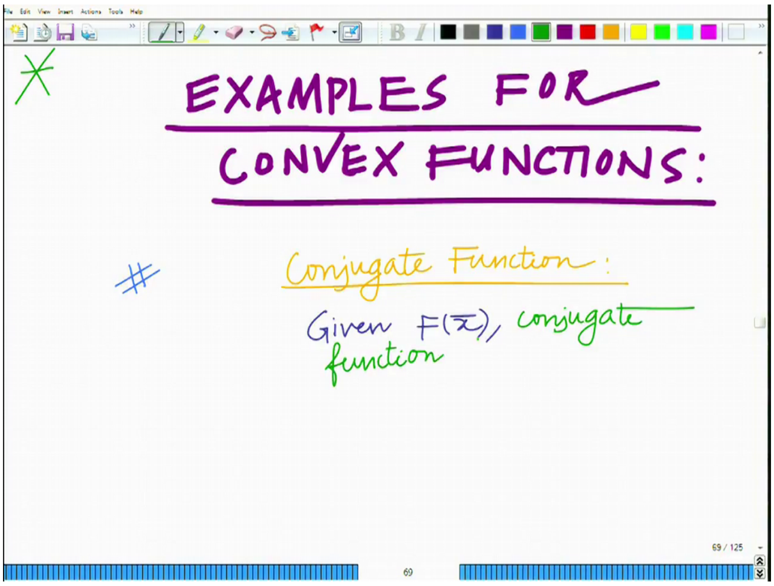
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture - 30
Conjugate Function and Examples to prove Convexity of various Functions

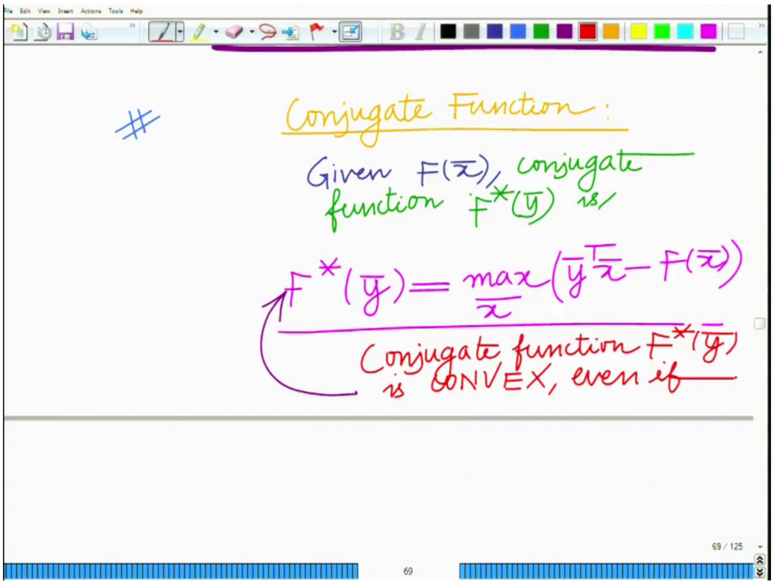
Hello, welcome to another module in this massive open online course. So, we are looking at convex functions, we have completed the basic discussion on the basic aspects including illustration of several practical applications. Let us now focus on some examples to understand these concepts better all right.

(Refer Slide Time: 00:29)



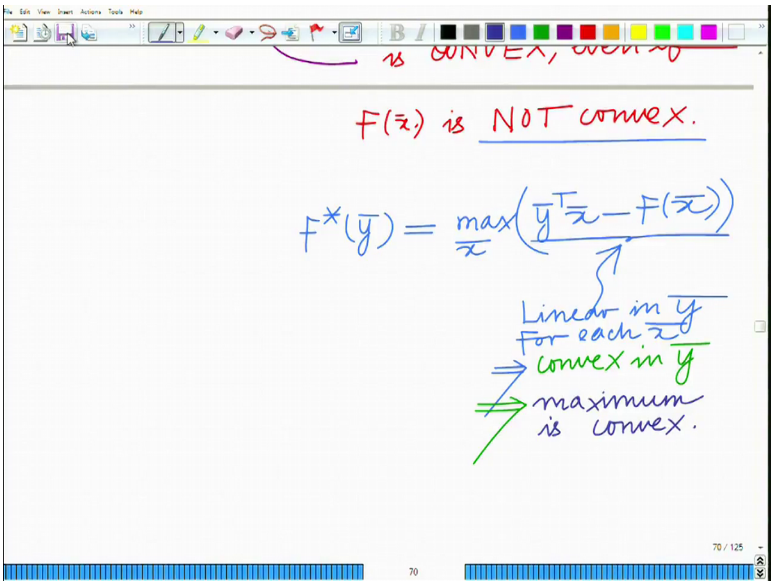
So, what we want to start looking at is examples for convex functions. The first example we want to look at a new concept which is that of a conjugate that of a conjugate. This is a very interesting concept for the following reason which I am going to describe shortly, but the definitions, so given F of given a function, given a function F of x bar the conjugate denoted by the conjugate function F conjugate y bar is this is given as the maximum over the vectors x bar of y bar transpose x bar minus F of x bar; it has a rather very interesting definition.

