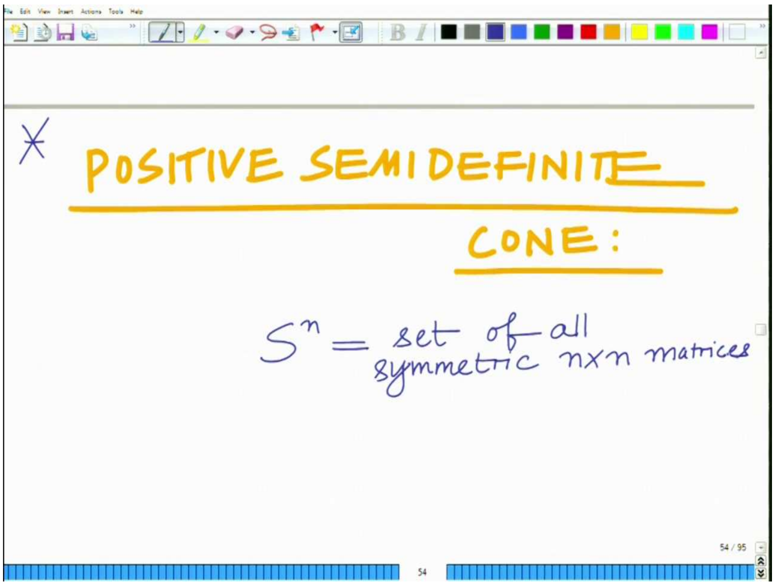
Applied Optimization for Wireless, Machine Learning, Big Data

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

Lecture – 17
Positive Semi Definite Cone And Positive Semi Definite (PSD) Matrices

Hello. Welcome to another module in this massive open online course. Let us continue this discussion by looking at the set of Positive Semi Definite Matrices which is also known as the Positive Semi Definite Cone.

(Refer Slide Time: 00:30)

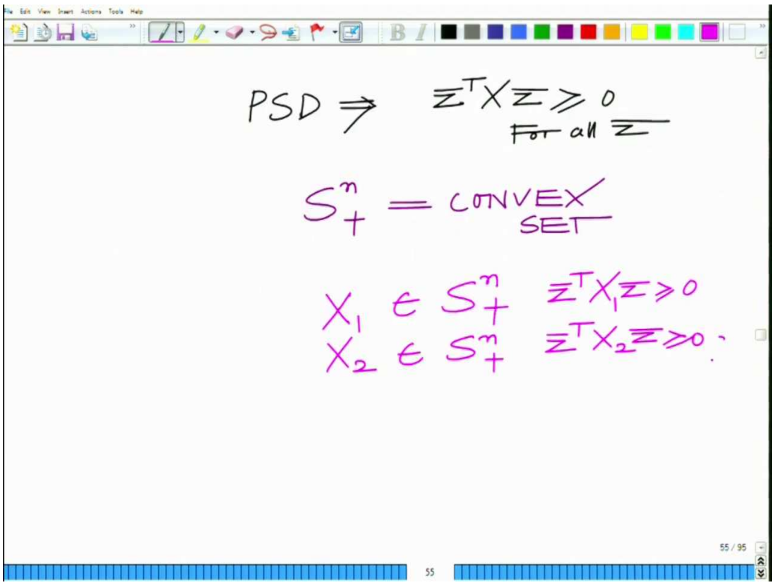


Consider a set of all symmetric $n \times n$ matrices as S^n. Let us define another set S_+^n as the set of all those element of S^n which are greater than or equal to zero. So S_+^n is defined as

$$S_+^n = \left\{X \in S^n \mid X \geq 0\right\}$$

Hence X defines positive semi definite (PSD) matrix and S_+^n is the set of all positive semi definite $n \times n$ matrices.

(Refer Slide Time: 03:40)



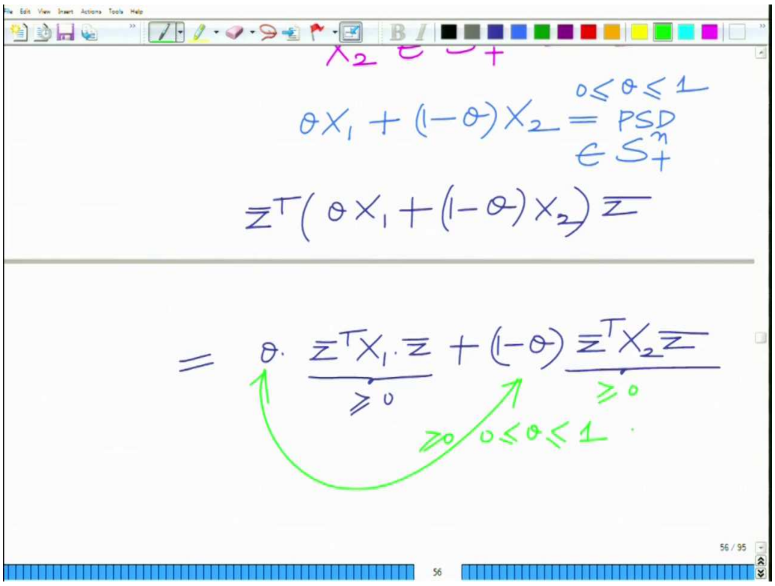
So, PSD matrix implies that for all n dimensional vector $\bar{Z}$;

