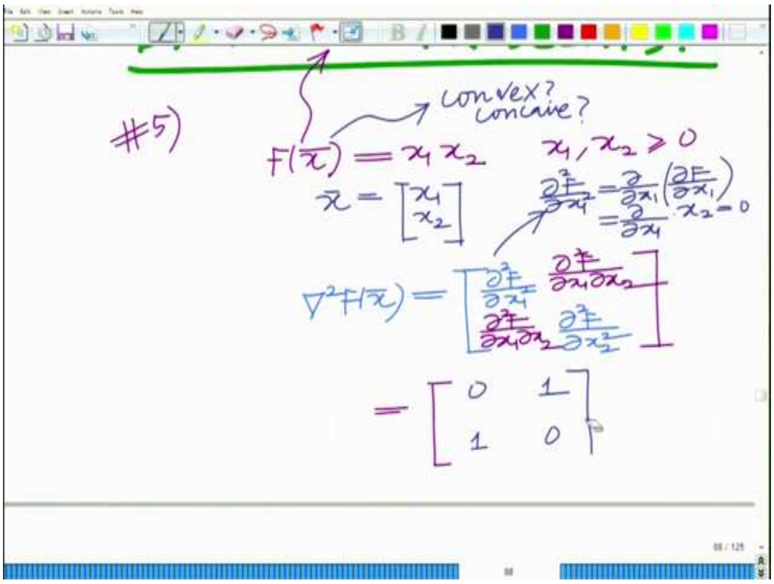
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 32
Example Problems: verify Convexity, Quasi-Convexity and Quasi-Concavity of functions

Hello. Welcome to another module in this massive open online course. So, we are looking at example problems in convex functions and convexity. Let us continue our discussion.

(Refer Slide Time: 01:44)



Next function to check the convexity is

$$F(\bar{x}) = x_1 x_2 \qquad \text{for } x_1, x_2 \geq 0$$

Here $\bar{x}$ is a 2 dimensional vector.

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Let us take a simple test for convexity. The hessian of this function is

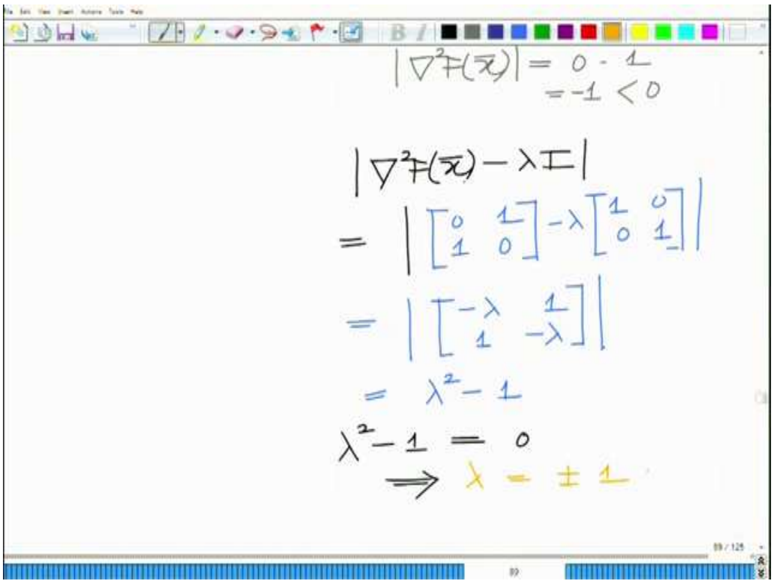
$$\nabla^2 F(\bar{x}) = \begin{bmatrix} \dfrac{\partial^2 F}{\partial x_1^2} & \dfrac{\partial^2 F}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 F}{\partial x_1 \partial x_2} & \dfrac{\partial^2 F}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

So the hessian is a symmetric matrix. Now the determinant of this hessian is

$$\left|\nabla^2 F(\bar{x})\right| = 0 - 1 = -1$$

Remember, the determinant of a positive semi definite matrix has to be a positive quantity because the determinant is the product of the eigenvalues, and for positive semi definite, all of these eigenvalues are non-negative. Therefore, as $\left|\nabla^2 F(\bar{x})\right| < 0$ so this means the hessian of above matrix is not positive semi definite.

