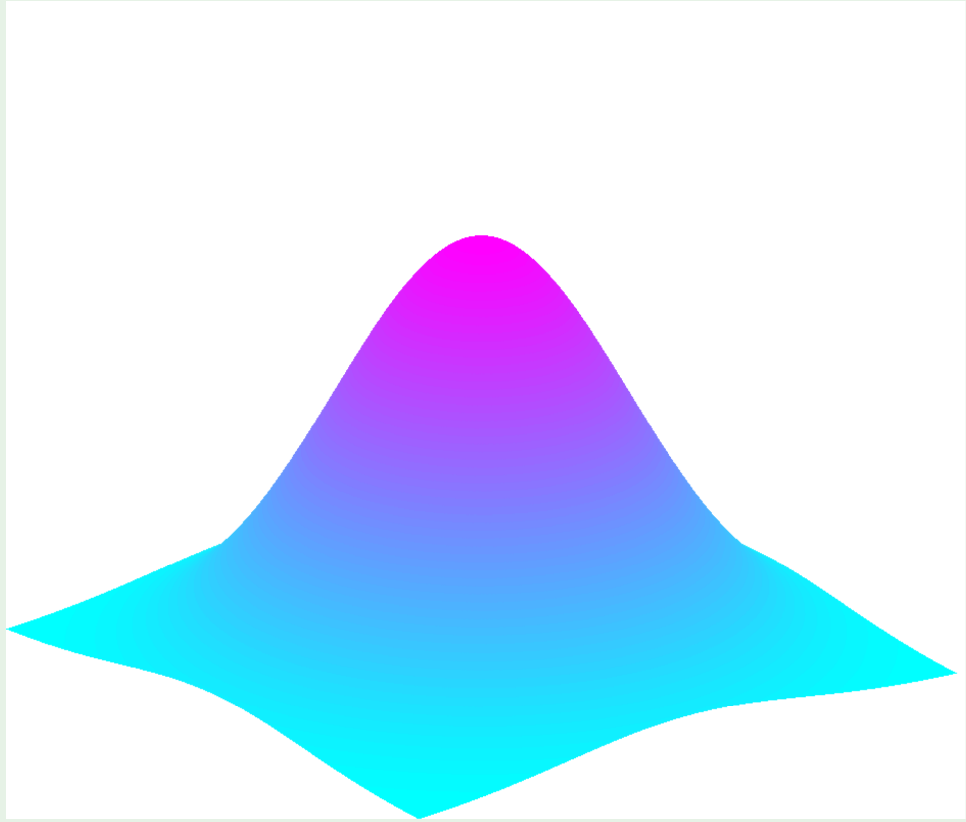
Basic RBF model

Each $(\mathbf{x}_n, y_n) \in \mathcal{D}$ influences $h(\mathbf{x})$ based on $\underbrace{\|\mathbf{x} - \mathbf{x}_n\|}_{\text{radial}}$

Standard form:

$$h(\mathbf{x}) = \sum_{n=1}^{N} w_n \underbrace{\exp\left(-\gamma \|\mathbf{x} - \mathbf{x}_n\|^2\right)}_{\text{basis function}}$$



The learning algorithm

Finding $w_1, \cdots, w_N$:

$$h(\mathbf{x}) = \sum_{n=1}^{N} w_n \exp\left(-\gamma \left\|\mathbf{x} - \mathbf{x}_n\right\|^2\right)$$

based on $\mathcal{D} = (\mathbf{x}_1, y_1), \cdots, (\mathbf{x}_N, y_N)$

$E_{\text{in}} = 0$: $\quad h(\mathbf{x}_n) = y_n$ for $n = 1, \cdots, N$:

$$\sum_{m=1}^{N} w_m \exp\left(-\gamma \left\|\mathbf{x}_n - \mathbf{x}_m\right\|^2\right) = y_n$$

The solution

$$\sum_{m=1}^{N} w_m \ \exp\left(-\gamma \left\|\mathbf{x}_n - \mathbf{x}_m\right\|^2\right) \ = \ y_n$$