$$\bar{Z}^T X \bar{Z} \geq 0$$

To show that the set of all symmetric positive semi definite matrices S_+^n is a convex set, take any two positive semi definite matrix X_1 and X_2, both belong to S_+^n. So;

$$X_1 \in S_+^n \Rightarrow \bar{Z}^T X_1 \bar{Z} \geq 0,$$
$$X_2 \in S_+^n \Rightarrow \bar{Z}^T X_2 \bar{Z} \geq 0$$

(Refer Slide Time: 05:10)



So, to prove that S_+^n is a convex set, show that a convex combination of these X_1 and X_2 is a PSD matrix and hence it belong to S_+^n. Therefore for $0 \leq \theta \leq 1$; $\theta X_1 + (1-\theta) X_2$ is a convex combination of X_1 and X_2. Hence

$$
\begin{aligned}
&\overline{Z}^T \left(\theta X_1 + (1-\theta) X_2\right) \overline{Z} \\
&= \overline{Z}^T \theta X_1 \overline{Z} + \overline{Z}^T (1-\theta) X_2 \overline{Z} \\
&= \theta \overline{Z}^T X_1 \overline{Z} + (1-\theta) \overline{Z}^T X_2 \overline{Z}
\end{aligned}
$$

As θ is a positive scalar and it could have maximum value of 1, then $(1-\theta)$ will also be a positive scalar. Therefore considering the definition of PSD matrix for X_1 and X_2, the convex combination

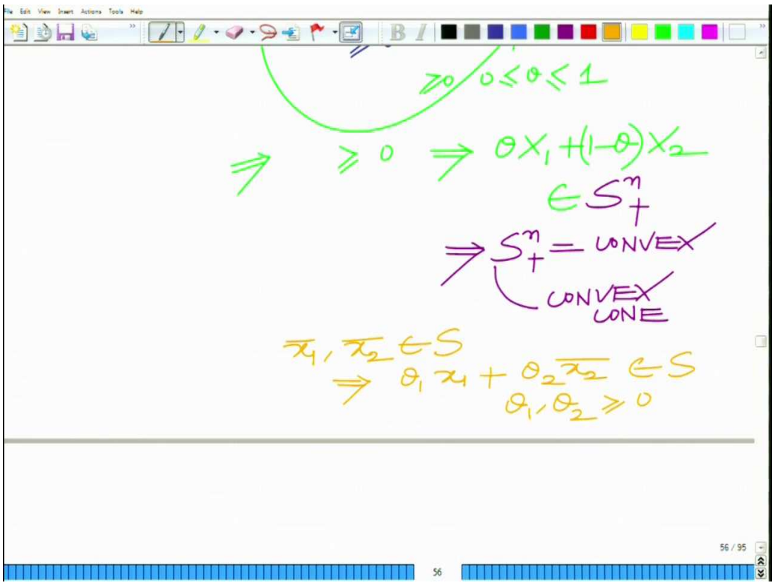
$$\overline{Z}^T \left(\theta X_1 + (1-\theta) X_2\right) \overline{Z} \quad \geq \quad 0$$

And it implies that this convex combination is a PSD matrix and hence it belongs to the set S_+^n.

$$\theta X_1 + (1-\theta) X_2 \in S_+^n$$

Therefore it verifies that S_+^n is convex.

(Refer Slide Time: 06:27)

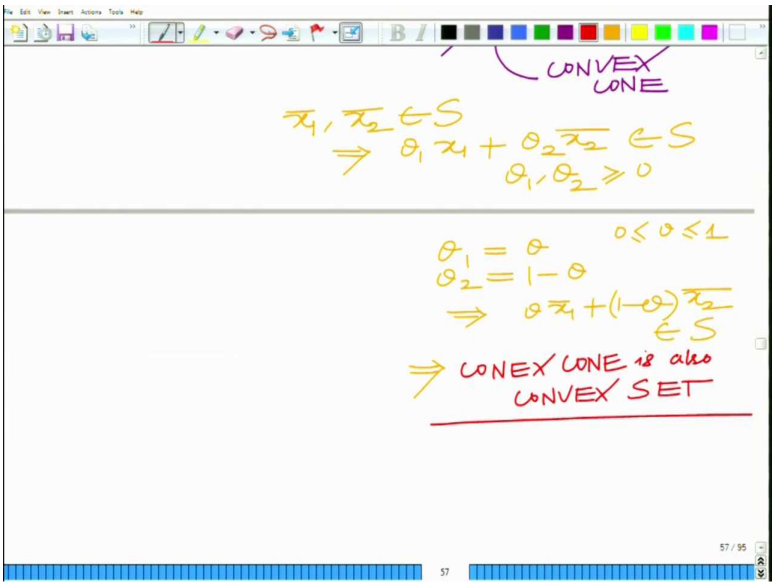


In fact, S_+^n is a convex cone. The definition of a convex cone is that if two convex vectors $\overline{x}_1, \overline{x}_2 \in S$ then set S is a convex cone only if for all $\theta_1, \theta_2 \geq 0$;