(Refer Slide Time: 01:54)



And so this is the definition the conjugate function and the interesting aspect of this is this conjugate function is convex even when F of x, the original function is not convex, this is important aspect. So, the important aspect of this is that the conjugate function F conjugate y bar is convex.

(Refer Slide Time: 03:10)

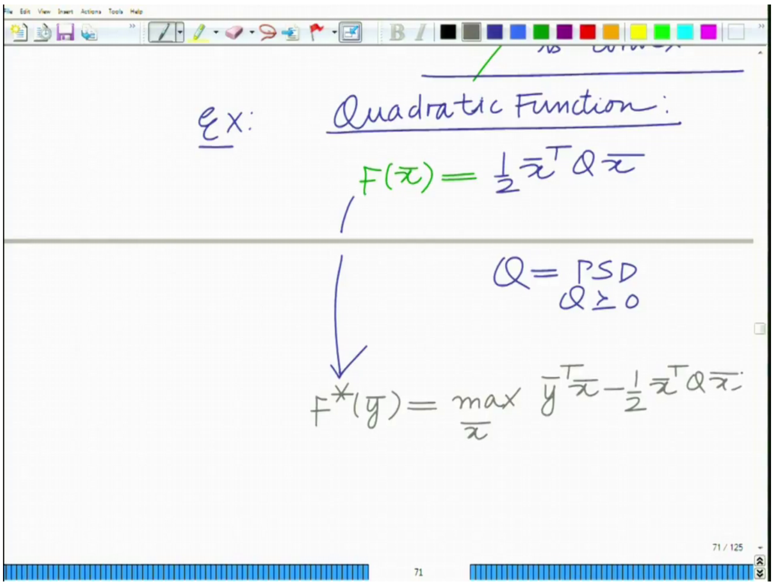


Even if F of x bar is not even is F of x bar is not convex, so corresponding to every function any function F of x bar convex or non convex, one can construct an associated convex function which is a conjugate function and that is very easy to see.

For instance, if you look at F conjugative of y bar let me just write this again which is y bar transpose x bar minus F of x bar. Well you can see that this is a linear, linear in y bar for each x; for each value of x bar this is a linear function in y bar.

So, you are taking which is basically which implies that this is a convex in y bar. And therefore, we need to taking the set anyway you take the maximum, you are taking the maximum of your set of convex functions and therefore, from the property that we have discussed for convex functions, the maximum of a set of convex functions is convex, therefore, you are taking the set the maximum of the set of convex functions that is one function for each x bar all right which is convex in y bar for each x bar, you are taking the maximum and therefore, the resulting function is convex, implies the maximum, implies the maximum implies the maximum is convex ok.

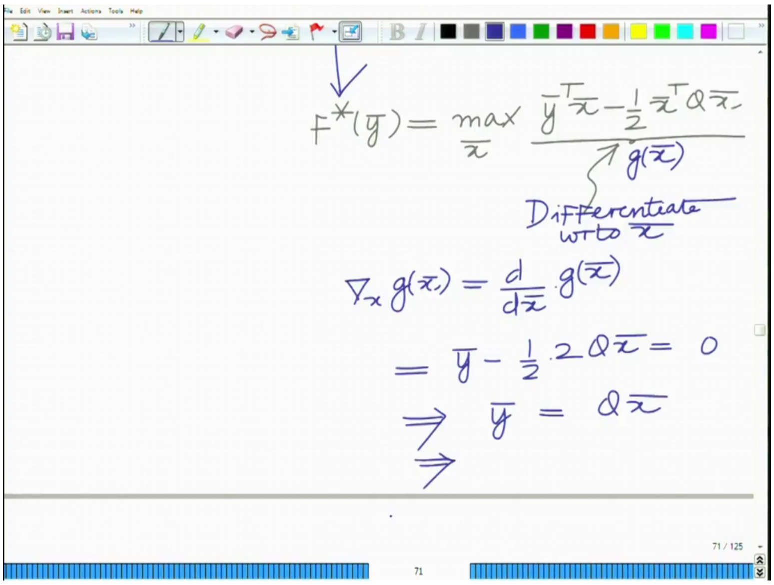
(Refer Slide Time: 05:11)



