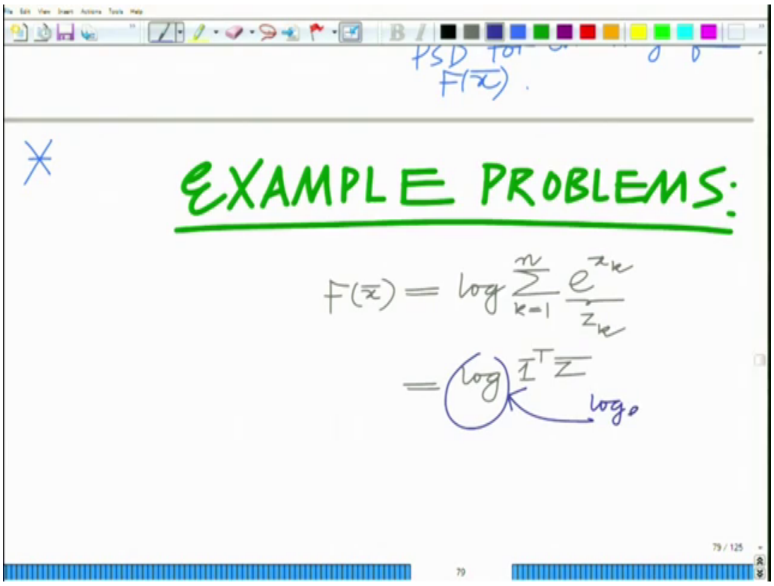
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 31
Example Problems: Operations preserving Convexity (log-sum-exp, average) and Quasi-Convexity

Hello. Welcome to another module in this massive open online course. So, we are looking at example problems for convex function, all right let us continue our discussion.

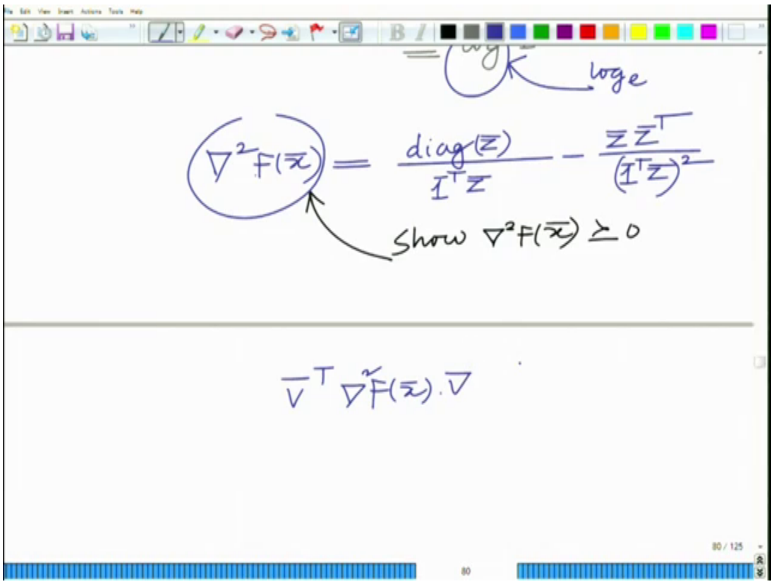
(Refer Slide Time: 00:23)



So, what we are looking at is examples or rather example problems and in particular, we are considering this interesting function which is the log sum exponential; that is, you have F of x bar is log of summation k equals 1 to n e raise to x k.

And if we denote this by Z k; that is e raise to x k by Z k, then you can write this as log 1 bar transpose Z bar. Just to be just to clarify this, this is the natural logarithm. You also you can write this as ln ok, log to the base this is log to the base e that we are considering alright.

