Problem 1

Solution
We denote $k(x, y) =< \phi(x), \phi(y) >$
(A).

Proof.

$$K_{ij} = k(x_i, x_j) = \phi(x_i)\phi(x_j)$$

$$\begin{aligned} c^T K c &= \sum_i \sum_j c_i c_j K_{ij} \\ &= \sum_i \sum_j c_i c_j \phi(x_i)\phi(x_j) \\ &= (\sum_i c_i \phi(x_i))(\sum_j c_j \phi(x_j)) \\ &= || \sum_i c_i \phi(x_i)||_2^2 \geq 0 \end{aligned}$$

□

(a).

Proof.

$$\begin{aligned} k(x, y) &= \alpha k_1(x, y) + \beta k_2(x, y) \\ &=< \sqrt{\alpha}\phi_1(x)\sqrt{\alpha}\phi_1(y) > + < \sqrt{\beta}\phi_2(x)^T \sqrt{\beta}\phi_2(y) > \\ &=< [\sqrt{\alpha}\phi_1(x), \sqrt{\beta}\phi_2(x)], [\sqrt{\alpha}\phi_1(y), \sqrt{\beta}\phi_2(y)] > \end{aligned}$$

□

(b).

Proof. Let $f_i(x)$ be the ith feature value under the feature map ϕ_1, $g_i(x)$ be the ith feature value under the feature map ϕ_2. Then:

$$\begin{aligned} k(x, y) &= k_1(x, y)k_2(x, y) \\ &= (\phi_1(x)\phi_1(y))(\phi_2(x)\phi_2(y)) \\ &= (\sum_{i=0}^{\infty} f_i(x)f_i(y))(\sum_{j=0}^{\infty} g_j(x)g_j(y)) \\ &= \sum_{i,j} f_i(x)f_i(y)g_j(x)g_j(y) \\ &= \sum_{i,j} (f_i(x)g_j(x))(f_i(y)g_j(y)) \\ &=< \phi_3(x), \phi_3(y) > \end{aligned}$$

where ϕ_3 has feature $h_{i,j}(x) = f_i(x)g_j(x)$. □

(c). Since each polynomial term is a product of kernels with a positive coefficient, the proof follows from part a and b.
(d). By Taylor expansion:

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$$

The proof follows part c.

(B).

Proof. We wish to show that the kernel $k(\mathbf{x},\mathbf{y}) = \exp(-\frac{1}{2}||\mathbf{x}-\mathbf{y}||^2)$ can be written as an inner product form, $k(\mathbf{x},\mathbf{y}) =< \phi(\mathbf{x}), \phi(\mathbf{y}) >$. Consider the formula for $\phi_z(x) = (\pi/2)^{-d/4}\exp(-||\mathbf{x}-\mathbf{z}||^2)$, $\phi_z(y) = (\pi/2)^{-d/4}\exp(-||\mathbf{y}-\mathbf{z}||^2)$, $\mathbf{x},\mathbf{y},\mathbf{z} \in \mathbb{R}^d$. We define the kernel as $k(\mathbf{x},\mathbf{y}) =< \phi_z(\mathbf{x}), \phi_z(\mathbf{y}) >= \int_z \phi_z(\mathbf{x}) \times \phi_z(\mathbf{y}) d\mathbf{z}$.

$$
\begin{aligned}
k(\mathbf{x},\mathbf{y}) &= \int_z (\pi/2)^{-d/4}\exp(-||\mathbf{x}-\mathbf{z}||^2) \times (\pi/2)^{-d/4}\exp(-||\mathbf{y}-\mathbf{z}||^2) d\mathbf{z} \\
&= (\pi/2)^{-d/2}\int_z \exp(-\mathbf{x}^T\mathbf{x} - \mathbf{z}^T\mathbf{z} + 2\mathbf{x}^T\mathbf{z})\exp(-\mathbf{y}^T\mathbf{y} - \mathbf{z}^T\mathbf{z} + 2\mathbf{y}^T\mathbf{z}) d\mathbf{z} \\
&= (\pi/2)^{-d/2}\exp(-\mathbf{x}^T\mathbf{x} - \mathbf{y}^T\mathbf{y})\int_z \exp(-2\mathbf{z}^T\mathbf{z} + 2(\mathbf{y}+\mathbf{x})^T\mathbf{z}) d\mathbf{z}
\end{aligned}
$$

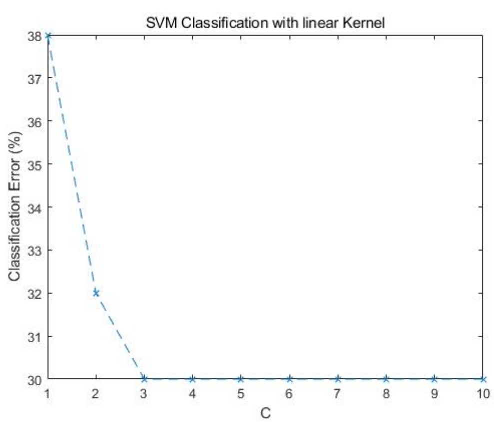
Denote $\mathbf{p} = (\mathbf{y}+\mathbf{x})/2$

