# MS 04/12/24

## MODELLO LINEARE BAYESIANO

Il 09/10/24 abbiamo visto che se

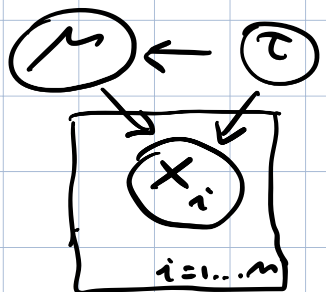
$$X_1, \dots X_n \overset{ciid}{\sim} N(\mu, \tau^{-1}) \qquad \tau = \frac{1}{\sigma^2} = \text{precisione}$$

$$(*) \begin{cases} \mu \mid \tau \sim N(\mu_0, (\tau \tau_0)^{-1}) \\ \tau \sim Gamma(a, \lambda) \end{cases}$$

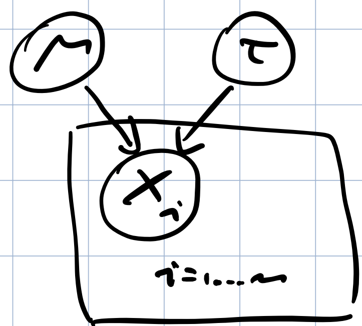
allora

$$\mu \mid \tau\, x_1, \dots x_n \sim N( \quad , \quad )$$
$$\tau \mid x_1, \dots x_n \sim Gamma( \quad , \quad )$$

cioè questa a priori, normal gamma (*), è coniugata al campionamento normale.




equivalente rappresentazione DAG, risolvibile esplicitamente




Non coniugato, risolvibile in JAGS (MCMC)

Con MCMC e le reti bayesiane abbiamo:

1) modularità (flessibilità)

2) capacità di risoluzione computazionale (MCMC)

Comunque, anche nel modello lineare abbiamo una a priori coniugata che generalizza (*)

$$\underline{Y} \sim \mathcal{N}(X\beta, \sigma^2 I)$$

$\cancel{\theta = (\beta, \sigma^2)}$

$\theta = (\beta, \tau = 1/\sigma^2)$

Vedremo ora che una a priori coniugata è

$$\beta / \tau \sim \mathcal{N}_p\left(\beta_0, \frac{V_0}{\tau}\right) \qquad V_0 \text{ nota}$$

$$\tau \sim \text{Gamma}\left(\frac{c_0}{2}, \frac{d_0}{2}\right)$$

Come al solito, $\beta_0, V_0, c_0, d_0$ sono iperparametri che esplicitano delle informazioni a priori.

Abbiamo visto che la verosimiglianza è

$$\mathcal{L}(\beta, \overset{\tau}{\cancel{\sigma^2}}; \underline{y}) \propto \tau^{n/2} \exp\left\{-\frac{\tau}{2} (\underset{\pm X\hat{\beta}}{y - X\beta})'(\underset{\pm X\hat{\beta}}{y - X\beta})\right\}$$

$$= (\text{un po' di algebra})$$

$$= \tau^{n/2} \exp\left\{-\frac{\tau}{2}\left[\overbrace{(y - X\hat{\beta})'}^{E'}\overset{E}{(y - X\hat{\beta})} + (\hat{\beta} - \beta)' X'X (\hat{\beta} - \beta)\right]\right\}$$

$\frac{E'E}{n-p}$

$$= \tau^{n/2} \exp\left\{-\frac{\tau}{2}\left[(n-p)s^2 + (\beta - \hat{\beta})' X'X (\beta - \hat{\beta})\right]\right\}$$

quindi la densità a posteriori è

$$\pi(\beta, \tau / y) \propto \pi(\beta, \tau) \times \mathcal{L}(\beta, \tau; y)$$

a priori gamma su $\tau$ — a priori normale su $\beta | \tau$

$$\propto \tau^{\frac{c_0}{2}-1} e^{-\frac{d_0}{2}\tau} \tau^{p/2} \exp\left\{-\frac{\tau}{2}(\beta-\beta_0)' V_0^{-1}(\beta-\beta_0)\right\}$$

$$\tau^{n/2} \exp\left\{-\frac{\tau}{2}\left[(n-p)s^2 + (\beta-\hat{\beta})' X'X(\beta-\hat{\beta})\right]\right\}$$

verosimiglianza

cominciamo a vedere una gamma per $\tau$

$$\propto \tau^{\frac{c_0}{2}+\frac{n}{2}+\frac{p}{2}-1} \exp\left\{-\frac{\tau}{2}\left[d_0 + (n-p)s^2 + (\beta-\beta_0)' V_0^{-1}(\beta-\beta_0) + (\beta-\hat{\beta})' X'X (\beta-\hat{\beta})\right]\right\}$$

