Dodatek 1. Bayesowskie modele VAR

MODELOWANIE POLSKIEJ GOSPODARKI z R

Formuła Bayesa

Formuła Bayesa:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Dlaczego:

$$P(A, B) = P(A|B)P(B) = P(B|A)P(A)$$

Formuła Bayesa

Dla modelu ekonomicznego:

$$P(\Theta|Y) = \frac{P(Y|\Theta)P(\Theta)}{P(Y)}$$

$$P(\Theta|Y) \propto P(Y|\Theta)P(\Theta),$$

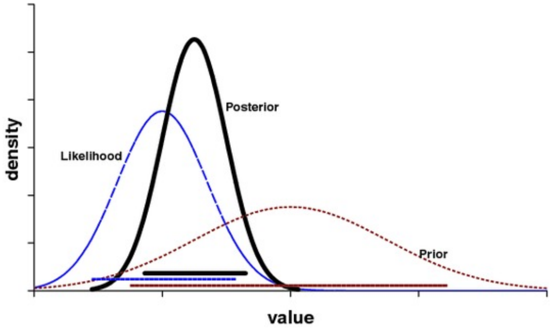
gdzie:

$P(\Theta)$: rozkład a-priori dla parametru Θ

$P(Y|\Theta)$: funkcja wiarygodności danych Y

$P(\Theta|Y)$: rozkład a-posteriori dla parametru Θ

Ilustracja



Rozkłady a-posteriori

- Istnieją klasy modeli oraz rozkładów a-priori dla których rozkład a-posteriori można uzyskać w sposób analityczny. W takim przypadku mówimy o sprzężonym rozkładzie a-priori (conjugate prior)
- W większości przypadków rozkład a-posteriori możemy uzyskać jedynie wykorzystując metody numeryczne (Gibbs sampling, Metropolis Hastings algorithm oraz inne)

Sprzężony rozkład a-priori - model liniowy

Rozważmy model liniowy:

$$y_t = \beta x_t + \epsilon_t, \ \ \epsilon_t \sim N(0, 1/h)$$

Rozkład a-priori typu Normal-Gamma:

$$\beta|h \sim N(\underline{\beta}, h^{-1}\underline{V}); \quad h \sim \Gamma(\underline{s}^{-2}, \underline{v}), \quad \beta, h \sim NG(\underline{\beta}, \underline{V}, \underline{s}^{-2}, \underline{v})$$

Oszacowania MNW:

$$\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2}, \quad v = T - 1, \quad s^2 = \frac{\sum (y_i - \hat{\beta} x_i)^2}{v}$$

Rozkład a-prosteriori:

$$\beta, h|y \sim NG(\overline{\beta}, \overline{V}, \overline{s}^{-2}, \overline{v}); \quad \beta|y \sim t(\overline{\beta}, \overline{s}^2 \overline{V}, \overline{v}),$$

gdzie:

$$\overline{V} = \frac{1}{\underline{V}^{-1} + \sum x_i^2}; \quad \overline{\beta} = \overline{V}(\underline{V}^{-1}\underline{\beta} + \hat{\beta}\sum x_i^2); \quad \overline{v} = \underline{v} + N$$

$$\overline{vs}^2 = \underline{vs}^2 + \frac{(\hat{\beta} - \underline{\beta})^2}{\underline{V} + (\sum x_i^2)^{-1}}$$

Gibbs sampling

- Interesuje nas łączny rozkład $p(\beta, \sigma^2|y)$
- Znamy postać analityczną $p(\beta|\sigma^2, y)$ oraz $p(\sigma^2|\beta, y)$
- Ustalamy wartości początkowe $\beta_{(0)}$ i $\sigma^2_{(0)}$
- Losujemy rekurencyjnie dla $i = 1, 2, \ldots, N + N_0$: $\beta_{(i)}$ z rozkładu $p(\beta|\sigma^2_{(i-1)}, y)$ oraz $\sigma^2_{(i)}$ z rozkładu $p(\sigma^2|\beta_{(i)}, y)$
- W rezultacie mamy N losowań z rozkładu łącznego $p(\beta, \sigma^2|y)$, tj. $(\beta_{(i)}, \sigma^2_{(i)})$
- Usuwamy N_0 początkowych losowań (burn-in sample)
- Możemy estymować $g(\beta, \sigma^2|y) \approx \frac{1}{N} \sum g(\beta_{(i)}, \sigma^2_{(i)})$

