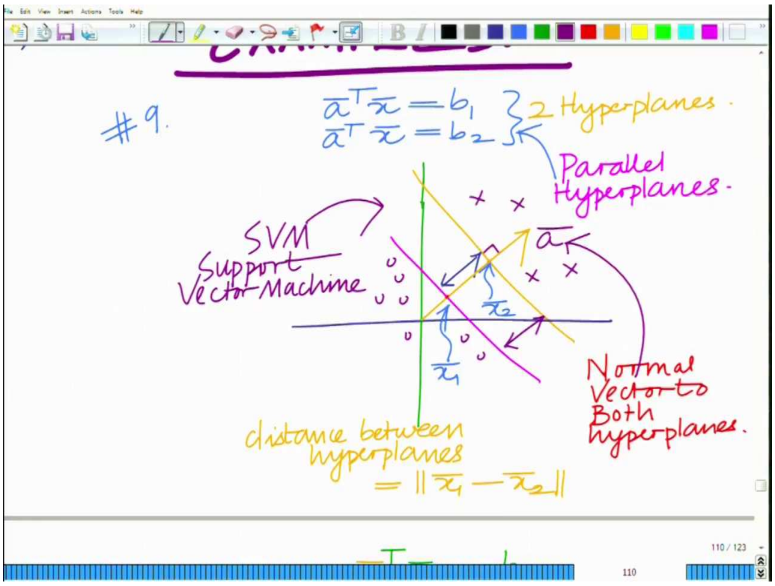
Applied Optimization for Wireless, Machine Learning, Big data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture-22
Problems on Convex Sets (contd.)

Hello welcome to another module in this massive open online course. So, we are looking at examples for convex sets and various properties of matrices. Let us discuss the properties of hyperplanes.

(Refer Slide Time: 00:25)



Consider two parallel hyperplanes given by

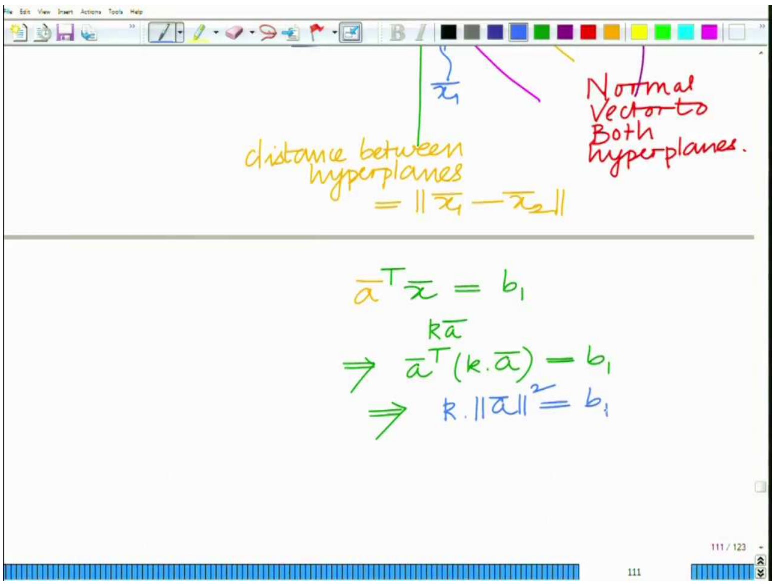
$\bar{a}^T\bar{x} = b_1$ Hyperplane 1

$\bar{a}^T\bar{x} = b_2$ Hyperplane 2

Both of these hyperplanes have the same normal vector. The distance between both the hyperplanes dist_{12} is equal to the distance between the intersection points of normal vector on both hyperplanes. If $\bar{x}_1$ is the intersection point of normal vector on first hyperplane and $\bar{x}_2$ is the intersection point of normal vector on second hyperplane, then dist_{12} can be calculated as follows.

