Source task and conventions

Prove rank, coordinate and new-point formulas, retaining the required centring convention.

Conventional solution and proof

Let $J = I - \mathbf{1}\mathbf{1}^\top/N$ and $K_c = JKJ = UD^2U^\top$, keeping only r positive eigenvalues d_m^2. The feature-centred data operator has SVD with these singular values, so exactly $r = \operatorname{rank} K_c$ nonzero principal components occur. Without centring this is $\operatorname{rank} K$; with centring it need not be. A constant kernel has rank one before centring and zero after centring. Set $\alpha_m = u_m/d_m$. Then $K_c\alpha_m = d_m u_m = z_m$, establishing the coordinate formula. To see its meaning, write $K_c = \Phi_c\Phi_c^*$, where the rows of Φ_c are the centred feature vectors. The feature-space direction $v_m = \Phi_c^* u_m/d_m$ has norm one, and $\Phi_c v_m = d_m u_m$. A new point has coordinate $z_{0m} = \langle \phi(x_0) - \bar{\phi}, v_m \rangle = k_c(x_0)^\top u_m/d_m$, where

$$[k_c(x_0)]_i = K(x_i, x_0) - \frac{1}{N}\sum_j K(x_j, x_0) - \frac{1}{N}\sum_j K(x_i, x_j) + \frac{1}{N^2}\sum_{j,l} K(x_j, x_l).$$

All terms use the training mean. Zero eigenvalues are omitted, so no division by zero is performed. Eigenvalue multiplicities permit rotations within the corresponding principal subspace.

Structural or analytic derivation

This is the finite-vector/SVD argument applied to an implicit Hilbert-space feature map. The kernel supplies the inner products needed by the projection without constructing all feature coordinates. The new-point centring formula is necessary to preserve the same affine origin.

Comparative cost and preferred method