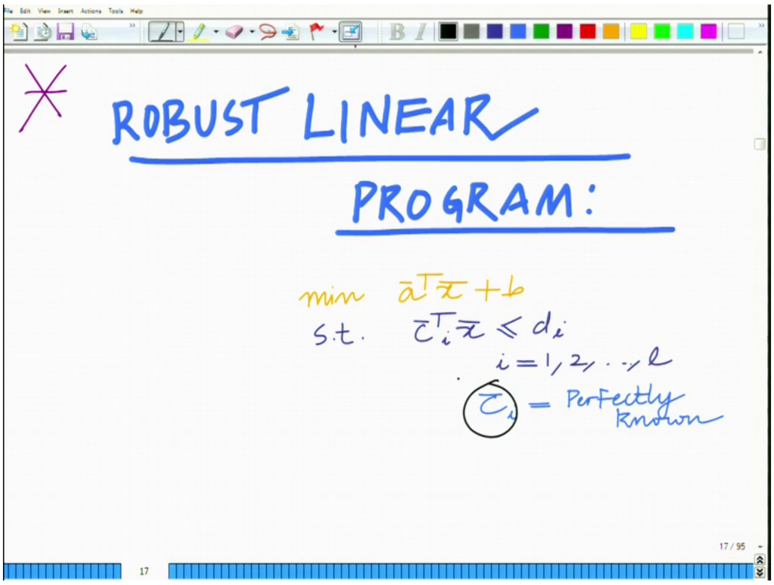
Applied Optimization for Wireless, Machine Learning, Big Data

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

Lecture - 49
Stochastic Linear Program, Gaussian Uncertainty

Hello, welcome to another module in this Massive Open Online Course. So, we are looking at different kinds, of we are looking at convex optimization problems and in almost specifically we are looking at the specific class of convex optimization problems this is basically linear programs ok. We have looked at linear programs and also demonstrated the practical application of linear program all right. And in this module let us start looking at our variation of that or an extension of it know known as the robust linear program all right.

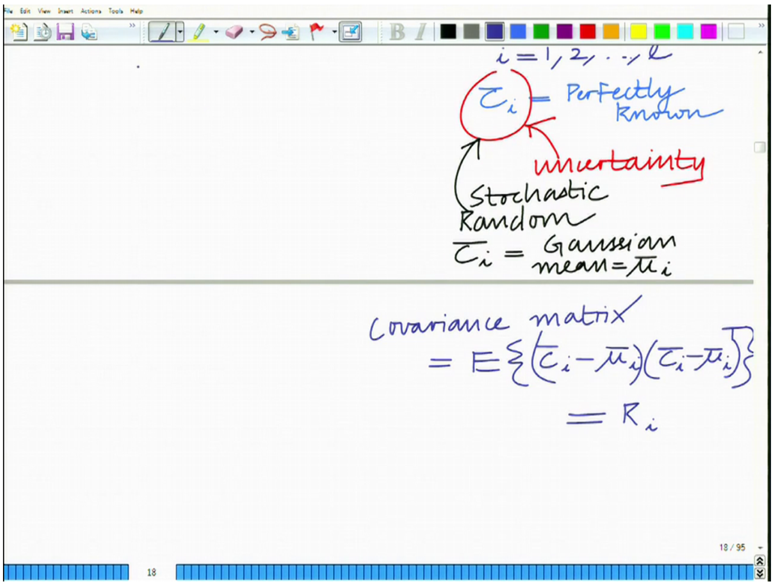
(Refer Slide Time: 00:49)



So, what you want to look at in this module is another interesting extension of a linear program and which is also very useful practically that is robust linear program. And linear program if you remember can be formulated as follows that is we want to minimise a bar transpose x bar these are objective function which in a fine objective. And this is subject to the constraint up again an a fine constraints that is C bar transpose i x bar less than or equal to d i or i equals to 1 to l.

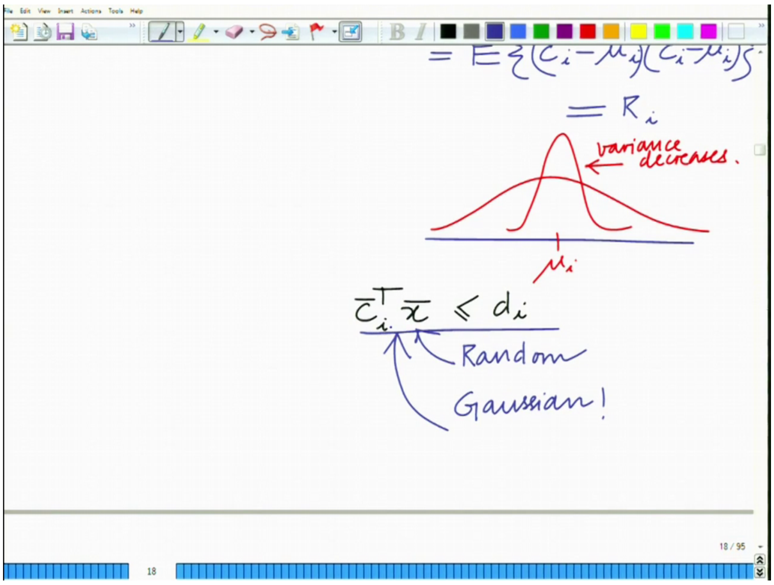
And now, however, while formulating this linear program, we have assumed this constraint that is this C bar i that is if you look at these vectors ok, we have assume these to be perfectly known. Now, remember this can be perfectly known. Now, remember this C bar i, this vector C bar i which characterize this problem, they depend on the problem right, for instance in our base station corporation, there they are basically the power gains all right, the gains between the base stations and the various use all right. So, these have to be estimated in practice for any particular problem which means naturally that this going to be a certain level of uncertain or there can be a certain level of uncertainty in this. So, what happens in as we have seen many times before and practical scenarios and this is specially that is what we have said it is very, it has a lot of practical relevance.