$\theta_1 \overline{x}_1 + \theta_2 \overline{x}_2 \in S$

Remember, there is no restriction of $\theta_1 + \theta_2 = 1$ which is there both in the definition of convex and as well as in affine. Therefore; any convex cone is also called a convex set but vice versa is not true.

(Refer Slide Time: 08:27)



To show that convex cone is also a convex set, set θ_1 and θ_2 as

$\theta_1 = \theta$ and $\theta_2 = 1-\theta$

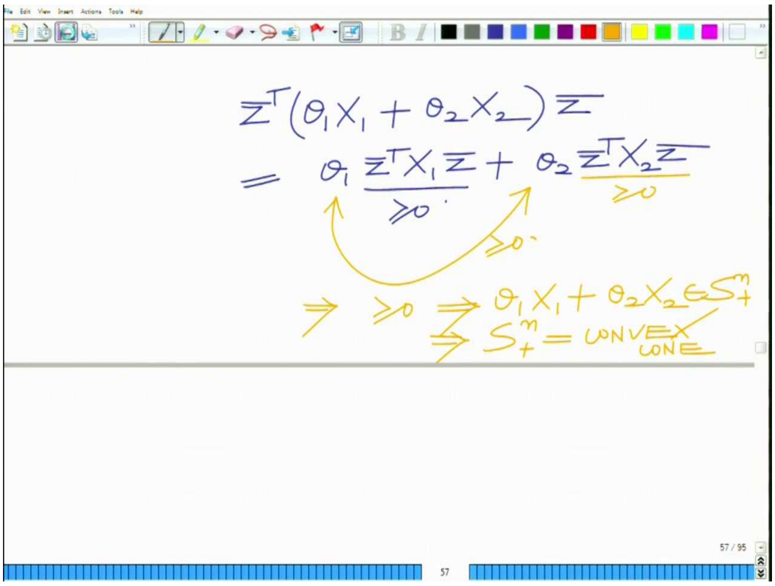
In this way,

$\theta \overline{x}_1 + (1-\theta)\,\overline{x}_2 \in S$

And this proves that convex cone is also a convex set. So, convex set is a subclass of convex cones.

So, the set of positive semi definite matrices is a convex cone it is known as a convex cone.

(Refer Slide Time: 09:52)



To further verify that $\theta_1 X_1 + \theta_2 X_2$ belongs to ;

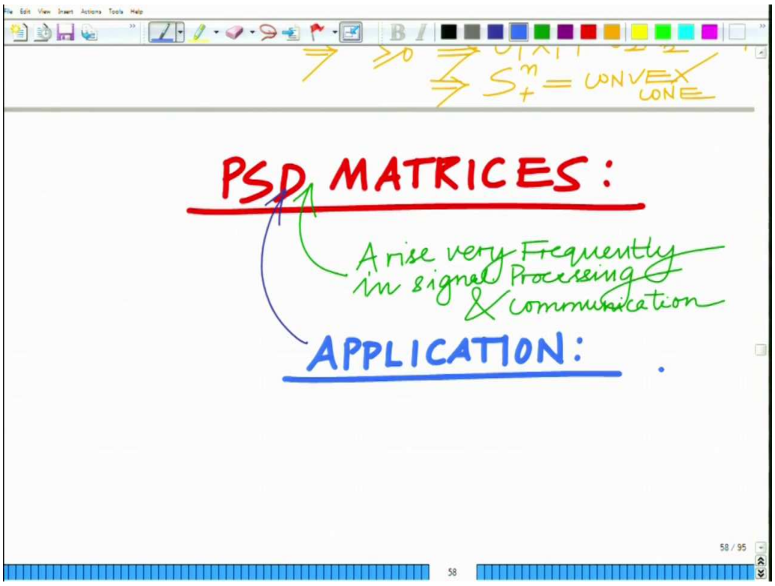
$$\overline{Z}^T\left(\theta_1 X_1 + \theta_2 X_2\right)\overline{Z}$$
$$= \overline{Z}^T \theta_1 X_1 \overline{Z} + \overline{Z}^T \theta_2 X_2 \overline{Z}$$
$$= \theta_1 \overline{Z}^T X_1 \overline{Z} + \theta_2 \overline{Z}^T X_2 \overline{Z}$$

