Theorem 21.1

If $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ is a convex function,

Then Ω is convex.

To complete the proof, proceed by contradiction.

Suppose that Ω is not convex. Then for $y_1, y_2 \in \Omega$, and $\alpha \in (0,1)$

$$w = \alpha y_1 + (1-\alpha) y_2 \notin \Omega$$

$\notin$ not belongs to

Theorem 21.1

If $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ is a convex function,

Then Ω is convex.

To complete the proof, proceed by contradiction.

Suppose that Ω is not convex. Then for $y_1, y_2 \in \Omega$, and $\alpha \in (0, 1)$

$$w = \alpha y_1 + (1-\alpha) y_2 \notin \Omega$$

$\notin$ → not belongs to

Now, let, $f(y_1) = \beta_1$ and $f(y_2) = \beta_2$

$\forall\ y_1, y_2 \in \Omega$; $\beta_1\ \beta_2 \in \mathbb{R}$

Then,

$$\text{graph } f = \left\{ (y_1, \beta_1)^T, (y_2, \beta_2)^T \right\}$$

~~where~~ which also belongs to epigraph of f

$$\text{epi}(f) = \left\{ (x, \beta)^T : f(x) \leq \beta \right\}$$

Now,

$$z = \alpha\,(y_1, \beta_1)^T + (1-\alpha)\,(y_2, \beta_2)^T$$

$$= \begin{pmatrix} \alpha . y_1 + (1-\alpha) . y_2 \\ \alpha \beta_1 + (1-\alpha) \beta_2 \end{pmatrix}$$

$$= \begin{pmatrix} \omega \\ \alpha\beta_1 + (1-\alpha)\beta_2 \end{pmatrix} \notin \text{epi}(f)$$

since, $\omega \notin \Omega$

This, implies that epi(f) is not a convex set. Therefore, f is not a convex function, which is a contradiction.

Theorem 21.2:

A function $f: \Omega \to \mathbb{R}$ defined as a convex set $\Omega \subseteq \mathbb{R}^n$ is convex if and only if

$$f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$$

$$\forall\ x, y \in \Omega, \text{ and } \alpha \in (0, 1)$$

Proof:

Final ; suppose that, $f(\alpha x + (1-\alpha)y) \leq$

$\alpha f(x) + (1-\alpha) f(y)$.

we need to show that f is convex function.

Let, $(x, \alpha)^T$ and $(y, \beta)^T \in epi(f)$, $\alpha, \beta \in \mathbb{R}$.

Then we have $f(x) \leq \alpha$, $f(y) \leq \beta$

Now, $f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$

$\leq \lambda \alpha + (1-\lambda)\beta$.

This, implies that

$\lambda x + (1-\lambda) y \in \Omega$ and hence f is convex function.

$\begin{pmatrix} f(\lambda x + (1-\lambda)y) \\ \lambda \alpha + (1-\lambda)\beta \end{pmatrix} \in epi f$

Converscely, suppose that f is convex function.

We have to show that,

$$f(\alpha x+(1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y); \forall x, y \in \Omega;$$
$$\alpha \in (0,1)$$

Let, $x, y \in \Omega$ and $f(x) = a$, $f(y) = b$

Then, $\begin{pmatrix} x \\ a \end{pmatrix}, \begin{pmatrix} y \\ b \end{pmatrix} \in epi(f)$

The proof will be complete, if we can show that $epi(f)$ is a convex set.

$$\alpha \begin{pmatrix} x \\ a \end{pmatrix} + (1-\alpha) \begin{pmatrix} y \\ b \end{pmatrix} = \begin{pmatrix} \alpha x + (1-\alpha) y \\ \alpha a + (1-\alpha) b \end{pmatrix} \in epi(f)$$

This implies that,

$$f(\alpha x + (1-\alpha) y) \le a\alpha + (1-\alpha) b$$

$$= \alpha f(x) + (1-\alpha) f(y)$$

i.e. $f(\alpha x + (1-\alpha) y) \le \alpha f(x) + (1-\alpha) f(y)$
which completes the proof.

Ex. 21.4.

Let, $f(x) = x_1 x_2$. Is f convex over $\mathbb{R}^2 \to \mathbb{R}$

$\Omega = \{x : x_1 \geq 0, x_2 \geq 0\}$?

