**ANALISE BAYESIANA DE
ALGUNS MODELOS DE
SÉRIES TEMPORAIS**

Thelma Sáfadi

TESE APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA A OBTENÇÃO DO GRAU
DE
DOUTOR EM ESTATÍSTICA

Área de concentração: **Estatística**
Orientador: **Prof. Dr. Pedro A. Morettin**

São Paulo, Maio de 1997

ANALISE BAYESIANA DE ALGUNS MODELOS DE SÉRIES TEMPORAIS

Este exemplar corresponde à redação final
da tese devidamente corrigida,
apresentada por Thelma Sáfy e
aprovada pela Comissão Julgadora.

São Paulo, 22 de Maio de 1997

Banca Examinadora

-Prof. Dr. Pedro Alberto Morettin (Presidente)	IME-USP
-Prof. Dr. Sérgio Wechsler	IME-USP
-Prof. Dr. Basílio de Bragança Pereira	IM-UFRJ
-Prof. Dr. Jorge Alberto Achcar	ICMSC-USP
-Prof. Dr. Luiz Koodi Hotta	IMECC-UNICAMP

*Aos meus sobrinhos
Brenna e Henrique*

Agradecimentos

Mais uma etapa concluída!

Agradeço a Deus por sua presença constante trazendo tranquilidade e fazendo com que os obstáculos fossem um a um superados.

Agradeço ao Prof. Pedro A. Morettin pela excelente orientação e por ter aceito o desafio de orientar uma matemática com a pretensão de se tornar uma estatística Bayesiana.

Agradeço aos Profs. Ruben Delly Veiga e Marinho G. Andrade pelo apoio computacional nos Capítulos 4 e 5, respectivamente.

Resumo

O objetivo deste trabalho é a análise Bayesiana dos modelos autoregressivos e de médias móveis com limiares (TARMA), médias móveis assimétricas (ASMA) e autoregressivo com coeficientes aleatórios (RCA). Apesar de definido por Tong (1983), ainda não havia na literatura uma análise para o modelo TARMA. A análise para este modelo é feita através de dois procedimentos distintos. Para o primeiro, consideramos conhecidos os valores de limiares e retardo, e para o segundo fazemos uma extensão para o caso geral onde métodos de simulação como o Amostrador de Gibbs e o Algoritmo de Metropolis-Hastings são utilizados para a inferência. Para a análise do modelo ASMA, a inovação é considerada uma variável limiar. Neste caso a inferência é feita diretamente das distribuições marginais a posteriori. Os coeficientes do modelo RCA são supostos normalmente distribuídos e a análise é feita através do Amostrador de Gibbs. Para as análises envolvendo o Algoritmo de Metropolis-Hastings e o Amostrador de Gibbs, utilizamos métodos para verificação da convergência.

Abstract

The purpose of this work is to make a Bayesian analysis of some non-linear discrete time series models, namely TARMA (threshold autoregressive moving average), ASMA (asymmetric moving average) and RCA (random coefficients autoregressive) models. Eventhough defined by Tong(1983), TARMA models have not been analyzed in the literature. The analysis for this model is done through two different procedures. For the first, we considered known the values of the thresholds and delay and for the second we give an extension for the general case for which simulation methods, like the Gibbs sampler and Metropolis-Hastings are necessary for inference. For the analysis of the ASMA model the innovations are considered threshold variables and the inference is made directly from the posterior marginal distributions. The coefficients of the RCA model are assumed to be normaly distributed and the analysis carried out with the Gibbs sampler. Procedures are used to check convergence of the simulated chains using the Gibbs sampler and the Metropolis-Hastings algorithm.

