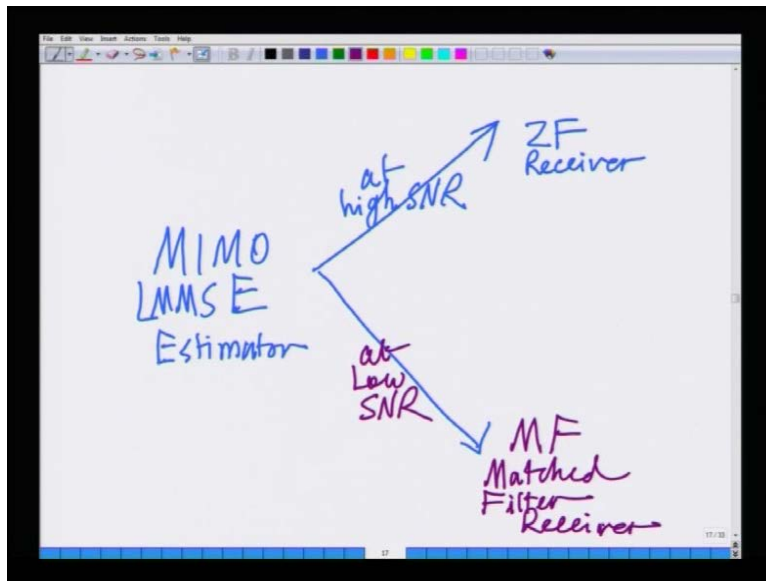
And let me find out, let me now go to one final property of this MIMO MMSE estimator $\hat{\mathbf{x}}$ as we already noted, $\hat{\mathbf{x}}$ is given as $\mathbf{P}_d (\mathbf{P}_d \mathbf{H}^H \mathbf{H} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{H}^H \bar{\mathbf{y}}$. Now, in high signal to noise power ratio let us assume that \mathbf{P}_d is at high SNR let me write this down at high at high SNR \mathbf{P}_d is much larger compare to σ_n^2 which means this approximately becomes \mathbf{P}_d into $\mathbf{P}_d \mathbf{H}^H \mathbf{H}$ σ_n^2 is negligible inverse $\mathbf{H}^H \mathbf{H}$ $\bar{\mathbf{y}}$ this \mathbf{P}_d is a scalar. So, \mathbf{P}_d into \mathbf{P}_d inverse is nothing but, 1. So, this is approximately equal to $\mathbf{H}^H \mathbf{H}$ inverse $\mathbf{H}^H \mathbf{H}$ $\bar{\mathbf{y}}$ and we have seen this before this is nothing but, the zero forcing receiver this is nothing but, the zero forcing receiver.

(Refer Slide Time: 31:04)

The image shows a whiteboard with handwritten mathematical expressions. The top expression is $\hat{x} = P_d (P_d H^H H + \sigma_n^2 I)^{-1} H^H \bar{y}$. Below it, a note says "at low SNR". The next expression is $\approx P_d (\sigma_n^2 I)^{-1} H^H \bar{y}$. This is followed by an approximation $\approx \frac{P_d}{\sigma_n^2} H^H \bar{y}$, where the fraction $\frac{P_d}{\sigma_n^2} H^H \bar{y}$ is circled in blue. An arrow points from the text "Matched Filter" below to the circled expression.

So, what happens at low SNR at low SNR we have first let us write the expression $P_d P_d^H H^H H + \sigma_n^2 I$ inverse into $H^H H$ hermitian H plus sigma n squared I inverse into H^H hermitian. At low s n r P_d is negligible compare to sigma n square, hence this results in P_d into sigma n squared I inverse into H^H hermitian y bar which is nothing but, $P_d I$ inverse is $I P_d$ over sigma n squared into H^H hermitian y bar. If you remember this, this is looks similar to the MRC that is we had for a single input multiple receive we had received x is nothing but, $H^H y$. So, this is nothing but, a match filter, so I claim at high SNR this is an approximation let me write that this is an approximation at low SNR at low SNR. So, at low SNR it reduces to a matched filter. So, the MIMO MMSE estimator has a very interesting property at SNR if I look at the MIMO MMSE estimator.

