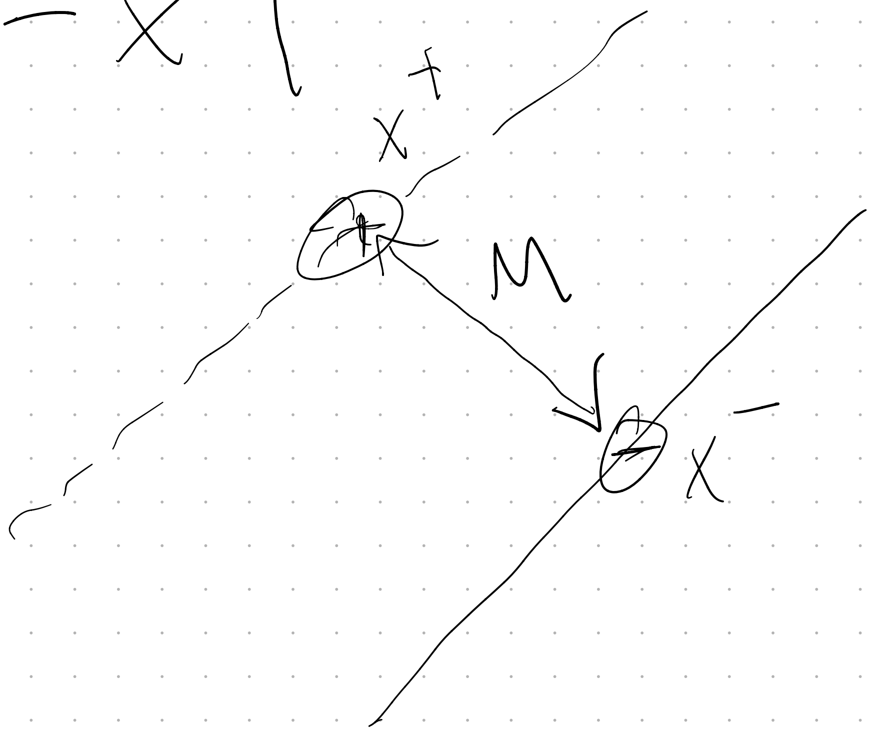
$M = |x^+ - x^-|$



$\lambda w = M$

$M = 2 / \sqrt{w^T w}$

① correctly classify all points

② maximize $M = 2 / \sqrt{w^T w}$

minimize $\frac{1}{2} w^T w$

$\min \ w^T w/2$

subj to

$$w^T x_i + b \geq 1 \quad \text{for } y_i = 1$$

$$w^T x_i + b \leq -1 \quad y_i = -1$$

$$y_i(w^T x_i + b) \geq 1$$

$$y_i(w^T x_i + b) \geq 1$$

same

$\min \ w^T w/2$

subj to

$$y_i(w^T x_i + b) \geq 1$$

Dual form

Lagrange multipliers

$$\min \frac{w^T w}{2}$$

subj to $y_i(w^T x_i + b) \geq 1$

lagrange multipliers

inequalities less than eq 0 $g(w)$

equalities eq 0 $h(w)$

$$\mathcal{L}(w, \alpha, \beta) = f(w) + \sum_{i=1}^{k} \alpha_i \underbrace{g_i(w)}_{\text{ineq.}} + \sum_{i=1}^{l} \beta_i \underbrace{h_i(w)}_{\text{eq.}}$$

$$\min_{w} \max_{\alpha \beta, \ \alpha \geq 0} \mathcal{L} = \begin{cases} f(w) & \text{if const sat.} \\ \infty & \text{if const. not satisfied} \end{cases}$$

① $$\mathcal{L}(w,b,\alpha) = \frac{1}{2}w^Tw - \sum_i \underbrace{\alpha_i}_{\text{non zero only for supp vecs}} \underbrace{\left[y_i(w^Tx+b)-1\right]}_{\text{0 for supp vecs}}$$