$\text{dist}_{12} = \|\bar{x}_1 - \bar{x}_2\|$

(Refer Slide Time: 03:12)



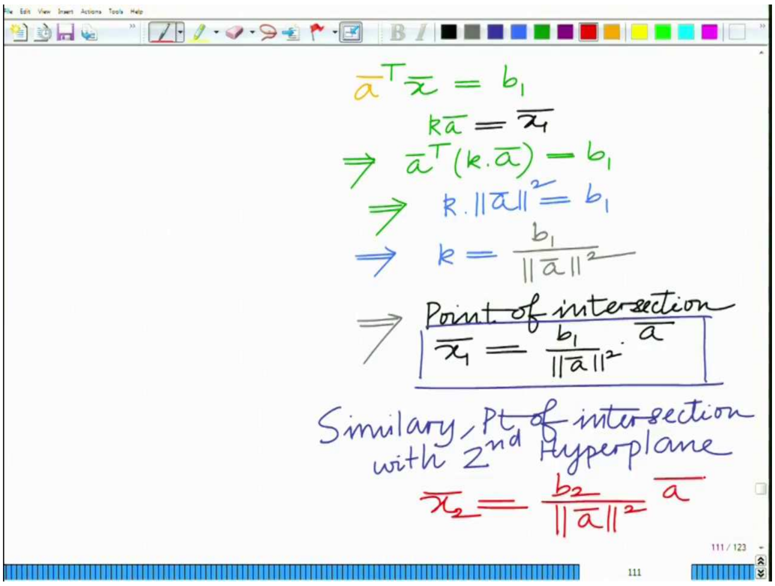
So, if $\bar{x}$ is some constant k times the vector $\bar{a}$; i.e.

$\bar{x} = k\bar{a}$

Then the equation of first hyperplane becomes

$$\bar{a}^T(k\bar{a}) = b_1$$
$$k\|\bar{a}\|^2 = b_1$$

Refer Slide Time: 04:21)



And from here the constant k is equal to

$$k = \frac{b_1}{\|\bar{a}\|^2}$$

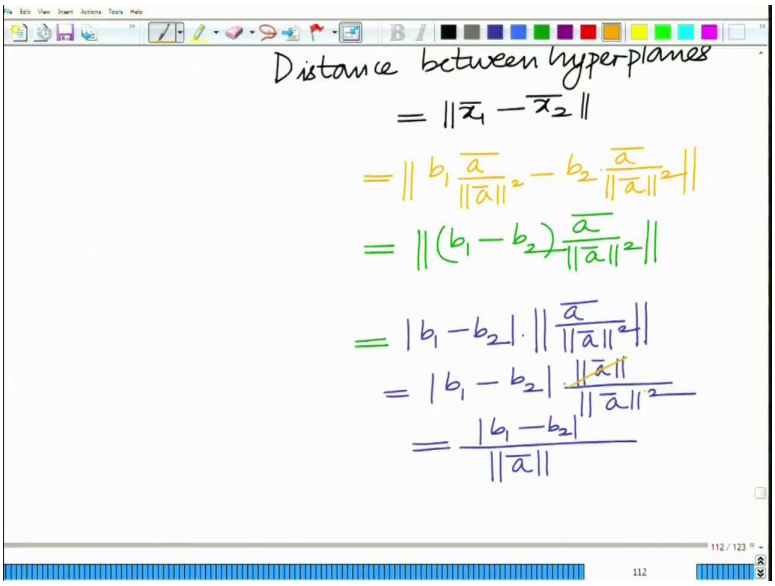
So, the point of intersection of normal vector to the first hyperplane is

$$\bar{x}_1 = k\bar{a} = \frac{b_1}{\|\bar{a}\|^2} \cdot \bar{a}$$

And similarly the point of intersection of normal vector to the second hyperplane is

$$\bar{x}_2 = k\bar{a} = \frac{b_2}{\|\bar{a}\|^2} \cdot \bar{a}$$

(Refer Slide Time: 06:08)