(Refer Slide Time: 32:39)

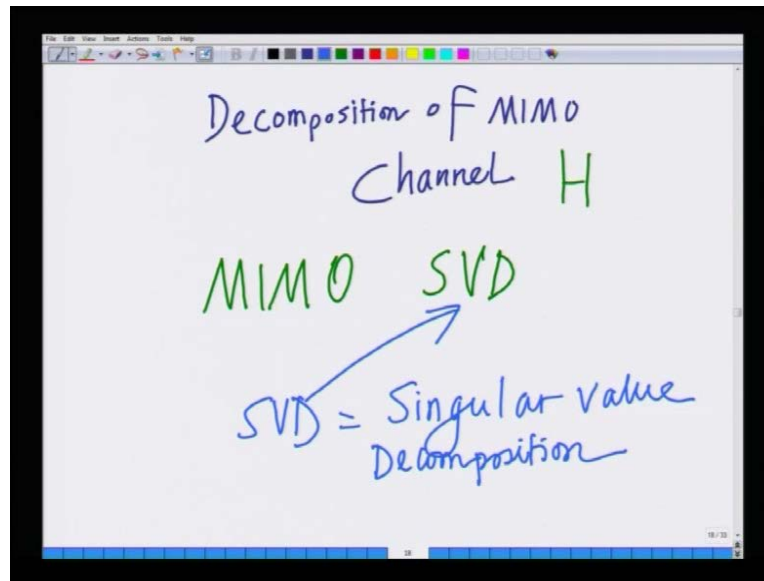


So, I write at the MIMO MMSE or, in fact, the LMMSE estimator at low SNR and high SNR at high SNR it reduces to the zero forcing receiver. At low SNR it reduces to the MF the maximal ratio combiner but, the maximal ratio combiner for a MIMO system which is simply $H^H H$ hermitian the matched filter it reduces to the matched filter the matched filter receiver.

So, we have seen in summary in linear receivers we have two receivers one is the zero forcing receiver we said that it is slightly it is simple but, slightly inferior because it results in noise enhancement. And then we have proposed a robust MIMO MMSE receiver in this MIMO MMSE receiver as this interesting property where at high SNR first it does not result in noise enhancement.

And second it has interestingly high SNR it reduces to the zero forcing receiver, low SNR it reduces to the matched filter. So, that is complete the discussion on MIMO linear receivers there are also MIMO non-linear receivers. However, I will differ that discussion to slightly later first we have to cover something that is slightly more important which is the decomposition of a MIMO channel. So, next we are going to start with a very critical aspect of MIMO the very key aspect of MIMO, which is a key to understanding all the properties of a MIMO wireless communication system, which essentially is this singular value decomposition of a MIMO communication system.

(Refer Slide Time: 34:58)



So, let me talk about, so let me start with, so let me start with the decomposition of the MIMO channel. So, we are going to talk about a decomposition of the MIMO communication channel of the MIMO channel channel matrix H . In fact, specifically we are going to look at one specific decomposition that is the singular value decomposition we are going to look at MIMO SVD, where SVD equals Singular Value Decomposition. So, we are going to talk about MIMO SVD, where SVD stands for the Singular Value Decomposition.

(Refer Slide Time: 36:08)

$r \geq t$

$$H = U \Sigma V^H$$

t columns

t rows

$$U = \begin{bmatrix} | & | & \dots & | \\ u_1 & u_2 & \dots & u_t \\ | & | & \dots & | \end{bmatrix}$$
$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_t \end{bmatrix}$$
$$V^H = \begin{bmatrix} v_1^H \\ v_2^H \\ \vdots \\ v_t^H \end{bmatrix}$$

