(1) We need to center the data and compute the inner products of pairs of centered data. More precisely, if the centroid of $\varphi(S)$ is

$$\mu = \frac{1}{n}(\varphi(x_1) + \cdots + \varphi(x_n)),$$

then we need to compute the inner products $\langle \varphi(x) - \mu, \varphi(y) - \mu\rangle$.

(2) Let us assume that $F = \mathbb{R}^d$ with the standard Euclidean inner product and that the data points $\varphi(x_i)$ are expressed as *row vectors* X_i of an $n \times d$ matrix X (as it is customary). Then the inner products $\kappa(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j)\rangle$ are given by the *kernel matrix* $\mathbf{K} = XX^\top$. Be aware that with this representation, in the expression $\langle \varphi(x_i), \varphi(x_j)\rangle$, $\varphi(x_i)$ is a d-dimensional column vector, while $\varphi(x_i) = X_i^\top$. However, the jth component $(Y_k)_j$ of the principal component Y_k (viewed as a n-dimensional column vector) is given by the projection of $\widehat{X}_j = X_j - \mu$ onto the direction u_k (viewing μ as a d-dimensional row vector), which is a unit eigenvector of the matrix $(X-\mu)^\top(X-\mu)$ (where $\widehat{X} = X - \mu$ is the matrix whose jth row is $\widehat{X}_j = X_j - \mu$), is given by the inner product

$$\langle X_j - \mu, u_k\rangle = (Y_k)_j;$$

see Definition 23.2 and Theorem 23.11. The problem is that we know what the matrix $(X-\mu)(X-\mu)^\top$ is from (1), because it can be expressed in terms of $\mathbf{K}$, but we don't know what $(X-\mu)^\top(X-\mu)$ is because we don't have access to $\widehat{X} = X - \mu$.

Both difficulties are easily overcome. For (1) we have

$$\begin{aligned}\langle \varphi(x) - \mu, \varphi(y) - \mu\rangle &= \left\langle \varphi(x) - \frac{1}{n}\sum_{k=1}^n \varphi(x_k), \varphi(y) - \frac{1}{n}\sum_{k=1}^n \varphi(x_k)\right\rangle \\ &= \kappa(x,y) - \frac{1}{n}\sum_{i=1}^n \kappa(x, x_i) - \frac{1}{n}\sum_{j=1}^n \kappa(x_j, y) + \frac{1}{n^2}\sum_{i,j=1}^n \kappa(x_i, x_j).\end{aligned}$$

For (2), if $\mathbf{K}$ is the kernel matrix $\mathbf{K} = (\kappa(x_i, x_j))$, then the kernel matrix $\widehat{\mathbf{K}}$ corresponding to the kernel function $\widehat{\kappa}$ given by

$$\widehat{\kappa}(x,y) = \langle \varphi(x) - \mu, \varphi(y) - \mu\rangle$$

can be expressed in terms of $\mathbf{K}$. Let $\mathbf{1}$ be the column vector (of dimension n) whose entries are all 1. Then $\mathbf{1}\mathbf{1}^\top$ is the $n \times n$ matrix whose entries are all 1. If A is an $n \times n$ matrix, then $\mathbf{1}^\top A$ is the row vector consisting of the sums of the columns of A, $A\mathbf{1}$ is the column vector consisting of the sums of the rows of A, and $\mathbf{1}^\top A\mathbf{1}$ is the sum of all the entries in A. Then it is easy to see that the kernel matrix corresponding to the kernel function $\widehat{\kappa}$ is given by

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