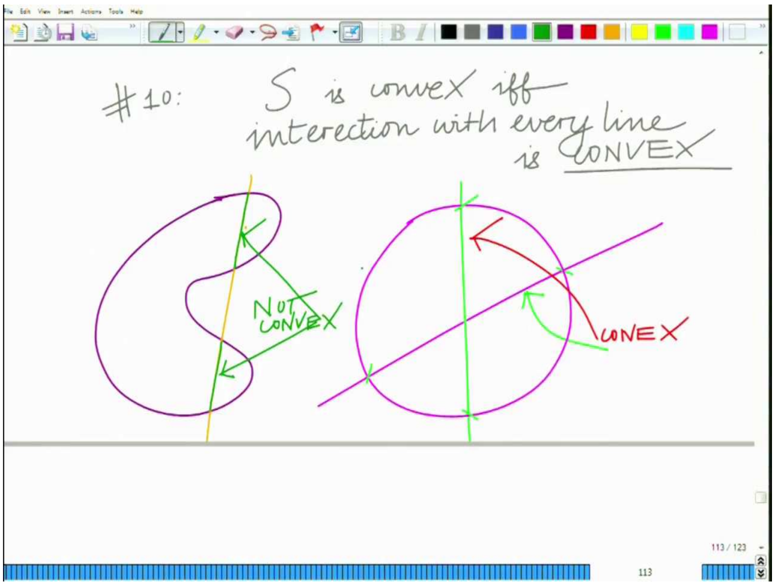


Thus; the distance between both the hyperplanes dist_{12} is

$$
\begin{aligned}
\text{dist}_{12} &= \|\bar{x}_1 - \bar{x}_2\| \\
&= \left\| \frac{b_1}{\|\bar{a}\|^2} \cdot \bar{a} - \frac{b_2}{\|\bar{a}\|^2} \cdot \bar{a} \right\| \\
&= \left\| (b_1 - b_2) \frac{\bar{a}}{\|\bar{a}\|^2} \right\| \\
&= |b_1 - b_2| \cdot \frac{\|\bar{a}\|}{\|\bar{a}\|^2} \\
&= \frac{|b_1 - b_2|}{\|\bar{a}\|}
\end{aligned}
$$

This is the distance between the set of parallel hyperplanes which have the same normal vector $\bar{a}$. This property has a lot of applications such as the support vector machine classifier (SVM). So, this forms the basis for the SVM which is to maximize the distance between hyperplanes by minimizing $\|\bar{a}\|$ and this makes the classifier more effective by effectively separating the two different classes of objects.

(Refer Slide Time: 11:00)

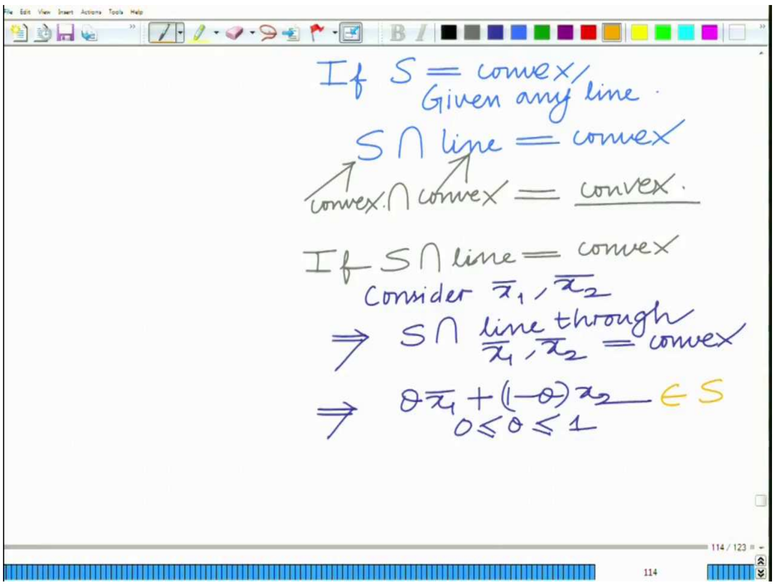


Let us look at another interesting problem that the set *S* is convex if and only if its intersection with every line is convex.

So, if there is a circle then all the lines intersecting a circle are convex and this means that circle is a convex set which is true. So, the intersection with any line is a convex set.

On the other hand let us take a non convex region as shown in the above figure. Now let us take any line as shown in figure, then the intersection of this line with this non convex region is has two disjointed line segments which is not convex.

