

Kernel Ridge Regression (KRR)

$$\underbrace{\mathbf{K}_{ij}}_{\text{Kernel matrix element}} = \underbrace{k(\mathbf{x}_i, \mathbf{x}_j)}_{\text{Kernel function}} \quad \begin{matrix} \text{Features of training points} \\ \text{Kernel function} \end{matrix} \quad (1)$$

$$\mathbf{w} = (\mathbf{K} + \lambda \mathbf{I}_N)^{-1} \mathbf{y} \quad \begin{matrix} \text{Model weights} & \text{Training labels} \\ \mathbf{K} & \mathbf{y} \\ \text{Kernel matrix } (N \times N) & \text{Identity matrix } (N \times N) \\ \lambda & \text{Regularization} \end{matrix} \quad (2)$$

$$\hat{y}(\mathbf{x}_q) = \sum_{i=1}^N \underbrace{w_i}_{\text{Weight of } i\text{-th training point}} k(\mathbf{x}_i, \mathbf{x}_q) \quad \begin{matrix} \text{Query} & \text{Training point features} \\ \mathbf{x}_q & \mathbf{x}_i \\ \text{Prediction} & \text{Weight of } i\text{-th training point} \end{matrix} \quad (3)$$