And therefore, it is convex even if the original function F is not convex ok. Let us take an example of this conjugate function. Consider the quadratic, consider the quadratic function which is F of x bar equals half x bar transpose Q x bar with Q equals symmetric positive semi definite ok, which we have also written as follows Q greater than equal to 0.

Now, if you take now to conjugate the to construct the conjugate function, we have F conjugate y bar equals maximum of over x bar y bar transpose x bar minus F of x bar which has half x bar transpose.

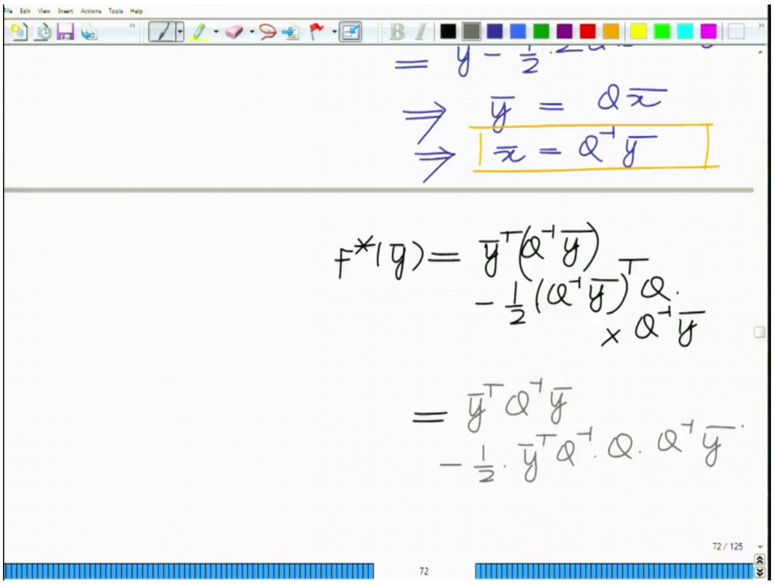
(Refer Slide Time: 06:35)



And now to maximize this let us differentiate this ok, to maximize this differentiate and set it equal to 0. We are going to differentiate with respect to x bar; that is considered its gradient of this quantity ok.

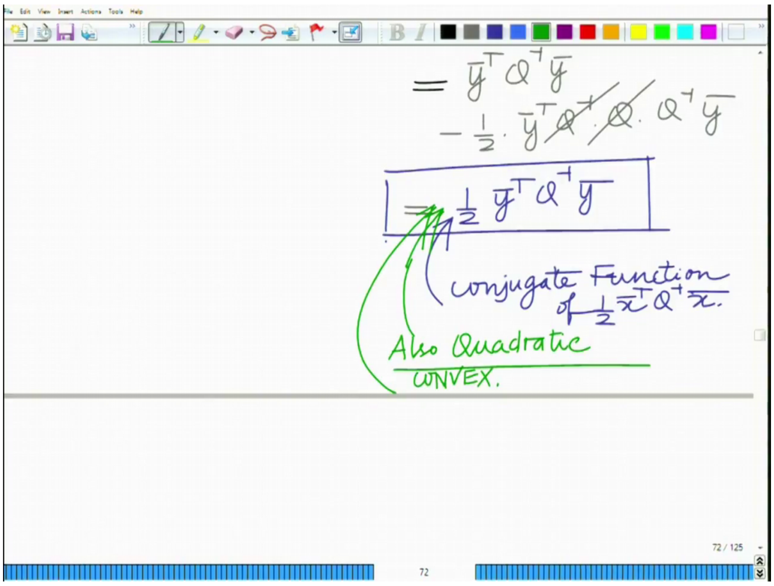
Let us call this as g of x bar, which has differentiated g derivative of g with respect to the vector x bar you and remember derivative of y bar transpose x bar, it is of the form c bar transpose x bar derivative F c bar which is basically y bar minus half derivative of x bar transpose 2 x bar; remember we said is twice derivative of x bar transpose p x bar is twice p x bar. So, here we have twice Q x bar, which is equal to 0, you are setting the derivative equal to 0; this implies y bar equals Q x bar; this implies x bar equals Q inverse x bar equals Q inverse y bar ok.

