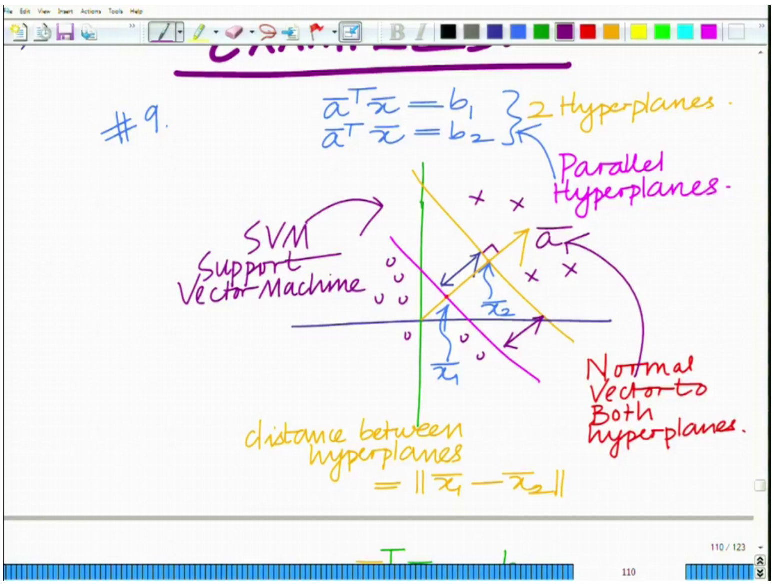
Applied Optimization for Wireless, Machine Learning, Big data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

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 continue our discussion.

(Refer Slide Time: 00:25)

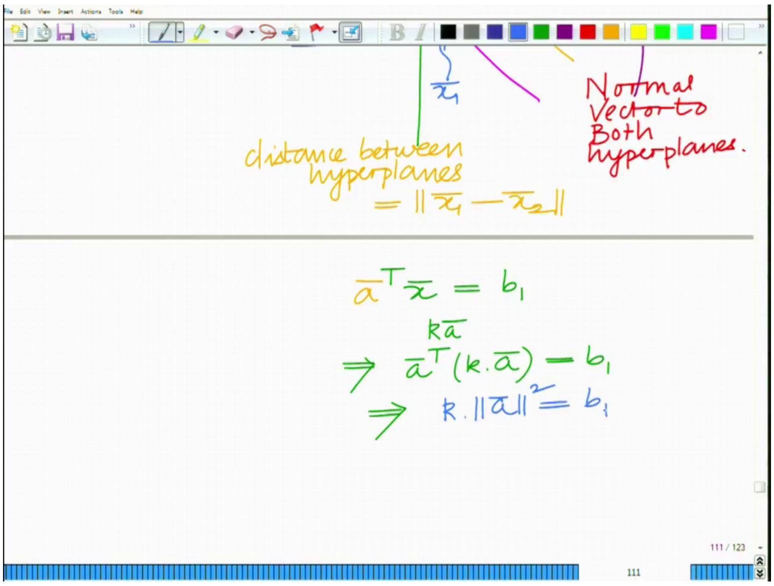


And what you want to look at in today's module is you want to look at the properties of hyperplane. So, this is example number let us call this is example number 9. So, consider two hyperplanes given by a bar transpose x bar equals b 1 and a bar transpose x bar equals b 2, recall that this is equation of hyperplane these are two hyperplanes.

And in fact, these hyperplanes you can see these are parallel you will realize that these are two parallel hyperplanes, these are parallel hyperplsanes; so, that you can draw a figure to denote this ok. So, if you represent this pictorially you find these are hyperplane one and this is your hyperplane two and both have the same normal vector, this a bar this vector is the normal vector correct this vector a bar is the normal to both the hyperplanes these are the normal, this is the normal vector to both hyperplanes and in fact, the distance between both these hyperplanes can now be calculated as follows.

If you look at this point of intersection of the normal vector with the hyperplane, that is if you call this points as x 1 bar and x 2 bar this is from your follows from a simple knowledge of high school level geometry your coordinate geometry, that is these two hyperplanes are parallel look, if you look at the point of intersection of this normal with these two hyperplanes. And if you look at the distance of these two points of intersection the distance between these two points of intersection that is the distance between these two hyperplanes. So, the distance between these two hyperplanes is the distance between these two points of intersection of the normal a bar with a hyperplane.

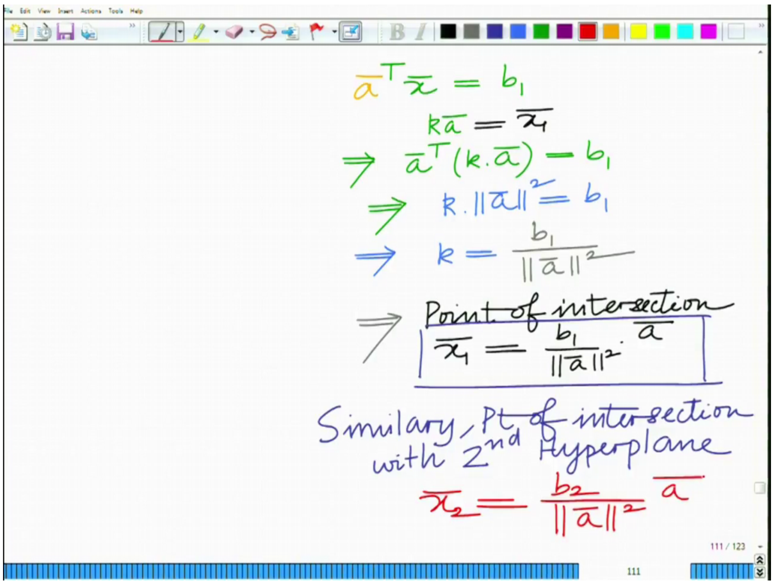
(Refer Slide Time: 03:12)



So, distance between the hyperplanes is the distance between, these points of intersection ok. And now what are these points of intersection remember the points of intersection are along the normal. So, we have x so, we have let us look at the first hyperplane that is your a bar transpose x bar equals b 1 the point along the normal, if you call that as k times a bar some constant times a bar.

