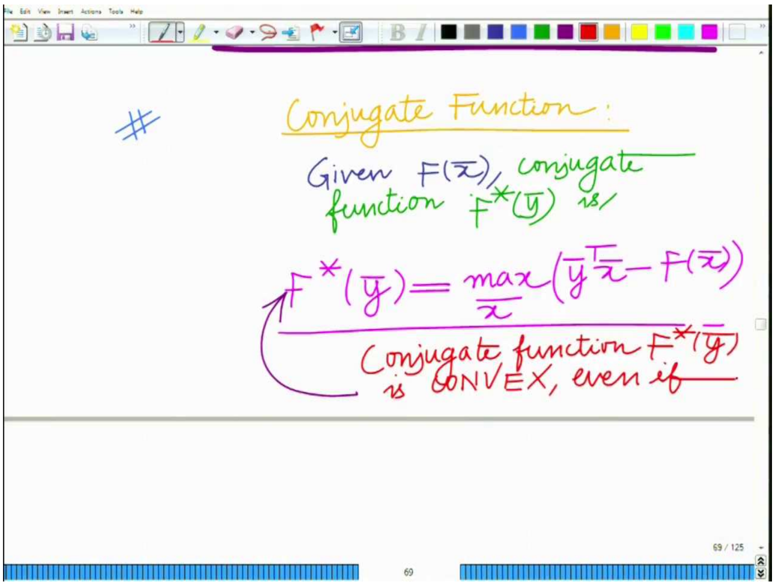
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

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

Hello, welcome to another module in this massive open online course. Let us now focus on some examples to understand the concepts discussed later.

(Refer Slide Time: 01:54)



So, the first example is the conjugate function of function $F(\bar{x})$ which is denoted by $F^*(\bar{y})$. It is this is given as

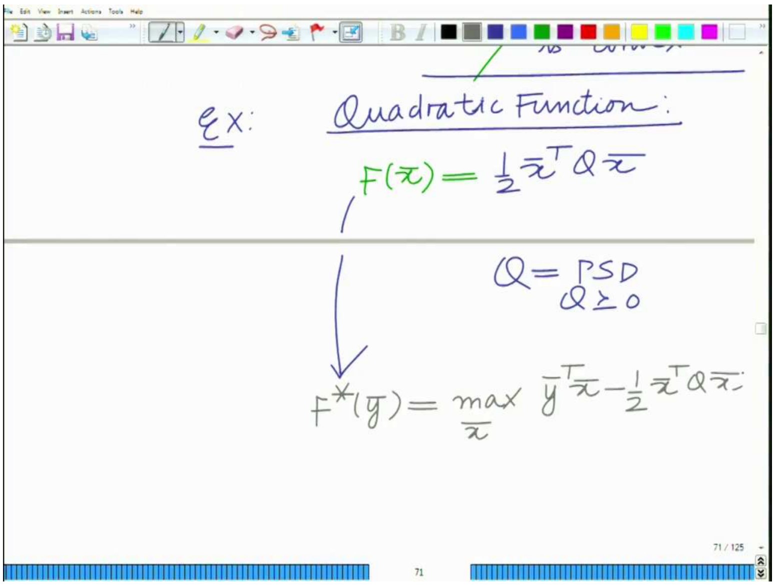
$$F^*(\bar{y}) = \max_{\bar{x}} \left(\bar{y}^T \bar{x} - F(\bar{x})\right)$$

This conjugate function $F^*(\bar{y})$ is defined such that it is convex irrespective of the convexity of $F(\bar{x})$. It is the interesting aspect of this conjugate function.

So corresponding to any function $F(\bar{x})$ being convex or non-convex, one can construct an associated convex function which is a conjugate function.

For instance, $\bar{y}^T \bar{x} - F(\bar{x})$ is a linear function in $\bar{y}$ for each value of $\bar{x}$. So the maximum of a set of convex functions is also convex.

