MAURICE DIESENDRUCK

Lin Mod M2 : 1

1A (i) Under M1, $S(Y) = \dfrac{Y'(M-M_1)Y / r(M-M_1)}{Y'(I-M)Y / r(I-M)}$

$$\sim F\left(r(M-M_1),\ r(I-M),\ \beta'X'(M-M_1)X\beta / 2\sigma^2 \right)$$

according to Theorem 3.2.1.

(*) Here, $r(M-M_1) = (p+q) - p = q$, and $r(I-M) = n-(p+q) = n-p-q$.

So $S(Y) = \dfrac{Y'(M-M_1)Y/q}{Y'(I-M)Y/(n-p-q)} \sim F\left(q,\ n-p-q,\ \beta'X'(M-M_1)X\beta / 2\sigma^2 \right)$

(ii) Under M∅, (*) still applies, and by the same theorem,

$$S(Y) \sim F\left(r(M-M_1),\ r(I-M),\ \emptyset \right) = F\left(q,\ n-p-q,\ \emptyset \right).$$

1B

Show that the LRT $\equiv$ a test that rejects M$\emptyset$ for large values of $S(Y)$.

LRT rejects M$\emptyset$ when $L(\hat{\theta}_0)$ is not sufficiently bigger than $L(\hat{\theta})$. Using c, the LRT states that it will reject M$\emptyset$ if the likelihood under M$\emptyset$ is lower than the likelihood under M1.

$$\text{LRT: Reject if } \frac{L(\hat{\theta}_0)}{L(\hat{\theta})} < c$$

Lower $L(\hat{\theta}_0)$ means that M$\emptyset$ didn't describe the data as well as the other model, and has larger errors than the other model.

Larger errors under M$\emptyset$ $\Rightarrow$ Larger $Y'(M-M_1)Y$ $\Rightarrow$ Larger $S(Y)$

$\Rightarrow$ More likely to reject M$\emptyset$

Lin Mod M2 : 2

MB

For M1 : $Y = X\beta + \varepsilon$, $\varepsilon \sim N(0, \sigma^2 I)$

Consider constraint $\Lambda'\beta = 0$ with $\Lambda' = P'X \Rightarrow P'X\beta = \emptyset$.

Show that $\beta' X' M_{MP} X\beta = \emptyset$ iff $\Lambda'\beta = \emptyset$.

"$\Leftarrow$" $\Lambda'\beta = P'X\beta = P'MX\beta = \emptyset$.

$\Rightarrow X\beta \in C(MP)^{\perp} \Rightarrow \beta' X' M_{MP} X\beta = \beta' X' \cdot \emptyset = \emptyset$

More details

Given U s.t. $C(U) = C(\Lambda)^{\perp}$, $\beta \in C(U)$ and $Y = X\beta + \varepsilon = XU\gamma + \varepsilon$

for some γ. XU can be re-written as X_1 , where $X = [X_1 \vdots X_2]$

and $C(X_1) \subset C(X)$. $X\beta = XU\gamma \in C(MP)^{\perp}$.

Therefore, $C(MP) = C(XU)^{\perp}_{C(X)} = C(X_1)^{\perp}_{C(X)} = C(M - M_1)$.

$X\beta \in C(MP)^{\perp} \equiv X\beta \in C(M - M_1)^{\perp} \Rightarrow \forall g \in C(M - M_1)$, $X\beta \cdot g = \emptyset$

$\beta' X' M_{MP} X\beta = \underbrace{\beta' X' (M - M_1) X\beta}_{\emptyset} = \emptyset$ ∎

"$\Rightarrow$" $\beta' X' M_{MP} X\beta = \emptyset$

M_{MP} is projection matrix onto $C(MP)$, so $\beta' X' (M_{MP})' M_{MP} X\beta = \emptyset$

$\Rightarrow (M_{MP} X\beta)' M_{MP} X\beta = \emptyset \Rightarrow M_{MP} X\beta = \emptyset \Rightarrow MP(P'MP)^{-} P'\underbrace{M \cdot X\beta}_{X\beta} = \emptyset$

$\Rightarrow MP(P'MP)^{-} P'X\beta = \emptyset \Rightarrow MP(P'MP)^{-} \Lambda'\beta = \emptyset$

$\Rightarrow \Lambda'\beta = \emptyset$ ∎

MD.

3A Find $p(\beta_i \mid y, \lambda_i, \tau, \sigma^2)$. First find joint distribution over vector forms of $\underline{\beta}, \underline{y}$, and $\underline{\lambda}$, then get full conditional of $\underline{\beta}$ by finding the kernel w.r.t. $\underline{\beta}$, and finally isolate just the β_i element.