$\|u_i\|^2 = 1$ $u_i^H u_j = 0$ if $i \neq j$

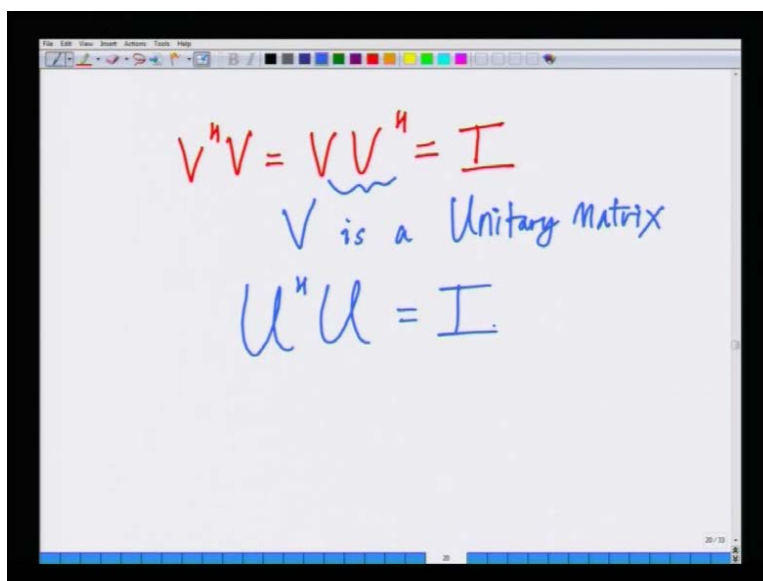
$\|v_i\|^2 = 1$ $v_i^H v_j = 0$ if $i \neq j$

And the singular value decomposition of a matrix of any matrix. In fact, it can be expressed as follows H the matrix H is given as $U \Sigma V^H$ hermitian where these matrices are such that we have U is nothing is columns $u_1 u_2$, so on up to U_t . That is this has t columns into the singular matrix Σ $\Sigma_1 \Sigma_2$ up to Σ_t into and this is along the main diagonal into $v_1^H v_2^H v_t^H$ hermitian and these are the different rows these are the different rows.

So, this has t columns remember I am assuming here that r is greater than equal to t that the number of receive antennas is greater than or equal to the number of transmit antennas this has t columns, this matrix has t rows. And the properties and these are not any t columns just any t columns and t rows but, the properties of these matrices are such that. That first of all this columns what can we say about this columns the columns are ortho normal that is $u_i^H u_j = \delta_{ij}$ while $u_i^H u_j = 0$ if $i \neq j$.

That is the norm of each column is 1 and cross product the dot product between different column that is $u_i^H u_j$ which is $u_1^H u_2 u_1^H u_3 u_2^H u_3$ so on and so forth are 0. That is the cross product dot product is orthogonal they are unit norm these are known as these are known as ortho normal columns. Similarly, for the case of the rows of the matrix V which is we say $v_i^H v_j = \delta_{ij}$ $v_i^H v_j = 0$ if $i \neq j$. So, the rows column and rows of matrix U and V are ortho normal and then you can verify it. So, this is an $r \times t$ matrix U is an $r \times t$ matrix V is a $t \times t$ matrix.

(Refer Slide Time: 39:27)