Przykład zastosowania Gibbs sampling - rozkład t-Studenta

- Chcemy obliczyć wartość oczekiwaną i wariancję zmiennej losowej $X \sim t(v)$, ale nie mamy generatora liczb z rozkładu t-Studenta (teoretyczne wartości to 0 i $v/(v-2)$)
- Wiemy, że jeżeli $\sigma^2 \sim Inv\chi^2(v)$ oraz $X|\sigma^2 \sim N(0, \sigma^2)$ to $X \sim t(v)$
- Dla $i = 1, 2, \ldots, N$ losujemy $\sigma^2_{(i)}$ z rozkładu $Inv\chi^2(v)$ oraz $X_{(i)}$ z rozkładu $N(0, \sigma^2_{(i)})$.
- Szacujemy wartość oczekiwaną jako $E(X) = \frac{1}{N}\sum X_{(i)}$, zaś wariancję jako $Var(X) = \frac{1}{N}\sum X^2_{(i)}$.

Przykład zastosowania Gibbs sampling - rozkład t-Studenta

```
v    = 10;               #  liczba stopni swobody
N    = 1000;             #  liczba ciągnięć
sig2 = NULL;
X    = NULL;

for (i in 1:N){
   Y    =  rchisq(1, v);      # losujemy z rozkładu chi2(v)
   sig2 =  c(sig2,v/Y);
   X    =  c(X , rnorm(1,0,sqrt(sig2[i])));
   }

> sum(X)/N                    # teoretyczna wartość 0
0.03229785

> sum(X^2)/N                  # teoretyczna wartość v/(v-2)
1.287174
```

Algorytm Metropolis-Hastings

- Chcemy obliczyć momenty rozkładu a-posteriori $\theta|y$ (który nie jest standardowy)
- Startujemy w $\theta_{(0)}$ dla którego liczymy $p(\theta_{(0)}|y)$
- Kolejne ciągnięcia z rozkładu a-posteriori dla $i = 1, 2, \ldots, N + N_0$ liczymy rekurencyjnie
- Ustalamy $\theta^* = \theta_{(i)} + \delta\epsilon$, gdzie δ to stała skalująca, zaś $\epsilon \sim N(0, \Sigma)$.
- Losujemy liczbę u z rozkładu jednostajnego $U(0, 1)$
- Jeżeli $\frac{p(\theta^*)}{p(\theta_{(i)})} > u$ ustalamy $\theta_{(i+1)} = \theta^*$. W przeciwnym przypadku $\theta_{(i+1)} = \theta_{(i)}$
- Usuwamy N_0 początkowych obserwacji (burn-in sample)
- Możemy estymować $g(\beta, \sigma^2|y) \approx \frac{1}{N} \sum g(\beta_{(i)}, \sigma^2_{(i)})$

Model BVAR z rozkładem apriori Littermana

- Zapis modelu VAR:

$$y_t = \begin{bmatrix} A_0 & A_1 & \dots & A_P \end{bmatrix} \begin{bmatrix} 1 & y'_{t-1} & \dots & y'_{t-P} \end{bmatrix}' + \epsilon_t = Ax_t + \epsilon_t$$

- Rozkład a-priori dla elementów macierzy A_p:

$$\underline{a}_{ij}^{(p)} = \begin{cases} 1 & \text{dla } i = j, p = 1 \\ 0 & \text{dla pozostałych} \end{cases}$$

$$\underline{\sigma}_{ij}^{(p)} = \lambda_1 \times p^{-\lambda_3} \times \sigma_i / \sigma_j$$

$$\underline{\sigma}_{i}^{(0)} = \lambda_4 \times \sigma_i$$

- Oznaczenia:

  λ_1: Overall tightness
  λ_3: Lag decay
  λ_4: constant tightness
  σ_i: Wariancja reszt modelu AR(P) dla zmiennej y_{it}

BVAR - Litterman prior

- Zapis dla *i*-tego równania (każde równanie jest szacowane oddzielnie)

$$y_{it} = a_i' x_t + \epsilon_{it}$$

- Oszacowanie MNW:

$$\hat{a}_i = (X'X)^{-1}X'y, \ \sigma^2_{\epsilon_i} = 1/T \sum e_t^2$$

- Rozkład a-priori:

$$a_i \sim N(\underline{a}_i, \underline{\Sigma}_i)$$

- Rozkład a-posteriori:

$$a_i | y \sim N(\overline{a}_i, \overline{\Sigma}_i),$$

gdzie:

$$\begin{aligned} \overline{\Sigma}_i =& (\underline{\Sigma}_i^{-1} + \sigma_{\epsilon_i}^{-2} X'X)^{-1} \\ \overline{a}_i =& \overline{\Sigma}_i (\underline{\Sigma}_i^{-1} \underline{a}_i + \sigma_{\epsilon_i}^{-2} X'y) \end{aligned}$$

BVAR - Litterman prior dla $N = 2$ i $P = 2$

Rozkład a-priori dla dwuwymiarowego modelu VAR(2):

$$a_1 = \begin{bmatrix} a_1^{(0)} \\ a_{11}^{(1)} \\ a_{12}^{(1)} \\ a_{11}^{(2)} \\ a_{12}^{(2)} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} (\lambda_4 \sigma_1)^2 & 0 & 0 & 0 & 0 \\ 0 & (\lambda_1)^2 & 0 & 0 & 0 \\ 0 & 0 & (\lambda_1 \times \sigma_1/\sigma_2)^2 & 0 & 0 \\ 0 & 0 & 0 & (\lambda_1 \times 2^{-\lambda_3})^2 & 0 \\ 0 & 0 & 0 & 0 & (\lambda_1 \times 2^{-\lambda_3} \times \sigma_1/\sigma_2)^2 \end{bmatrix} \right)$$

$$a_2 = \begin{bmatrix} a_2^{(0)} \\ a_{21}^{(1)} \\ a_{22}^{(1)} \\ a_{21}^{(2)} \\ a_{22}^{(2)} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} (\lambda_4 \sigma_2)^2 & 0 & 0 & 0 & 0 \\ 0 & (\lambda_1 \times \sigma_2/\sigma_1)^2 & 0 & 0 & 0 \\ 0 & 0 & (\lambda_1)^2 & 0 & 0 \\ 0 & 0 & 0 & (\lambda_1 \times 2^{-\lambda_3} \times \sigma_2/\sigma_1)^2 & 0 \\ 0 & 0 & 0 & 0 & (\lambda_1 \times 2^{-\lambda_3})^2 \end{bmatrix} \right)$$

BVAR - Sims-Zha prior

- W przypadku rozkładu a-priori Littermana parametry każdego równania są estymowane oddzielnie i nie bierze się pod uwagę korelacji składnika losowego Σ
- Rozkład a-priori zaproponowany w opracowaniu Sims-Zha (1998) uwzględnia występowanie zależności między składnikami losowymi
- Autorzy proponują rozkład a-priori typu Normal-Wishart, który w połączeniu z funkcją wiarygodności dla wielowymiarowego rozkładu losowego daje rozkład a-posteriori typu Normal-Wishart

BVAR - Sims-Zha prior

Zapiszmy model VAR:

$$y_t = Ax_t + \epsilon_t, \epsilon_t \sim N(0, \Sigma)$$

Jako:

$$y_t = \begin{bmatrix} x_t' & 0 & \dots & 0 \\ 0 & x_t' & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x_t' \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_N \end{bmatrix} + \epsilon_t = Z_t a + \epsilon_t, \epsilon_t \sim N(0, \Sigma)$$

