Stat 542
Homework 2
Assigned 01/20/10
Due 01/29/10

1. Covariance and independence 1. Let $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2)^T$ be a multivariate normal random vector and let $\Sigma_{1,2}$ be the upper-right corner of $\text{Cov}[\boldsymbol{X}]$ as in the notes. Using characteristic functions, show that $\Sigma_{1,2} = \boldsymbol{0}$ implies that $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are independent.

   Write out the characteristic function of $\boldsymbol{X}$ from Theorem 2.6.1 from the book, partition the terms, and show the joint characteristic function of $\{\boldsymbol{X}_1, \boldsymbol{X}_2\}$ factors into the the (marginal) characteristic functions of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. From the notes, this implies the joint density $f(\boldsymbol{X}_1, \boldsymbol{X}_2)$ can be factored into $f(\boldsymbol{X}_1)f(\boldsymbol{X}_2)$.

2. Covariance and independence 2. Let (X, Y) be two random variables with joint density

$$f(x,y) = (2\pi^3)^{-1/2}(x^2+y^2)^{-1/2}e^{-(x^2+y^2)/2}.$$

   (a) Make a sketch of the contours of this density.
   (b) Are X and Y independent? Prove why or why not.
   (c) Compute $\text{Cov}[X, Y]$.

   (a) The density depends on x an y only through $r^2 = x^2+y^2$, and so the density is spherically symmetric and has spherical (circular) contours

   (b) X and Y are not independent. Most people just said the density could not be factored as $f(x)f(y)$, but did not prove this. One way to prove they are not independent is to show $f(x|y)$ depends on y. A simple way to do this is to plot $f(x|y_1)$ and $f(x|y_2$ and show that they have different shapes. Basically, the bigger y is, the less concentrated $f(x|y)$ is around zero.

   (c) The covariance is zero. One way to show this is to break up $\mathbb{R}^2$ into the four quadrants $Q_1 = (x > 0, y > 0)$, $Q_2(x > 0, y < 0)$, $Q_3 = (x < 0, y > 0)$ and $Q_4 = (x < 0, y < 0)$, and show that

$$\int_{Q_1} xyf(x,y)dxdy = -\int_{Q_2} xyf(x,y)dxdy = -\int_{Q_3} xyf(x,y)dxdy = \int_{Q_4} xyf(x,y)dxdy.$$

   This shows that $E(XY) = 0$. A similar argument shows that $E(X) = E(Y) = 0$, and so the covariance is zero, even though X and Y are not independent.

3. Graphical models 1. The following correlation matrix comes from a copula model:

$$\Sigma = \begin{pmatrix} \text{age} & \text{edu} & \text{bmi} & \text{vocab} & \text{income} \\ 1.00 & -0.03 & 0.70 & 0.00 & 0.57 \\ -0.03 & 1.00 & 0.00 & 0.73 & 0.56 \\ 0.70 & 0.00 & 1.00 & 0.03 & 0.43 \\ 0.00 & 0.73 & 0.03 & 1.00 & 0.43 \\ 0.57 & 0.56 & 0.43 & 0.43 & 1.00 \end{pmatrix}$$

"edu" is educational level, "bmi" is body mass index, "vocab" is a score on a vocabulary test.

(a) Draw a graph in which the nodes are the five variable names. Draw a line between each pair (j, k) of nodes for which $|\sigma_{j,k}| > 0.25$.

(b) Assuming that the marginal means are zero and normality on some scale, $\mathrm{E}[X_j|\boldsymbol{X}_{-j}] = \boldsymbol{\beta}_j^T \boldsymbol{X}_{-j}$. Based on the above numerical correlation matrix, compute $\boldsymbol{\beta}_j$ for each $j \in \{1, \ldots, 5\}$.

(c) Draw a graph in which the nodes are the five variable names. Draw a line between each pair (j, k) of nodes for which $|\beta_{j,k}| > 0.25$. Compare to your graph in (a).

You can use the following R-code to compute the regression coefficients:

```
B[j, -j] <- C[j, -j] %*% solve(C[-j, -j])
```

Let me provide some interpretation of the graph in (c). First of all, it is not a correlation graph - it is a conditional dependence graph, that is, a graph representing how the conditional distribution of a given variable depends on the other variables. Two variables can have a positive correlation but be conditionally independent. For example, arm length and leg length may be correlated, but conditionally independent given height.

