Um Estudo sobre Modelos para Volatilidade Estocástica

Projeto para Qualificação em Mestrado

André Silva de Queiroz Orientadora: Prof.^a Cibele Queiroz da Silva, Ph.D.

> Departamento de Estatística Universidade de Brasília

18 de junho de 2015





A importância dos modelos de séries temporais.





- A importância dos modelos de séries temporais.
- Dados com a variância variável:
 - Tradicionalmente: ARCH e GARCH.
 - Nova proposta: Modelos de Volatilidade Estocástica.
- Inferência Bayesiana.

Introdução

Modelos Dinâmicos.





- A importância dos modelos de séries temporais.
- Dados com a variância variável:
 - Tradicionalmente: ARCH e GARCH.





- A importância dos modelos de séries temporais.
- Dados com a variância variável:
 - Tradicionalmente: ARCH e GARCH.
 - Nova proposta: Modelos de Volatilidade Estocástica.





- A importância dos modelos de séries temporais.
- Dados com a variância variável:
 - Tradicionalmente: ARCH e GARCH.
 - Nova proposta: Modelos de Volatilidade Estocástica.
- Inferência Bayesiana.





- A importância dos modelos de séries temporais.
- Dados com a variância variável:
 - Tradicionalmente: ARCH e GARCH.
 - Nova proposta: Modelos de Volatilidade Estocástica.
- Inferência Bayesiana.
- Modelos Dinâmicos.





- Família mais simples dentre os modelos dinâmicos.

$$Y_t = \mathbf{F}_t' \theta_t + \nu_t, \qquad \nu_t \sim \mathcal{N}(0, V_t), \tag{1}$$

$$\theta_t = \mathbf{G}_t \theta_{t-1} + \omega_t, \qquad \omega_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t).$$
 (2)



Modelo Linear Dinâmico (MLD)

- Família mais simples dentre os modelos dinâmicos.
- Segundo West e Harrison (1997) a definição do modelo linear dinâmico univariado é:

$$Y_t = \mathbf{F}_t' \mathbf{\theta}_t + \nu_t, \qquad \nu_t \sim \mathcal{N}(0, V_t),$$
 (1)

$$\theta_t = \mathbf{G}_t \theta_{t-1} + \omega_t, \qquad \omega_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t).$$
 (2)





Modelo Linear Dinâmico (MLD)

- Família mais simples dentre os modelos dinâmicos.
- Segundo West e Harrison (1997) a definição do modelo linear dinâmico univariado é:
 - Equação das Observações:

$$Y_t = \mathbf{F}_t' \mathbf{\theta}_t + \nu_t, \qquad \nu_t \sim \mathcal{N}(0, V_t),$$
 (1)

Equação do Sistema

$$\theta_t = \mathbf{G}_t \theta_{t-1} + \omega_t, \qquad \omega_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t).$$
 (2)





Modelo Linear Dinâmico (MLD)

- Família mais simples dentre os modelos dinâmicos.
- Segundo West e Harrison (1997) a definição do modelo linear dinâmico univariado é:
 - Equação das Observações:

$$Y_t = \mathbf{F}_t' \mathbf{\theta}_t + \nu_t, \qquad \nu_t \sim \mathcal{N}(0, V_t),$$
 (1)

Equação do Sistema:

$$\theta_t = \mathbf{G}_t \theta_{t-1} + \omega_t, \qquad \omega_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t).$$
 (2)





- Se { F, G, V, W}, então o modelo é chamado de constante.

- A informação inicial sobre θ_t é representada por:

$$(\theta_0|D_0) \sim \mathcal{N}(\mathbf{m}_0, \mathbf{C}_0).$$
 (3)



- O MLD é caracterizado pela quádrupla { \mathbf{F}_t, \mathbf{G}_t, V_t, \mathbf{W}_t \}.
- Se {F, G, V, W}, então o modelo é chamado de constante.

- A informação inicial sobre θ_t é representada por:

$$(\theta_0|D_0) \sim \mathcal{N}(\mathbf{m}_0, \mathbf{C}_0).$$
 (3)





- O MLD é caracterizado pela quádrupla $\{F_t, G_t, V_t, W_t\}$.
- Se {**F**, **G**, **V**, **W**}, então o modelo é chamado de constante.
- O modelo linear dinâmico constante engloba essencialmente todos os modelos lineares tradicionais de séries temporais.
- Modelos mais gerais podem ser propostos com ν_t e ω_t correlacionados. Porém esses novos modelos podem sempre ser reescritos satisfazendo as condições de independência, West e Harrison (1997).
- A informação inicial sobre θ_t é representada por:
 - Informação Inicial:

 $(\boldsymbol{\theta}_0|D_0) \sim \mathcal{N}(\boldsymbol{m}_0, \boldsymbol{C}_0).$ (3)





- O MLD é caracterizado pela quádrupla $\{F_t, G_t, V_t, W_t\}$.
- Se {**F**, **G**, **V**, **W**}, então o modelo é chamado de constante.
- O modelo linear dinâmico constante engloba essencialmente todos os modelos lineares tradicionais de séries temporais.
- Modelos mais gerais podem ser propostos com ν_t e ω_t correlacionados. Porém esses novos modelos podem sempre ser reescritos satisfazendo as condições de independência, West e Harrison (1997).
- A informação inicial sobre θ_t é representada por:
 - Informação Inicial:

 $(oldsymbol{ heta}_0|D_0) \sim \mathcal{N}(oldsymbol{m}_0, oldsymbol{\mathcal{C}}_0).$