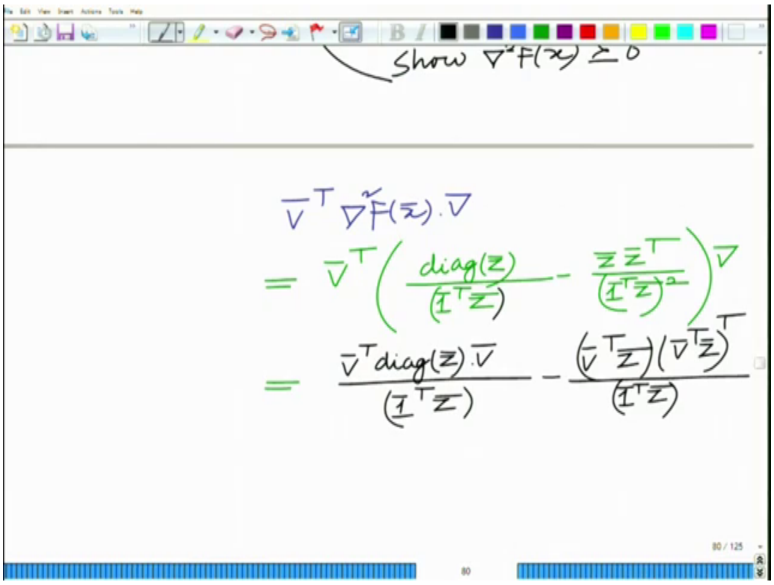
(Refer Slide Time: 01:30)



So, by default, if you are not mentioning anything, then you can assume that it is log to the base e. We have computed the hessian of this function and that has an interesting structure. So, if you look at the ,hessian, we have seen that this is diagonal of Z bar divided by 1 bar transpose Z bar minus Z bar Z bar transpose divided by 1 bar transpose Z bar whole square.

And, now what we want to do is, we want to show that this is positive semi definite that the hessian is positive semi definite ok. Remember, we have already defined the symbol to which basically indicates that this matrix the hessian is a positive semi definite matrix. Now, to show that we employed the straightforward approach, that is consider any vector V bar and multiply this V bar transpose the hessian times V bar.

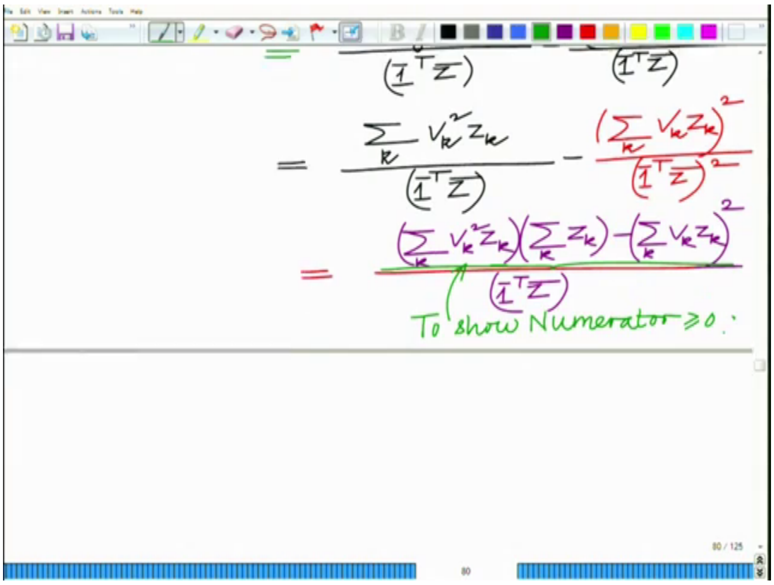
(Refer Slide Time: 03:03)



And, this you can see this is therefore, substituting the hessian, it is diagonal of Z bar by 1 bar transpose Z bar square minus Z bar Z bar transpose divided by 1 bar transpose Z bar square times V bar which is well, this is V bar transpose diagonals Z bar into V bar divided by 1 bar transpose Z bar, I am sorry. There is no square here. 1 bar transpose Z bar correct, 1 bar transpose Z bar minus V bar transpose Z bar times well Z bar transpose V bar. But, you can think of this as V bar transpose Z bar transpose and this is a scalar quantity, right.