(Refer Slide Time: 04:24)



In fact,

$$\left|\nabla^2 F(\bar{x}) - \lambda I\right| = 0$$

$$\left|\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right| = 0$$

$$\left|\begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix}\right| = 0$$

$$\lambda^2 - 1 = 0$$

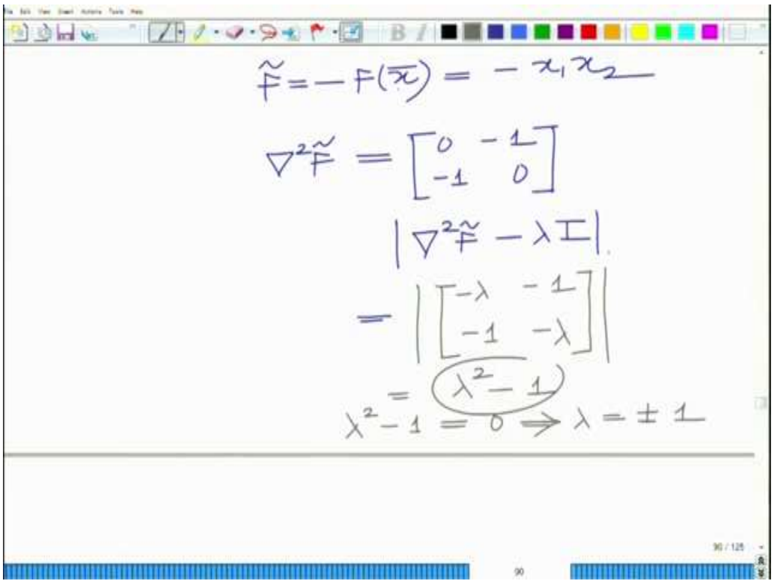
$$\lambda = \pm 1$$

So it has the eigenvalues both positive and negative. This again shows that hessian of this function is not positive semi definite and hence this function is not convex.

Let us check this function's concavity. So, consider a new function $\tilde{F}$ such that

$\tilde{F} = -F(\bar{x}) = -x_1 x_2$ for $x_1, x_2 \geq 0$

(Refer Slide Time: 06:39)



The hessian of this function $\tilde{F}$ is

$$\nabla^2 \tilde{F} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

Let us check the eigenvalues of this function.

$$\left|\nabla^2\tilde{F}-\lambda I\right|=0$$

$$\left\|\begin{bmatrix}-\lambda & -1\\ -1 & -\lambda\end{bmatrix}\right\|=0$$

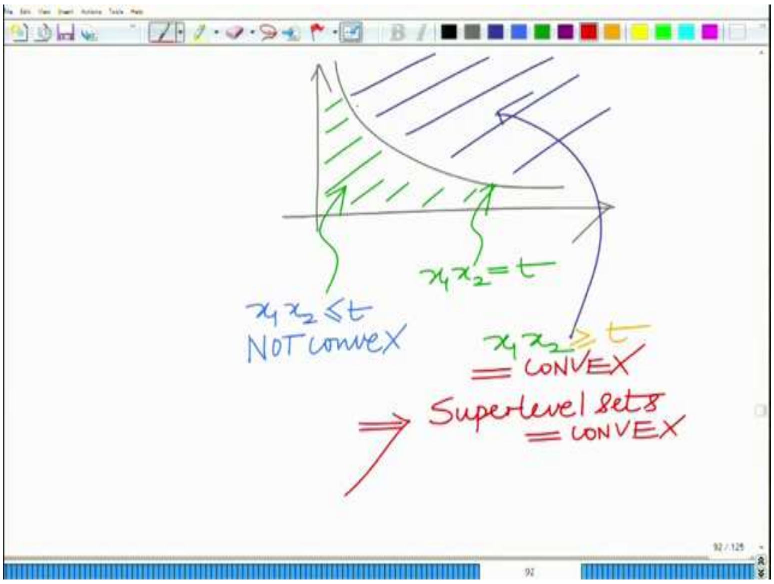
$$\lambda^2-1=0$$

$$\lambda=\pm 1$$

So function $\tilde{F}$ has the eigenvalues both positive and negative. Again the hessian of this function is not positive semi definite and hence this new function $\tilde{F}$ is not convex. This means that function $F(\bar{x})$ is not convex. This implies that function $F(\bar{x})$ is neither convex nor concave.