(Refer Slide Time: 12:27)

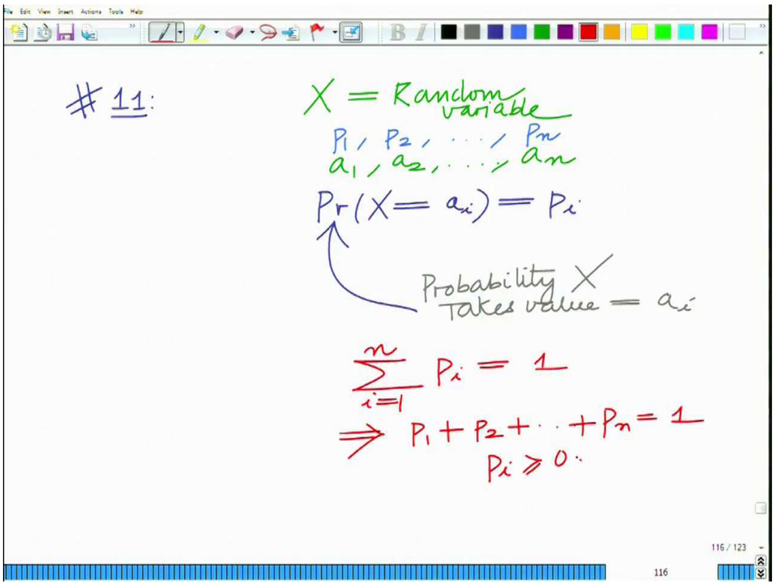


So, it is easy to verify this property. Let us have a convex set *S*. Now consider any two points $\overline{x}_1$ and $\overline{x}_2$ such that the intersection of line passing both of these points with set *S* is convex. So, $\overline{x}_1$ and $\overline{x}_2$ belongs to the line that lies within *S* and $\overline{x}_1, \overline{x}_2 \in S$; Therefore, $\overline{x}_1$ and $\overline{x}_2$ also belongs to the intersection of *S* with the line and this implies that for $0 \le \theta \le 1$ a convex combination of $\overline{x}_1$ and $\overline{x}_2$;

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

Hence a convex combination belongs to set *S*. This simply infers that *S* is a convex set. This verifies the above property.

(Refer Slide Time: 17:04)

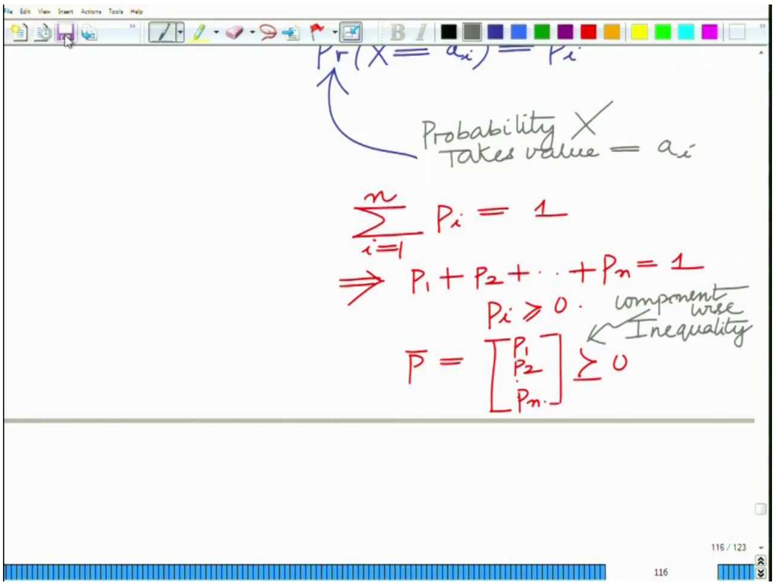


Let us look at another interesting problem that pertains to probabilities. Let X be a random variable which takes *n* values as $a_1, a_2, \ldots, a_n$. The probability that X takes the value a_i is equal to P_i.

$$\Pr(X = a_i) = P_i$$

Therefore there are *n* such probabilities $P_1, P_2, \ldots, P_n$. Also that each P_i is greater than or equal to zero i.e. $P_i \ge 0$.

