Lagrangian formulation
[illegible]

1) Today's agenda
2) Applications
3) Mathematical definition
4) Review of undergraduate solution
5) Cannot complete this
6) Solved example

Applications

Suppose you have a function
$f(x) = 3x + 2x^2$
x is the variable
$f(x)$ is the output
Now, suppose you want to find the minimum value of the function $f(x)$
$f(x) = 3x + 2x^2$
$\frac{df(x)}{dx} = 3 + 4x = 0$
$x = -\frac{3}{4}$
$f(x) = 3(-\frac{3}{4}) + 2(-\frac{3}{4})^2$
$= -\frac{9}{8}$

Now, if we want to find the minimum value of the function with constraint $x = 1$

$f(x) = 3x + 2x^2$
$\frac{df(x)}{dx} = 3 + 4x$
$\boxed{x = 1}$
$\therefore f(x) = 3(1) + 2(1)^2 = 5$

However, when you place constraints on the above equation, the minimum value of the above function subject to constraint might be different.

If an upper constraint that $x \le 1$, then the minimum value will become $x = 1$, the minimum value of the function will become $f(x) = 5$

This is where Lagrangian formulation comes into picture. Lagrangian L takes the constraints into consideration when formulating the objective function.

1) Lagrangian L takes the constraints into consideration when formulating the objective function.

2) This formulation helps in applying traditional calculus methods

Lagrangian formulation

Theorem: Let $f \& g^j : j = 1 \dots m$, $m \le n$, be continually differentiable on an open set $X \subset \mathbb{R}^n$. Suppose that x^* is a local maximum point of f for all $x \in X$ $g^j(x^*)$ that also satisfy $g^j(x) = 0$, $j = 1, \dots m$. Also suppose that the Jacobian matrix $J = [\nabla g^1(x^*) \dots \nabla g^m(x^*)]$ does exist. Then there exists a unique vector of real numbers $\lambda = (\lambda_1^*, \dots \lambda_m^*)$ called "Lagrange", such that

$$\boxed{-\nabla f(x^*) = \sum_{j=1}^{m} \lambda_j^* \nabla g^j(x^*)}$$

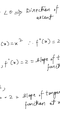
$$-[\nabla f(x^*)] / \nabla x^*] = \sum_{j=1}^{m} \lambda_j^* [\nabla g^j(x^*) / \nabla x^*], \text{ for } j = 1, \dots m$$

$$\boxed{\mathcal{L}(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g^j(x)} \rightarrow \text{Lagrangian equation}$$

$\lambda_j \rightarrow$ Lagrangian multiplier

Geometric Meaning

Gradients are orthogonal to level curves



Concept Review

1) Partial derivative → derivative with respect to one variable while keeping other variables constant.

Ex:- Let $f(x, y) = x^2 y + y^3$

$\frac{\partial f}{\partial x} = 2xy$; $\frac{\partial f}{\partial y} = x^2 + 3y^2$

2) If $\frac{\partial f(x)}{\partial x_i} > 0$, then to increase the value of the function, we must should in the direction of increasing the value of the variable x_i which we are differentiating

3) If $\frac{\partial f(x)}{\partial x_i} < 0$, then to increase the value of the function, we should move in the direction of decreasing value of variable x_i which we are differentiating

4) $\frac{\partial f(x)}{\partial x_i} = 0$, then additional point on local minimum or maximum.

5) Direction of gradient → Direction of steepest ascent

If $\frac{\partial f(x)}{\partial x_i} > 0 \Rightarrow$ Direction of steepest ascent = Direction of gradient

$\frac{\partial f(x)}{\partial x_i} < 0 \Rightarrow$ Direction of steepest ascent = Direction of gradient

Ex:- $f(x) = x^2$ $\therefore f'(x) = 2x$

Now at $x = 1$, $f'(x) = 2 = 2$ slope of tangent to the function at $x = 1$ = Direction of gradient = Direction of increasing x

At $x = -1$, $f'(x) = -2 = -2$ slope of tangent to the function at $x = -1$ = Direction of gradient = Direction of decreasing x

6) For multivariate function, such as $f(x) \times \mathbb{R}^n$

gradient of f $= \nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix} \rightarrow$ vector of gradients

7) Now we know that gradient points in the direction of steepest ascent. Therefore, gradient of a multivariate function gives the direction of steepest ascent.

To find out arbitrary directional rate of change, we limited to orthogonal or steepest ascent directions, we use "Directional gradient".

