1. Implement an SVM to classify 400 patterns drawn from the $[-1,1]^2$ area. The patterns are assigned to classes according to the following specification:

$$\text{If } x^{(1)} + x^{(2)} \leq -0.5, \text{ or } x^{(1)} + x^{(2)} \geq 0.5, \text{ then } x \in C_1,$$

$$\text{If } -0.5 < x^{(1)} + x^{(2)} < 0.5, \text{ then } x \in C_2.$$  

Draw 200 patterns from class 1 and 200 patterns from class 2. Use a polynomial kernel of degree 2 classifier. You do not need to add a bias dimension (dummy feature of one) for each pattern. Consequently, if the original feature vector $x = [x^{(1)}; x^{(2)}]^T$, then $\Phi(x) = [x^{(1)}x^{(1)}, x^{(2)}x^{(2)}, \sqrt{2}x^{(1)}x^{(2)}]$. Plot $(\psi_0 \ast \Phi(x)) + b_0 \forall x \in [-1,1]^2$ as an image in both 2D and 3D. [For each $x \in \mathbb{R}^2, \Phi(x) \in \mathbb{R}^3$.] What conclusions can you draw from the plots? I suggest you use one of the SVM implementations available at http://www.support-vector-machines.org/SVM_soft.html. There are also MATLAB SVM implementations that you can use.

2. Verify the results of (6.13) to (6.22) for constructing valid kernels.

3. Bishop 6.15. The $2 \times 2$ Gram matrix in the question is $G = \begin{bmatrix} \langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle \end{bmatrix}$.

4. Bishop 6.3