② $$\nabla_w \mathcal{L}(w,b,\alpha) = w - \sum_i \alpha_i y_i x_i = 0$$

$$\Rightarrow w = \sum_i \alpha_i y_i x_i$$

③ $$\nabla_b \mathcal{L} = \sum_i \alpha_i y_i = 0$$

Facts 1 & 2

$$\mathcal{L}(w,b,\alpha) = \frac{1}{2}\sum_{ij} \alpha_i \alpha_j y_i y_j x_i x_j$$

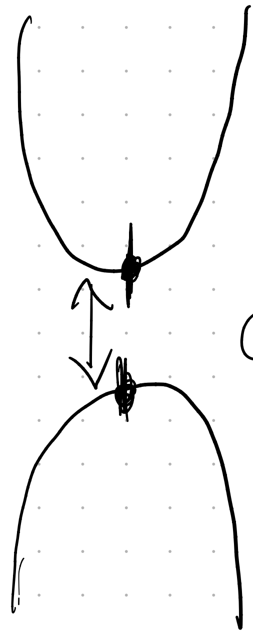
$$- \sum_i \alpha_i \left[y_i \left(\underbrace{\left[\sum_j \alpha_j y_j x_j\right]}_{w} x_i + b \right) - 1 \right]$$

$$= \sum_i \alpha_i - \frac{1}{2}\sum_{ij} y_i y_j \alpha_i \alpha_j x_i^T x_j - b \underbrace{\sum_i \alpha_i y_i}_{\text{Fact 3} \; \sum \alpha_i y_i = 0}$$

$$= \sum_i \alpha_i - \frac{1}{2}\sum_{ij} y_i y_j \alpha_i \alpha_j \boxed{x_i^T x_j}$$

