(Zobecněné) lineární smíšené modely 0

→ částečné opakování z Pokročilých regresních modelů

1. Lineární smíšený model 1

– některá písmenka jiná než v PRM 2

+ $E b_i = \mu$ (≡ podmnožina pevných efektů)

$$Y_i = \underset{n_i \times}{X_i} \beta + \underset{n_i \times q}{Z_i} b_i + \varepsilon_i, \quad i = 1, \dots, N$$

plné matice

$$b_i \overset{iid}{\sim} N_q(\mu, D)$$

$$\varepsilon_i \sim N_{n_i}(0, \Sigma_i)$$

$b_1, \dots, b_N, \varepsilon_1, \dots, \varepsilon_N$ nezávislé

„obecné" parametry modelu:

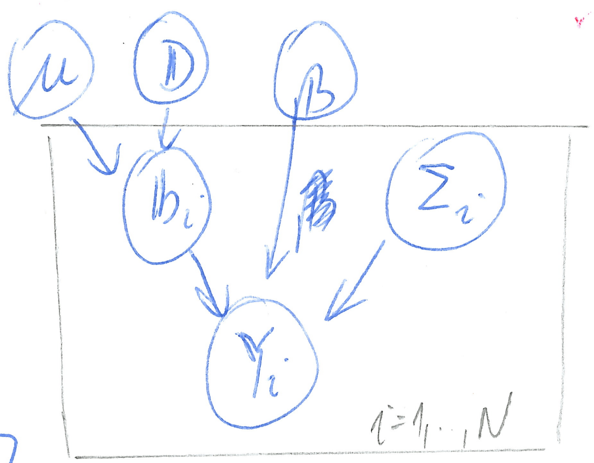
$$\psi = (\beta, \mu, D, \mathrm{parm}(\Sigma_i))$$

obvykle $\Sigma_i = \sigma^2 I_{n_i}$ $\forall i$

LMM hierarchicky 3

$$Y_i \mid b_i; \beta, \Sigma_i \overset{nezáv.}{\sim} N_{n_i}(X_i\beta + Z_i b_i, \Sigma_i)$$

$$b_i \mid \mu, D \overset{iid}{\sim} N_q(\mu, D)$$



2. Příklady LMM

4

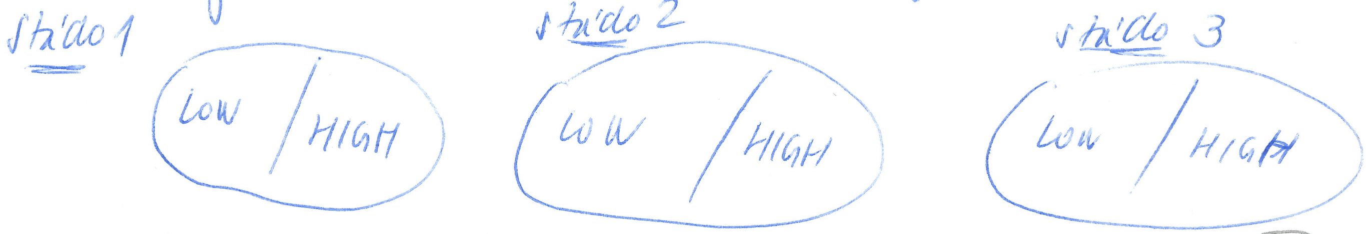
Příklad 1 „ANOVA" LMM 5

$$Y_{ij} = b_i + \varepsilon_{ij}, \qquad \mathbb{E}\, b_i = \mu, \quad \mathrm{var}\, b_i = d^2$$

$$i = 1, \dots, N \qquad \mathbb{E}\, \varepsilon_{ij} = 0, \quad \mathrm{var}\, \varepsilon_{ij} = \sigma^2$$