(Refer Slide Time: 02:49)



There can be uncertainty in this C bar all right. So and there can be various levels of uncertainty. Now, there can be various models for this. Now, one interesting way or one practically useful model for such scenarios where is uncertainty C bar i is to assume that these C bar i's are random in nature all right. And this also gives rise to a stochastic linear program or stochastic version of this linear program. So, one can assume that the C bar i's random in nature in particular you can assume that the c bar i's are Gaussian these are Gaussian random vectors with their nominal value which is their mean; mean equals mu bar i and the covariance that is if you look at this C bar i's C bar i minus mu bar i C bar i minus mu bar i transpose this is equal to R i. This is the covariance matrix you can say this is the covariance matrix.

(Refer Slide Time: 04:29)



So, and remember the covariance characterizes the spread around the mean all right. So, this is a vector I am simplifying this for a scalar, for instance we have a Gaussian all right and this is your mean mu i. And as the variance decreases, as the variance decreases, it becomes more and more concentrated on the mean all right. So, C bar i is a Gaussian random vector the nominal value or let us say the estimated value of this is mu bar i that is the vector. And the covariance matrix characterizes the spread around that. If the covariance matrix let us say the eigenvalues or if the variances of these different elements of C bar i are low, it means that it is very close to the mu; if the variances are large it means that it has a large spread around the mu ok.

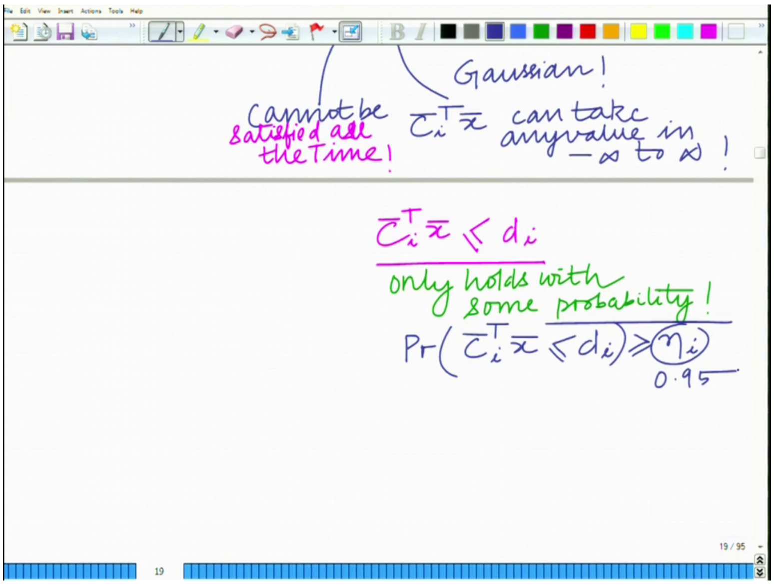
So, basically that characterizes this problem. And of course, when the covariance matrix tends to 0 or becomes very close to 0, it reduces the previous version, because in that case C bar i becomes equal to mu bar i that it is equal to the mean mu bar i with probability 1; and again reduced to the deterministic linear program that we have seen before ok.

So, now, for this stochastic linear program or this you know linear program in which the C bar i's are random, how do you formulate that. Now, if you look at the constraints let us go back and take a look at the constraint the constraint is C bar transpose i into x bar is less than or equal to d i. Now, if you look at this C bar is random, in fact, C bar is

Gaussian x bar is a vector correct. Now, therefore, C now C bar C bar i transpose x bar this is also random because C bar is a random vector. So, this implies this is random ok.

And in fact, you realize something interesting that this is we can also say that this is Gaussian reason being C bar i is a Gaussian random vector C bar i transpose x bar is a linear combination of the components of this vector C bar i. When you linearly combine Gaussian random variables, you get another Gaussian random. Therefore, c bar there is a linear transformation of Gaussian random variables yields another Gaussian random variable. Therefore, C bar i transpose x bar is a Gaussian random variable in fact all right.