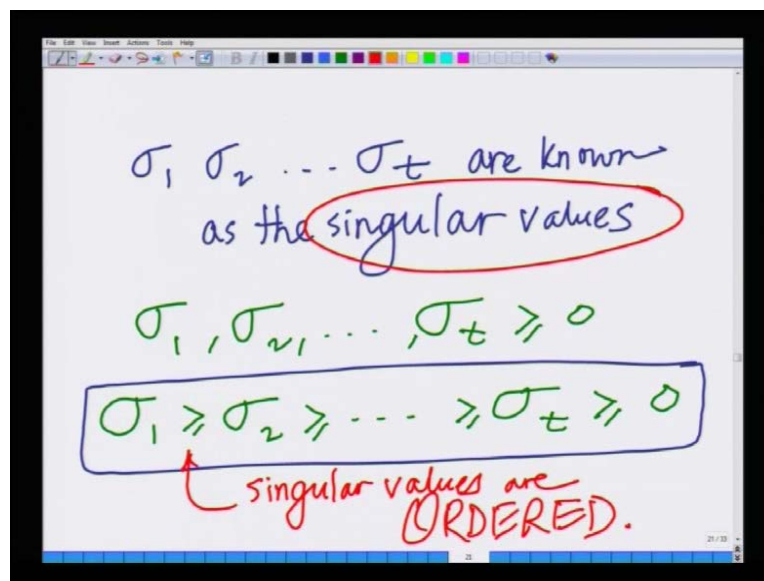
$$V^H V = \underline{V V^H} = I$$

V is a Unitary Matrix

$$U^H U = I$$

So, we have the following property we have V hermitian V equals V V hermitian equals identity, this is known as a unitary matrix this is known as V is a unitary V is known as a unitary matrix. While U the matrix U satisfies the property that U hermitian U equals identity observe that U U hermitian is not equal to identity, because we have assumed r is greater than or equal to t . If r is equal to t then U U hermitian is also identity but, the more general result is that U hermitian U is identity. And the other important aspect is this singular values, what is the structure of this singular values.

(Refer Slide Time: 10:30)



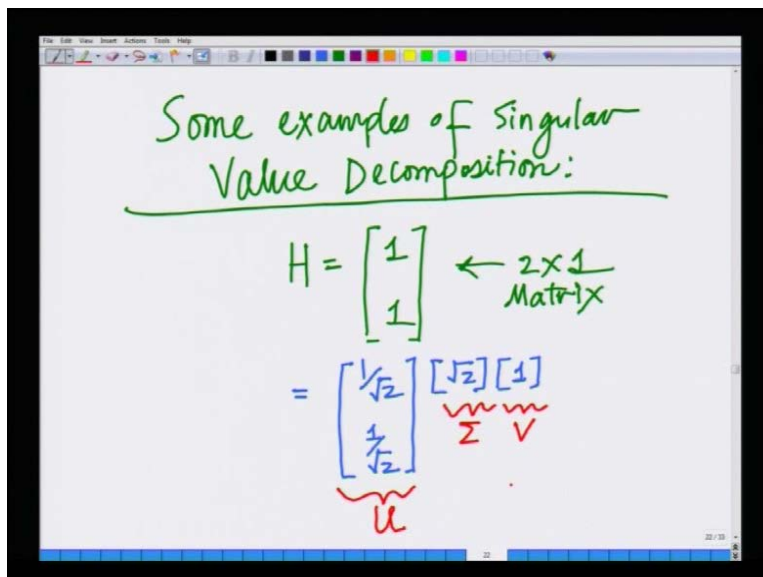
The structure of this singular values is such that first $\sigma_1, \sigma_2, \dots, \sigma_t$ are known as the singular value these are known as the these are known as the singular values of the matrix h there they are important for any matrix h . Such that further they satisfy the property they are such that $\sigma_1, \sigma_2, \dots, \sigma_t$ are greater than or equal to 0, that is they are non negative; either all the singular values either zero or greater than zero that is they are non negative. Further this singular values are order that is I cannot place them any order but, the order is such that $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_t \geq 0$.

So, these singular values are, so this is the decreasing order, so these singular values are ordered this is the order property of singular values. So, we say it as singular values are ordered alright. So, the singular values have to be given in that degrees I got, so these singular

values are non negative and they are in decreasing order and this singular value decomposition exist any matrix including the non square matrix.

For instance you might remember another related another decomposition which is known as the Eigen value decomposition. However, the Eigen value decomposition exist only for square matrices, unlike an Eigen value decomposition a singular value decomposition exist for matrix of any dimension that is also a non square matrices. As we are going to see some examples, for instance let me take a simple example of a 2 cross 1 matrix, so let me take a simple example of a two cross.

