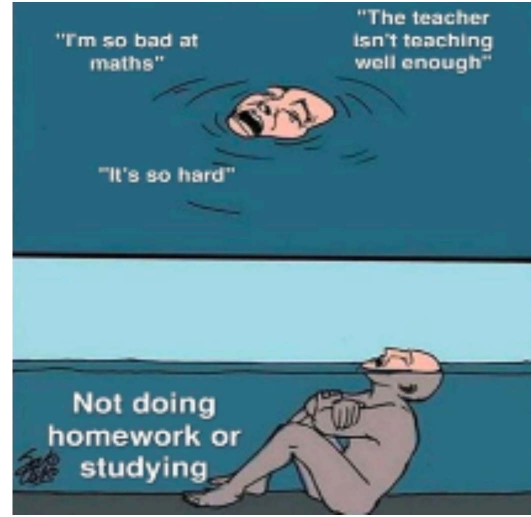
Constraint Optimisation



Agenda :

→ Recap

→ Lagrange Multipliers

→ Unconstraint Opt.

→ Constraint Opt.

Recap

Derivatives : $f(x) \rightarrow$ Continous & differentiable

$$\frac{d\,f(x)}{dx} = f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta y}{\Delta x}$$

Maxima/minima : To find Candidate Points for maxima & minima.

find x, wh: $\boxed{f'(x) = 0}$

if $f''(x) > 0$: minima

$f''(x) < 0$: maxima

Rules:

1. Linearity
2. Product
3. Quotient
4. Chain

Partial Derivatives:

$$\boxed{f(x) = x^2 + 2x}$$

function with multiple inputs.

$$f(x,y) = x^2 + y^2$$

$$\frac{d f(x)}{dx} \quad \text{Vs.} \quad \frac{\partial f(x,y)}{\partial x} \ \& \ \frac{\partial f(x,y)}{\partial y}$$

$f(w_1, w_2, w_0)$

$$\text{gradient} = \nabla_{\bar{w}} f(\bar{w})$$

$\hookrightarrow$ derivative w.r.t a vector of inputs.

Gradient Descent: $\rightarrow$

$\swarrow$

optimization

$\downarrow$

minima

1. init Variables $(\bar{w})$ randomly
2. repeate {

$$\bar{w} = \bar{w} - \underbrace{\eta}_{0.1} \cdot \boxed{\frac{\partial L}{\partial \bar{w}}}$$

}

$$\nabla_w L = \begin{bmatrix} \frac{\partial L}{\partial w_1} \\ \frac{\partial L}{\partial w_2} \end{bmatrix}$$

In what direction does the gradient of a function point at a specific point?