(Refer Slide Time: 07:01)

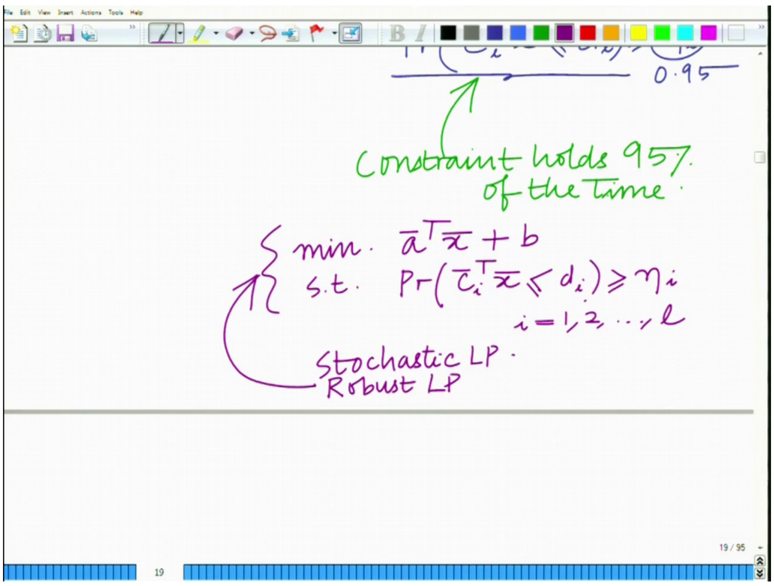


And what you see is more interestingly since this is C bar i transpose x bar i is Gaussian random variable C bar i transpose x bar x bar this belongs or this can lie can take any value can take any value in minus infinity to infinity that is the other interesting thing. So, remember C bar i transpose x bar previously once a deterministic one, and now it is a random. So, you can take any value between minus infinity and infinity. Which means that it need not be less than or equal to di for some value, that is there can always exist some random some values of this vector C bar i where this exceeds di. One cannot because of the random nature of this constraint one cannot hope that this constraint is always satisfied in the optimization problem alright.

So, because this is random need not or cannot be satisfied all the time, because it is varying randomly, cannot be satisfied. This cannot be satisfied all, it cannot be satisfied all the time, which means this only holds with a certain probability I can expect this to hold. So, this holds only with some probability. So, only holds this only holds with some probability let us say the probability is eta. So, what eta i, what we mean by that or let us say this is beta i.

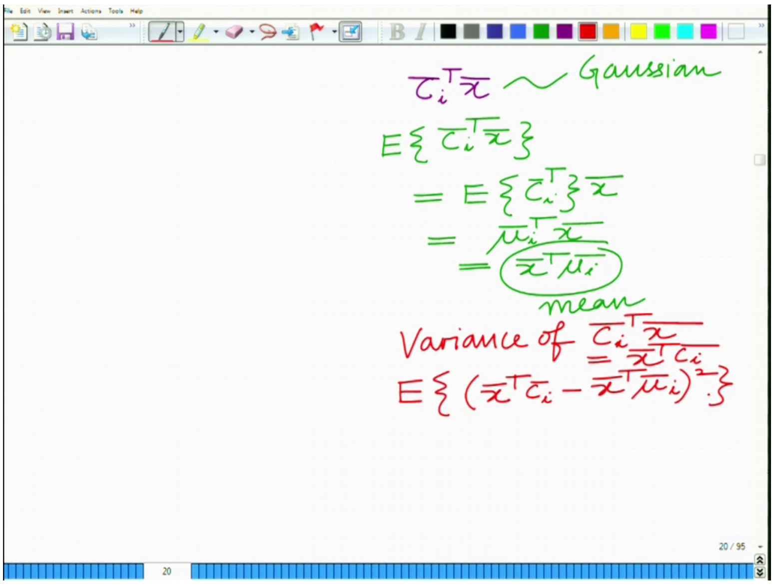
So, what or eta i, let us say the probability that C bar transpose i x bar less than or equal to d i has to be greater than or equal to eta i which means the, what we are saying is very interesting the probability with which the constraint holds that this you can only talk in terms of probability. So, this constraint holds with probability eta i that a what we are saying is this constraint need not hold all the time, but it can hold with a very high probability that is the probability with which this constraint holds is has to be greater than or equal to eta i. For instance, eta i can be let us say 95 percent just take a simple example which means that this constraint has to hold 95 percent of the time ok. So, eta i can be let us make it 0.95 95 percent or 0.95 which means this constraint has to hold; which means this constraint which means this constraint holds 95 percent of the time..

(Refer Slide Time: 09:59)



And therefore, now I can modify this linear program, now I can modify this optimization problem. I can write this as minimise a bar transpose x bar plus b subject to the constraint that the probability C bar i transpose x bar less than or equal to d i greater than or equal to eta i i equals 1, 2 up to l. So, this is your modified stochastic, you call this as robust LP or you call it as a stochastic because a constraint is random in nature or holds with a certain probability. You can also called this as robust LP because you are ensuring that you are taking the uncertainty into C bar i uncertainty in C bar i into account.

