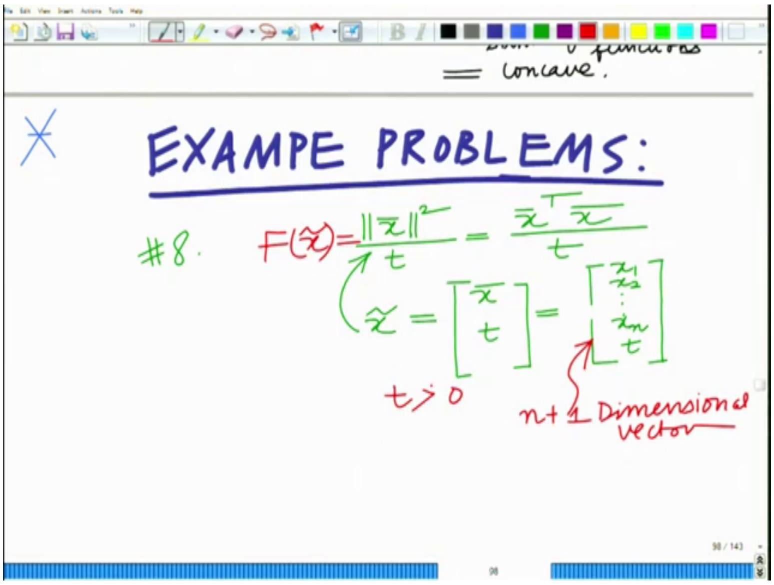
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 33
Example Problems: Perspective function, Product of Convex functions, Pointwise Maximum is Convex

Hello, welcome to another module in this massive open online course. So, we are looking at Example Problems for Convex Functions. Let us continue our discussion, alright.

(Refer Slide Time: 00:23)

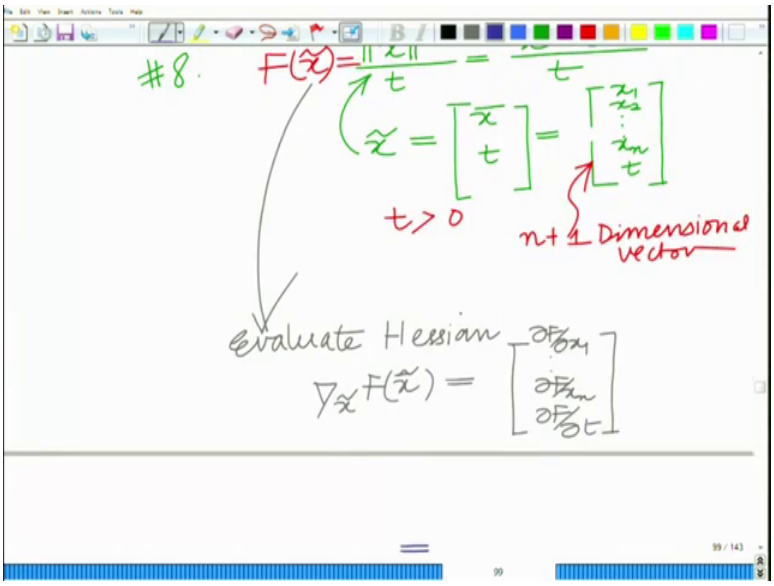


We are looking at example problems and let us look at problem number 8 and we want to consider the function norm x bar square by t which you can also write as a remember norm of a vector square is vector transpose time means itself, that is x bar transpose x bar divided by t you can think of this as an n plus 1 dimensional function.

So, we have if you define the vector x tilde equals well the vector x bar t augmented with t which is basically you think of this as x 1, x 2, up to x n and then one additional element. So, this is basically your n plus 1 dimensional vector ok. This is an n plus 1 dimensional vector we have this function F of x tilde ok. So, this is x tilde is the vector x bar that is augmented with t and we can consider further that t is greater than 0, that is t is a positive quantity.

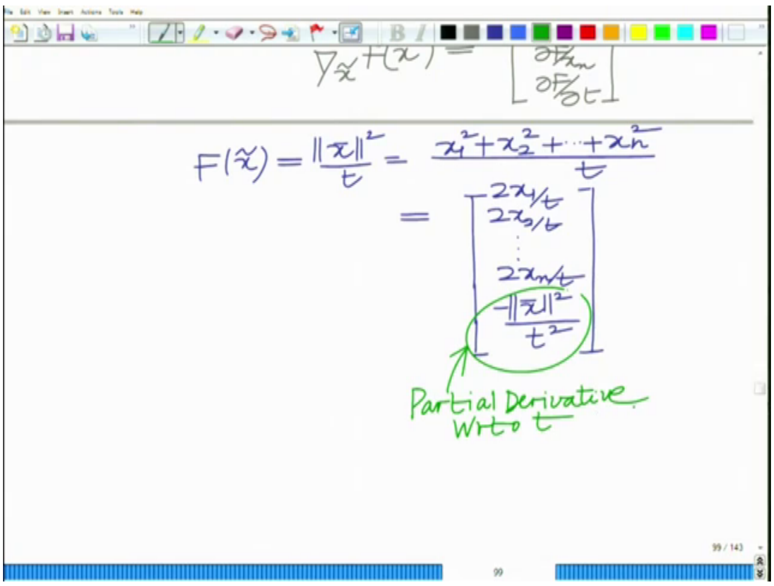
Now, we want to show that this function is indeed a convex function. We will follow the approach that we have shown before that is the test for convexity which is to evaluate the Hessian and demonstrate that it is indeed a positive semi definite matrix.