(Refer Slide Time: 05:11)



For example consider the quadratic function.

$$F(\bar{x}) = \frac{1}{2}\bar{x}^T Q\bar{x}$$

Q is a symmetric positive semi definite matrix. This means

$$Q \geq 0$$

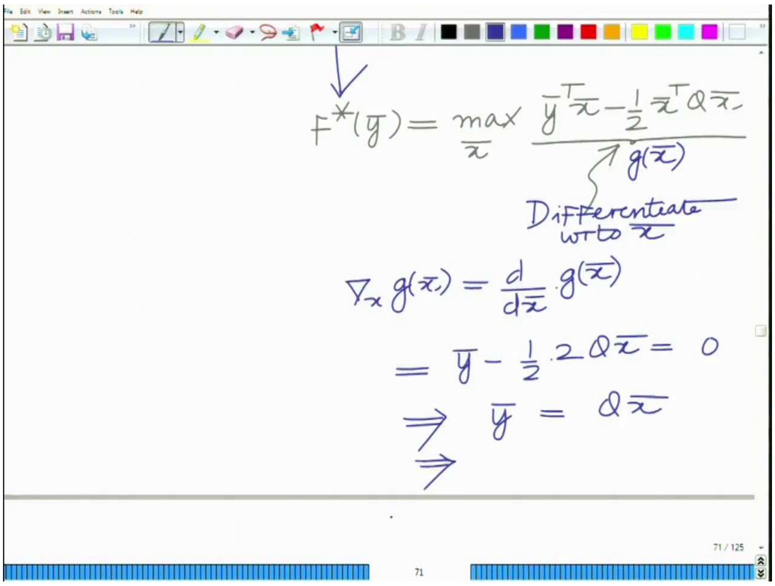
So the conjugate function is constructed as,

$$F^*(\bar{y}) = \max_{\bar{x}} \left(\bar{y}^T \bar{x} - \frac{1}{2}\bar{x}^T Q\bar{x} \right)$$

Let us say

$$g(\bar{x}) = \left(\bar{y}^T \bar{x} - \frac{1}{2}\bar{x}^T Q\bar{x} \right)$$

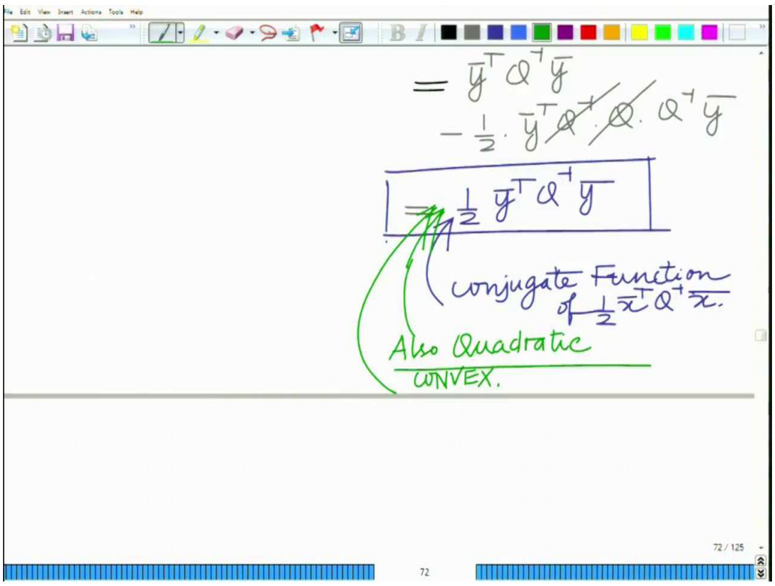
(Refer Slide Time: 06:35)



Now to maximize this, differentiate $g(\bar{x})$ with respect to $\bar{x}$ and make it equal to zero.

$$\nabla_{\bar{x}} g(\bar{x}) = 0$$
$$\bar{y}^T - \frac{2Q\bar{x}}{2} = 0$$
$$\bar{y}^T - Q\bar{x} = 0$$
$$\bar{x} = Q^{-1}\bar{y}$$

(Refer Slide Time: 09:17)



