

Modelo dinâmico para um processo auto-regressivo de primeira ordem, aplicando metodologia Bayesiana

Leonilce Mena^{1*} e Marinho G. Andrade²

¹Departamento de Estatística, Universidade Estadual de Maringá, Av. Colombo, 5790, 87020-900, Maringá, Paraná, Brasil.

²Instituto de Computação de Matemática Computacional, Universidade de São Paulo, C.P 668.13560-970, São Carlos, São Paulo, Brasil *Author for correspondence. e-mail: lmena@uem.br

RESUMO. Neste artigo apresentamos uma ramificação do processo auto-regressivo de primeira ordem com coeficiente aleatório e variante no tempo, assumindo uma estrutura de dependência dos coeficientes aleatórios, que leva a um modelo de filtro de Kalman adaptado. Embora o modelo de filtro de Kalman considerado seja uma generalização do filtro de Kalman Ordinário, sua análise produz dificuldades técnicas, porque não é possível encontrar uma forma fechada para o filtro, assim aplicamos simulação de Monte Carlo em Cadeia de Markov utilizando os algoritmos Amostrador de Gibbs e Metropolis-Hasting para fazer inferência quanto aos parâmetros do modelo e também fazer previsões de dados de uma série temporal de índice de preços de ações e preço do boi gordo.

Palavras-chave: processo Auto-regressivo, inferência Bayesiana, modelo dinâmico, filtro de Kalman, Gibbs-Sampling, Metropolis-Hastings.

ABSTRACT. Dynamic model for a first order autoregression process with Bayesian methodology. A ramification of a first order autoregression process is provided. It comprises randomized and variant coefficients in time and assumes a structure of dependency of randomized coefficients that leads towards adapted Kalman's Filter. Although the Kalman Filter model is a generalization of the ordinary Kalman Filter, its analysis produces technical difficulties. It does not seem to be impossible to find a closed form for the filter. Monte Carlo's simulation was applied to Markov's Chain by Gibbs-Sampling and Metropolis-Hasting algorithms to infer parameters of model and work out forecasts of data for a time series of indexes of shares and meat prices.

Key words: auto regression process, dynamic model, Kalman filter, Bayesian inference Gibbs-Sampling, Metropolis-Hasting.

O propósito deste trabalho é formular um procedimento bayesiano, utilizando simulação de Monte Carlo em Cadeia de Markov (MCMC), para análise e previsão de séries temporais modeladas por processos auto-regressivos com coeficientes aleatórios denominados processos auto-regressivos generalizados. Aqui, apresenta-se uma ramificação do processo auto-regressivo de primeira ordem com coeficiente aleatório assumindo uma estrutura de dependência dos coeficientes do processo estocástico (θ_t, s) tão forte quanto sua permutabilidade. Descreve-se o modelo para uma série temporal em função de seus componentes não observáveis θ_t, s quando θ_t é expresso em função dos valores desses mesmos componentes no instante anterior.

Concentramo-nos no procedimento pelo qual obtemos estimadores atualizados dos componentes

não observáveis a todo instante de tempo, a partir da informação dada pelo componente observável do sistema, ou seja, y_t . Dada a natureza linear dinâmica do modelo, esta atualização seqüencial pode ser obtida representando o modelo sob a forma de espaço de estados e, então, utilizar o modelo Filtro de Kalman, cujo procedimento é recursivo.

O filtro de Kalman, segundo (Meinhold e Singpurwalla, 1990), proporciona a elaboração de um algoritmo de estimação recursivo, que tem sido utilizado com sucesso por engenheiros, na teoria de controle, e, mais recentemente por estatísticos e econometristas. Descreve-se a utilização do filtro de Kalman em problemas de previsão, onde a série temporal é modelada por uma média, que varia no tempo, superposta a um ruído aditivo. Essa média é por hipótese, uma combinação linear de funções conhecidas cujos coeficientes são desconhecidos.