Joint "J" $= P(\underline{y}, \underline{\beta}, \underline{\lambda}, \tau, \sigma^2) = N(\underline{y} \mid \underline{\beta}, \sigma^2 I) \cdot N(\underline{\beta} \mid \underline{0}, \underline{\lambda}^2 I)$

$$\cdot\, C^+(\underline{\lambda} \mid \phi, \tau) \; C^+(\tau \mid \phi, \sigma) \cdot P(\sigma^2)$$

Get kernel w.r.t. $\underline{\beta}$:

$$P(\underline{\beta} \mid \underline{y}, \underline{\lambda}, \tau, \sigma^2) \propto J \propto \exp\left\{-\frac{1}{2}\left[(\underline{y}-\underline{\beta})'(\sigma^2 I)^{-1}(\underline{y}-\underline{\beta})\right]\right\}$$

$$\times \exp\left\{-\frac{1}{2}\left[\underline{\beta}'(\underline{\lambda}^2 I)^{-1}\underline{\beta}\right]\right\}$$

$$= \exp\left\{-\frac{1}{2}\left[\underbrace{(\underline{y}-\underline{\beta})'}_{1\times n}\underbrace{(\sigma^2 I)^{-1}}_{n\times n}\underbrace{(\underline{y}-\underline{\beta})}_{n\times 1} + \underbrace{\underline{\beta}'}_{1\times n}\underbrace{(\underline{\lambda}^2 I)^{-1}}_{n\times n}\underbrace{\underline{\beta}}_{n\times 1}\right]\right\}$$

Taking just β_i, $\propto \exp\left\{-\frac{1}{2}\left[\frac{(y_i-\beta_i)^2}{\sigma^{2}} + \frac{\beta_i^2}{\lambda_i^2}\right]\right\}$

$$= \exp\left\{-\frac{1}{2}\left[\frac{y_i^2}{\sigma^2} - \frac{2y_i\beta_i}{\sigma^2} + \frac{\beta_i^2}{\sigma^2} + \frac{\beta_i^2}{\lambda_i^2}\right]\right\}$$

$$= \exp\left\{-\frac{1}{2}\left[\left(\frac{1}{\sigma^2}+\frac{1}{\lambda_i^2}\right)\beta_i^2 - 2\left(\frac{y_i}{\sigma^2}\right)\beta_i + C_1\right]\right\}$$

By completing the square, $\sim N\left(\left(\frac{1}{\sigma^2}+\frac{1}{\lambda_i^2}\right)^{-1}\left(\frac{y_i}{\sigma^2}\right), \left(\frac{1}{\sigma^2}+\frac{1}{\lambda_i^2}\right)^{-1}\right)$

$$= P(\beta_i \mid \underline{y}, \lambda_i, \tau, \sigma^2)$$

$\boxed{3B}$ Assume $\sigma^2 = \tau = 1$, let $m_i = E[\beta_i \mid y, \lambda]$. From 3A, found

$$m_i = \left(\frac{1}{\sigma^2} + \frac{1}{\lambda_i^2}\right)^{-1}\left(\frac{y_i}{\sigma^2}\right) = \left(1 + \frac{1}{\lambda_i^2}\right)^{-1} \cdot y_i = \left(\frac{1}{1 + \frac{1}{\lambda_i^2}}\right) \cdot y_i$$

$$= \left(\frac{\lambda_i^2}{\lambda_i^2 + 1}\right) \cdot y_i = \left(1 - \frac{1}{1+\lambda_i^2}\right) \cdot y_i = (1 - k_i) \cdot y_i, \quad \text{for } k_i = \frac{1}{1+\lambda_i^2}.$$

Find $P(k_i)$ by change of variable and plot it.

$$k_i = \frac{1}{1+\lambda_i^2} \Rightarrow \lambda_i = \pm\left(\frac{1}{k_i} - 1\right)^{\frac{1}{2}} \Rightarrow \frac{\partial \lambda_i}{\partial k_i} = \frac{1}{2}\left(\frac{1}{k_i} - 1\right)^{-\frac{1}{2}} (-1)(k_i^{-2})$$

$$= \frac{-1}{2k_i^2\left(\frac{1}{k_i} - 1\right)^{\frac{1}{2}}}$$

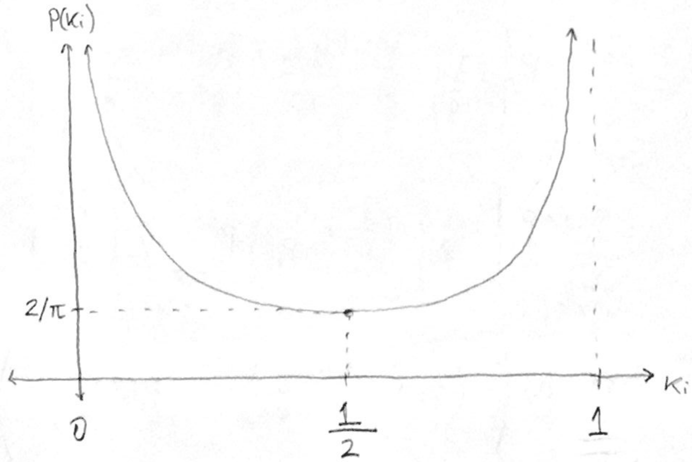
$$P_{k_i}(k_i) = P_{\lambda_i}\left(\lambda_i = \left(\tfrac{1}{k_i} - 1\right)^{\frac{1}{2}}\right) \cdot \left|\frac{\partial \lambda_i}{\partial k_i}\right|$$