- O MLD é caracterizado pela quádrupla {F_t, G_t, V_t, W_t}.
- Se {F, G, V, W}, então o modelo é chamado de constante.
- O modelo linear dinâmico constante engloba essencialmente todos os modelos lineares tradicionais de séries temporais.
- Modelos mais gerais podem ser propostos com ν_t e ω_t correlacionados. Porém esses novos modelos podem sempre ser reescritos satisfazendo as condições de independência, West e Harrison (1997).
- A informação inicial sobre θ_t é representada por:

$$(\boldsymbol{\theta}_0|D_0) \sim \mathcal{N}(\boldsymbol{m}_0, \boldsymbol{C}_0).$$
 (3)





- O MLD é caracterizado pela quádrupla $\{F_t, G_t, V_t, W_t\}$.
- Se {**F**, **G**, **V**, **W**}, então o modelo é chamado de constante.
- O modelo linear dinâmico constante engloba essencialmente todos os modelos lineares tradicionais de séries temporais.
- Modelos mais gerais podem ser propostos com ν_t e ω_t correlacionados. Porém esses novos modelos podem sempre ser reescritos satisfazendo as condições de independência, West e Harrison (1997).
- A informação inicial sobre θ_t é representada por:
 - Informação Inicial:

$$(\boldsymbol{\theta}_0|D_0) \sim \mathcal{N}(\boldsymbol{m}_0, \boldsymbol{C}_0).$$
 (3)





Equações de Atualização

• Seja, para algum m_t e C_t , a posteriori de θ_t :

$$(\boldsymbol{\theta}_t|D_t) \sim \mathcal{N}(\boldsymbol{m}_t, \boldsymbol{C}_t).$$
 (4)

 A equação (2) permite, então, calcular a distribuição a priori de θ_{t+1} dada a informação em t que é:

$$\theta_{t+1} = \mathbf{G}\theta_t + \omega_{t+1},$$

$$(\theta_{t+1}|D_t) \sim \mathbf{G}\mathcal{N}_{(\theta_t|D_t)}(\mathbf{m}_t, \mathbf{C}_t) + \mathcal{N}_{\omega_{t+1}}(\mathbf{0}, \mathbf{W}),$$

$$(\theta_{t+1}|D_t) \sim \mathcal{N}(\mathbf{a}_{t+1}, \mathbf{R}_{t+1}),$$
(5)

onde $\pmb{a}_{t+1} = \pmb{G}\pmb{m}_t$ e $\pmb{R}_{t+1} = \pmb{G}\pmb{C}_t\pmb{G}' + \pmb{W}$



Equações de Atualização

• Seja, para algum m_t e C_t , a posteriori de θ_t :

$$(\boldsymbol{\theta}_t|D_t) \sim \mathcal{N}(\boldsymbol{m}_t, \boldsymbol{C}_t).$$
 (4)

• A equação (2) permite, então, calcular a distribuição a priori de θ_{t+1} dada a informação em t que é:

$$egin{aligned} oldsymbol{ heta}_{t+1} &= oldsymbol{G}oldsymbol{ heta}_{t} + oldsymbol{\omega}_{t+1}, \ &(oldsymbol{ heta}_{t+1}|D_{t}) \sim oldsymbol{G}\mathcal{N}_{(oldsymbol{ heta}_{t}|D_{t})}(oldsymbol{m}_{t}, oldsymbol{C}_{t}) + \mathcal{N}_{oldsymbol{\omega}_{t+1}}(oldsymbol{0}, oldsymbol{W}), \ &(oldsymbol{ heta}_{t+1}|D_{t}) \sim \mathcal{N}(oldsymbol{a}_{t+1}, oldsymbol{R}_{t+1}), \end{aligned}$$

onde $a_{t+1} = Gm_t$ e $R_{t+1} = GC_tG' + W$.





• A equação (1) é utilizada para calcular a distribuição da previsão da observação de Y_{t+1} dada a informação em t, ou a primeira previsão:

$$Y_{t+1} = \mathbf{F}' \boldsymbol{\theta}_{t+1} + \nu_{t+1},$$

$$(Y_{t+1}|D_t) \sim \mathbf{F}' \mathcal{N}_{(\boldsymbol{\theta}_{t+1}|D_t)}(\boldsymbol{a}_{t+1}, \boldsymbol{R}_{t+1}) + \mathcal{N}_{\nu_{t+1}}(0, V),$$

$$(Y_{t+1}|D_t) \sim \mathcal{N}(f_{t+1}, Q_{t+1}),$$
(6)

onde $f_{t+1} = \mathbf{F}' \mathbf{a}_{t+1}$ e $Q_{t+1} = \mathbf{F}' \mathbf{R}_{t+1} \mathbf{F} + V$.





• A distribuição *a posteriori* de θ_{t+1} dada a informação em t+1, que será *a priori* na próxima iteração, será:

$$(\boldsymbol{\theta}_{t+1}|D_{t+1}) \sim \mathcal{N}(\boldsymbol{m}_{t+1}, \boldsymbol{C}_{t+1}), \tag{7}$$

onde,
$$m{m}_{t+1} = m{a}_{t+1} + m{A}_{t+1} e_{t+1}$$
 e $m{C}_{t+1} = m{R}_{t+1} - m{A}_{t+1} Q_{t+1} m{A}'_{t+1}$, com $m{A}_{t+1} = m{R}_{t+1} m{F} Q_{t+1}^{-1}$ e $e_{t+1} = Y_{t+1} - f_{t+1}$.



- Para definir o k-ésimo valor predito, é necessário antes encontrar a distribuição de θ_{t+k} .
- Para um valor de $k \geq 2$, Pole et al. (1994) mostra que:

$$(\boldsymbol{\theta}_{t+k}|D_t) \sim \mathcal{N}(\boldsymbol{a}_{t+k}, \boldsymbol{R}_{t+k}),$$
 (8)

$$\mathbf{R}_{t+k} = \mathbf{G}^{k-1}\mathbf{R}_{t+1}\left(\mathbf{G}^{k-1}\right)' + \sum_{j=2}^{k} \mathbf{G}^{k-j}\mathbf{W}\left(\mathbf{G}^{k-j}\right)'.$$

$$(Y_{t+k}|D_t) \sim \mathcal{N}(f_{t+k}, Q_{t+k}), \tag{9}$$

onde $f_{t+k} = \mathbf{F}' \mathbf{a}_{t+k}$ e $Q_{t+k} = \mathbf{F}' \mathbf{R}_{t+k} \mathbf{F}' + V$.



