# Lecture Six

$$y_t = \beta_0 + \beta_1 t + \beta_2 (t-c)_+ + \varepsilon_t$$

## Parameter Estimation

① Calculate RSS $(c)$ for a bunch of values of $c$.

$$c \in [1, n]$$

$$c \in \{1, 2, \ldots, n\}$$

② $\hat{c}$ : value which minimizes RSS $(c)$

③ Fix $c = \hat{c}$ & run linear regression of $y$ on $[1, t, (t-\hat{c})_+] = X$ to estimate $\beta_0, \beta_1, \beta_2$ (also $\sigma$)

Least Squares: $\sum_{t=1}^{n} \left[y_t - \beta_0 - \beta_1 t - \beta_2 (t-c)_+\right]^2$

$$RSS(c) = \min_{\beta_0, \beta_1, \beta_2}$$

## Uncertainty Quantification

$$\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$$

Likelihood :

$$\prod_{t=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}\left(y_t - \beta_0 - \beta_1 t - \beta_2 (t-c)_+\right)^2\right)$$

$$\propto \sigma^{-n} \exp\left[-\frac{\sum_{t=1}^{n}\left(y_t - \beta_0 - \beta_1 t - \beta_2 (t-c)_+\right)^2}{2\sigma^2}\right]$$

$$S(\beta, c) = \sum_{t=1}^{n} \left(y_t - \beta_0 - \beta_1 t - \beta_2 (t-c)_+\right)^2$$

$$\sigma^{-n} \exp\left[-\frac{S(\beta, c)}{2\sigma^2}\right]$$ } SAME AS FOR LINEAR REGRESSION

Remark: MLE of $\beta$ & $c$ = Least squares of $\beta$ & $c$

prior: $\beta_0, \beta_1, \beta_2, \sigma, c$

$\beta_0, \beta_1, \beta_2, \log \sigma \overset{iid}{\sim} \text{Unif}[-C, C]$

$C \to \infty$

previous prior that we used in linear regression

$c \in \{$

$$y_t = \beta_0 + \beta_1 t + \beta_2 (t-c)_+ + \varepsilon_t$$

$c = 0$ $\quad \beta_0 + \beta_1 t + \beta_2 t$

$\beta_0 + (\beta_1 + \beta_2) t$

$c = 1$ : $\quad \beta_0 + \beta_1 t + \beta_2 (t-1)$

$= \beta_0 - \beta_2 + (\beta_1 + \beta_2) t$

$c = n$ $\quad (t-c)_+ = 0 \to \beta_0 + \beta_1 t$

Natural to restrict $c \in (1, n) \to$ wide region

$c \in (10, n-10) \to$ reasonable

our interest mainly lies in $c$ away from the edges.

$c \sim \text{Unif}(1, n)$

$c \sim \text{Unif}\{2, 3, \ldots, n-1\}$

reasonable

Prior: $\beta_0, \beta_1, \beta_2, \log\sigma, c$ independent

$\beta_0, \beta_1, \beta_2, \log\sigma \overset{iid}{\sim} \text{Unif}(-\infty, \infty)$, $c \sim \text{Unif}(1, n)$

$$f_{\beta_0,\beta_1,\beta_2,\sigma,c}(\beta_0, \beta_1, \beta_2, \sigma, c) \propto \frac{1}{\sigma} I\{1 < c < n\} I\{\sigma > 0\}$$

posterior $\propto$ likelihood $\times$ prior

$$\propto \sigma^{-n} \exp\left[-\frac{S(\beta, c)}{2\sigma^2}\right] \frac{1}{\sigma} I\{1<c<n\} I\{\sigma>0\}$$

$$= \sigma^{-n-1} \exp\left[\frac{-S(\beta, c)}{2\sigma^2}\right] I\{1 < c < n\} I\{\sigma > 0\}$$

$\underbrace{\beta}_{\beta_0,\beta_1,\beta_2}, c, \sigma$ To get posterior for $c$ alone, need to integrate $\beta$ & well as $\sigma$.

$$S(\beta, c) = \sum_{t=1}^{n} \left[y_t - \beta_0 - \beta_1 t - \beta_2 (t-c)_+\right]^2$$

$$= \|y - X_c \beta\|^2$$

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad X_c = \begin{bmatrix} 1 & 1 & (1-c)_+ \\ \vdots & 2 & (2-c)_+ \\ \vdots & \vdots & \vdots \\ 1 & n & (n-c)_+ \end{bmatrix}$$

$t, t = 1 \ldots, n$

Pythagorean Identity:

$$\underbrace{\|y - X_c\beta\|^2}_{S(\beta, c)} = \underbrace{\|y - X_c\hat{\beta}_c\|^2}_{S(\hat{\beta}_c, c)} + (\beta - \hat{\beta}_c)^T X_c^T X_c (\beta - \hat{\beta}_c)$$

$$S(\beta, c) = RSS(c) + (\beta - \hat{\beta}_c)^T X_c^T X_c (\beta - \hat{\beta}_c)$$

posterior for all parameters $\propto \sigma^{-n-1} \exp\left[-\frac{S(\beta, c)}{2\sigma^2}\right] I(\sigma > 0)\, I(c \in (1, n))$

$$= \sigma^{-n-1} \exp\left[-\frac{RSS(c)}{2\sigma^2}\right] \exp\left[-\frac{(\beta - \hat{\beta}_c)^T X_c^T X_c (\beta - \hat{\beta}_c)}{2\sigma^2}\right]$$