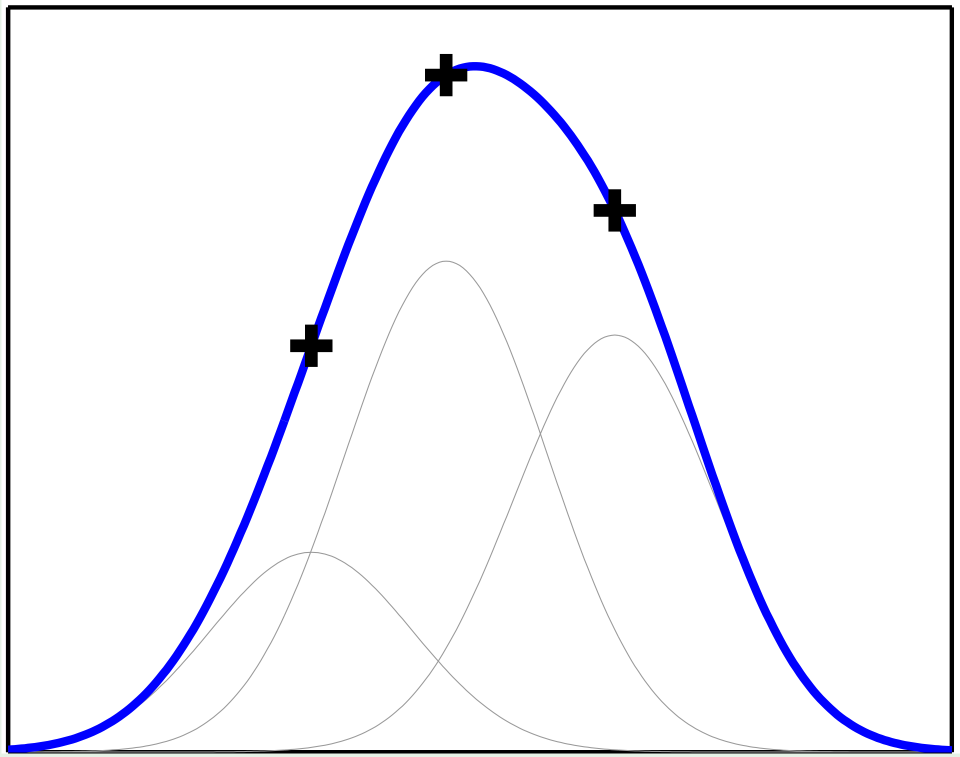
N equations in N unknowns

$$\underbrace{\begin{bmatrix} \exp(-\gamma \left\|\mathbf{x}_1 - \mathbf{x}_1\right\|^2) & \dots & \exp(-\gamma \left\|\mathbf{x}_1 - \mathbf{x}_N\right\|^2) \\ \exp(-\gamma \left\|\mathbf{x}_2 - \mathbf{x}_1\right\|^2) & \dots & \exp(-\gamma \left\|\mathbf{x}_2 - \mathbf{x}_N\right\|^2) \\ \vdots & \vdots & \vdots \\ \exp(-\gamma \left\|\mathbf{x}_N - \mathbf{x}_1\right\|^2) & \dots & \exp(-\gamma \left\|\mathbf{x}_N - \mathbf{x}_N\right\|^2) \end{bmatrix}}_{\Phi} \underbrace{\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}}_{\mathbf{w}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}}_{\mathbf{y}}$$

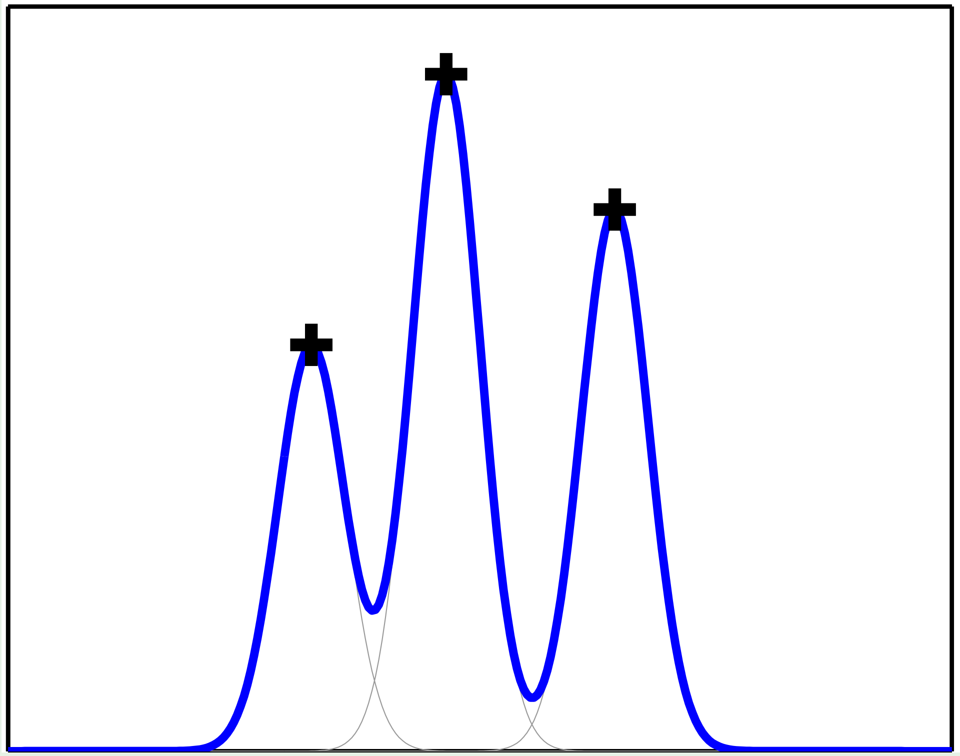
If Φ is invertible, $\boxed{\mathbf{w} = \Phi^{-1}\mathbf{y}}$ "exact interpolation"