Previsões

- Para definir o k-ésimo valor predito, é necessário antes encontrar a distribuição de θ_{t+k} .
- Para um valor de $k \ge 2$, Pole et al. (1994) mostra que:

$$(\boldsymbol{\theta}_{t+k}|D_t) \sim \mathcal{N}(\boldsymbol{a}_{t+k}, \boldsymbol{R}_{t+k}),$$
 (8)

onde $m{a}_{t+k} = m{G}^{k-1}m{a}_{t+1}$ e $m{R}_{t+k} = m{G}^{k-1}m{R}_{t+1}\left(m{G}^{k-1}
ight)' + \sum\limits_{i=2}^k m{G}^{k-j}m{W}\left(m{G}^{k-j}
ight)'.$

 Com isso, a distribuição da k-ésima predição dada a informação D_t será:

$$(Y_{t+k}|D_t) \sim \mathcal{N}(f_{t+k}, Q_{t+k}), \tag{9}$$

onde $f_{t+k} = \mathbf{F}' \mathbf{a}_{t+k}$ e $Q_{t+k} = \mathbf{F}' \mathbf{R}_{t+k} \mathbf{F}' + V$.





Previsões

- Para definir o k-ésimo valor predito, é necessário antes encontrar a distribuição de θ_{t+k} .
- Para um valor de $k \ge 2$, Pole et al. (1994) mostra que:

$$(\boldsymbol{\theta}_{t+k}|D_t) \sim \mathcal{N}(\boldsymbol{a}_{t+k}, \boldsymbol{R}_{t+k}),$$
 (8)

onde $a_{t+k} = G^{k-1} a_{t+1}$ e

$$\mathbf{R}_{t+k} = \mathbf{G}^{k-1}\mathbf{R}_{t+1}\left(\mathbf{G}^{k-1}\right)' + \sum_{j=2}^{k} \mathbf{G}^{k-j}\mathbf{W}\left(\mathbf{G}^{k-j}\right)'.$$

Com isso, a distribuição da k-ésima predição dada a informação D_t será:

$$(Y_{t+k}|D_t) \sim \mathcal{N}(f_{t+k}, Q_{t+k}), \tag{9}$$

onde $f_{t+k} = \mathbf{F}' \mathbf{a}_{t+k}$ e $Q_{t+k} = \mathbf{F}' \mathbf{R}_{t+k} \mathbf{F}' + V$.





- MVE vs ARCH e GARCH.
- Bos (2012) cita que o MVE é pouco amigável.
 - ARCH e GARCH: Muitas variações do modelo, basicamente uma maneira de se estimar os parâmetros.
 - MVE: Basicamente uma definição do modelo, muitas maneiras de se estimar os parâmetros.

Kim et al. (1998)	
Nguyen (2007)	





- MVE vs ARCH e GARCH.
- Bos (2012) cita que o MVE é pouco amigável.
 - ARCH e GARCH: Muitas variações do modelo, basicamente uma maneira de se estimar os parâmetros.
 - MVE: Basicamente uma definição do modelo, muitas maneiras de se estimar os parâmetros.

Kim et al. (1998)	
Nguyen (2007)	





- MVE vs ARCH e GARCH.
- Bos (2012) cita que o MVE é pouco amigável.
 - ARCH e GARCH: Muitas variações do modelo, basicamente uma maneira de se estimar os parâmetros.
 - MVE: Basicamente uma definição do modelo, muitas maneiras de se estimar os parâmetros.

Kim et al. (1998)	
Nguyen (2007)	





- MVE vs ARCH e GARCH.
- Bos (2012) cita que o MVE é pouco amigável.
 - ARCH e GARCH: Muitas variações do modelo, basicamente uma maneira de se estimar os parâmetros.
 - MVE: Basicamente uma definição do modelo, muitas maneiras de se estimar os parâmetros.

Kim et al. (1998)	
Nguyen (2007)	





- MVE vs ARCH e GARCH.
- Bos (2012) cita que o MVE é pouco amigável.
 - ARCH e GARCH: Muitas variações do modelo, basicamente uma maneira de se estimar os parâmetros.
 - MVE: Basicamente uma definição do modelo, muitas maneiras de se estimar os parâmetros.

Método	Referência	Paradigma
Quasi-Maximum Likelihood (QML)	Harvey et al. (1994)	Clássico
Gaussian Mixture Sampling (GMS)	Kim et al. (1998)	Bayesiano
Simulated Method of Moments (SMM)	Gallant e Tauchen (1996)	Clássico
Importance Sampling (IS)	Durbin e Koopman (1997)	Clássico
Efficient Importance Sampling (EIS)	Richard e Zhang (2007)	Clássico
Improved Importance Sampling (IIS)	Nguyen (2007)	Clássico
Single Site Sampler (SSS)	Carter e Kohn (1994)	Bayesiano
MultiMove Sampler (MMS)	Shephard e Pitt (1997)	Bayesiano





Modelo Canônico

- Modelagem inicialmente proposta por Taylor (1982).
- Kim et al. (1998) definem o modelo canônico como:

$$\textbf{Modelo Canônico}: \begin{cases} Y_t &= e^{\frac{h_t}{2}} \delta_t, \\ h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \\ h_0 &\sim \mathcal{N}\left(\mu, \frac{\sigma_\eta^2}{1 - \phi^2}\right). \end{cases} \tag{10}$$

• Os termos δ_t e η_t são ruídos gaussianos, independentes no tempo e entre si, com distribuições:

$$\delta_t \sim \mathcal{N}(0,1)$$
 e $\eta_t \sim \mathcal{N}\left(0,\sigma_\eta^2\right)$. (11)

