Métodos computacionais para inferência estatística - modelos dinâmicos

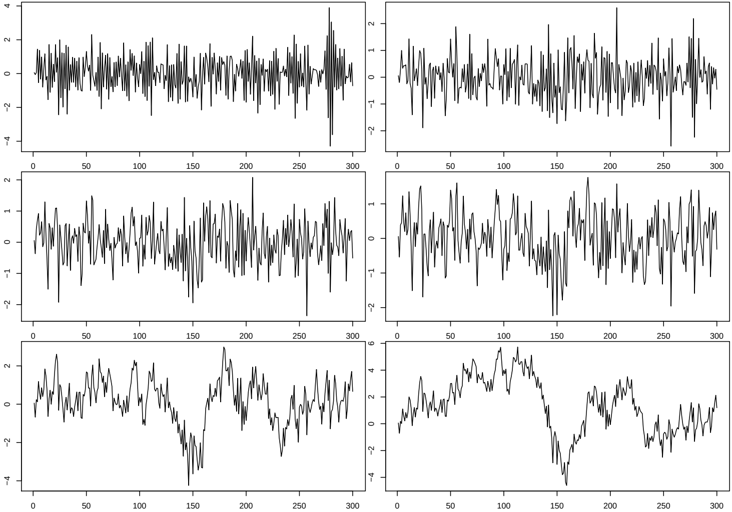
Elias Teixeira Krainski

LEG: Laboratório de Estatística e Geoinformação
Universidade Federal do Paraná

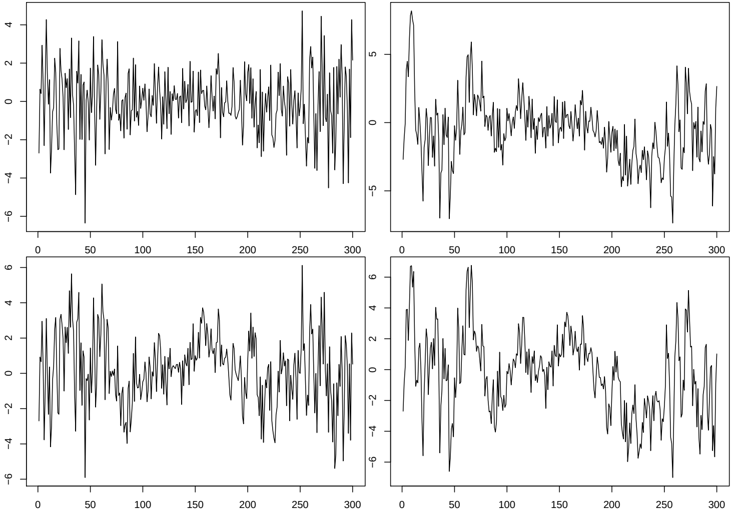
30 de julho de 2012

Sumário

Laboratório de Estatística e Geoinformação



Laboratório de Estatística
e Geoinformação


4
2
0
−2
−4
−6
0
50
100
150
200
250
300
5
0
−5
6
4
2
0
−2
−4
−6

O modelo dinâmico mais simples

1. $y_t = \theta_t + v_t$, com $v_t \sim N(0, V_t)$
2. $\theta_t = \phi\theta_{t-1} + w_t$, com $w_t \sim N(0, W_t)$

Laboratório de Estatística e Geoinformação

O modelo dinâmico mais simples

1. $y_t = \theta_t + v_t$, com $v_t \sim N(0, V_t)$
2. $\theta_t = \phi\theta_{t-1} + w_t$, com $w_t \sim N(0, W_t)$

- Com $V_t = 0$ e $0 \leq \phi \leq 1$ temos um modelo AR1
- Com $\phi = 1$ temos um passeio aleatório para θ
- Com $\phi = 1$ e $\frac{W}{V} \approx 0$, y_t será parecido com y_{t-1}.