(Refer Slide Time: 07:57)



And the conjugate function F conjugate y equals y bar transpose Q inverse y bar minus that is y bar transpose into x bar minus half, x bar transpose Q inverse x bar which is nothing but basically Q inverse is of x bar and substituting y bar Q inverse y bar transpose Q times Q inverse y bar, which is equal to basically y bar transpose Q inverse y bar minus half y bar transpose Q inverse into Q into Q inverse y bar.

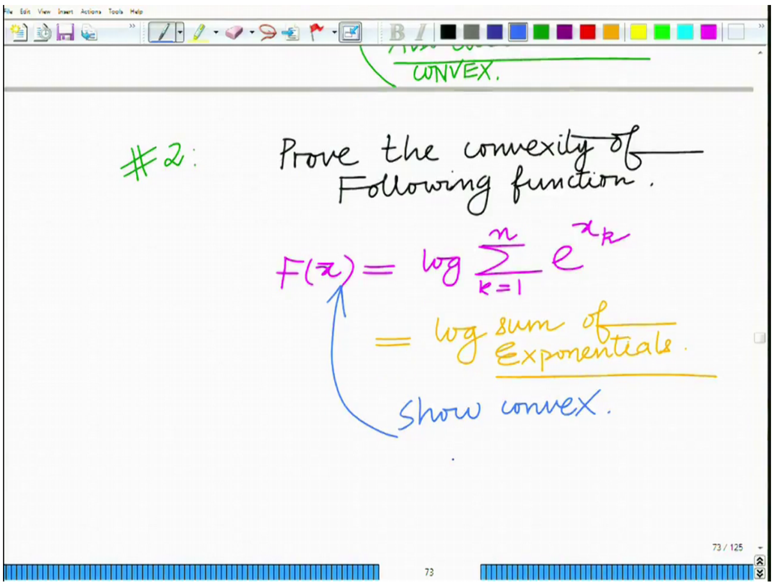
(Refer Slide Time: 09:17)



So, Q inverse and Q cancels, so you again have minus half y bar transpose Q inverse y bar. So, that is basically half, y bar transpose Q inverse y bar which is basically the conjugate function conjugate function of.

And you can see this conjugate function is also quadratic is also quadratic and in fact, you can say this is convex, because Q is positive semi definite which implies Q inverse is also positive semi definite remember that is an important result property of positive semi definite matrices, if that is the metrics is positive semi definite, the inverses is also positive semi definite. And now you have a quadratic function with a positive semi definite metrics and therefore, that is Q inverse and therefore, that is also a convex, it is a convex function that half y bar transpose y bar Q inverse y bar is a convex function.

(Refer Slide Time: 10:49)

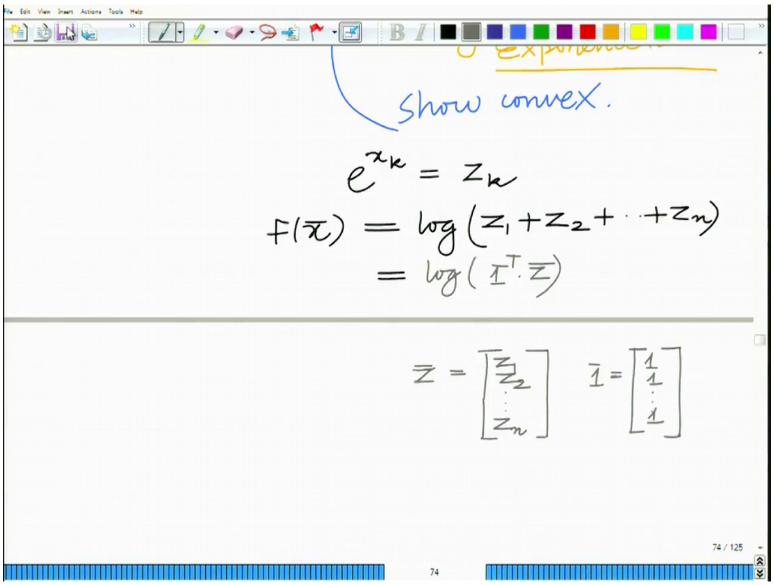


And so therefore, this is convex and this the conjugate function of x bar transpose Q inverse x bar. So, this is something that we have. So, one can readily derive the conjugate function ok.