That shows that any function does not always need to be either convex or concave. Some functions are neither convex nor concave.

(Refer Slide Time: 09:34)



Let us check the quasi convexity of function $F(\bar{x}) = x_1x_2$. The sublevel set of this function is

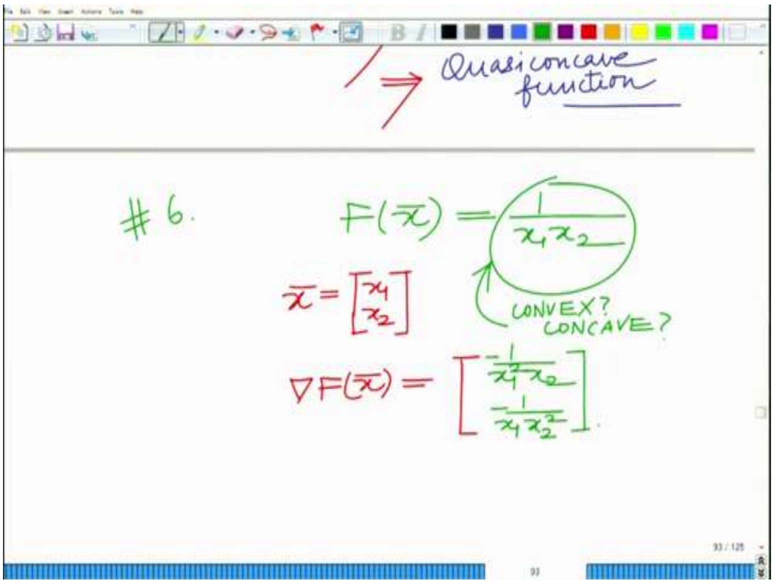
$$S_t = \{\bar{x} \mid x_1x_2 \leq t\}$$

If we draw the plot of S_t, it is clearly seen that this sublevel set S_t is not convex. On the other hand, observe the counterpart of this set that is the super level set. It is defined as

$$\tilde{S}_t = \{\bar{x} \mid x_1 x_2 \geq t\}$$

This super level set is a convex set. This means that sublevel set is concave. This implies function $F(\bar{x})$ is quasi-concave. Thus, if the super level sets are convex then the corresponding function is a quasi-concave function.

(Refer Slide Time: 11:16)



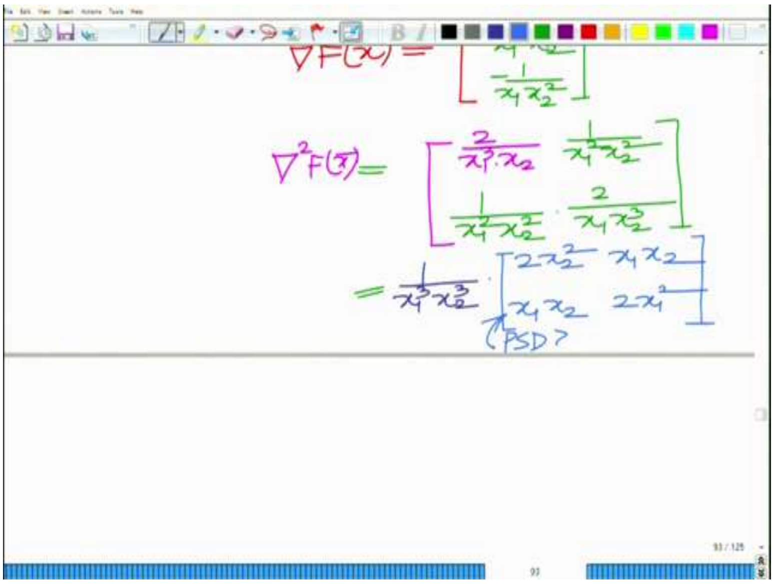
Let us consider the reciprocal of the previous function.

$$F(\bar{x}) = \frac{1}{x_1 x_2} \qquad \text{for } x_1, x_2 > 0$$

Vector $\bar{x}$ is a 2 dimensional vector.