(Refer Slide Time: 42:45)



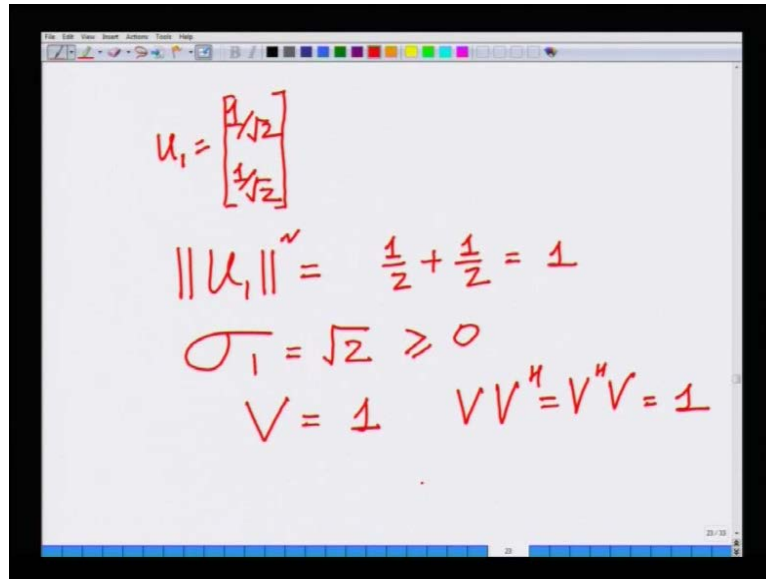
Some examples of Singular Value Decomposition:

$$H = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \leftarrow 2 \times 1 \text{ Matrix}$$

$$= \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}}_U \underbrace{\begin{bmatrix} \sqrt{2} \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_V$$

So, let me start with some examples singular value. So, let me start with some examples of singular value decomposition let us consider with a basic example that is I consider a matrix H equals 1 comma 1 this is a 2 cross 1 matrix this is a essentially a 2 cross 1 matrix corresponds to when you have two receive antennas and one transmit antenna. Remember this matrix is a 2 cross 1 matrix it is non square it does not have an Eigen value decomposition which means; however, it still has a singular value decomposition that is what we are going to show and this also satisfies their criterion that r is greater than or equal to t . I can write this as equals 1 over root 2 1 over root 2 into the single the single term matrix square root of 2 into the single dimensional matrix 1 alright. Now I claim this is my matrix u , this is my matrix σ this is my matrix this is my matrix v alright.

(Refer Slide Time: 44:38)



The image shows a whiteboard with handwritten mathematical derivations in red ink. The derivations are as follows:

$$u_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$
$$\|u_1\|^2 = \frac{1}{2} + \frac{1}{2} = 1$$
$$\sigma_1 = \sqrt{2} \geq 0$$
$$V = 1 \quad VV^H = V^H V = 1$$

So, I claim this is my matrix u this is my matrix v now why is this my matrix, why is this satisfies ortho normality. For instance if we take the vector 1 over root 2 1 over root 2 we can see that the norm of that vector if I take u_1 equals this thing I can see that norm u_1 square equals 1 over root 2 square that is half plus half equals 1 .

So, this is unit norm further if you go since the singular value of this is square root of 2 . So, the singular value σ_1 equals square root of 2 there is only one singular value and it is greater than equal to 0 . And further the matrix p it is trivially a unitary matrix, because if you look at this matrix v equals 1 , which means $VV^H = V^H V = 1$ which is essentially identity matrix. Hence this is a trivial singular value decomposition, we have shown that singular value decomposition exists for a non square matrix. So, this is a singular value decomposition, a simple singular value decomposition.

(Refer Slide Time: 45:59)

Ex 2:

$$H = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{5} \end{bmatrix}$$

NOT a valid SVD

Singular Values are NOT ordered

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$\sigma_1 = 1$
 $\sigma_2 = \sqrt{5}$
 $\sigma_2 \geq \sigma_1$

Let me take a slightly more slightly more what is shuttle example of a singular value decomposition. Let me take an example another example two of a singular value decomposition which is the diagonal matrix 1 comma 0 0 comma square root of 5. This is the matrix H that I am considering now one might tempted to think that the singular value since this is the diagonal matrix is a simpler value decomposition can simply be written as 1 0 0 1 that is the identity matrix times 1 0 0 square root of 5 times 1 0 0 1.

