Cap 9: Regresión lineal

Objetivo de estudio

Modelo $y_i = \beta_1 + \beta_2 (\text{Programa}_i) + \beta_3 \text{Edad}_i + \beta_4 \text{Edad}_i^2 + \varepsilon_i$

$\hookrightarrow SSR(\beta) = \sum_{i=1}^{n} (y_i - \beta^T x_i)^2 = (Y - X\beta)^T (Y - X\beta)$

$\hat{\beta}_{MCO} = (X^T X)^{-1} X^T Y \longrightarrow Var(\beta) = \sigma^2 (X^T X)^{-1}$

$\hat{\sigma} = \frac{SSR(\beta_{ols})}{n-p}$

Bayesiano para una regresión

Posterior para β (Full conditional)

Receta: i) Verosimilitud → Evaluar el modelo

ii) Prior → Modelo

$Pr(\beta | Y, X, \Sigma) = \frac{\overbrace{Pr(Y | \beta, X, \Sigma)}^{1} \overbrace{Pr(\beta | X, \Sigma)}^{2}}{\underbrace{Pr(Y, X, \Sigma)}_{3}} \to$ Objetivo

② $\sim N(\beta_0, \Sigma_0)$

$Y \in \mathbb{R}^n$

De ① $Pr(Y | \beta, X, \Sigma) = (2\pi)^{-n/2} |\Sigma|^{-n/2} \exp\left\{-\frac{1}{2} (Y - X\beta)^T \Sigma^{-1} (Y - X\beta)\right\}$

De ② $Pr(\beta | X, \Sigma) = (2\pi)^{-p/2} |\Sigma_0|^{-1/2} \exp\left\{-\frac{1}{2} (\beta - \beta_0)^T \Sigma_0^{-1} (\beta - \beta_0)\right\}$

① * ② $= K \exp\left\{-\frac{1}{2}\left[(Y - X\beta)^T \Sigma^{-1} (Y - X\beta) + (\beta - \beta_0)^T \Sigma_0^{-1} (\beta - \beta_0)\right]\right\}$

$\boxed{\text{Con } K = (2\pi)^{-n/2} (2\pi)^{-p/2} |\Sigma_0|^{-1/2} |\Sigma|^{-n/2}}$

Desarrollo de exponencial

$\Rightarrow Y^T \Sigma^{-1} Y - Y^T \Sigma^{-1} X\beta - \beta^T X^T \Sigma^{-1} Y + \beta^T X^T \Sigma^{-1} X\beta + \beta^T \Sigma_0^{-1} \beta - \beta^T \Sigma_0^{-1} \beta_0 - \beta_0^T \Sigma_0^{-1} \beta - \beta_0^T \Sigma_0^{-1} \beta_0$

$\hookrightarrow = \beta^T (X^T \Sigma^{-1} X + \Sigma_0^{-1}) \beta - 2\beta^T (X^T \Sigma^{-1} Y + \Sigma_0^{-1} \beta_0) + Y^T \Sigma^{-1} Y - \beta_0^T \Sigma_0^{-1} \beta_0$

$\hookrightarrow = [\beta - \underbrace{(X^T \Sigma^{-1} X + \Sigma_0^{-1})^{-1} (X^T \Sigma^{-1} Y + \Sigma_0^{-1} \beta_0)}_{\mu_n}]^T \underbrace{(X^T \Sigma^{-1} X + \Sigma_0^{-1})}_{\Lambda_n^{-1}} [\beta - \underbrace{(X^T \Sigma^{-1} X + \Sigma_0^{-1})^{-1} (X^T \Sigma^{-1} Y + \Sigma_0^{-1} \beta_0)}_{\mu_n}]$

$$+Y^T\Sigma^{-1}Y - B_0^T\Sigma_0^{-1}B_0 - \mu_m \Lambda_m^{-1}\mu_m$$

$$\underbrace{\qquad\qquad\qquad}_{H}$$

Reescribimos ① * ② : $K \exp\left\{-\frac{1}{2}\left[(Y-X\beta)^T\Sigma^{-1}(Y-X\beta) + (\beta-\beta_0)^T\Sigma_0^{-1}(\beta-\beta_0)\right]\right\}$

