Optimization - 4

Agenda :-

1. Recap
2. Constrained optimization
3. GD $\longrightarrow$ code
4. Computing gradients of function

* Recap :-

$$\min_{\vec{w}, w_0} \left[-\sum y_i * \left(\frac{w^T x + w_0}{\|\vec{w}\|} \right) \right]$$

$\downarrow$

* Gradient Descent

$$\boxed{w_i^{t+1} = w_i^t - \eta * \frac{\partial f}{\partial w_i^t}}$$

$$f(w_0, w_1, w_2 \cdots w_n) = -\sum y_i * \left(\frac{w^T x_i + w_0}{\|\vec{w}\|} \right)$$

$$\frac{\partial f}{\partial w_1} = \frac{f(w_1+\Delta, w_2, w_3 \cdots w_n) - f(w_1, w_2 \cdots w_n)}{\Delta}$$

$$f(w_0, w_1 \cdots w_n) = \boxed{-\sum y_i * \left(\frac{w^T x + w_0}{\|\vec{w}\|}\right)}$$

$$= -\sum y_i * \left(\frac{[w_1 x_1 + w_2 x_2 \cdots + w_n x_n] + w_0}{\|\vec{w}\|}\right)$$

$$= -\sum y_i * \left(\frac{w_1 x_1 + w_2 x_2 \cdots + w_n x_n + w_0}{\sqrt{w_1^2 + w_2^2 + \cdots + w_n^2}}\right)$$

$$f = -\sum y_i * \left[\frac{h(w_i)}{g(w_i)}\right]$$

h/g

Idea

$\rightarrow$ Computationally expensive.

Soln :-

eq :- $$\frac{\overset{w_1}{3x} + \overset{w_2}{4y} + \overset{w_0}{4}}{||w||}$$

$$||w|| = \sqrt{w_1^2 + w_2^2}$$

$$= \sqrt{3^2 + 4^2} = \boxed{5}$$

$$\Rightarrow \frac{3x + 4y + 4}{||\vec{w}||}$$

$$= \frac{3x + 4y + 4}{5}$$

$$= \left(\frac{3}{5}\right)x + \left(4/5\right)y + \left(4/5\right)$$

w_1' $\quad$ w_2' $\quad$ w_0'

$$= \boxed{w_1' x + w_2' y + w_0'}$$

$$w_1' = \frac{w_1}{\|w\|} \qquad w_2' = \frac{w_2}{\|w\|} \qquad w_0' = \frac{w_0}{\|w\|}$$

$$h / g \rightarrow 1 \qquad \Rightarrow$$

$$\min_{\vec{w}, w_0} \; - \sum y_i \left(\frac{w^T x + w_0}{\boxed{\|\vec{w}\|}} \right)$$

$$\|\vec{w}\| \rightarrow 1$$

$$f/g$$

$$\min_{\vec{w}, w_0} = - \sum y_i (w^T x + w_0)$$

<u>such that</u> $\|\vec{w}\| = 1$ → <u>condition</u>

Constrained Optimization

$$f = -\sum y_i * (w^T x_i + w_0)$$

$$f = -\sum y_i * (w_1 x_1 + \cancel{w_2 x_2} + \cdots + \cancel{w_n x_n} + \cancel{w_0})$$

0 0

$$\boxed{\frac{\partial f}{\partial w_1} = -\sum y_i * x_1}$$

$$\boxed{\frac{\partial f}{\partial w_2} = -\sum y_i * x_2}$$

$\vdots$

$$\boxed{\frac{\partial f}{\partial w_j} = -\sum y_i * x_j}$$

$$w_j^{t+1} = w_j^t - \eta * \frac{\partial f}{\partial w_j^t}$$

$$\boxed{w_j^{t+1} = w_j^t - \eta * \left(-\sum y_i * x_j^i\right)}$$

$j \neq 0$

1

$$f = -\sum y_i * (\cancel{w_1 x_1} + \cancel{w_2 x_2} + \cdots + \cancel{w_n x_n} + w_0)$$

$$\underset{\downarrow}{w_0}$$
$$1$$

$$\boxed{\frac{\partial f}{\partial w_0} = -\sum y_i}$$

$$w_0^{t+1} = w_0^t - \eta * \frac{\partial f}{\partial w_0^t}$$

$$\boxed{w_0^{t+1} = w_0^t - \eta * \left(-\sum y_i\right)} \rightarrow 2$$

* Computing gradients of a function

(Q) $f(x_1, x_2, x_3) = a_1 x_1 + a_2 x_2 + a_3 x_3$

where, $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$

$$\nabla_{\vec{x}} f(\vec{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \\ \frac{\partial f}{\partial x_2} \\ \\ \frac{\partial f}{\partial x_3} \end{bmatrix}$$

$$f = a_1 x_1 + a_2 \cancel{x_2} + a_3 \cancel{x_3}$$

$$\boxed{\frac{\partial f}{\partial x_1} = a_1}$$

$$\frac{\partial f}{\partial x_2} = a_2$$

$$\frac{\partial f}{\partial x_3} = a_3$$

$$\nabla_{\vec{x}} f = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \vec{a}$$