Desse modo, a série pode ser representada por um sistema linear cujo vetor de estados é formado pelos coeficientes desconhecidos e pelo valor da média do processo no instante t . Nessas circunstâncias, o filtro de Kalman pode ser utilizado para obter-se estimativas ótimas do vetor de estados, com a vantagem de permitir a variação dos coeficientes através do tempo.

Seja o processo estocástico X_t ; $t = 1, 2, \dots$ com distribuição log-normal denotado por:

$$X_t = X_{t-1}^{\theta_t}; \quad t = 1, 2, \dots \quad (1)$$

onde θ_t é o coeficiente cujos valores descrevem um crescimento ou decrescimento no processo.

Para produzir uma generalização no modelo, introduziu-se um erro multiplicativo δ_t , assim temos:

$$X_t = X_{t-1}^{\theta_t} \delta_t; \quad t = 1, 2, \dots \quad (2)$$

Convenientemente, assumiu-se que δ_t também com distribuição log-normal e, em alguns problemas, é conveniente supor que o coeficiente θ_t tem característica aleatória. (Diaz, 1990; Liu, 1995).

Se for aplicado o logaritmo natural, em ambos os lados, em (2) e seja $\varepsilon_t = \log(\delta_t)$ então, o modelo torna-se:

$$y_t = \theta_t y_{t-1} + \varepsilon_t; \quad t = 1, 2, \dots \quad (3)$$

onde, $y_t = \log(X_t)$ com y_t e ε_t normalmente distribuídos, ε_t tem média zero e variância τ^2 .

A partir desse ponto, enfocaram-se as variáveis y_t ao invés das variáveis X_t , com o objetivo de estimar θ_t e de fazer previsão para observações futuras y_{t+1} .

A previsão de θ_t permite analisar o comportamento do processo que depende dos valores dos θ_t , ou seja, o crescimento do processo ocorre quando $\theta_t \geq 1$ e o decrescimento ocorre

quando $\theta_t < 1$. Um caso particular ocorre quando $\theta_t = 0$, pois teremos da equação (3) que $y_t = \varepsilon_t$ o que implica y_t não depender de y_{t-1} , ou seja, y_t é um ruído branco.

Assumindo os θ_t 's com densidade de probabilidade normal de média λ e variância γ^2 , a estratégia para assegurar que os θ_t 's sejam dependentes é assumirmos um modelo auto-regressivo de primeira ordem para θ_t e também uma distribuição de probabilidade para λ isto é, assumimos que λ tem distribuição normal com média m e variância S^2 .

Pode-se, assim, resumir o modelo da seguinte maneira:

$$\begin{aligned} y_t &= \theta_t y_{t-1} + \varepsilon_t \\ \theta_t &= \lambda \theta_{t-1} + \omega_t \\ \omega_t &\sim N(0; \delta^2) \\ \lambda &\sim N(m; S^2); \quad m, s \text{ conhecidos} \end{aligned} \quad (4)$$

Na equação (4), assumiu-se θ_t independente de ε_t e dependente de seus valores passados, então será determinado $y^{(t)} = (y_1, \dots, y_t)$. Dado $y^{(t)}$, nosso objetivo é fazer inferências sobre λ , θ_t e fazer previsão de valores futuros de y_{t+k} , $k \geq 1$.

Considerando todas as quantidades como escalares e fazendo $G_t = \lambda$, $F_t = y_{t-1}$, $V_t = \varepsilon_t$ e $W_t = \omega_t$. Uma extensão desse procedimento, para considerar os valores de θ_t dependentes de seus valores passados e assumindo o modelo auto-regressivo de primeira ordem para θ_t , pode-se escrever o modelo como:

$$\begin{aligned} y_t &= \theta_t y_{t-1} + \varepsilon_t \\ \theta_t &= \lambda \theta_{t-1} + \omega_t \end{aligned} \quad (5)$$

$$\text{onde } \begin{pmatrix} \varepsilon_t \\ \omega_t \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} \sigma^2 & 0 \\ 0 & \gamma^2 \end{pmatrix} \right]$$