Therefore the conjugate function is

$$F^*(\bar{y}) = \bar{y}^T\left(Q^{-1}\bar{y}\right) - \frac{1}{2}\left(Q^{-1}\bar{y}\right)^T Q\left(Q^{-1}\bar{y}\right)$$

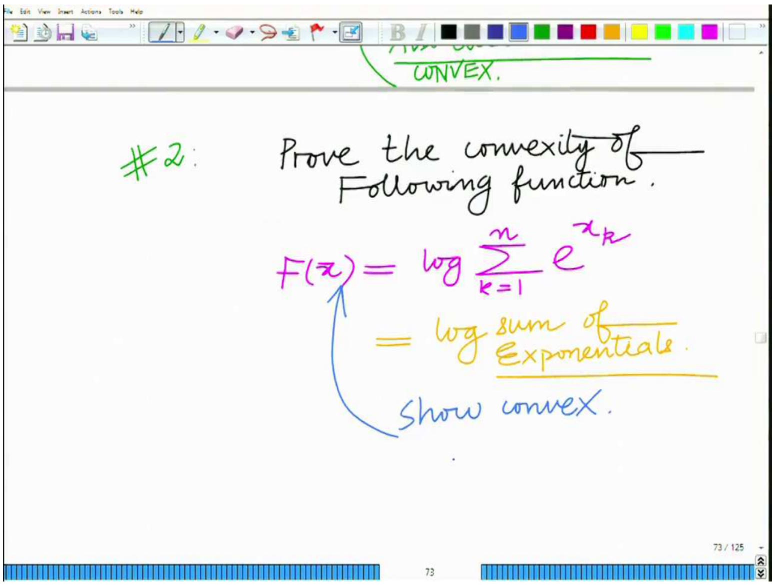
$$= \bar{y}^T Q^{-1}\bar{y} - \frac{1}{2}\bar{y}^T Q^{-1} Q Q^{-1}\bar{y}$$

$$= \bar{y}^T Q^{-1}\bar{y} - \frac{1}{2}\bar{y}^T Q^{-1}\bar{y}$$

$$= \frac{1}{2}\bar{y}^T Q^{-1}\bar{y}$$

Here it can be seen that this conjugate function is also quadratic. Hence this is convex because Q is positive semi definite which implies inverse of this matrix Q is also positive semi definite.

(Refer Slide Time: 10:49)

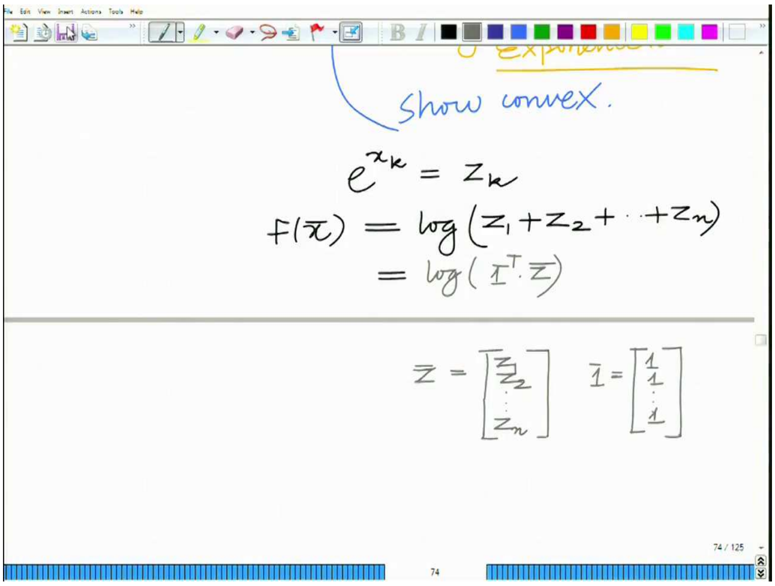


Next example is to prove the convexity of the log of sum of exponentials which is as follows.

$$F(\bar{x}) = \log\left|\sum_{k=1}^{n} e^{x_k}\right|$$

This function arises in several applications.