Again θ_1 and θ_2 are positive scalars. Also it was considered that $X_1, X_2 \in S_+^n$. Therefore,

$$\bar{Z}^T(\theta_1 X_1 + \theta_2 X_2)\bar{Z}$$
$$= \theta_1 \bar{Z}^T X_1 \bar{Z} + \theta_2 \bar{Z}^T X_2 \bar{Z} \quad \geq \quad 0$$

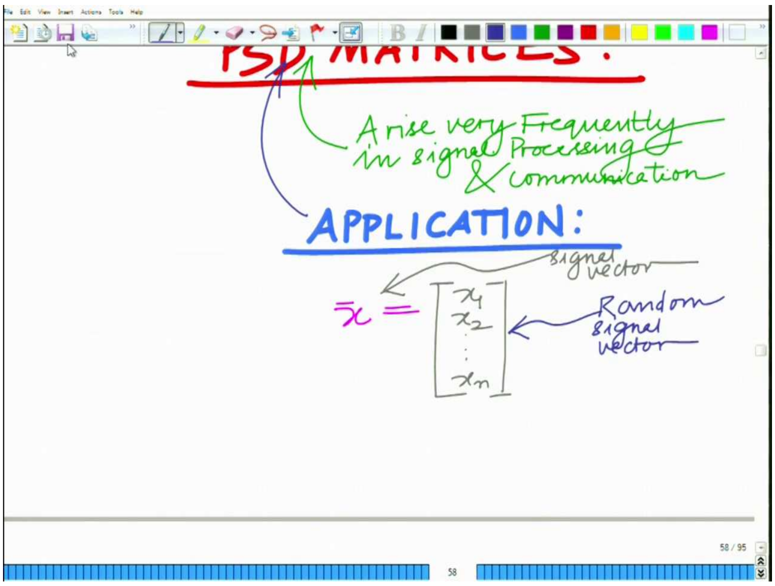
This verifies that $(\theta_1 X_1 + \theta_2 X_2) \in S_+^n$. And this means that S_+^n is a convex cone.

(Refer Slide Time: 11:36)



Now, Positive Semi Definite matrices are very important in signal processing and communication and these have a lot of applications in signal processing and communication. For instance a simple application can be demonstrated as follows.

(Refer Slide Time: 12:50)



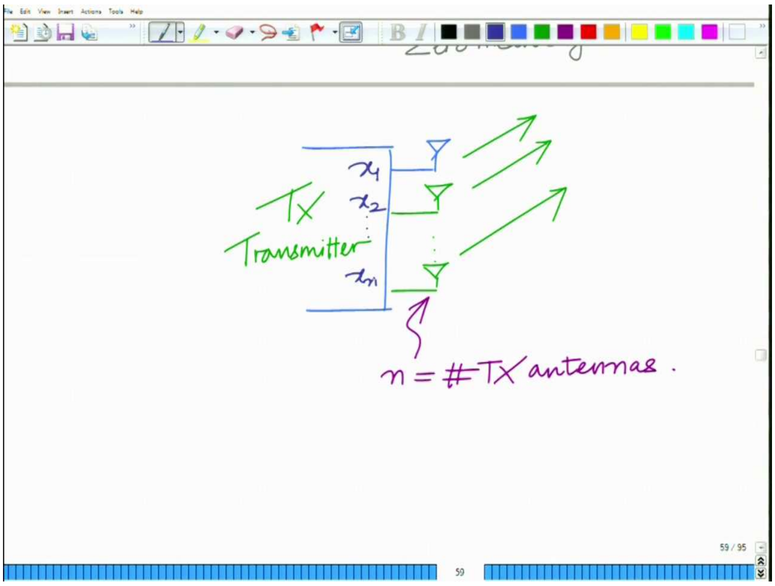
Let us consider a discrete signal vector given as follows.

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

This $\bar{x}$ is a random signal vector of size n assuming that the mean of this signal is zero.

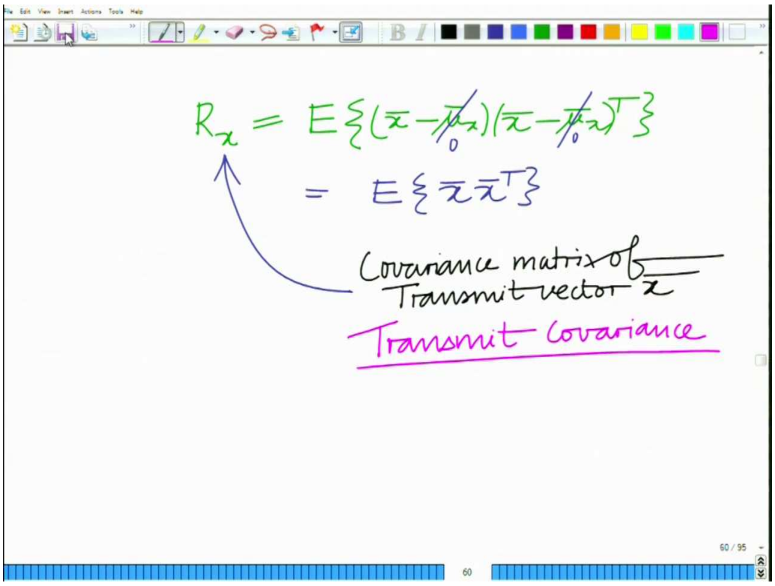
$$\bar{\mu}_x = E\{\bar{x}\} = 0$$

(Refer Slide Time: 13:59)



Again let us go back to multi antenna system. In this case consider a set of n transmit antennas transmitting n symbols $x_1, x_2, \ldots, x_n$.