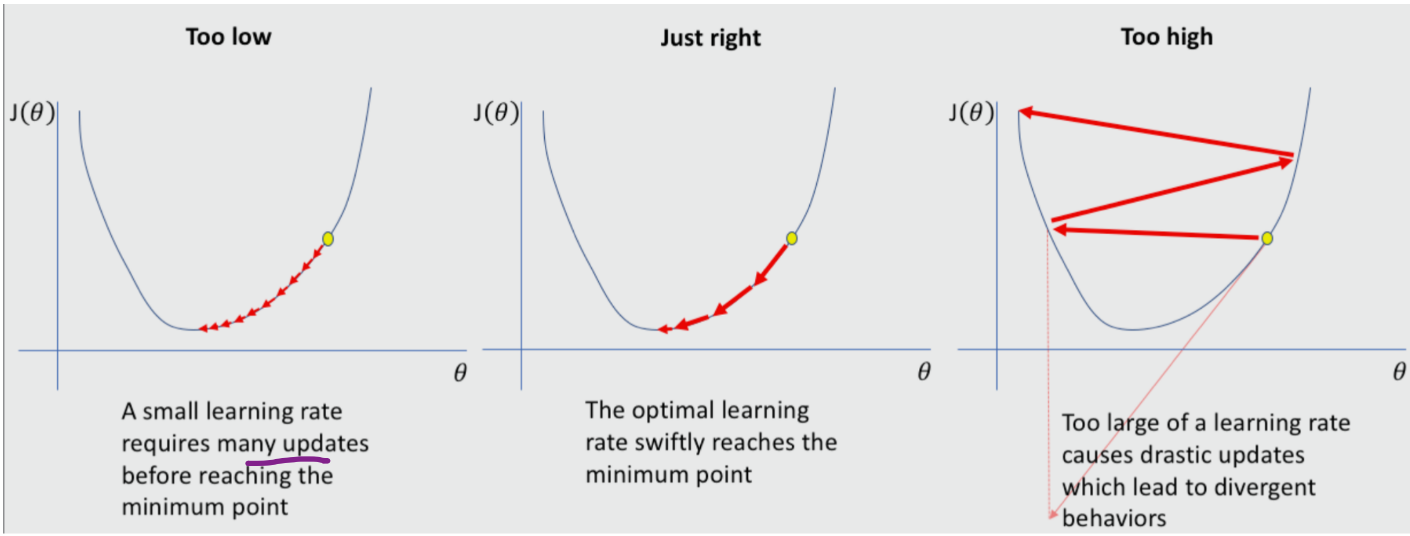
24 users have participated

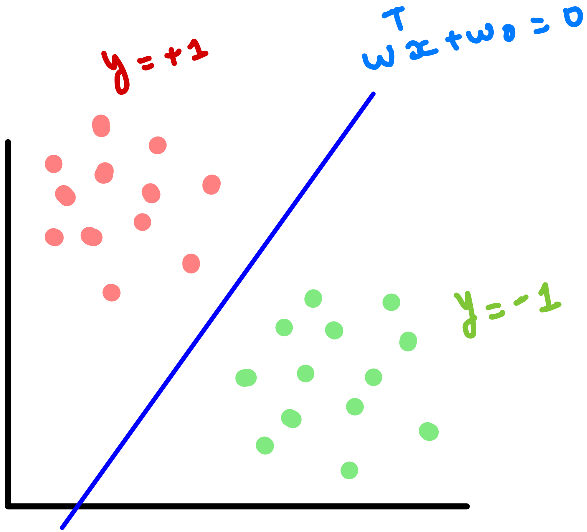
A	It points towards the nearest maximum value of the function.	0%
B	It points towards the nearest minimum value of the function.	25%
C	It points in the direction of the steepest increase of the function at that point.	75%
D	It points in the direction of the x-axis.	0%

End Quiz Now

η

0.000001   0.01   5.0





$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

$\downarrow$

Gain function

$\downarrow$

Loss function

$$w^*, w_0^* = \underset{w, w_0}{\operatorname{argmin}} \text{ Loss func}$$

Computing Gradient / Vector Calculus

$$f(x_1, x_2, x_3) = a^T x$$
$$= a_1 x_1 + a_2 x_2 + a_3 x_3$$

$$a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_1 \\ x_3 \end{bmatrix}$$

$$\nabla_{\bar{x}} f(x) = \begin{bmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \partial f / \partial x_3 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = a$$

$$\boxed{\begin{array}{l} f(x) = a^T x \\ \nabla_{\bar{x}} f(x) = \bar{a} \end{array}}$$

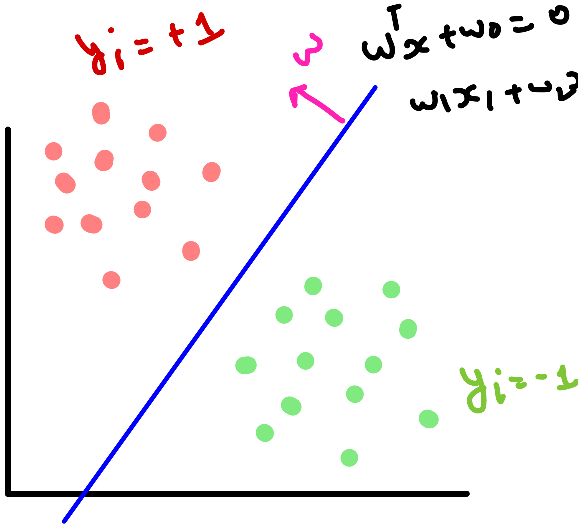
2. $f(x) = x^T x$

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

$$f(x_1, x_2 \dots x_d) = \sum_{i=1}^{d} x_i^2 = x_1^2 + x_2^2 + \dots x_d^2$$

$$\nabla_{\bar{x}} f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_d} \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_d \\ \vdots \\ 2x_d \end{bmatrix} = 2 \cdot \bar{x}$$

$$\boxed{\begin{aligned} f(x) &= x^T x \\ \nabla_x f &= 2\bar{x} \end{aligned}}$$



$w^T x + w_0 = 0$

$w_1 x_1 + w_2 x_2 + w_0 = 0$

$$\text{Gain}(w, w_0) = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{(w^T x^i + w_0)}{\|w\|} \cdot y^i$$

$\downarrow$

$$w^*, w_0 = \underset{w, w_0}{\text{argmax}} \ \text{Gain}(w, w_0)$$

$\downarrow$

$$\text{Loss}(w, w_0) = -\frac{1}{n} \sum_{i=1}^{n} \frac{(w^T x^i + w_0)}{\|w\|} \cdot y^i$$

$\downarrow$

$$\boxed{w^*, w_0^* = \underset{w, w_0}{\text{argmin}} \ -\frac{1}{n} \sum_{i=1}^{n} \frac{(w^T x^i + w_0)}{\|w\|} y^i}$$

$$w^*, w_0^* = \underset{w, w_0}{\operatorname{argmin}} -\frac{1}{n} \sum_{i=1}^{n} \frac{(w^T x^i + w_0)}{\|w\|} y^i$$