(Refer Slide Time: 12:30)



To prove that this function is convex, consider

$$z_k = e^{x_k}$$

Thus the function can be represented as

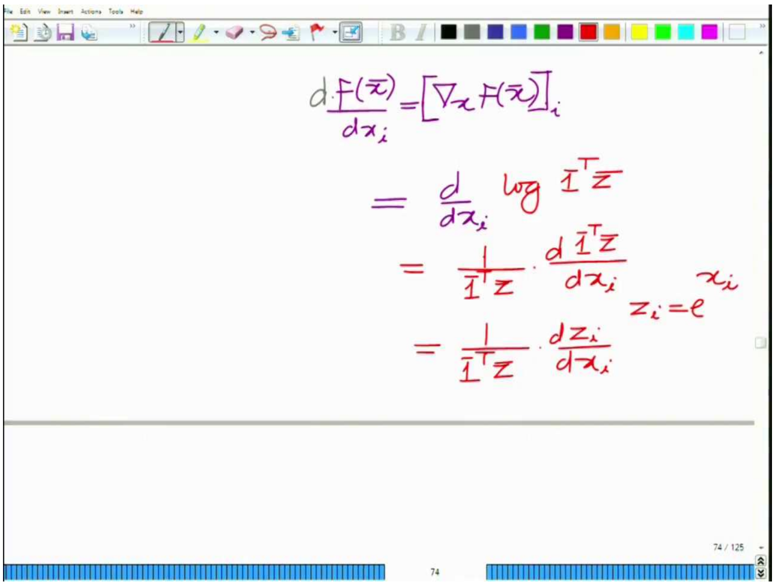
$$\begin{aligned} F(\bar{x}) &= \log|z_1 + z_2 + \ldots + z_n| \\ &= \log|\bar{1}^T \bar{z}| \end{aligned}$$

Where $\bar{z}$ and $\bar{1}$ are $n \times 1$ matrices and are defined as follows.

$$\bar{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \quad \text{and} \quad \bar{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

So, if the Hessian of this function is a positive semi definite matrix then this function would be a convex function.

(Refer Slide Time: 13:56)



So, the first order derivative of this function is

$$\frac{dF(\bar{x})}{dx_i} = \left[\nabla_x F(\bar{x})\right]_i$$

$$= \frac{d}{dx_i} \log\left|\bar{1}^T \bar{z}\right|$$

$$= \frac{1}{\bar{1}^T \bar{z}} \frac{d\left|\bar{1}^T \bar{z}\right|}{dx_i}$$

$$= \frac{1}{\bar{1}^T \bar{z}} \frac{dz_i}{dx_i}$$

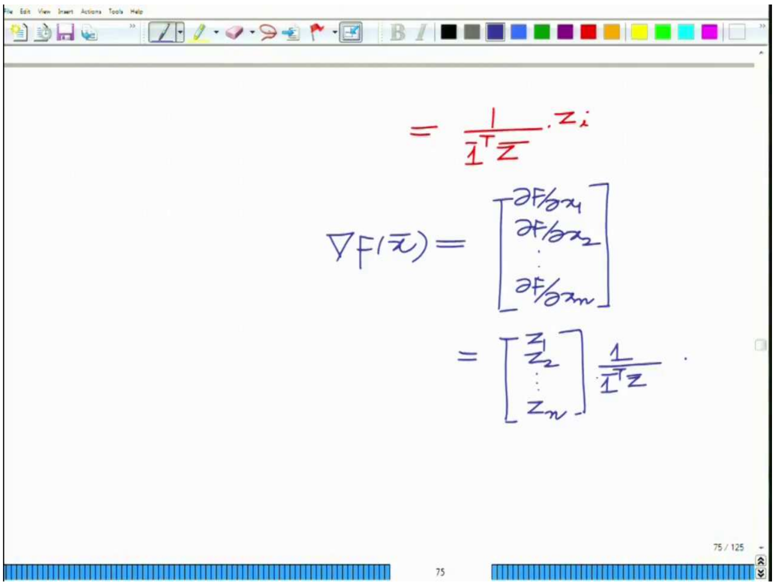
As $z_i = e^{x_i}$ thus the differentiation of z_i with respect to x_i is