(Refer Slide Time: 02:19)



So, what we want to do is we want to evaluate the Hessian of this. First let us start with the gradient; gradient with respect to x tilde of F of x tilde which will contain first all the partials with respect to all the x's followed of course, by the partial with respect to t.

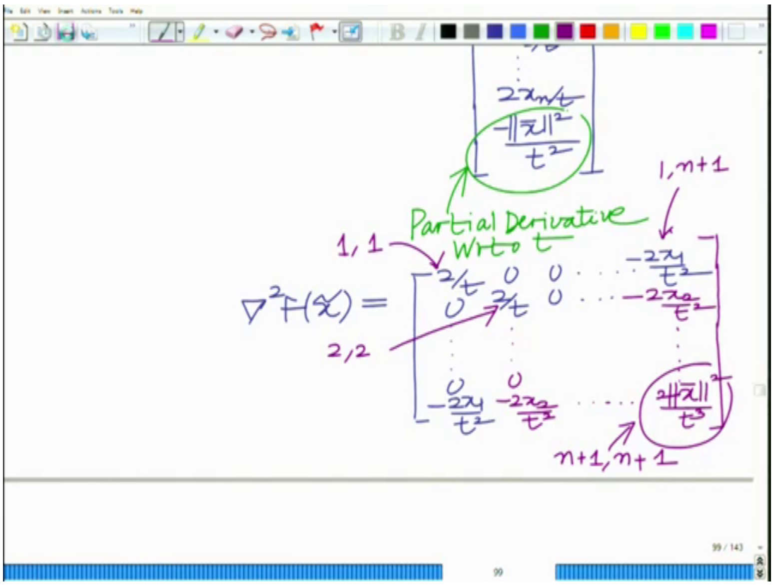
(Refer Slide Time: 03:05)



And, well, if you look at this well you can simplify this x bar transpose x bar so, F of x tilde norm x bar square by t which is also x 1 square plus x 2 square plus x n square divided by t you can see the partial with respect to x 1 is simply 2x 1. 2x 1 divided by t partial with respect to x 2 is 2x 2 divided by t partial with respect to x n is 2x n divided by t and the partial with respect to t is norm of x bar square into differentiate this with respect to 3. So, that will give us minus 1 over t square, ok.

So, the partial with respect to t is minus norm x bar square divided by t square. So, this is your partial with respect to partial derivative with respect to with respect to t. And, now, we have to compute the Hessian for this, ok. This is basically the gradient we have to compute the Hessian, ok.

(Refer Slide Time: 04:37)



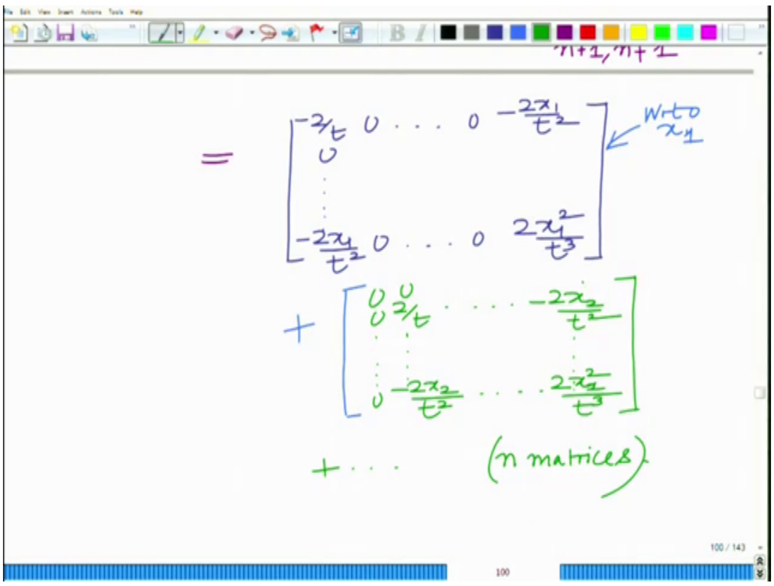
And, that is also it has an interesting structure it is fairly straightforward you just have to pay attention to each element. Now, first well the 1 cross 1 element is the partial with respect to x 1 that is a second order partial with respect to x 1 square. So, you take 2x 1 over t divide it with differentiate with respect to x 1. So, that gives you 2 over t. Now, you take 2x 1 over t differentiate with respect to x 2 that gives you 0 differentiate with respect to x 3 0. In fact, differentiate it with respect to. Now, the last element will be the derivative that is 1 comma n plus 1-th element will be the derivative with respect to t and that will be minus 2x 1 over t square.

Similarly, you have all these elements are 0 and the last element will be once again minus 2x 1 over t square. Now, look at the 2 cross 2 element that is the derivative partial second order partial with respect to x 2. So, we take 2x 2 or t differentiate with respect to x 2 that is 2 over t of course, rest of the elements will be 0 and this last element will be once again minus 2x 2 or t square this element here will be minus 2x 2 or t square you will have so on and so forth, so on and so forth and here the last element here will be partial with respect to second order partial with respect to t.