$$j = 1, \dots, n_i \qquad \Sigma_i = \sigma^2 I_{n_i}$$

Příklad 2 Y_{ij} = PCV přírůstek u j-té krávy i-tého stáda 6



obrázky $\boxed{7}$ – $\boxed{18}$

(a) dávka má stejný účinek ve všech stádech 9

$$Y_{ij} = b_i + \varepsilon_{ij}, \quad \text{pokud LOW}$$

$$= b_i + \beta + \varepsilon_{ij}, \quad \text{pokud HIGH}$$

$$Y_{ij} = b_i + \beta x_{ij} + \varepsilon_{ij} \qquad x_{ij} = \begin{cases} 0 & \text{LOW} \\ 1 & \text{HIGH} \end{cases}$$

$$\mathbb{E}\, b_i = \mu, \quad \mathrm{var}\, b_i = d^2$$

$$\mathbb{E}\, \varepsilon_{ij} = 0, \quad \mathrm{var}\, \varepsilon_{ij} = \sigma^2$$

$$Y_i = X_i \beta + Z_i b_i + \varepsilon_i, \qquad \mathrm{var}\, \varepsilon_i = \underbrace{\sigma^2 I_{n_i}}_{\Sigma_i}$$

$$X_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \hline 1 \\ \vdots \\ 1 \end{pmatrix} \qquad Z_i = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$ 11

(b) dávka funguje různě v různých stádiích 12

→ efekt dávky (β) závisí na i

$$Y_{ij} = b_{i,1} + b_{i2} z_{ij} + \varepsilon_{ij}$$

$z_{ij} \equiv x_{ij}$ v předch. případě

$$E\begin{pmatrix} b_{i,1} \\ b_{i,2} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \equiv \beta \text{ z předchozího}$$

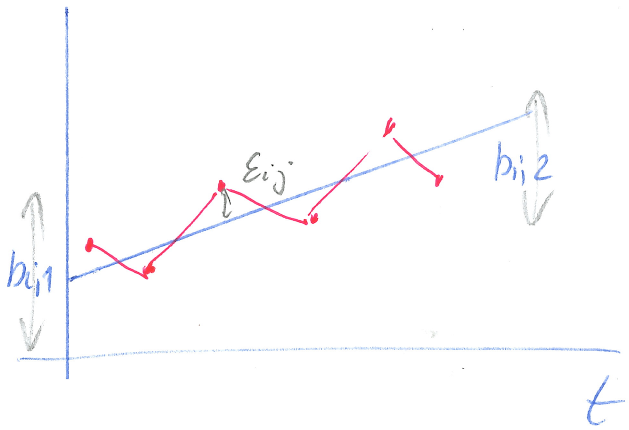
$$\operatorname{var}\begin{pmatrix} b_{i,1} \\ b_{i,2} \end{pmatrix} = D \quad , \operatorname{var} \varepsilon_{ij} = \sigma^2$$

13, 14

Příklad 3: Longitudinální data 15

Obrázek 16

MODEL? lineární trend pro každého pac. 17



18

2 skupiny – liší se STŘEDNÍ HODN. směrnice 19

β = rozdíl středních směrnic

20

$$X_i = \begin{pmatrix} t_{i1} \\ \vdots \\ t_{in_i} \end{pmatrix}, \text{ pokud } x_{ij} = x_i = 1$$

$$= \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \text{ pokud } x_{ij} = x_i = 0$$

$$Z_i = \begin{pmatrix} 1 & t_{i1} \\ \vdots & \vdots \\ 1 & t_{in_i} \end{pmatrix} \qquad \Sigma_i = \sigma^2 I_{n_i}$$

---

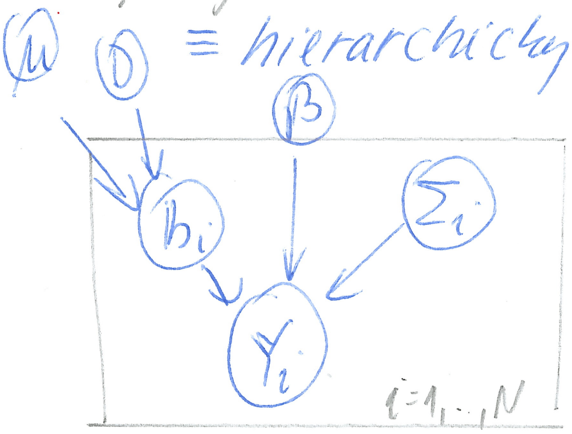
3. Lineární smíšený model bayesovsky

22

opakuje se ...