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(Refer Slide Time: 12:52)



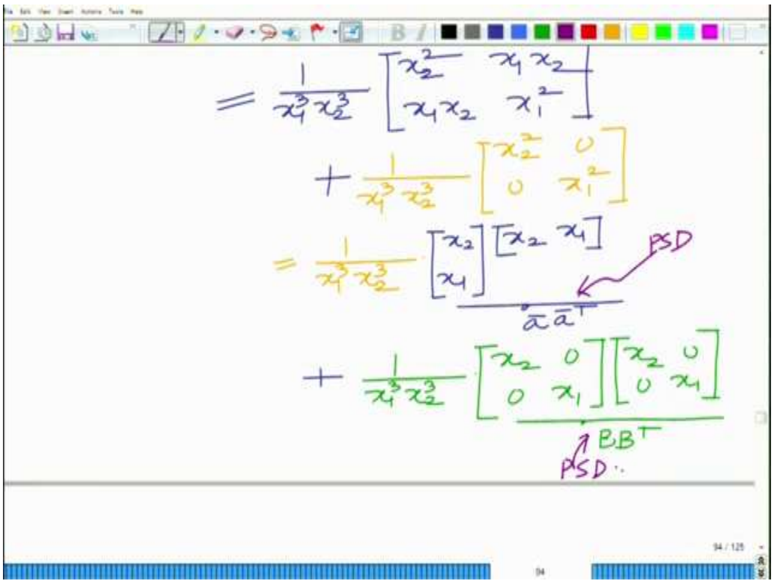
Again take a simple test for convexity. The hessian of this function is

$$
\begin{aligned}
\nabla^2 F(\bar{x}) &= \begin{bmatrix} \dfrac{\partial^2 F}{\partial x_1^2} & \dfrac{\partial^2 F}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 F}{\partial x_1 \partial x_2} & \dfrac{\partial^2 F}{\partial x_2^2} \end{bmatrix} \\
&= \begin{bmatrix} \dfrac{2}{x_1^3 x_2} & \dfrac{1}{x_1^2 x_2^2} \\ \dfrac{1}{x_1^2 x_2^2} & \dfrac{2}{x_1 x_2^2} \end{bmatrix} \\
&= \frac{1}{x_1^3 x_2^3} \begin{bmatrix} 2x_2^2 & x_1 x_2 \\ x_1 x_2 & 2x_1^2 \end{bmatrix}
\end{aligned}
$$

Further on decomposing this function,

$$
\begin{aligned}
\nabla^2 F(\bar{x}) &= \frac{1}{x_1^3 x_2^3} \begin{bmatrix} x_2^2 & x_1 x_2 \\ x_1 x_2 & x_1^2 \end{bmatrix} + \frac{1}{x_1^3 x_2^3} \begin{bmatrix} x_2^2 & 0 \\ 0 & x_1^2 \end{bmatrix} \\
&= \frac{1}{x_1^3 x_2^3} \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \begin{bmatrix} x_2 & x_1 \end{bmatrix} + \frac{1}{x_1^3 x_2^3} \begin{bmatrix} x_2 & 0 \\ 0 & x_2 \end{bmatrix} \begin{bmatrix} x_2 & 0 \\ 0 & x_2 \end{bmatrix}
\end{aligned}
$$

(Refer Slide Time: 16:35)



Let

$$\bar{a} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} x_2 & 0 \\ 0 & x_2 \end{bmatrix}$$