So, V bar transpose Z bar is the same thing as Z bar transpose V bar or in other words, Z bar V bar transpose Z bar transpose is the same thing as V bar transpose Z bar, all right. So, it is a scalar quantity V bar transpose Z bar times itself or rather it is basically V bar transpose Z bar square, ok. So, this quantity is V bar transpose Z bar square.

(Refer Slide Time: 04:39)

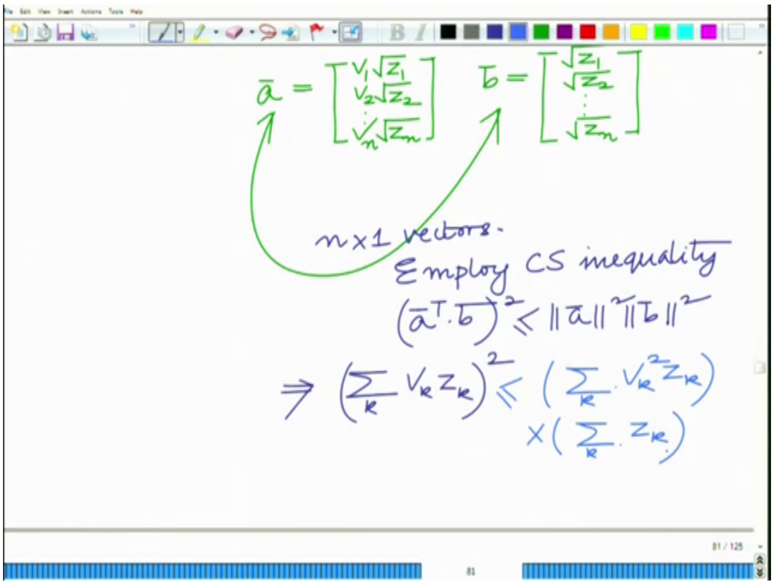


And now if you look at this quantity V bar transpose diagonal Z bar into V bar you can clearly see or you can it is very easy to see that this will be nothing but the summation over k V k square times Z k divided by 1 bar transpose Z bar minus V bar transpose Z bar square which is summation over k V k V k Z k whole square divided by 1 bar transpose Z bar whole square.

And, if you simplify this therefore, now further what you have is in the denominator you will have 1 bar transpose Z bar times summation k, k square Z k into 1 bar transpose Z bar which is nothing but summation k over Z k minus summation k V k Z k whole square and this is the quantity.

And now, to demonstrate that this is positive semi definite we have to demonstrate that V bar transpose the hessian times V bar is greater than equal to 0 for any vector V bar which in term implies that the numerator of this expression has to be greater than equal to 0 and that is something that we are going to show in a straight forward fashion now. So, we want to show that this numerator quantity to show where the numerator is greater than equal to 0.

(Refer Slide Time: 06:57)