Let us now look at another example slightly complicated one which we want to prove the convexity, prove the convexity of you want to prove the convexity of the following function, where F of x bar equals log k equal to 1 to n e raise to summation e raise to; so, this is log some of exponentials; so also referred to as the log of sum of exponentials.

So, very interesting function that arises in several logs, sum of this is the log sum of exponentials, which rises in several applications. What we want to show that show that this is convex, show that this function is convex.

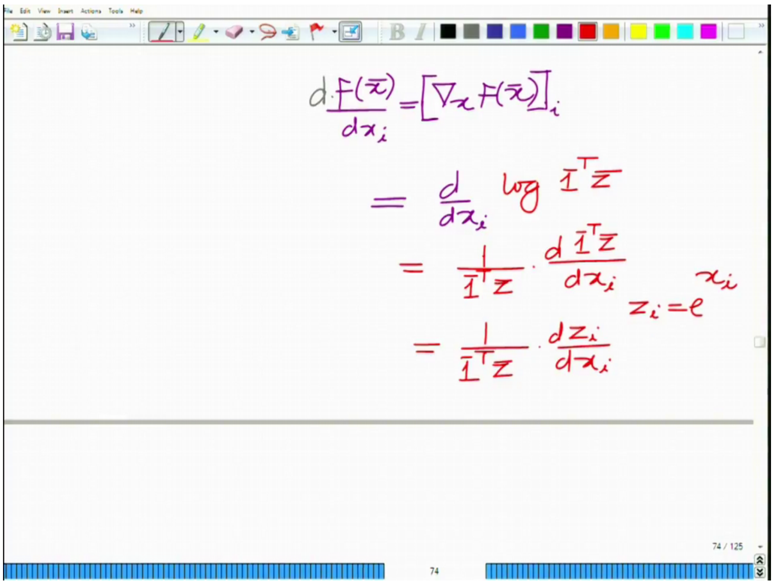
(Refer Slide Time: 12:30)



And for this what we will do is will compute the Hessian all right; so remember that has to be convexity, whenever its differentiable is yet you can compute the Hessian and demonstrate that the Hessian is a positive semi definite matrix. And that is what we want to do in this problem, I want to compute the Hessian and demonstrate that it is indeed a positive semi definite matrix.

So, all right the process is slightly involved, so, I urge you to be patient ok. So, first let us set for convenience e raise to x k; let us say z this equal to z k. So, I have F of x bar equals log sum of exponentials that is log that is basically you can see this is simply log of z 1 plus z 2 plus z n which is log I can write this as 1 bar transpose the vector z bar, where z bar is the vector is z 1 z 2 up to z n and 1 bar is basically the vector of all 1's. So, we are taking simply the when you are multiplying one bar transpose z bar, you are take simply taking the sum of all elements of z bar ok.

(Refer Slide Time: 13:56)

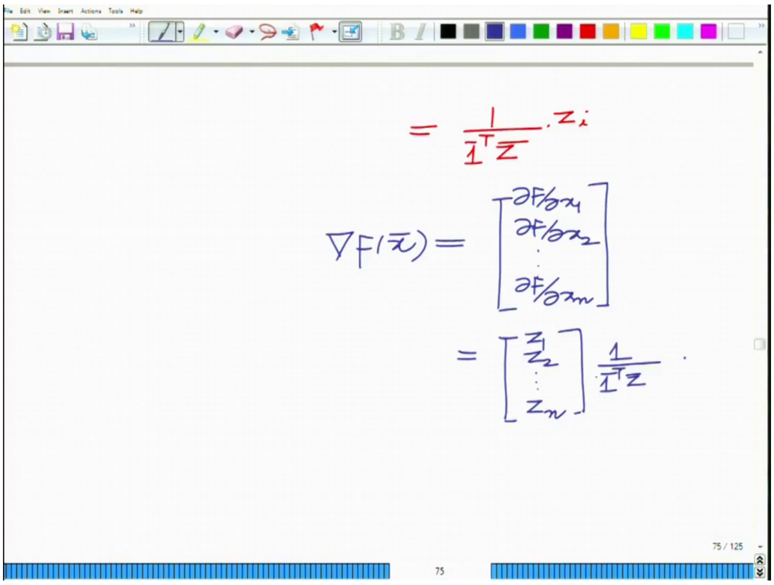


And now, if you differentiate, that is if you differentiate this with respect to each x i all right, that is basically you are computing the gradient with respect to x of F of x bar or rather the ith element, if you are looking at the gradient, you are looking at the ith element of the gradient ok. Now, that would be d by d x or rather d by d x derivative with respect to i of your log 1 bar transpose z bar. The derivative of log is log x is 1 over; so 1 over 1 bar transpose, times the derivative of 1 bar transpose z bar with respect to each x i.