Then this implies a bar transpose k times a bar equals b 1 which implies k times, now a bar transpose a bar that is norm of a bar square equals b 1, this implies the constant k equals b 1 divided by norm of a bar square.

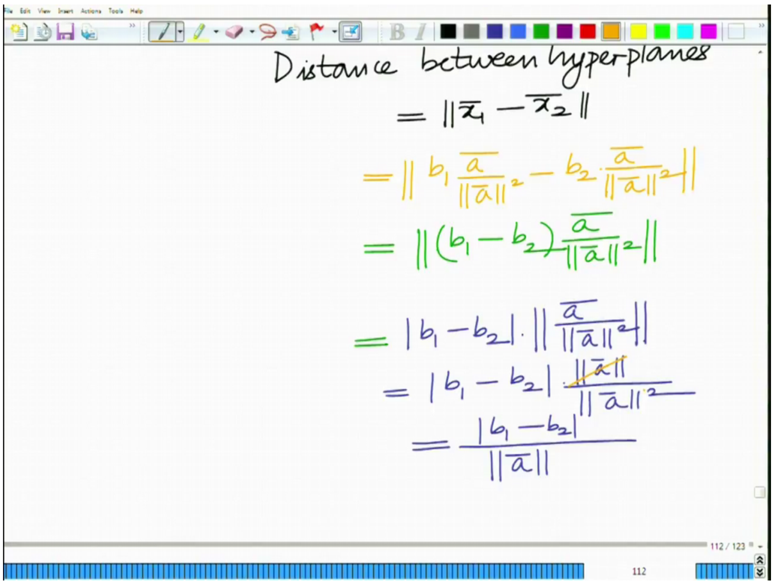
(Refer Slide Time: 04:21)



And this implies point of intersection of a bar, that is a point of intersection x 1 bar is this is your point of intersection x 1 bar is k times a bar which is b 1 divided by norm a bar square into that is k times a bar times a bar.

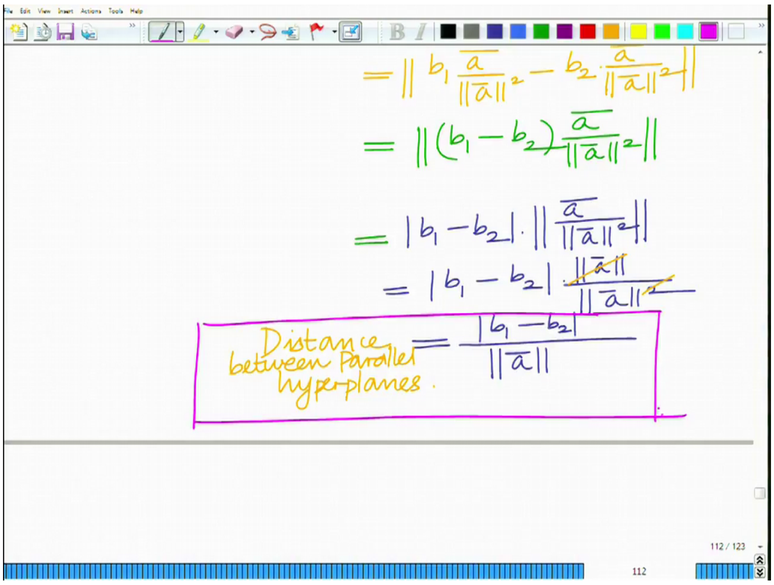
So, this is the point of intersection of the normal vector a bar with the first hyperplane. Similarly the point of intersection with the 2nd hyperplane, that is if you look at x 2 bar x 2 bar is all you have to do is replace b 1 by b 2 that is b 2 divided by norm a bar square into a bar ok. And therefore, we have found the two points of intersection with these hyperplanes of the normal vector a bar of this hyperplane and therefore, the distance between the hyperplanes is the distance between these point of intersection.

(Refer Slide Time: 06:08)



The distance norm of x 1 bar minus x 2 bar equals norm of b 1 a bar divided by norm a bar square minus b 2 a bar divided by norm a bar square, the norm of this which is equal to norm of b 1 minus b 2 times a bar divided by norm of a bar square which is equal to magnitude of b 1 minus b 2 times norm of a bar divided by norm of a bar square, which is magnitude of b 1 minus b 2 times norm of a bar divided by norm of a bar square, which is equal to magnitude of b 1 minus b 2 divided by norm of a bar. So, you have this a bar cancelling with this a bar square in the and you have norm magnitude of b 1 minus b 2 divided by norm of a bar.

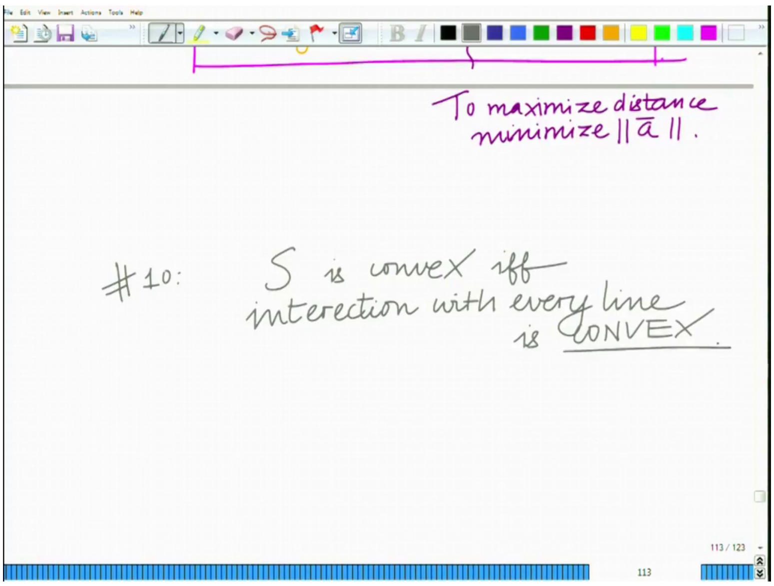
(Refer Slide Time: 07:51)