$$\frac{dz_i}{dx_i} = e^{x_i} = z_i$$

Therefore

$$\frac{dF(\bar{x})}{dx_i} = \frac{1}{\bar{1}^T \bar{z}} z_i$$

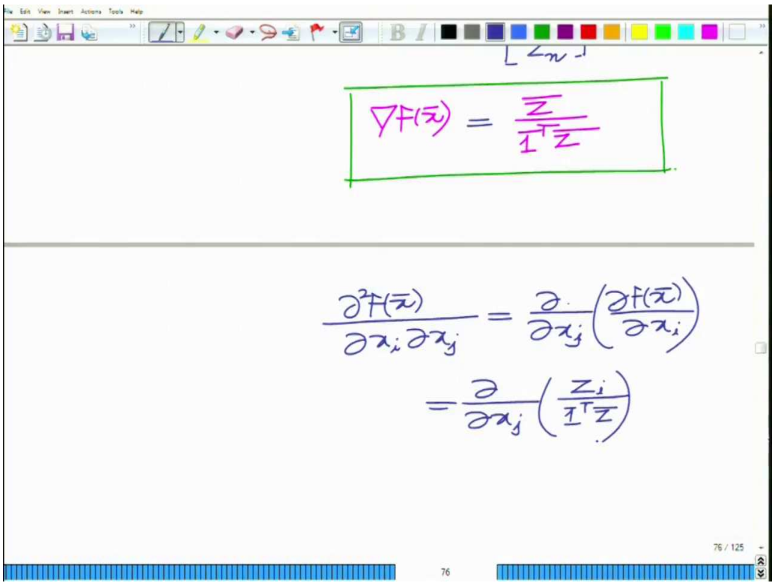
(Refer Slide Time: 15:36)



So, the gradient of above function can be expanded as

$$\nabla_x F(\bar{x}) = \begin{bmatrix} \dfrac{\partial F}{\partial x_1} \\ \dfrac{\partial F}{\partial x_2} \\ \vdots \\ \dfrac{\partial F}{\partial x_n} \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \frac{1}{1^T \bar{z}}$$

(Refer Slide Time: 16:35)



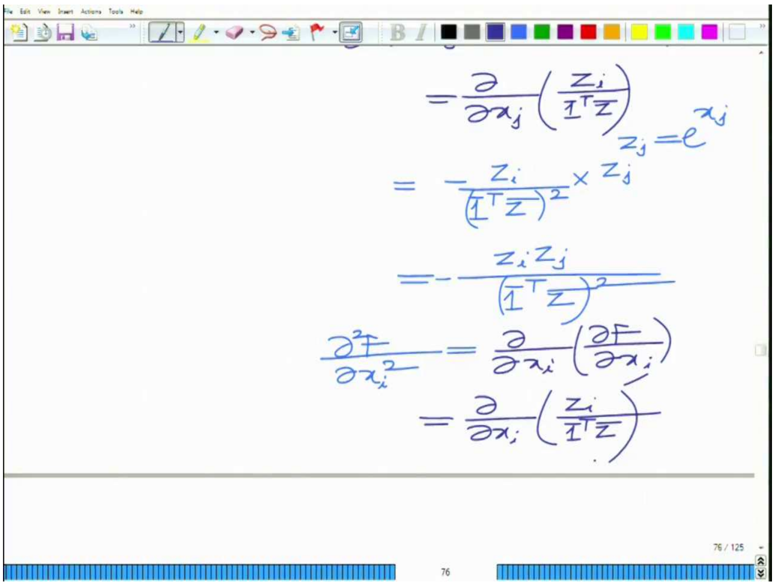
Therefore one can write it as

$$\nabla F(\bar{x}) = \frac{\bar{z}}{1^T \bar{z}}$$

Similarly the second order derivative (Hessian) of the function is

$$\frac{\partial^2 F(\bar{x})}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j}\left(\frac{\partial F(\bar{x})}{\partial x_i}\right)$$
$$= \frac{\partial}{\partial x_j}\left(\frac{z_i}{1^T \bar{z}}\right)$$