Ans:

The answer is negative. To show the counter example, consider

$x = (\underset{x_1}{1}, \underset{x_2}{2})^T$ and $y = (2,1)^T \in \Omega$

Note;

$f(\alpha x + (1-\alpha) y)$

$\alpha \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (1-\alpha) \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

$= \begin{pmatrix} \alpha + 2 - 2\alpha \\ 2\alpha + 1 - \alpha \end{pmatrix}$

$= \begin{pmatrix} 2-\alpha \\ \alpha + 1 \end{pmatrix}$

CamScanner

Hence,

$$f(\alpha x + (1-\alpha) y) = (2-\alpha)(\alpha+1)$$

$$= 2 + \alpha - \alpha^2 \quad \text{--- (1)}$$

and, $\alpha f(x) + (1-\alpha) f(y) = \alpha f\begin{pmatrix}1\\2\end{pmatrix} + (1-\alpha) f\begin{pmatrix}2\\1\end{pmatrix}$

$$= \alpha \cdot 1 \cdot 2 + (1-\alpha) \cdot 2 \cdot 1$$

$$= 2\alpha + 2 - 2\alpha$$

$$= 2 \quad \text{--- (2)}$$

From, eq (1) and (2) we can see that,

$$f(\alpha x + (1-\alpha) y) \geq \alpha f(x) + (1-\alpha) f(y)$$

Thus, f is not convex over Ω.

Theorem 21.3

Q. What do you mean by convex optimization problem?

Theorem 21.5:

Let, $f: \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ be a convex function on a convex set Ω. Then, a point is a global ~~opti~~ minimizer of f over Ω if and only if it is a local minimizer.

Suppose that, x^* is not a global minimizer. Then for same $y \in \Omega$, we have,

$$f(y) \leq f(x^*)$$

But, f is convex:

$\therefore \forall \alpha \in (0,1)$, $f(\alpha y + (1-\alpha) x^*) \leq \alpha f(y) + (1-\alpha) f(x^*)$

$\leq f(x^*)$

Hence, $\exists$ points that are arbitrary close to x^* and have objective value.

consider a sequence of points $\{y_n\}$ given by $y_n = \frac{1}{n} y + \left(1 - \frac{1}{n}\right) x^*$ converges to x^* and $f(y_n) < f(x^*)$

which implies that x^* is not a local minimizer.

$\{x_n\}$ converge to x^* means $\lim_{n \to \infty} x_n = x^*$

Lagrange condition:

Consider the optimization problem

$$\text{minimize } f(x)$$
$$\text{subject to } h(x) = 0$$
$$g(x) \leq 0 \text{ , where}$$

$f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$; $g: \mathbb{R}^n \rightarrow \mathbb{R}^p$

Def: ***

Any point satisfying the constraints are called feasible point. The set of feasible points can be defined by

$$\Omega = \{x : h(x) = 0 \text{ ; } g(x) \leq 0\}$$

$f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$

$$h(x) = [h_1(x), h_2(x) \dots h_m(x)]^T$$

regular point: ***

$$\text{minimize } f(x)$$

$$\text{s.t } h(x) = 0$$

if x^* only satisfy $h(x)=0$ তবে x^* regular point.

$$\nabla h(x) = \left[\nabla h_1(x), \nabla h_2(x) \dots \nabla h_m(x)\right]^T$$

if $\nabla h_1(x) \dots \nabla h_m(x)$ are linear independent

Regular point; ~~Linear independent~~

Def: A point x^* is called regular point if it satisfyies the equality constraint $h(x)=0$ and The gradient vectors $\nabla h_1(x), \nabla h_2(x) \dots h_m(x)$ are linearly independent.

Linear independent:

$u_1, u_2, ..., u_n \in V$

k - scalar

$a_1 \; a_2 \cdots a_n \in k$

$v = a_1u_1 + a_2u_2 \cdots a_nu_n.$

if $a_1u_1 + a_2u_2 + \cdots a_nu_n = 0$
then we can called v is linearly in dependent.