This is the distance between the between the parallel, this is the distance between the set of your parallel hyperplanes, which have the same normal vector a bar ok. And this is important because this has a lot of applications this interesting property. So, if you look at the normal if you look at these two hyperplanes, these two hyperplanes can be used for classifications. So, you can have a set of points on one side a set of points on the other side and you can use these two hyperplanes to separate these sets of points alright. So, this is a know classifier in like this is the basis for what is known as the support vector machine classifier. So, this forms this simple principle of maximizing the distance between hyperplanes this forms the basis for the SVM.

That is your support vector machine this forms the basis for the support vector machine. And this basically this maximizes maximizing the distance between this hyperplanes, basically makes the classifier more effective thereby effectively separating these two different classes of objects alright. So, been such problems your interested in maximizing the distance between the hyperplanes, which is given by you can see magnitude b 1 minus b 2 divided by norm of a bar. And therefore, if you want to maximize the distance between hyperplanes you have to minimize norm of a bar.

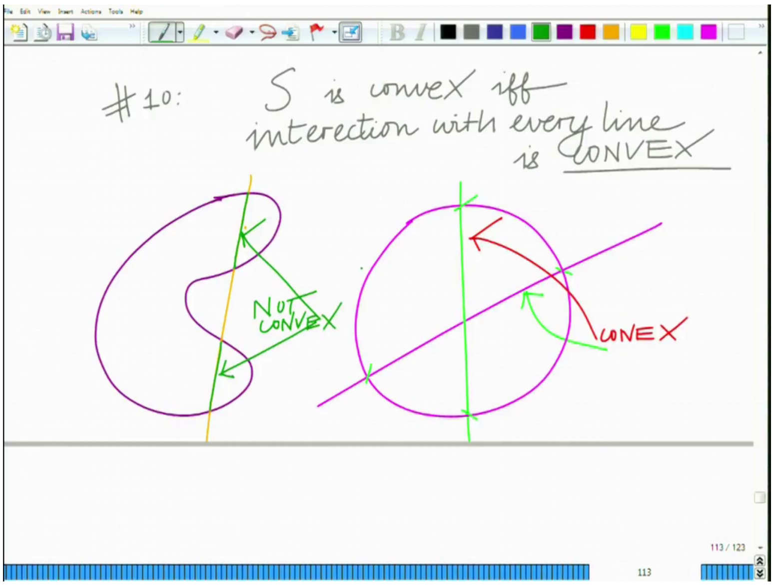
(Refer Slide Time: 09:55)



So, if you minimize to maximize distance that is if b 1 and b 2 are fixed you have to minimize normal if in fact, this is a very interesting property that pertains to classification ok. Now, let us look at another in very interesting problem this is number

10 example number 10, you want to show the set S is convex if and only if its intersection with every line is convex, what this means is that if a set alright consider any set convex set ok.

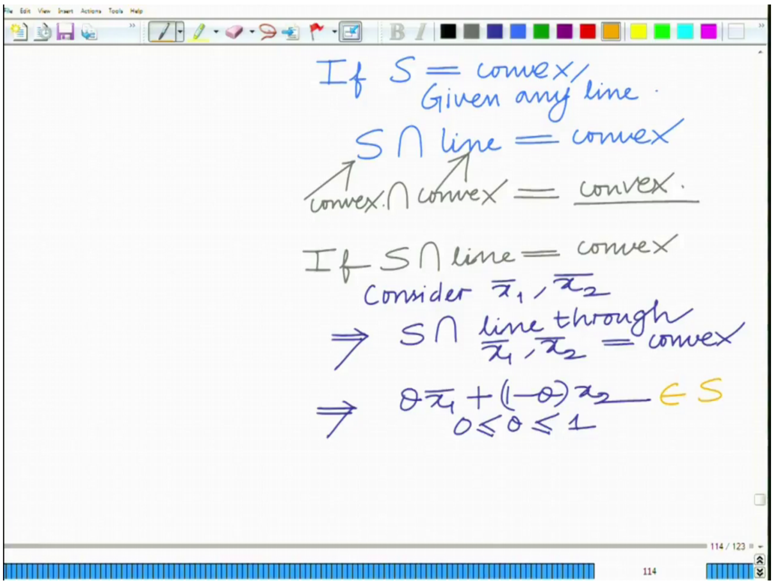
(Refer Slide Time: 11:00)



Now, if its intersection with every line is convex that is you take any line and, if you look at its intersection with the line, you can clearly see that the intersection with every line is a line segment which is itself is a convex set ok. So, the intersection with any line is a convex set.

On the other hand if you take if you take for instance a region like this our non convex set. And if you take any line, then the intersection with respect to this is this two disjointed line segments and this is not convex. So, this is an if and only if statement, but it says something very interesting. If the intersection of a set the set is convex then it intersection with every line is convex.