23

(μ) (D) ≡ hierarchický model



$$Y_i \mid b_i;\ \beta, \Sigma_i \overset{\text{nezáv.}}{\sim} N_{n_i}(X_i\beta + Z_i b_i, \Sigma_i)$$

$$b_i \mid \mu, D \overset{iid}{\sim} N_q(\mu, D)$$

obvykle $\underline{\Sigma_i = \sigma^2 I_{n_i}}$ 24

„skutečné" parametry:

$$\psi \equiv (\beta, \sigma^2, \mu, D)$$

$$L_F(\psi) = p(y \mid \psi) = \int p(y \mid b, \psi)\, p(b \mid \psi)\, db = \qquad 25$$

$$= \prod_{i=1}^{N} \int \underbrace{p(y_i \mid b_i, \psi)}_{N(\dots)}\ \underbrace{p(b_i \mid \psi)}_{N(\dots)}\, db_i \overset{\text{integrál umíme}}{=}$$

$$= \prod_{i=1}^{N} \varphi(y_i \mid X_i\beta + Z_i\mu,\ V_i)$$

$$V_i = Z_i D Z_i^T + \Sigma_i$$

- teď by stačilo specifikovat apriorní $p(\psi)$ pro ψ a můžeme odvozovat aposteriorní rozdělení.

$$p(\psi \mid y) \propto \underbrace{p(y \mid \psi)}_{L_F(\psi)}\ p(\psi)$$

Pro Bayes. počítání:

$(b_1, \dots, b_N) \equiv \xi$ (skrytá data)

$\rightarrow$ další „parametry" modelu

věrohodnost pro Bayes. počítání:

$$L(\xi, \psi) = p(y / \xi, \psi) = \prod_{i=1}^{N} p(y_i / b_i, \psi)$$

$$= \prod_{i=1}^{N} \varphi(y_i / X_i\beta + Z_i b_i, \Sigma_i)$$

apriorní rozdíl:

$$p(\xi, \psi) = p(\xi / \psi)\, p(\psi)$$

$$= \prod_{i=1}^{N} p(b_i / \psi) \quad p(\psi)$$

$$= \prod_{i=1}^{N} \varphi(b_i / \mu, D)\; p(\psi)$$

$$\rightarrow \quad p(\xi, \psi / y) \propto \underbrace{p(y / \xi, \psi)}_{N} \; \underbrace{p(\xi / \psi)}_{N} \; \underbrace{p(\psi)}_{\text{nutno nějak zvolit}}$$

hierarchický princip

$\Rightarrow$

$$p(\psi / y) = \int p(\xi, \psi / y)\, d\xi$$

(z minulé strany)

4. Zobecněný lineární smíšený model

Generalized linear mixed model - GLMM

[27]

Příklad: opět myši [28]

$$Y_{ij} = \begin{cases} 0 & j\text{-tý zárodek } i\text{-té myši BEZ vady} \\ 1 & \text{s vadou} \end{cases}$$

Primární cíl: inference o $\pi = \mathbb{E}\,Y_{ij} = P(Y_{ij} = 1)$

obrázek [29]

MOŽNÝ MODEL 1: $Y_{ij} \overset{iid}{\sim} A(\pi)$ [30]

asi ne...

MOŽNÝ MODEL 2 (hierarchický)

pro myš i: $P(Y_{ij} = 1) = \pi_i$

tj: $Y_{ij} \mid \pi_i \sim A(\pi_i)$

$\pi_i \overset{iid}{\sim}$ ~~rh~~ rozdělení na $(0,1)$ (s parametry ψ)

$0 < \pi_i < 1$ omezení jsou obtěžující...

