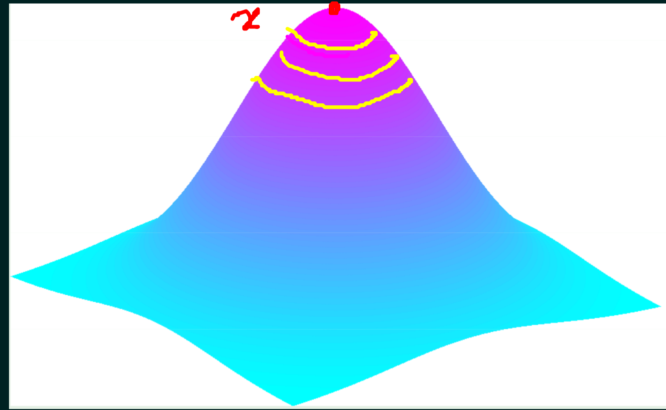
Radial Basis Function (RBF)

$$(x_n, y_n) \rightsquigarrow h(x).$$

$$e^{\left(-\gamma \overbrace{\|x - x_n\|^2}^{\text{Radial}}\right)} \quad \underbrace{\forall n \in [1, N]}_{\text{basis}}$$



$$h(x) = \sum_{i=1}^{N} \boxed{w_n} e^{-\gamma \|x - x_n\|^2}$$

Linear Regression: $\quad \underline{\underline{E_{in} \approx 0}} \rightarrow \displaystyle\sum_{i=1}^{N} w_n e^{-\gamma \|x - x_n\|^2} = y_n$

$\boxed{N \text{ eq}^n, \quad N \text{ unknown}}$

$$\underbrace{\begin{bmatrix} e^{-\gamma\|x_1 - x_1\|^2} & \cdots & e^{-\gamma\|x_1 - x_N\|^2} \\ e^{-\gamma\|x_2 - x_1\|^2} & \cdots & e^{-\gamma\|x_2 - x_N\|^2} \\ e^{-\gamma\|x_3 - x_1\|^2} & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix}}_{\phi} \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ \vdots \\ w_N \end{bmatrix}}_{w} = \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_N \end{bmatrix}}_{y}$$

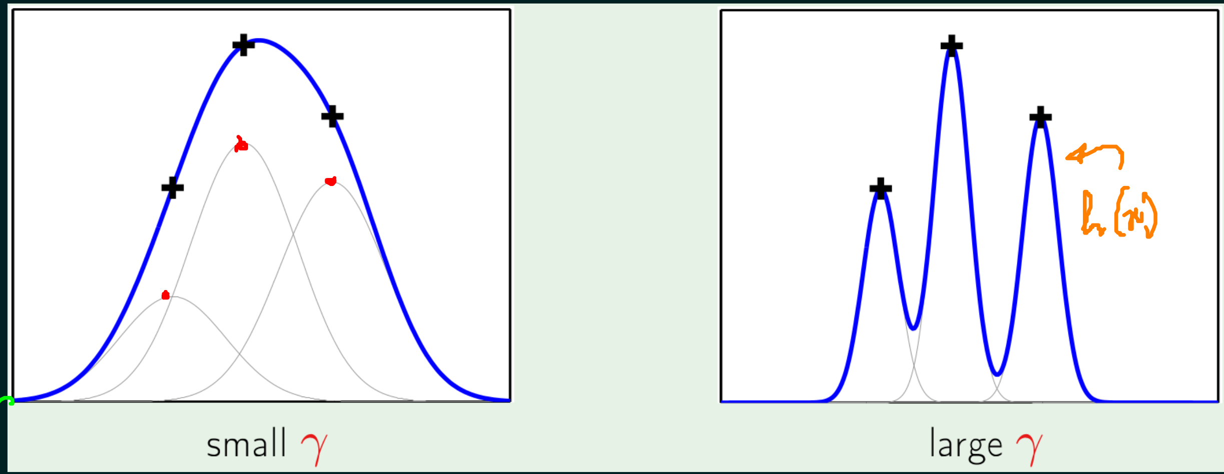
$\phi \rightarrow$ invertible

$\Rightarrow \quad w = \phi^{-1} y. \quad \leftarrow \underline{\underline{\text{exact}}}$ interpolation $\rightarrow$ overfitting??

