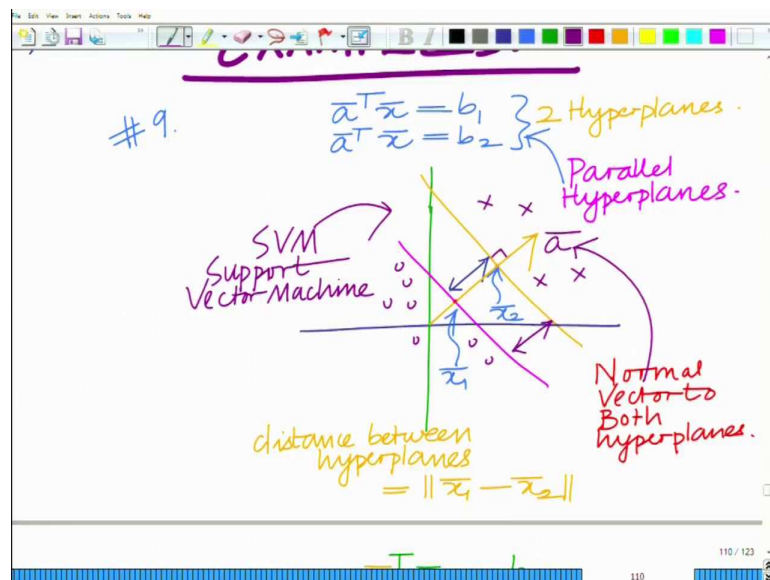


**Applied Optimization for Wireless, Machine Learning, Big data**  
**Prof. Aditya K. Jagannatham**  
**Department of Electrical Engineering**  
**Indian Institute of Technology, Kanpur**

**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

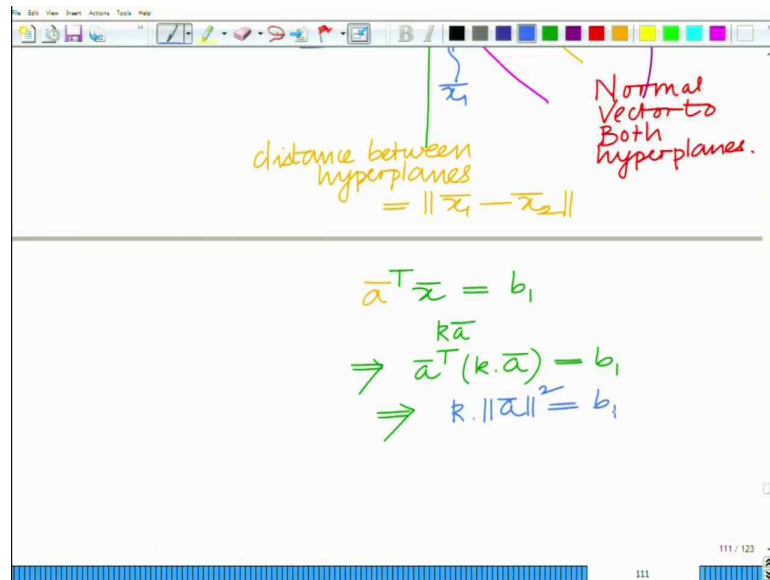
$$\vec{a}^T \vec{x} = b_1 \quad \text{Hyperplane 1}$$

$$\vec{a}^T \vec{x} = b_2 \quad \text{Hyperplane 2}$$

Both of these hyperplanes have the same normal vector. The distance between both the hyperplanes  $\text{dist}_{12}$  is equal to the distance between the intersection points of normal vector on both hyperplanes. If  $\vec{x}_1$  is the intersection point of normal vector on first hyperplane and  $\vec{x}_2$  is the intersection point of normal vector on second hyperplane, then  $\text{dist}_{12}$  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)

The image shows a digital whiteboard with a toolbar at the top. The handwritten text is as follows:

$$\begin{aligned}\bar{a}^T \bar{x} &= b_1 \\ k\bar{a} &= \bar{x}_1 \\ \Rightarrow \bar{a}^T (k\bar{a}) &= b_1 \\ \Rightarrow k \cdot \|\bar{a}\|^2 &= b_1 \\ \Rightarrow k &= \frac{b_1}{\|\bar{a}\|^2} \\ \Rightarrow \text{Point of intersection} \\ \bar{x}_1 &= \frac{b_1}{\|\bar{a}\|^2} \bar{a}\end{aligned}$$

Below this, it says "Similar, Pt of intersection with 2<sup>nd</sup> Hyperplane" and then:

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

The bottom right corner of the whiteboard shows "111 / 123".