(Refer Slide Time: 18:56)



Naturally the sum of the probabilities must be one.

$$\sum_{i=1}^{n} P_i = 1$$

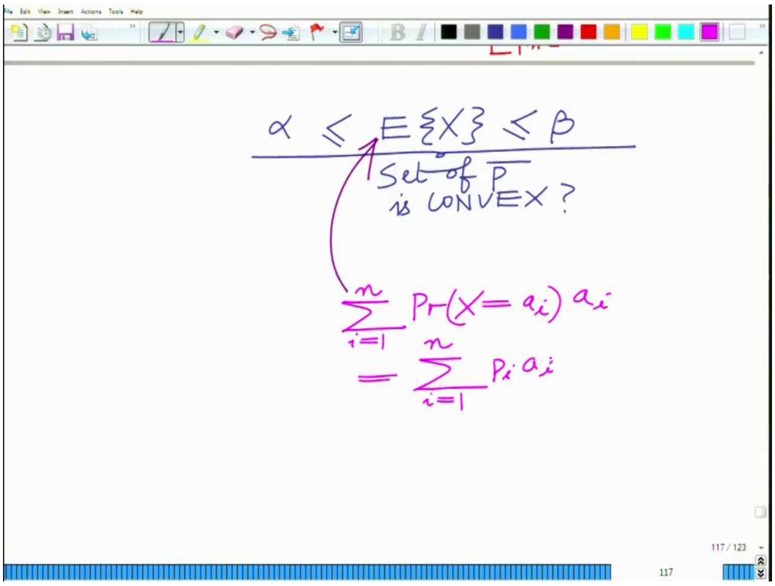
$$P_1 + P_2 + \ldots + P_n = 1$$

So, the probability vector $\bar{P}$ is given as

$$\bar{P} = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix} \geq 0$$

This is the component wise inequality. Each component of vector $\bar{P}$ must be greater than equal to 0.

(Refer Slide Time: 19:45)

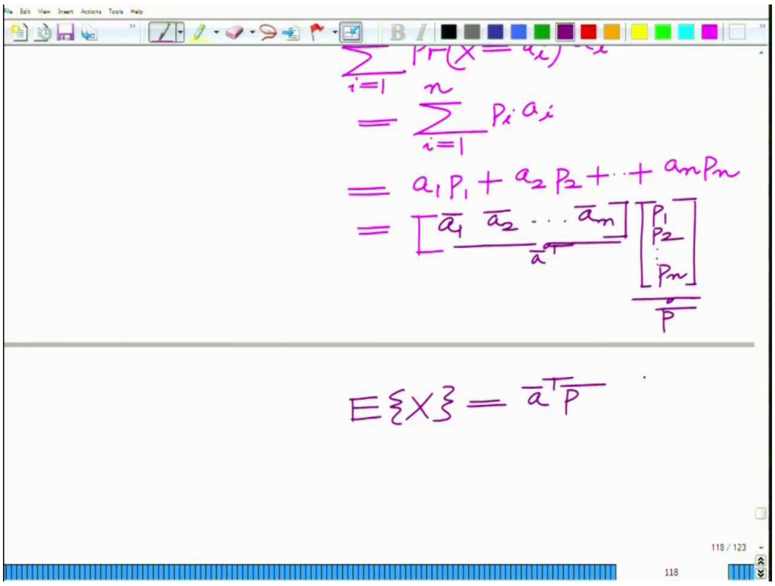


Let us look at the first property of probabilities that if X is a random variable such that for two constants α and β;

$$\alpha \leq E\{x\} \leq \beta$$

Then set of all $\bar{P}$ is a convex set.