$$\sum_{i=1}^{N} (h(x_i) - y_i)^2 \leftarrow \text{Minimize}$$

$$h(x) = \sum_{n=1}^{N} w_n e^{-\gamma\|x-x_n\|^2}$$

RBF



Linear Regression

→ Linear Classification

$$h(x) = \text{sign}\left(\underbrace{\sum_{n=1}^{N} w_n e^{-\gamma\|x-x_n\|^2}}_{s}\right)$$

$$\mathcal{D}\left\{(x_1, \pm 1),\ (x_2, \pm 1), \ldots, (x_N, \pm 1)\right.$$

[minimize $(s-y)^2$ on $\mathcal{D}$ $y_i = \pm 1$

→ $h(x) = \text{sign}(s)$ where s minimize error → $f(w_1, \ldots, w_N)$

$w = \phi^{-1} y \to \boxed{\pm 1}$

RBF as a Model → Linear Reg.

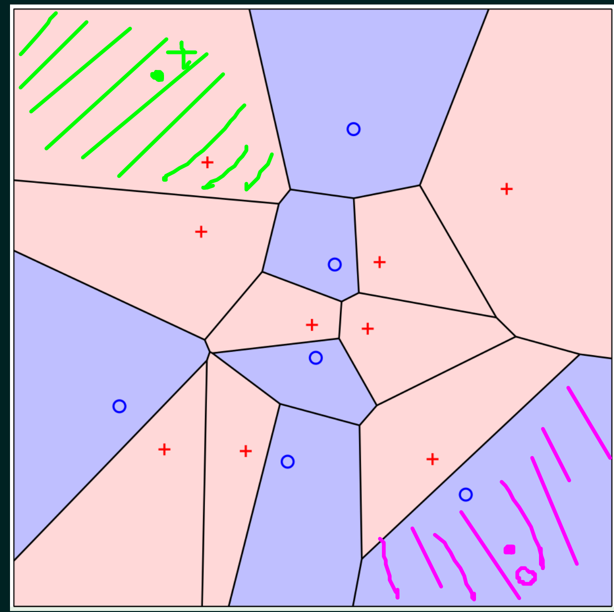
→ Linear classify.

Nearest Neighbour

gaussian

K - Centres $K \ll N$

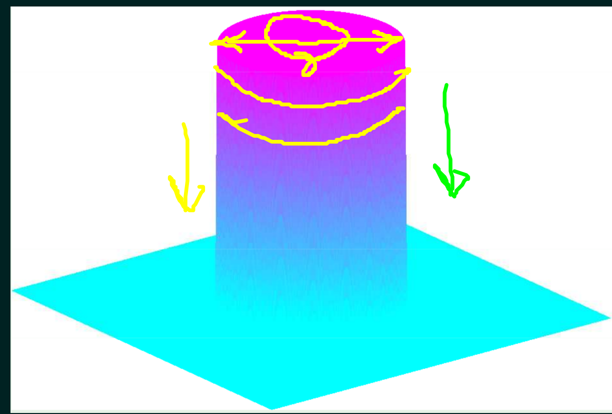
1-NN

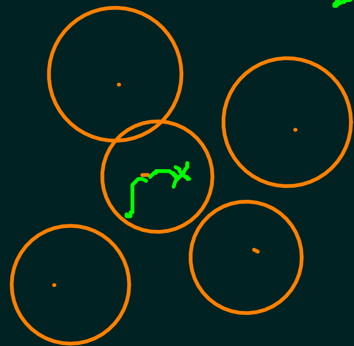


$\mu_1 \quad \mu_2 \quad \cdots \quad \mu_K$

$$h(x) = \sum_{k=1}^{K} \underbrace{w_k}_{?\ \mathcal{X}_K} e^{-\gamma \|x - \underbrace{\mu_k}_{?}\|^2}$$

Obj: $(x_1 \cdots\cdots x_N)$ datapoint $S_1 \cdots\cdots S_K$ Classes





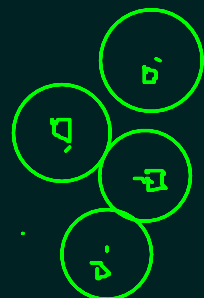
$\rightarrow$ min. $\displaystyle\sum_{k=1}^{K} \sum_{x_n \in S_k} \|x_n - \mu_k\|^2$

