1.

$z_1 = (-2, -2)$ -1

$z_2 = (4, -5)$ -1

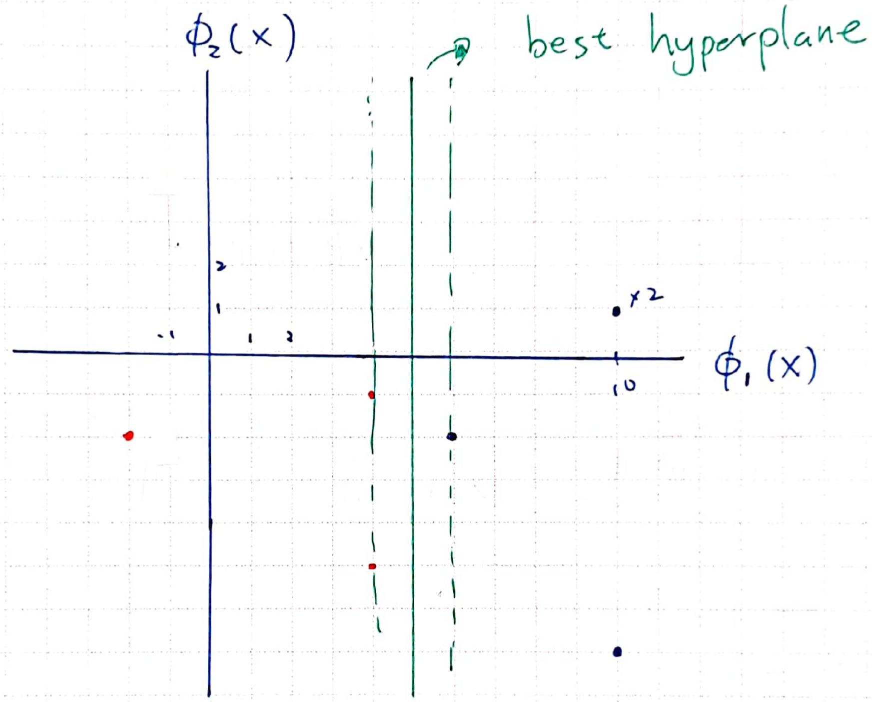
$z_3 = (4, -1)$ -1

$z_4 = (6, -2)$ 1

$z_5 = (10, -7)$ 1

$z_6 = (10, 1)$ 1

$z_7 = (10, 1)$ 1



$\phi_1(x) = 5$ is the best hyperplane since it expands the thickest margin.

6.

$$\cos(x, x') = \frac{x^T x'}{\|x\| \cdot \|x'\|}$$

Mercer condition

let $K_{ij} = \cos(x_i, x_j)$

$$K = \begin{bmatrix} \frac{x_1^T x_1}{\|x_1\| \|x_1\|} & \frac{x_1^T x_2}{\|x_1\| \|x_2\|} & \cdots & \frac{x_1^T x_N}{\|x_1\| \|x_N\|} \\ \vdots & & \ddots & \vdots \\ \frac{x_N^T x_1}{\|x_N\| \|x_1\|} & & & \frac{x_N^T x_N}{\|x_N\| \|x_N\|} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{x_1}{\|x_1\|} & \frac{x_2}{\|x_2\|} & \cdots & \frac{x_N}{\|x_N\|} \end{bmatrix}^T \begin{bmatrix} \frac{x_1}{\|x_1\|} & \frac{x_2}{\|x_2\|} & \cdots & \frac{x_N}{\|x_N\|} \end{bmatrix}$$

$$= ZZ^T$$

this is always P.S.D and symmetric

let

$$Z = \begin{bmatrix} \frac{x_1^T}{\|x_1\|} \\ \frac{x_2^T}{\|x_2\|} \\ \vdots \\ \frac{x_N^T}{\|x_2\|} \end{bmatrix}$$

$\cos(x, x')$ is a valid kernel

5. $\exp(-x^2) = \dfrac{1}{\exp(x^2)} \overset{?}{=} \dfrac{1}{\|\tilde{\Phi}(x)\|}$

show $\exp(x^2) = \|\tilde{\phi}(x)\|$

$$\|\tilde{\Phi}(x)\| = \sqrt{1^2 + \left(\sqrt{\tfrac{2}{1!}}\,x\right)^2 + \left(\sqrt{\tfrac{2^2}{2!}}\,x^2\right)^2 + \cdots}$$

$$\Rightarrow \|\tilde{\Phi}(x)\|^2 = 1 + \frac{2}{1!}x^2 + \frac{2^2}{2!}x^4 + \cdots$$

$$= 1 + \frac{2x^2}{1!} + \frac{(2x^2)^2}{2!} + \cdots$$

$$= \exp(2x^2) = \left[\exp(x^2)\right]^2$$

$$\Rightarrow \|\tilde{\Phi}(x)\| = \exp(x^2)$$

6.

9.