So, this will be minus norm x bar square derivative over 1 over t square that is 2 over or minus 2 over t cube. So, that will be 2 over t q. This is basically the Hessian. So, you can see this is basically you can see each element. So, this is basically the 1 cross 1 element this is basically the 2 cross 2 element this is your 1 I am sorry I should say 1 comma 1 element 1 comma 1 element. This is the 2 comma 2 element that is the diagonal element this is your 1 comma n plus 1 element and so on, and this last element here this is your n plus 1 comma n plus 1 element, ok.

So, this is a Hessian it has an interesting structure. So, the first n diagonal elements are all 2 over t the last n plus 1 comma n plus 1-th element is 2 norm x bar square divided by t cube. And, along the last row and the last column you have the elements which are of the form entries of the form minus 2 x i over t square. So, that is the structure of the Hessian.

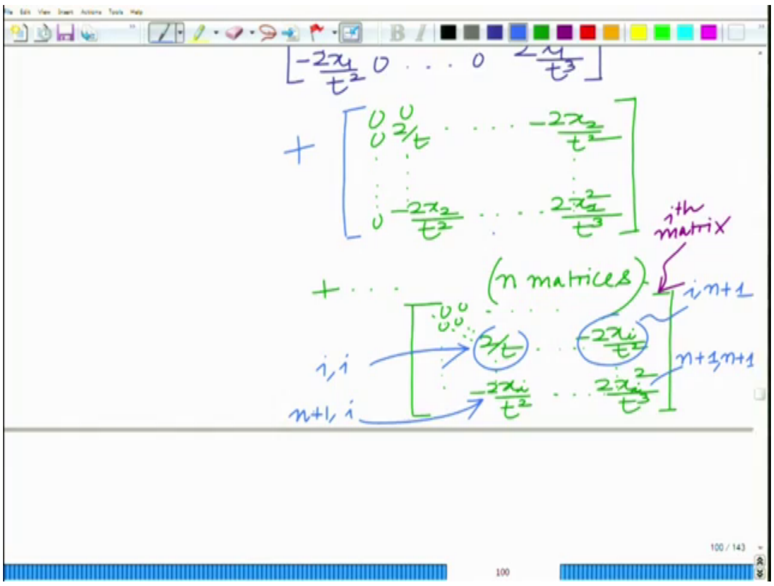
(Refer Slide Time: 07:59)



And, now you can divide this into the sum of several matrices each matrix with respect to one of the x i's. So, the first matrix will be with respect to x 1. So, you can take this and this has the particular structure. So, this will be 2 over t last element will be minus 2x 1 or t square 0 0 minus 2x 1 over t square and the last element out of the norm x bar square you simply take x 1 square. So, this will be 2x 1 square divided by t cube. So, this you can think of this as a corresponding to with respect to x 1 or corresponding to the x 1.

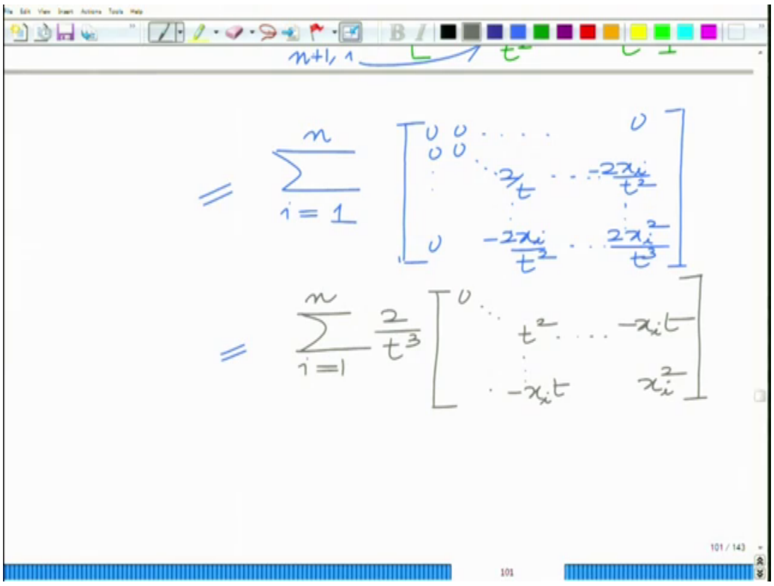
Similarly, corresponding to x 2 you will have a matrix which is of the form 2 cross 2 element is 2 over t the last element is minus 2x 2 the last element in the second row is minus 2x 2 over t square this is 0 n plus 1 comma 2. The last the second element in the last row that is again minus 2x 2 over t square and again you take n plus 1 comma n plus 1-th element you take the component corresponding to x 2. So, this will be 2x 2 square over t cube plus so on. So, total of n you have a total of n such matrices, ok.

(Refer Slide Time: 10:03)



Well, for the i-th matrix you will have if you look at the structure of the i-th matrix ok, the i-th matrix corresponding to x i we will have 2 over t in the i comma i-th element this will be minus 2x i over t square. This is n plus 1 comma i similarly here you will have minus 2x i over t square and n plus 1 comma n plus 1 element that will be 2x i square divided by t cube and the rest of all ok. So, this is the matrix you can think of this as the i-th matrix this is basically your i comma n plus one-th element this is your i comma i-th element, this is your n plus 1 comma i-th element and this is your n plus 1 comma n plus one-th element, ok. So, this is the structure you can decompose it into such matrices.