$$\Rightarrow K\exp\left\{-\frac{1}{2}\left[(\beta-\mu_m)^T\Lambda_m^{-1}(\beta-\mu_m) + H\right]\right\} = \hat{k}\exp\left\{-\frac{1}{2}(\beta-\mu_m)^T\Lambda_m^{-1}(\beta-\mu_m)\right\}$$

$$\hat{k} = K\exp\left\{-\frac{1}{2}H\right\}$$

Escribimos el posterior de β

$$P(\beta|Y,X,\Sigma) = \hat{k}\exp\left\{-\frac{1}{2}(\beta-\mu_m)^T\Lambda_m^{-1}(\beta-\mu_m)\right\} / Pr(Y,X,\Sigma)$$

$$Pr(\beta|Y,X,\Sigma) = \frac{(2\pi)^{-n/2}|\Sigma|^{-m/2}|\Sigma_0|^{-1/2}|\Lambda_m|^{1/2}\exp\left\{-\frac{1}{2}H\right\}(2\pi)^{-p/2}|\Lambda_m|^{-1/2}\exp\left\{-\frac{1}{2}(\beta-\mu_m)^T\Lambda_m^{-1}(\beta-\mu_m)\right\}}{Pr(Y,X,\Sigma)}$$

Integrando sobre β

$$Pr(Y,X,\Sigma) = (2\pi)^{-n/2}|\Sigma|^{-m/2}|\Sigma_0|^{-1/2}|\Lambda_m|^{1/2}\exp\left\{-\frac{1}{2}H\right\}$$

Finalmente:

$$Pr(\beta/Y,X,\Sigma) = (2\pi)^{-p/2}|\Lambda_m|^{-1/2}\exp\left\{-\frac{1}{2}(\beta-\mu_m)^T\Lambda_m^{-1}(\beta-\mu_m)\right\} \Rightarrow \text{Posterior Full conditional}$$

C : $\mu_m = (X^T\Sigma^{-1}X+\Sigma_0^{-1})^{-1}(X^T\Sigma^{-1}Y+\Sigma_0^{-1}\beta_0) \Rightarrow$ si $\Sigma^{-1} = \sigma^{2^{-1}} \Rightarrow \mu_m = (\Sigma_0^{-1}+X^TX/\sigma^2)^{-1}(\Sigma_0^{-1}\beta_0 + X^TY/\sigma^2)$

$\Lambda_m = (X^T\Sigma^{-1}X+\Sigma_0^{-1})^{-1} \rightarrow$ " $\Rightarrow \Lambda_m = (\Sigma_0^{-1}+X^TX/\sigma^2)^{-1}$

$$P(\beta|Y,X,\Sigma) \sim N(\mu_m, \Lambda_m)$$

Posterior para la varianza (γ) - Full conditional

$$\underbrace{\gamma \sim \text{gamma}(a,b) = \frac{b^a}{\Gamma(a)}\gamma^{a-1}\exp\{-\gamma b\}}, \quad \gamma \sim \text{gamma}(V_0/2, V_0\sigma_0^2/2)$$

$$\gamma = \frac{1}{\sigma^2}$$

$$\Rightarrow P(\gamma|\beta,Y,X) = P(Y|X,\beta,\gamma)\,Pr(\gamma|X,\beta) \,/\, Pr(\beta,Y,X)$$

$$= \frac{(2\pi)^{-m/2}|\Sigma|^{-m/2}\exp\left\{-\frac{1}{2}(Y-X\beta)^T\Sigma^{-1}(Y-X\beta)\right\}\frac{(V_0\sigma_0^2/2)^{V_0/2}}{\Gamma(V_0/2)}\gamma^{V_0/2-1}\exp\{-\gamma V_0\sigma_0^2/2\}}{Pr(\beta,Y,X)}$$

$$= (2\pi)^{-m/2}\gamma^{+m/2}\frac{(V_0\sigma_0^2/2)^{V_0/2}}{\Gamma(V_0/2)}\gamma^{\frac{V_0}{2}-1}\exp\left\{-\frac{\gamma}{2}\left[V_0\sigma_0^2/2 + SSR(\beta)\right]\right\} / Pr(\beta,Y,X)$$

$$P(\gamma|\beta,Y,X) = (2\pi)^{-m/2}(V_0\sigma_0^2/2)^{V_0/2}\gamma^{\frac{V_0+m}{2}-1}\exp\{-\gamma[V_0\sigma_0^2 + SSR(\beta)]/2\} / P_r(\beta,Y,X)$$