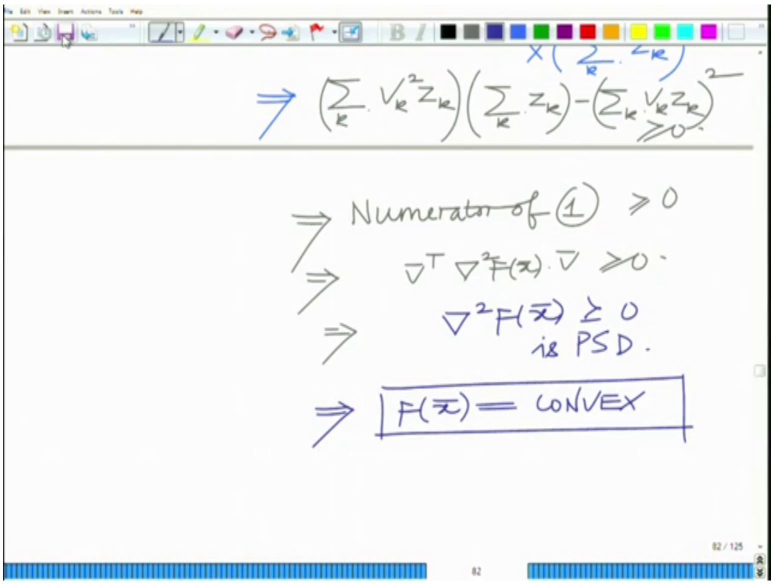


Now, for that what we are going to do is, we are going to define 3 vectors we have a bar equals V 1 square root of Z 1 V 2 square root of Z 2, so on V n square root of Z n. We want to define another vector b bar which is square root of Z 1 square root of Z 2, so on square root of Z n. These are 2 n dimensional vectors or you can also say these are n cross 1 real vectors. And now, employ the Cauchy Schwarz inequality or the inequality for the inner product of vectors which says that a bar transpose b bar whole square that is less than or equal to norm a bar square into norm b bar square, all right.

And, now if you look at norm a bar transpose b bar whole square, that is nothing but summation k V 1 square root of Z 1 into square root of Z 1, that is nothing but V 1 Z 1 similarly V 2 Z 2 and so on. So, this is summation over k V k Z k whole square this is less than or equal to norm a bar square which is V 1 square Z 1 plus V 2 square Z 2 plus V n square Z n.

So, this will be summation k V k square Z k times norm b square which is Z 1 plus Z 2 plus Z n which is summation k Z k. So, this quantity is less than quantity on the right is less than equal to quantity on the left.

(Refer Slide Time: 09:13)

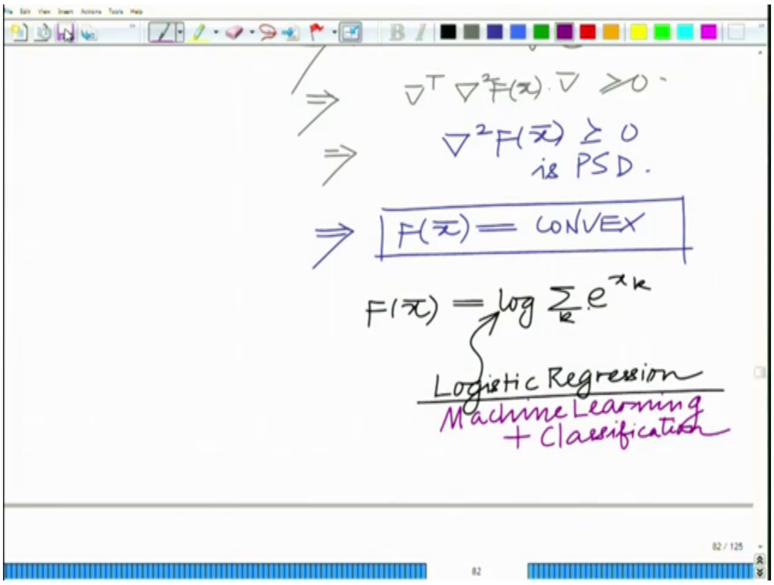


Which means, now you can simplify this as summation k V k square Z k times summation k Z k minus summation k V k Z k square this is greater than equal to 0 implies the numerator is greater than equal to 0 or numerator of one let us call this expression this is what we have said to prove. So, let us call this expression as one implies numerator of 1 is greater than equal to 0 implies V bar transpose the hessian times V bar greater than equal to 0 implies the hessian is positive semi definite, alright. This is the chain of arguments.

So, implies the hessian; the hessian is positive semi definite implies F of x bar equals convex function ok. So, F of x bar which is the log some exponential is convex because we have demonstrated that there is hessian first we have derived the hessian and again in turn demonstrated the hessian is positive semi definite.

So, the proof is a little lengthy and tedious, but it has some very interesting aspects that can be used in general to demonstrate the convexity of functions especially the especially the convexity of functions of vectors and sometimes these proofs indeed tend to be slightly involved.