[K - means Clustering]

(NP-hard)

(1-NN) Basis fn

Lloyd: (1) $\mu_k \leftarrow \frac{1}{|S_k|} \sum_{x_n \in S_k} x_n$ $\leftarrow$ updation of mean



(2) $S_k \leftarrow \{ x_n : \|x_n - \mu_k\| \leq \text{all } \|x_n - \mu_l\| \}$

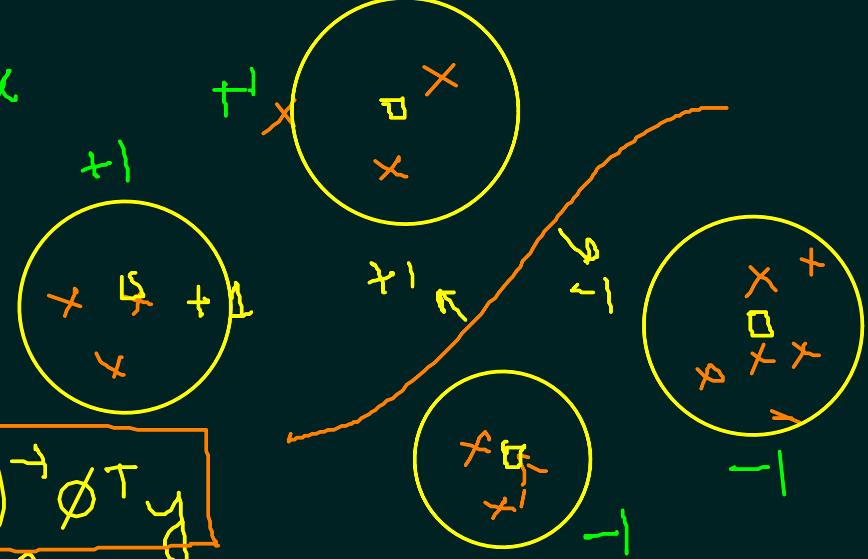
$$\sum_{k=1}^{K} w_k\, e^{-\gamma \|x_n - \mu_k\|^2} \approx y_n \quad \forall n$$

$$\begin{bmatrix} \leftarrow \phi \rightarrow \end{bmatrix} \begin{bmatrix} w \end{bmatrix} \approx \begin{bmatrix} y \end{bmatrix}$$

K unk

N eq.

