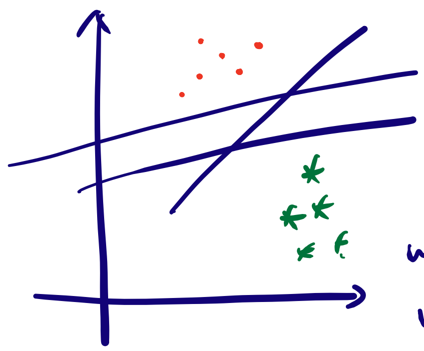
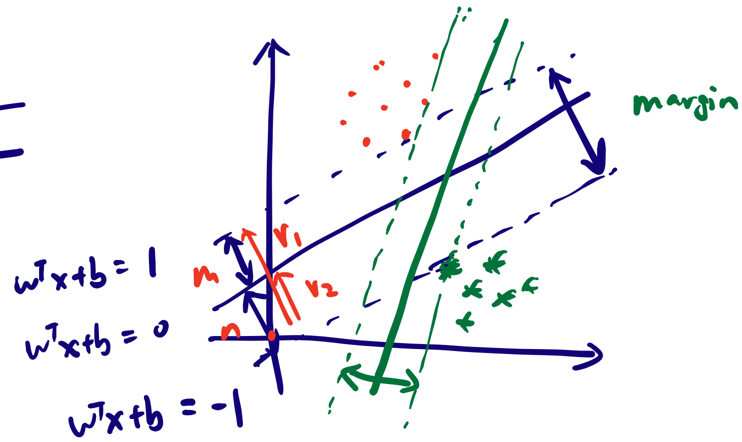
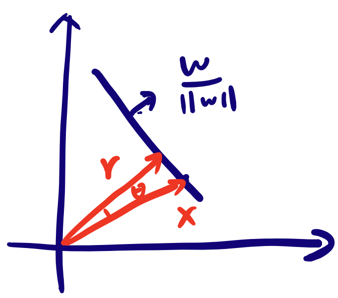
Oct 10, 2024





Maximum margin



proj x to the direction of $\hat{w}$

$$\vec{x}\cdot\vec{w} = \|\vec{w}\|\cdot\underbrace{\|\vec{x}\|\cdot\cos\theta}_{r}$$

$$r = \frac{\vec{x}\cdot\vec{w}}{\|w\|}$$

$$m = n = \frac{1}{\|w\|} \qquad r_1 - r_2 = \frac{1-b}{\|w\|} - \frac{-b}{\|w\|} = \frac{1}{\|w\|}$$

$$\text{margin} = \frac{2}{\|w\|} = m+n$$

SVM

$$\max \frac{2}{\|w\|} \Rightarrow \min \frac{\|w\|^2}{2}$$

subject to : $y_n(w^Tx_n+b) \geq 1 \quad n=1,\cdots N$

Optimization

$$\min\ f(x,y) = 2 - x^2 - 2y^2$$

$$\text{s.t. } h(x,y) = x + y - 1 = 0$$ equality constraint

Lagrange multiplier, β

$$\min\ L(x,y,\beta) = f(x,y) + \beta\, h(x,y)$$

$$\frac{\partial L(x,y,\beta)}{\partial x} = -2x + \beta = 0$$

$$\frac{\partial L(x,y,\beta)}{\partial y} = -4y + \beta = 0$$

$$\frac{\partial L(x,y,\beta)}{\partial \beta} = x + y - 1 = 0$$

$\beta = 2x = 4y$

$x = 2y$

$3y - 1 = 0$

$$x = \frac{2}{3} \qquad y = \frac{1}{3} \qquad \beta = \frac{4}{3}$$

$$\min\ f(x,y) = x^3 + y^2$$

$$\text{s.t. } g(x) : x^2 - 1 \leq 0$$ inequality constraints

$$\min\ L(x,y,\alpha) = f(x,y) + \alpha \cdot g(x)$$

$$\text{s.t. } \alpha \geq 0$$

$$\frac{\partial L(x,y,\alpha)}{\partial x} = 3x^2 + 2\alpha x = 0$$

$$\frac{\partial L(x,y,\alpha)}{\partial y} = 2y = 0$$

$$\frac{\partial L(x, y, \alpha)}{\partial \alpha} = x^2 - 1 = 0$$

Lagrange Multipliers

$$\min f(w)$$

$$\text{s.t.} \quad g_i(w) \le 0 \qquad i = 1 \cdots k$$

$$h_j(w) = 0 \qquad j = 1, \cdots l$$

$$\Rightarrow \quad L(w, \alpha, \beta) = f(w) + \sum_{i=1}^{k} \alpha_i g_i(w) + \sum_{j=1}^{L} \beta_j h_j(w)$$

$$\text{s.t.} \quad \alpha_i \geq 0$$

primal formulation

$$\theta_p(w) = \max_{\alpha, \beta\, \alpha_i \ge 0} L(w, \alpha, \beta)$$

$$\theta_p(w) = \begin{cases} f(w), & \text{constraints are satified.} \\ & g_i(w) \le 0, \quad h_j(w) = 0 \\ \infty, & \text{otherwise.} \end{cases}$$