* <u>general</u>

$$f = \vec{a} * \vec{x}$$

$$\nabla_{x} f = \vec{a}$$

$$f = a * x$$

$$\frac{df}{dx} = a$$

eg 2 :

$$f(x_1, x_2 \cdots x_d) = \vec{x}^T * \vec{x}$$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

$$\nabla_{\vec{x}} f = ?$$

Soln :

$$f = \vec{x}^T * \vec{x}$$

$$= \begin{bmatrix} x_1 & x_2 & \cdots & x_d \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

$$= x_1 * x_1 + x_2 * x_2 \cdots + x_d * x_d$$

$$f = \underbrace{x_1^2 + x_2^2 \cdots + x_d^2}$$

$$\nabla_{\vec{x}} f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{x_d} \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ 2x_d \end{bmatrix}$$

$$= 2 * \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

$$\boxed{\nabla_x f = 2 * \vec{x}}$$ ✓

general :

$$f = x * x = x^2$$

$$\boxed{\frac{df}{dx} = 2x}$$ ✓

(*) general form of optimization problem.

optimize $\rightarrow \boxed{f(\theta)}$

such that $g_1(\theta)$ $\rightarrow$ $\|w\|=1$

$g_2(\theta)$

$\vdots$

eg:- $\min\limits_{\vec{w}, w_0} -\sum y_i * (w^T x_i + w_0)$

such that $\|w\| = 1$

(Q) Can above problem be solved using Gradient Descent?

→ NO

$$\boxed{w_{new} = w_{old} - \eta * \Delta_w f}$$

→ Iterative process

initially, $\|w\| = 1$

eg: $w_1 = 3/5$ ✓

$w_2 = 4/5$ ✓

$$
\begin{tikzcd}
\text{unconstrained} \arrow[d] \\
\text{Constrained} \arrow[d] \\
\text{unconstrained}'
\end{tikzcd}
$$

* Lagrange's Multiplier

eg:- Constrained prob :-

$$\min_{\theta} f(\theta)$$

st, $g_1(\theta) = 0$
$g_2(\theta) = 0$
$g_3(\theta) = 0$
⋮

$\|w\| = 1$

$\boxed{\|w\| - 1 = 0}$

$$\min_{\theta} \left\{ f(\theta) + \lambda_1 * g_1(\theta) + \lambda_2 * g_2(\theta) + \lambda_3 g_3(\theta) \right\}$$

Lagrange's Multiplier

eg :

$$L = \min_{x,y} \left(x^2 + y^2 \right)$$

Such that $x + 2y - 1 = 0$

$$L = \min_{x,y,\lambda} \left(x^2 + y^2 + \lambda \left(x + 2y - 1 \right) \right)$$

$$\frac{\partial L}{\partial x} = 2x + \lambda = 0$$

$$\boxed{\lambda = -2x} \longrightarrow 1$$

,

$$\frac{\partial L}{\partial y} = 2y + 2\lambda = 0$$

$$\boxed{\lambda = -y} \longrightarrow 2$$

$$\frac{\partial L}{\partial \lambda} = x + 2y - 1 = 0 \quad — 3$$

using 1 & 2 in 3

$\lambda = -2x$

$$\boxed{x = -\frac{\lambda}{2}}$$

$\lambda = -y$

$$\boxed{y = -\lambda}$$

$x + 2y - 1 = 0$

$-\frac{\lambda}{2} + 2(-\lambda) - 1 = 0$

$$-\lambda\left(\frac{1}{2}+2\right)-1=0$$

$$\lambda\left(1/2+2\right)=-1$$

$$\lambda\left(5/2\right)=-1$$

$$\boxed{\lambda=-2/5}$$

using this in eqn 1

$$y = -\lambda$$

$$\boxed{y = 2/5}$$ ✓

$$x = -\lambda/2$$

$$= -\left(-2/5\right) \times \frac{1}{2}$$

$$\boxed{x = 1/5}$$ ✓

(Q) $f(x_1, x_2, x_3) = \vec{x}^T * (\vec{x} + \vec{b})$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$$

value of gradient of $f(\vec{x})$ w.r.t $\vec{x}$?

