EE380L: Data Mining

Homework 0

(Due Th, Jan 19th, 11:55 PM central)

1. Have you read through the class syllabus, noted the important dates, and the class policies? (yes or no).

2. Which of the following courses have you taken?
   (i) Have you taken any course on Probability/Statistics? If yes, please write down the course department and course name.
   (ii) Have you taken any course on Linear Algebra? If yes, please write down the course department and course name.
   (iii) Have you taken any course on Optimization? If yes, please write down the course department and course name.
   (iv) Have you taken any courses on Data Mining/Pattern Recognition/Machine Learning? If yes, please write down the course department(s) and course name(s).

3. Consider a 3-valued random variable X such that $P(X = 1) = 0.4$, $P(X = 0) = 0.4$ and $P(X = -1) = 0.2$. Assume you have access to a program A that generates a number in $[0, 1]$ uniformly at random. Describe how you can use A to draw random samples of X.

4. Let $X \in \mathbb{R}^{n \times p}$ and $y \in \mathbb{R}^n$ be given. The goal is to find a $w^* \in \mathbb{R}^p$ which solves the following problem:
$$\min_{w \in \mathbb{R}^p} \frac{1}{2}\|y - Xw\|^2 + \frac{c}{2}\|w\|^2 ,$$
where $c > 0$ is a constant. Give a closed form expression for w^* in terms of X, y and c. (Consult the *Matrix Cookbook* if you want to look up expressions for derivatives in matrix/vector form.)

5. Let A be a $n \times n$ positive definite matrix. The solutions to the following problems
$$\max_{w \in \mathbb{R}^n : w^T w \leq 1} w^T A w \quad \text{and} \quad \min_{w \in \mathbb{R}^n : w^T w \leq 1} w^T A w \tag{1}$$
have well known names—do you know what the solutions to these problems are called? (You can refer back to your Linear Algebra course if needed)

6. What is the probability density function $p(x; \mu, \Sigma)$ of a multivariate Gaussian distribution with mean μ and covariance Σ? Please provide an expression in terms of $\mathbf{x}, \mu, \Sigma$, and clearly define any special function you use in the expression.

   Give one example each of (i) a marginal distribution and (ii) a conditional distribution, that can be obtained from a a bivariate Gaussian distribution.

   Let $\Theta = \Sigma^{-1}$ be the precision or inverse covariance matrix. What is expression of the probability density function $p(x; \mu, \Theta^{-1})$ of a multivariate Gaussian distribution in terms of the mean μ and precision matrix Θ?