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)

A screenshot of a presentation slide showing a handwritten derivation of the distance between two parallel hyperplanes. The text is written in various colors (black, orange, green, blue) on a white background. The derivation starts with the title 'Distance between hyperplanes' and proceeds through several steps:  $= \|\bar{x}_1 - \bar{x}_2\|$ ,  $= \left\| b_1 \frac{\bar{a}}{\|\bar{a}\|^2} - b_2 \frac{\bar{a}}{\|\bar{a}\|^2} \right\|$ ,  $= \left\| (b_1 - b_2) \frac{\bar{a}}{\|\bar{a}\|^2} \right\|$ ,  $= |b_1 - b_2| \cdot \left\| \frac{\bar{a}}{\|\bar{a}\|^2} \right\|$ ,  $= |b_1 - b_2| \cdot \frac{\|\bar{a}\|}{\|\bar{a}\|^2}$ , and finally  $= \frac{|b_1 - b_2|}{\|\bar{a}\|}$ . The slide has a standard presentation interface with a toolbar at the top and a status bar at the bottom showing '112 / 123'.

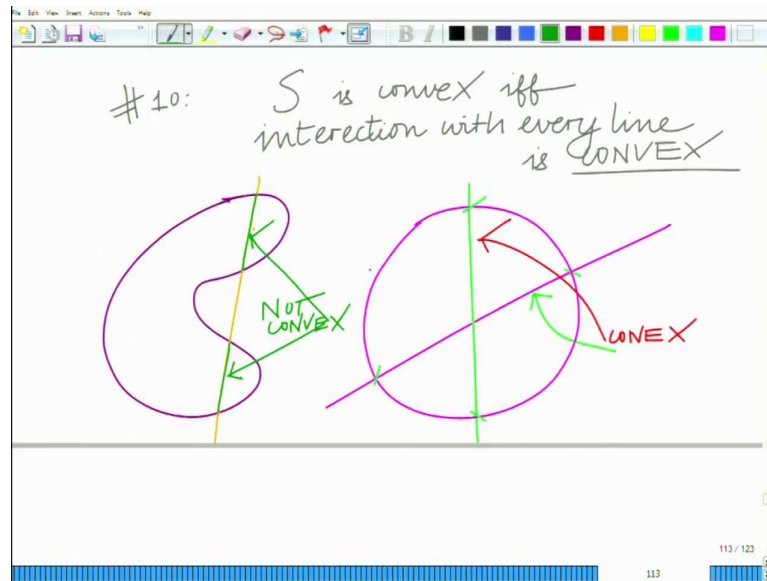
$$\begin{aligned} &\text{Distance between hyperplanes} \\ &= \|\bar{x}_1 - \bar{x}_2\| \\ &= \left\| b_1 \frac{\bar{a}}{\|\bar{a}\|^2} - b_2 \frac{\bar{a}}{\|\bar{a}\|^2} \right\| \\ &= \left\| (b_1 - b_2) \frac{\bar{a}}{\|\bar{a}\|^2} \right\| \\ &= |b_1 - b_2| \cdot \left\| \frac{\bar{a}}{\|\bar{a}\|^2}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

Thus; the distance between both the hyperplanes  $\text{dist}_{12}$  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)

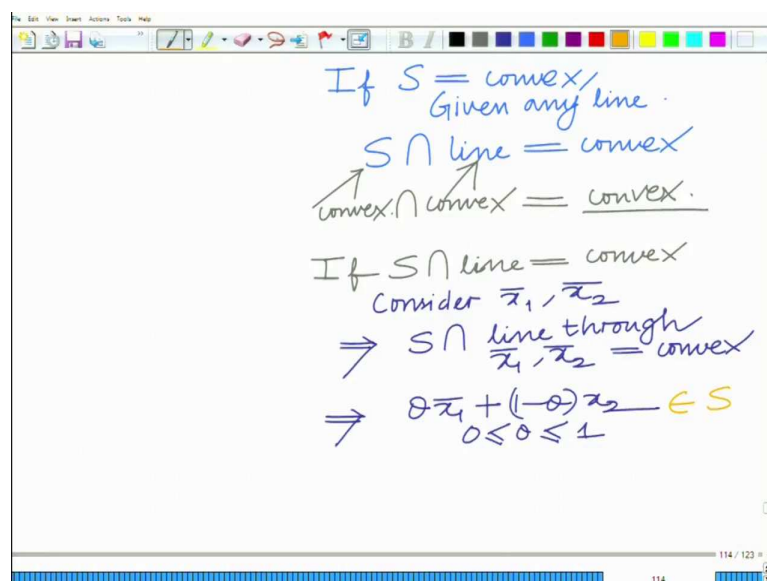


Let us look at another interesting problem that the set  $S$  is convex if and only if its intersection with every line is convex.