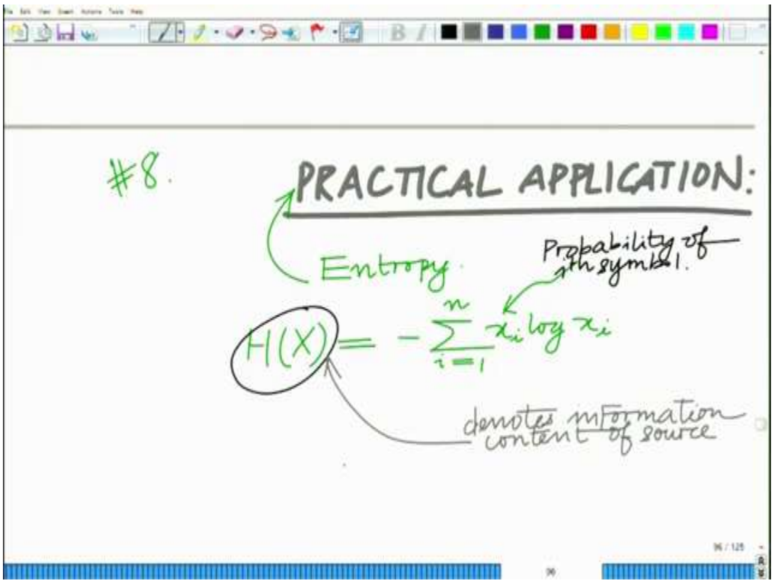
So

$$\overline{aa}^T = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \begin{bmatrix} x_2 & x_1 \end{bmatrix},$$

$$BB^T = \begin{bmatrix} x_2 & 0 \\ 0 & x_2 \end{bmatrix} \begin{bmatrix} x_2 & 0 \\ 0 & x_2 \end{bmatrix}$$

This shows that the above hessian of function is the sum of two positive semi definite matrices. Therefore, the hessian of the function is positive semi definite. Hence function $\frac{1}{x_1 x_2}$ is convex and therefore it is quasi convex also.

(Refer Slide Time: 19:49)



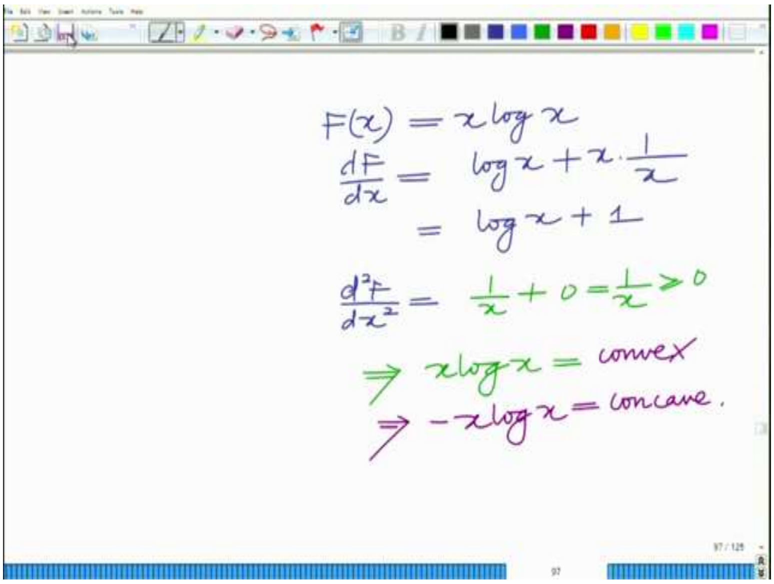
Let us look at as a practical application of quasi convex functions from the information theory which is the entropy function.

$$H(X) = -\sum_{i=1}^{n} x_i \log x_i$$

This entropy denotes the information content of the source with n symbols which have probabilities $x_1, x_2, \cdots, x_n$. So, if the entropy is high, then it means the information content of the source is the high. And therefore, this entropy needs to be maximized.

So, this is a very important quantity in information theory and by extension, also in wireless communication and signal processing and even in machine learning.

(Refer Slide Time: 23:16)



Consider

$$F(x) = x \log x$$

Take the first derivative of $F(x)$.

$$\frac{dF(x)}{dx} = \log x + x\frac{1}{x}$$
$$= \log x + 1$$

Take the second derivative of $F(x)$.

$$\frac{d^2F(x)}{dx^2} = \frac{1}{x} + 0 \geq 0$$

This implies that $x \log x$ is convex. Thus $-x \log x$ is concave.

As entropy function $H(X)$ is the sum of concave function, hence it is a concave function.