$\gamma(v_0/2)$

$$P(\gamma|\beta,y,X) = (2\pi)^{-n/2}\frac{(v_0\sigma_0^2/2)^{v_0/2}}{\Gamma(v_0/2)}\gamma^{\frac{v_0+n}{2}-1}\exp\left\{-\gamma[v_0\sigma_0^2+SSR(\beta)]/2\right\} / P_r(\beta,y,X)$$

Integrado ambos

$$P_r(\beta,y,X) = (2\pi)^{-n/2}\frac{(v_0\sigma_0^2/2)^{v_0/2}}{\Gamma(v_0/2)}\frac{\Gamma([v_0+n]/2)}{\{[v_0\sigma_0^2+SSR(\beta)]/2\}^{\frac{v_0+n}{2}}}\underbrace{\int_0^{+\infty}\frac{\{[v_0\sigma_0^2+SSR(\beta)]/2\}^{\frac{(v_0+n)}{2}}\gamma^{\frac{v_0+n}{2}}\exp\{-\gamma[v_0\sigma_0^2+SSR(\beta)]/2\}}{\Gamma([v_0+n]/2)}d\gamma}_{=1}$$

Entonces:

$$P_r(\gamma|\beta,y,X) = \frac{([v_0\sigma_0^2+SSR(\beta)]/2)^{\frac{v_0+n}{2}}}{\Gamma([v_0+n]/2)}\gamma^{\frac{v_0+n}{2}-1}\exp\{-\gamma[v_0\sigma_0^2+SSR(\beta)]/2\}$$

$$P(\gamma|\beta,y,X)\sim \text{gamma}([v_0+n]/2,\ [v_0\sigma_0^2+SSR(\beta)]/2)$$

Puedo hacer Gibbs-Sampling

1. $P(\beta|y,X,\Sigma)\sim N(\mu_n,\Lambda_n)$
2. $P(\gamma|\beta,y,X)\sim \text{gamma}([v_0+n]/2\ \ [v_0\sigma_0^2+SSR(\beta)]/2)$

Caso particular

Se utilizan los datos del libro

$\rightarrow \boxed{\beta_0=0,\ \Sigma_0 = g\sigma^2(X^TX)^{-1},\ \Sigma=\sigma^2 I}$

Veamos el posterior de β

$$P_r(\beta/y,X,\Sigma) = (2\pi)^{-p/2}|\Lambda_n|^{-1/2}\exp\left\{-\frac{1}{2}(\beta-\mu_n)^T\Lambda_n^{-1}(\beta-\mu_n)\right\}$$

Con: $\mu_n = (X^T\Sigma^{-1}X+\Sigma_0^{-1})^{-1}(X^T\Sigma^{-1}y+\cancel{\Sigma_0^{-1}\beta_0}^{\,0}) = \frac{g}{g+1}\sigma^2(X^TX)^{-1}\frac{1}{\sigma^2}X^Ty = \frac{g}{g+1}(X^TX)^{-1}X^Ty$

$\Lambda_n = (X^T\Sigma^{-1}X+\Sigma_0^{-1})^{-1} = \frac{g}{g+1}\sigma^2(X^TX)^{-1}$

Veamos el posterior de γ

$$P(\sigma^2|y,X) = \underbrace{P(y/X,\sigma^2)}_{1}\ \underbrace{P_r(\sigma^2/X)}_{2}\ /\ \underbrace{P_r(y,X)}_{3}$$

De 1 $P(y|X,\sigma^2) = \int^{+\infty} P(y|X,\beta,\sigma^2)P(\beta/X\sigma^2)\,d\beta$