(Refer Slide Time: 12:27)



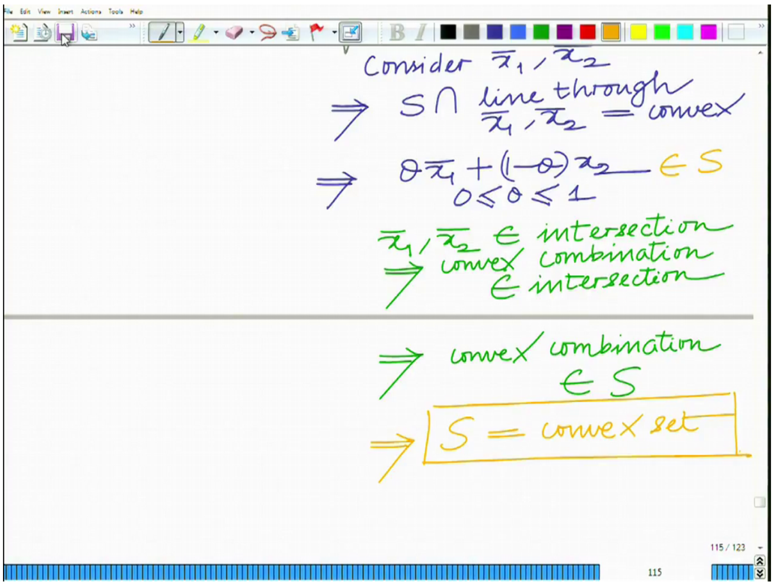
Similarly, if the intersection of set with every line is convex, then the set as is also convex and this is easy to verify you can see this as follows, let us start in one direction if S is convex. And the intersection and the, if S is convex now consider any line given any line, now S intersection line is convex if S is convex.

The set S intersection the line is convex, this is true because S is convex given that S is convex any line is a convex set we know that. So, convex intersection convex this is convex this is trivia ok. So, S is convex the set S is convex, then a line any line is also a convex set. So, S intersection with the line convex set intersection with another convex so, that is also convex, let us move it in the other direction. If S intersection with any line is convex ok, consider any two points x 1 bar x 2 bar

This implies S intersection with line through x 1, because that is also a line a line through consider a line through x 1 bar comma x 2 bar, this is also a line this implies the intersection with the line through x 1 bar plus x 2 bar is convex. This implies that if you look at any convex combination theta times x 1 bar plus 1 minus theta times x 2 bar 0 less than equal to theta less than equal to 1. This must belongs must belong to as the reason is the following, because x 1 bar x 2 bar belongs to S, x 1 bar x 2 belongs to the line through x 1 bar x 2 bar. Therefore, x 1 bar x 2 bar belongs to the intersection of us with the line alright.

Now, if you look at any convex combination that has to naturally belong to this intersection, because of the convex combination does not belong to the intersection; that means, the intersection is not convex combination.

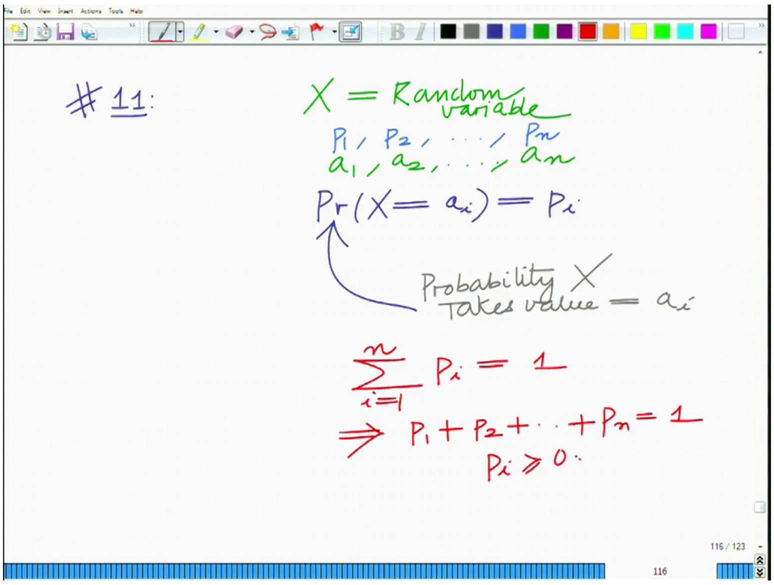
(Refer Slide Time: 15:21)



Therefore, the convex combination of x 1 bar, because x 1 bar x 2 bar belong to the intersection. The convex combination of x 1 bar x 2 bar must also belong to the intersection correct, x 1 bar x 2 bar implies their convex combination must also belong to the intersection, otherwise the intersection is not going to be convex, which intern implies that the convex combination belongs to S.

Because the convex combination belongs to the intersection, which implies that S is the which implies that S a convex set. So, the convex set if you consider an intersection with any line is convex. And the other direction on other hand for conversely, if the intersection of set S with any line is convex, then the set it S itself is a must be a convex set alright, this is a very interesting property and often very useful in demonstrating the convexity of sets alright.

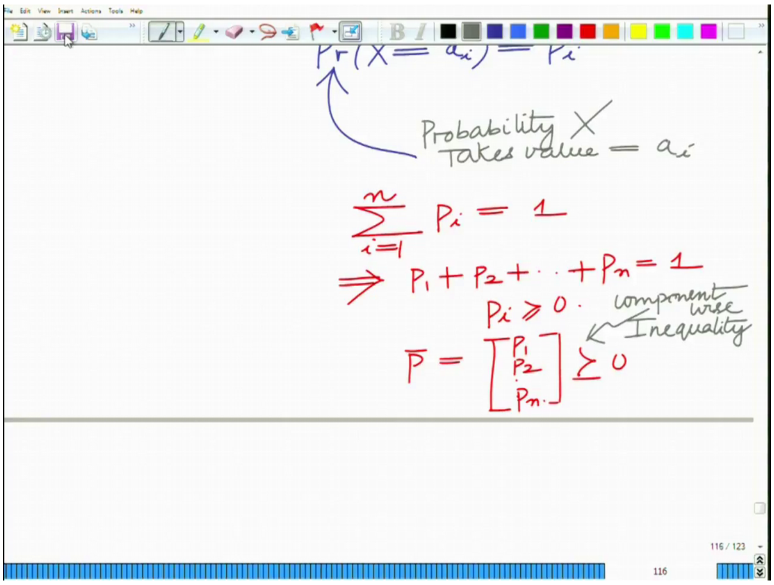
(Refer Slide Time: 17:04)



Let us look at another interesting problem that pertains to probabilities problem number 11. That is let X be a random variable and it will take values a 1 a 2 up to a n and the probability that X takes the value a I, this is given by the this is equal to P i. So, this is the probability X random variable x takes the value a i probability X takes the value ok. Now, naturally if you can look at the set of all this probabilities. So, we have probabilities a 1 a 2 a n corresponding probability P 1 P 2 P n.