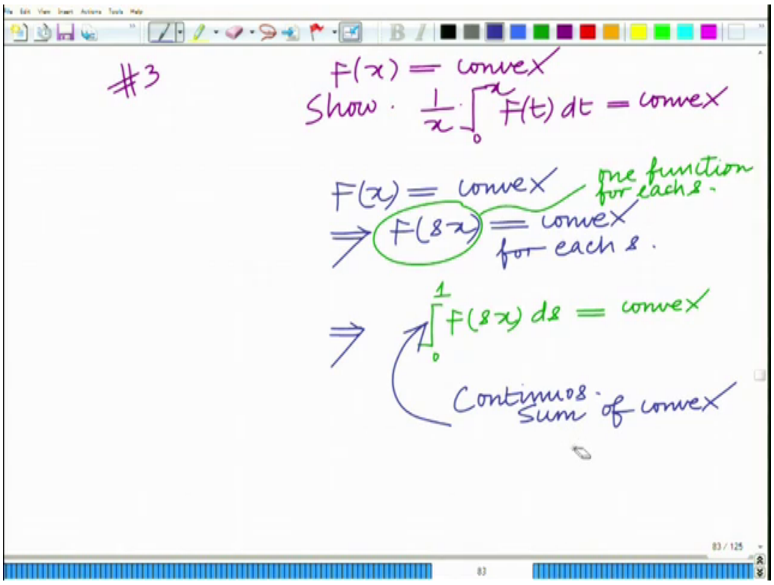
(Refer Slide Time: 11:16)



Let us proceed to the next example and in fact, this is several very interesting applications this is used for in if you look at the log sum exponential this is the original function we can this has a lot of. In fact, this can be used to logistic regression that is to fit a curve to a given set of points, all right. And this has applications in machine learning and classification as we are going to see later in this course to classify to classify a set of data points divided them into 2 sets; one which gives corresponds to a response of 1, other corresponds to response of 0.

So, this can be used for machine learning or classification of data sets ok, all right. On that note, let us move to the next example which is the following.

(Refer Slide Time: 12:29)

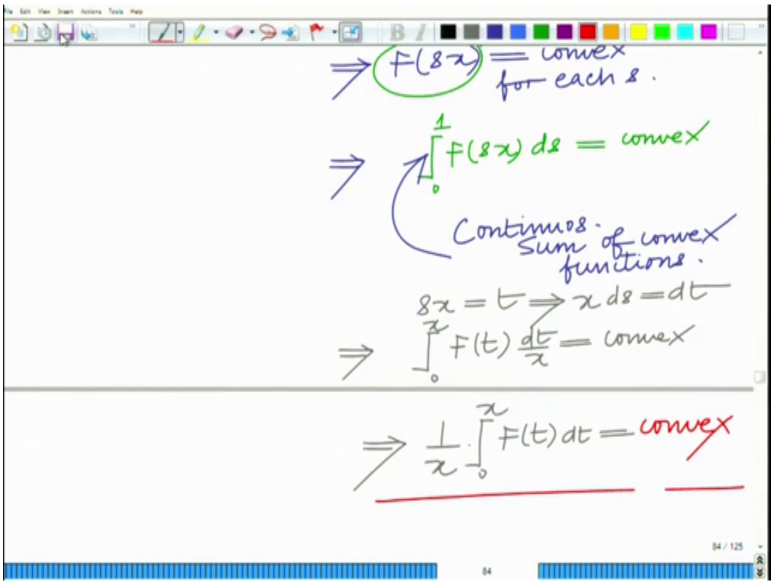


We want to demonstrate that if F of x is convex, we want to show that 1 over x integral 0 to x F of t d t this is convex. We want to show that this is convex or that we will use a simple procedure F of x equals convex implies the fine p composition implies F of s x equals convex for each s.

Now, we will use the property of the sum that is a function is convex functions of convex right, several if you have several functions which is convex their sum is convex. In fact, here we are going to use a continuous sum. So, this implies that if you take the integral, I can treat this as one function for each s ok. This is one function for each value of s.