$$\nabla_{\vec{x}} f(\vec{x}) = ?$$

$$f = \underline{\vec{x}^T *} \left(\vec{x} + \vec{b}\right)$$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \vec{x}^T = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$$

$$\vec{x} + \vec{b} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$$

$$\vec{x} + \vec{b} = \begin{bmatrix} x_1 + 1 \\ x_2 + 3 \\ x_3 + 5 \end{bmatrix}$$

$$f = \vec{x}^T * (\vec{x} + \vec{b})$$

$$= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} * \begin{bmatrix} x_1 + 1 \\ x_2 + 3 \\ x_3 + 5 \end{bmatrix}$$

$$= x_1 * (x_1 + 1) + x_2 * (x_2 + 3) + x_3 * (x_3 + 5)$$

$$f = \underbrace{\left[x_1^2 + x_1 + x_2^2 + 3x_2 + x_3^2 + 5x_3 \right]}$$

$$\nabla_{\vec{x}} f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 2x_1 + 1 \\ 2x_2 + 3 \\ 2x_3 + 5 \end{bmatrix}$$

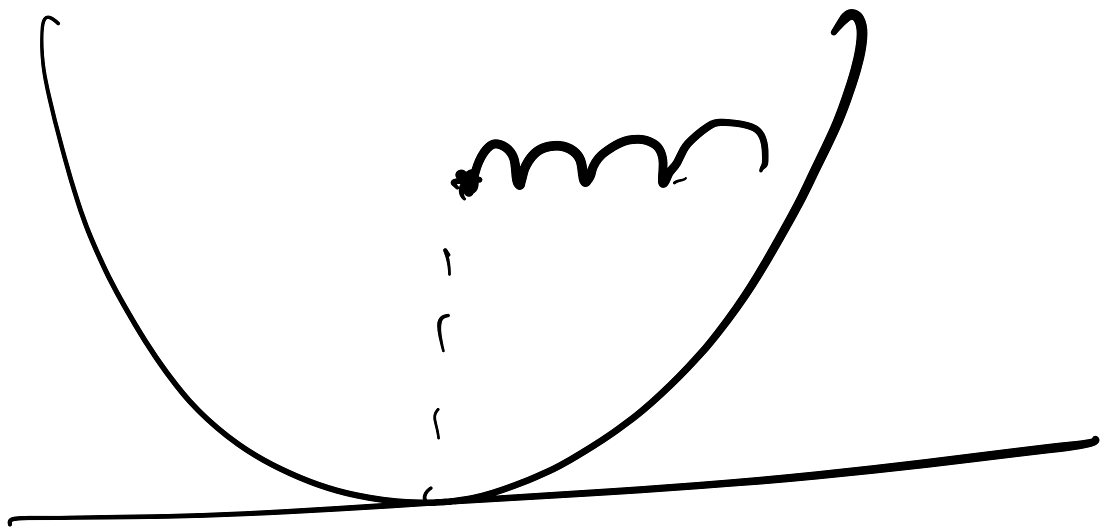
$$= 2\vec{x} + \vec{b}$$

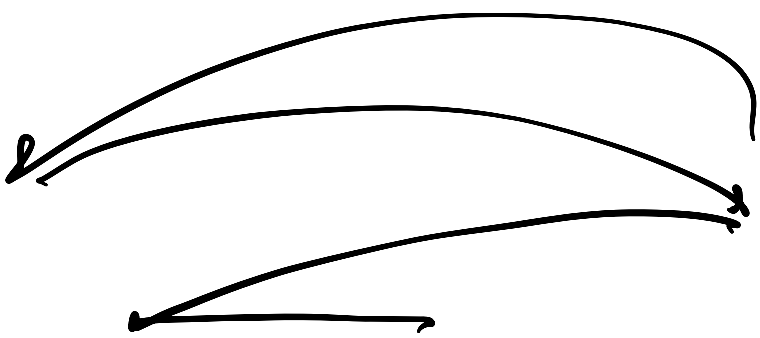
general :

$f = x(x+b)$

$f = x^2 + bx$

$$\boxed{\frac{df}{dx} = 2x + b}$$





Quiz :- $f = \underbrace{w_1 x_1 + w_2 x_2 + w_3 x_3}$

$$\nabla_{\vec{x}} f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{bmatrix}$$

$$= \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \quad \vec{w}$$