$\rightarrow$ přeparametrizuj

$$b_i := \operatorname{logit} \pi_i = \log\left(\frac{\pi_i}{1-\pi_i}\right)$$

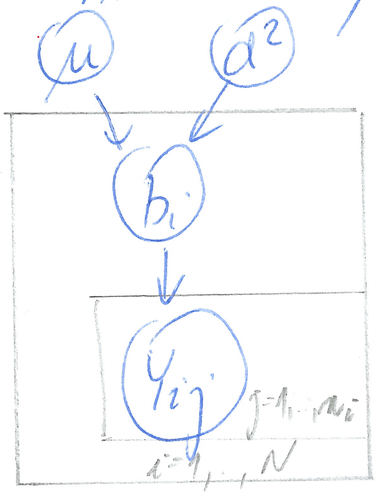
$$\text{tj: } \pi_i = \frac{e^{b_i}}{1+e^{b_i}}$$

$\pi_i \overset{iid}{\sim}$ rozdel. na $(0,1)$ $\Longleftrightarrow$ $b_i \overset{iid}{\sim}$ rozdel. na $\mathbb{R}$

$\rightarrow$ logistická regrese s náhodnými efekty

Logistická regrese s náhodným interceptem

Hierarchický model:



$$y_{ij} \mid b_i \sim A(\pi_i) \qquad \pi_i = \frac{e^{b_i}}{1+e^{b_i}}$$

$$b_i \overset{iid}{\sim} \underline{\mathcal{N}}(\underbrace{\mu,\sigma^2}_{\psi})$$

např. ψ: „skutečné" parametry

$$p(\mathbb{y}\mid\mathbb{B}) = \prod_{i=1}^{N} p(\mathbb{y}_i\mid\mathbb{b}_i) = \prod_{i=1}^{N}\prod_{j=1}^{n_i} \underbrace{p(y_{ij}\mid b_i)}_{\pi_i^{y_{ij}}(1-\pi_i)^{1-y_{ij}}}$$

$\mathbb{B} = (\mathbb{b}_1,\dots,\mathbb{b}_N)$

$$p(\mathbb{B}\mid\underbrace{\mu,\sigma^2}_{\psi}) = \prod_{i=1}^{N} \underbrace{p(b_i\mid\mu,\sigma^2)}_{\sim\mathcal{N}(\mu,\sigma^2)}$$

$$L_F(\psi) = p(\mathbb{y}\mid\psi) = \int p(\mathbb{y}\mid\mathbb{B},\cancel{\psi})\, p(\mathbb{B}\mid\psi)\, d\mathbb{B} =$$

$$= \prod_{i=1}^{N} \int p(\mathbb{y}_i\mid b_i)\, p(b_i\mid\psi)\, db_i =$$

$$= \prod_{i=1}^{N} \int \prod_{j=1}^{n_i} p(y_{ij}\mid b_i)\, p(b_i\mid\psi)\, db_i =$$

$$= \prod_{i=1}^{N} \underbrace{\int_{\mathbb{R}}}_{\text{good luck}} \frac{e^{b_i(\sum_j y_{ij})}}{(1+e^{b_i})^{n_i}}\, \varphi(b_i\mid\mu,\sigma^2)\, db_i$$

Bayesovský: přidej $p(\psi)$, pracuj s

$$p(\mathbb{B},\psi\mid\mathbb{y}) \propto p(\mathbb{y}\mid\mathbb{B},\cancel{\psi})\, p(\mathbb{B}\mid\psi)\, p(\psi)$$

$$= \underbrace{\prod_i \Big(\prod_j p(y_{ij}\mid b_i)\Big)\, p(b_i\mid\psi)\, p(\psi)}_{\text{žádný integrál}}$$

i když pozor, stále neznáme normující konstantu!

GLMM obecně

$Y_{ij} | \mathbb{b}_i \sim$ exponenciální třída , $i=1,\dots,N$

$Y_{ij} | \mathbb{b}_i$ nezávislé pro $j=1,\dots,n_i$ $\quad j=1,\dots,n_i$

$\mathbb{E}(Y_{ij} | \mathbb{b}_i) = h(\underline{x}_{ij}^T \beta + \underline{z}_{ij}^T \mathbb{b}_i) =: \mu(\mathbb{b}_i)$

↑ inverzní link → vektory konstant

$\operatorname{var}(Y_{ij} | \mathbb{b}_i) = \phi V(\mu(\mathbb{b}_i))$, ϕ = disperzní param.