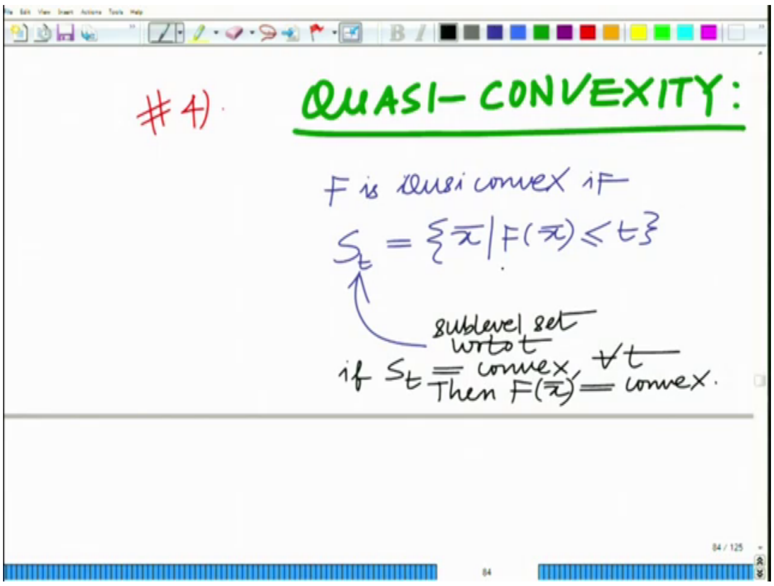
So, this implies that integral 0 to 1 F of s x dx d s. This is convex. Why is this convex? Because, this is a continuous sum one function for each s continuous sum over s for s lying in the interval 0 to s. So,,this is basically a continuous. So, instead of having a discrete sum you have a continuous sum, all right.

(Refer Slide Time: 14:45)



The integral is nothing but a continuous sum and this implies that. Now, in this you set s x equals t which implies that x times d s equals d t. So, this implies now substituting this. So, what you have is basically integral 0 to upper limit becomes s times x which is which is t becomes equal to 1 times x. So, this is 0 to x F of s x is t ds is d t by x is this implies that this is convex which in turn implies that 1 over x because x is a constant integral is with respect to t 0 to x F of t d t. This is convex, alright.

(Refer Slide Time: 16:09)

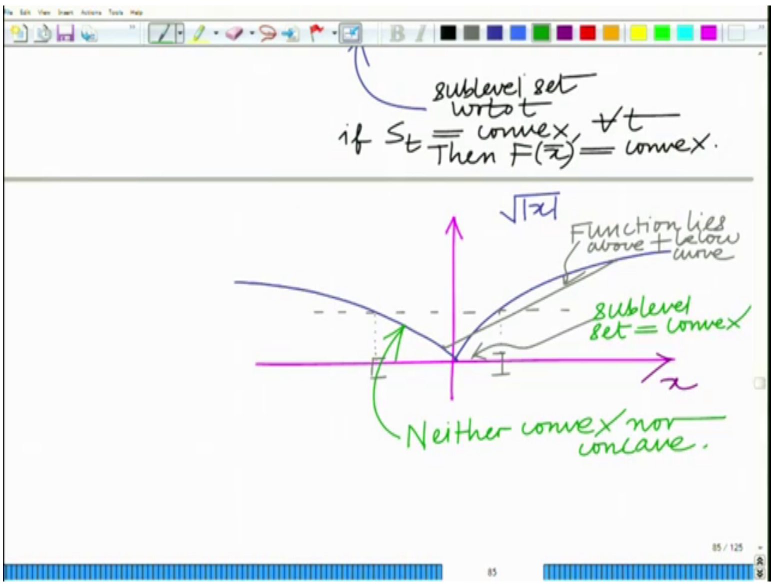


So, this demonstrates that for if F is a convex function 1 over x integral 0 to x F of t d t is also a convex function. Let us go to the next example and in this, we want to look at an interesting concept and this is the concept of quasi convexity. And quasi concavity, we have seen the definition of convexity similarly one can define a set of functions which are quasi convex.