Now in z bar, you can see only z i which is e raise to x i depends on x i. So, this is 1 over 1 bar transpose z bar into derivative of z i with respect to x i, but recall z i is e raise to x i so, e raise to x i.

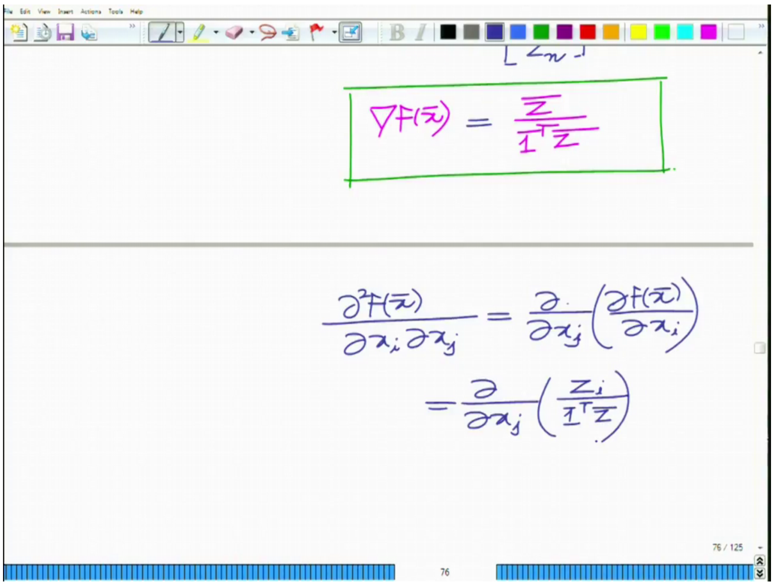
So, derivative with respect to x i is e raise to x i only which is basically z I, so this reduces to 1 over 1 bar transpose z bar times z i ok.

(Refer Slide Time: 15:36)



So, you have if you look at the gradient; which is basically the derivative of F of x with respect to each component of derivative of F of x with respect to each component that is equal to well z 1 z 2 up to z n 1 divided by 1 bar transpose z bar, which is basically you can see this is z bar transpose z bar over 1 bar transpose z bar, that is the gradient of z bar over 1 bar transpose z bar that is the gradient.

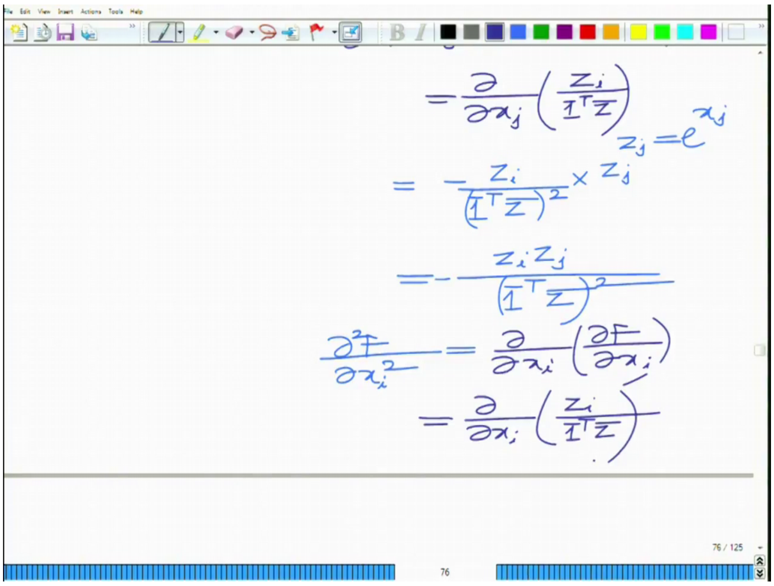
(Refer Slide Time: 16:35)



Now let us compute the Hessian, Hessian means the second order derivative. So, now, for the Hessian first let us look at the derivative with respect to each component x j; that is we have compute the second order components of the form derivative with partial with respect to x i, partial with respect to x j.

This is nothing but partial with respect to x j of partial of F of x bar with respect to x I, which is partial with respect to x j of partial with respect to x j of well, what is this quantity? This is we note the partial with respect to x i that is your z i divided by 1 bar transpose z bar.