(Refer Slide Time: 20:48)



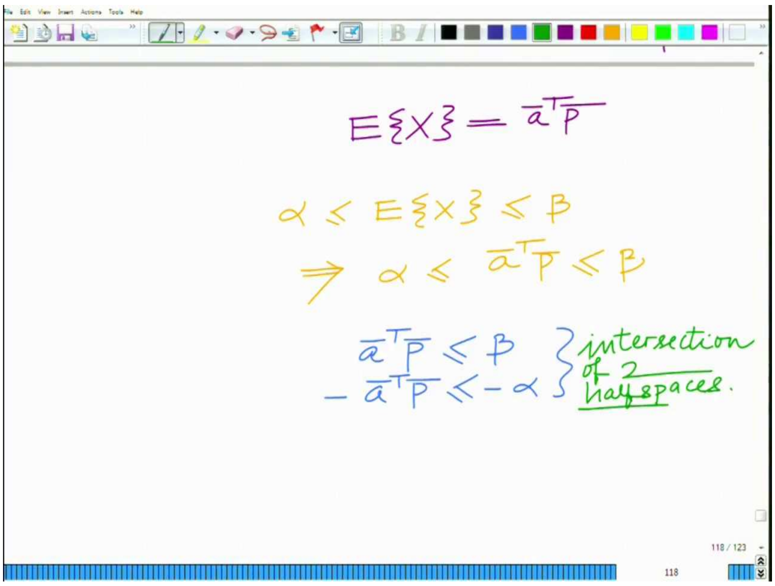
To verify this, look at the expected value of a random variable X.

$$
\begin{aligned}
E\{X\} &= \sum_{i=1}^{n} \Pr(X = a_i)\, a_i \\
&= \sum_{i=1}^{n} P_i a_i \\
&= a_1 P_1 + a_2 P_2 + \ldots + a_n P_n \\
&= \underbrace{\begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \cdots & \bar{a}_n \end{bmatrix}}_{\bar{a}^T} \underbrace{\begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix}}_{\bar{P}}
\end{aligned}
$$

So, expected value of the random variable X becomes

$$E\{X\} = \bar{a}^T \bar{P}$$

(Refer Slide Time: 21:36)



So if

$$\alpha \le E\{x\} \le \beta$$

Then

$$\alpha \le \bar{a}^T \bar{P} \le \beta$$

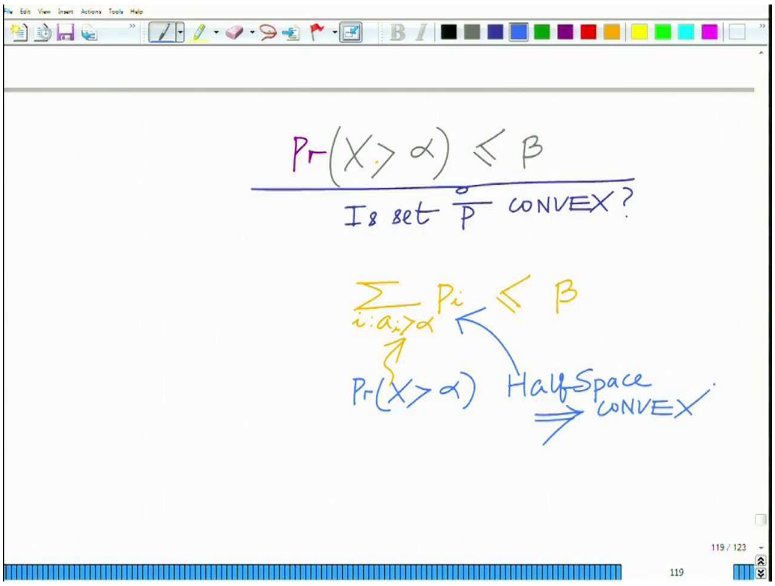
Therefore, in terms of half space, this can be splitted as

$\bar{a}^T \bar{P} \le \beta$ Halfspace 1

$-\bar{a}^T \bar{P} \le -\alpha$ Halfspace 2

Hence set of all $\bar{P}$ is the intersection of two halfspaces and this verifies that the set of all $\bar{P}$ is a convex set.

