Exercise 1

Based on our observation, the parameter C punishes the existence of noisy data points and separates these points from the surrounding points.

Exercise 2

a) rbf-kernel, good classifier (all points are classified correctly)
b) polynomial-kernel, bad classifier (many points not classified correctly, structure of "classification-surface" is not the same as "data-surface"), can be avoided by choosing higher degree of kernel and higher C In Fig.2-a the empty white circles my be the support vectors (but not sure).

Exercise 3

A kernel (K) must be a positive definite function, thus we have the property that there exists a c such that $c^T K c > 0$. Therefore we have to check if this property holds also for the new kernels given it holds for the original kernel(s).
1) If $c^T K c > 0$ holds, then $\alpha c^T K c > 0$ also holds for $\alpha > 0$
2) $c^T (K_1 + K_2) c = c^T K_1 c + c^T K_2 c$ By definition, both terms are > 0 and therefore the new kernel is positive definite.
3) $K_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $K_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ $c = \begin{pmatrix} a & b \end{pmatrix}$
$c^T K_1 c = a^2 + b^2$
$c^T K_2 c = 2a^2 + 2b^2$ but
$c^T (K_1 - K_2) c = -1a^2 + (-1b^2)$