Def 1) $$P(Y|X,\sigma^2) = \int_{\Theta}^{+\infty} \underbrace{P(Y|X,\beta,\sigma^2)\,P(\beta|X,\sigma^2)}_{\text{Conocido}}\, d\beta$$

$$P(Y|X,\sigma^2) = \int_{\Theta}^{+\infty} \frac{(2\pi)^{-n/2}\,|\Sigma|^{-M/2}\,|\Sigma_0|^{-1/2}\,|\Lambda_m|^{1/2}\exp\left\{-\frac{1}{2}H\right\}(2\pi)^{-p/2}\,|\Lambda_m|^{-1/2}\exp\left\{-\frac{1}{2}(\beta-\mu_m)^T\Lambda_m^{-1}(\beta-\mu_m)\right\}}{P_r(Y,X,\Sigma)}\, d\beta$$

$$P_r(Y|X,\sigma^2) = (2\pi)^{-M/2}\,|\Sigma|^{-n/2}\,|\Sigma_0|^{-1/2}\,|\Lambda_m|^{1/2}\exp\left\{-\frac{1}{2}H\right\} \Big/ P_r(Y,X,\Sigma)$$

$\beta_0 = 0$

$$P(Y|X,\sigma^2) = (2\pi)^{-M/2}(\sigma^2)^{-M/2}\left(\frac{1}{g\cancel{\sigma^2}}\right)^{1/2}\cancel{|X^TX|}^{+1/2}\left(\frac{g}{g+1}\right)^{1/2}\cancel{(\sigma^2)^{1/2}}\,\cancel{|X^TX|}^{-1/2}\exp\left\{-\frac{1}{2\sigma^2}\,Y^T\Sigma Y - \mu_m^T\Lambda_m^{-1}\mu_m\right\}$$

$$P(Y|X,\sigma^2) = (2\pi)^{-M/2}(\sigma^2)^{-M/2}(1+g)^{-1/2}\exp\left\{-\frac{1}{2}\,Y^T\Sigma Y - \cancel{\frac{g}{g+1}}Y^TX\,\cancel{(X^TX)^{-1}}\,\cancel{\frac{g+1}{g\sigma^2}}\,\cancel{(X^TX)}\,\frac{g}{g+1}(X^TX)^{-1}X^TY\right\}$$

$$P(Y|X,\sigma^2) = (2\pi)^{-M/2}(\sigma^2)^{-M/2}(1+g)^{-1/2}\exp\left\{-\frac{1}{2\sigma^2}\underbrace{Y^T\left(I - \frac{g}{g+1}X(X^TX)^{-1}X^T\right)Y}_{SSR_g}\right\}$$

Recordar:

$$\beta_{OLS} = (X^TX)^{-1}X^TY \Rightarrow (Y-X\beta)^T(Y-X\beta) = Y^TY - \beta^TX^TY - Y^TX\beta + \beta^TX^TX\beta$$

$$= Y^TY - 2Y^TX(X^TX)^{-1}X^TY + \underbrace{Y^TX\cancel{(X^TX)^{-1}}}_{\beta}\cancel{X^TX}\underbrace{(X^TX)^{-1}X^TY}_{\beta}$$

$$= Y^TY - Y^TX(X^TX)^{-1}X^TY$$

$$SSR(\beta_{OLS}) = Y^T(I - X(X^TX)^{-1}X^T)Y = Y^TMY$$

En este caso

$$SSR_g = Y^T\left(I - \frac{g}{g+1}X(X^TX)^{-1}X^T\right)Y$$

Continuando

Diferente del libro

$$P(Y|X,\sigma^2) = (2\pi)^{-M/2}(\sigma^2)^{-M/2}(1+g)^{-1/2}\exp\left\{-\frac{1}{2\sigma^2}SSR_g\right\}$$

Lo q' yo quiero (ojo: $\gamma = 1/\sigma^2$)

$$P(\gamma|Y,X) = P(Y|X,\gamma)\,P_r(\gamma)\,/\,P_r(Y,X)$$

$$P(\gamma|Y,X) = \frac{(2\pi)^{-M/2}\gamma^{+M/2}(1+g)^{-1/2}\exp\left\{-\frac{1}{2}\gamma\,SSR_g\right\}\dfrac{(V_0\sigma_0^2/2)^{V_0/2}}{\Gamma(V_0/2)}\gamma^{\frac{V_0}{2}-1}\exp\left\{-(V_0\sigma_0^2/2)\gamma\right\}}{P_r(Y,X)}$$