σ^2 e γ^2 conhecidos

No modelo Filtro de Kalman ordinário, as quantidades F_t, G_t, V_t e W_t , denominadas matriz de medida, matriz de transição, erro observado e erro do sistema, respectivamente, são consideradas conhecidas. Elas podem variar no tempo, mas somente se a natureza da variação for considerada conhecida. Em muitas aplicações é difícil especificar algum ou todos esses parâmetros, devido à informação inadequada sobre o processo, ou porque os parâmetros são permutáveis, mas a natureza exata dessa permuta é desconhecida. Nesses casos, o filtro de Kalman é aplicado para estimar os parâmetros desconhecidos dos dados em cada estágio. Na literatura esse procedimento é conhecido como filtro de Kalman adaptado, cujos estimadores dos parâmetros podem ser obtidos na forma do estimador bayesiano

A estrutura de dependência imposta para a seqüência θ_t no modelo filtro de Kalman é especificada via equação de estado. O modelo descreve agora os θ_t 's permutáveis de um estágio para outro. Além disso, na prática, a natureza exata da permuta é desconhecida e conseqüentemente G_t é desconhecido. Desde que não seja possível especificar um modelo definitivo para θ_t , é razoável assumir G_t desconhecido, mas variante no tempo.

Por hipótese ε_t e ω_t são independentes entre si e entre estágios e λ desconhecido e independente de ω_t .

Desde que λ é desconhecido, assumiu-se uma distribuição *a priori* normal de média m e variância S^2 conhecidos para o mesmo, no qual, aplicando o teorema de Bayes, o filtro de Kalman e, utilizando alguns resultados da estatística multivariada, dados por Johnson e Wichern (1992), obteve-se:

Função de verossimilhança

$$L(\lambda | y^{(i-1)}) \propto \prod_{i=1}^t (y_{i-1}^2 \tau_i + \sigma^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2(y_{i-1}^2 \tau_i + \sigma^2)} (y_i - \lambda \hat{\theta}_{i-1} y_{i-1})^2 \right\} \quad (6)$$

A distribuição *a posteriori* θ_t dado $y^{(t)}$ e λ é dada por

$$p(\theta_t | y^{(t)}, \lambda) \propto \sum_i \frac{1}{2} \exp \left\{ -\frac{1}{2 \sum_i} (\theta_t - \hat{\theta}_t)^2 \right\} \quad (7)$$

onde:

$$\hat{\theta}_t = \lambda \hat{\theta}_{t-1} + \frac{\tau_t y_{t-1}}{\sigma^2 + y_{t-1}^2 \tau_t} (y_t - y_{t-1} \lambda \hat{\theta}_{t-1})$$

$$\tau_t = \lambda^2 \sum_{t-1} + \gamma^2$$

$$\sum = \tau_t - \tau_t^2 y_{t-1}^2 (\sigma^2 + y_{t-1}^2 \tau_t)^{-1}$$

$\hat{\theta}_t$ e τ_t são funções de λ .

A *Posteriori* de λ dado $y^{(t)}$ é dada por,

$$p(\lambda | y^{(t)}) \propto \prod_{i=1}^t (y_{i-1}^2 \tau_i + \sigma^2)^{-\frac{1}{2}} \times \exp \left\{ -\frac{1}{2(y_{i-1}^2 \tau_i + \sigma^2)} (y_i - \lambda \hat{\theta}_{i-1} y_{i-1})^2 \right\} S^{-1} \times \exp \left\{ -\frac{1}{2S^2} (\lambda - m)^2 \right\} \quad (8)$$

As estimativas de $E(\theta_t | y^{(t)})$ e $E(\lambda | y^{(t)})$ são obtidas pelos estimadores MCMC, com amostras gerada pelos algoritmos Amostrador de Gibbs (Casella e Gerorg, 1992) e Metropolis-Hasting (Chib e Greemberg, 1995)