Oszacowania MNW dla parametru a:

$$\hat{a} = (\sum_{t=1}^{T} Z_t' Z_t)^{-1} (\sum_{t=1}^{T} Z_t' y_t)$$

BVAR - Sims-Zha prior

$$y_t = Z_t a + \epsilon_t, \epsilon_t \sim N(0, \Sigma)$$

Rozkład a-priori:

$$\Sigma \sim InvW(\underline{\Sigma}, \underline{v}) \text{ oraz } a|\Sigma \sim N(\underline{a}, \underline{V}),$$

$$\underline{\Sigma} = diag((\sigma_1/\lambda_0)^2, (\sigma_2/\lambda_0)^2, \ldots, (\sigma_N/\lambda_0)^2)$$
$$\underline{v} = N + 2$$
$$\underline{a}_{ij}^{(p)} = 1 \text{ dla } i = j, p = 1 (0 \text{ dla pozostałych })$$
$$\underline{V} = \Omega \otimes \Sigma$$
$$\Omega_{ij}^{(p)} = \left(\frac{\lambda_0 \lambda_1}{\sigma_i p^{\lambda_3}}\right)^2 \text{ oraz } \Omega_i^{(0)} = \lambda_0 \lambda_4$$

σ_i : odchylenie std. reszt modelu AR(P) dla y_{it}

λ_0 : overall tightness

λ_1 : tightness around AR coefficients

λ_4 : tightness around intercept

λ_3 : lag decay

BVAR - Sims-Zha prior

Dodatkowo, Sims-Zha zaproponowali dodanie do macierzy obserwacji zmiennych zero-jedynkowych postaci (dla dwuwymiarowego modelu VAR(2)):
„Sum of coefficients" dummies:

$$\begin{bmatrix}\mu_5\bar{y}_1 & 0\end{bmatrix}' \text{ oraz } \begin{bmatrix}0 & \mu_5\bar{y}_1 & 0 & \mu_5\bar{y}_1 & 0\end{bmatrix}'$$

$$\begin{bmatrix}0 & \mu_5\bar{y}_2\end{bmatrix}' \text{ oraz } \begin{bmatrix}0 & 0 & \mu_5\bar{y}_2 & 0 & \mu_5\bar{y}_2\end{bmatrix}'$$

„Cointegration" dummy:

$$\begin{bmatrix}\mu_6\bar{y}_1 & \mu_6\bar{y}_2\end{bmatrix} \text{ oraz } \begin{bmatrix}\mu_6 & \mu_6\bar{y}_1 & \mu_6\bar{y}_2 & \mu_6\bar{y}_1 & \mu_6\bar{y}_2\end{bmatrix}'$$

BVAR - Sims-Zha prior

Model z dummies: $\tilde{y}_t = \tilde{Z}_t a + \tilde{\epsilon}_t, \tilde{\epsilon}_t \sim N(0, \Sigma)$

Rozkład a-priori: $\Sigma \sim InvW(\underline{\Sigma}, \underline{v})$ oraz $a|\Sigma \sim N(\underline{a}, \underline{V})$

Oszacowanie MNW: $\hat{a} = (\sum_{t=1}^{\tilde{T}} \tilde{Z}_t' \tilde{Z}_t)^{-1} (\sum_{t=1}^{\tilde{T}} \tilde{Z}_t' \tilde{y}_t)$

Rozkład a-posteriori: $\Sigma|y \sim InvW(\overline{\Sigma}, \overline{v})$ oraz $a|\Sigma, y \sim N(\overline{a}, \overline{V})$

Zależności:

$$\overline{\Sigma} = \underline{\Sigma} + \sum_{t=1}^{\tilde{T}} (\tilde{y}_t - \tilde{Z}_t \hat{a})(\tilde{y}_t - \tilde{Z}_t \hat{a})'$$

$$\overline{v} = \underline{v} + T$$

$$\overline{a} = \overline{V}(\underline{V}^{-1} \underline{a} + \sum_{t=1}^{\tilde{T}} \tilde{Z}_t \Sigma^{-1} \tilde{y}_t)$$

$$\overline{V} = (\underline{V}^{-1} + \sum_{t=1}^{\tilde{T}} \tilde{Z}_t \Sigma^{-1} \tilde{Z}_t)^{-1}$$