(Refer Slide Time: 17:14)



So, the transmit covariance is given as

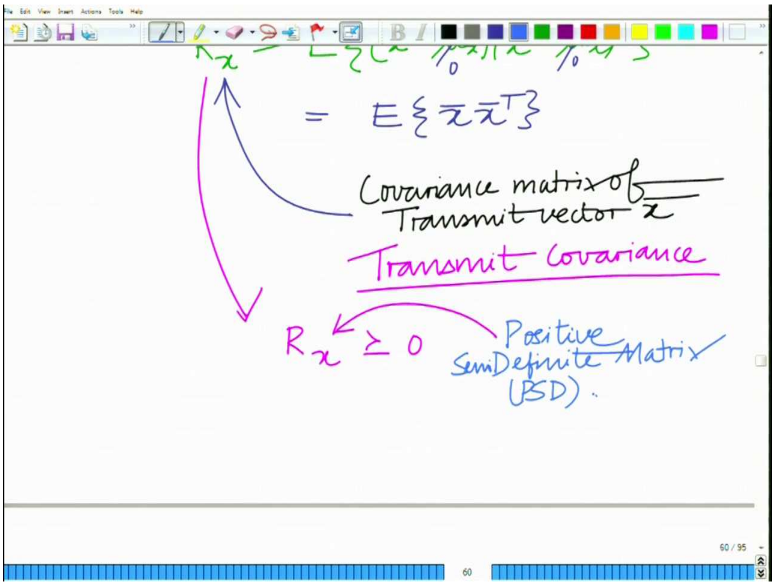
$$R_x = E\left\{(\bar{x}-\bar{\mu}_x)(\bar{x}-\bar{\mu}_x)^T\right\}$$

As this is a zero mean signal, therefore

$$R_x = E\left\{\overline{xx}^T\right\}$$

This is also termed as the covariance matrix of the transmit vector $\bar{x}$ or transmit covariance.

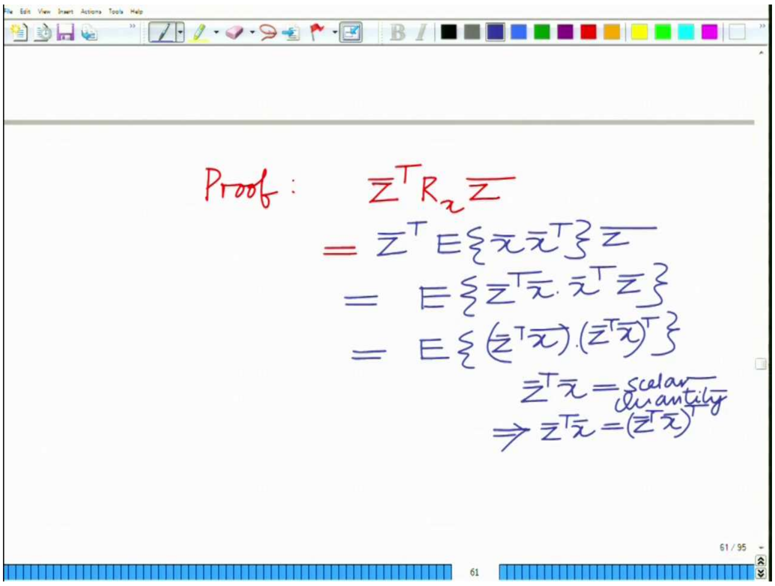
(Refer Slide Time: 18:10)



This transmit covariance matrix R_x is positive semi definite matrix. So

$$R_x \geq 0$$

(Refer Slide Time: 19:07)