$\mu_k \downarrow w_k$

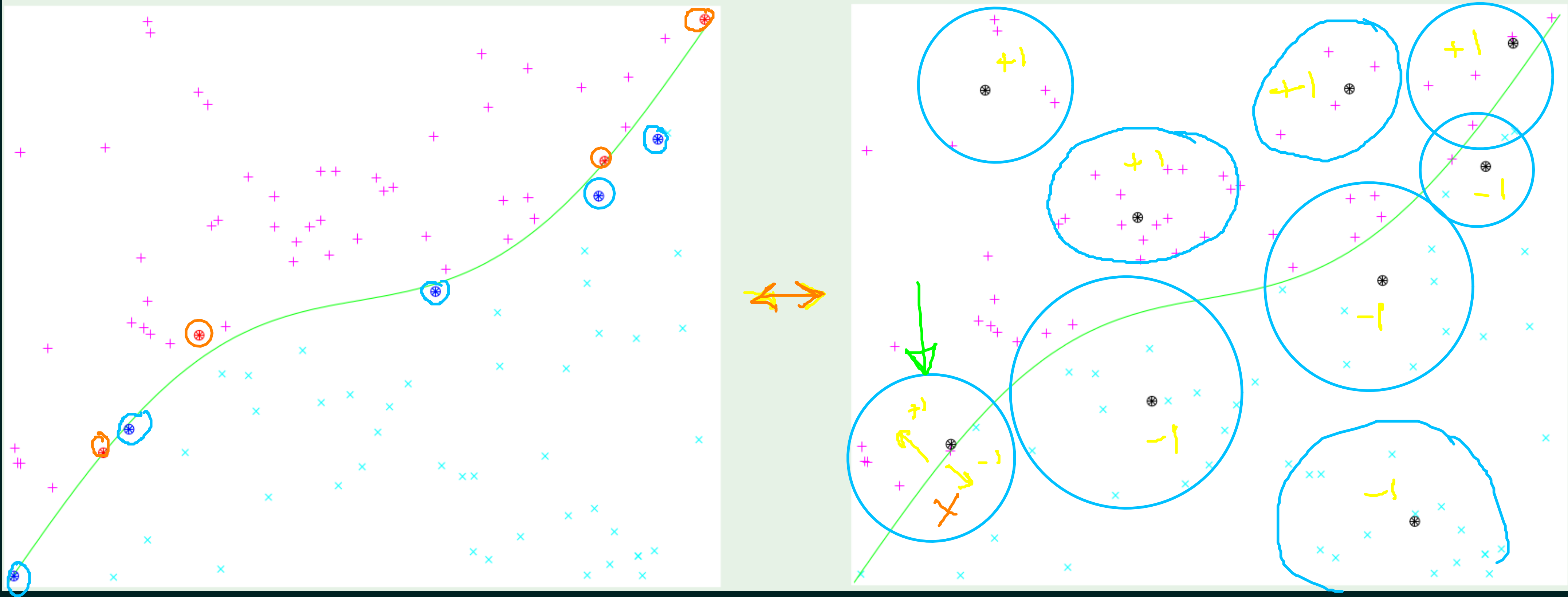


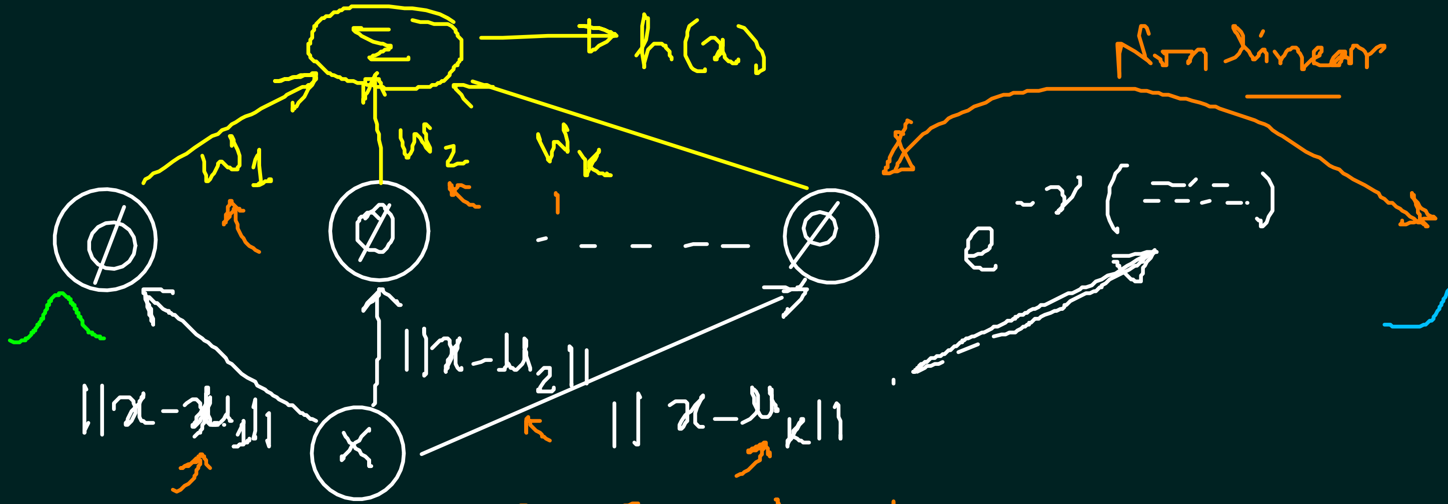
$\phi^T\phi$ is invertible $\Rightarrow$ $\boxed{w = (\phi^T\phi)^{-1}\phi^T y}$

