269
Homework 1 Solutions

Question 1

See file: list data and regression.R

Comments:

1. Much applied work has to worry about singular or near singular X'X matrices. This applies even outside of linear models. As Bayesians, we <u>never</u> have to worry <u>if</u> we use proper priors, i.e. X'X + A is always well conditioned
2. I used the QR decomposition to do the least squares computation. Generally, a good idea, but slower than the "direct" method using Cholesky roots.

Cholesky root method

Compute

i.) $\hat{B} = (X'X)^{-1} X'Y$

ii.) root of $(X'X)^{-1}$ (simulation)

iii.) $E = Y - X\hat{B}$ and $S = E'E$

Compute Cholesky root of $(X'X)$

$(X'X) = U'U$ `chol(crossprod(X))`

$(X'X)^{-1} = U^{-1}(U^{-1})'$

Find U^{-1} by backsolving triangular system.

```
IU = backsolve(U,diag(ncol(U)))
```

i.) $\hat{B} = IUIU' X'Y$ `crossprod(t(IU))%*%crossprod(X,Y)`

ii.) IU can be used for simulation

iii.) $E = Y - X\hat{B}$ `S = crossprod(E)`

QR method

$X_{nxk} = Q_{nxk} R_{nxk}$ X has full column rank

Columns of Q are orthogonal

$Q = [q_1 ... q_k]$ $q_i' q_i = 1$

$q_i' q_j = 0$

R is the Cholesky root up to a sign transformation.

$$
\begin{aligned}
(Y - XB)'(Y - XB) &= Y'Y - Y'XB - B'X'Y + B'X'XB \\
&= Y'Y - Y'QRB - B'R'Q'Y + B'R'Q'QRB \\
&= Y'Y - Y'QQ'Y + Y'QQY - Y'QRB - B'R'Q'Y + B'R'RB \\
&= Y'(I - QQ')Y + (RB - QY)'(RB - QY)
\end{aligned}
$$

$\Rightarrow$ $\hat{B}$ solves

$$
\begin{aligned}
R\hat{B} &= QY \\
\hat{B} &= R^{-1}QY
\end{aligned}
$$

i.) $\hat{B} = R^{-1}QY$ [`q=qr(X)`, `B̂=qr.coef(q,Y)`]

ii.) $S = Y'(I - QQ')Y$ [`S = crossprod(qr.resid(q, Y)`]

iii.) root of $(X'X)^{-1}$ - additional computations!

Question 2

$$
\begin{vmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{vmatrix} = l_{11}l_{22}
$$

so true for 2 x 2 case.

Assume also true for (n-1) x (n-1) case.

expanding first row by co-factors

$$
\begin{vmatrix} l_{11} & 0 & 0 \\ & \dots & 0 \\ l_{n1} & & l_{nn} \end{vmatrix} = l_{11} \begin{vmatrix} l_{22} & 0 & 0 \\ & \dots & 0 \\ l_{n2} & & l_{nn} \end{vmatrix} + 0 \left| \quad \right| + \dots + 0 \left| \quad \right|
$$

$\Rightarrow$ true for n x n case

Question 3

a. $Z \sim N(0, I)$ $\quad$ $y = Lz$ $\quad$ $Var(y) = LIL' = LL'$

b.

$$
\begin{bmatrix} l_{11} & & \\ l_{21} & l_{22} & \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ & l_{22} & l_{32} \\ & & l_{33} \end{bmatrix} = \begin{bmatrix} \sigma_{11} & & \\ \sigma_{21} & \sigma_{22} & \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}
$$

$\sigma_{11} = l_{11}^2$ $\quad$ $l_{11} = \sqrt{\sigma_{11}}$

$$l_{21}l_{11} = \sigma_{21} \qquad l_{21} = \frac{\sigma_{21}}{l_{11}} = \frac{\sigma_{21}}{\sqrt{\sigma_{11}}}$$

$$l_{21}^2 + l_{22}^2 = \sigma_{22} \qquad l_{22}^2 = \sigma_{22} - l_{21}^2 \qquad l_{22} = \sqrt{\sigma_{22} - \frac{\sigma_{21}^2}{\sigma_{11}}}$$

$$l_{31}l_{11} = \sigma_{31} \qquad l_{31} = \frac{\sigma_{31}}{\sqrt{\sigma_{11}}}$$

$$l_{31}l_{21} + l_{32}l_{22} = \sigma_{32} \qquad l_{32} = \frac{\sigma_{32} - l_{31}l_{21}}{l_{22}} \qquad l_{32} = \frac{\sigma_{32} - \frac{\sigma_{31}}{\sqrt{\sigma_{11}}}\frac{\sigma_{21}}{l_{11}}}{\sqrt{\sigma_{22} - \frac{\sigma_{21}^2}{\sigma_{11}}}}$$

$$l_{31}^2 + l_{32}^2 + l_{33}^2 = \sigma_{33}$$
$$l_{33}^2 = \sigma_{33} - l_{31}^2 - l_{32}^2$$
$$l_{33} = \sqrt{\sigma_{33} - l_{31}^2 - l_{32}^2}$$

Note that l_{ij} are defined recursively, i.e. 2nd row depends on the first etc.

c.

We can orthogonalize $y_1, \dots, y_p$ $Var\left(\underline{y}\right) = \Sigma$ via a series of regressions

$$y_1 = e_1$$

regress

$$y_2 \; on \; y_1 \quad \rightarrow \quad e_{2.1}$$
$$y_3 \; on \; y_1, y_2 \quad \rightarrow \quad e_{3.1,2}$$
$$\dots$$
$$y_p \; on \; y_1, \dots, y_{p-1} \quad \rightarrow \quad e_{p.1,2,\dots,p-1}$$

$\left(e_1, e_{2.1}, e_{3.1,2}, \dots, e_{p,1,2,\dots,p-1}\right)$ are uncorrelated (Gram-Schmidt)

But not with unit variance.

$$y_1 = \sigma_1 z_1$$
$$-\beta_{21}y_1 + y_2 = \sigma_2 z_2$$
$$\dots$$
$$\underbrace{-\beta_{p.1}y_1 - \beta_{p.2}y_2 - \dots - \beta_{p.p-1}y_{p-1} + y_p}_{e_{p.1,2,\dots,p-1}} = \sigma_p z_p$$

This can be written:

$\underline{Ty} = Dz$

$$T = \begin{pmatrix} 1 & \cdots & & 0 \\ \vdots & \ddots & \vdots & \\ -\beta_{p,1} & \cdots & -\beta_{p,p-1} & 1 \end{pmatrix}$$

$$D = diag\left(\sigma_1, \ldots, \sigma_p\right)$$

$$\underline{Ty} = D\underline{z}$$

$$y = \underbrace{T^{-1}}_{L \text{ by direct construction}} Dz$$

How do we know T^{-1} is lower triangular?
Write out $TT^{-1} = I$

$$\begin{pmatrix} t_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ t_{p1} & \cdots & t_{pp} \end{pmatrix} \begin{pmatrix} t^{11} & \cdots & t^{1p} \\ \vdots & \ddots & \vdots \\ t^{p1} & \cdots & t^{pp} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$$

row1, col2:

$$t_{11}t^{12} = 0$$

$$\Rightarrow t^{12} = 0 \quad t_{11} \neq 0$$

row1, col3:

$$t_{11}t^{13} = 0$$

$$\Rightarrow t^{13} = 0$$

$\cdots$

row2, col3:

$$t_{21}t^{23} = 0 \qquad \left(t^{13} = 0\right)$$

$$\Rightarrow t^{23} = 0$$

...