$D_v f(x_0) = \nabla f(x_0) \cdot v$ → Dot product

$\vec{v}$ → unit vector

$\nabla f(x_0)$ → gradient of f at $f(x_0)$

8) $\vec{v} \cdot \vec{w} = \|v\| \cdot \|w\| \cos\theta$; θ = angle between the vectors

∴ When $\vec{v} \cdot \vec{w} = 1$, $\theta = \frac{\pi}{2}$

9) Consider a set of vectors $\{\vec{v}_1, \vec{v}_2, \dots \vec{v}_n\}$;

Again $c_1\vec{v}_1 + c_2\vec{v}_2 + \dots c_n\vec{v}_n = c_1\vec{v}_1$

→ Any vector in the space can be described as combination of original set of vectors, where $c_1, c_2 \dots c_n$ are scalars.

10) Suppose there are m constraint equations; $g^j(x) : j = 1 \dots m$, $x \in \mathbb{R}^n$ then

$$J(x) = \begin{bmatrix} \frac{\partial g^1}{\partial x_1} & \frac{\partial g^2}{\partial x_1} & \cdots & \frac{\partial g^m}{\partial x_1} \\ \frac{\partial g^1}{\partial x_2} & \frac{\partial g^2}{\partial x_2} & \cdots & \frac{\partial g^m}{\partial x_2} \\ \frac{\partial g^1}{\partial x_3} & \frac{\partial g^2}{\partial x_3} & \cdots & \frac{\partial g^m}{\partial x_3} \\ \frac{\partial g^1}{\partial x_4} & \frac{\partial g^2}{\partial x_4} & \cdots & \\ \vdots & \vdots & \cdots & \\ \frac{\partial g^1}{\partial x_n} & \frac{\partial g^2}{\partial x_n} & \cdots & \frac{\partial g^m}{\partial x_n} \end{bmatrix}_{n \times m}$$

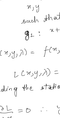
Full rank of Jacobian simply means that

1) The matrix is invertible

2) There is no redundancy in constraint functions

3) No one equation cannot be written in terms of another.

Back to Lagrangian formulation



Consider a function
$f(x, y) = \sin(\sqrt{x^2 + y^2})$

When you cut the above graph by a plane parallel to X-Y plane at some particular value of c, and project the top view on X-Y plane, then you get level curve.

$$\nabla f(x, y) = \begin{pmatrix} \frac{\partial f(x,y)}{\partial x} \\ \frac{\partial f(x,y)}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} \cos\left(\sqrt{x^2+y^2}\right) \\ \frac{y}{\sqrt{x^2+y^2}} \cos\left(\sqrt{x^2+y^2}\right) \end{pmatrix}$$

Now the above gradient gives us the direction of the steepest ascent is the direction for increasing the value of the f by maximum amount at that point.

To find arbitrary directional rate of change we use directional gradient

$D_v f(x, y) = \nabla f(x, y) \cdot \vec{v}$

When $D_v f(x, y) = 0$, $\Rightarrow$ rate of change for that arbitrary direction is 0

This is possible on a level curve, as you can travel on a level curve without changing "c" value.

Thus when you travel on a tangential vector at any point on a level curve, there is no change in $f(x, y)$.

∴ gradient of function ⊥ Tangent vector of function = 0

Direction of maximum change ⊥ Direction of no change = 0

Now, you look at the following diagram,

When the level curve & constraint function touch at (x_0, y_0), they will share a tangent.

The normal to that tangent will the directional gradient of the objective function and constraint function.

$$-[\nabla f(x^*) / \nabla x^*] = \sum_{j=1}^{m} \lambda_j^* [\nabla g^j(x^*) / \nabla x^*], \text{ for } j = 1, \dots m$$

This is the geometric and intuitive understanding of Lagrange multiplierS.

Ex:- $\max_{x,y} \; xy$

such that

$g_1 : x + y - 10 = 0$

$\mathcal{L}(x, y, \lambda_1) = f(x, y) + \lambda_1 \mathcal{L} g_1(x, y) - 0]$

$\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda_1 g_1(x, y) - 0]$

Finding the stationary points of the above equation

$\frac{\partial \mathcal{L}}{\partial x} = 0 \;\therefore\; y + \lambda_1 = 0 \quad y = -\lambda_1$

$\frac{\partial \mathcal{L}}{\partial y} = 0 \quad \therefore x + \lambda_1 = 0 \quad x = -\lambda_1$

$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 \quad \therefore 0 + x + y - 10 = 0 \quad \therefore x + y = 10$

$\therefore -\lambda_1 + (-\lambda_1) = 10$

$-2\lambda_1 = 10 \quad \boxed{\lambda_1 = -5}$

$\therefore x = y = 5$; $\lambda_1 = -5$

To calculate function whether it is maximum or minimum or saddle point, we find the Hessian matrix.