 O conjunto de parâmetros do modelo a serem estimados é definido pelo vetor:

$$= (\mu, \phi, \sigma_{\eta}^2). \tag{12}$$

Modelo Canônico

- Modelagem inicialmente proposta por Taylor (1982).
- Kim et al. (1998) definem o modelo canônico como:

• Os termos δ_t e η_t são ruídos gaussianos, independentes no tempo e entre si, com distribuições:

$$\delta_t \sim \mathcal{N}(0,1)$$
 e $\eta_t \sim \mathcal{N}\left(0,\sigma_\eta^2\right)$. (11)

 O conjunto de parâmetros do modelo a serem estimados é definido pelo vetor:

$$= (\mu, \phi, \sigma_{\eta}^2). \tag{12}$$

- Modelagem inicialmente proposta por Taylor (1982).
- Kim et al. (1998) definem o modelo canônico como:

(10)

$$\delta_t \sim \mathcal{N}(0,1)$$
 e $\eta_t \sim \mathcal{N}\left(0,\sigma_\eta^2\right)$. (11)

$$= (\mu, \phi, \sigma_{\eta}^2). \tag{12}$$

- Modelagem inicialmente proposta por Taylor (1982).
- Kim et al. (1998) definem o modelo canônico como:

• Os termos δ_t e η_t são ruídos gaussianos, independentes no tempo e entre si, com distribuições:

$$\delta_t \sim \mathcal{N}(0,1)$$
 e $\eta_t \sim \mathcal{N}\left(0,\sigma_\eta^2\right)$. (11)

$$=(\mu,\phi,\sigma_{\eta}^2).$$

Modelo Canônico

- Modelagem inicialmente proposta por Taylor (1982).
- Kim et al. (1998) definem o modelo canônico como:

Modelo Canônico :
$$\begin{cases} Y_t &= e^{\frac{h_t}{2}} \delta_t, \\ h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \\ h_0 &\sim \mathcal{N}\left(\mu, \frac{\sigma_\eta^2}{1 - \phi^2}\right). \end{cases} \tag{10}$$

• Os termos δ_t e η_t são ruídos gaussianos, independentes no tempo e entre si, com distribuições:

$$\delta_t \sim \mathcal{N}(0,1)$$
 e $\eta_t \sim \mathcal{N}\left(0,\sigma_\eta^2\right)$. (11)

O conjunto de parâmetros do modelo a serem estimados é definido pelo vetor:

$$\psi = (\mu, \phi, \sigma_{\eta}^2). \tag{12}$$

Um Modelo de Espaço-Estado

- Os MVE podem ser reescritos na notação de espaço-estado apresentada por West e Harrison (1997) estendendo a proposta de Zivot e Yollin (2012).

$$Y_t = e^{\frac{h_t}{2}} \delta_t,$$

$$\ln Y_t^2 = h_t + \ln \delta_t^2,$$

$$\ln Y_t^2 = h_t + \mathbb{E}(\ln \delta_t^2) + \varepsilon_t.$$
(13)





Um Modelo de Espaço-Estado

- Os MVE podem ser reescritos na notação de espaço-estado apresentada por West e Harrison (1997) estendendo a proposta de Zivot e Yollin (2012).
- Pela definição em (10):

$$Y_t = e^{\frac{h_t}{2}} \delta_t,$$

$$\ln Y_t^2 = h_t + \ln \delta_t^2,$$

$$\ln Y_t^2 = h_t + \mathsf{E}(\ln \delta_t^2) + \varepsilon_t.$$
(13)

Assim $\varepsilon_t \sim (0, \text{Var}(\ln \delta_t^2))$ é um ruído em torno de zero.





$$\ln \delta_t^2 \approx \xi_t = \sum_{i=1}^{10} w_i \mathcal{N}\left(m_j, s_j^2\right). \tag{14}$$





- Durbin e Koopman (2012) afirmam que isso não impede a utilização das técnicas associadas aos modelos gaussianos.
- Omori et al. (2007) sugere uma mistura de r = 10 normais que se aproxima da distribuição do erro:

$$\ln \delta_t^2 \approx \xi_t = \sum_{j=1}^{10} w_j \mathcal{N}\left(m_j, s_j^2\right). \tag{14}$$





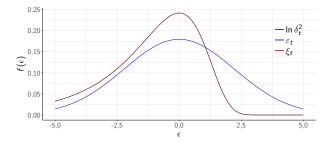


Figura: Comparação das distribuições de $\ln \delta_t^2$, ε_t e ξ_t .





• A equação que define a evolução da variável latente, h_t , em (10) pode ser reescrita:

$$h_{t} = \mu + \phi(h_{t-1} - \mu) + \eta_{t},$$

$$h_{t} = \mu - \phi\mu + \phi h_{t-1} + \eta_{t},$$

$$h_{t} = (1 - \phi)\mu + \phi h_{t-1} + \eta_{t}.$$
(15)

Portanto (13) e (15) definem a equação das observações e a





• A equação que define a evolução da variável latente, h_t , em (10) pode ser reescrita:

$$h_{t} = \mu + \phi(h_{t-1} - \mu) + \eta_{t},$$

$$h_{t} = \mu - \phi\mu + \phi h_{t-1} + \eta_{t},$$

$$h_{t} = (1 - \phi)\mu + \phi h_{t-1} + \eta_{t}.$$
(15)

Portanto (13) e (15) definem a equação das observações e a equação do sistem, respectivamente.