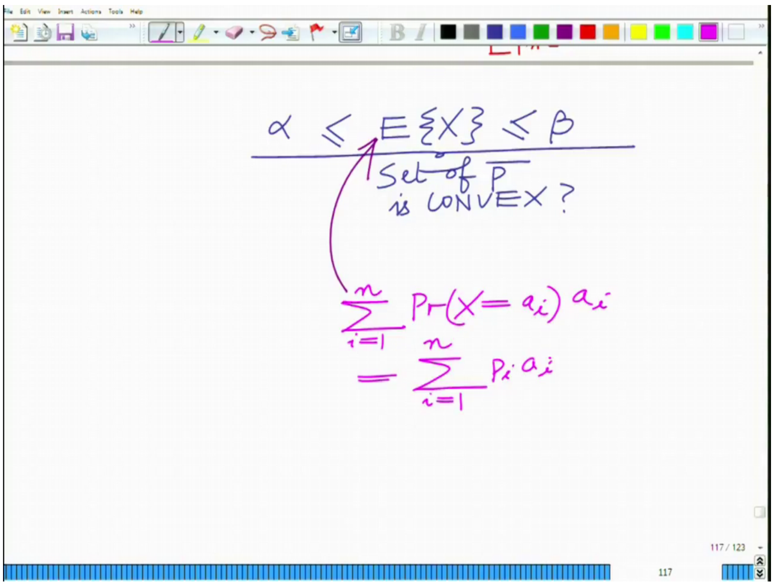
Now, one naturally the sum of the probabilities must be one first this probability each of these properties, because their probabilities they must be greater than equal to 0. And further the sum of all this probabilities must be equal to 1. Therefore, we must have summation i equal to n P i equals to 1, which is basically also represented by P 1 plus P 2 plus P n equals to 1. And we must also have each P i is greater than or equal to 0.

(Refer Slide Time: 18:56)



And, if you denote this by the vector P bar P 1 P 2 P n and then remember, you can use the component wise inequality symbol to say this vector P bar, each component remember this is our component wise inequality, this is the component wise inequality. Each component of this vector P bar must be greater than equal to 0. Now, I want to examine some properties of this set that contains of this set of vectors P bar let us look at the first one ok.

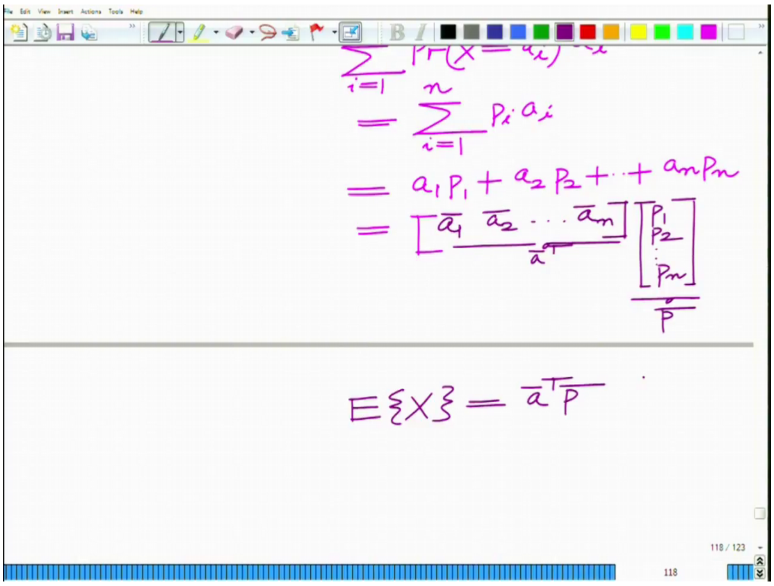
(Refer Slide Time: 19:45)



Let us look at all the probability vectors P bar, such that alpha is less than or equal to the expected value of a random variable X less than or equal to beta. Set of all P bar that satisfy this is this set convex, we would not ask this question is the set of all probabilities P bar, such that the expected value of a random variable X lies between alpha and beta is this set convex well that is easy to figure out.

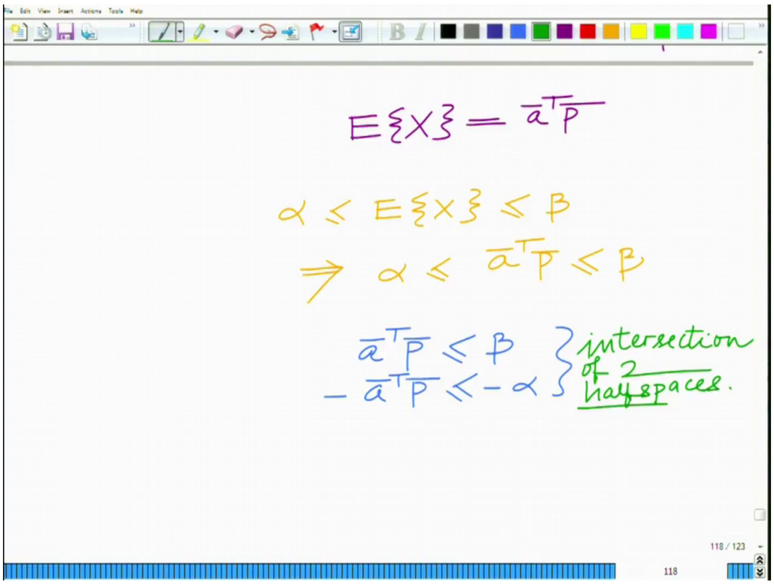
If you look at the expected value of a random variable X, that when we calculated as follows that a summation i equals to 1 to n. The expected value of a random variable is probability that it takes the value a i into a i which is summation i equals to 1 to n P i a i.