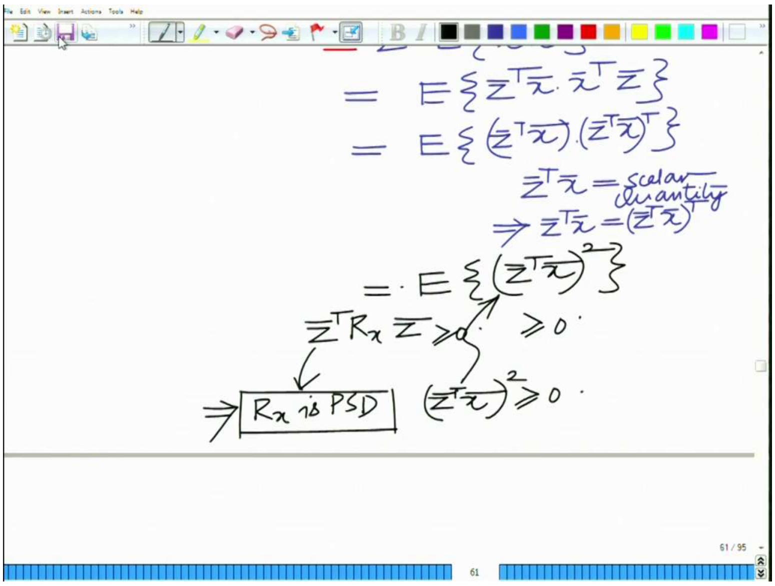
This is a very important property and this can be shown simply as follows.

$$\bar{Z}^T R_x \bar{Z}$$
$$= \bar{Z}^T E\left\{\bar{x}\bar{x}^T\right\}\bar{Z}$$
$$= E\left\{\bar{Z}^T \bar{x}.\bar{x}^T \bar{Z}\right\}$$
$$= E\left\{\left(\bar{Z}^T \bar{x}\right)\left(\bar{Z}^T \bar{x}\right)^T\right\}$$

As the transpose of a vector times another vector is a scalar quantity, so $\left(\bar{Z}^T \bar{x}\right)$ is a scalar quantity. Also; for a scalar quantity the transpose, the transpose of the quantity is the quatity itself.

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

(Refer Slide Time: 20:34).



Therefore,

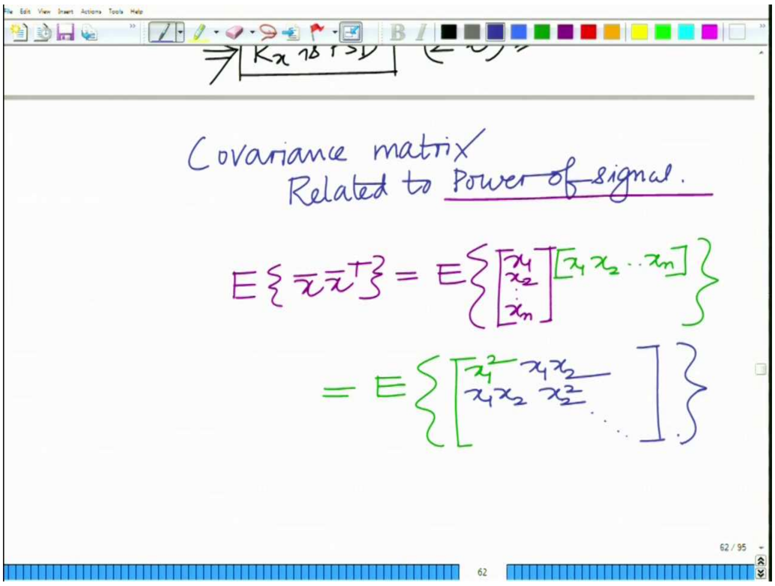
$$\bar{Z}^T R_x \bar{Z}$$
$$= E\left\{\left(\bar{Z}^T \bar{x}\right)^2\right\}$$

Again the expected value of a positive quantity is also positive so,

$$\bar{Z}^T R_x \bar{Z}$$
$$= E\left\{\left(\bar{Z}^T \bar{x}\right)^2\right\} \geq 0$$

This finally implies that the covariance matrix R_x is positive semi definite matrix. This shows that any covariance matrix of a random vector is a positive semi definite matrix.

(Refer Slide Time: 21:43)

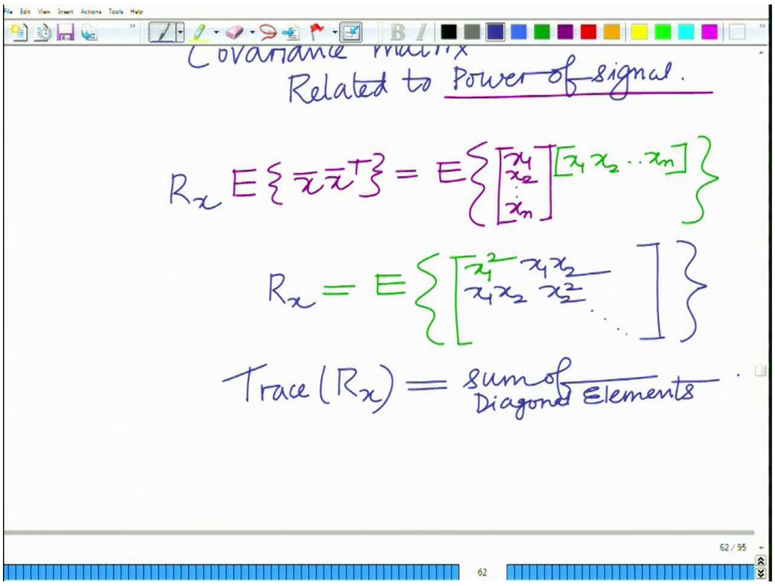


Covariance matrix has an important role to play. In fact, the covariance matrix is related to the transmit power of the signal. The covariance matrix can be expanded as follows.