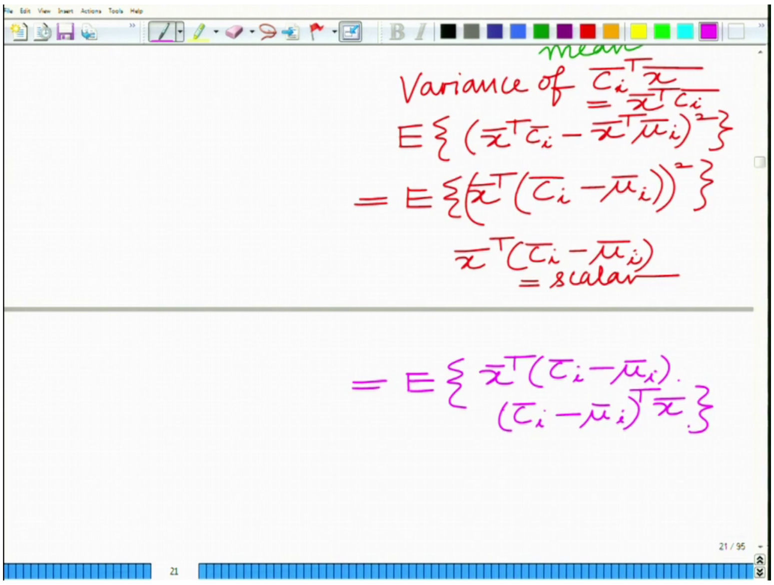
(Refer Slide Time: 11:47)



Now, let us modify this problem further. Now, let us look at this quantity C bar i transpose x bar; if you look at this quantity C bar i transpose x bar, we have already said that this quantity this is a Gaussian random variable. Now, let us find the mean and variance of this Gaussian random variable, we have expected mean is simple expected value of C bar i transpose x bar this is expected value of C bar i transpose times x bar or which is equal to expected value of C bar i is mu bar i. So, this is mu bar i transpose x bar which is equal to x bar transpose mu bar i. So, this is the mean, this is the mean.

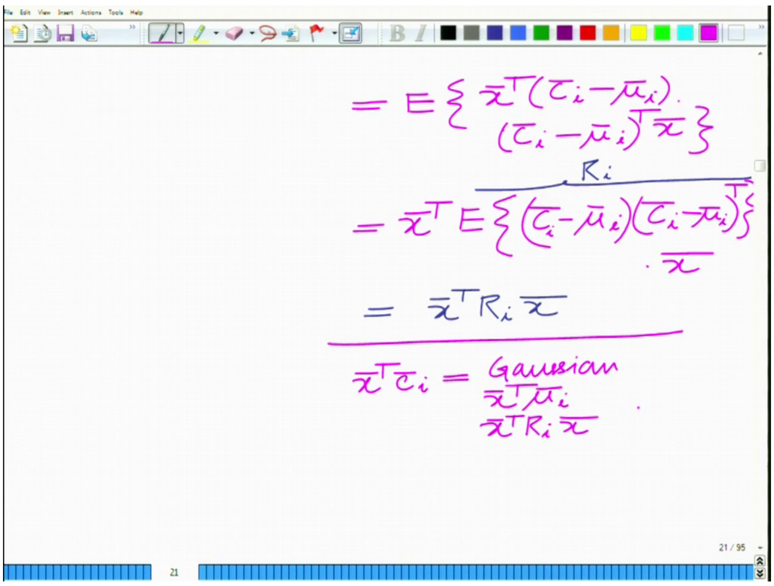
Now, what is the variance of this, what is the variance? Now, variance is the expected value of the random variable that is C bar i transpose x bar or you can write it as x bar transpose C bar i. So, x bar transpose C bar i minus its mean, which is x bar transpose mu bar i whole square.

(Refer Slide Time: 13:23)



Now, I can write this as expected value of x bar transpose C bar i minus mu bar i whole square. Now, this is a scalar quantity. So, I can write this as expected, so I can write it as a quantity into itself or quantity into its transpose because this is a scalar quantity x bar transpose C bar i minus mu bar i this is the scalar quantity or basically it is simply a number. So, this is simply a number. So, this is therefore equal to expected value of x bar transpose C bar i minus mu bar i times its transpose which is C bar i minus mu bar i transpose times x bar.

