

Optimization - EE5327

Srujana B - MA17BTECH11001

J. Sai Vyshnavi - MA17BTECH11005

20/02/2019

Question 3.3

- Define $L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x})$ and $\nabla = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial \lambda} \end{bmatrix}$.

Solve the equations $\nabla L(\mathbf{x}, \lambda) = 0$.

$$\text{where } f(\mathbf{x}) = (x_1 - 8)^2 + (x_2 - 6)^2 \text{ and } g(\mathbf{x}) = x_1 + x_2 - 9 = 0, \mathbf{x} = (x_1 \ x_2)^T.$$

How is this related to problem 3.1? What is the sign of λ ?

L is known as the Lagrangian and the above technique is known as the Method of Lagrange Multipliers,

Solution

$$L(\mathbf{x}, \lambda) = (x_1 - 8)^2 + (x_2 - 6)^2 - \lambda(x_1 + x_2 - 9)$$

$$\nabla L(\mathbf{x}, \lambda) = \begin{bmatrix} 2x_1 - 16 - \lambda \\ 2x_2 - 12 - \lambda \\ x_1 + x_2 - 9 \end{bmatrix}$$

Equating $\nabla L(\mathbf{x}, \lambda)$ to 0

Solution

$$\begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} 16 \\ 12 \\ 9 \end{bmatrix}$$

This is of the form $A\mathbf{x} = \mathbf{b}$

Solving the above equations, we get $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} \frac{11}{2} \\ \frac{7}{2} \\ -5 \end{bmatrix}$

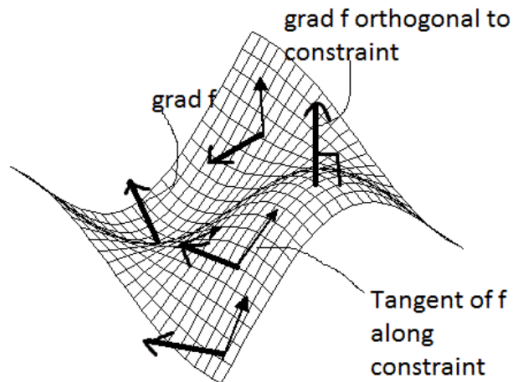
$\lambda = -5$ i.e. it is negative

Solution

- ▶ Let, maximize or minimize a function $f(x,y)$ under a constraint path $g(x,y)=c$.
- ▶ Now,
Gradient is the multidimensional rate of change of given function.

Figure

(27).png



Solution

- ▶ So, $\nabla f(x,y)$ is a vector that indicates direction to maximize $f(x,y)$. But due to constraint $g(x,y)=c$ allowed direction is to move along the tangents to constraint $g(x,y)=c$ which is orthogonal to $\nabla g(x,y)$.
- ▶ Therefore, to maximize $f(x,y)$, directional derivative along the path df must be zero, which means $\nabla f(x,y,z)$ must be orthogonal to $g(x,y,z)=c$ i.e. parallel to $\nabla g(x,y,z)$.

Solution

- ▶ Now, if two vectors are parallel they can be defined linearly (scalar multiply). Thus, $\nabla f = \lambda \nabla g$ where λ is the Lagrange multiplier. Lagrange function is defined by $L(x,y) = f(x,y) - \lambda g(x,y)$
- ▶ So, $\nabla L = 0$