(Refer Slide Time: 20:48)



Which is basically a 1 times P 1 plus a 2 times P 2 plus a n times P n, which you can find as a 1 bar a 2 bar a n bar time P 1 P 2 P n, you can given this as a bar transpose this is your P bar. So, expected value of the random variable X, now becomes a bar transpose P bar.

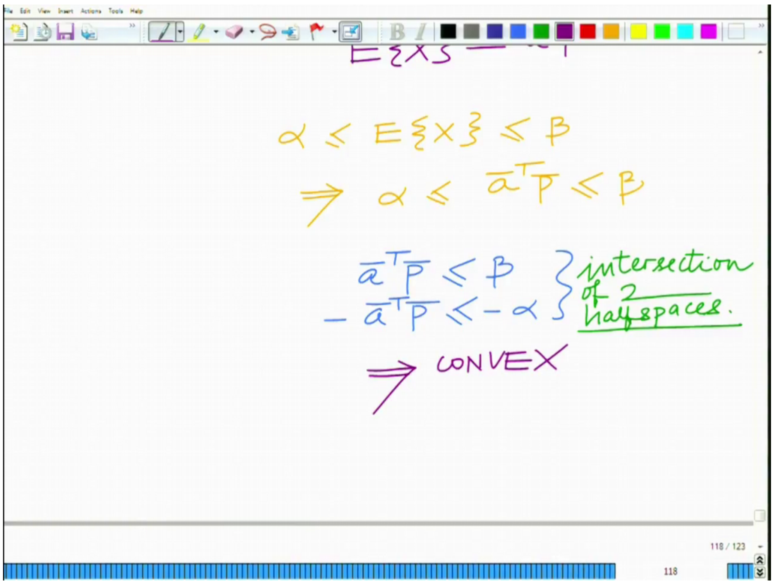
(Refer Slide Time: 21:36)



And therefore, if you look at this problem alpha is less than or equal to expected value of X less than or equal to beta this implies alpha is less than or equal to a bar transpose P bar less than or equal to beta. So, this is the intersection of two hyperplanes, you can readily see that this is the first hyper plane is a bar transpose P bar less than or equal to beta.

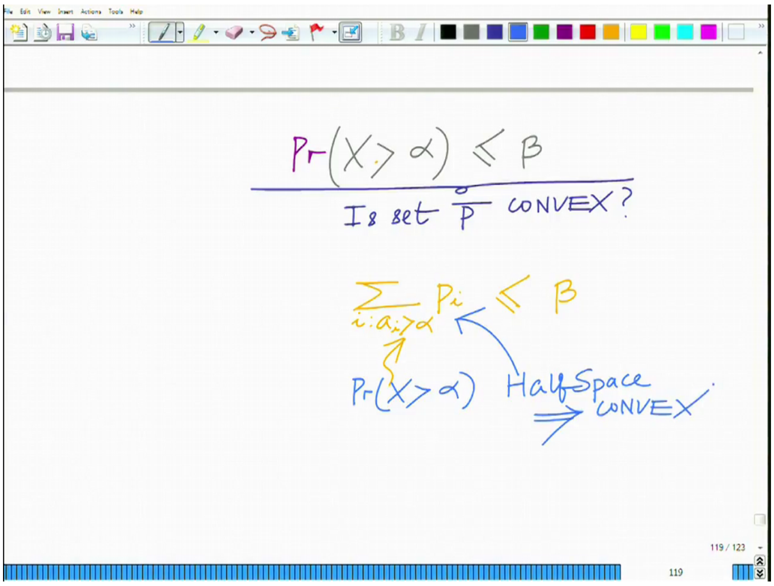
Second hyper plane is minus a bar a bar transpose P bar greater than equal to alpha, which can be written as a bar transpose P bar less than equal to minus a bar transpose P bar less than equal to minus alpha. So, this is the intersection of two this is the in fact, the intersection of two half spaces, I am sorry intersection of two this is the intersection of two half spaces implies that this is convex.

(Refer Slide Time: 22:48)



Each is half space expected value of X is less than equal to alpha, that is can be represented as the half space a bar transpose P bar is less than equal to beta, expected value of x greater than equal to alpha is another half space. So, this forms the intersection of two half spaces and therefore, this is indeed this set is indeed this is indeed a convex set alright.

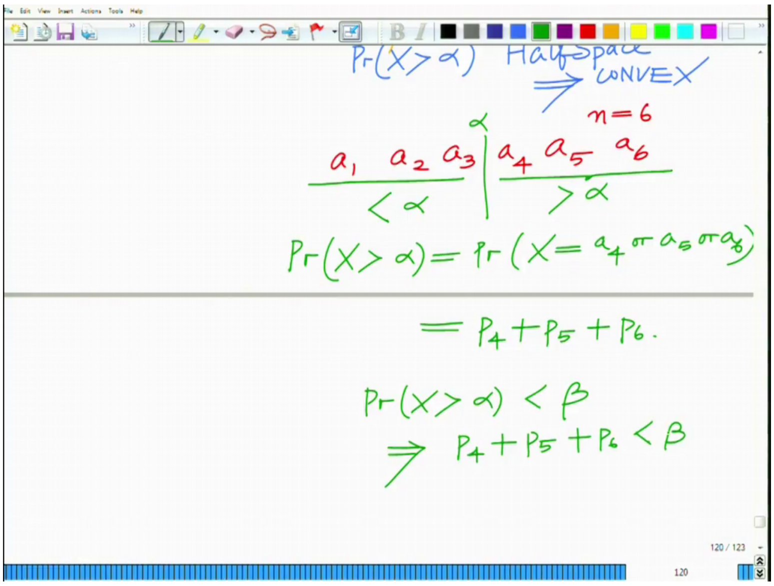
(Refer Slide Time: 23:23)