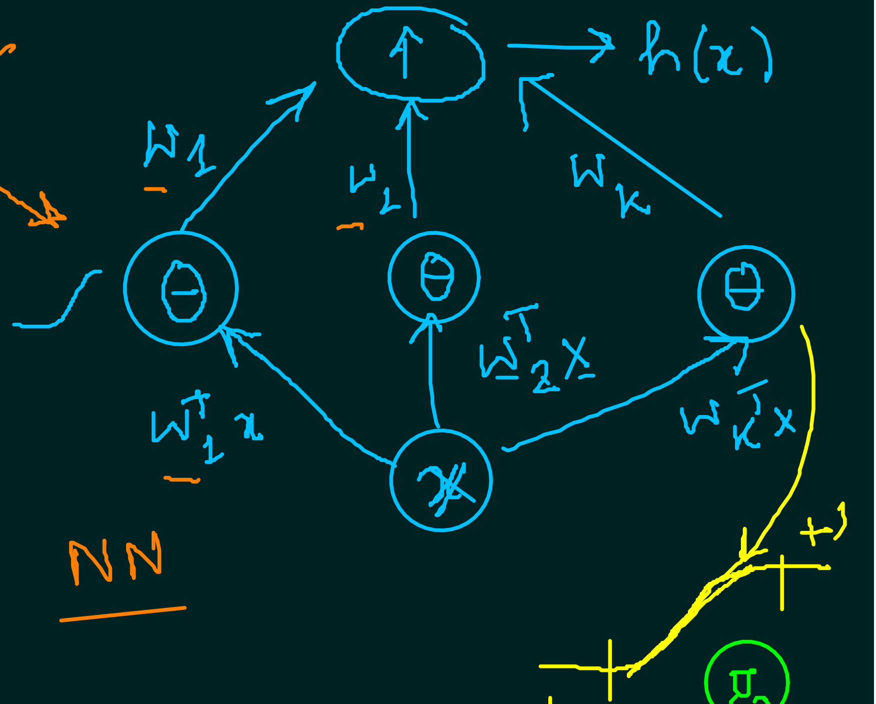
(Pseudo inverse)

support vectors

RBF centers







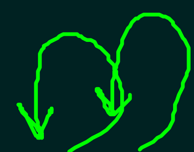
RBF Network ⟷ NN

minimize the error

→ Choose γ : $h(x) = \sum_{k=1}^{K} w_k e^{-\gamma \|x-\mu_k\|^2}$

↳ ~ EM algorithm as mixture of Gaussians

↳ ① Fix γ and ~~choose~~ solve $w_1 \ldots w_K$



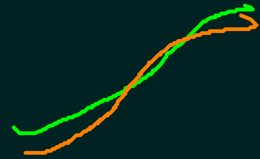
② For $w_1 \ldots w_K$, min error w.r.t. γ

Converge $\gamma_1 \ldots \gamma_K$ local min.

$K(x, x') = e^{-\gamma\|x-x'\|^2}$ ← RBF as Kernel

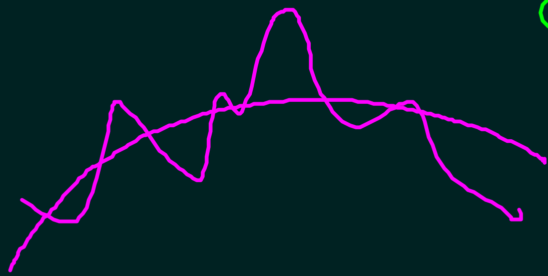
SVM → $\text{sign}\left(\sum_{\alpha_n>0} \alpha_n y_n e^{-\gamma\|x-x_n\|^2} + b\right)$

$\text{sign}\left(\sum_{k=1}^{K} w_k e^{-\gamma\|x-\mu_k\|^2} + \cancel{b}\right)$ $\quad b \to 0$



▷ RBF as Regularizers:

↳ "smooth" hyp.



$\left(\frac{\partial^k h}{\partial x^k}\right)$

$$E = \underbrace{\sum_{n=1}^{N} \left(h(x_n) - y_n\right)^2}_{\text{error min}} + \lambda a_k \underbrace{\int_{-\infty}^{\infty} \left(\frac{d^k h}{dx^k}\right)^2 dx}_{\text{introduce smoothness}}$$

$E \to$ min

λ: Reg. const

[more abrupt is your deriv. less smooth is h(x)]

→ drive → Regularize

(smooth interpolation)

⇝ $\boxed{\text{RBF}} = h(x)$