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  $\bar{x}_1$  and  $\bar{x}_2$  such that the intersection of line passing both of these points with set  $S$  is convex. So,  $\bar{x}_1$  and  $\bar{x}_2$  belongs to the line that lies within  $S$  and  $\bar{x}_1, \bar{x}_2 \in S$ ; Therefore,  $\bar{x}_1$  and  $\bar{x}_2$  also belongs to the intersection of  $S$  with the line and this implies that for  $0 \leq \theta \leq 1$  a convex combination of  $\bar{x}_1$  and  $\bar{x}_2$ ;

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

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)

# 11:

$X = \text{Random variable}$

$P_1, P_2, \dots, P_n$   
 $a_1, a_2, \dots, a_n$

$\Pr(X = a_i) = P_i$

Probability  $X$  Takes value  $= a_i$

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

$\Rightarrow P_1 + P_2 + \dots + P_n = 1$   
 $P_i \geq 0$

Let us look at another interesting problem that pertains to probabilities. Let  $X$  be a random variable which takes  $n$  values as  $a_1, a_2, \dots, a_n$ . 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  $P_1, P_2, \dots, P_n$ . Also that each  $P_i$  is greater than or equal to zero i.e.  $P_i \geq 0$ .

(Refer Slide Time: 18:56)

The image shows a digital whiteboard with handwritten notes in red and blue ink. At the top, the formula  $\Pr(X = a_i) = p_i$  is written in blue. A blue arrow points from this formula to the text "Probability X Takes value =  $a_i$ ". Below this, the sum of probabilities is given as  $\sum_{i=1}^n p_i = 1$  in red. This is followed by the implication  $\Rightarrow p_1 + p_2 + \dots + p_n = 1$  in red. Then, the probability vector is defined as  $\bar{P} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}$  in red. To the right of the vector, the condition  $p_i \geq 0$  is written in red, with a blue arrow pointing to it from the text "component wise Inequality". Below  $p_i \geq 0$ , the vector is shown with a greater-than-or-equal-to zero symbol:  $\bar{P} \geq 0$ . The whiteboard interface includes a toolbar at the top and a status bar at the bottom showing "116 / 123".

Naturally the sum of the probabilities must be one.

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

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

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)

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


---

 Set of  $\bar{P}$  is CONVEX?
 
$$\sum_{i=1}^n \Pr(X=a_i) a_i$$

$$= \sum_{i=1}^n p_i a_i$$

Let us look at the first property of probabilities that if  $X$  is a random variable such that for two constants  $\alpha$  and  $\beta$ ;

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

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

(Refer Slide Time: 20:48)

$$\sum_{i=1}^n \Pr(X=a_i) a_i$$

$$= \sum_{i=1}^n p_i a_i$$

$$= a_1 p_1 + a_2 p_2 + \dots + a_n p_n$$

$$= \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}$$

$$\quad \quad \quad \bar{a}^T \quad \quad \bar{P}$$


---


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

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

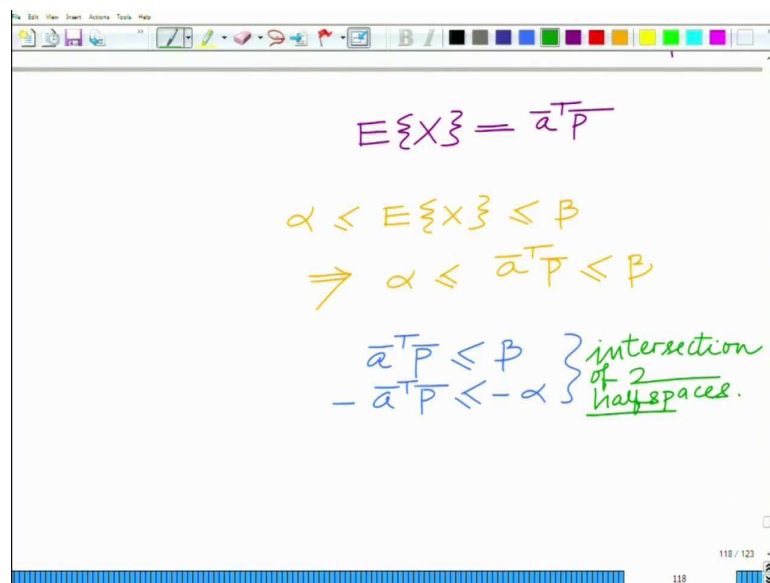


$$\begin{aligned}
 E\{X\} &= \sum_{i=1}^n \Pr(X = a_i) a_i \\
 &= \sum_{i=1}^n P_i a_i \\
 &= a_1 P_1 + a_2 P_2 + \dots + a_n P_n \\
 &= \underbrace{[\bar{a}_1 \quad \bar{a}_2 \quad \dots \quad \bar{a}_n]}_{\bar{a}^T} \underbrace{\begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix}}_{\bar{P}}
 \end{aligned}$$