Let us look at another set probability of X greater than alpha, that is the set of all probability vectors such that the probability of X greater than equal to alpha, is less than or equal to another constant beta or property of X greater than alpha is less than or equal to beta. Is this set of all P bar that satisfies this is the resulting set of P bar convex well what is the probability that X is greater than alpha the probability that X is greater than alpha.

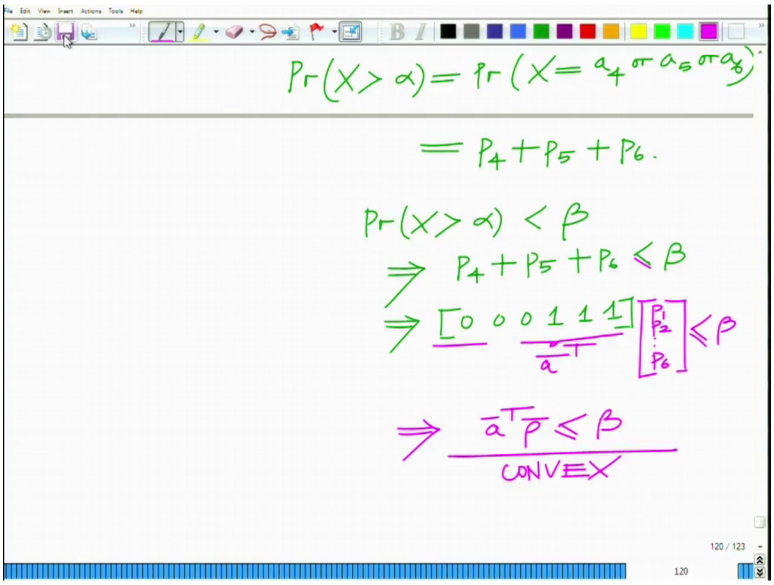
A simply summation of all probabilities P i such that the corresponding a i are greater than alpha. And this since this probability X greater than alpha has to be less than equal to beta, this is less than equal to beta this is probability X greater than alpha that is you simply have to sum the probabilities of corresponding to all a i S. That is greater than alpha and you can clearly see, this is the linear sum this is the half space, this is the half space implies, once again this is convex.

(Refer Slide Time: 24:54)



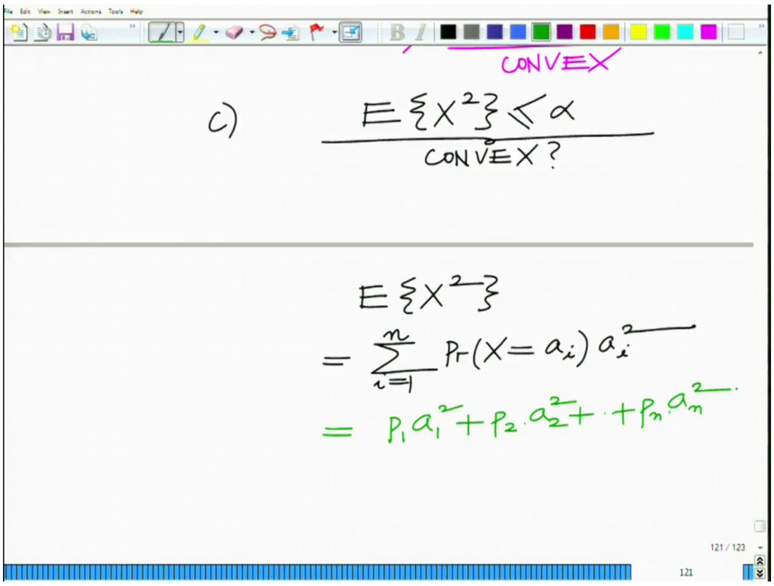
For example, let us take an example to understand this better for instance let us say you have a 1 a 2 a 2 a 4 or up to a 5 a 6, take the simple example n equal to 6. Now, let us say your alpha is here lies between a 3 and a 4. So, these are less than alpha so, these are greater than alpha and these are less than alpha. So, the probability X is greater than alpha equals the probability X equals either a 4 or a 5 or a 6 equals P 4 plus P 5 plus P 6. So, probability X greater than alpha less than beta implies P 4 plus P 5 plus P 6 less than beta.

(Refer Slide Time: 26:02)



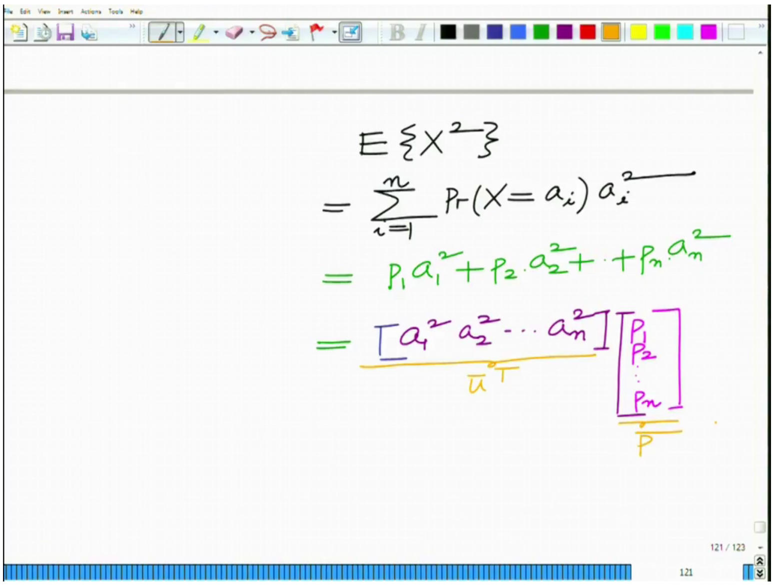
This implies basically 0 0 0 1 1 1.This is your a bar transpose times P 1 P 2 up to P six this is I am sorry less than equal to beta, this is your a bar transpose this implies you can write this as a bar transpose P bar less than equal to beta and this is a convex set ok. So, the set of all probability vectors P bar such that the probability X is greater than alpha some quantity fixed constant alpha is less than equal to beta is a convex set ok. Now, what about the second moment what about expected value of X square and this is an interesting aspect.

