**Multivariate Statistical Analysis Mid Term**

**November 9, 2012**

**2 pages, 14 problems, 100 points**

1. Find the eigenvalues and normalized eigenvectors for $S=\left[\begin{matrix}5&4\\4&5\end{matrix}\right]$. (7%)
2. Verify the spectral decomposition theorem for Problem 1. (7%)
3. Find $\max\_{x\ne 0}\frac{x'Sx}{x'x}$ and $\min\_{x\ne 0}\frac{x'Sx}{x'x}$, where **S** is defined in Problem 1. (7%)
4. Under what situations will the generalized variance be zero? (7%)
5. Find the maximum likelihood estimate of the $2×1$ mean vector $μ$ and the $2×2$ covariance matrix $Σ $based on the random sample $\left[\begin{matrix}3&6\\4&4\\5&7\\4&7\end{matrix}\right]$ from a bivariate normal population. (7%)
6. Set up a null hypothesis  and an alternative hypothesis, with  for bivariate normal distribution. Will the Hotelling’s test rejects the null hypothesis at the level of significance? Suppose that the sample size is 30, the sample mean, and the sample covariance matrix . (7%)
7. Find the semi-major and semi-minor axes and directions of the Hotelling’s 90% confidence region for Problem 6. (7%)
8. Determine the 90% simultaneous  confidence intervals for Problem 6. (7%)
9. Determine the 90% Bonferroni simultaneous confidence intervals for Problem 6. Note that if your *t* corresponding to the desired ** can not be found directly in the table, use the following linear interpolation formula to get an approximate value: , where  and  are the two *t* values corresponding to **1 and **2, respectively, with **1 and **2 being the two ** values in the table and closest to the desired **. (7%)
10. Explain briefly why the lengths of the one-at-a-time confidence intervals are often shorter than those of the corresponding  simultaneous confidence intervals. (7%)
11. Given the data  with missing components, use the prediction-estimation algorithm (the EM algorithm) to estimate  and . Determine the initial estimate and iterate to find the *first* revised estimates. (7%)
12. Bars of soap are manufactured in each of two ways. The characteristics *X*1 = lather and *X*2 = mildness are measured. The summary statistics for 61 and 61 bars produced by methods 1 and 2, respectively, are , , , . . Test the hypothesis that the population means ******1 = ******2 at 10% significance level, using the spooled covariance matrix to estimate the population covariance matrix ****and assuming that both populations are with multivariate normal distribution. (7%)
13. Find the 90% Bonferroni simultaneous confidence interval for the individual mean difference in Problem 12. (7%)
14. Explain why paired comparison of univariate means is more accurate than unpaired one by considering their confidence intervals of significance ** for a pair of paired samples. In other words, use individual formulae of the confidence intervals for the same paired samples, and compare their lengths. Assume that the number of pair of samples *n* is not too small. (9%)