Constraint optimization

Recap

$$\min_{\vec{w}, w_0} - \sum y_i \cdot \frac{w^T x_i + w_0}{\|w\|}$$

Gradient Descent

learning rate (control)

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

$$f(w_0, w_1, \ldots, w_n) = - \sum_{i=1}^{n} y_i \cdot \frac{w^T x_i + w_0}{\|w\|}$$

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

$$\Delta \sim 0$$

$$(y_i - \hat{y}_i)^2$$

$$\frac{\partial f}{\partial w_1} = - \sum y_i \cdot \frac{(w_1 x_1 + w_2 x_2 + \dots + w_0)}{\sqrt{w_1^2 + w_2^2 + \dots + w_n^2}}$$

$$= - \sum y_i \cdot \frac{f(w_i)}{g(w_i)}$$

$$\Downarrow$$

Computationally expensive

Idea!

$$\frac{3x + 4y + 4 = 0}{||w||}$$

$$w_1 x_1 + w_2 x_2 + w_0 = 0$$

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

$$\sqrt{3^2 + 4^2} = \sqrt{25} = 5$$

$$\frac{3}{5}x + \frac{4}{5}y + \frac{4}{5} = 0$$

$$\frac{3}{5} \downarrow w_1' \qquad \frac{4}{5} \downarrow w_2' \qquad \frac{4}{5} \downarrow w_0'$$

$$w_1' x_1 + w_2' x_2 + w_0' = 0$$

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

$$w_0' = \frac{w_0}{\|w\|}$$

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

$$\text{s.t.} \quad \underbrace{\|w\| = 1}_{\text{Condition}}$$

$$\frac{\partial f}{\partial w_1} \Rightarrow \frac{d}{dw_1} - \Sigma\, y_i \,(\cancel{w_1} x_1 + \overset{0}{\cancel{w_2 x_2}} + \overset{0}{\cancel{\cdots}} + \overset{0}{\cancel{w_d}})$$

$$\Rightarrow -\Sigma\, y_i x_1$$

$$\boxed{w_j^{(t+1)} = w_j^{(t)} - \eta\left(-\sum_{i=1}^{n} y_i x_i\right)}$$

$j \neq 0$

$$\frac{\partial J}{\partial w_0} = -\sum y_i$$

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

General form of optimization problem

min/max

optimisation fn $\rightarrow$ $f(\theta)$

θ

such that,

constraint fn $\rightarrow$ $g_1(\theta)$, $g_2(\theta)$

$$\downarrow$$

$$\|w\| = 1$$

$$\downarrow$$

$$\|w\| - 1 = 0$$

$$\min_{\theta_1, \theta_2} f(\theta_1, \theta_2)$$

$$\text{s.t.} \quad g_1(\theta_1), \; g_2(\theta_2)$$

ex:

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

$$\text{s.t.} \quad \|w\| - 1 = 0$$

Q ⇒ Can above problem be solved by Gradient Descent ?

$$w_{new} = w_{old} - \eta \nabla_{\vec{w}} f$$

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

w_1, w_2

$\frac{3}{5}, \frac{4}{5}$

0.55 ← $w_1 = w_1 - \frac{\partial f}{\partial w_1}$

0.71 ← $w_2 = w_2 - \frac{\partial f}{\partial w_2}$

⇒ Lagrange Multiplier

constraint problem ⟶ unconstraint problem

$$\min - \sum y_i \frac{w_1 x_1 + w_2 x_2 + \cdots w_0}{\|w\|}$$ → unconstraint

⇓

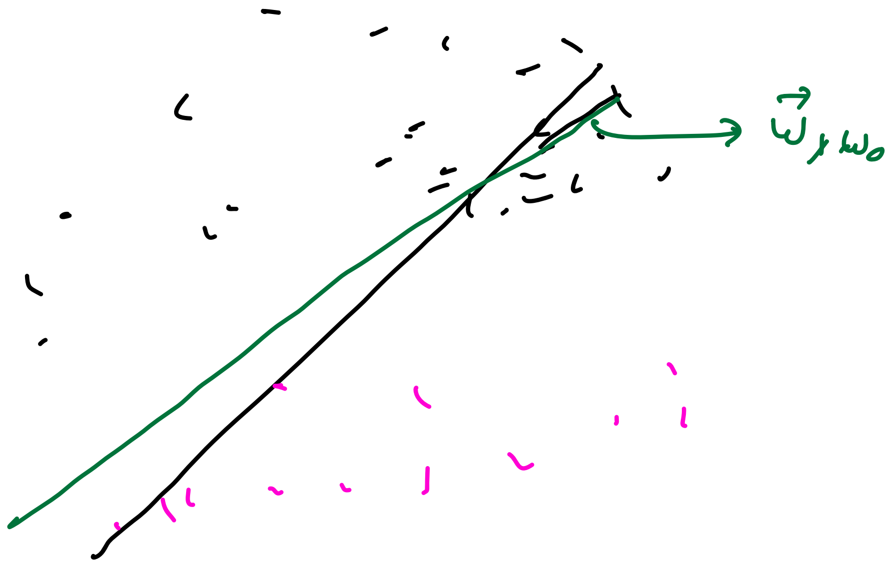
$$\min - \sum y_i (w^T x + w_0)$$
$$\text{s.t} \quad \|w\| = 1$$ → constraint prob