$$\text{st} \quad \alpha_i \geq 0$$

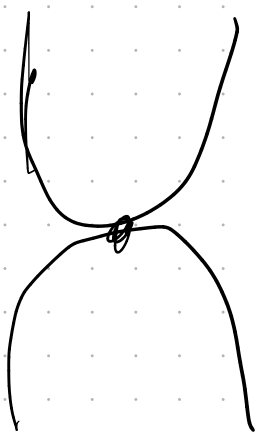
$$\sum \alpha_i y_i = 0$$



primal form

duality gap

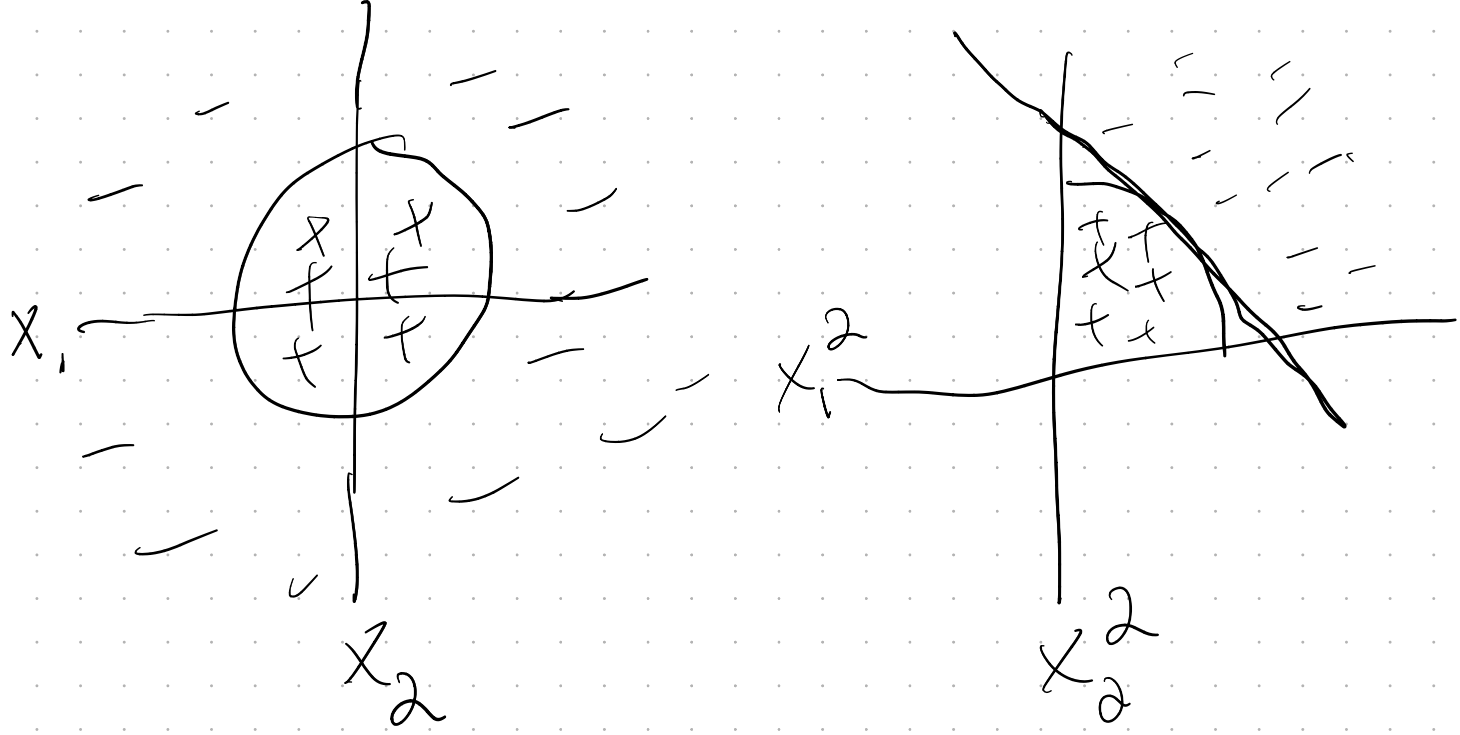
dual form



$$wx + b \Rightarrow \sum_i \alpha_i y_i \langle x_i, x \rangle + b$$

only over supp vecs

inner prod



Kernel transforms data into a new space where it may be linearly sep.

$$\phi(x) = x^2$$

$$\phi(x) = x^2$$

$$\max_{\alpha} \mathcal{L} = \sum_i \alpha_i - \frac{1}{2} \sum_{ij} y_i y_j \alpha_i \alpha_j x_i x_j$$

w/ Kernel

$$= \sum_i \alpha_i - \frac{1}{2} \sum_{ij} y_i y_j \alpha_i \alpha_j \underline{\phi(x_i)\phi(x_j)}$$

$$K(x, z) \rightarrow \phi(x)\phi(z)$$

$$= \sum_i \alpha_i - \frac{1}{2} \sum_{ij} y_i y_j \alpha_i \alpha_j K(x_i, x_j)$$

new example

$K(x,z) = (x^T z)^2$ p terms

p dim x, z

can be shown to be eq. to

$$(x^T z)^2 = \underbrace{\sum_{i,j=1}^{p} x_i x_j z_i z_j}_{p^2 \text{ terms}}$$

$$\phi(x) = \begin{bmatrix} x_1 x_1 \\ x_1 x_2 \\ \vdots \\ x_3 x_3 \end{bmatrix}$$

p^2 terms

$p=3$

$p^2=9$

Gaussian / RBF

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

equal the infinite sum of all polynomial kernels of all degree

---

classification w/ kernels

s = # sup vecs

$$w = \sum_{k=1}^{s} y_k \alpha_k \phi(x_k$$

$$b = y_l - w\,\phi(x_l)$$

l index of a supp vec

$$\text{sign}(wx+b)$$

copy from prev page

$$= \text{sign}(w\phi(x)+b)$$

$$= \text{sign}\left(\sum_{k=1}^{S} y_k \alpha_k \underbrace{\phi(x_k)\phi(x)}_{\text{innerprod}} + b\right)$$

$$= \text{sign}\left(\sum_{k=1}^{S} y_k \alpha_k K(x_k, x) + b\right)$$