Conditional dependence graphs like the one in (c) are often used when people want to do "causal modeling." Suppose education has a causal effect on vocabulary and on income, but income and vocabulary do not affect one another. Then we expect to see a correlation between vocab and income, even though they are conditionally independent given education.

4. Graphical models 2. Let $\boldsymbol{X} \in \mathbb{R}^p$ be a random vector with density $f(\boldsymbol{x}) = f(x_1, \ldots, x_p)$. We say that X_1 and X_2 are conditionally independent given $X_3, \ldots, X_p$ if

$$f(x_1, x_2|x_3, \ldots, x_p) = f(x_1|x_3, \ldots, x_p)f(x_2|x_3, \ldots, x_p).$$

(a) Show that $f(x_1|x_2, x_3, \ldots, x_p)$ not being a function of x_2 is a necessary and sufficient condition for X_1 and X_2 being conditionally independent.

First suppose $f(x_1|x_{-1})$ doesn't depend on x_2. In this case, $f(x_1|x_2, \ldots, x_p) = f(x_1|x_3, \ldots, x_p)$. Therefore

$$\begin{aligned} f(x_1, x_2|x_3, \ldots, x_p) &= f(x_1|x_2, \ldots, x_p)f(x_2|x_3, \ldots, x_p) \quad \text{(always)} \\ &= f(x_1|x_3, \ldots, x_p)f(x_2|x_3, \ldots, x_p) \quad \text{(from the assumption)} \end{aligned}$$

which shows conditional independence. Going the other way, assume conditional independence. Then

$$\begin{aligned} f(x_1|x_2, \ldots, x_p) &= \frac{f(x_1, x_2|x_3, \ldots, x_p)}{f(x_2|x_3, \ldots, x_p)} \quad \text{(always)} \\ &= \frac{f(x_1|x_3, \ldots, x_p)f(x_2|x_3, \ldots, x_p)}{f(x_2|x_3, \ldots, x_p)} \quad \text{(by conditional independence)} \\ &= f(x_1|x_3, \ldots, x_p). \end{aligned}$$

which gives the result.

Now suppose $\boldsymbol{X}$ is multivariate normal with mean zero and covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}^{-1}$.

(b) Express $\mathrm{E}[X_j|\boldsymbol{X}_{-j}]$ and $\mathrm{Var}[X_j|\boldsymbol{X}_{-j}]$ in terms of the elements of $\boldsymbol{\Lambda}$ and $\boldsymbol{X}_{-j}$.

(c) Letting $\beta_{j,k}$ be the regression coefficient for X_k in the conditional expectation $\mathrm{E}[X_j|\boldsymbol{X}_{-j}]$, find the relationship between $\beta_{j,k}$ and $\beta_{k,j}$ in terms of the elements of $\boldsymbol{\Lambda}$. Must these coefficients always have the same sign? Why or why not?

(c) Find conditions on $\boldsymbol{\Lambda}$ such that X_1 and X_2 are conditionally independent, given $X_3, \ldots, X_p$.

For part (b), use the identity on p.638 in the appendix to relate $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}^{-1}$ to various quantities that make up the conditional mean and variance formulas. From this, you should get that "regression coefficient" can be written $-\lambda_{j,j}^{-1}\boldsymbol{\Lambda}_{j,-j}$ and the conditional variance is $\lambda_{j,j}^{-1}$. The coefficient for variable k in the conditional expectation of j is then $-\lambda_{j,k}/\lambda_{j,j}$. and the coefficient for j in that of k is $-\lambda_{j,k}/\lambda_{k,k}$. Since $\boldsymbol{\Sigma}$ is positive definite, so is $\boldsymbol{\Lambda}$, and so $\lambda_{j,j}$ and $\lambda_{k,k}$ are both positive. Thus the coefficients must have the same sign. Finally, from this we also see that X_j and X_k are conditionally independent if $\lambda_{j,k} = 0$.

Some interpretation: $\boldsymbol{\Lambda}$ is called the *precision* matrix. Our results say that, even if we knew all variables except j, our variance (uncertainty) about j is $\lambda_{j,j}^{-1}$, or that the "maximal precision" for X_j is $\lambda_{j,j}$. Our last result says that zeros in the precision matrix directly indicate conditional independencies in the data.