(Refer Slide Time: 17:58)

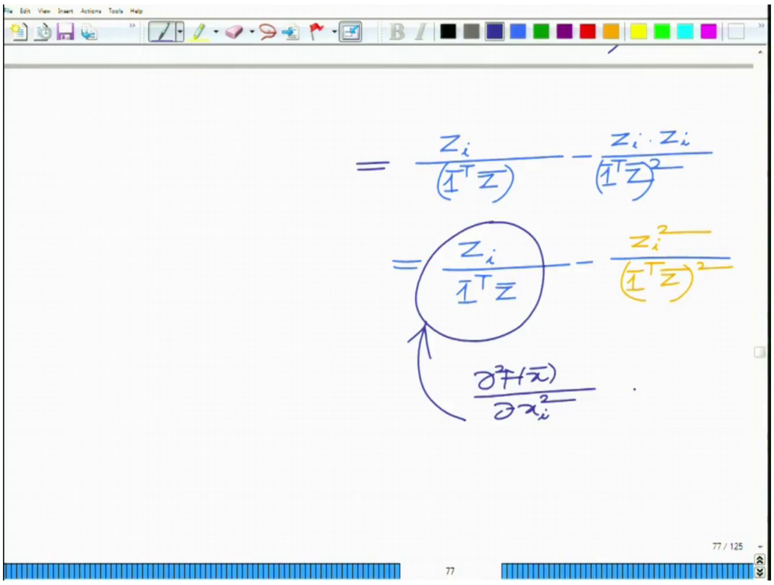


Well partial with respect to x j of z i is 0; so, this is simply you can write this as minus 1 over 1 bar transpose z bar square into partial of 1 bar transpose z bar with respect x j that leads on this x j z j, which depends on x j.

So, that is when you differentiate z j that is remember z j equals e raise to x j, when you differentiate with respect to x j you are left with e raise to x j which is again z j. So, this is this quantity component is minus z i z j by 1 over transpose z bar square ok.

Now on the other hand if you compute the second order derivative with respect to x i itself, that is terms of the form partial with second order partial with respect to x i that is dou square F by dou x i square that would be second order partial with respect to x i of the first order partial with respect to x I, which is dou by dou i z i divided by 1 bar transpose z bar.

(Refer Slide Time: 19:25)

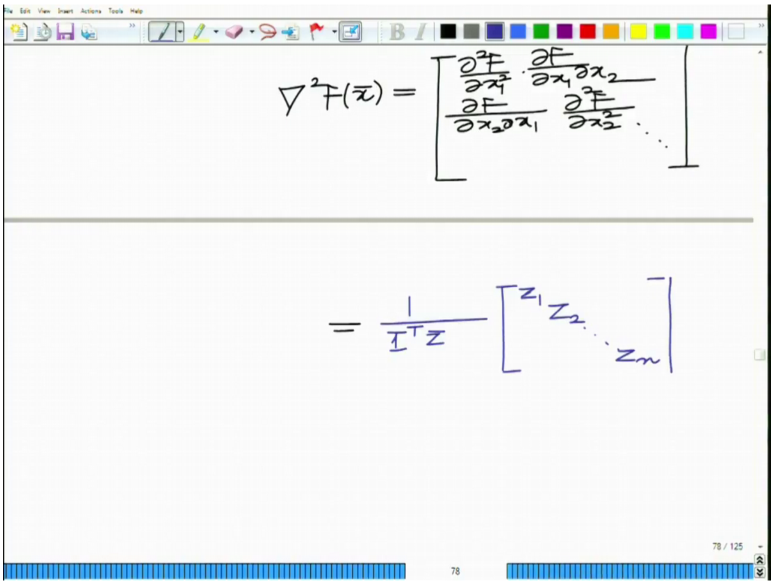


Now, if you differentiate the numerator we already see in derivative of z i with respect x i is z i itself. So, that will be z i divided by well that will be z i divided by 1 bar transpose z bar minus, now derivative of the numerator. So, 1 bar derivative of the denominator that is minus 1 over 1 bar transpose z bar square z i times the derivative of z i with respect to x i or 1 bar transpose z bar with respect to x i which is again z i.

So, this is you have a z i square in the numerator. So you will have z i divided by 1 bar transposes z bar minus z i square divided by 1 bar transfer z bar whole square, ok.

And now, if I will therefore, so only, so these terms are only present in terms of the form dou square F x bar dou x i square all right. They are not present if you can look at it they are not present, when you are in the second order terms of the form partial with respect to x i x j ok.