Laboratório de Estatística e Geoinformação

O modelo dinâmico mais simples na forma geral

- Modelo Dinâmico Linear
  1. Observações

$$y_t = F_t\theta_t + v_t, \quad v_t \sim N(0, V_t)$$

  2. Estados

$$\theta_t = G_t\theta_{t-1} + w_t, \quad w_t \sim N(0, W_t)$$

O modelo dinâmico mais simples na forma geral

- Modelo Dinâmico Linear
  1. Observações
  $$y_t = F_t\theta_t + v_t, \quad v_t \sim N(0, V_t)$$
  2. Estados
  $$\theta_t = G_t\theta_{t-1} + w_t, \quad w_t \sim N(0, W_t)$$
- O modelo simples
  1. Observações: $y_t = \theta_t + v_t$, com $v_t \sim N(0, V_t)$
  2. Estados: $\theta_t = \phi\theta_{t-1} + w_t$, com $w_t \sim N(0, W_t)$

Laboratório de Estatística e Geoinformação

O modelo dinâmico mais simples na forma geral

- Modelo Dinâmico Linear
  1. Observações
  $$y_t = F_t\theta_t + v_t, \quad v_t \sim N(0, V_t)$$
  2. Estados
  $$\theta_t = G_t\theta_{t-1} + w_t, \quad w_t \sim N(0, W_t)$$
- O modelo simples
  1. Observações: $y_t = \theta_t + v_t$, com $v_t \sim N(0, V_t)$
  2. Estados: $\theta_t = \phi\theta_{t-1} + w_t$, com $w_t \sim N(0, W_t)$

  - y_t: dados observáveis
  - $F_t = 1$, escalar fixado
  - θ_t: intercepto variando no tempo
  - v_t: erro de observação no tempo t
  - w_t: erros do estado latente
  - $G_t = \phi$, escalar controlando evolução de θ

Laboratório de Estatística e Geoinformação

Modelo de crescimento

1. Observações

$$y_t = \alpha_t + v_t, \quad v_t \sim N(0, V)$$

2. Estados
   - $\alpha_t = \alpha_{t-1} + \beta_{t-1} + w1_t, \ w1_t \sim N(0, \sigma_1^2)$
   - $\beta_t = \beta_{t-1} + w2_t, \ w2_t \sim N(0, \sigma_2^2)$

Laboratório de Estatística e Geoinformação

Modelo de crescimento

1. Observações

$$y_t = \alpha_t + v_t, \quad v_t \sim N(0, V)$$

2. Estados
   - $\alpha_t = \alpha_{t-1} + \beta_{t-1} + w1_t, \; w1_t \sim N(0, \sigma_1^2)$
   - $\beta_t = \beta_{t-1} + w2_t, \; w2_t \sim N(0, \sigma_2^2)$

- Na formulação geral

$$F = [1 \; 0], \quad \theta_t = \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \; W = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}.$$

- Lembrando
  1. Observações

  $$y_t = F_t \theta_t + v_t$$

  2. Estados

  $$\theta_t = G_t \theta_{t-1} + w_t$$

Laboratório de Estatística e Geoinformação

Regressão dinâmica

1. Observações

$$y_t = F_t\theta_t + v_t$$

2. Estados

$$\theta_t = G_t\theta_{t-1} + w_t$$

No contexto de regressão dinâmica, nomeamos:

- F_t pode ser uma matriz de covariáveis
- θ_t vetor de coeficientes de regressão
- v_t vetor de erros das observações, $v_t \sim N(0, V_t)$
- w_t vetor de erros dos estados, $w_t \sim N(0, W_t)$
- G_t matriz de evolução dos estados

Regressão dinâmica

1. Observações

$$y_t = F_t\theta_t + v_t$$

2. Estados

$$\theta_t = G_t\theta_{t-1} + w_t$$

No contexto de regressão dinâmica, nomeamos:

- F_t pode ser uma matriz de covariáveis
- θ_t vetor de coeficientes de regressão
- v_t vetor de erros das observações, $v_t \sim N(0, V_t)$
- w_t vetor de erros dos estados, $w_t \sim N(0, W_t)$
- G_t matriz de evolução dos estados

Com $G = I$ e $W_t = 0 \rightarrow$ regressão linear usual.