$$\rightarrow p^* = \min_{w} \theta_p(w) = \min_{w} f(w) = \min_{w} \max_{\alpha, \beta, \alpha_i \ge 0} L(w, \alpha, \beta)$$

dual formulation

$$\theta_d(\alpha, \beta) = \min_{w} L(w, \alpha, \beta)$$

$$\rightarrow d^* = \max_{\alpha, \beta\, \alpha_i \ge 0} \theta_d(\alpha, \beta) = \max_{\alpha, \beta, \alpha_i \ge 0} \min_{w} L(w, \alpha, \beta)$$

$$d^* \le p^* \qquad \text{"} \max_{x} . \min_{y} \text{"} \le \text{"} \min_{y} \max_{x} \text{"}$$

$$d^* = p^* \ ?$$

when $\begin{cases} f(w) & \text{convex} \\ g_i(w) & \text{convex} \\ h_i(w) & \text{linear} \end{cases}$

KKT condition. $d^* = p^* = L(w^*, \alpha^*, \beta^*)$

SVM optimization satisfies KKT conditions

$$\frac{\partial L(w^*, \alpha^*, \beta^*)}{\partial w} = 0$$

$$\frac{\partial L(w^*, \alpha^*, \beta^*)}{\partial \beta_j} = 0 \qquad j = 1 \cdots l$$

$\rightarrow$ $\alpha_i^* g_i(w^*) = 0 \qquad i = 1, \cdots k$

$g_i(w^*) \leq 0 \qquad i = 1 \cdots k$

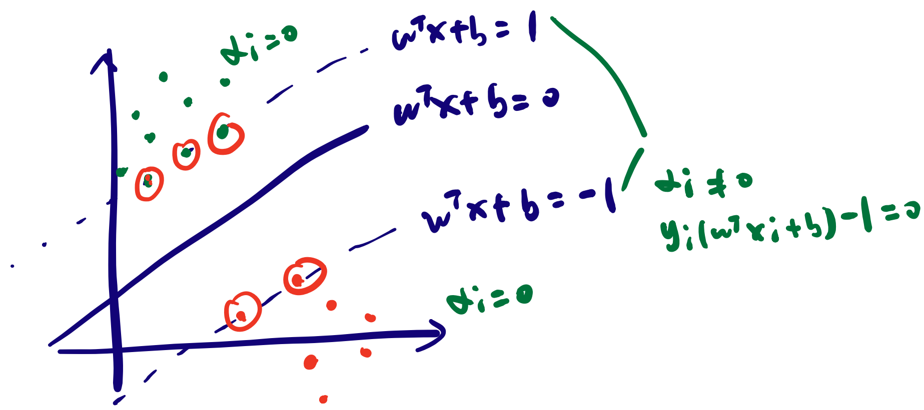
$\alpha_i^* \geq 0 \qquad i = 1 \cdots k$

$\alpha_i^* g_i(w^*) = 0 \Rightarrow$ either $\alpha_i^* = 0$

$g_i(w^*) = 0$

if $\alpha_i^* \neq 0$, $\alpha_i^* > 0 \Rightarrow g_i(w^*) = 0$

if $g_i(w^*) \neq 0 \Rightarrow \alpha_i^* = 0$



KKT
$$\begin{cases} \alpha_i \left(y_i(w^T x_i + b) - 1 \right) = 0 \\ y_i(w^T x_i + b) - 1 \geq 0 \qquad g_i(w^*) \geq 0 \\ \alpha_i \geq 0 \end{cases}$$

$\alpha_i \neq 0 \Rightarrow y_i(w^T x_i + b) - 1 = 0$ Support vectors

$\underline{y_i(w^T x_i + b) - 1 \neq 0} \Rightarrow \underline{\alpha_i = 0}$

$\tilde{x}$. new testing sample

$\underline{w^T \tilde{x} + b}$, only need to compare with the Support vectors

> 0

< 0

SVM

$$\min_{w, b} \frac{1}{2} w^T w$$

s.t. $\underline{y_i(w^T x_i + b) \geq 1}$, $i = 1 \cdots N$, $g_i(w) \leq 0$

$$\rightarrow L(w, b, \alpha) = \frac{1}{2} w^T w + \sum_{i=1}^{N} \alpha_i \left(1 - y_i(w^T x_i + b) \right)$$

s.t. $\alpha_i \geq 0 \quad i = 1 \cdots N$

$\rightarrow$ $w^* = \sum_{i=1}^{N} \alpha_i y_i x_i$ 　　 many $\alpha_i = 0$

when $y_i(w^T x_i + b) - 1 = 0$

$w^T \tilde{x} + b = \sum_{i=1}^{N} \alpha_i y_i x_i \tilde{x} + b$

replace $k(x_i, \tilde{x}) = \exp(-\frac{\|x_i - \tilde{x}\|^2}{2\sigma^2})$

$x_i \tilde{x} = k(x_i, \tilde{x})$

inner product

$x \Rightarrow \phi(x) = \begin{bmatrix} x \\ x^2 \\ \vdots \\ x^d \end{bmatrix}$

$\langle x, z \rangle \Rightarrow \langle \phi(x), \phi(z) \rangle$

$k(x, z)$

non-seperable

kernel trick

Gaussian kernel

$$k(x, z) = \exp\left(-\frac{\|x - z\|^2}{2\sigma^2}\right)$$