Densidade preditiva para θ_{t+1}

$$[(\theta_{t+1} | y^{(t)}) \sim N(\lambda \hat{\theta}_t, \lambda^2 \hat{\sigma}_t + \gamma^2)] \quad (9)$$

onde;

$$\hat{\theta}_t = \frac{\lambda \hat{\theta}_{t-1} \sigma^2 + \tau_t y_t y_{t-1}}{\sigma^2 + y_{t-1}^2 \tau_t}$$

$$\tau_t = \lambda^2 \sum_{t-1} + \gamma^2$$

e

$$\hat{\sum}_t = \frac{\tau_t \sigma^2}{\sigma^2 + y_{t-1}^2 \tau_t}$$

Densidade preditiva para y_{t+1}

$$(y_{t+1} | y^{(t)}, \lambda) \sim N \left[\lambda y_t \hat{\theta}_t; y_t^2 \left(\lambda^2 \hat{\sum}_t + \gamma^2 \right) + \sigma^2 \right] \quad (10)$$

A partir das condicionais *a posteriori*, aplicando os algoritmos Gibbs e Metropolis-Hasting, (Chib e Greenberg,1995), geraram-se amostras dessas séries, considerando 5 cadeias com 3000 iterações cada uma. Para garantir a independência entre os valores simulados de cada parâmetro, tomaram-se 50% das iterações, isto é, desprezaram-se os primeiros 1500 valores e, a partir dos valores considerados deu-se um salto a cada 5 valores selecionando, assim, uma amostra com 300 valores em cada cadeia, totalizando uma amostra final de 1500 valores.

Como exemplo de aplicação do modelo, são apresentados os resultados obtidos na estimação e previsão dos dados de duas séries reais da área econômica.

Primeiramente, considera-se a série correspondente ao preço de ações da Siderúrgica Vale do Rio Doce, em dólares, negociadas no período de 02 de janeiro de 1996 a 01 de fevereiro de 1999, pela BOVESPA (Bolsa de Valores de São Paulo).

Na Tabela 1, são apresentados os valores aplicados a partir de distribuições *a priori* normal para λ e θ_t para gerar amostras a partir das condicionais *a posteriori*

Tabela 1. Valores *a Priori*

m	s	λ	σ	y_0
1.0	0,5	0,65	0,3	10

As Figura s 1, 2 e 3 mostram a convergência dos parâmetros de interesse, partindo de condições iniciais arbitrárias dos dados da série.

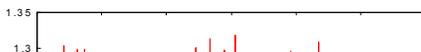


Figura 1. θ_t

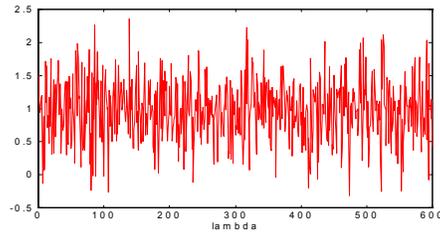


Figura 2. λ

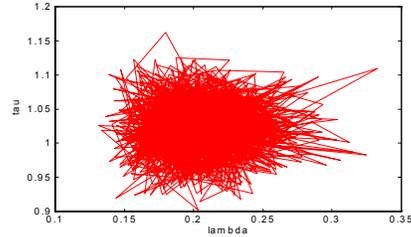


Figura 3. $(\lambda; \theta_t)$

aqui, a convergência das amostras geradas pelo algoritmo Gibbs com Metropolis-Hasting foi monitorada utilizando-se o método proposto por Gelman e Rubin (1992).

Na Figura 4, é visualizada a evolução de preços reais da série, enquanto que na Figura 5, são comparados os preços reais com os preços previstos, considerando um passo (dia) à frente, $(y_{t+k} | y^{(t+k-1)})$; $k = 1,2,...,10$, onde os últimos 10 dias foram deixados fora do conjunto de observações usadas no ajuste do modelo.