Laboratório de Estatística e Geoinformação

Resumindo

- Modelo dinâmico: **Duas equações**
  1. Observações: Modelo assumido para os dados, condicional a (ou função de) um estado latente não observável
  2. Estados: Modelo assumido para o estado latente

Resumindo

- Modelo dinâmico: **Duas equações**
  1. Observações: Modelo assumido para os dados, condicional a (ou função de) um estado latente não observável
  2. Estados: Modelo assumido para o estado latente

1. Observações

$$y_t = F_t\theta_t + v_t$$

2. Estados

$$\theta_t = G_t\theta_{t-1} + w_t$$

Laboratório de Estatística e Geoinformação

Resumindo

- Modelo dinâmico: **Duas equações**
  1. Observações: Modelo assumido para os dados, condicional a (ou função de) um estado latente não observável
  2. Estados: Modelo assumido para o estado latente

  1. Observações
$$y_t = F_t\theta_t + v_t$$
  2. Estados
$$\theta_t = G_t\theta_{t-1} + w_t$$

- v_t e w_t são erros Normais e independentes (DLM)
- θ_t não é observável (latente)
- θ_t vetor de estados (modelo de espaço de estados)
- F_t e G_t define o tipo de modelo

Laboratório de Estatística e Geoinformação

Suposições

- Independência condicional das variáveis observáveis

$$[y_t|y_{0:t}, \theta_{0:t}] = [y_t|\theta_t]$$

- Estrutura Markoviana dos estados

$$[\theta_t|\theta_{0:t}, y_{0:t}] = [\theta_t|\theta_{t-1}]$$

- No caso linear Gaussiano, temos

$$\begin{aligned}[y_t|\theta_t] =& N(F_t\theta_t, V_t)\\ [\theta_t|\theta_{-t}] =& N(G_t\theta_{t-1}, W_t).\end{aligned}$$

Verossimilhança: Produto das condicionais

- F_t conhecido
- $\psi_t = \{V_t, W_t, G_t\} = \psi$ parâmetros de variância desconhecidos
- Distribuição conjunta

$$[y_{1:T};\psi] = [y_1;\psi][y_2|y_1;\psi]...[y_T|y_1,...,y_{T-1};\psi]$$

- DLM: Assumimos $[y_t|...]$ gaussiano
- médias e variâncias condicionais de $[y_t|...]$ (Filtro de Kalman)

Laboratório de Estatística e Geoinformação

Verossimilhança: Modelo simples

1. $y_t = \theta_t + v_t \quad \rightarrow \quad y_t \sim N(\theta_t, V)$
2. $\theta_t = \phi\theta_{t-1} + w_t \quad \rightarrow \quad \theta_t \sim N(\phi\theta_{t-1}, W)$

Laboratório de Estatística e Geoinformação

Verossimilhança: Modelo simples

1. $y_t = \theta_t + v_t \quad \rightarrow \quad y_t \sim N(\theta_t, V)$
2. $\theta_t = \phi\theta_{t-1} + w_t \quad \rightarrow \quad \theta_t \sim N(\phi\theta_{t-1}, W)$

Filtro de Kalman (Predição):

- considerando ($\theta_0 \sim N(\theta_0^0, P_0^0)$)
- para t=1,2,...,T, calcular:
    1. Predição (θ_t): $[\theta_t|y_{1:t-1}]$ (condicional ao passado)
        - $E(\theta_t|y_{1:t-1}) = \phi\theta_{t-1}^{t-1} = \theta_t^{t-1}$
        - $V(\theta_t|y_{1:t-1}) = \phi^2 P_{t-1}^{t-1} + W = P_t^{t-1}$