často $\phi = 1$

Pozn: v LMM $\phi = \sigma^2$ = (reziduální) rozptyl pro ε_{ij}

$\mathbb{b}_i \overset{iid}{\sim} N(\mu, D)$

---

primární (skutečné) parametry $\psi = (\beta, \underbrace{\phi}_{\text{často není}}, \mu, D)$

skryté data: $\xi \equiv \mathbb{B} = (\mathbb{b}_1, \dots, \mathbb{b}_N)$

Pouze „drobné" rozdíly oproti LMM:

$p(y | \mathbb{B}, \psi) = \prod_{i=1}^{N} p(y_i | \mathbb{b}_i, \psi) =$

$= \prod_i \prod_j \underbrace{p(y_{ij} | \mathbb{b}_i, \psi)}_{\text{hustota exponenciální třídy, ozn. } g(\bullet | \mu(\mathbb{b}_i), \phi)} \overset{LMM}{=} \prod_i \prod_j \varphi(y_{ij} | \dots)$

$\sim N(\underbrace{x_{ij}^T \beta + z_{ij}^T \mathbb{b}_i}_{\mu(\mathbb{b}_i)}, \sigma^2)$

↑ ϕ

věrohodnost pro frekventistické výpočty

$$L_F(\psi) = p(y/\psi) = \int p(y/B,\psi)\, p(B/\psi)\, dB =$$

$$= \prod_i \int p(y_i/b_i,\psi)\, p(b_i/\psi)\, db_i =$$

$$= \prod_i \underbrace{\int_{R^q} \prod_j g(y_{ij}/\mu(b_i),\phi)\, \varphi(b_i/\mu,D)\, db_i}_{\text{typicky nelze analyticky} \ldots}$$

Bayesovský: přidej $p(\psi)$, pracuj s

$$p(B,\psi/y) \propto p(y/B,\psi)\, p(B/\psi)\, p(\psi)$$

$$= \underbrace{\prod_i \left(\prod_j p(y_{ij}/b_i,\psi)\right) p(b_i/\psi)\, p(\psi)}$$

žádný integrál,
ale neznáme normující konstantu,
takže ani přímá simulace
není možná!

→ MCMC (později)

5. Apriorní rozdělení pro LMM / GLMM

$$\psi = (\underbrace{\beta}_{\text{pevné efekty}}, \underbrace{\mu}_{\text{stř. hodnoty nah. efektů}}, \tau, Q)$$

$\tau = \phi^{-1}$ = inverzní disperz. param
$= \sigma^{-2}$ v LMM

$Q = D^{-1}$ = inverzní var. matice nah. efektů

Připomínka:

$Y_{ij} \mid b_i \sim$ exp. třída$(\mu(b_i), \phi)$

$\mu(b_i) = h(x_{ij}^T \beta + z_{ij}^T b_i)$

$b_i \overset{iid}{\sim} N(\mu, D)$

$\xi = B = (b_1, \dots, b_N)$

$$p(\xi, \psi) = \underbrace{p(\xi \mid \psi)}_{\text{viz model}} p(\psi)$$

$= \prod_i p(b_i \mid \psi)$
$\sim N(\mu, D)$

NYNÍ: Jak volit $p(\psi)$?

$p(\psi)$: případně též hierarchicky

$$p(\psi) = \int p(\psi \mid \zeta)\, p(\zeta)\, d\zeta$$

ζ: hyperparametry (= další param. modelu)

$$p(\xi, \psi, \zeta) = p(\xi \mid \psi) \underbrace{p(\psi \mid \zeta)\, p(\zeta)}_{\text{Jak tohle?}}$$

přeskoč

$\psi = (\beta, \mu, \tau, Q)$ $\qquad \tau = \Phi^{-1}$, $Q = D^{-1}$ 36

VOLBA $p(\psi)$, resp. $p(\psi|\xi)$:

obvyklý postup: apriorní nezávislost pro sady „různých" param.

$$\to p(\psi) = p(\beta, \mu, \tau, Q) = p(\beta)\, p(\mu)\, p(\tau)\, p(Q)$$

resp.

$$p(\psi|\xi) = p(\beta|\xi^{(1)}) \cdot p(\mu|\xi^{(2)}) \cdot p(\tau|\xi^{(3)})\, p(Q|\xi^{(4)})$$