B nel lemma

A nel lem

c'è una somma di due forme quadratiche che dobbiamo riscrivere come una sola per avere una normale a posteriori su $\beta$ (è il vecchio completamento dei quadrati)

LEMMA $x, a, b \in \mathbb{R}^k$, $\underset{k\times k}{A}$, $\underset{k\times k}{B}$ simmetriche tali che $(A+B)^{-1}$ esista, allora

$$(x-a)'A(x-a) + (x-b)'B(x-b)$$

$$= (x-c)'(A+B)(x-c) + (a-b)'A(A+B)^{-1}B(a-b)$$

dove $c = (A+B)^{-1}(Aa + Bb)$

( dimostrazione in calce*)

si dimostra che la a posteriori è di nuovo normal-gamma con i seguenti aggiornamenti bayesiani:

$$V_0 \longrightarrow V_* = (X'X + V_0^{-1})^{-1}$$

$$\beta_0 \longrightarrow \beta_* = V_* (X'y + V_0^{-1}\beta_0)$$

prior guess per $\beta$

$$c_0 \longrightarrow c_* = c_0 + n$$

$$d_0 \longrightarrow d_* = d_0 + (n-p)s^2 + (\hat\beta - \beta_0) X'X \, V_* V_0^{-1} (\hat\beta - \beta_0)$$

Se prendiamo il caso limite di distribuzione a priori impropria ( "informazione a priori nulla" )

$$c_0 = d_0 = 0$$

$$\underset{p\times p}{V_0^{-1}} = \underset{p\times p}{0}$$

otteniamo tuttavia una a posteriori propria

$$\tau / y \sim \mathrm{Gamma}\left(\frac{n}{2}, \frac{(n-p)s^2}{2}\right)$$

$$E(\tau) = \frac{n/2}{\frac{(n-p)s^2}{2}} = \frac{n}{E'E} = \frac{1}{\sigma}$$

MLE di $\beta$

$$\beta / \tau, y \sim N_p\left(\hat\beta, \tau^{-1}(X'X)^{-1}\right)$$

nota: $$E(\tau/y) = \frac{n/2}{\frac{(n-p)s^2}{2}} = \frac{n}{(n-p)s^2} = \frac{n}{E'E} = \frac{1}{\hat\sigma^2}$$ MLE di $\sigma^2$

* DIMOSTRAZIONE

$$(x-a)'A(x-a) + (x-b)'B(x-b) =$$

$$= x'Ax - \overset{-x'Aa}{\cancel{a'Ax}} - x'Aa + a'Aa$$

$$x'Bx \underbrace{- b'Bx}_{\substack{\text{uguale} \\ \text{al trasposto} \\ = x'Bb}} - x'Bb + b'Bb$$

$$= x'(A+B)x - 2x'\overbrace{\phantom{.}}^{(A+B)(A+B)^{-1}}(Aa+Bb) + a'Aa + b'Bb$$

$$= x'(A+B)x - 2x'(A+B)\underbrace{(A+B)^{-1}(Aa+Bb)}_{= c} \pm c'(A+B)c + a'Aa + b'Bb$$

$$= (x-c)'(A+B)(x-c) - c'(A+B)c + a'Aa + b'Bb$$

Notiamo ora che

$$c'(A+B)c = (\underset{\pm Ab}{Aa} + Bb)'(A+B)^{-1}\cancel{(A+B)(A+B)^{-1}}(Aa + \underset{\pm Ba}{Bb})$$

$$= (A(a-b)+(A+B)b)'(A+B)^{-1}((A+B)a - B(a-b))$$

$$= (a-\cancel{b})'A\cancel{(A+B)^{-1}(A+B)}a + b'(\cancel{A}+B)\cancel{(A+B)^{-1}(A+B)}a$$

$$- (a-b)'A(A+B)^{-1}B(a-b) - b'\cancel{(A+B)(A+B)^{-1}}B(\cancel{a}-b)$$

$$= a'Aa - (a-b)'A(A+B)^{-1}B(a-b) + b'Bb$$

e infine

$$(x-a)'A(x-a) + (x-b)'B(x-b) = (x-c)(A+B)(x-c)$$

$$- \cancel{a'Aa} + (a-b)'A(A+B)^{-1}B(a-b) - \cancel{b'Bb} + \cancel{a'Aa} + \cancel{b'Bb}$$

C.V.D.