Laboratório de Estatística e Geoinformação

Verossimilhança: Modelo simples

1. $y_t = \theta_t + v_t \quad \rightarrow \quad y_t \sim N(\theta_t, V)$
2. $\theta_t = \phi\theta_{t-1} + w_t \quad \rightarrow \quad \theta_t \sim N(\phi\theta_{t-1}, W)$

Filtro de Kalman (Predição):

- considerando ($\theta_0 \sim N(\theta_0^0, P_0^0)$)
- para t=1,2,...,T, calcular:
  1. Predição (θ_t): $[\theta_t|y_{1:t-1}]$ (condicional ao passado)
     - $E(\theta_t|y_{1:t-1}) = \phi\theta_{t-1}^{t-1} = \theta_t^{t-1}$
     - $V(\theta_t|y_{1:t-1}) = \phi^2 P_{t-1}^{t-1} + W = P_t^{t-1}$
  2. Predição (y_t): $[y_t|y_{1:t-1}]$ (condicional ao passado)
     - $E(y_t|y_{1:t-1}) = \theta_t^{t-1}$, $e_t = y_t - \theta_t^{t-1}$ (erro de previsão)
     - $V(y_t|y_{1:t-1}) = P_t^{t-1} + V = Q_t$

Laboratório de Estatística e Geoinformação

Verossimilhança: Modelo simples

1. $y_t = \theta_t + v_t \quad \rightarrow \quad y_t \sim N(\theta_t, V)$
2. $\theta_t = \phi\theta_{t-1} + w_t \quad \rightarrow \quad \theta_t \sim N(\phi\theta_{t-1}, W)$

Filtro de Kalman (Predição):

- considerando ($\theta_0 \sim N(\theta_0^0, P_0^0)$)
- para t=1,2,...,T, calcular:
  1. Predição (θ_t): $[\theta_t|y_{1:t-1}]$ (condicional ao passado)
     - $E(\theta_t|y_{1:t-1}) = \phi\theta_{t-1}^{t-1} = \theta_t^{t-1}$
     - $V(\theta_t|y_{1:t-1}) = \phi^2 P_{t-1}^{t-1} + W = P_t^{t-1}$
  2. Predição (y_t): $[y_t|y_{1:t-1}]$ (condicional ao passado)
     - $E(y_t|y_{1:t-1}) = \theta_t^{t-1}$, $e_t = y_t - \theta_t^{t-1}$ (erro de previsão)
     - $V(y_t|y_{1:t-1}) = P_t^{t-1} + V = Q_t$

log-verossimilhança : $l(\psi|y_{1:t}) = -\frac{n}{2}\log(2\pi Q_t) - \frac{1}{2Q_t}\sum_{t=1}^{n} e_t^2$

Laboratório de Estatística e Geoinformação

Verossimilhança: Caso geral

1. $y_t = F_t\theta_t + v_t \quad \rightarrow \quad y_t \sim N(\theta_t, V)$
2. $\theta_t = G\theta_{t-1} + w_t \quad \rightarrow \quad \theta_t \sim N(G\theta_{t-1}, W)$

Filtro de Kalman (Predição):

- considerando $(\theta_0 \sim N(\theta_0^0, P_0^0)\ )$
- para t=1,2,...,T, calcular:
  1. Predição (θ_t): $[\theta_t|y_{1:t-1}]$ (condicional ao passado)
     - $E(\theta_t|y_{1:t-1}) = G\theta_{t-1}^{t-1} = \theta_t^{t-1}$
     - $V(\theta_t|y_{1:t-1}) = GP_{t-1}^{t-1}G' + W = P_t^{t-1}$
  2. Predição (y_t): $[y_t|y_{1:t-1}]$ (condicional ao passado)
     - $E(y_t|y_{1:t-1}) = F_t\theta_t^{t-1}$, $e_t = y_t - F_t\theta_t^{t-1}$ (erro de previsão)
     - $V(y_t|y_{1:t-1}) = F_tP_t^{t-1}F_t' + V = Q_t$

log-verossimilhança

$$l(\psi|y_{1:t}) = -\frac{1}{2}\sum_{t=1}^{n}\log|Q_t| - \frac{1}{2}\sum_{t=1}^{n}e_t'Q_t^{-1}e_t. \qquad (1)$$