$\{x(t)\}$ → continuous curve, differiable

$t \in (a, b)$

$x(t) = [x_1(t) \quad x_2(t)]^T$ $\quad x : \mathbb{R}^2 \to \mathbb{R}$

$x(t^*) = x^*$, $t^* \in (a, b)$

$x'(t^*) \neq 0$

$\nabla h(x^*)$ is orthogonal to $x'(t^*)$

prove

h → constant

curve

defined by

$h(x(t)) = 0$

$\forall t$??

respect a

differentiate

করো।

$\cancel{\nabla h(x(t)) \cdot \frac{d}{dt}(x(t))}$

$\nabla h(x(t))^T \cdot x(t) = 0$

$\nabla h(x^*)^T \cdot x'(t^*) = 0$

at x^*

$\phi(t) = f(x(t)) = 0$

$\phi'(t) = 0$

$\nabla f(x(t))^T \cdot x'(t) = 0$

$\nabla f(x^*)^T x'(t^*) = 0$

$\frac{d}{dt} h(x(t)) = 0$

Lagrange theorem ***

Let, x^* be a regular point and x^* be the minimizer of $f: \mathbb{R}^2 \to \mathbb{R}$ s.t $h(x) = 0$, $h: \mathbb{R}^2 \to \mathbb{R}$

Then, $\nabla f(x^*)$ and $\nabla h(x^*)$ are parallel i.e. if $\nabla h(x^*) \neq 0$

Then, $\exists$ a scalar $\lambda^* \in \mathbb{R}$ s.t.

$$\boxed{\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0}$$

$f : \mathbb{R}^n \to \mathbb{R}^m \quad \exists \; \lambda^* \in \mathbb{R}^m$

$h : \mathbb{R}^n \to \mathbb{R}^m$

$$\boxed{Df(x^*) + \lambda^{*T} Dh(x^*) = 0}$$

Lagrange function:

A function $l : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ is said to be a langrange function if it can be defined by

$$l(x, \lambda) = f(x) + \lambda^T h(x), \quad x \in \mathbb{R}^n, \; \lambda \in \mathbb{R}^m$$

if x^* is a local minimizer, Then

$$D\,l(x^*, \lambda^*) = 0$$

$$\Rightarrow Df(x^*) + \lambda^T Dh(x^*) = 0$$

$$D_x l(x, \lambda) = Df(x) + \lambda^T Dh(x)$$

$$D_\lambda l(x, \lambda) = h(x)^T$$

$$D_x l(x^*, \lambda^*) = 0 \quad \text{--- i)}$$

$$D_\lambda l(x^*, \lambda^*) = 0 \quad \text{--- ii)}$$

These two equations are called lagrangere equations.

Problem: 19.6 :

Consider the problem of extremizing the objective function.

$$f(x) = x_1^2 + x_2^2$$

on the ellipse

$$\{(x_1, x_2)^T : h(x) = x_1^2 + 2x_2^2 - 1 = 0\}$$

Solve:

$$\nabla f(x) = \begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix}^T$$

$D^T f(x)$

$= \nabla f(x)$

$= \begin{pmatrix} \nabla f(x_1) \\ \nabla f(x_2) \end{pmatrix}$

$$\nabla h(x) = \begin{bmatrix} 2x_1 & 4x_2 \end{bmatrix}$$

$$D_x\, l(x, \lambda) = D f(x) + \lambda^T . Dh(x)$$

$$= \begin{pmatrix} 2x_1 + 2\lambda x_1 & 2x_2 + 4\lambda x_2 \end{pmatrix}$$

Here,

$$D f(x) = \begin{pmatrix} D f(x_1) \\ D f(x_2) \end{pmatrix}^T = \begin{pmatrix} 2x_1 & 2x_2 \end{pmatrix}$$

$$\lambda^T D h(x) = \begin{pmatrix} \lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} 2 & \lambda_1 \\ x_1 & x_2 \end{pmatrix}$$

$$= \begin{pmatrix} 2\lambda_1 x_1 & 4\lambda x_2 \end{pmatrix}$$

$$D_\lambda l(x,\lambda) = h(x)^T = x_1^2 + 2x_2^2 - 1$$

Lagrange eq.

