(c) $k(x,x') = (\langle x,x'\rangle + \vartheta)^d = (x^T x' + \vartheta)^d$ $x^T x'$ is the inner product in space R^d and $\langle x,x'\rangle$ is a Mercer kernel. From (b)(i) we know that the sum of two kernels is a Mercer kernel, and therefore $(\langle x,x'\rangle + \vartheta)$ is also a Mercer kernel. Since $(\langle x,x'\rangle + \vartheta)^d$ is $(\langle x,x'\rangle + \vartheta)$ multiplied d times, it follows from (b)(ii) that $(\langle x,x'\rangle + \vartheta)^d$ is a Mercer kernel.

(d) A Gaussian kernel of width σ can be expanded into the following two power series:

$$\cancel{k(x,y)}\ k(x,x') = \exp\left\{-\frac{\|x\|^2}{2\sigma^2}\right\}\left(\sum_{i=0}^{\infty} c_i \left(\frac{\langle x,x'\rangle}{\sigma^2}\right)^i\right)\exp\left\{-\frac{\|x'\|^2}{2\sigma^2}\right\}.$$

using the results above it is easy to see that $k(x,x') = \exp\left\{-\frac{\|x-x'\|^2}{2\sigma^2}\right\}$ is a Mercer kernel.

Exercise 2: The Feature Map

(a) $k(x,y) = \langle \varphi(x)\cdot\varphi(y)\rangle = \begin{pmatrix} x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \end{pmatrix} (y_1^2\ \ \sqrt{2}y_1y_2\ \ y_2^2) = x_1^2y_1^2 + 2x_1y_1x_2y_2 + x_2^2y_2^2$

$= \left(\sum_{i=1}^{2} x_iy_i\right)^2 = \langle x,y\rangle^2$

so that $\varphi\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \end{pmatrix}$ are possible choices for feature space and feature map

$F = R^3$ and $\varphi: R^2 \to R^3$ and

(b)