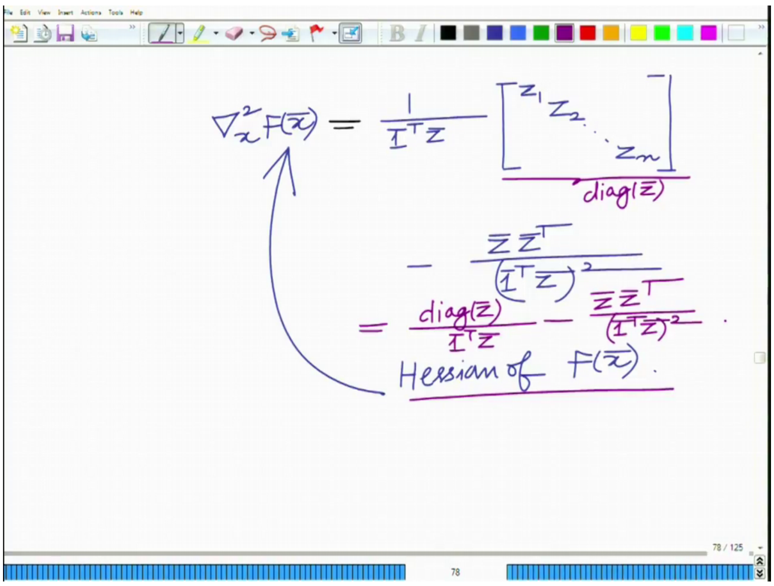
(Refer Slide Time: 21:03)



So, when you write you can readily see that if you look at the Hessian that will be well what is the hessian? Hessian is partial of F with respect to x 1 square partial of F with respect to x 1 x 2 partial of x with respect to x 2, partial with respect to x 1 and you have the partial with respect to dou square F by dou x 2 square and so on.

And if you put all these terms together what you will see is you will see that this is given as 1 over 1 bar transpose z bar, times the diagonal terms, remember the diagonal terms of the forms z 1 z 2 which are only there for the partial, that is terms of the from dou square F by dou x i square.

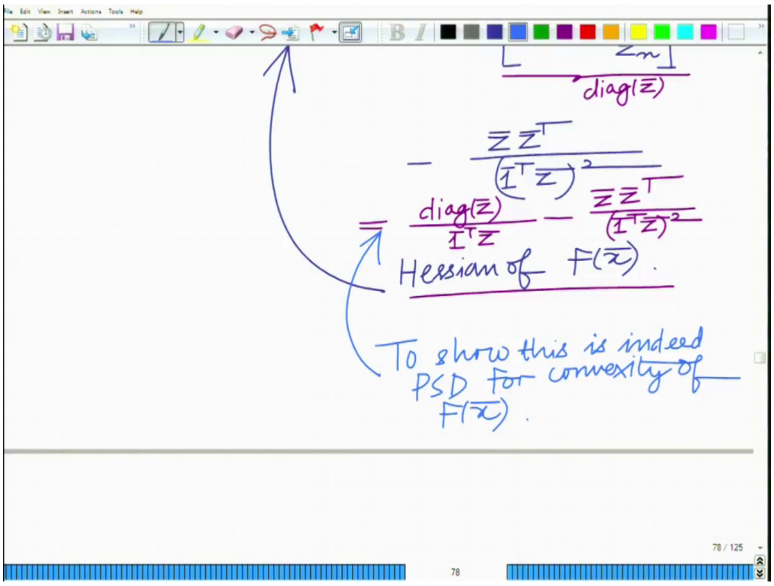
(Refer Slide Time: 22:13)



So, z 1 z 2 up to z n minus terms of the forms z i z j, which can be represented as z bar z bar transpose divided by 1 bar transpose z bar square. So, this is the hessian, Hessian of F of x bar; this is the Hessian of the F x bar.

Now we have to demonstrate to show that this is a positive semi; so I can also write this as just one last example; this is a diagonal metrics; diagonal metrics that contains z bar, the vector z bar on its diagonal. I can simply write this as diagonal z bar. So, this is diagonal z bar divided by 1 bar transpose z bar minus z bar z bar transpose divided by 1 bar transpose z bar whole square all right.

(Refer Slide Time: 23:24)



This is a Hessian and we have to show to show that this is positive semi definite, to show this is indeed PSD, for convexity of F of x bar.

That is we want to show that is this is indeed a positive semi definite metrics to verify the complexity of this function F of x all right, which we will do in the subsequent module; so, we will stop here.

Thank you very much.