$$D_x l(x,\lambda) = 0$$

and $D_\lambda l(x,\lambda) = 0$

$$2x_1 + 2\lambda x_1 = 0 \quad \text{—(i)}$$ is valid for $x_1 = 0$ or $\lambda = -1$

$$2x_2 + 4\lambda x_2 = 0 \quad \text{—(ii)}$$

$$x_1^2 + 2x_2^2 - 1 = 0 \quad \text{—(iii)}$$

x_1, x_2, λ এর মান বের করে function এ মান বসাতে হবে।

$x_1 = 0$	$\lambda = -1$ হলে
$2x_2^2 = 1$	$x_2 = 0$
$x_2 = \pm \frac{1}{\sqrt{2}}$	$x_1 = \pm 1$

$$x^1 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad x^2 = \begin{pmatrix} 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \quad x^3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad x^4 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

$f(x_1) = ?$

$f(x_2) = ?$

$f(x_3) = ?$

$f(x_4) = ?$

যেটাতে maximum, যেটাতে minimum. আর।

minimum value at x_1 & x_2

maximum $\rightarrow$ x_3 and x_4

1\. Optimization Problems and its classifications

Jacobian, Hessian matrix and some problems

2 FONC & SONC and their properties. with some examples.

local minimizers and global and global minimizer and some related results and

Numerical examples

3 Unimodal function

and Golden section search method

Gradient method, specially steepest descent method

4 Convex function and convex optimization

lagrange conditions, lagrange functions and lagrange equations.

KKT condition.

KKT condition:

consider the optimization problems,

minimize $f(x)$

subject to the constraints

$$\left.\begin{aligned} h(x) &= 0 \\ g(x) &\leq 0 \end{aligned}\right\}$$

$$\boxed{\begin{aligned} f &: \mathbb{R}^n \to \mathbb{R} \\ h &: \mathbb{R}^n \to \mathbb{R}^m \\ g &: \mathbb{R}^n \to \mathbb{R}^p \end{aligned}}$$

$h_i(x) = 0$
$g_j(x) \leq 0$

Def:

An inequality constraints $g_j(x) \leq 0$ is said to be active if $g_j(x) = 0$ and inactive if $g_j(x) < 0$.

Note :

Index set;

$$J(x) = \{ j : g_j(x) = 0 \}$$

$i \in I$, integer

it is also called active index set.

Th. 20.1 : (KKT)

Let, $f, h, g \in C^1$, let x^* be a regular point and a local minimizer for the problem of minimizing f s.t $h(x) = 0$, $g(x) \leq 0$. Then $\exists \ \lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^p$ such that :

1. $\mu^* \geq 0$

2. $Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T$

3. $\mu^{*T} g(x^*) = 0$

4. $h(x^*) = 0$

5. $g(x^*) \leq 0$

Page: 404:

Ex : 20.4:

Consider the optimization problem

$$\text{minimize } f(x_1, x_2)$$
$$\text{s.t. } x_1, x_2 \geq 0$$

where $f(x_1, x_2) = x_1^2 + x_2^2 + x_1 x_2 - 3x_1$

$f: \mathbb{R}^n \to \mathbb{R}$

Find the minimizer using KKT conditions.

Solve:

1. $\mu^* \geq 0$

$\mu^* = [\mu_1, \mu_2]^T \leq 0$

$g(x^*) = x$

$\mu^{*T} g(x^*$

2. $Df(x^*) + \mu^{*T} = 0$

3. $\mu^{*T} x = 0$

4. $x \geq 0$

$Df(x) + \mu^T = 0$

(2) $Df(x) + \mu^T = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix}^T + [\mu_1 \ \ \mu_2]$

$= \begin{bmatrix} 2x_1 + x_2 - 3 , & x_1 + 2x_2 \end{bmatrix}^T + [\mu_1 \quad \mu_2]^T$

$$\left.\begin{aligned} 2x_1 + x_2 + \mu_1 &= 3 && \text{(i)} \\ x_1 + 2x_2 + \mu_2 &= 0 && \text{(ii)} \\ \cancel{\mu_1 \ \mu_2 +} & \\ \mu_1 x_1 + \mu_2 x_2 &= 0 && \text{(iii)} \end{aligned}\right\}$$

(3) $\mu^T x = 0$

$\begin{pmatrix} \mu_1 & \mu_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

x_1 x_2 μ_1 μ_2

Let μ_1 be the free variable and choose $\mu_1 = 0$, then

From 3rd eq. $x_2 = 0$

Then, from 1st and 2nd equation gives

$x_1 = 3/2$, $\mu_2 = -3/2$

CS CamScanner