- Let r be the rank of $\widehat{\mathbf{K}}$, where

$$\widehat{\mathbf{K}} = \mathbf{K} - \frac{1}{n}\mathbf{1}\mathbf{1}^\top\mathbf{K} - \frac{1}{n}\mathbf{K}\mathbf{1}\mathbf{1}^\top + \frac{1}{n^2}(\mathbf{1}^\top\mathbf{K}\mathbf{1})\mathbf{1}\mathbf{1}^\top,$$

  let $\sigma_1^2 \geq \cdots \geq \sigma_r^2$ be the nonzero eigenvalues of $\widehat{\mathbf{K}}$, and let $v_1, \ldots, v_r$ be corresponding unit eigenvectors. The notation

$$\alpha_k = \sigma_k^{-1} v_k$$

  is often used, where the α_k are called the *dual variables*.

- The column vector Y_k $(1 \leq k \leq r)$ defined by

$$Y_k = \left(\sum_{i=1}^{n} (\alpha_k)_i \widehat{\mathbf{K}}_{ij} \right)_{j=1}^{n}$$

  is called the *kth kernel principal component* (for short *kth kernel PCA*) of the data set $S = \{x_1, \ldots, x_n\}$ in the direction $u_k = \sum_{i=1}^{n} \sigma_k^{-1} (v_k)_i \widehat{X}_i^\top$ (even though the matrix $\widehat{X}$ is *not known*).

53.5 Summary

The main concepts and results of this chapter are listed below:

- Feature map, feature embedding, feature space.
- Kernel function.
- Positive definite kernel, real positive definite kernel.
- Gram matrix.
- Hadamard product, Schur product.
- Gaussian kernel.
- Reproducing kernel Hilbert space (RKHS).
- Reproducing property.
- Intersection kernel, union complement kernel, agreement kernel.
- Kernel PCA.
- k-th kernel PCA.