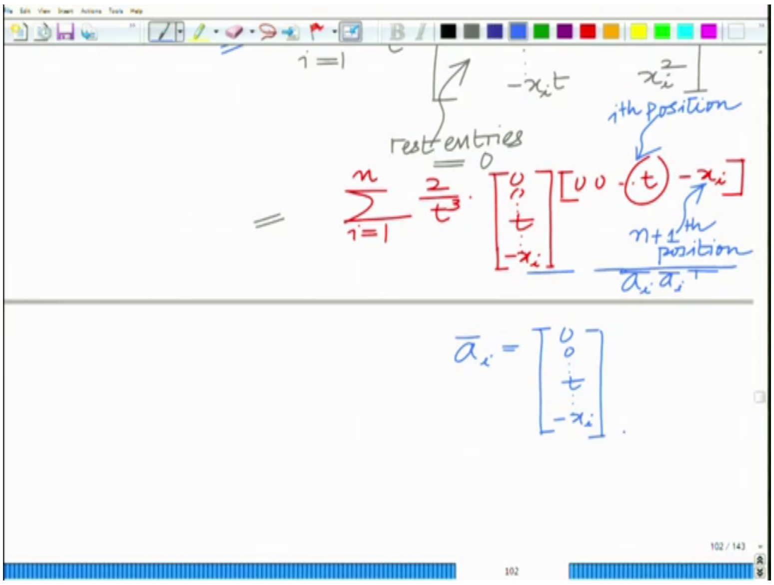
(Refer Slide Time: 11:35)



And, now so, I can write it as the summation. So, I can write this or rather let us now I can write this as the summation over i of such matrices i equals 1 to n you have minus 2x i or t square minus 2x i over t square and 2x i square over t cube, and the rest all the rest of the entries are zeros.

And, now, if I take 1 over that is 2 over t cube as common so, I can write this as i equals 1 to n take 2 over t cube as common this will become again very simple. So, if you take 2 or t cube common this will be t square this will be minus 2x i times t minus 2x i times t and this will be well, this will simply be I am sorry the 2 will go because we have taken the 2 outside. So, minus x i t and this will be simply x i square and rest of the entries are 0, rest of the entries are 0.

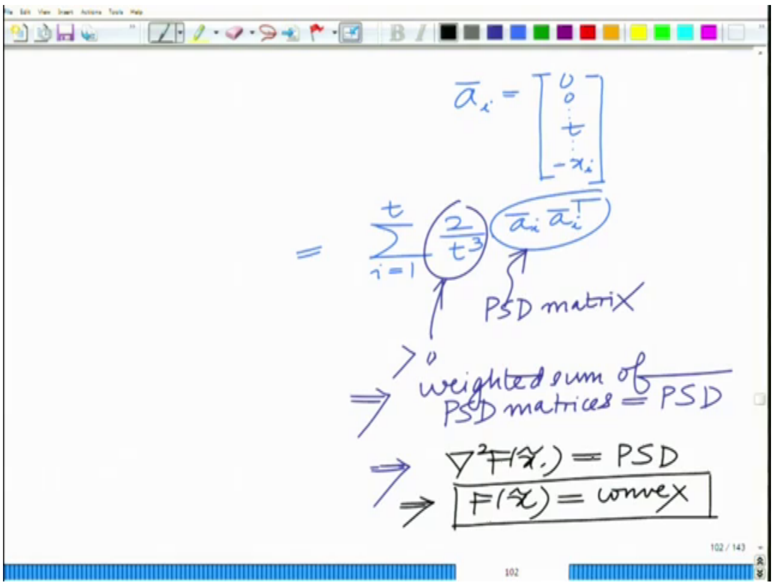
(Refer Slide Time: 13:25)



And, now you can see I can decompose this as i equal to 1 to n 2 over t cube I can write this as 0, 0 then you have a t in the i-th position minus x i in the n minus n plus one-th position times 0 0 t again you have 0's and minus x i in the nth position.

So, this t is in the i-th a position this minus x i is in the n plus one-th n plus one-th position and now, you can see I am writing this basically decomposing this as a i bar a i bar transpose, ok. So, a i bar is this vector a i bar is basically your vector which has t in the i-th position and minus x i in the in the in the n plus 1-th position.

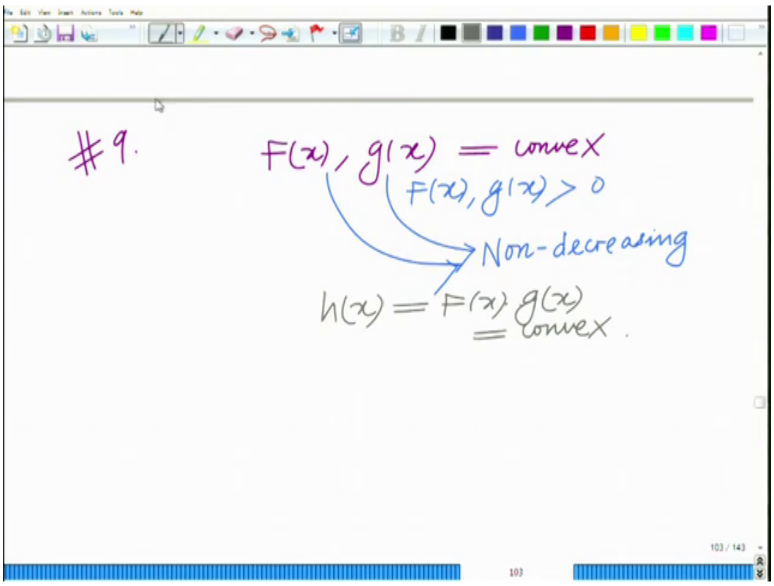
(Refer Slide Time: 15:03)



And therefore, I can write this as summation i equals 1 to t 2 over t cube a bar a bar transpose. Now, each of this is a positive semi definite matrix. Remember, whenever a matrix can be decomposed as a transpose correct, it is a positive semi definite matrix that is what we have seen you are weighing it by a positive coefficient because remember t is greater than 0. So, 2 over t cube is greater than 0. So, sum of positive semi definite matrices weighted by positive coefficients, the resulting resultant matrix is also positive semi definite. Therefore, the Hessian is positive semi definite and hence the function is convex ok.

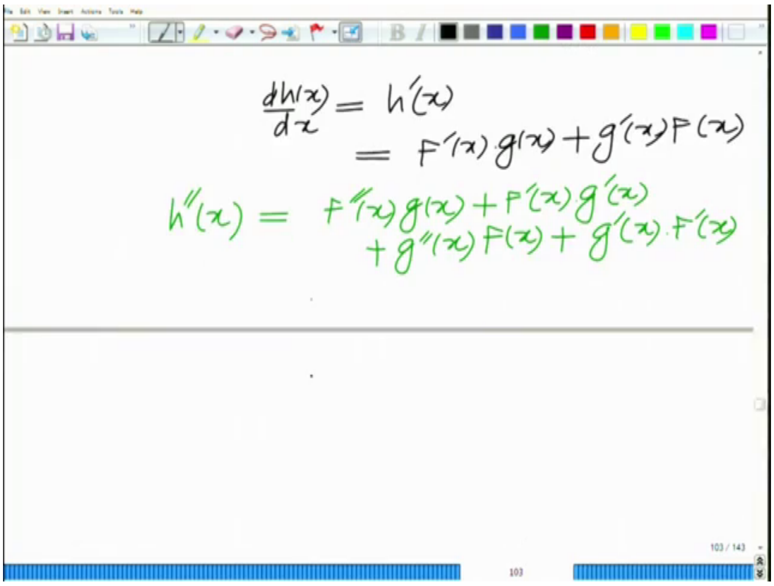
So, this is greater than 0 implies the weighted sum of PSD matrices is PSD which implies that your the Hessian F of x tilde equals is a PSD matrix which implies F of x tilde is indeed therefore, x tilde is indeed a convex function, alright. So, that is tells that is what. It is a slightly it is a slightly involved and lengthy proof, but as we have seen some of these tend to be a bit involved, ok. So, we have demonstrated that norm x bar square divided by t that is x bar transpose x bar divided by t considering this as a function of the n plus 1 dimensional vector x bar augmented with t this is a convex function. Let us proceed to the next problem that is problem number 9.

(Refer Slide Time: 17:17)



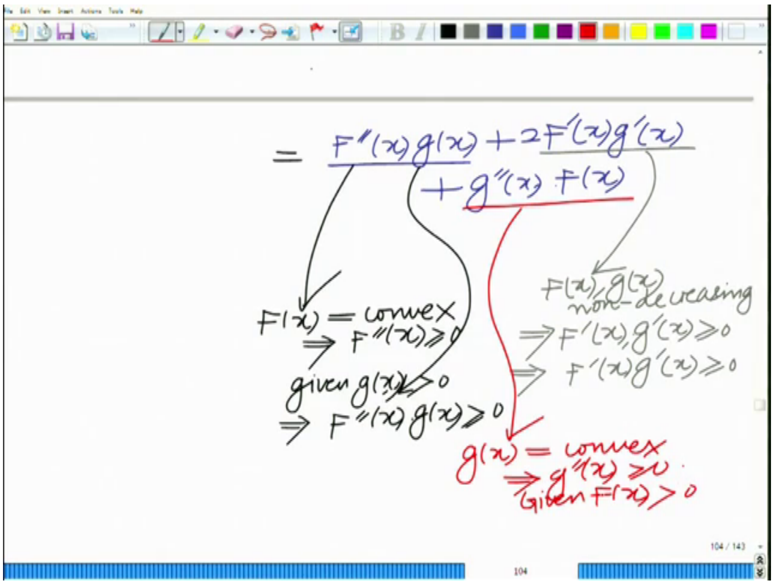
Given two functions F of x g of x which are convex and these are greater than 0, that is F of x comma g of x greater than 0 and further these are non-decreasing. We want to show that h of x equals F of x into g of x is also a convex function, and that is easy to show.

(Refer Slide Time: 18:17)



So, we consider first the first order derivative of h of x which we denote by h prime of x using the product rule that is F prime of x g of x plus g prime of x into F of x. Now, considering the second order derivative it is h double prime of x which is F prime of F double prime of x g of x plus F prime of x g prime of x plus plus plus well, g double prime of x F, I am sorry this has to be g prime of x g prime of x F x plus g prime of x into F prime of x which you can now simplify as follows.

(Refer Slide Time: 19:31)

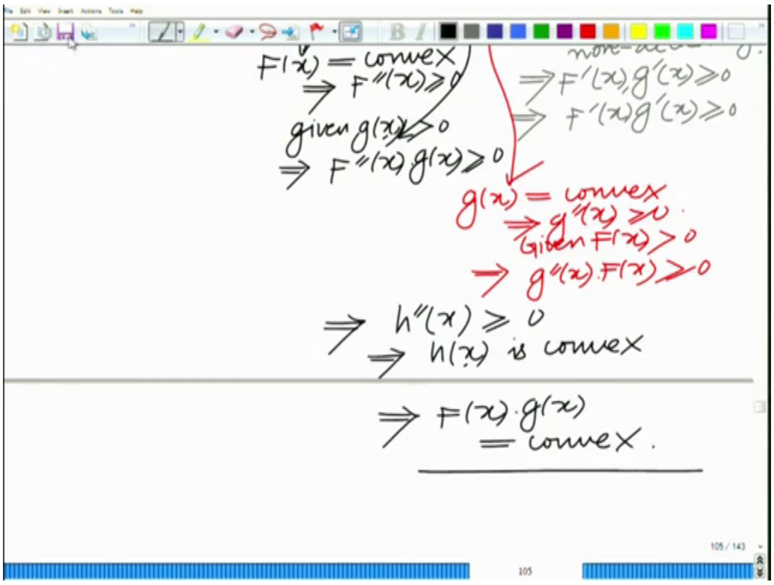


You can write this as this is equal to, well F double prime of x g of x plus twice you can combine these two terms. So, that gives you twice F prime of x g prime of x plus g double prime of x F of x. Now, let us dissect this term by term if you look at this quantity here you can see F of x is convex. So, this implies F double prime of x is greater than equal to 0. Now, g of x is given to be greater than or equal to 0 or greater than 0.

So, this implies F double prime of x into g of x is greater than or equal to 0. Now, F of x and g of x are non-decreasing are non decreasing this implies F prime of x comma g prime of x greater than equal to 0 this implies the product that is since they are non decreasing both of them are non-negative. So, the product is also non-negative and now the last term is similar g is convex. This implies g double prime of x is greater than equal to 0 given F of x greater than 0. So, this implies g double prime x into F of x greater than equal to 0.

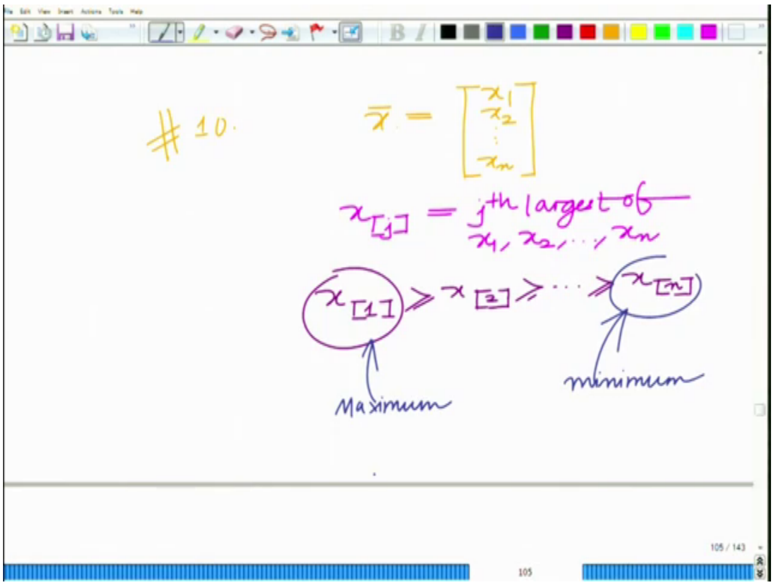
So, all the three components in the some are non-negative. Therefore, the sum is non-negative which implies the second order derivative h double prime x is non-negative or it is basically greater than equal to 0, which implies essentially that h of x is convex or the product F of x into g of x is convex.

(Refer Slide Time: 21:33)



So, this implies h prime of x greater than equal to 0, this implies h of x is convex, which is nothing, but F of x into g of x. This implies that F of x into g of x is convex. Let us now move on to another problem number 10.

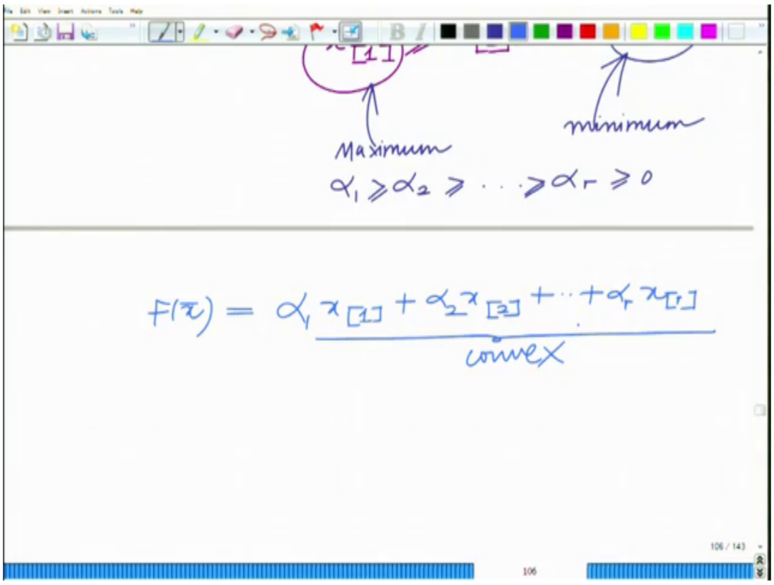
(Refer Slide Time: 22:33)



Now, consider a set of variables x 1, x 2, x n or a vector or let us say we have x bar this is an n dimensional vector, and we have x j is the j-th largest of x 1, x 1, x 2 up to x n which implies that basically you are sorting this. So, what this means is the largest is x 1 which is greater than if you sort this x 2 greater than or equal to x n; n with the square bracket remember this is actually different from x n.