(Refer Slide Time: 27:01)



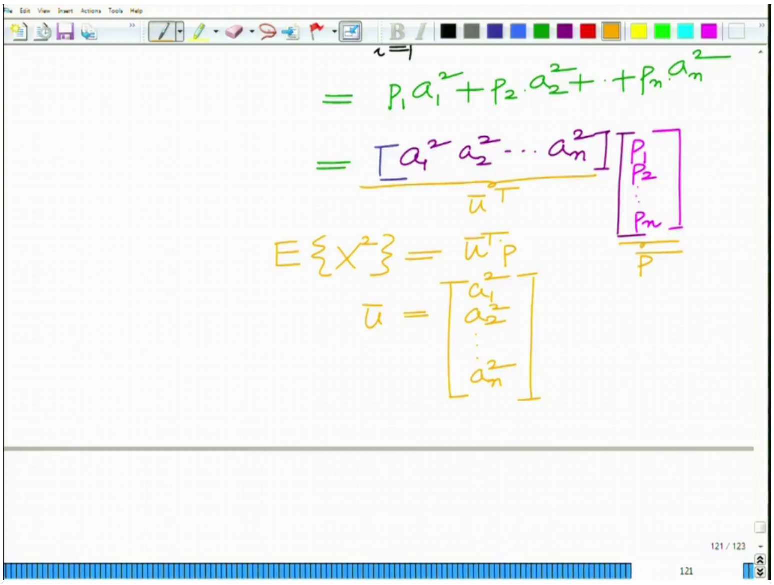
Let us look at the set of all vectors P bar such that expected value of X square is less than or equal to alpha. Now, we want to ask the question is this convex well, what is the expected value of X square, this might seem a little confusing, because X square is non-linear ok. But, what is but look at expected value of X square this is summation i equals 1 to n probability X equals, well probability X equals probability X equals a i into a i square that is the expected value of X square ok, which is equal to P 1 times a 1 square probability X equal to a 2 times a 2 square plus P n times a n square ok, which is now you can again write it as a different vector transpose times P bar.

(Refer Slide Time: 28:09)



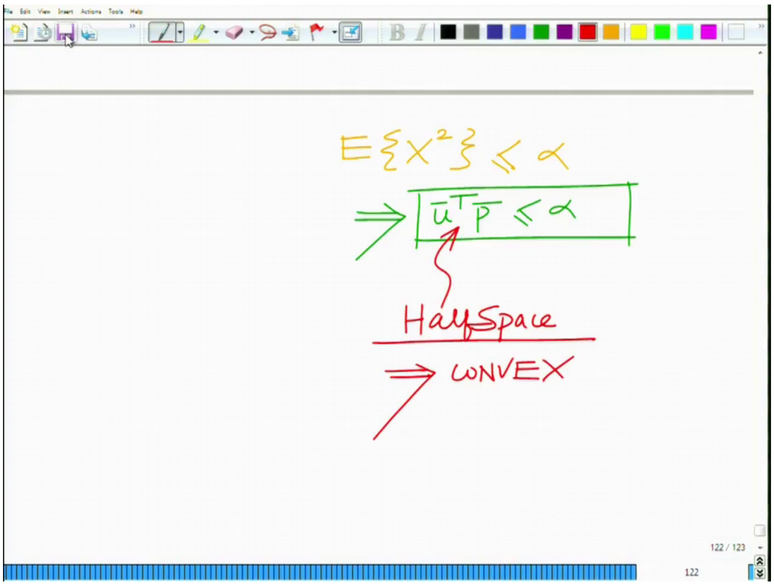
So, I can write this as a 1 square it is very interesting, I can write this as a 1 square a 2 square times a n square times the vector P bar P 1 P 2 up to P n ok.

(Refer Slide Time: 28:46)



Now, this if you look at this, this is a different vector you call this as well let us call this as u bar transpose, this is your vector P bar. where what is u bar; u bar is this vector, we are now calling this vector a 1 square a 2 square so, on a n square by this vector u bar. So, expected value of X square if you look at that is, u bar transpose P bar and expected value of X square less than equal to alpha.

(Refer Slide Time: 29:11)



This implies u bar transpose P bar less than equal to alpha. And you can see this is once again corresponds to a half space implies, this is yes this is therefore, convex. And therefore, once again the set of all probability vectors once again the set of all probability vectors P bar such that expected value of X square is less than or equal to alpha the set of all such vectors P bar is once again convex alright.

So, these are some interesting applications of the notion of convexity, convex sets some of the properties of convex sets and so, on which have heavy application or which are going to be used very frequently in our discussion on optimization theory, on our discussion on the practical applications of optimization bit in the context of wireless communication, or signal processing or several other fields.

So, these form the these principles these examples that we are so, far seen from the basic building blocks of several large problems, or several large how do you put it several large paradigms or frameworks, that we are going to explore in the future with respect to optimization and its application in several areas of interest.

Thank you very much.