Linear Regression

$$Y \mid \beta, \sigma^2 \sim N(X\beta, \sigma^2 I_n)$$

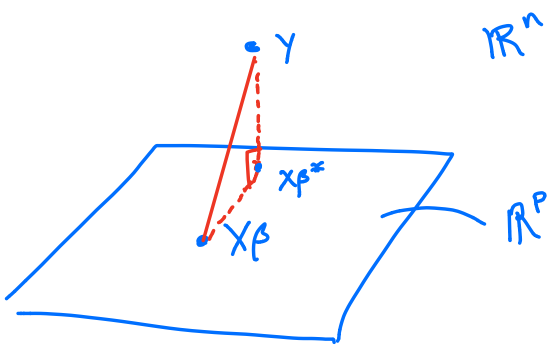
$$p(\beta, \sigma^2 \mid Y) \propto (\sigma^2)^{-n/2} \exp\left\{ \frac{-1}{2\sigma^2} (Y - X\beta)^T (Y - X\beta) \right\} \exp\left\{ -\frac{1}{2} (\beta - \beta_0)^T \Sigma_0^{-1} (\beta - \beta_0) \right\}$$

$$\times\, \sigma^{2-(\frac{\nu}{2}+1)} \exp\left(\frac{-\nu\tau^2}{2\sigma^2} \right)$$

$$p(\beta \mid \sigma^2, y) \propto \exp\left\{ -\frac{1}{2} \left[\frac{1}{\sigma^2} (Y - X\beta)^T (Y - X\beta) + (\beta - \beta_0)^T \Sigma_0^{-1} (\beta - \beta_0) \right] \right\}$$

Use centering trick with $X^T X \beta^* = X^T y$

$$(Y - X\beta)^T (Y - X\beta) = (Y - X\beta^*)^T (Y - X\beta^*) + (\beta - \beta^*)^T X^T X (\beta - \beta^*)$$



$$\frac{1}{\sigma^2} (\beta - \beta^*)^T X^T X (\beta - \beta^*) + (\beta - \beta_0)^T \Sigma_0^{-1} (\beta - \beta_0)$$

$$= \frac{1}{\sigma^2} (\beta^* - \hat{\beta})^T X^T X (\beta^* - \hat{\beta}) + (\beta_0 - \hat{\beta})^T \Sigma_0^{-1} (\beta_0 - \hat{\beta}) + (\beta - \hat{\beta})^T \Sigma_\beta^{-1} (\beta - \hat{\beta})$$

$$\Sigma_\beta^{-1} = \frac{1}{\sigma^2} X^T X + \Sigma_0^{-1} \qquad \hat{\beta} = \Sigma_\beta \left(\frac{1}{\sigma^2} X^T y + \Sigma_0^{-1} \beta_0 \right)$$

which reduces the posterior by

$$\mathbb{P}(\beta|\sigma^2,Y) \propto \exp\left\{-\frac{1}{2}(\beta-\hat{\beta})^T\Sigma_\beta^{-1}(\beta-\hat{\beta})\right\}$$

So $\beta|\sigma^2,Y \sim N(\beta,\Sigma_\beta)$

Rmk: Regularizer

Rmk: Cannot sample from $\mathbb{P}(\sigma^2|Y)$

Idea: Gibbs

1. $\beta^{(t)} \sim N(\hat{\beta},\Sigma_\beta(\sigma^{2(t)}))$

2. Sample σ^2

$$\mathbb{P}(\sigma^2|\beta,Y) \propto (\sigma^2)^{-\left(\frac{n+r}{2}+1\right)}\exp\left(-\frac{1}{2\sigma^2}(RSS(\beta)+r\tau^2)\right)$$

$$\sim \text{Inv-}\chi^2\left(n+r, \frac{RSS(\beta)+r\tau^2}{n+r}\right)$$

If we want point estimates

- Post. Mode MAP. estimator $(\beta,\sigma^2)_{MAP} = \underset{\Theta}{\arg\max}\log\mathbb{P}(\beta,\sigma^2|X)$

- Find it using "expectation maximization" (σ^2 unknown)

$$\beta^{(t+1)} = \left(\frac{1}{\sigma^{2(t)}}X^TX+\Sigma_0^{-1}\right)^{-1}\left(\frac{1}{\sigma^{2(t)}}X^Ty+\Sigma_0^{-1}\beta_0\right) \qquad \sigma^2(t+1) = \frac{RSS(\beta^{(t)})+r\tau^2}{n+r}$$

Compute $\hat{\beta}^{(t)}$

(a) $\Sigma_\beta^{-1} = C^T C$

(b) $z = \frac{1}{\sigma^{2(t)}} X^T y + \Sigma_0^{-1} \beta_0$

(c) $\hat{\beta}^{(t+1)} = \Sigma_\beta \cdot z = \underbrace{C^{-1}}\underbrace{C^{-T} z}_{u = \text{backsolve}(C,\, z,\, \text{trans} = T)}$

$C^{-1} u = \text{backsolve}(C, u)$

To sample $\beta \mid \sigma^2, Y \sim N(\hat{\beta}, \Sigma_\beta) = N\left(\Sigma_\beta \left(\underbrace{\frac{1}{\sigma^2} X^T Y + \Sigma_0^{-1} \beta_0}_{w}\right), \Sigma_\beta\right)$

$C\beta \mid \sigma^2, Y \sim N(C^{-T} w, I_n) = C^{-T} w + \underbrace{N(0, I_n)}_{\text{sample this}}$

$\beta = C^{-1}(u + z), \; z \sim N(0, I_n)$