Integrando

$$P(Y,X)\underbrace{\int_0^{+\infty}P(\gamma|Y,X)\,d\gamma}_{=1} = (2\pi)^{-M/2}(1+g)^{-1/2}\frac{(V_0\sigma_0^2/2)^{V_0/2}}{\Gamma(V_0/2)}\int_0^{+\infty}\gamma^{\frac{V_0+M}{2}-1}\exp\left\{-\gamma(V_0\sigma_0^2+SSR_g)/2\right\}d\gamma$$

$$P(Y,X) = (2\pi)^{-M/2}(1+g)^{-1/2}\frac{(V_0\sigma_0^2/2)^{V_0/2}}{\Gamma(V_0/2)}\,\frac{\Gamma([V_0+M]/2)}{([V_0\sigma_0^2+SSR_g]/2)^{\frac{V_0+M}{2}}}\underbrace{\int_0^{+\infty}\frac{([V_0\sigma_0^2+SSR_g]/2)^{\frac{V_0+M}{2}}}{\Gamma([V_0+M]/2)}\gamma^{\frac{V_0+M}{2}-1}\exp\left\{-\gamma(V_0\sigma_0^2+SSR_g)/2\right\}d\gamma}_{=1}$$

$$P(y,x) = (2\pi)^{-m/2}(1+g)^{-1/2}\frac{(\nu_0\sigma_0^2/2)^{\nu_0/2}}{\Gamma(\nu_0/2)}\frac{\Gamma([\nu_0+m]/2)}{([\nu_0\sigma_0^2+SSR_g]/2)^{\frac{\nu_0+m}{2}}}\int_0^{\infty}\cancel{\frac{([\nu_0\sigma_0^2+SSR_g]/2)^{\frac{\nu_0+m}{2}}}{\Gamma([\nu_0+m]/2)}\gamma^{\frac{\nu_0+m}{2}-1}\exp\{-\gamma(\nu_0\sigma_0^2+SSR_g)/2\}\,d\gamma}$$

$$P(y,x) = (2\pi)^{-m/2}(1+g)^{-1/2}\frac{(\nu_0\sigma_0^2/2)^{\nu_0/2}}{\Gamma(\nu_0/2)}\frac{\Gamma([\nu_0+m]/2)}{([\nu_0\sigma_0^2+SSR_g]/2)^{\frac{\nu_0+m}{2}}}$$

Finalmente

$$P(\gamma|y,x) = \frac{\cancel{(2\pi)^{-m/2}}\gamma^{+m/2}\cancel{(1+g)^{-1/2}}\exp\{-\frac{1}{2}\gamma\, SSR_g\}\cancel{\frac{(\nu_0\sigma_0^2/2)^{\nu_0/2}}{\Gamma(\nu_0/2)}}\gamma^{\frac{\nu_0}{2}-1}\exp\{-(\nu_0\sigma_0^2/2)\gamma\}}{\cancel{(2\pi)^{-m/2}}\cancel{(1+g)^{-1/2}}\cancel{\frac{(\nu_0\sigma_0^2/2)^{\nu_0/2}}{\Gamma(\nu_0/2)}}\frac{\Gamma([\nu_0+m]/2)}{([\nu_0\sigma_0^2+SSR_g]/2)^{\frac{\nu_0+m}{2}}}}$$

$$\Rightarrow P(\gamma|y,x) = \frac{([\nu_0\sigma_0^2+SSR_g]/2)^{\frac{\nu_0+m}{2}}}{\Gamma([\nu_0+m]/2)}\gamma^{\underbrace{\frac{\nu_0+m}{2}-1}_{a}}\exp\{-\gamma\underbrace{[\nu_0\sigma_0^2+SSR_g]/2}_{b}\}$$

$$\Rightarrow P(\gamma/y,x) \sim \text{Gamma}([\nu_0+m]/2,\ [\nu_0\sigma_0^2+SSR_g]/2)$$

FIN