Inferência sobre θ

- Consideremos $\theta_s | y_{1:t}$
  - $s > t$ predição
  - $s = t$ filtragem
  - $s < t$ suavização: cálculo de $[\theta_{1:t} | y_{1:t}]$
- Considerar (condicionar) $\hat{\psi}$

Laboratório de Estatística e Geoinformação

Filtragem: Modelo simples

- considerando ($\theta_0 \sim N(\theta_0^0, P_0^0)$)
- para t=1,2,...,T, calcular:
  1. Predição (θ_t): $[\theta_t | y_{1:t-1}]$ (condicional ao passado)
     - $E(\theta_t | y_{1:t-1}) = \phi\theta_{t-1}^{t-1} = \theta_t^{t-1}$
     - $V(\theta_t | y_{1:t-1}) = \phi^2 P_{t-1}^{t-1} + W = P_t^{t-1}$
  2. Predição (y_t): $[y_t | y_{1:t-1}]$ (condicional ao passado)
     - $E(y_t | y_{1:t-1}) = \theta_t^{t-1}$, $e_t = y_t - \theta_t^{t-1}$ (erro de previsão)
     - $V(y_t | y_{1:t-1}) = P_t^{t-1} + V = Q_t$

Laboratório de Estatística e Geoinformação

Filtragem: Modelo simples

- considerando ($\theta_0 \sim N(\theta_0^0, P_0^0)$)
- para t=1,2,...,T, calcular:
  1. Predição (θ_t): $[\theta_t | y_{1:t-1}]$ (condicional ao passado)
     - $E(\theta_t | y_{1:t-1}) = \phi\theta_{t-1}^{t-1} = \theta_t^{t-1}$
     - $V(\theta_t | y_{1:t-1}) = \phi^2 P_{t-1}^{t-1} + W = P_t^{t-1}$
  2. Predição (y_t): $[y_t | y_{1:t-1}]$ (condicional ao passado)
     - $E(y_t | y_{1:t-1}) = \theta_t^{t-1}$, $e_t = y_t - \theta_t^{t-1}$ (erro de previsão)
     - $V(y_t | y_{1:t-1}) = P_t^{t-1} + V = Q_t$
  3. Filtragem (θ_t): $[\theta_t | y_{1:t}]$ (condicional ao passado e presente)
     - $K_t = P_t^{t-1} / Q_t$
     - $E(\theta_t | y_{1:t}) = \theta_t^{t-1} + e_t = \theta_t^t$
     - $V(\theta_t | y_{1:t}) = P_t^{t-1} - P_t^{t-1} = P_t^t$

  Se $V = 0$, então: $Q_t = P_t^{t-1}$ e $\theta_t^t = y_t$

Laboratório de Estatística e Geoinformação

Filtro de Kalman: Caso geral

- considerando ($\theta_0 \sim N(\theta_0^0, P_0^0)$)
- para t=1,2,...,T, calcular:
  1. Predição (θ_t): $[\theta_t | y_{1:t-1}]$ (condicional ao passado)
     - $E(\theta_t | y_{1:t-1}) = G\theta_{t-1}^{t-1} = \theta_t^{t-1}$
     - $V(\theta_t | y_{1:t-1}) = GP_{t-1}^{t-1}G^{'} + W = P_t^{t-1}$
  2. Predição (y_t): $[y_t | y_{1:t-1}]$ (condicional ao passado)
     - $E(y_t | y_{1:t-1}) = F_t\theta_t^{t-1}$, $e_t = y_t - F_t\theta_t^{t-1}$ (erro de previsão)
     - $V(y_t | y_{1:t-1}) = F_tP_t^{t-1}F_t^{'} + V = Q_t$