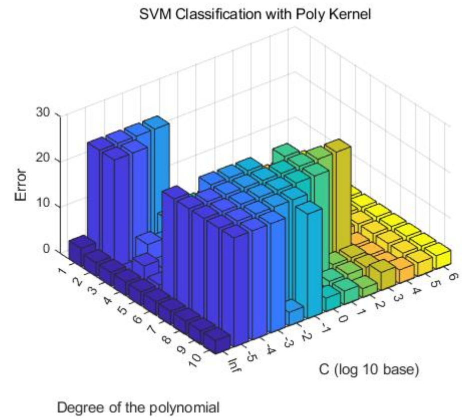
$$
\begin{aligned}
k(\mathbf{x},\mathbf{y}) &= (\pi/2)^{-d/2}\exp(-\mathbf{x}^T\mathbf{x} - \mathbf{y}^T\mathbf{y})\int_z \exp(-2\mathbf{z}^T\mathbf{z} + 4\mathbf{p}^T\mathbf{z}) d\mathbf{z} \\
&= (\pi/2)^{-d/2}\exp(-\mathbf{x}^T\mathbf{x} - \mathbf{y}^T\mathbf{y})\exp(2\mathbf{p}^T\mathbf{p})\int_z \exp(-2\mathbf{z}^T\mathbf{z} + 4\mathbf{p}^T\mathbf{z} - 2\mathbf{p}^T\mathbf{p}) d\mathbf{z} \\
&= (\pi/2)^{-d/2}\exp(-\mathbf{x}^T\mathbf{x} - \mathbf{y}^T\mathbf{y})\exp(2\mathbf{p}^T\mathbf{p})(\pi/2)^{d/2} \\
&= \exp(-\mathbf{x}^T\mathbf{x} - \mathbf{y}^T\mathbf{y})\exp(\frac{1}{2}\mathbf{x}^T\mathbf{x} + \frac{1}{2}\mathbf{y}^T\mathbf{y} + \mathbf{x}^T y) \\
&= \exp(-\frac{1}{2}\mathbf{x}^T\mathbf{x} - \frac{1}{2}\mathbf{y}^T\mathbf{y} + \mathbf{x}^T y) \\
&= \exp(-\frac{1}{2}||x-y||^2)
\end{aligned}
$$

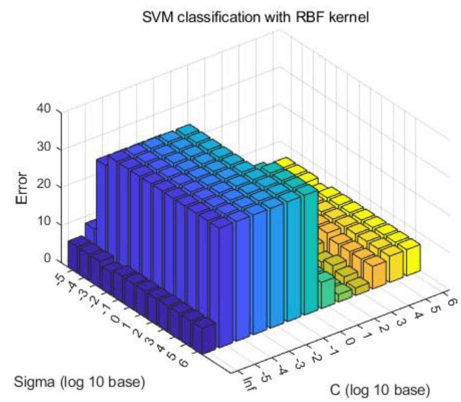
□

Problem 2

Solution After randomly splitting the data set into two halves for cross validation, train SVM using linear, polynomial and RBF kernel with different polynomial degree, sigma and C.







Problem 3

Solution

$$L(\alpha) = \prod_{i=1}^{n} \alpha^{-2} x e^{-x/\alpha} = \alpha^{-2n} \prod_{i=1}^{n} x e^{-x/\alpha} = \alpha^{-2n} e^{-\frac{1}{\alpha}\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} x$$

$$\log L(\alpha) = -2n \log \alpha - \frac{1}{\alpha}(x_1 + x_2 + ... x_n) + \log \prod_{i=1}^{n} x_i$$

$$\frac{d}{d\alpha} \log L(\alpha) = \frac{-2n}{\alpha} + \frac{1}{\alpha^2} \sum_{i=1} n x_i = 0$$

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} x_i}{2n} \quad (\hat{\alpha} > 0)$$

$$\frac{d^2}{d\alpha^2} \log L(\hat{\alpha}) = \frac{2n}{\hat{\alpha}^2} - \frac{2}{\hat{\alpha}^3} \sum_{i=1}^{n} x_i = \frac{2n}{\hat{\alpha}^2} - \frac{4n}{\hat{\alpha}^2} = -\frac{2n}{\hat{\alpha}^2} < 0$$

we then have the estimator $\hat{\alpha} = \frac{1}{2n} \sum_{i=1}^{n} x_i$, and for the given data, the estimate:

$$\hat{\alpha} = \frac{1}{10}(x_1 + x_2 + x_3 + x_4 + x_5) = \frac{1}{10}(0.25 + 0.75 + 1.50 + 2.5 + 2.0) = 0.7$$