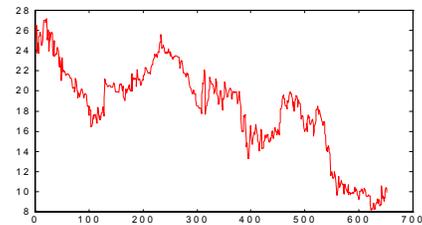


Figura 4. Evolução de preços reais



Figura 5. Comparação de preços reais e preços previstos

Na Figura 5, é possível observar que os valores previstos estão bem próximos dos valores reais, com pequeno erro de previsão, como mostra a Tabela 2.

Tabela 2. Valores previstos, um passo à frente, para a série de ações da Vale do Rio Doce, no período de 02/01/96 a 01/02/99

$t+1$	y_t	$E(\theta_{t+1} y_t)$	$E(y_{t+1} y_t)$	e_t (%)
643	9,4319	0,8932	8,4249	10,67
644	9,3710	1,0333	9,6831	3,33
645	9,5964	1,0611	10,1832	6,11
646	9,0351	0,9423	8,5138	5,76
647	9,3847	1,0443	9,8005	4,43
648	9,3234	0,9931	9,2591	0,68
649	10,0604	1,0794	10,8594	7,94
650	10,4134	1,0357	10,7853	3,57
651	10,3368	1,0288	10,6345	2,88
652	9,9297	0,9994	9,2400	6,94

O erro médio percentual de previsão é dado por:

$$\bar{e}_t = \frac{1}{N} \sum_{i=1}^N \frac{|(y_{obs} - \hat{y})|}{y_{obs}} 100 = 5,53\%$$

Mediante os resultados obtidos nessa aplicação, vale ressaltar que a dependência das variáveis (coeficiente) do modelo auto-regressivo não apresenta mudanças muito acentuadas na previsão das observações futuras.

Como uma segunda ilustração para o processo consideramos uma série que representa o preço da arroba do boi gordo, em dólares, negociadas pela Tortuga Cia Zootécnica Agrária, no período de janeiro de 1998 a agosto de 1999.

Na Tabela 3, são apresentados os valores aplicados a partir de distribuições *a priori* normal de λ e θ_t , e geraram-se amostras a partir das distribuições condicionais *a posteriori* para os parâmetros.

Tabela 3. Valores *a Priori*

m	s	λ	σ	y_0
1,0	0,3	0,12	2,5	10

As Figuras 6, 7 e 8 mostram graficamente a convergência dos parâmetros de interesse, partindo de condições iniciais arbitrárias dos dados da série.

Aqui, a convergência das amostras geradas pelo algoritmo Gibbs com Metropolis-Hasting foi também monitorada utilizando-se o método proposto por Gelman e Rubin (1992).

Na Figura 9, visualiza-se a evolução de preços reais da série enquanto que na Figura 10, visualiza-se a comparação entre os preços real e o preço previstos, considerando um passo (dia) à frente,

$(y_{t+k} | y^{(t+k-1)})$, nos últimos 10 dias do período considerado, sendo que esses não foram computados no conjunto dos dados usados para ajustar o modelo.