$$= \frac{\cancel{2}}{\pi\left(1 + \left[\left(\frac{1}{k_i} - 1\right)^{\frac{1}{2}}\right]^2\right)} \cdot \frac{1}{\cancel{2} k_i^2 \left(\frac{1}{k_i} - 1\right)^{\frac{1}{2}}}$$

$$= \frac{1}{\pi\left(\cancel{1} + \frac{1}{k_i} - \cancel{1}\right) k_i^2 \left(\frac{1}{k_i} - 1\right)^{\frac{1}{2}}} = \frac{1}{\pi k_i \left(\frac{1}{k_i} - 1\right)^{\frac{1}{2}}} \cdot \frac{(k_i)^{\frac{1}{2}}}{(k_i)^{\frac{1}{2}}} = \frac{1}{\pi k_i^{+\frac{1}{2}} (1 - k_i)^{\frac{1}{2}}}$$

$$= \frac{1}{\pi} k_i^{-\frac{1}{2}} (1 - k_i)^{-\frac{1}{2}} = \text{Beta}\left(\tfrac{1}{2}, \tfrac{1}{2}\right) \blacksquare$$

Plot.



$$P_{k_i}\left(\tfrac{1}{2}\right) = \frac{1}{\pi}\left(\tfrac{1}{2}\right)^{-\frac{1}{2}}\left(1 - \tfrac{1}{2}\right)^{-\frac{1}{2}}$$

$$= \frac{2}{\pi}$$

Linear Models Midterm 2

Maurice Diesendruck

April 7, 2015

1 Question 3C

1.1 Strategy

Find $P(\beta_i|\tau, \sigma^2)$, still assuming $\tau = \sigma^2 = 1$. To do this, first find joint distribution $P(\beta_i, \lambda_i|\tau, \sigma)$, then marginalize out λ_i.

1.2 Joint Distribution of β_i, λ_i

$$
\begin{aligned}
P(\beta_i, \lambda_i, \tau, \sigma^2) &= P(\beta_i|\lambda_i)P(\lambda_i|\tau)P(\tau|\sigma)P(\sigma^2) \\
&\text{Note that the assumptions on } \tau, \sigma^2 \text{ imply } P(\tau|\sigma) = P(\sigma^2) \propto 1 \\
P(\beta_i, \lambda_i|\tau, \sigma^2) &\propto N(\beta_i|0, \lambda_i^2)C^+(\lambda_i|0, \tau) \\
&\propto \left((\lambda_i^2)^{-1/2} exp\Big(-\frac{1}{2}\frac{\beta_i^2}{\lambda_i^2} \Big) \right) \left(\frac{1}{1+\lambda_i^2} \right) \\
&\propto (\lambda_i^2)^{-1/2}(1+\lambda_i^2)^{-1} exp\Big(-\frac{1}{2}\frac{\beta_i^2}{\lambda_i^2} \Big)
\end{aligned}
$$

1.3 Marginalize out λ_i

$$
\begin{aligned}
P(\beta_i|\tau, \sigma^2) &= \int P(\beta_i, \lambda_i|\tau, \sigma^2) d\lambda_i \\
&\propto \int (\lambda_i^2)^{-1/2}(1+\lambda_i^2)^{-1} exp\Big(-\frac{1}{2}\frac{\beta_i^2}{\lambda_i^2} \Big) \\
&\propto \int \frac{1}{\lambda_i + \lambda_i^3} exp\Big(-\frac{1}{2}\frac{\beta_i^2}{\lambda_i^2} \Big)
\end{aligned}
$$

1.4 Numerical Integration

```
# Create grid of Betas.
betas <- seq(-3, 3, length=1000)

# Define function to integrate. Integrate function for each Beta_i.
INTEGRAND <- function(beta.i) {
  integrand <- function(x) {
    (1/(x+x^3))*(exp((-1/2)*beta.i^2/x^2))
  }
  return (integrand)
}

# Do integration on Lambda from 0 to Inf, and plot results.
len <- length(betas)
results <- matrix(0, nrow=len, ncol=1)
for (i in 1:len) {
  results[i,] <- integrate(INTEGRAND(betas[i]), lower=0, upper=Inf)$value
}

plot(betas,results, xlab="Betas",
     main=expression(paste("Numerical Evaluation of P(", beta[i], "|",
                           tau, ",", sigma^2, ")")),
     ylab=expression(paste("P(", beta[i], "|", tau, ",", sigma^2, ")")))
```