So, this satisfies all criteria for instance this has ortho normal columns this as ortho normal columns the similar values are non negative. However this is not a valid singular value decomposition this is not a valid decomposition this is not a valid s v d, why is this not a valid singular value decomposition. Because, look at the singular values these are sigma 1 equals 1 sigma 2 equals square root of 5 these are not ordered, in fact we have sigma 2 greater than or equal to sigma 1.

(Refer Slide Time: 47:39)

$$H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Hence the matrix is unitary.

$$\sigma_1 = \sqrt{5} \quad \sigma_2 = 1$$

$$\sigma_1 \geq \sigma_2$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$VV^H = I$$

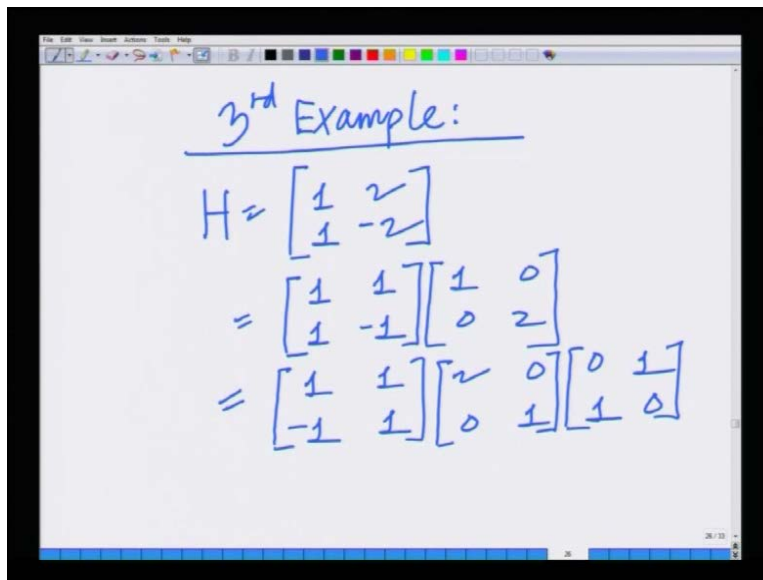
So, the singular values, so the singular values are not order hence this is not a valid SVD instead the valid SVD of this can be written as and you can verify this the valid SVD of this can be written as $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ that is nothing but the identity matrix that is whose columns are flipped. Look at this, this is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is identity matrix this is $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ times now square root of 5 comma $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ again $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Now, you can see the singular values σ_1 equals square root of 5 σ_2 equals 1 $\sigma_1 \geq \sigma_2$.

Now, you can see the singular values order further you can see that the matrix this is of course, ortho normal columns, in fact, it is unitary. In fact, you can also see $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ that is $V V^H$ hermitian equals $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ the product this is nothing but, $V V^H$ hermitian this is equal to $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ this is nothing but identity. Hence we have $V V^H$ hermitian equals identity, hence this matrix is unitary hence the matrix is unitary. So, that is the slightly more tricky example of a singular value decomposition; let me take a third and final example to illustrate the singular of entire decomposition.

Normally the singular values decomposition that is complicated you cannot compute it ordinarily by on a pen on a paper am just illustrating this simple examples. So, has to give you a feel of what the singular values decomposition looks like. If you want to compute the singular value decomposition of matrices in in the cases of its not straight forward there is a mat lab command known as SVD.

If you go into mat lab and you type SVD or you type the help section of SVD you will find out how to use this command. So, typing that command SVD you can get this singular value decomposition of the matrices, which you can which you can more matrices in general, which you cannot compute by a simple paper and pen kind of computation.