(Refer Slide Time: 23:23)



Another property of probability set is that if for two constants α and β;

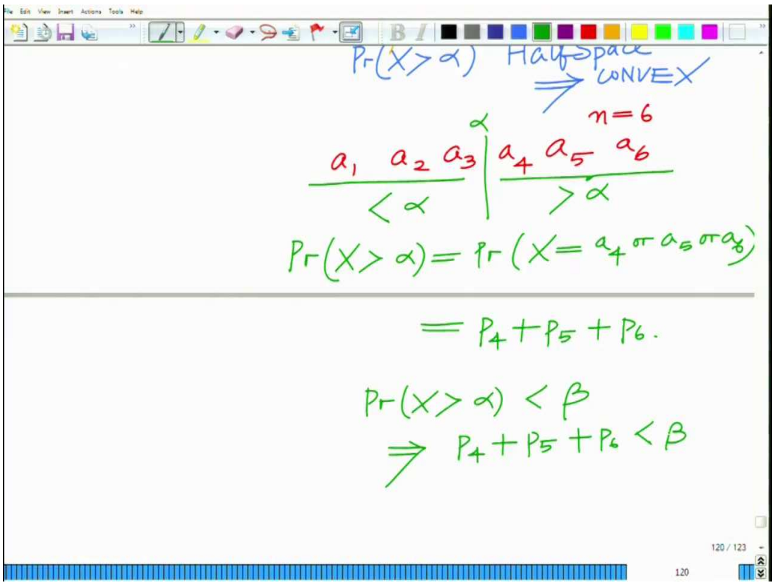
$\Pr(X > \alpha) \le \beta$

Then the set of probability vector $\bar{P}$ of random variable X is a convex set.

So $\Pr(X > \alpha)$ is the summation of all probabilities P_i such that the corresponding $a_i > \alpha$.

As this is a linear sum this means that it is the half space. And this implies that this set is convex.

(Refer Slide Time: 24:54)



For example, let us take an example. Consider X has 6 values as $a_1, a_2, a_3, a_4, a_5, a_6$. Let α lies between a_3 and a_4. Therefore

$$a_1, a_2, a_3 < \alpha$$

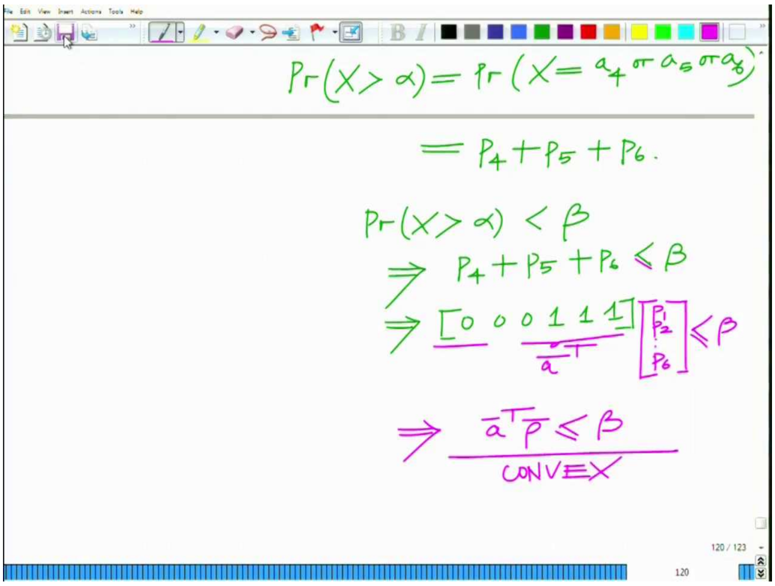
And

$$a_4, a_5, a_6 > \alpha$$

The probability that X takes the value greater than α is

$$\begin{aligned}\Pr(X > \alpha) &= \Pr(X = a_4 \text{ or } a_5 \text{ or } a_6)\\ &= P_4 + P_5 + P_6\end{aligned}$$

(Refer Slide Time: 26:02)



The vector $\bar{a}$ for this situation is

$$\begin{aligned}\bar{a} &= \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}\end{aligned}$$