Assim, a equação das observações do MVE será:

$$Y_{t} = \mathbf{F}_{t}' \boldsymbol{\theta}_{t} + \nu_{t},$$

$$\ln Y_{t}^{2} = h_{t} + \mathsf{E}(\ln \delta_{t}^{2}) + \varepsilon_{t},$$

$$\ln Y_{t}^{2} = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_{t} \\ \mu \\ \mathsf{E}(\ln \delta_{t}^{2}) \end{bmatrix} + \varepsilon_{t}.$$
(16)





$$\theta_{t} = \mathbf{G}_{t}\theta_{t-1} + \omega_{t},$$

$$h_{t} = \phi h_{t-1} + (1 - \phi)\mu + \eta_{t},$$

$$\begin{bmatrix} h_{t} \\ \mu \\ \mathsf{E}(\ln \delta_{t}^{2}) \end{bmatrix} = \begin{bmatrix} \phi & 1 - \phi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_{t-1} \\ \mu \\ \mathsf{E}(\ln \delta_{t}^{2}) \end{bmatrix} + \begin{bmatrix} \eta_{t} \\ 0 \\ 0 \end{bmatrix}. \tag{17}$$





 O modelo de volatilidade estocástica é definido como um modelo de espaço-estado constante através das matrizes:

$$\mathbf{F}_{t} = \mathbf{F} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \text{e} \quad \mathbf{G}_{t} = \mathbf{G} = \begin{bmatrix} \phi & 1 - \phi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$
(18)

$$V_t = V = Var(\varepsilon_t)$$
 e $\mathbf{W}_t = \mathbf{W} = \begin{bmatrix} \sigma_{\eta}^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. (19)



Definição do Modelo Bayesiano

- Como o paradigma Bayesiano sugere, deve-se definir as distribuições a priori dos parâmetros. Foram assumidas as distribuições a priori conforme Kastner e Frühwirth-Schnatter (2014) e Kim et al. (1998).
- O nível μ da volatilidade tem seu suporte em R, e será atribuída uma distribuição a priori com densidade gaussiana para o seu valor:

$$\mu \sim \mathcal{N}(a_{\mu}, B_{\mu}).$$
 (20)





Definição do Modelo Bayesiano

- Como o paradigma Bayesiano sugere, deve-se definir as distribuições a priori dos parâmetros. Foram assumidas as distribuições a priori conforme Kastner e Frühwirth-Schnatter (2014) e Kim et al. (1998).
- O nível μ da volatilidade tem seu suporte em \mathbb{R} , e será atribuída uma distribuição *a priori* com densidade gaussiana para o seu valor:

$$\mu \sim \mathcal{N}(\mathsf{a}_{\mu}, \mathsf{B}_{\mu}).$$
 (20)





- O parâmetro ϕ determina a persistência da volatilidade, e $|\phi| < 1$. Seja uma nova variável ϕ_0 cuja distribuição *a priori* é $\mathcal{B}(a_\phi, b_\phi)$.
- A persistência então será $\phi = 2\phi_0 1$ e sua distribuição *a priori* será:

$$\pi(\phi) = \frac{\Gamma(a_{\phi} + b_{\phi})}{2\Gamma(a_{\phi})\Gamma(b_{\phi})} \left(\frac{1+\phi}{2}\right)^{a_{\phi}-1} \left(\frac{1-\phi}{2}\right)^{b_{\phi}-1}.$$
 (21)



- O parâmetro ϕ determina a persistência da volatilidade, e $|\phi| < 1$. Seja uma nova variável ϕ_0 cuja distribuição *a priori* é $\mathcal{B}(a_{\phi},b_{\phi}).$
- A persistência então será $\phi = 2\phi_0 1$ e sua distribuição a priori será:

$$\pi(\phi) = \frac{\Gamma(a_{\phi} + b_{\phi})}{2\Gamma(a_{\phi})\Gamma(b_{\phi})} \left(\frac{1+\phi}{2}\right)^{a_{\phi}-1} \left(\frac{1-\phi}{2}\right)^{b_{\phi}-1}.$$
 (21)





• O terceiro e último parâmetro do modelo é σ_n^2 , a variância da volatilidade, e seus possíveis valores estão em \mathbb{R}^+ . Sua distribuição a priori será:

$$\sigma_{\eta}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2B_{\sigma}}\right).$$
 (22)





- Foi realizado um estudo com dados simulados para avaliar um processo de estimação dos parâmetros.
- Foram geradas 64 populações (n = 1461) com cada combinação dos valores resultantes do produto cruzado de:

```
* \mu \in \{-5,4\},

* \phi \in \{0,0;0,65;0,99\},

* \sigma_n^2 \in \{0,1^2;0,3^2;0,5^2\}
```





- Foi realizado um estudo com dados simulados para avaliar um processo de estimação dos parâmetros.
- Foram geradas 64 populações (n = 1461) com cada combinação dos valores resultantes do produto cruzado de:

```
 \begin{array}{ll} & \mu \in \{-5,4\}, \\ & \phi \in \{0,0;0,65;0,99\}, \\ & \sigma_{\eta}^2 \in \{0,1^2;0,3^2;0,5^2\}. \end{array}
```





- Foi realizado um estudo com dados simulados para avaliar um processo de estimação dos parâmetros.
- Foram geradas 64 populações (n = 1461) com cada combinação dos valores resultantes do produto cruzado de:
 - $\mu \in \{-5,4\}$,
 - $\phi \in \{0,0;0,65;0,99\},$
 - $\sigma_{\eta}^2 \in \{0,1^2;0,3^2;0,5^2\}.$





- Foi realizado um estudo com dados simulados para avaliar um processo de estimação dos parâmetros.
- Foram geradas 64 populações (n = 1461) com cada combinação dos valores resultantes do produto cruzado de:
 - $\mu \in \{-5,4\},$
 - $\phi \in \{0.0; 0.65; 0.99\}$,
 - $\sigma_{\eta}^2 \in \{0,1^2;0,3^2;0,5^2\}.$





- Foi realizado um estudo com dados simulados para avaliar um processo de estimação dos parâmetros.
- Foram geradas 64 populações (n = 1461) com cada combinação dos valores resultantes do produto cruzado de:
 - $\mu \in \{-5,4\}$,
 - $\phi \in \{0.0, 0.65, 0.99\}$,
 - $\sigma_n^2 \in \{0,1^2;0,3^2;0,5^2\}.$





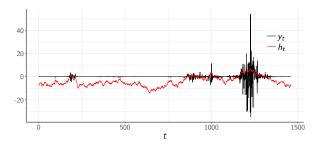


Figura: Conjunto de dados simulados a partir do modelo (10), $\mu = -5.4$, $\phi = 0.99$ e $\sigma_n^2 = 0.5^2$.