Laboratório de Estatística e Geoinformação

Filtro de Kalman: Caso geral

- considerando ($\theta_0 \sim N(\theta_0^0, P_0^0)$)
- para t=1,2,...,T, calcular:
  1. Predição (θ_t): $[\theta_t | y_{1:t-1}]$ (condicional ao passado)
     - $E(\theta_t | y_{1:t-1}) = G\theta_{t-1}^{t-1} = \theta_t^{t-1}$
     - $V(\theta_t | y_{1:t-1}) = GP_{t-1}^{t-1}G^{'} + W = P_t^{t-1}$
  2. Predição (y_t): $[y_t | y_{1:t-1}]$ (condicional ao passado)
     - $E(y_t | y_{1:t-1}) = F_t\theta_t^{t-1}$, $e_t = y_t - F_t\theta_t^{t-1}$ (erro de previsão)
     - $V(y_t | y_{1:t-1}) = F_t P_t^{t-1} F_t^{'} + V = Q_t$
  3. Filtragem (θ_t): $[\theta_t | y_{1:t}]$ (condicional ao passado e presente)
     - $K_t = P_t^{t-1} F_t^{'} Q_t^{-1}$
     - $E(\theta_t | y_{1:t}) = \theta_t^{t-1} + K_t e_t = \theta_t^t$
     - $V(\theta_t | y_{1:t}) = P_t^{t-1} - K_t F_t P_t^{t-1} = P_t^t$

Suavização: Caso geral

- Suavização (retrospectiva): $[\theta_t|\theta_{t+1}, y_{1:T}]$
- via teorema de Bayes

$$[\theta_t|\theta_{t+1}, y_{1:T}] = \frac{[\theta_{t+1}|\theta_t][\theta_t|y_{1:t}]}{[\theta_{t+1}|y_{1:t}]}$$

Suavização: Caso geral

- Suavização (retrospectiva): $[\theta_t|\theta_{t+1},y_{1:T}]$
- via teorema de Bayes

$$[\theta_t|\theta_{t+1},y_{1:T}] = \frac{[\theta_{t+1}|\theta_t][\theta_t|y_{1:t}]}{[\theta_{t+1}|y_{1:t}]}$$

- Obtem θ_t^T e P_t^T
- Inicia com $\theta_t^T = \theta_t^t$ e $P_n^T = P_t^t$, em $t = T$
- Para $t = T, T-1, ..., 1$:

$$\begin{aligned}
J_{t-1} &= P_{t-1}^{t-1} G^{'} (P_t^{t-1})^{-1} \\
\theta_{t-1}^T &= \theta_{t-1}^{t-1} + J_{t-1}(\theta_t^T - \theta_t^{t-1}) \\
P_{t-1}^T &= P_{t-1}^{t-1} + J_{t-1}(P_t^T - P_t^{t-1})J_{t-1}^{'}
\end{aligned}$$

Inferência Bayesiana

- A inferência por Verossimilhança sobre θ é condicional a $\hat{\psi}$
- A inferência Bayesiana
  - possibilidade de inclusão de informação *a priori*
  - considerar incerteza de ψ na estimação de θ

Laboratório de Estatística e Geoinformação

Inferência Bayesiana

- A inferência por Verossimilhança sobre θ é condicional a $\hat{\psi}$
- A inferência Bayesiana
  - possibilidade de inclusão de informação *a priori*
  - considerar incerteza de ψ na estimação de θ
- Verossimilhança: Calculada como visto anteriormente
- Nos exemplos: G e W diagonais
  - $\psi_1 = diag\{G\} = \{\psi_1[1], ..., \psi_1[k]\}$, $k = nrow(G)$
  - $\psi_2 = \{diag\{W\}, V\} = \{\psi_2[1], ..., \psi_2[k], \psi_2[k+1]\}$