(Refer Slide Time: 17:58)



As $z_j = e^{x_j}$ therefore

$$\frac{\partial^2 F(\bar{x})}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j}\left(\frac{z_i}{1^T \bar{z}}\right)$$
$$= -\frac{z_i z_j}{\left(1^T \bar{z}\right)^2}$$

Also if comparing this second order derivative with $\frac{\partial^2 F(\bar{x})}{\partial x_i^2}$ then

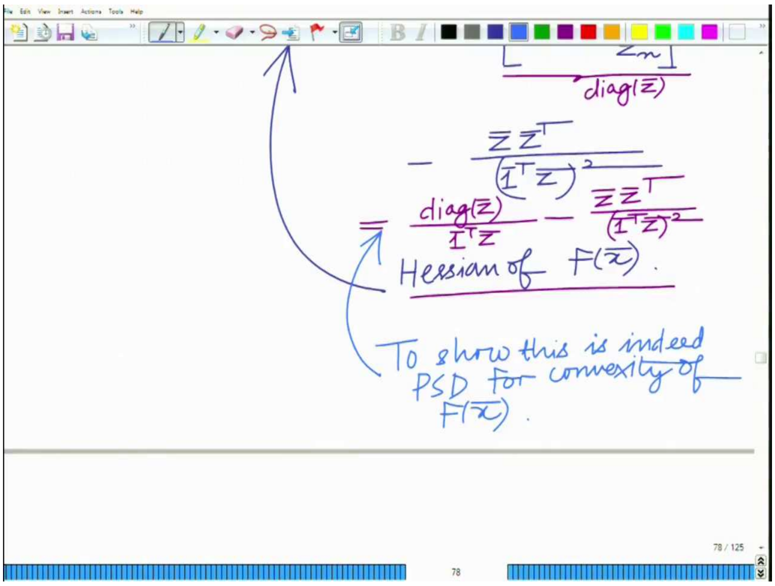
$$\frac{\partial^2 F(\bar{x})}{\partial x_i^2} = \frac{\partial}{\partial x_i}\left(\frac{\partial F(\bar{x})}{\partial x_i}\right)$$

$$= \frac{\partial}{\partial x_i}\left(\frac{z_i}{\bar{1}^T \bar{z}}\right)$$

$$= \frac{z_i}{\bar{1}^T \bar{z}} - \frac{z_i^2}{\left(\bar{1}^T \bar{z}\right)^2}$$

So, here one can observe that the first term of this derivative i.e. $\frac{z_i}{\bar{1}^T \bar{z}}$ is only present in the second order derivative form of function with respect to x_i and are not present in the second order derivative form of the function with respect to $x_i x_j$.

(Refer Slide Time: 23:24)



Hence the Hessian of this function $F(\bar{x})$ is of the following form.

$$\nabla^2 F(\bar{x}) = \begin{bmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \\ \frac{\partial^2 F}{\partial x_1 \partial x_2} & \frac{\partial^2 F}{\partial x_2^2} & \\ & & \ddots \\ & & \end{bmatrix}$$

So the Hessian of this function $F(\bar{x})$ is

$$\nabla^2 F(\bar{x}) = \frac{1}{\bar{1}^T \bar{z}} \underbrace{\begin{bmatrix} z_1 & & & \\ & z_2 & & \\ & & \ddots & \\ & & & z_n \end{bmatrix}}_{\mathrm{diag}(\bar{z})} - \frac{\bar{z}\bar{z}^T}{\left(\bar{1}^T \bar{z}\right)^2}$$

$$= \frac{\mathrm{diag}(\bar{z})}{\bar{1}^T \bar{z}} - \frac{\bar{z}\bar{z}^T}{\left(\bar{1}^T \bar{z}\right)^2}$$

The matrix $\mathrm{diag}(\bar{z})$ is the diagonal matrix of function $\bar{z}$. So this is the Hessian of the given function and we need to prove that this is positive semi definite to show that the given function $F(\bar{x})$ is convex. This we will do in the subsequent module.