- O método de estimação testado será baseado no amostrador de Gibbs via JAGS.
- JAGS (Just Another Gibbs Sampler).
- Por que JAGS?
 - Desenvolvimento
 - Integração
- Sintaxe parecida com R





- O método de estimação testado será baseado no amostrador de Gibbs via JAGS.
- JAGS (Just Another Gibbs Sampler).
- Por que JAGS?
 - Desenvolvimento.
 - Integração.
- Sintaxe parecida com R.





- O método de estimação testado será baseado no amostrador de Gibbs via JAGS.
- JAGS (Just Another Gibbs Sampler).
- Por que JAGS?
 - Desenvolvimento.
 - Integração
- Sintaxe parecida com R.





- O método de estimação testado será baseado no amostrador de Gibbs via JAGS.
- JAGS (Just Another Gibbs Sampler).
- Por que JAGS?
 - Desenvolvimento.
 - Integração.
- Sintaxe parecida com R





- O método de estimação testado será baseado no amostrador de Gibbs via JAGS.
- JAGS (Just Another Gibbs Sampler).
- Por que JAGS?
 - Desenvolvimento.
 - Integração.
- Sintaxe parecida com R.





- O método de estimação testado será baseado no amostrador de Gibbs via JAGS.
- JAGS (Just Another Gibbs Sampler).
- Por que JAGS?
 - Desenvolvimento.
 - Integração.
- Sintaxe parecida com R.





- burn-in = 5.000
- iterações = 10.000
- thin = 10.
- Distribuições a priori dos parâmetros:

$$\mu \sim \mathcal{N}(0, 100), \quad \phi_0 \sim \mathcal{B}(1, 1) \quad \text{e} \quad \sigma_\eta^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right). \quad (23)$$



- burn-in = 5.000.
- iterações = 10.000
- thin = 10
- Distribuições a priori dos parâmetros:

$$\mu \sim \mathcal{N}(0, 100), \quad \phi_0 \sim \mathcal{B}(1, 1) \quad \text{e} \quad \sigma_{\eta}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right). \quad (23)$$



- burn-in = 5.000.
- iterações = 10.000.
- thin = 10
- Distribuições a priori dos parâmetros:

$$\mu \sim \mathcal{N}(0, 100), \quad \phi_0 \sim \mathcal{B}(1, 1) \quad \text{e} \quad \sigma_{\eta}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right). \quad (23)$$



- burn-in = 5.000.
- iterações = 10.000.
- thin = 10.
- Distribuições a priori dos parâmetros:

$$\mu \sim \mathcal{N}(0, 100), \quad \phi_0 \sim \mathcal{B}(1, 1) \quad \text{e} \quad \sigma_{\eta}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right). \quad (23)$$



Procedimento:

- burn-in = 5.000.
- iterações = 10.000.
- thin = 10.
- Distribuições a priori dos parâmetros:

$$\mu \sim \mathcal{N}(0, 100), \quad \phi_0 \sim \mathcal{B}(1, 1) \quad \text{e} \quad \sigma_{\eta}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right). \quad (23)$$





Resultados

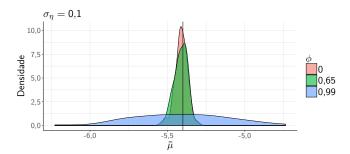


Figura: Valores de $\bar{\mu}$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.





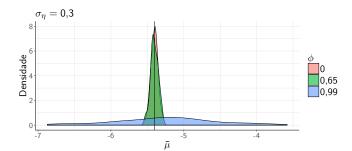


Figura: Valores de $\bar{\mu}$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.





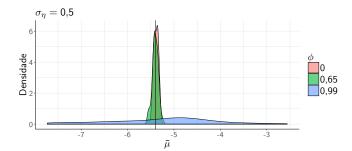


Figura: Valores de $\bar{\mu}$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.





ϕ_{real}	$\sigma_{\eta_{real}}$	média $(ar{\mu})$	mín $(ar{\mu})$	$Md(ar{\mu})$	máx $(ar{\mu})$	$\sigma_{ar{\mu}}$
0	0,1	-5,410	-5,521	-5,409	-5,312	0,039
0	0,3	-5,403	-5,527	-5,401	-5,284	0,049
0	0,5	-5,388	-5,555	-5,386	-5,293	0,053
0,65	0,1	-5,414	-5,533	-5,411	-5,308	0,043
0,65	0,3	-5,402	-5,529	-5,408	-5,292	0,052
0,65	0,5	-5,402	-5,558	-5,401	-5,242	0,065
0,99	0,1	-5,411	-6,228	-5,413	-4,736	0,297
0,99	0,3	-5,240	-6,872	-5,241	-3,581	0,699
0,99	0,5	-5,277	-7,728	-4,995	-2,560	1,142





Resultados



Figura: Valores de $\bar{\phi}$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.





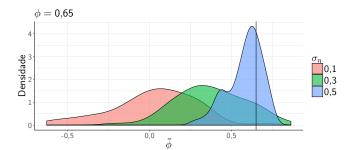


Figura: Valores de $\bar{\phi}$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.





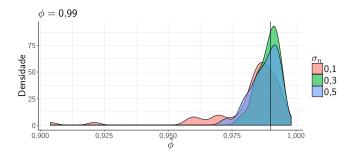


Figura: Valores de $\bar{\phi}$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.