Laboratório de Estatística e Geoinformação

Inferência Bayesiana

- A inferência por Verossimilhança sobre θ é condicional a $\hat{\psi}$
- A inferência Bayesiana
  - possibilidade de inclusão de informação *a priori*
  - considerar incerteza de ψ na estimação de θ
- Verossimilhança: Calculada como visto anteriormente
- Nos exemplos: G e W diagonais
  - $\psi_1 = diag\{G\} = \{\psi_1[1], ..., \psi_1[k]\}$, $k = nrow(G)$
  - $\psi_2 = \{diag\{W\}, V\} = \{\psi_2[1], ..., \psi_2[k], \psi_2[k+1]\}$
- *Priori*s, conveniências: independência e conjugação
  - $\psi_1[i] \sim N(m_i, s2_i)$
  - $1/\psi_2[i] \sim Gamma(a_i, b_i)$, $i = 1, 2, ..., k$

Laboratório de Estatística e Geoinformação

Distribuições condicionais completas

- $\psi_1[i] | y_{1:T}, \theta_{1:T} \sim N(m_i^*, s2_i^*)$, $i = 1,\ldots,k$, onde

$$s2_i^* = (\frac{1}{s2_i} + \psi_2[i] \sum_{t=1}^{n} \theta_t^2[i])^{-1}$$

$$m_i^* = s2_i^* (\frac{m_i}{s2_i^*} + \psi_2[i] \sum_{t=1}^{n} \theta_t[i] \theta_{t-1}[i]).$$

Laboratório de Estatística e Geoinformação

Distribuições condicionais completas

- $\psi_1[i] | y_{1:T}, \theta_{1:T} \sim N(m_i^*, s2_i^*)$, $i = 1, \dots, k$, onde

$$s2_i^* = (\frac{1}{s2_i} + \psi_2[i] \sum_{t=1}^{n} \theta_t^2[i])^{-1}$$

$$m_i^* = s2_i^* (\frac{m_i}{s2_i^*} + \psi_2[i] \sum_{t=1}^{n} \theta_t[i] \theta_{t-1}[i]).$$

- $\psi_2[i]$, $i = 1, \dots, k$

$$Gamma(a_i + \frac{n}{2}, b_i + \frac{1}{2} \sum_{t=1}^{n} (\theta_t[i] - G_{(i,i)} \theta_{t-1}[i])).$$

- $\psi_2[i] | y_{1:T}, \theta_{1:T}$, $i = k+1$

$$Gamma(a_{k+1} + \frac{n}{2}, b_i + \frac{1}{2} \sum_{t=1}^{n} (y_t - F_t \theta_t)).$$

Laboratório de Estatística e Geoinformação

Amostrador de Gibbs

- Dois blocos, para simular de ψ,$\theta|y_{1:T}$:
  1. considere valores iniciais de ψ, ψ^0
  2. para $i = 1, 2, ..., N$
     1. simule $\theta^{(i)}$ de $[\theta|\psi^{(i-1)}, y_{1:T}]$ (usando o algoritmo FFBS)
     2. simule $\psi^{(i)}$ de $[\psi|\psi^{(i-1)}, \theta^{(i)}, y_{1:T}]$

Laboratório de Estatística e Geoinformação

Algoritmo FFBS

- Na suavização temos $[\theta_t | y_{1:T}]$
- De forma recursiva $[\theta_t | \theta_{t+1}, y_{1:T}]$
- FFBS: **F***orward* **F***iltering* **B***ackward* **S***ampling*
- o algoritmo FFBS segue os passos:
  1. execute o passo de filtragem
  2. simule θ_n de $\theta_n | y_{1:T}$
  3. para $i = n-1, n-2, ..., 0$ simule θ_i de $\theta_i | \theta_{i+1}, y_{1:T}$.