Quasi convex basically means, so, F is quasi convex. If we define the set S of t equals the set of all x or x bar such that F of x bar is less than or equal to t ok. This is called a sublevel set with respect to t. If the sublevel set with respect to t, if is convex for all t, then F of x bar is a convex function that is, if you look at the sublevel sets of this function, what is the sublevel set.

That is, if you look at any parameter value t, consider the set of all points x bar such that F of x takes values less than or equal to t. And, if these sublevel sets with respect to each p are convex the function is said to be quasi convex. This is important because, there are several functions which are not necessarily convex. But, qualify as quasi convex and can also and also have a lot of utility in practical applications for instance, let us take a simple example.

(Refer Slide Time: 18:36)

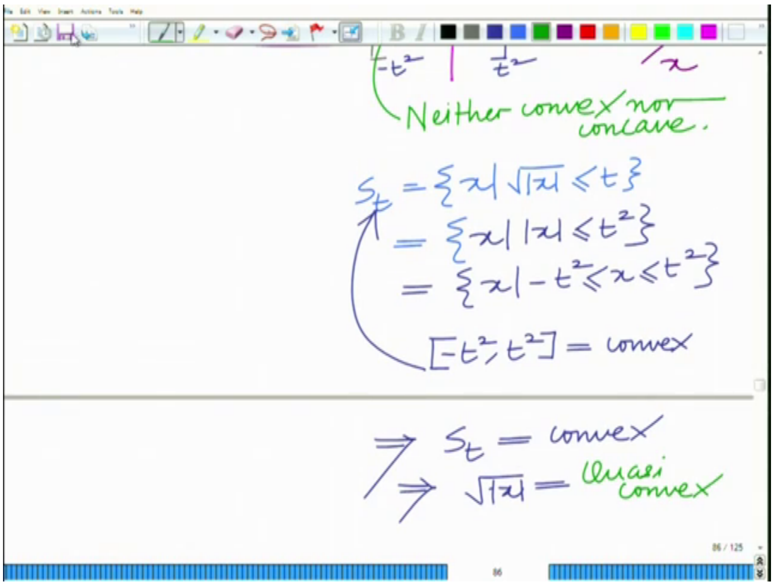


Now, if you look at square root of course, square root of magnitude of x. Now, you can clearly see this is not convex, correct. Because, if you take 2 points let us say right and join them join the chord. Now, function lies part of it lies above a part of it lies below.

So, it is neither convex nor concave ok. So, you can say function straddles the chord or function lies both above plus below the curve. And therefore, this is neither convex nor this is neither convex.

Because, remember we said if the chord lies below it is concave with the chord lies above the function, then it is convex. So, this is neither convex nor concave. However, if you look at the sublevel sets, that is if you look at any t that is you take the all the set all the points such that F of x is less than or equal to t you look at the sublevel set, now this set you can see sublevel set, the sublevel set is convex.

(Refer Slide Time: 20:40)

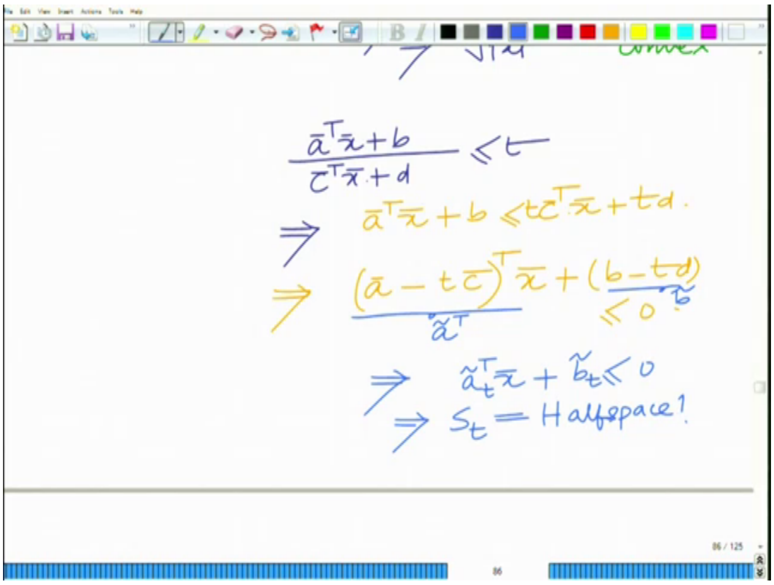


And, you can easily see that for instance, if you look at S of t equals set of all x such that square root of magnitude of x is less than equal to t. This is equal to set of all x such that magnitude of x less than or equal to t square. This is equal to set of all x which is basically minus t square less than equal to x less than equal to t square.