0,99

0,99

0,3

0,5

0,989

0,988

amostrador de Gibbs para o conjunto de populações de dados simulados.

		•	,	,		
ϕ_{real}	$\sigma_{\eta_{real}}$	média $(ar{\phi})$	mín $(ar{\phi})$	$Md(ar{\phi})$	máx $(ar{\phi})$	$\sigma_{ar{\phi}}$
0	0,1	0,012	-0,657	0,017	0,503	0,248
0	0,3	-0,006	-0,547	-0,013	0,810	0,271
0	0,5	-0,027	-0,465	-0,032	0,448	0,218
0,65	0,1	0,025	-0,629	0,054	0,708	0,259
0,65	0,3	0,373	-0,209	0,364	0,863	0,213
0,65	0,5	0,582	0,277	0,602	0,750	0,105
0,99	0,1	0,982	0,904	0,986	0,998	0,016

0,977

0,972

0,990

0,989





0,004

0,005

0,997

0,995

Resultados

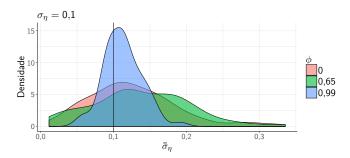


Figura: Valores de $\bar{\sigma}_{\eta}$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.





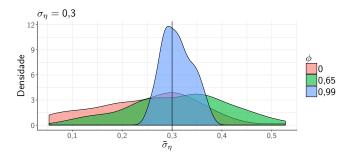


Figura: Valores de $\bar{\sigma}_{\eta}$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.





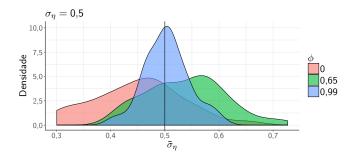


Figura: Valores de $\bar{\sigma}_{\eta}$ estimados através do amostrador de Gibbs para o conjunto das populações de dados simulados.





Tabela: Principais estatísticas dos valores estimados $\bar{\sigma}_{\eta}$ através do amostrador de Gibbs para o conjunto de populações de dados simulados.

ϕ_{real}	$\sigma_{\eta_{real}}$	média $(ar{\sigma}_{\eta})$	mín $(ar{\sigma}_{\eta})$	$Md(ar{\sigma}_\eta)$	máx $(ar{\sigma}_{\eta})$	$\sigma_{\bar{\sigma}_{\eta}}$
0	0,1	0,126	0,012	0,118	0,323	0,061
0	0,3	0,247	0,053	0,271	0,527	0,104
0	0,5	0,451	0,300	0,454	0,658	0,082
0,65	0,1	0,137	0,013	0,128	0,336	0,067
0,65	0,3	0,314	0,097	0,333	0,527	0,100
0,65	0,5	0,539	0,412	0,546	0,727	0,072
0,99	0,1	0,113	0,055	0,112	0,192	0,026
0,99	0,3	0,310	0,258	0,308	0,372	0,030
0,99	0,5	0,504	0,404	0,501	0,604	0,041





Exemplo com Dados Reais

- Replicação de (Achcar et al., 2011).
- Estudar a concentração de ozônio na Cidade do México
- 16 anos (1/1/1990 a 31/12/2005) de medições da média semanal do nível diário máximo de ozônio em cinco regiões da cidade.
- Os autores propõe dois modelos.





Exemplo com Dados Reais

- Replicação de (Achcar et al., 2011).
- Estudar a concentração de ozônio na Cidade do México.
- 16 anos (1/1/1990 a 31/12/2005) de medições da média semanal do nível diário máximo de ozônio em cinco regiões da cidade
- Os autores propõe dois modelos





Exemplo com Dados Reais

- Replicação de (Achcar et al., 2011).
- Estudar a concentração de ozônio na Cidade do México.
- 16 anos (1/1/1990 a 31/12/2005) de medições da média semanal do nível diário máximo de ozônio em cinco regiões da cidade.
- Os autores propõe dois modelos





- Replicação de (Achcar et al., 2011).
- Estudar a concentração de ozônio na Cidade do México.
- 16 anos (1/1/1990 a 31/12/2005) de medições da média semanal do nível diário máximo de ozônio em cinco regiões da cidade.
- Os autores propõe dois modelos.





• No tempo t = 1, ..., N e na região j = 1, ..., K, como:

Modelo I :
$$\begin{cases} Y_{j,t} \sim \mathcal{N}\left(0, e^{h_{j,t}}\right) \\ h_{j,t} = \begin{cases} \mu_{j} + \eta_{j,t}, & t = 1, \\ \mu_{j} + \phi_{j}(h_{j,t-1} - \mu_{j}) + \eta_{j,t}, & t = 2, \dots, N. \end{cases}$$
(24)





Modelo II

• No tempo $t=1,\ldots,N$ e na região $j=1,\ldots,K$, como:

$$\textbf{Modelo II}: \begin{cases} Y_{j,t} \sim \mathcal{N}\left(0, e^{h_{j,t} + \omega_t}\right). \\ h_{j,t} = \begin{cases} \mu_j + \eta_{j,t}, & t = 1, \\ \mu_j + \phi_j(h_{j,t-1} - \mu_j) + \eta_{j,t}, & t = 2, \dots, N, \\ \omega_t \sim \mathcal{N}\left(0, \sigma_\omega^2\right). \end{cases}$$

$$(25)$$

• A distribuição *a priori* de σ_{ω}^2 é:

$$\sigma_{\omega}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2B_{\sigma}}\right).$$
 (26)





Modelo II

• No tempo $t=1,\ldots,N$ e na região $j=1,\ldots,K$, como:

$$\textbf{Modelo II}: \begin{cases} Y_{j,t} \sim \mathcal{N}\left(0, e^{h_{j,t} + \omega_t}\right). \\ h_{j,t} = \begin{cases} \mu_j + \eta_{j,t}, & t = 1, \\ \mu_j + \phi_j(h_{j,t-1} - \mu_j) + \eta_{j,t}, & t = 2, \dots, N, \\ \omega_t \sim \mathcal{N}\left(0, \sigma_\omega^2\right). \end{cases}$$
(25)

• A distribuição *a priori* de σ_{ω}^2 é:

$$\sigma_{\omega}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2B_{\sigma}}\right).$$
 (26)





- O procedimento foi idêntico ao anterior com os dados simulados.
- Distribuição *a priori* de σ_{ω}^2 :

$$\sigma_{\omega}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right).$$
 (27)



- O procedimento foi idêntico ao anterior com os dados simulados.
- Distribuição *a priori* de σ_{ω}^2 :

$$\sigma_{\omega}^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right).$$
 (27)





Resultados

Tabela: Comparação entre valores estimados dos parâmetros do modelo I em Achcar et al. (2011) e do modelo canônico para a região NW.

Região	Parâmetro	Modelo	Parame-	Valor	Desvio	95% Int. de
			trização	estimado	padrão	Credibilidade
NW	μ	Modelo I	σ^2	-2,817	0,075	(-2,972; -2,670)
		Modelo I	$1/\sigma^2$	-2,796	0,080	(-2,950; -2,635)
		M. Canônico	$1/\sigma^2$	-2,789	0,078	(-2.947; -2.631)
	ϕ	Modelo I	σ^2	0,690	0,086	(0,498; 0,830)
		Modelo I	$1/\sigma^2$	0,791	0,076	(0,628; 0,912)
		M. Canônico	$1/\sigma^2$	0,775	0,093	(0,546; 0,903)
	σ_{η}	Modelo I	σ^2	5,413	1,420	(3,092; 8,610)
		Modelo I	$1/\sigma^2$	0,336	0,077	(0,203; 0,500)
		M. Canônico	$1/\sigma^2$	0,327	0,075	(0,214; 0,485)





Tabela: Comparação entre valores estimados dos parâmetros do modelo II em Achcar et al. (2011) e do modelo canônico para a região NW.

	D A .		<u> </u>		.	050/ 1
Região	Parâmetro	Modelo	Parame-	Valor	Desvio	95% Int. de
			trização	estimado	padrão	Credibilidade
NW	μ	Modelo II	σ^2	-3,008	0,072	(-3,148; -2,854)
		Modelo II	$1/\sigma^2$	-2,993	0,045	(-3,063; -2,902)
		M. Canônico	$1/\sigma^2$	-3,000	0,055	(-3.094; -2.871)
	ϕ	Modelo II	σ^2	0,300	0,184	(0,021; 0,700)
		Modelo II	$1/\sigma^2$	0,009	0,504	(-0,755; 0,892)
		M. Canônico	$1/\sigma^2$	-0,254	0,544	(-0,915; 0,833)
	σ_{η}	Modelo II	σ^2	8,755	2,172	(5,027; 13,490)
		Modelo II	$1/\sigma^2$	0,036	0,025	(0,010; 0,086)
		M. Canônico	$1/\sigma^2$	0,055	0,027	(0,020; 0,126)
-	σ_{ω}	Modelo II	σ^2	1,279	0,106	(1,091; 1,512)
		Modelo II	$1/\sigma^2$	0,912	0,036	(0,843; 0,983)
		M. Canônico	$1/\sigma^2$	0,916	0,034	(0,848; 0,983)





Referências Bibliográficas I

- Achcar, J. A., Rodrigues, E. R., e Tzintzun, G. (2011). Using stochastic volatility models to analyse weekly ozone averages in Mexico City. *Environmental and Ecological Statistics*, 18(2):271–290.
- Bos, C. S. (2012). Handbook of Volatility Models and their Applications, chapter Relating Stochastic Volatility Estimation Methods, pages 147–174. John Wiley & Sons.
- Carter, C. K. e Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81:541–553.
- Durbin, J. e Koopman, S. J. (1997). Monte Carlo maximum likelihood estimation for non-gaussian state-space models. *Biometrika*, 89:603–615.
- Durbin, J. e Koopman, S. J. (2012). *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford, 2nd edition.



Referências Bibliográficas II

- Gallant, A. R. e Tauchen, G. (1996). Which moments to match? *Econometric Theory*, 12:657–681.
- Harvey, A., Ruiz, E., e Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies*, 61:247–264.
- Kastner, G. e Frühwirth-Schnatter, S. (2014).
 Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models.

 Computational Statistics and Data Analysis, 76:408–423.
- Kim, S., Shepard, N., e Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies*, 65(3):361–393.
- Nguyen, T. M. (2007). A new efficient algorithm for the analysis with non-linear and non-gaussian state space models. Master's thesis, VU University Amsterdam.



Referências Bibliográficas III

- Omori, Y., Chib, S., Shephard, N., e Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140:425–449.
- Pole, A., West, M., e Harrison, J. (1994). *Applied Bayesian Forecasting and Time Series Analysis*. Chapman & Hall/CRC, Boca Raton, 1st edition.
- Richard, J. F. e Zhang, W. (2007). Efficient high-dimensional importance sampling. *Journal of Econometrics*, 141:1385–1411.
- Shephard, N. e Pitt, M. K. (1997). Likelihood analysis of non-gaussian measurement time series. *Biometrika*, 84:653–667.
- Taylor, S. (1982). Financial returns modelled by the product of two stochastic processes a study of daily sugar prices 1961-75. Time Series Analysis: Theory and Practice, 1:203–226.
- West, M. e Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. Springer, New York, 2nd edition.



Referências

Referências Bibliográficas IV

Zivot, E. e Yollin, G. (2012). Time series forecasting with state space models. In *R/Finance Workshop, University of Chicago*.