$$E\left\{\overline{x}\overline{x}^T\right\} = E\left\{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}\right\}$$

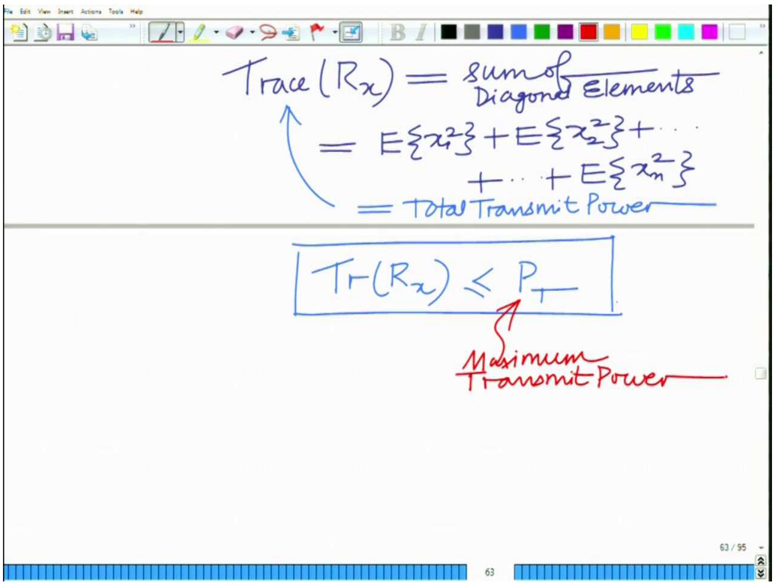
$$= E\left\{\begin{bmatrix} x_1^2 & x_1x_2 & \cdots \\ x_1x_2 & x_2^2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}\right\}$$

(Refer Slide Time: 23:03)



So, the trace of the covariance matrix is evidently equal to the sum of the diagonal elements which are square of each symbol.

(Refer Slide Time: 23:27)



So this trace is

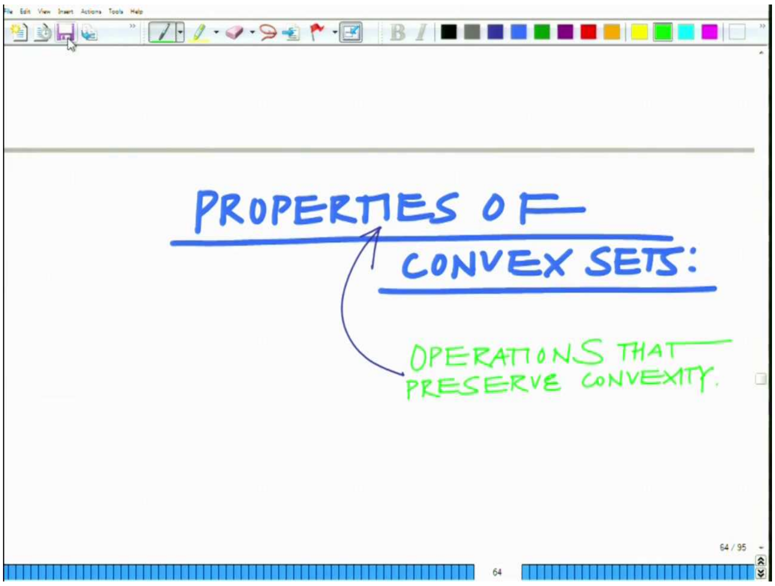
$$Tr(R_x) = E\{x_1^2\} + E\{x_2^2\} + \cdots + E\{x_n^2\}$$
$$= \text{Total Transmit Power}$$

So, the trace of the covariance matrix is equal to the total transmit power. Thus, the trace of the covariance matrix has to be less than or equal to the maximum transmit power P_T at the transmitter.