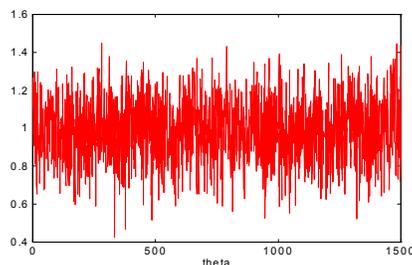


Figura 6. θ_t

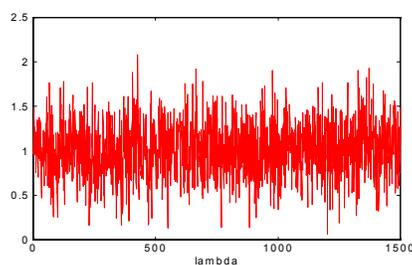


Figura 7. λ

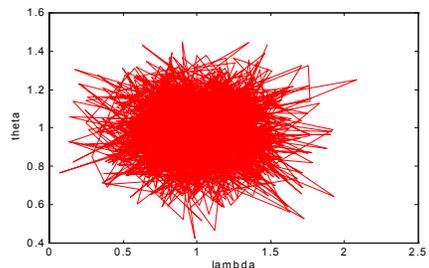


Figura 8. $(\lambda; \theta_t)$

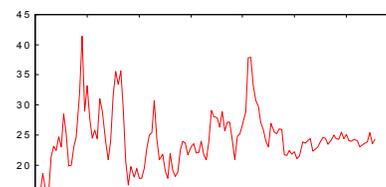


Figura 9. Evolução de preços reais

Na Figura 10, é possível observar que os valores previstos, dado o preço real do dia anterior, estão bem próximos dos valores reais com pequeno erro de previsão. Isso é comprovado pelos valores obtidos na Tabela 4.

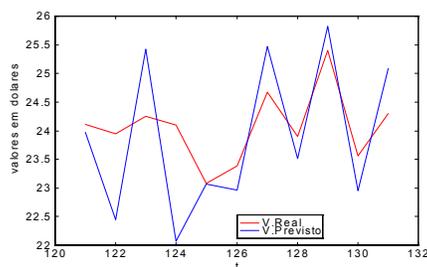


Figura 10. Comparação de preços reais e preços estimados em 10 dias

Tabela 4. Valores previstos, um passo (dia) à frente, para a série de preço de arroba de boi gordo, no período de 01/1998 a 08/1999

$t+1$	y_t	$E(\theta_{t+1} y_t)$	$E(y_{t+1} y_t)$	e_t (%)
	24,1100	0,9941	23,9665	0,59
122	23,9500	0,9372	22,4461	6,27
123	24,2500	1,0486	25,4284	4,85
124	24,1000	0,9157	22,0683	8,43
125	23,0800	0,9950	23,0676	0,05
126	23,3800	0,9824	22,9676	1,70
127	23,6800	0,9830	23,2780	2,60
128	23,9000	0,9840	23,5183	1,59
129	25,4000	1,0165	25,8195	1,65
130	23,5000	0,9743	22,9550	2,56
131	24,3000	1,0323	25,0852	3,23

Fonte: Valores de y_t fornecidos por Tortuga Cia Zootécnica Agrária

O erro médio percentual de previsão é dado por:

$$\bar{e}_t = \frac{1}{N} \sum_{i=1}^N \frac{|(y_{\text{obs}} - \hat{y})|}{y_{\text{obs}}} 100 = 3,53\%$$

Através dos resultados obtidos neste estudo de casos, podemos concluir que o modelo dinâmico é uma alternativa eficiente para modelagem de processos auto-regressivos de primeira ordem com coeficiente aleatório para fazer análise e previsões futuras de evolução de valores econômicos.

Referências

- CASELLA, G.; GEORGE, E.J. Explaining the Gibbs Sampler. Belmont. *Am. Stat. Soc.*, Alexandria, v. 46, n.3, p. 167-174, 1992.
- CHIB, S.; GREENBERG, E. J. Understanding the Metropolis-Hasting Algorithm. *Am. Stat. Soc.*, Alexandria, v. 49, n.4, p. 35-51, 1995.
- DIAZ, J. Bayesian forecasting for AR(1) models with normal coefficients. *Stat. Theor. Meth. Abstr.*, Voorburg, v.196, p.2229-2246, 1990.
- GELMAM, A.E.; RUBIN, D. Inference from iterative application using multiple sequences. *Stat. Sci.*, Hayward, v.7, p.457-511, 1992.
- JOHNSON, R.A.; WECHERN, D.W. Applied Multivariate Statistical Analysis, 3. ed. New York: Prentice Hall, 1992
- LIU, S.I. Comparison of forecastes for ARMA models between a random coeficent appro ach and a bayesian approach. *Comm. Stat. Theor. Meth.*, Madison, v.24, n. 2, p.319-333, 1995.
- MEINHOLD, R. J.; SINGPURWALLA, N. D. Understanding the Kalman Filter. *Am. Stat. Soc.*, Alexandria, v.37, n.2, p.229-246, 1990

Received on October 08, 2001.

Accepted on November 23, 2001.