The effect of γ

$$h(\mathbf{x}) = \sum_{n=1}^{N} w_n \exp\left(-\gamma \, \|\mathbf{x} - \mathbf{x}_n\|^2\right)$$



small γ



large γ

RBF for classification

$$h(\mathbf{x}) \; = \operatorname{sign}\left(\sum_{n=1}^{N} w_n \; \exp\left(-\gamma \, \|\mathbf{x}-\mathbf{x}_n\|^2\right)\right)$$

Learning: $\sim$ linear regression for classification

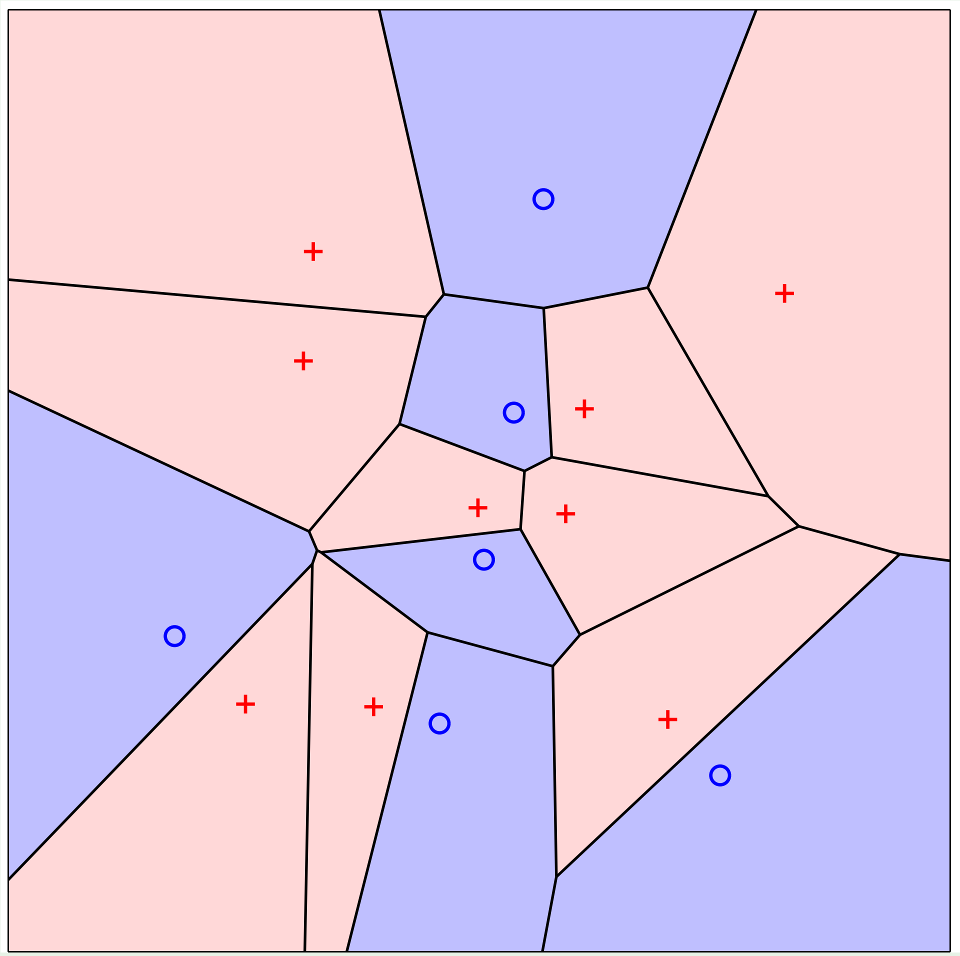
$$s \; = \sum_{n=1}^{N} w_n \; \exp\left(-\gamma \, \|\mathbf{x}-\mathbf{x}_n\|^2\right)$$

Minimize $(s-y)^2$ on $\mathcal{D}$ $\qquad y = \pm 1$

$h(\mathbf{x}) = \operatorname{sign}(s)$

Relationship to nearest-neighbor method

Adopt the y value of a nearby point:



similar effect by a basis function: