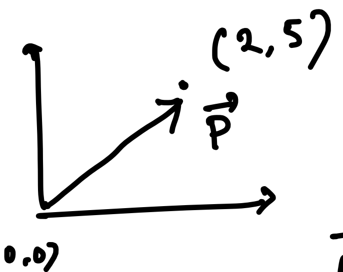
Sep, 19

Support Vector Machines (SVM)
Support Vector Networks

Vladimir Vapnik (1995 – 1999)

- linear classification
- Kernel tricks.
- is a part of max - margin models

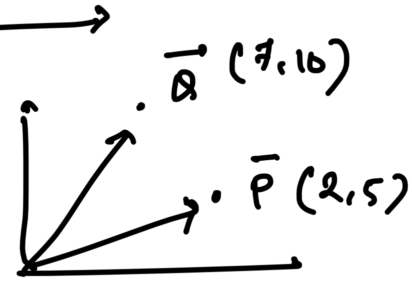
Coordinate Geometry



magnitude = $\sqrt{2^2+5^2} = \sqrt{29}$.

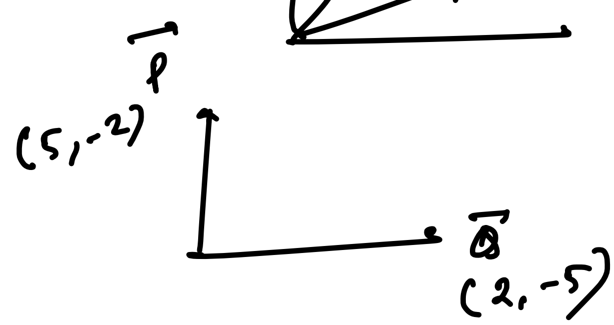
Angle / direction = $\tan\theta = y/x$

$\theta = \tan^{-1}(y/x)$



$\overrightarrow{PQ} = ((7-2), (10,5))$

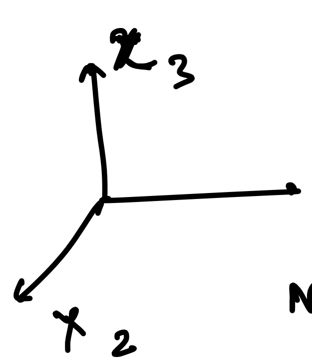
$= (5,5)$



$\bar{P}\cdot\bar{Q} = 0$ perpendi

$= 10-10 = 0$

Plane



Eq. of a plane:

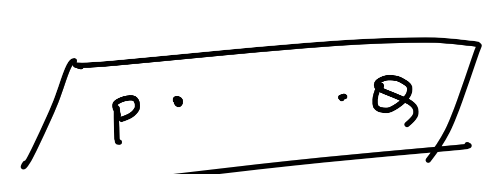
$w_1x_1 + w_2x_2 + w_3x_3 + \underline{w_0} = 0$

$w^T x + w_0 = 0$

Normal vector of the plane:

(w_1, w_2, w_3)

$w^T x + w_0 = 0$



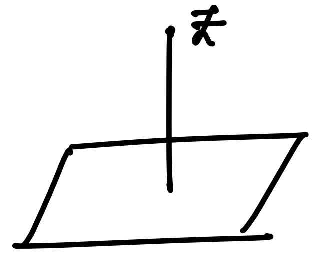
$w^T P + w_0 = 0$ $w^T Q + w_0 = 0$

$w^T P - w^T Q = 0$

$w^T(\underline{\vec{P} - \vec{Q}}) = 0$

distance of a point from the plane

z



$$\frac{w^T z + w_0}{\|w\|} = d = \gamma$$

= geometric margin

Functional margin = $w^T z + w_0 = \hat{\gamma}$

$$\boxed{\gamma = \frac{\hat{\gamma}}{\|w\|}}$$

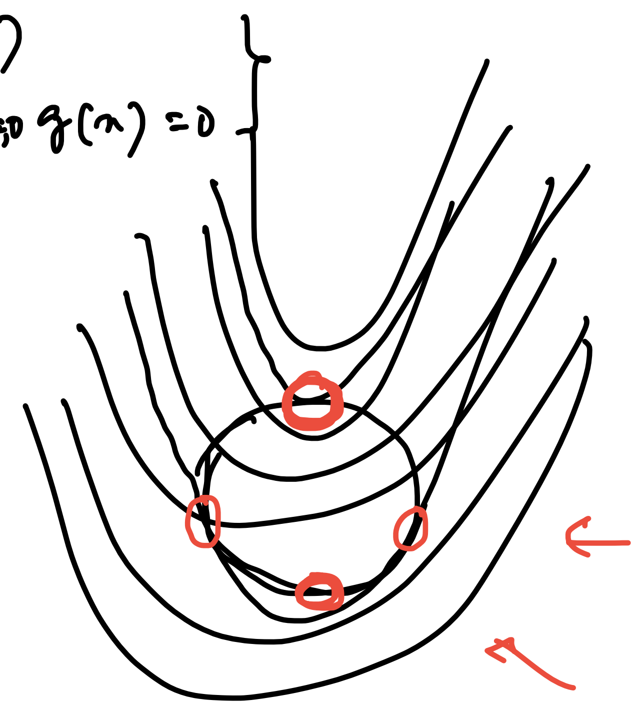
$$d = \begin{cases} +ve : & \text{if } z \text{ is in the directn of the normal vector.} \\ -ve : & \text{otherwise.} \end{cases}$$

Lagrange Multiplier

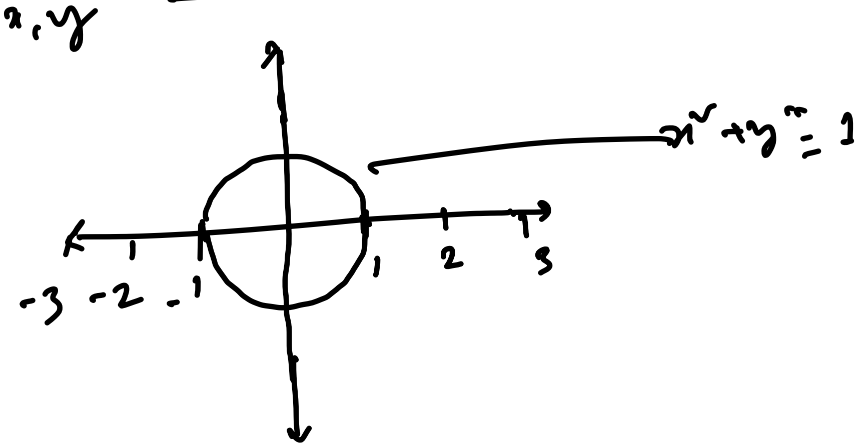
Linear Regression : Sq loss

logistic " : log loss

$$\left.\begin{array}{l} \max f(x) \\ h(x) \le 0 \quad g(x) = 0 \end{array}\right\}$$



E.g. $\max_{x,y} \; z = x^2 y$ s.t. $x^2 + y^2 = 1$



1) Convert a constraint O.F to an unconstraint O.F.

$f(h,s)$

$H(h,s)$

$$L(h,s,\lambda) = f(h,s) - \lambda H(h,s)$$

λ = Lagrange multiplier.

$$= f(h,s) - \lambda_1 H_1(h,s) - \lambda_2 H_2(h,s)$$

n

1

$$\left\{\begin{array}{l} \dfrac{\partial L}{\partial h} = 0 \\[2ex] \dfrac{\partial L}{\partial s} = 0 \\[2ex] \dfrac{\partial L}{\partial \lambda} = 0 \end{array}\right.$$

E.g. $\underline{f}(h,s) = 200\, h^{2/3} s^{1/3}$ s.t. $H(h,s) \Rightarrow 20h + 170s = 20{,}000$

$$L(h,s,\lambda) = 200\, h^{2/3} s^{1/3} - \lambda\,(20h + 170s - 20{,}000)$$

$$\left.\begin{aligned} \frac{\partial L}{\partial h} &= 200 \cdot \frac{2}{3} h^{-1/3} s^{1/3} - \lambda \cdot 20 = 0 \\ \frac{\partial L}{\partial s} &= 200\, \frac{1}{3} h^{2/3} s^{-2/3} - 170\lambda = 0 \\ \frac{\partial L}{\partial \lambda} &= -20h - 170s + 20000 = 0 \end{aligned}\right\}$$

$h = 666.66$; $s = 39.12$; $\lambda = 3.59$

Max $f = 51777$

if two graphs are tangent at the point then their normal vectors must be parallel. — the two normal vectors must be scalar multiples of each other.

$$\nabla f(\ ) = \lambda \quad g(\ )$$

$$\nabla f(x,y,z) = \lambda\, g(x,y,z)$$

$$\underline{\nabla f_x = \lambda \nabla g_x} \qquad \underline{\nabla f_y = \lambda \nabla g_y} \qquad \nabla f_z = \lambda \nabla g_z$$

$$\nabla f_x - \lambda \nabla g_x = 0$$

$f(x)$ $x^2 + y^2 - 4 = 0$ $\qquad g(x):$ $\frac{x^2}{4} + y^2 - 1 = 0$