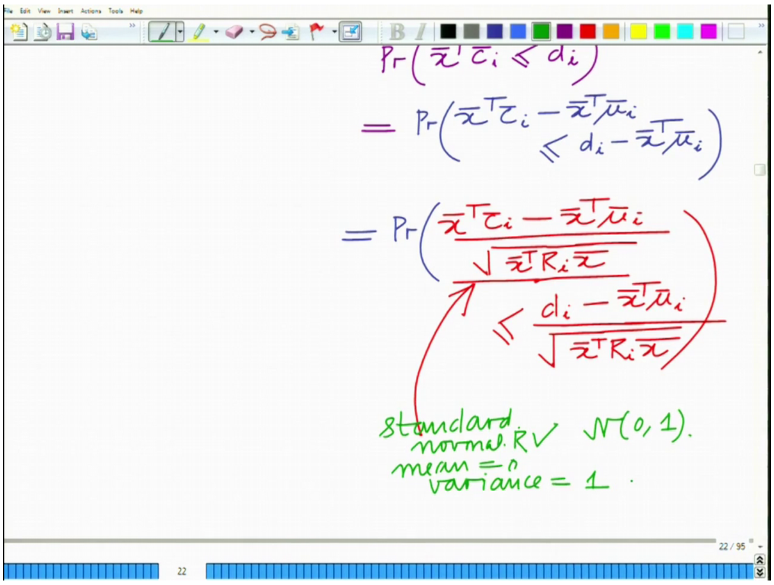
(Refer Slide Time: 14:27)



Which is basically expected value of I can take the x bar outside. So, this is x bar transpose and now you have something interesting expected value of C bar i minus mu bar i time C bar i minus mu bar i transpose into x bar. And this is equal to x bar transpose R i into x bar, where R i you can see this is the covariance that is expected value of C bar i minus mu bar i into C bar i minus mu bar i transpose. So, basically the variance of this C bar i transpose x bar or x bar transpose C bar i, this is x bar transpose R i into x bar that is the variance of this quantity. And therefore, and therefore, now what we have seen is x bar transpose C bar i, this is Gaussian with mean x bar transpose mu bar i and variance x bar transpose R i into x bar ok.

And now let us find remember our constraint involves the probability. So, now, let us find what is the probability that x bar transpose mu bar i is less than or equal to. So, therefore, x bar transpose C bar i is Gaussian alright. Now, let us go back to the constraint and find what is the probability that this let us simplify the probability that x bar transpose C bar i is indeed less than or equal to d i.

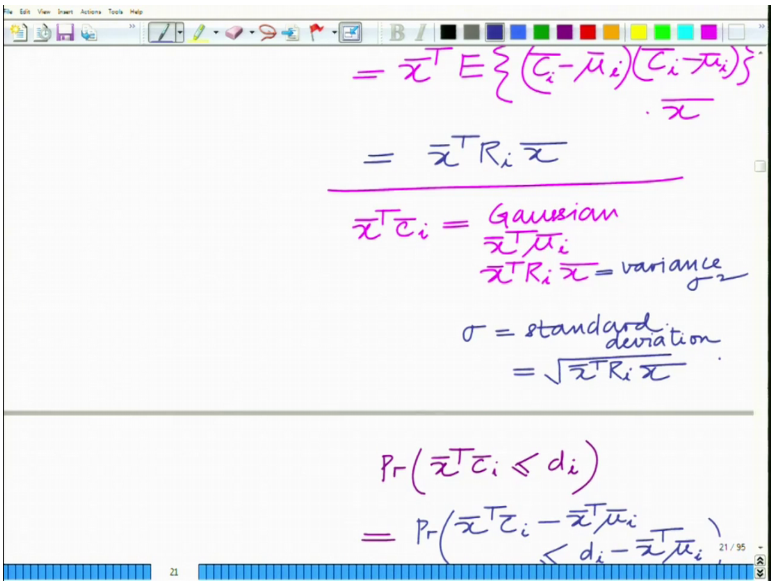
(Refer Slide Time: 16:33)



Now, if you look at that the probability, probability x bar transpose C bar i less than equal to d i. This is equal to the probability, now x bar transpose let us me a subtract the mean x bar transpose C bar i minus x bar transpose mu bar i less than or equal to d i minus x bar transpose mu bar i alright. I am subtracting the mean x bar transpose mu bar

i. Now, I am going to denote divide by the variance, remember the divided by the square root of the variance.

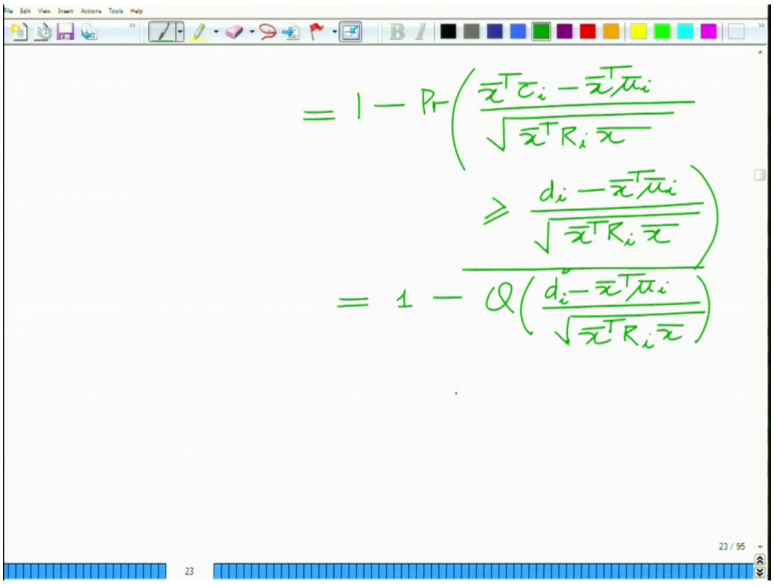
(Refer Slide Time: 17:21)