7.

$$\mathcal{L}(R, c, \lambda)$$

$$= R^2 + \sum_{n=1}^{N} \lambda_n \left(\|z_n - c\|^2 - R^2 \right)$$

8. KKT, Conditions

$$\begin{cases} \text{primal feasible.} & \|z_n - c\|^2 \le R^2, \quad \forall n \\ \text{dual-feasible} & \lambda_n \ge 0 \quad \forall n \\ \text{and} & \\ \lambda_n \left(\|z_n - c\|^2 - R^2 \right) = 0 & \forall n. \end{cases}$$

And (D)

$$\max_{\alpha_n \ge 0} \left(\min_{R, c} \mathcal{L}(R, c, \lambda) \right)$$

inner optimal.

$$\frac{\partial L}{\partial R} = 0 \Rightarrow 2R - 2R \sum_{n=1}^{N} \lambda_n = 0$$

$$\Rightarrow R\left(1 - \sum_{n=1}^{N} \lambda_n\right) = 0 \quad \text{--- (a)}$$

$$\frac{\partial L}{\partial C_i} = 0 \Rightarrow -2 \sum_{n=1}^{N} \lambda_n \left(Z_n(i) - C_i \right) = 0$$

$$\Rightarrow C_i \sum_{n=1}^{N} \lambda_n = \sum_{n=1}^{N} \lambda_n Z_n(i)$$

$$\Rightarrow C = \frac{\sum^{N} \lambda_i Z_n}{\sum^{N} \lambda_i} \quad \left(\text{if } \sum_{n=1}^{N} \lambda_i \ne 0 \right) \quad \text{--- (b)}$$

9. $R > 0$. by (a) $\Rightarrow \sum_{n=1}^{N} \lambda_i = 1$

$$\Rightarrow \quad C = \sum_{n=1}^{N} \lambda_n z_n$$

Then $\quad L(R, C, \lambda) = \sum \lambda_n \| z_n - C \|^2$

(D')

$$= \sum_{n=1}^{N} \lambda_n \| z_n - \sum_{m=1}^{N} \lambda_m z_m \|^2$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \lambda_i = 1$$

10.

$$L(R, C, \lambda) = \sum_{n=1}^{N} \lambda_n \left(z_n - \sum_{m=1}^{N} \lambda_m z_m \right)^T \left(z_n - \sum_{m=1}^{N} \lambda_m z_m \right)$$

$$= \sum_{n=1}^{N} \lambda_n z_n^T z_n - 2 \sum_{n=1}^{N} \sum_{m=1}^{N} \lambda_n \lambda_m z_m^T z_n + \sum_{n=1}^{N} \lambda_n \sum_{m=1}^{N} \sum_{j=1}^{N} \lambda_m \lambda_j z_m^T z_j$$

$$= \sum_{n=1}^{N} \lambda_n K(x_n, x_n) - 2 \sum_{n=1}^{N} \sum_{m=1}^{N} \lambda_n \lambda_m K(x_m, x_n) + \sum_{n=1}^{N} \lambda_n \left[\sum_{m=1}^{N} \sum_{j=1}^{N} \lambda_m \lambda_j K(x_m, x_j) \right]$$

$$= \sum_{n=1}^{N} \lambda_n K(x_n, x_n) - \sum_{n=1}^{N} \sum_{n=1}^{N} \lambda_n \lambda_m K(x_m, x_n)$$

→ ~~In~~ to get minimized R^2.

We take any $\lambda_i = 0$. then

$$R = \sqrt{\| z_i - C \|} = \sqrt{K(z_i, z_i) - 2 \sum_{m=1}^{N} \lambda_m K(x_m, x_i) + \sum_{n=1}^{N} \sum_{m=1}^{N} \lambda_n \lambda_m K(x_n, x_m)}$$

11. In Soft-margin SVM

$\beta_n = C - \alpha_n \rightarrow$ multiplier for $(-\xi)$

If $C \geq \max \alpha_n^*$, α_n^* in hard-margin.

this imply all $\beta_n \geq 0$

$\Rightarrow$ all $\xi_i = 0$, this makes soft-margin SVM the same as hard-margin SVM.

12. Let original SVM with $K(x, x')$ is

$$g_{SVM}(x) = \text{sign}\left(\sum_{SV_n} \alpha_n y_n K(x_n, x) + \left(y_s - \sum \alpha_n y_n K(x_n, x_s)\right)\right)$$

New $$\tilde{g}_{SVM} = \text{sign}\left(\sum_{SV} \tilde{\alpha}_n y_n p K(x_n, x) + \left(y_s - \sum \tilde{\alpha}_n y_n p(x_n, x_s)\right)\right),$$

with $\tilde{K}(x, x') = p\, K(x, x')$

if let $\tilde{\alpha}_n = \frac{\alpha_n}{p}$, then $g_{SVM} = \tilde{g}_{SVM}$

and to make sure bounded SV and SV are the same

for $\Rightarrow \tilde{\alpha}_n = \frac{\alpha_n}{p} = \boxed{\frac{C}{p} = \tilde{C}}$