Apply Gradient Descent

1. randomly init w & w_0

2. repeat {

$$w = w - \eta \cdot \nabla_{\bar{w}} L(w, w_0)$$

$$w_0 = w_0 - \eta \cdot \frac{\partial L}{\partial w_0}$$

}

find

$$\bar{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_d \end{bmatrix}$$

```
def derivative(x, f):
  delta = 0.0001
  return (f(x + delta) - f(x)) / delta
```

$$L(w, w_0) = -\frac{1}{n} \sum_{i=1}^{n} \frac{(w^T x^i + w_0)}{\|w\|} \cdot y^i$$

s.t. $\boxed{\|w\| = 1} \rightarrow$ Constraint

such that

Q- $f(x) / y = x^2 - 3x - 3$

s.t $-x^2 + 2x + 3 = 0$

$x = -1 \quad f(x) = 7$

$x = 3 \quad f(x) = 3$

minima

Ans $\Rightarrow$ $\boxed{x = 3}$

Constraint optimization $\longrightarrow$ **Lagrange multipler** $\longrightarrow$ Unconstraint optimization

$$\underset{x}{\operatorname{argmin}}\ f(x) \quad \text{s.t. } g(x) = 0 \quad \Rightarrow \quad \underset{x,\lambda}{\operatorname{argmin}} \left[f(x) + \lambda g(x) \right]$$

λ $\downarrow$ lagrange multipler.

$$\underset{x}{\operatorname{argmin}}\ f(x)$$

s.t $g_1(x) = 0$

$g_2(x) = 0$ - - - - . $g_n(x) = 0$

$$\underset{x,\lambda_1,\lambda_2\ldots\lambda_n}{\operatorname{argmin}}\left[f(x)+\lambda_1\cdot g_1(x)+\lambda_2\cdot g_2(x)+\ldots+\lambda_n\cdot g_n(x)\right]$$

Example

Q-

$$f(x)=x^2-3x-3$$

$$\underset{x}{\operatorname{argmin}}\ f(x)$$

$$\text{s.t}\ -x^2+2x+3=0$$

$$\Downarrow$$

$$x^*,\lambda^*=\underset{x,\lambda}{\operatorname{argmin}}\left[(x^2-3x-3)+\lambda(-x^2+2x+3)\right]$$

$$\Downarrow \text{G.D}$$

$$\frac{\partial L}{\partial x} = 2x - 2\lambda x + 2\lambda - 3 = 0$$

$$\frac{\partial L}{\partial \lambda} = -x^2 + 2x + 3 = 0$$

$$\Rightarrow -x^2 + 3x - x + 3 = 0$$

$$= x(-x+3) + 1(-x+3) = 0$$

$$(x+1)(-x+3) = 0$$

$$\begin{bmatrix} x = -1 \\ x = 3 \end{bmatrix}$$

$$L(w, w_0) = -\frac{1}{n}\sum_{i=1}^{n} \frac{(w^T x^i + w_0)}{\|w\|} \cdot y^i$$

S.t. $\|w\| = 1$ $\longrightarrow$ Constraint

such that

$g(x) = 0$

$$\boxed{\sqrt{w^T w} - 1 = 0}$$

$$\Downarrow$$

$$w^*, w_0^*, \lambda^* = \underset{w, w_0, \lambda}{\operatorname{argmin}} -\frac{1}{n}\sum_{i=1}^{n} (w^T x^i + w_0) \cdot y^i + \lambda\left(\sqrt{w^T w} - 1\right)$$

$\downarrow$ Apply G.D

1. init randomly w, w_0, λ

2. repeat {

$$w = w - \eta \cdot \nabla_w L$$

$$w_0 = w_0 - \eta \cdot \frac{\partial L}{\partial w_0}$$

$$\lambda = \lambda - \eta \cdot \frac{\partial L}{\partial \lambda}$$

}

$$\frac{\partial L}{\partial \lambda} = \frac{\partial}{\partial \lambda} \lambda \left(\underbrace{\sqrt{w^T w} - 1}_{\text{const}} \right) \Rightarrow \sqrt{w^T w} - 1 = 0$$

$$w^T w - 1 = 0$$

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

$$\frac{\partial L}{\partial w} = -\frac{1}{n} \sum \left[\frac{\partial\, w^T x \cdot y}{\partial w} + \cancel{\frac{\partial\, w_0 \cdot y}{\partial w}} \right] + \lambda \cdot \frac{\partial (w^T w)^{1/2}}{\partial w} - \cancel{\frac{\partial \lambda}{\partial w}}$$

$$= -\frac{1}{n} \sum_{i=1}^{n} x_i \cdot y_i + \lambda \cdot \frac{1}{2} \frac{1}{\sqrt{w^T w}} \cdot \frac{\partial w^T w}{\partial w}$$

$$\lambda \cdot \frac{\cancel{2} \bar{w}}{\cancel{2} \|w\|}$$

$$\left[\frac{\partial L}{\partial w} = -\frac{1}{n} \sum_{i=1}^{n} x_i \cdot y_i + \frac{\lambda \bar{w}}{\|w\|} \right]$$

If the function is represented as f(x)=w1x1+w2x2+w3x3, what would be the gradient of this function with respect to x?

$\frac{\partial f}{\partial x}$

24 users have participated

- **A** t would be a scalar value. — 8%
- **B** It would be a matrix. — 4%
- **C** It would be a derivative of the function with respect to x — 33%
- **D** It would be a vector, w=[w1,w2,w3]. — 54%

End Quiz Now

Pytorch | tf

"$\boxed{\text{autograd}}$"