⇓

unconstraint prob → easier to solve

⇒ Break: 8:05 am

$\vec{w}, w_0$



Contraint prob

$$\left\{ \min_{\theta} f(\theta) \quad ; \quad s.t. \quad \begin{array}{l} g_1(\theta) = 0 \\ g_2(\theta) = 0 \\ g_3(\theta) = 0 \end{array} \right.$$

$$\Downarrow$$

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

lagrange multiplier

Ex1

$$\min_{x,y} \quad x^2 + y^2$$

$$s.t. \quad x + 2y - 1 = 0$$

$$\Downarrow$$

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

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

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

$$\frac{\partial L}{\partial \lambda} = x + 2y - 1$$

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

$$x + 2(2x) - 1 = 0$$

$$x = 1/5$$

$$y = 2/5$$

$$\lambda = -2/5$$

Intuition

$$\min_{x,y,\lambda} \quad L = f(x,y) + \lambda\,(g(x,y))$$

if $\lambda >> 0$, $g(x,y) > 0$

GD problem

$$w_i = w - \frac{\partial L}{\partial w_i}$$

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

$$\Rightarrow \quad \min_{\vec{w}, w_0} \; -\sum y_i (w^T x_i + w_0)$$

$$\text{s.t.} \quad ||w|| - 1 = 0$$

$$\mathcal{L} \Rightarrow \min_{\vec{w}, w_0, \lambda} \; -\sum y_i \sum w^T x_i + w_0) + \lambda\,(||w|| - 1)$$

Dimention Reduction

$\Downarrow$

PCA

Deep Learning

Q $\Rightarrow$ $f(x_1, x_2, x_3) \Rightarrow x^T(x+b)$

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

Value of gradient of $f(x_1, x_2, x_3)$ w.r.t. vector x ?

$\Rightarrow x^T(x+b)$

$$x^T = [x_1 \quad x_2 \quad x_3]$$

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

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

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

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

$$f \Rightarrow x_1^2 + x_1 + x_2^2 + 3x_2 + x_3^2 + 5x_3$$

$$\frac{\partial f}{\partial x_1} \Rightarrow 2x_1 + 1$$

$$Q2 \Rightarrow f(x_1, x_2) = x_2 \cdot e^{x_1} - 4 \cdot x_2^2$$

$f(x_1, x_2)$ w.r.t. x_1

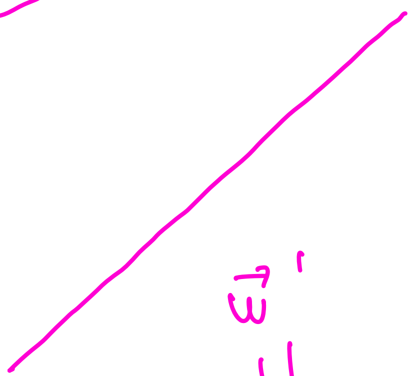
$$\frac{\cancel{\partial f(x_1, x_2)}}{\partial x_1}$$

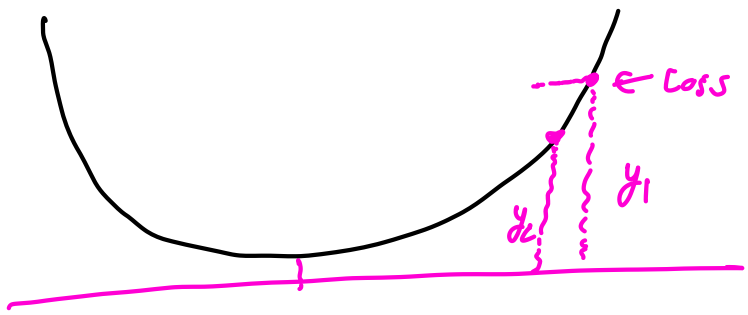
$$\frac{\partial e^{x_1}}{dx_1} = e^{x_1}$$

$$\frac{d}{dx_1}\left(x_2 \cdot e^{x_1}\right) = x_2 \cdot e^{x_1} + x_2' \cdot e^{x_1}$$

$$\frac{d}{dx_1}\left(-4x_2^{2}\right) = 0$$

$\underbrace{w_0, \vec{w}}$





$\vec{w}'$

$\Downarrow$

$\|\vec{w}'\| = 1$