And therefore, if you look at this set, this is the set between minus t square to t square all right. And therefore, this S of t which is basically simply the interval closed interval minus t square to t square. This is a convex set. Remember, if convex set if you take any 2 points in the set, join them by a line segment. It should lie in the set.

So, minus t square to t square in fact, any closed interval this is a convex set and therefore, this implies S of t is convex. And therefore, the sublevel sets are convex S of t is convex implies square root of magnitude of x this is a quasi-convex quasi; quasi means not exactly but, something that can pass for alright.

(Refer Slide Time: 22:20)

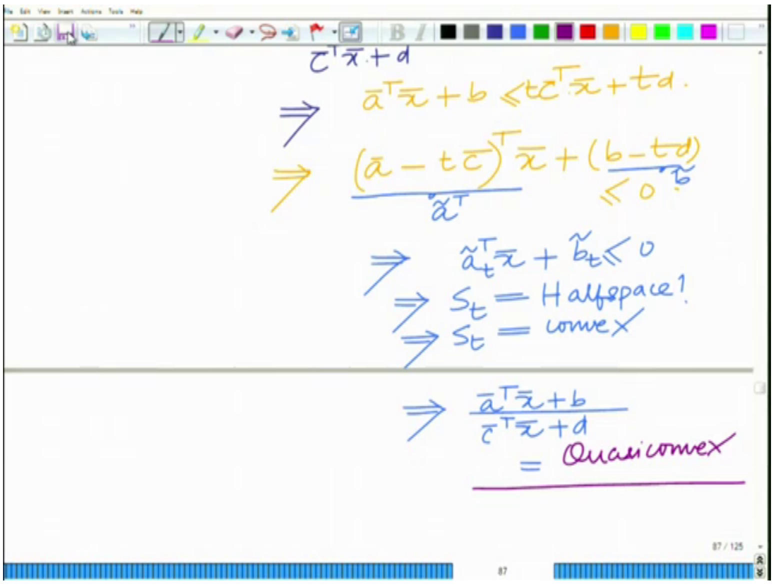


Something that is a quasi alright (Refer Time: 22:28) is a quasi-property. So, this is quasi convex function that is it not strictly speaking convex function, but it has some properties that are similar to the that of a convex function namely that the sublevel sets are convex. Let us look at another example. For instance, if you look at a bar transpose x bar plus b divided by c bar transpose x bar plus d.

Now, this is not a convex function, but if you consider the sublevel set this is less than equal to t. This implies that a bar transpose x bar plus p less than or equal to c bar transpose t times c bar transpose x bar plus t times t times d which basically implies that a bar transpose minus t or you can also say a bar minus t c bar transpose x bar plus b minus t d less than equal to 0.

Now, if you look at this, this is nothing but this is some a tilde transpose b tilde. So, this is basically of the form a tilde transpose x bar plus b tilde is less than or equal to 0 implies level set. In fact, this is a tilde of t this depends on t right. It depends of the parameter t implies and now if you look at this level set, this level set is nothing but a half space ok. So, the level set is a half space. We know that half space is convex. So, all the sublevel sets are convex and therefore, this is a quasi-convex function.

(Refer Slide Time: 24:29)



So, implies S of t is convex implies a bar transpose x bar plus b by c bar transpose x bar plus d. This is a quasi convex function. Similarly, one cannot come up with several other examples for quasi convex, all right. So, we will stop here and continue in the subsequent module.

Thank you very much.