STA561/CS571 — Fall 2015 **Homework 2** Due: September 23, 2015

Instructions:

Problem 1

Positive (semi) definite.

1. Consider a symmetric matrix $\mathbf{K}$ that is $n \times n$. One definition of the matrix being positive definite or positive semindefinite respectively is if for any vector $\alpha \in \mathbb{R}^n$ the following holds
$$\alpha^T \mathbf{K} \alpha > 0, \quad \alpha^T \mathbf{K} \alpha \geq 0.$$
Show that (a) the matrix $\mathbf{K}$ has real eigenvalues, (b) if all of the eigenvalues are strictly positive then $\alpha^T \mathbf{K} \alpha > 0$, and (c) if all of them are strictly non-negative then $\alpha^T \mathbf{K} \alpha \geq 0$.
What do the eigenvalues of the kernel matrix $\mathbf{K}$ tell us about the matrix being positive (semi) definite.

2. One definition for a symmetric kernel function being positive definite is that $k(u, v) = k(v, u)$ with $u, v \in \mathbb{R}^p$ is a map from $\mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$ with the following property
$$\forall n \in \mathbb{N}, \quad u_i, \ldots, u_n \in \mathbb{R}^p, \quad \mathbf{K} \text{ is positive (semi)definite where } \mathbf{K}_{ij} = k(u_i, u_j).$$
Explain the above in words, provide three examples of a kernel function $k(u, v)$ that satisfies the above condition, and provide an example of a function $h(u, v)$ that does not satisfy the above condition.

Problem 2

Reproducing properties.

Given the following penalized loss function
$$\arg \min_{f \in \mathcal{H}_K} \left[n^{-1} \sum_{i=1}^{n} \log(1 + \exp(-y_i \, f(x_i))) + \lambda \|f\|_K^2 \right].$$

(a) Show that the minimizer of the above takes the form
$$\hat{f}(x) = \sum_i \alpha_i k(x, x_i).$$

(b) Write out the optimization above in the form of matrix vector operations given the form the minimizer takes (a linear combination of kernel functions).

(c) Assume that $k(u,v) = (1+u^T v)^2$ and that $u, v \in \mathbb{R}^2$. Write out an expansion of

$$\hat{f}(x) = \sum_i \alpha_i k(x, x_i),$$

as a function of coordinates of x and x_i, there are two coordinates in this example since $x \in \mathbb{R}^2$.

Problem 3

Multivariate normal.

Given a vector $x \in \mathbb{R}^p$ the multivariate normal density is

$$f(x) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right).$$

Now split the vector into two parts

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \text{of size } \begin{bmatrix} q \times 1 \\ (p-q) \times 1 \end{bmatrix},$$

and

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \quad \text{of size } \begin{bmatrix} q \times q & q \times (p-q) \\ (p-q) \times q & (p-q) \times (p-q) \end{bmatrix}.$$

(a) Prove that the following are the the joint and marginal distributions

$$x_1 \sim \mathrm{N}(\mu_1, \Sigma_{11}), \quad x_2 \sim \mathrm{N}(\mu_2, \Sigma_{22}), \quad x \sim \mathrm{N}(\mu, \Sigma).$$

(b) Prove that the conditional density takes the form

$$x_1 \mid x_2 \sim \mathrm{N}\left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right).$$

Problem 4

Support vector machine.

Consider the SVM with the constant term

$$f(x) = \sum_i K(x, x_i) \text{ inxtead of } f(x) = \sum_i K(x, x_i) + b.$$

(a) Write the primal and dual programs for the support vector machine without an off set term.

(b) Write the KKT conditions.

Problem 5

Bayesian regression. For this problem, please download the data sets: test.txt, train.txt, samples.txt. This data consists of training data and independent data of gene expression for two types of leukemia (AML and ALL) the samples.txt describes the class labels.

(a) Build a Bayesian linear model on the training data and make predictions on the test data, also show the posterior predictive distribution for each point in the test set.

(b) Redo the analysis in the previous step for varying values of τ where τ is the hyper-parameter of the normal prior on β.