So, this is the variance and standard deviation, you can think of this as sigma square. So, sigma equals standard deviation equals square root of x bar transpose R i into x bar. And so therefore, this is equal to the probability divide both sides by the standard deviation. This is equal to the probability that x bar transpose C bar i minus x bar transpose mu bar i divided by square root of x bar transpose R i x bar is less than or equal to d i minus x bar transpose mu bar i divided by square root of x bar transpose R i into x bar.

And now if you look at this, if you look at this quantity, what we have done is from a Gaussian random variable we have basically subtracted the mean and divided by the standard deviation, so that gives us a zero mean unit variance Gaussian random variable which is nothing but the standard normal random variable. So, this gives us this manipulation has given us the standard normal, what we call the standard normal R means that is Gaussian random variable mean equal to 0. Now, the probability that this is less than equal to this is 1 minus the probability that this is greater than equal to this.

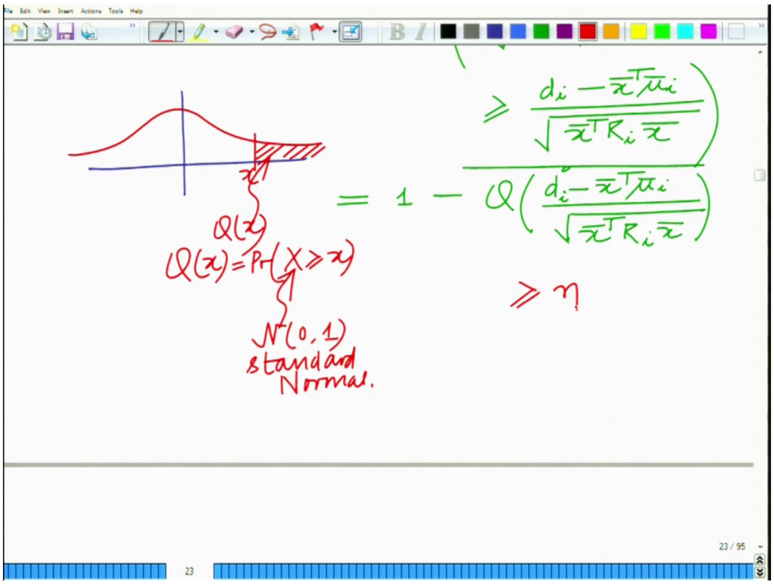
(Refer Slide Time: 19:29)



So, this is equal to 1 minus the probability that this quantity x bar transpose C bar i minus x bar transpose mu bar i divided by square root x bar transpose R i x bar is greater than or equal to x bar transpose mu bar i divided by x bar transpose R i into x bar. And now, we have probability that the standard the standard normal random variable Gaussian random variable mean 0, variance 1 is greater than equal to some threshold that is given by the Q function alright.

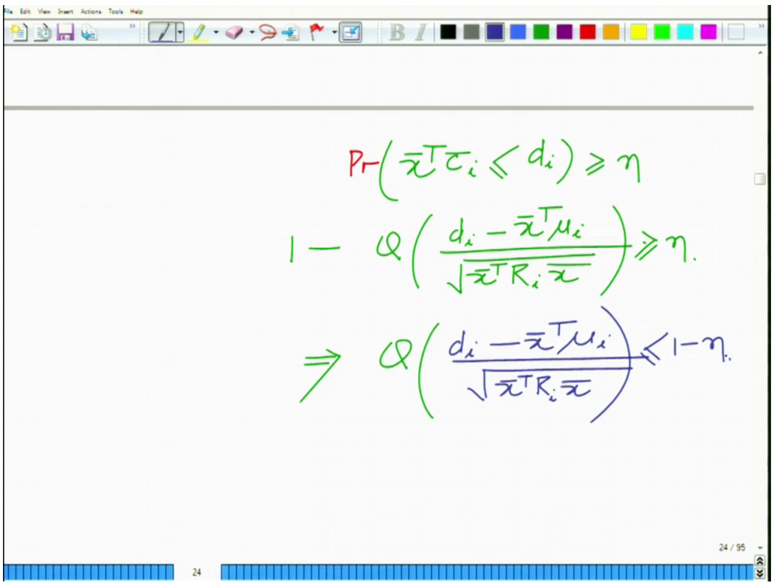
So, this quantity is nothing but the tail probability of the standard normal which is equal to Q of d minus x bar or d i minus x bar transpose mu bar i divided by under root x bar transpose R i x bar variance. And this quantity is equal to 1 minus that, 1 minus the tail problem.