- $\beta, \mu \equiv$ střední hodnota. 37

VOLBA 1

$$\left.\begin{array}{l} p(\beta) \propto 1 \\ p(\mu) \propto 1 \end{array}\right\} \text{POZOR, nevlastní!}$$

VOLBA 2

$$p(\beta) \sim N(\beta_0, \Sigma_0^\beta)$$

$$p(\mu) \sim N(\underbrace{\mu_0, \Sigma_0^\mu}_{\text{pevné/náhodné hyperparam.}})$$

pokud pevné:

např. $\left.\begin{array}{l} \beta_0 = 0 \\ \mu_0 = 0 \end{array}\right\}$ kromě abs. členů modelu

$\Sigma_0^\beta, \Sigma_0^\mu$: diag. matice, „velká" diagonála

- $\tau = \phi^{-1}$ ($\overline{\overline{\text{např}}}$ σ^{-2} v LMM)
  = inverzní „rozptyl"

VOLBA: $p(\tau) \sim Ga(\underbrace{c_\tau, d_\tau}_{\text{pevné / náhodné hyperparam.}})$

mj. zajišťuje $P(\tau > 0 \mid Y) = 1$

- Jak volit c_τ, d_τ?

  - $c_\tau \equiv$ stupně volnosti
    $c_\tau \in (0, 1]$: slabě informativní rozdel.
  - $E\tau = \frac{c_\tau}{d_\tau}$, $\operatorname{var}\tau = \frac{c_\tau}{d_\tau^2}$
    $d_\tau \gg 0$: „přesnost"
    $d_\tau \to 0$: slabě informativní
  - empirické studie:
    aposteriorní rozdelení bývá citlivé na volbu d_τ
    $\to$ d_τ často náhodný hyperparam.
    $d_\tau \sim Ga(\underbrace{\cdot, \cdot}_{\text{již pevné}})$
  - časté volby (pokud c_τ i d_τ pevné):
    $\tau \sim Ga(1, 0{,}005)$
    $\sim Ga(0{,}001, 0{,}001)$

39

- $Q = \mathbb{D}^{-1}$, $\mathbb{D} = \operatorname{var} b_i$
  - potřeba vícerozměrné rozdělení, které s.j. generuje pozitivně definitní matici

  $\rightarrow$ Wishartovo rozdělení

40

Wishartovo rozdělení

$\equiv$ zobecnění χ^2

$$Q \sim W_p(\nu, \Xi) \iff Q = \sum_{i=1}^{\nu} Z_i Z_i^T,$$

- $Z_i \overset{iid}{\sim} \mathcal{N}_p(0, \Xi)$
- $\nu > p-1$: stupně volnosti

$\rightarrow$ zřejmě : $P(Q > 0) = 1$

„SDC" : odvoď hustotu Wishartova rozdělení 41

$\rightarrow$ lze uvažovat též neceločíselné stupně voln.

- $\mathbb{E}\, Q = \nu\, \Xi$
- $W_1(\nu, 1) \equiv Ga\left(\frac{\nu}{2}, \frac{1}{2}\right) \equiv \chi^2_\nu$
- $W_1(\nu, \Xi) \equiv Ga\left(\frac{\nu}{2}, \frac{\Xi^{-1}}{2}\right)$

  (Ξ skalár)

ν, Ξ : hyperparametry $\left\langle \begin{array}{l} \text{pevné} \\ \text{náhodné} \end{array} \right.$ 42