$$Tr\left(R_x\right) \leq P_T$$

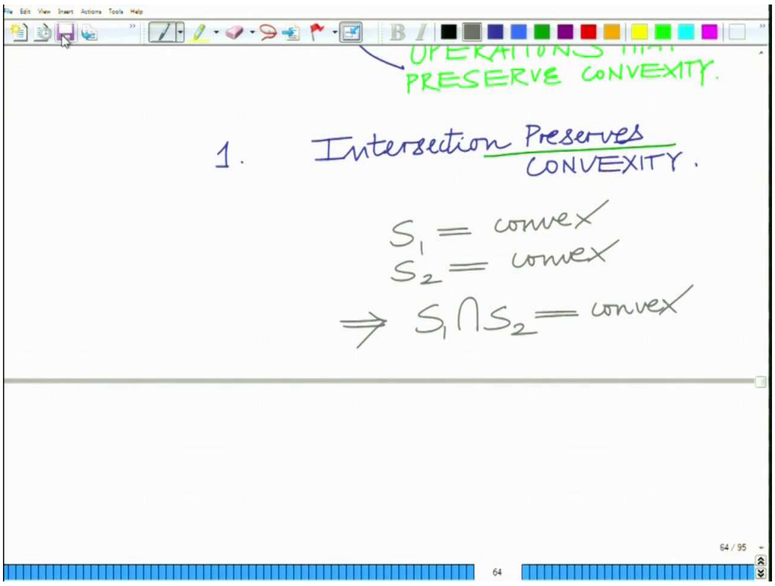
So, the covariance matrix plays a prominent role in wireless communications.

(Refer Slide Time: 25:50)



Let us now move on to another important concept which is the properties of convex sets. Now, the first property is that Intersection preserves Convexity.

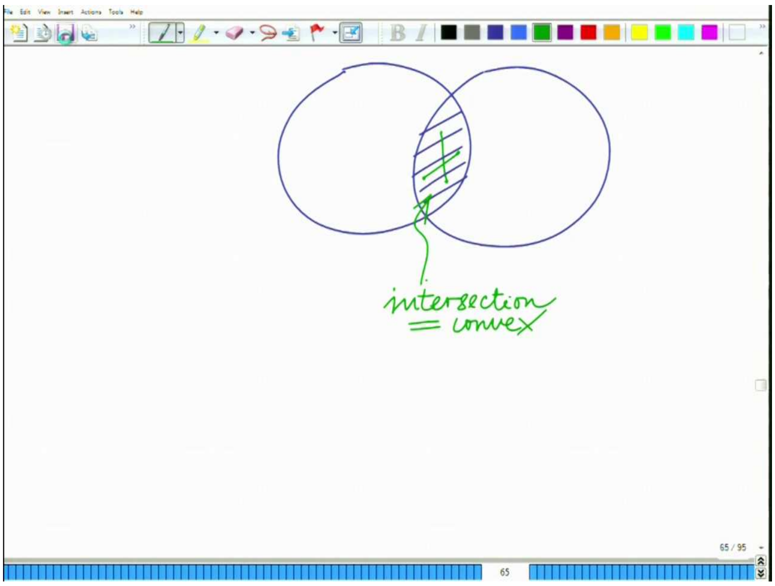
(Refer Slide Time: 26:41)



This property states that the intersection of two convex sets is also a convex set. To mathematically state this, let us consider two convex sets S_1 and S_2. So, according to this property, $S_1 \cap S_2$ is also a convex set.

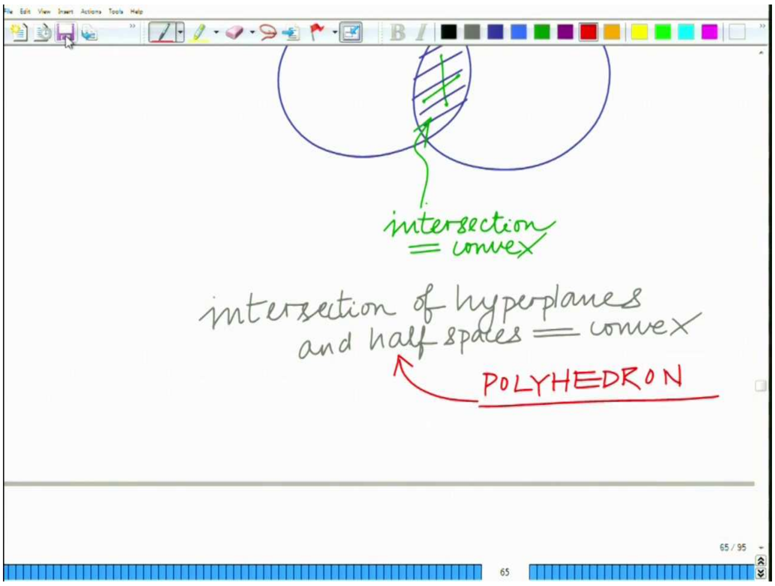
This property can be extended to any arbitrary number of convex sets, and the intersection of all these sets will also be convex. This is very simple to verify.

(Refer Slide Time: 27:56)



For instance take two circles such that they overlap a part to each other. It is known that a circle is a convex set. Now it can clearly be seen that if one take two points anywhere inside the intersection region of these circles, then the line segment drawn between these two points lie inside this intersection region only. And thus it is verified that the intersection of two circles is also convex.

(Refer Slide Time: 28:41)



A similar example to this is the intersection of hyperplanes and half spaces which is also convex because each hyperplane is convex, each half space is convex, and their intersection termed as the polyhedron is also convex.

Let us continue this discussion further in the subsequent modules.