(Refer Slide Time: 21:05)



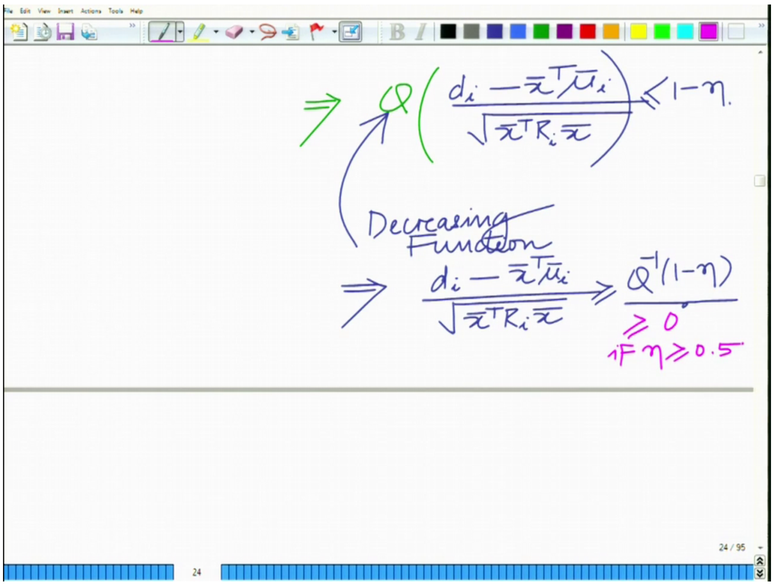
And therefore, the probability and that is what we are saying right. This is the tail probability that is a probability standard normal if you remember Q of x equals probability x greater than or equal to x, where x is your standard normal that is Gaussian random variable with mean 0 and variance unity. And we need this probability to be greater than or equal to eta. So, this means this quantity is greater than equal to eta. So, therefore what we have is, if you look at it what we have is the following thing.

(Refer Slide Time: 21:57)



The probability x bar transpose C bar i less than or equal to d i greater than equal to eta this implies the Q of di minus x bar transpose mu i divided by x bar transpose R i into x bar square root, this has to be or 1 minus Q I am sorry 1 minus Q has to be greater than or equal to eta ok. And this implies that Q of x bar transpose mu i divided by square root x bar transpose R i into x bar is less than or equal to 1 minus eta.

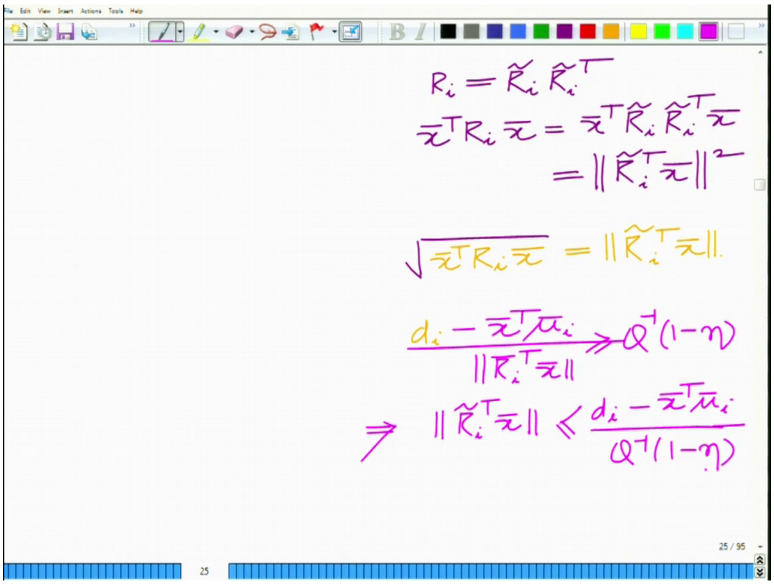
(Refer Slide Time: 23:13)



Now, the Q function is a decreasing function implies, this implies di minus x bar T transpose mu bar i, sorry this is has to be all mu bar i over mu bar i over x bar transpose R i x, bar this has to be greater than equal to q inverse 1 minus eta ok. This is the Q function is the decreasing function the equality gets reversed. So, if this has to be less than equal to this that implies that this quantity has to be greater than equal to Q inverse of 1 minus eta.

And note that if eta is greater than 0.5 that is the reliability with which this has to be hold is greater than 50 percent then this Q inverse 1 minus eta this is greater than equal to 0. So, this quantity is greater than equal to 0, eta is greater than equal to 0.5

