Přednáška 4 – Bayesovský odhad parametrů modelu

Obecný tvar

$f(y_t|\psi_t,\Theta)$

ψ_t – naměřená data

Θ – neznámé parametry:

- $\{\theta, r\}$ – regresní,
- $\Theta_{y_t|\psi_t}$ – diskrétní, θ – logistický

Bayesovské odhadování parametrů

Parametry – neznámá náhodná velicina

hp $f(\Theta|\text{průběžně měřená data})$ – bodové odhady

Označení:

$d_t = \{y_t, u_t\}$ – data v čase t

$$\underbrace{d(t) = \{d_0, d_1, \dots, d_t\}}_{\text{stará data } + \ d_t}$$

Průběžný vývoj hp:

$$f(\Theta|d(0)) \underbrace{\Rightarrow}_{d_1} f(\Theta|d(1)) \underbrace{\Rightarrow}_{d_2} f(\Theta|d(2)) \Rightarrow \dots \underbrace{\Rightarrow}_{d_t} f(\Theta|d(t))$$

Obecně:

$$\text{apriorní hp} \quad f(\Theta|d(t-1)) \underbrace{\Rightarrow}_{d_t} \text{aposteriorní hp} \quad f(\Theta|d(t))$$

Bayesovo pravidlo

$$\underbrace{f(\Theta|d(t))}_{\text{aposteriorní}} \propto \underbrace{f(y_t|\psi_t,\Theta)}_{\text{model}} \underbrace{f(\Theta|d(t-1))}_{\text{apriorní}}$$

Odvození Bayesova pravidla pro náhodné veličiny A, B, C

$$\underbrace{f(A,B|C)}_{\text{sdružená hp}} = \begin{cases} f(A|B,C)f(B|C) = \text{podmíněná} \times \text{marginální} \\ \text{nebo} \\ f(B|A,C)f(A|C) \end{cases}$$

Obě strany:

$$f(A|B,C)f(B|C) = f(B|A,C)f(A|C)$$

Podmíněná hp:

$$f(B|A,C) = \frac{f(A|B,C)f(B|C)}{f(A|C)} \qquad \text{– \underline{Bayesovo pravidlo}}$$

Význam:

$$\underline{\text{apriorní}}\ f(B|C) \Rightarrow \underline{\text{aposteriorní}}\ f(B|A,C)$$

Jmenovatel $f(A|C)$ je nezávislý na B (normalizační konstanta)

$$f(B|A,C) \underbrace{\propto}_{\underline{\text{úměrné}}} f(A|B,C)f(B|C)$$

Použití pro odhad parametrů Θ

- $A = y_t$
- $B = \Theta$
- $C = \{d(t-1), u_t\}$
- $\{A, C\} = \{d_t, d(t-1)\} = d(t)$

$f(\Theta|d(t-1)) \Rightarrow f(\Theta|d(t))$

$$\underbrace{f(\Theta|d(t))}_{\text{aposteriorní}} = \frac{f(y_t|\psi_t,\Theta) f(\Theta|d(t-1))}{f(y_t|d(t-1))}$$

$$\propto \underbrace{f(y_t|\psi_t,\Theta)}_{\text{model}} \underbrace{f(\Theta|d(t-1))}_{\text{apriorní}}$$

Předpoklad – přirozené podmínky řízení:

$\Theta \nwarrow$, $u_t \swarrow$ stejná data $d(t-1)$

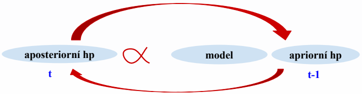
$$f(\Theta|\cancel{u_t}, d(t-1)) = f(\Theta|d(t-1))$$
$$f(u_t|\cancel{\Theta}, d(t-1)) = f(u_t|d(t-1))$$

Použití řídícího vstupu v regresním vektoru:

$$\underbrace{f(\Theta|d(t))}_{\text{aposteriorní}} \propto \underbrace{f(y_t|\psi_t,\Theta)}_{\text{model}} \cancel{f(u_t|d(t-1),\Theta)} \underbrace{f(\Theta|d(t-1))}_{\text{apriorní}}$$

Reproducukovatelnost apriorní a aposteriorní hp

Bayesovo pravidlo – rekurze



Pro normální regresní model:

- Volba reprodukující se apriorní hp (stejný tvar v čase t)
- Konjugovaná apriorní rozdělení (pro regresní a diskrétní modely)