$I(\sigma > 0)$
$I(1 < c < n)$

<u>Formula</u>:

$$\int_{\mathbb{R}^p} \exp\left[-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right] dx = (2\pi)^{p/2} \sqrt{\det \Sigma}$$

$$\Sigma = \sigma^2 (X_c^T X_c)^{-1}$$

posterior for $\sigma, c$ $\propto \sigma^{-n-1} \exp\left[-\frac{RSS(c)}{2\sigma^2}\right] (2\pi)^{p/2} \sqrt{\det\left(\sigma^2 (X_c^T X_c)^{-1}\right)}$

$I(\sigma > 0)$
$I(1 < c < n)$

$$\propto \sigma^{-n+p-1} |X_c^T X_c|^{-1/2} \exp\left[\frac{-RSS(c)}{2\sigma^2}\right]$$

$I(\sigma > 0)$
$I(1 < c < n)$

$$\boxed{\det(a \underset{p \times p}{A}) = a^p \det(A)}$$

posterior for $c$: Integrate over $\sigma$:

$$\propto \int_0^\infty \sigma^{-n+p-1} |X_c^T X_c|^{-1/2} \exp\left(-\frac{RSS(c)}{2\sigma^2}\right) d\sigma$$

$$\propto |X_c^T X_c|^{-1/2} I(1<c<n) \int_0^\infty \sigma^{-n+p-1} \exp\left(-\frac{RSS(c)}{2\sigma^2}\right) d\sigma$$

$$\propto |X_c^T X_c|^{-1/2} I(1<c<n) \left[\frac{1}{RSS(c)}\right]^{\frac{n-p}{2}}$$

$p$ : # columns in $X_c$
(in our case, $p=3$)

$$\int_0^\infty \sigma^{-n-1} \exp\left(-\frac{S(\beta)}{2\sigma^2}\right) d\sigma \propto \left(\frac{1}{S(\beta)}\right)^{\frac{n}{2}}$$

change of variable

$$\frac{\sigma}{\sqrt{S(\beta)}} = t$$

posterior of $c$

$$\propto |X_c^T X_c|^{-1/2} \left(\frac{1}{RSS(c)}\right)^{\frac{n-p}{2}} I(1<c<n)$$

$$X_c = \begin{bmatrix} 1 & 1 & (1-c)_+ \\ \vdots & 2 & \vdots \\ \vdots & \vdots & \vdots \\ 1 & n & (n-c)_+ \end{bmatrix}$$

If $c=0$

$$\begin{bmatrix} 1 & 1 & 1 \\ \vdots & 2 & 2 \\ \vdots & \vdots & \vdots \\ 1 & n & n \end{bmatrix}$$

$I(2 \le c \le n-1)$

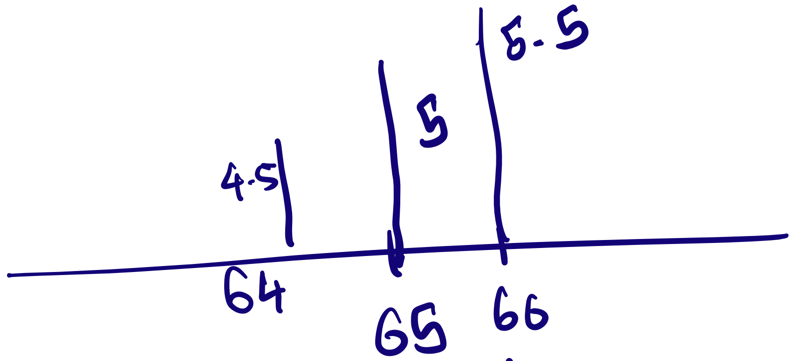
$c \in [2, n-1]$

① Take a grid of values of $c$ in $[2, n-1]$

② For each $c$ in the grid, calculate

$$|X_c^T X_c|^{-1/2} \left(\frac{1}{RSS(c)}\right)^{\frac{n-p}{2}}$$

This gives the unnormalized posterior




③ Normalize these values.

$\mathbb{P}(c = 65 \mid \text{data})$ $\quad \mathbb{P}(c = 65.05 \mid \text{data})$

---

$$y_t = \beta_0 + \beta_1 t + \beta_2 (t - c_1)_+ + \beta_3 (t - c_2)_+ + \varepsilon_t$$

$$RSS(c_1, c_2) = \min_{\beta} \|y - X_{c_1, c_2} \beta\|^2$$

$$\underset{(c_1, c_2)}{\text{posterior}} : \propto |X_{c_1,c_2}^T X_{c_1,c_2}|^{-1/2} \left(\frac{1}{RSS(c_1, c_2)}\right)^{\frac{n-p}{2}}$$

$p = 4$