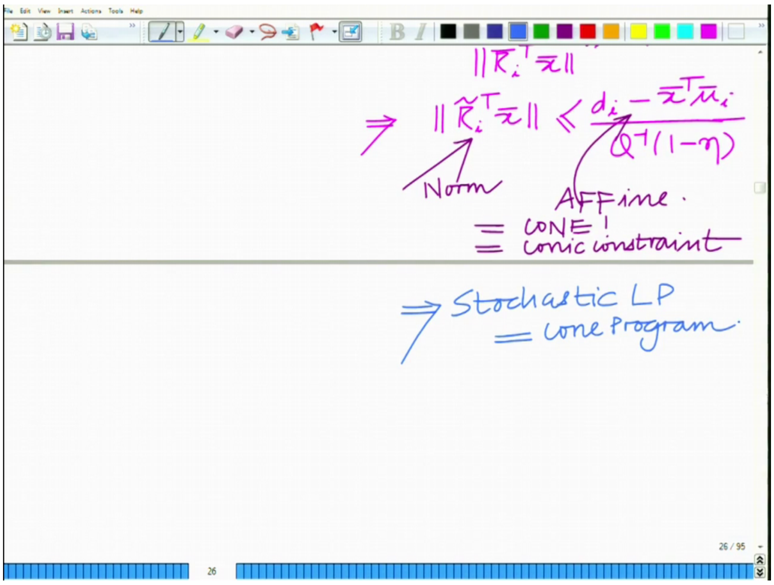
(Refer Slide Time: 24:27)



Let me just simplify this now further into something that is interesting. Remember R i is a covariance matrix. So, R i is a positive semi definite matrix. So, I can factor it as R i tilde R i tilde transpose this implies x bar x bar transpose R i x bar equals x x bar transpose R i tilde R i tilde transpose x bar. And if you look at this; this is nothing but R i tilde x bar norm square. And therefore, square root of x bar transpose R i x bar equals norm of R i tilde transpose x bar.

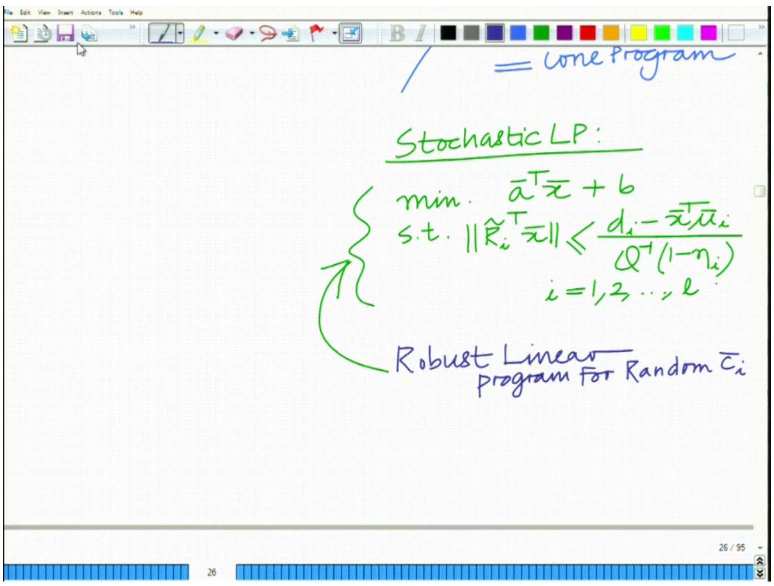
And therefore, the condition above this implies that d i minus x bar transpose mu bar i over square root of x bar transpose R i x bar which is nothing but norm R i tilde transpose x bar is greater than or equal to Q inverse 1 minus eta which basically implies that again Q inverse 1 minus eta is positive. So, you can write this as R i tilde transpose x bar is less than or equal to d i minus x bar transpose mu bar i divided by Q inverse of 1 minus eta ok.

(Refer Slide Time: 26:23)



And now if you look at this, you will have the something interesting. You have here a norm and here you have something that is affine. So, you have norm of vector x bar something all right, norm less than or equal to affine function of x bar, remember this represents the cone. So, this is the conic constraint ok. So, this represents a cone, and this is equal to a conic. So, this will become a cone program. So, this implies work with robust linear program that we are talking about, this becomes a cone program implies or the stochastic LP, you can call it the stochastic LP equals a cone program. It has conic constraints ok.

(Refer Slide Time: 27:29)



And therefore, now you can formulate this stochastic LP as follows that is basically minimise a bar transpose x bar plus b subject to the constraint that your norm R i tilde transpose x bar is less than or equal to d i minus x bar transpose mu bar i Q inverse 1 minus eta for i equal to 1, 2 to l. In fact, this can be eta i does not really matter, this can be eta i. I am just going to correct this over all places it can be a common eta or it can be eta i. You might want different reliabilities for different constraints alright.

So, this can be there you go. So, it is eta i and you can correct it everywhere to eta i; and this has to hold for i equal to 1, 2 up to l. So, this is your linear program. You can think of this as a robust version for the scenario when these vector C bar i are random in nature all right, so that is another interesting flavour.

In fact, I can say this is a practical flavour of the traditional linear program. The linear program itself has several practical applications, this in that sense it makes in more practical or immensely enhances its utility by making more relevant because several, in several scenarios these coefficients C bar i might not be known; might not be known accurately or might be known only with the certain degree of accuracy. In particular when they are random and they can be modelled as Gaussian random variables one can use this interesting framework to formulate the equivalent either stochastic or you can call it with the stochastic version of the robust stochastic LP or robust LP or the stochastic version of the robust. We will stop here.

Thank you very much.