Konjugovaná hp – inverzní Gauss-Wischartovo (GiW) rozdělení

$$f(\Theta|d(t-1)) \propto r^{-0.5\kappa_{t-1}} \exp\{-\frac{1}{2r}[-1 \;\; \theta'] V_{t-1} \begin{bmatrix} -1 \\ \theta \end{bmatrix}\}$$

V_{t-1} – informační matice, κ_{t-1} – počítadlo

Bayesovo pravidlo pro odhad parametrů regresního modelu

$$\underbrace{f(\Theta|d(t))}_{GiW} \propto \underbrace{f(y_t|\psi_t,\Theta)}_{\text{normální}} \underbrace{f(\Theta|d(t-1))}_{GiW}$$

Update statistik: $V_t = V_{t-1} + D_t, \quad \kappa_t = \kappa_{t-1} + 1$

Proč je to tak?

$$\underline{\text{Model:}}\ f(y_t|\psi_t,\Theta) = \frac{1}{\sqrt{2\pi}} r^{-0.5} \exp\{-\frac{1}{2r}(y_t - \psi_t'\theta)^2\}$$

$$\underline{\text{Úprava:}}\ y_t - \psi_t'\theta = -1[-1\ \ \theta']\begin{bmatrix} y_t \\ \psi_t \end{bmatrix} = -1\underbrace{[y_t\ \ \psi_t']}_{\text{data}} \underbrace{\begin{bmatrix} -1 \\ \theta \end{bmatrix}}_{\text{parametry}}$$

$$(y_t - \psi_t'\theta)^2 = [-1\ \ \theta'] \underbrace{\begin{bmatrix} y_t \\ \psi_t \end{bmatrix}[y_t\ \ \psi_t']}_{D_t\ \text{– datová matice}} \begin{bmatrix} -1 \\ \theta \end{bmatrix}$$

$$\underline{\text{Model:}}\ f(y_t|\psi_t,\Theta) = \frac{1}{\sqrt{2\pi}} r^{-0.5} \exp\{-\frac{1}{2r}[-1\ \ \theta']D_t \begin{bmatrix} -1 \\ \theta \end{bmatrix}\}$$

Model a GiW – do Bayesova pravidla:

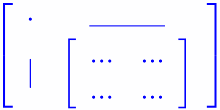
$V_t = V_{t-1} + D_t$ Algebraický přepočet statistik

$\kappa_t = \kappa_{t-1} + 1$ aposteriorní hp GiW

Po přepočtu statistik:

Rozklad informační matice:

$$V_t = \begin{bmatrix} V_y & V'_{y\psi} \\ V_{y\psi} & V_\psi \end{bmatrix}$$

$=$ 

Bodový odhad regresních koeficientů a rozptylu:

$$\hat{\theta}_t = V_\psi^{-1} V_{y\psi}$$

$$\hat{r}_t = \frac{V_y - V'_{y\psi} V_\psi^{-1} V_{y\psi}}{\kappa_t}$$

Algoritmus odhadu:

1. Pro čas $t = 0$ nastavíme počáteční statistiky V_0, κ_0
2. Pro čas $t = 1, 2, \ldots$
   1. Měříme data $d_t = \{y_t, u_t\}$
   2. Datová matice D_t
   3. Update statistik V_t, κ_t
   4. Jdeme na krok 2.1
3. Rozklad informační matice V_t
4. Výpočet bodových odhadů parametrů $\hat{\theta}_t$, $\hat{r}_t$

Využití pro odhad/predikci výstupu:

$$\hat{y}_t = \underbrace{\psi'_t \hat{\theta}_{t-1}}_{\text{střední hodnota}}$$

Program – odhad regresního modelu 1.řádu

```
nd=size(y,2); % počet dat
nps=5 % rozměr statistiky V
% počáteční statistiky
V=zeros(nps,nps); k=0;
for t=2:nd
    % rozšířený regresní vektor
    Ps=[y(t) u(t) y(t-1) u(t-1) 1];
    V=V+Ps'*Ps% update statistik
    k=k+1
end
% rozklad informační matice
Vy=V(1,1) % část Vy
Vyps=V(2:end,1) % část Vyps
Vps=V(2:end,2:end) % část Vps
% bodové odhady
theta_odhad=inv(Vps)*Vyps
r_odhad=(Vy-Vyps'*inv(Vps)*Vyps)/k
```