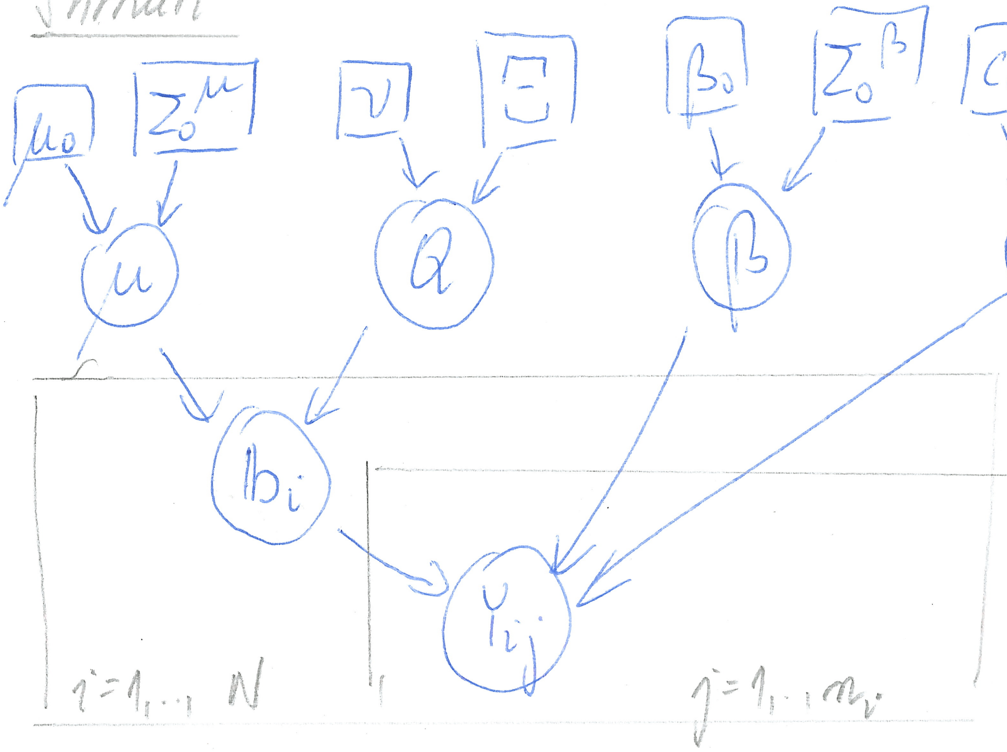
Obvykle voleno : $\nu \in [p-1, p]$

$\Xi^{-1} = \operatorname{diag}(\underline{\lambda_1}, \ldots, \underline{\lambda_p})$ — malá ($\rightarrow 0$)

případně : $\lambda_j \sim Ga(\cdot, \cdot)$

} motivace jako u τ

Shrnutí



$: \psi$

$: \xi$

$: Y$

6. Aposteriorní rozdělení 44

viz hierarchická struktura, resp. DAG 45

$$p(\xi, \psi \mid y) \propto p(y \mid \xi, \psi)\, p(\xi \mid \psi)\, p(\psi)$$

$$= \prod_i \prod_j p(y_{ij} \mid b_i, \beta, \tau) \cdot \prod_i p(b_i \mid \mu, Q) \cdot p(\mu)\, p(Q)\, p(\beta)\, p(\tau)$$

případně

$$p(\xi, \psi, \varsigma \mid y) \propto p(y \mid \xi, \psi)\, p(\xi \mid \psi)\, p(\psi \mid \varsigma)\, p(\varsigma)$$

$$= \prod_i \prod_j p(y_{ij} \mid b_i, \beta, \tau) \prod_i p(b_i \mid \mu, Q)\, p(\mu \mid \varsigma^{(1)})\, p(Q \mid \varsigma^{(2)})\, p(\beta \mid \varsigma^{(3)})\, p(\tau \mid \varsigma^{(4)})$$

!! neznáme normující konstantu

Díky hierarchické struktuře též platí 46

$$p(\psi|y) = \int p(\psi, \xi | y)\, d\xi$$ ← tady bez náhod. hyperparam.

$$= L_F(\psi)\, p(\psi), \text{ kde}$$

$$L_F(\psi) = p(y|\psi) = \int p(y|\underset{B}{\xi}, \psi)\, p(\underset{B}{\xi}|\psi)\, d\underset{B}{\xi}$$

7. Inference založená na aposteriorním rozdělení

47

povídání 48-50

→ bude: MCMC

Markov chain Monte Carlo

→ generuj Markovův řetězec

$\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(n)}$ takový, že

jeho stacionární rozdělení

je $p(\theta|y)$

(resp. rozdělení, z něhož bych měl rád náhodný výběr)