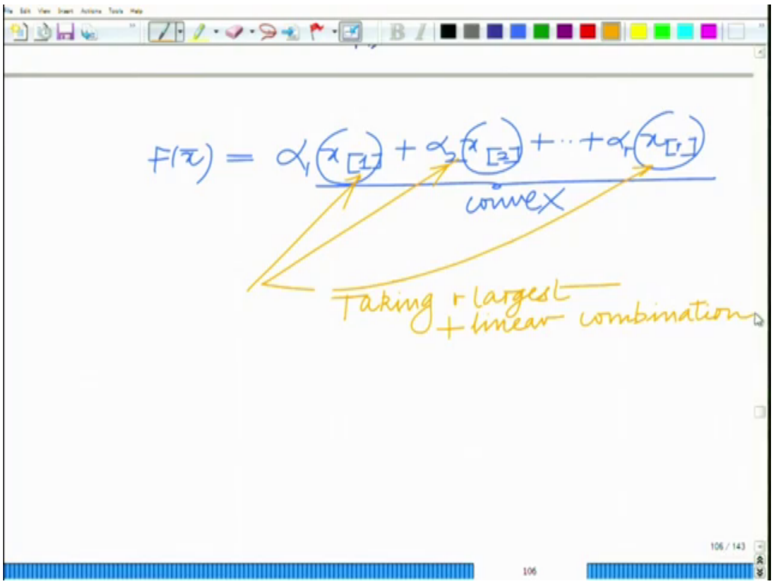
So, x 1, x square bracket or x subscript bracket you can say x subscript square bracket one this is the maximum or the largest and this is the this is the largest and this is the minimum, and x subscript square bracket j that is the j largest j-th largest that is you arrange them in descending order x subscript square bracket one is the largest followed by x subscript square bracket 2 and so on.

(Refer Slide Time: 24:35)



And, let us assume the non negative coefficients alpha 1 which are again arranged in decreasing order such that alpha 1 greater than alpha 2 greater than alpha r greater than is equal to 0. Now, what we want to show that this function of x bar which is alpha 1 x 1 plus alpha r x r, we want to show that this is convex and this can be seen as follows.

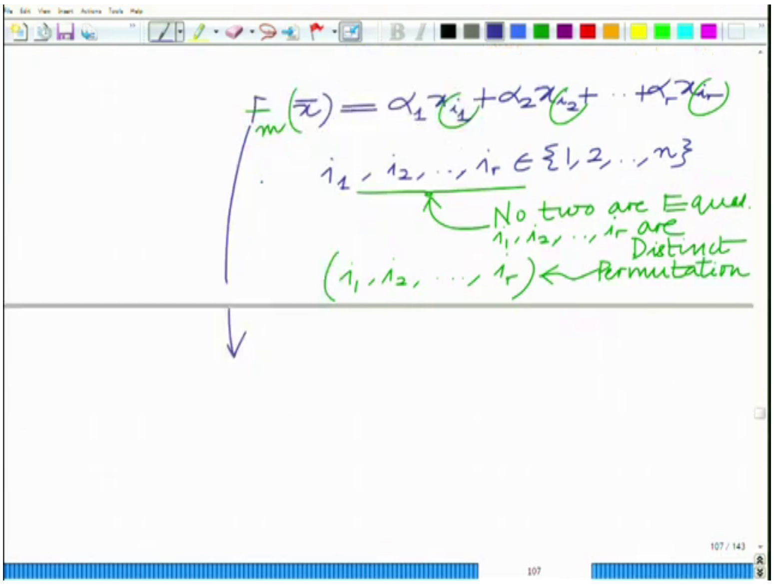
(Refer Slide Time: 25:33)



Now, this is an interesting function first you can see what you are doing is you are taking the r largest alright. This is very interesting and very complicated function. So, you are taking r largest plus you are taking the linear combination and this is a highly non-linear function because, when you look at the maximum, the maximum of the maximum is basically a non-linear function correct.

So, we are taking the r largest getting a non-linear getting a linear combination. So, this is basically it is a highly non-linear function because although you are taking a linear combination you are looking at the maximum of these elements, right. So, this is a highly non-linear interesting function yet you can demonstrate that this function is convex and that can be done as follows. In fact, it is very simple.

