Homework 3

Problem. 1. (10 points)

(A) Consider any Mercer kernel defined by $k(x, \tilde{x}) = \phi(x)^{\top}\phi(\tilde{x})$. We are given a sample $S = \{x_1, x_2, ..., x_n\}$ of n inputs. We can form the Kernel (Gram) matrix $\mathbf{K}$ as an $n \times n$ matrix of kernel evaluations between all pairs of examples i.e., $\mathbf{K}_{i,j} = k(x_i, x_j)$. **Mercer's Theorem** states that a symmetric function $k(.,.)$ is a kernel iff for any finite sample S the kernel matrix $\mathbf{K}$ is positive semi-definite. Recall that a matrix $\mathbf{K} \in \mathbb{R}^{n \times n}$ is positive semi-definite iff $\mathbf{c}^{\top}\mathbf{K}\mathbf{c} \geq 0$ for all real-valued vectors $\mathbf{c} \in \mathbb{R}^n$. Prove Mercer's theorem in one direction: for any Mercer kernel $k(.,.)$ and finite sample S, the kernel matrix $\mathbf{K}$ is positive semi-definite.

Given any two Mercer kernels $k_1(.,.)$ and $k_2(.,.)$, prove that the following are also Mercer kernels:
a) $k(x, \tilde{x}) = \alpha k_1(x, \tilde{x}) + \beta k_2(x, \tilde{x})$ for $\alpha, \beta \geq 0$
b) $k(x, \tilde{x}) = k_1(x, \tilde{x}) \times k_2(x, \tilde{x})$
c) $k(x, \tilde{x}) = f(k_1(x, \tilde{x}))$ where f is any polynomial with positive coefficients
d) $k(x, \tilde{x}) = \exp(k_1(x, \tilde{x}))$

(B) For the kernel $K(x, y) = \exp(-\frac{1}{2} \| x - y \|^2) = \varphi(x) \cdot \varphi(y)$, write an explicit formula for φ.

Problem. 2. (20 points)

You will build an SVM to classify data and use cross-validation to find the best SVM kernel and regularization value. Try different polynomials and RBF kernels (varying polynomial order from 1 to 5) and varying sigma in the RBF kernel. Also, try different values of C in the SVM. First, extract the support vector machine Matlab code from Steve Gunn's software package here:

```
http://www.isis.ecs.soton.ac.uk/resources/svminfo/
In svc.m replace [alpha lambda how] = qp(...);
with              [alpha lambda how] = quadprog(H,c,[],[],A,b,vlb,vub,x0);
```

Clearly denote the various components and the function calls or scripts that execute your Matlab functions. Note, to save the current figure in Matlab as a postscript file you can type:

```
print -depsc filename.eps
```

To test your SVM, you will build a simple object recognition system. Download the data file dataset.mat. This loads a matrix X of size 100 x 4 and a vector Y of size 100 x 1. The X matrix consists of 100 data points of dimensionality 4 that belong to one of two classes. Their actual labels are given in the Y vector. Train your SVM on half of the examples and test on the other half (or other random subsets of the examples if you see fit). Show performance of the SVM for linear, polynomial and RBF kernels with different settings of the polynomial order, sigma and C value. Try to hunt for a good setting of these parameters to obtain high recognition accuracy.

Problem. 3. (20 points)

Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. samples from a population with pdf

$$f(x|\alpha) = \alpha^{-2}xe^{-x/\alpha}, \quad x > 0, \alpha > 0.$$

Find the maximum likelihood estimator for α. Calculate an estimate using this estimator when $x_1 = 0.25, x_2 = 0.75, x_3 = 1.50, x_4 = 2.5, x_5 = 2.0$.