(Refer Slide Time: 50:01)



3rd Example:

$$H = \begin{bmatrix} 1 & 2 \\ 1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

So, let me illustrate a third example for SVD for a singular value decomposition H equals 1 comma 1 2 comma minus 2, this is clearly see this is a slightly more, slightly more general matrix. However, I can simplify this as 1 comma 1 1 comma minus 1 into 1 0 0 2 that is this matrix can be written as 1 1 1 minus 1 into 1 into the matrix 1 0 0 2. Now, again there is its tempting to simply claim the singular values are sigma 1 equals 1 sigma 2 equals 2. However, there are several things that I would be taken into account first of all this matrix is not unit, does not it is does not have ortho normal columns because you can clearly see the magnitude is not one.

Also this singular values are not ordered, so I will first take care of ordering these things what I will do is, I will first order this by simply reordering these matrices I can reorder this as follows. You can see and you can verify that this can be written as this matrix times again the same matrix that we had seen before the identity matrix, which is flipped.

Now, you can see that whatever these are these are not still the singular values but these are ordered. Now, I need now this matrix v here which is 0 1 1 0 this you can see is clearly a unitary matrix except this u matrix it does not have normal columns.

(Refer Slide Time: 52:01)

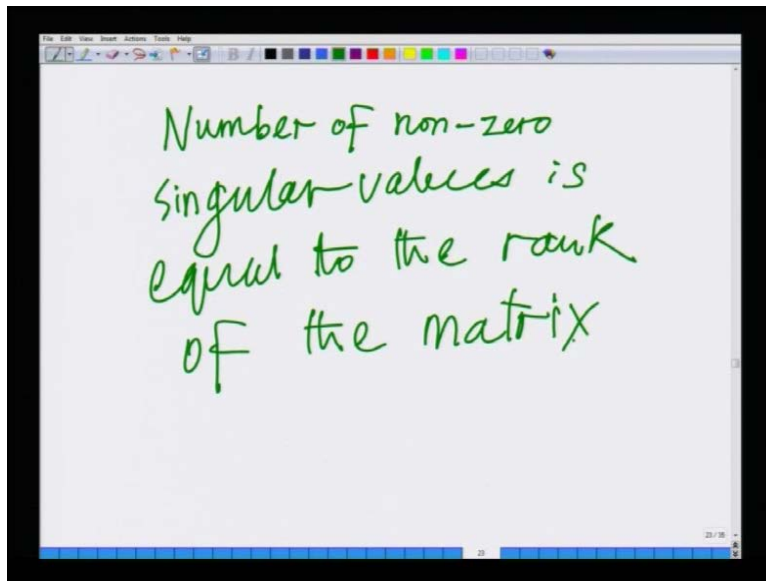
$$H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

ordered singular values $\left\{ \begin{array}{l} \sigma_1 = 2\sqrt{2} \geq 0 \\ \sigma_2 = \sqrt{2} \geq 0 \end{array} \right.$

$$\sigma_1 \geq \sigma_2$$

So, I will divide by the magnitude of each to make this as normal columns and I can write this as H equals $\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}}$. And now you can see the matrix U matrix V is orthogonal ortho unitary U is also unitary, in fact, it has normal columns now that is the norm of each column of U is 1. And also this singular values $\sigma_1 = \sqrt{2}$, $\sigma_2 = \sqrt{2}$, $\sigma_3 = 0$. And further these are ordered. In fact, we have $\sigma_1 \geq \sigma_2$, hence these are ordered, hence these are ordered singular values; hence this is a valid singular value decomposition.

(Refer Slide Time: 53:32)



So, we have looked at the singular values decomposition further I also forgot to mention one small point. In fact, the number of non zero singular values for a non zero singular values is equal to the rank of the number of non zero singular values is equal to the rank of the matrix. And you can clearly see the rank of the matrix is one number of non zero singular values is 1. And in this case, you can see the rank of the matrix is 2, so number of non zero singular values is 2. Beginning with this understanding of the singular value decomposition, we will explore the properties of the MIMO wireless communication channel in the next lecture.