Stat 542
Homework 3
Assigned 01/29/10
Due 02/08/10

1. Wishart generation:

   (a) Write an R function to sample from the Wishart distribution. The only inputs to your function should be a positive definite matrix `Sigma` and the degrees of freedom `m`. Send me your code in plain text in the body of an email (no attachments).

   (b) Using your function, generate 10,000 matrices Σ from the inverse-Wishart prior distribution, so that $\Sigma^{-1} \sim \text{Wishart}(\mathbf{I}_{p\times p}, p+1)$. For each Σ, compute $\mathbf{C}$, the associated correlation matrix. Store the off-diagonal elements of your 10,000 simulated $\mathbf{C}$-matrices, and then make a histogram of them. What can you say about the prior distribution of the correlations, when using an inverse-Wishart$(\mathbf{I}, p+1)$ prior distribution for Σ?

2. Correlations: The file `wasl_norm` contains sample data on the percentage of students passing the 10th grade math, reading, writing and science tests for 292 schools in the state of Washington. Also included are data on the student-teach ratio, the percentage of students on the free-lunch program, the average years of educational experience of teachers in each school, and the percentage of teachers with a masters degree. Read this data in using the `dget` command.

   (a) Compute the unbiased estimate of population covariance matrix, and the correlation matrix corresponding to this estimate.

   (b) Let $a = \{1, 2, 3, 4\}$ and $b = \{5, 6, 7, 8\}$ be two subsets of variable indices. Based on $\hat{\Sigma}$, compute an estimate of $\hat{\Sigma}_{a|b}$ and the corresponding correlation matrix. Compare to $\hat{\Sigma}_a$ and its corresponding correlation matrix, and comment on any differences.

   (c) Compute the multiple correlation coefficient for each variable in a conditional on the variables in b. Comment on the strength of the relationships between the test score variables (a), and the school characteristics (b).

3. Wishart matrices: Let $\mathbf{M} \sim \text{Wishart}_p(\Sigma, m)$

   (a) Let $\mathbf{A}$ be a $p \times q$ matrix with $q \leq p$. Find the distribution of $\mathbf{A}^T\mathbf{M}\mathbf{A}$.

   (b) Let $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^p$, with $\boldsymbol{b}^T\Sigma\boldsymbol{c} = 0$. Show that $\boldsymbol{b}^T\mathbf{M}\boldsymbol{b}$ is independent of $\boldsymbol{c}^T\mathbf{M}\boldsymbol{c}$.

   (c) Show that $m_{i,i}$ and $m_{j,j}$ are independent if $\sigma_{i,j} = 0$.

4. Matrix normal: Let $\mathbf{Z}$ be an $n \times p$ matrix whose rows are i.i.d. multivariate normal$(\mathbf{0}, \mathbf{I})$.

   (a) Show that the joint density of $\mathbf{Z}$ depends on $\mathbf{Z}$ only through $\text{tr}(\mathbf{Z}\mathbf{Z}^T)$

   (b) Let $\mathbf{X} = \mathbf{Z}\mathbf{A} + \mathbf{1}\boldsymbol{b}^T$, where $\mathbf{A}$ is a full-rank $p \times p$ matrix and $\boldsymbol{b}$ is $p \times 1$. Find the form of the joint density for $\mathbf{X}$. Describe the distribution of $\mathbf{X}$ in terms of the means and covariances of its elements.

(c) Let $\mathbf{X} = \mathbf{AZ} + \boldsymbol{b}\mathbf{1}^T$, where $\mathbf{A}$ is a full rank $n \times n$ matrix and $\boldsymbol{b}$ is $n \times 1$. Find the form of the joint density for $\mathbf{X}$. Describe the distribution of $\mathbf{X}$ in terms of the means and covariances of its elements.

(d) Let $\mathbf{X} = \mathbf{AZB} + \mathbf{C}$, where $\mathbf{A}$ and $\mathbf{B}$ are square matrices and $\mathbf{C}$ is $n \times p$. Find the form of the joint density for $\mathbf{X}$. Describe the distribution of $\mathbf{X}$ in terms of the means and covariances of its elements. In particular, find

i. $\mathrm{E}[x_{i,j}]$

ii. $\mathrm{Var}[x_{i,j}]$

iii. $\mathrm{Cov}[x_{i,j}, x_{i,k}]$, $\mathrm{Cov}[x_{i,j}, x_{k,j}]$, $\mathrm{Cov}[x_{i,j}, x_{k,l}]$