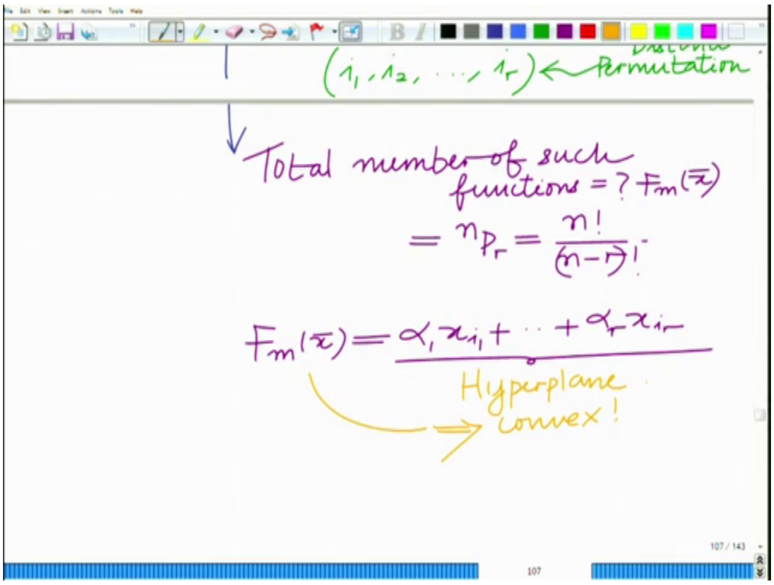
(Refer Slide Time: 26:35)



If you consider the function F i of x bar which is defined as alpha 1 x i 1 plus alpha 2 x i 2 index i 2 plus alpha r x i r where i 1 i r these belong to the set 1, 2 up to n and none of these two are equal or no two of these are equal or all of them are distinct no two are equal. So, which implies i 1 i 2 up to i r are distinct now how many ways can you choose these indices i 1, i 2, i r.

In fact, I can call this as F of let us say some index not I because we are using well F of let us say m just one particular combination. So, basically depends on total number of not even combinations because remember for I alpha 1 you have to choose one index i 1, alpha 2 you have to this is basically a problem of permutations, how many ways that is basically you are choosing the ordered pairs i 1 i 2 up to i r. So, this is basically the permutation and what is the total number of such you want to ask the questions what is the total number of such functions.

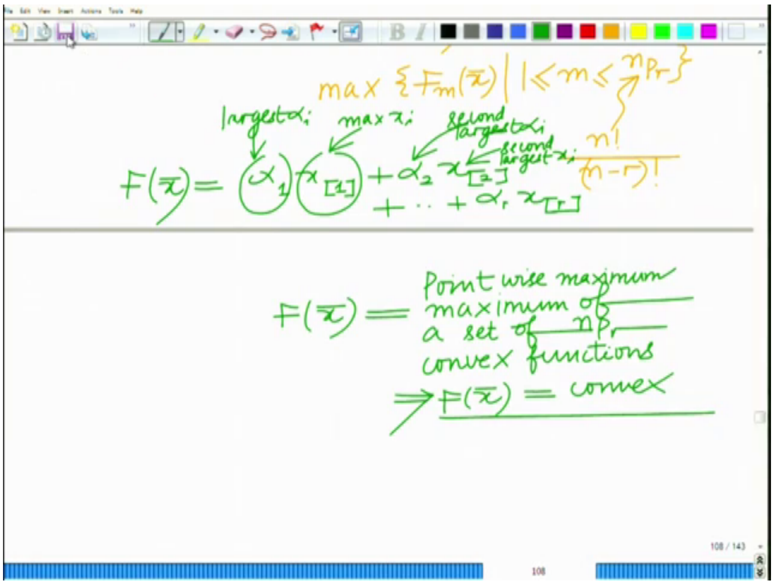
(Refer Slide Time: 28:59)



What is the total number of such functions that is the total number of permutations of r objects from a set of n objects that is n P r. So, total number of permutations will be n P r. You might have seen this is in this is in high school this is n factorial by n minus r factorial, ok. So, that is the total number of such functions F total number of such functions F m of x bar.

Now, you can see that each of these is a hyper plane each of these is a that is F m of x bar equals alpha 1 x i 1 plus alpha r x i r each of these very interestingly this is a which implies this is convex. So, each of these functions is a hyper plane each of this functions corresponding to a permutation of this r x r variables x i 1, x i 2, x r this is a hyper plane. So, each such function is convex.

(Refer Slide Time: 30:55)



And, now, therefore, if I take the maximum of this for all the you take the maximum of this for each x bar for all the m less than n P r that is you take the maximum of all these n P r or n factorial by n minus r factorial functions you can see that the maximum is nothing, but alpha 1 times x of x subscript square bracket 1 alpha 2 x subscript square bracket 2 plus so on alpha, r x subscript square bracket. That is a maximum occurs, remember we have said these coefficients alphas are arranged in the decreasing order so, alpha 1 is greater than equal to alpha 2 is greater than equal to alpha 3 greater than equals so on up to alpha r.

So, maximum occurs when the maximum alpha 1 is associated with the largest. So, this is the largest alpha this is the maximum x i, this is the second largest and second largest alpha i, this is the second largest x i and so on. Therefore, the maximum is nothing, but alpha 1 x substitute square bracket 1 so on summation alpha r x subscript square bracket r and this is nothing, but our F of x bar.

Now, you can see that F of x bar is the maximum of set of in fact, this is point wise maximum that is for each x you are taking the point wise maximum of a set of n P r convex functions, in fact, hyper planes implies that F of x bar is convex alright. So, basically you are taking n P r that is n factorial by n minus r factorial convex functions or hyper planes and you are take the maximum the point wise maximum of these n P r hyper planes and therefore, the resulting function is also indeed convex and that basically completes the proof for this interesting problem alright.

So, we will stop here and starting from the next module we will start looking at various convex optimization problems, the practical applications in various domains.

Thank you very much.