Thus,

$$\Pr(X > \alpha) < \beta$$

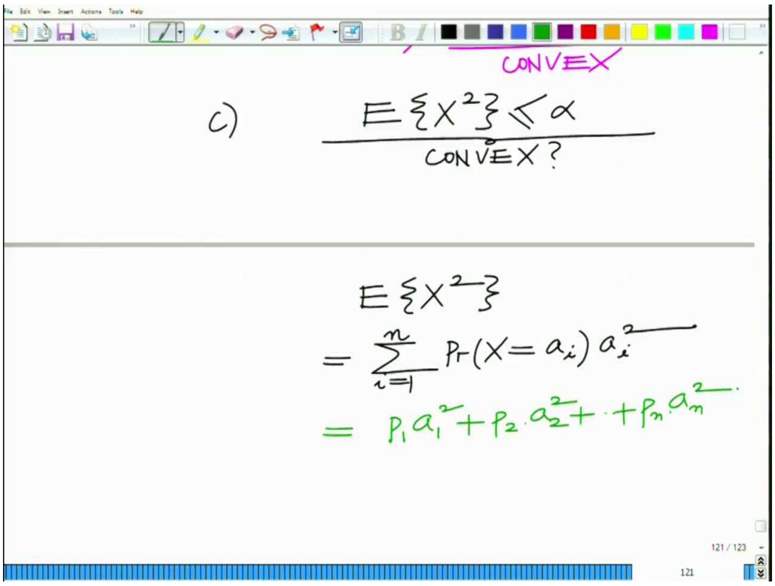
$$\Rightarrow P_4 + P_5 + P_6 < \beta$$

$$\Rightarrow \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}}_{\bar{a}^T} \underbrace{\begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ P_6 \end{bmatrix}}_{\bar{P}} \le \beta$$

$$\Rightarrow \bar{a}^T \bar{P} \le \beta$$

This is a half space which further shows that this is a convex set.

(Refer Slide Time: 27:01)



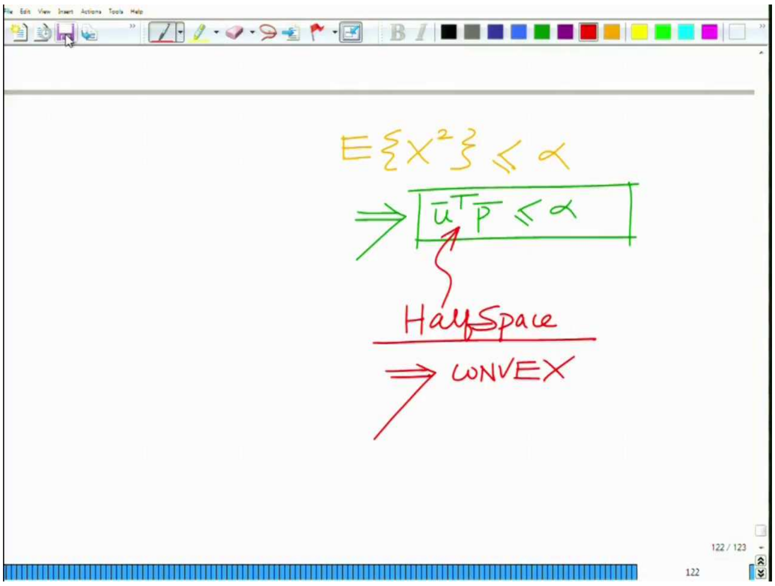
Secondly, if the second moment of random variable X that is the expected value of square of X is less than or equal to α

$$E\left\{X^2\right\} \leq \alpha$$

Then it is convex. To show this let us look at the set of all vectors $\bar{P}$ for such condition is.

$$\begin{aligned}
E\left\{X^2\right\} &= \sum_{i=1}^{n} \Pr\left(X = a_i\right) a_i^2 \\
&= P_1 a_1^2 + P_2 a_2^2 + \cdots + P_n a_n^2 \\
&= \underbrace{\begin{bmatrix} a_1^2 & a_2^2 & \cdots & a_n^2 \end{bmatrix}}_{\bar{u}^T} \underbrace{\begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix}}_{\bar{P}} \\
&= \bar{u}^T \bar{P}
\end{aligned}$$

(Refer Slide Time: 29:11)



Thus the second moment of random variable X is

$$E\left\{X^2\right\} = \bar{u}^T \bar{P}$$

This means

$$\bar{u}^T \bar{P} \leq \alpha$$

This is a half space which further shows that this set is a convex set.

So, these are some interesting applications of convexity which have huge application in optimization theory in the context of wireless communication or signal processing.