So, expected value of the random variable X becomes

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

(Refer Slide Time: 21:36)



$$\begin{aligned}
 E\{X\} &= \bar{a}^T \bar{P} \\
 \alpha &\leq E\{x\} \leq \beta \\
 \Rightarrow \alpha &\leq \bar{a}^T \bar{P} \leq \beta \\
 \left. \begin{aligned} \bar{a}^T \bar{P} &\leq \beta \\ -\bar{a}^T \bar{P} &\leq -\alpha \end{aligned} \right\} &\text{intersection of 2 halfspaces.}
 \end{aligned}$$

So if

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

Then

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

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

$$\begin{aligned}\bar{a}^T \bar{P} &\leq \beta && \text{Halfspace 1} \\ -\bar{a}^T \bar{P} &\leq -\alpha && \text{Halfspace 2}\end{aligned}$$

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)

The image shows a whiteboard with handwritten mathematical notes. At the top, it says  $\Pr(X > \alpha) \leq \beta$ . Below this, it asks "Is set  $\bar{P}$  CONVEX?". Then, it shows the summation  $\sum_{i: a_i > \alpha} P_i \leq \beta$ . A blue arrow points from this summation to the expression  $\Pr(X > \alpha)$ , which is also written below. A blue arrow then points from  $\Pr(X > \alpha)$  to the text "HalfSpace CONVEX".

Another property of probability set is that if for two constants  $\alpha$  and  $\beta$ ;

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

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)

Handwritten notes on a whiteboard:

$\Pr(X > \alpha)$  Half-space CONVEX  $\Rightarrow n=6$

Diagram showing values  $a_1, a_2, a_3$  on the left and  $a_4, a_5, a_6$  on the right, separated by a vertical line at  $\alpha$ . Below the left side is  $< \alpha$  and below the right side is  $> \alpha$ .

$\Pr(X > \alpha) = \Pr(X = a_4 \text{ or } a_5 \text{ or } a_6)$

$= P_4 + P_5 + P_6.$

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

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

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  $\alpha$  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  $\alpha$  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)

$$\begin{aligned}
 \Pr(X > \alpha) &= \Pr(X = a_4 + \sigma a_5 + \sigma a_6) \\
 &= P_4 + P_5 + P_6. \\
 \Pr(X > \alpha) &< \beta \\
 \Rightarrow P_4 + P_5 + P_6 &\leq \beta \\
 \Rightarrow \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ P_6 \end{bmatrix} &\leq \beta \\
 \Rightarrow \frac{\bar{a}^T \bar{p} \leq \beta}{\text{CONVEX}}
 \end{aligned}$$

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

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

Thus,

$$\begin{aligned}
 \Pr(X > \alpha) &< \beta \\
 \Rightarrow P_4 + P_5 + P_6 &< \beta \\
 \Rightarrow \underbrace{[0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 1]}_{\bar{a}^T} \underbrace{\begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ P_6 \end{bmatrix}}_{\bar{p}} &\leq \beta \\
 \Rightarrow \bar{a}^T \bar{p} &\leq \beta
 \end{aligned}$$

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

(Refer Slide Time: 27:01)

c)  $E\{X^2\} \leq \alpha$   
 $\text{CONVEX?}$

---


$$E\{X^2\} = \sum_{i=1}^n \Pr(X=a_i) a_i^2$$

$$= p_1 a_1^2 + p_2 a_2^2 + \dots + p_n a_n^2$$

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

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

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

$$E\{X^2\} = \sum_{i=1}^n \Pr(X=a_i) a_i^2$$

$$= P_1 a_1^2 + P_2 a_2^2 + \dots + P_n a_n^2$$

$$= \underbrace{\begin{bmatrix} a_1^2 & a_2^2 & \dots & a_n^2 \end{bmatrix}}_{\bar{u}^T} \underbrace{\begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix}}_{\bar{P}}$$

$$= \bar{u}^T \bar{P}$$

(Refer Slide Time: 29:11)

The image shows a whiteboard with handwritten mathematical derivations. At the top, the equation  $E\{X^2\} \leq \alpha$  is written in yellow. Below it, an arrow points to the boxed equation  $\bar{u}^T \bar{P} \leq \alpha$  in green. A red arrow points from the boxed equation to the text "HalfSpace" written in red. Below "HalfSpace", another red arrow points to the word "CONVEX" written in red. The whiteboard interface includes a toolbar at the top and a status bar at the bottom right showing "122 / 123" and "122".

Thus the second moment of random variable X is

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

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.