Conteúdo

1	Introdução	3
2	Análise Bayesiana	7
2.1	Inferência Bayesiana	7
2.2	Amostrador de Gibbs	10
2.3	Algoritmo de Metropolis-Hastings	14
3	Alguns Modelos de Séries Temporais	19
3.1	Introdução	19
3.2	Modelo TARMA	20
3.3	Modelo ASMA	21
3.4	Modelo RCA	24
4	Modelo TARMA	26
4.1	Análise considerando uma priori própria	26
4.2	Análise a posteriori com priori imprópria de Jeffreys	35
4.3	Distribuição preditiva	36
4.4	Aplicações	42

5	Modelo TARMA Rearranjado	48
5.1	Descrição do Modelo	48
5.2	Análise a posteriori	51
5.3	Previsão	55
5.4	Aplicação	56
6	Modelo ASMA	64
6.1	Análise do Modelo ASMA	64
6.2	Distribuição preditiva	69
6.3	Aplicações	72
7	Modelo RCA	77
7.1	Análise a posteriori	77
7.2	Densidade preditiva	81
7.3	Aplicação	82
	Discussões e Conclusões	86
	Bibliografia	88
	Apêndice 1: Algumas Distribuições	92
	Apêndice 2: Séries de dados reais	97

Capítulo 1

Introdução

A análise estatística de séries temporais tem aplicações em diversas áreas. A literatura é ampla e na sua maior parte envolve métodos não-Bayesianos. Algumas referências de tratamento Bayesiano de séries temporais são Zellner(1971), Broemeling(1985), West e Harrison(1989) e Pole, West e Harrison(1994) .

Apesar de Jeffreys(1939) ter feito um estudo da teoria espectral, foi apenas a partir de 1980 que as técnicas Bayesianas tornaram-se uma alternativa atrativa para a metodologia de Box-Jenkins. Monahan(1983) dá início a essa tendência ao usar integração numérica para implementar uma análise Bayesiana completa para modelos de séries temporais envolvendo identificação, estimação e previsão.

Com o avanço das técnicas computacionais, os modelos não-lineares de séries temporais têm despertado grande interesse nos últimos anos. Apesar disso, a literatura para métodos Bayesianos aplicados aos modelos de séries temporais é ainda muito limitada. Os métodos são desenvolvidos prin-

principalmente para modelos autoregressivos (AR), já que estes são facilmente analisados com distribuições normal-gama.

O desafio durante a década de 80 era encontrar uma análise puramente analítica para modelos de médias móveis (MA) e mistos.

Para o modelo MA, a função de verossimilhança não é linear nos coeficientes do processo, o que forçaria o uso de um procedimento numérico, como feito por Monahan, ou uma aproximação analítica da função de verossimilhança, como feito por Newbold (1973).

Broemeling e Shaarawy (1984, 1986) e Shaarawy e Broemeling (1985) propuseram uma aproximação da função de verossimilhança para a análise de modelos de MA e ARMA (autoregressivo e de médias móveis) por distribuições normal-gama combinando assim facilmente a priori com a verossimilhança.

Assim como para modelos AR, a literatura está repleta de artigos com análise Bayesiana de modelos autoregressivos com limiares ("Threshold Autoregressive", TAR). O modelo TAR é detalhadamente descrito em Tong (1983, 1990), incluindo aplicações físicas relacionadas a ciclos limites.

Levando-se em consideração o fato do ajustamento de um modelo ARMA poder ter um número menor de parâmetros do que de um modelo puro AR ou MA, consideramos aqui a análise Bayesiana para o modelo Autoregressivo e Médias Móveis com limiares ("Threshold Autoregressive Moving Average" TARMA). Embora tenha sido definido por Tong (1983), o modelo TARMA ainda não havia sido estudado, seja com enfoque clássico ou Bayesiano. Neste trabalho nos propusemos a analisá-lo com a metodologia Bayesiana.

A análise Bayesiana para o modelo TARMA é desenvolvida através de

dois procedimentos diferentes. O primeiro é uma extensão dos trabalhos de Broemeling and Shaarawy(1988) e Broemeling and Cook(1992) para a análise Bayesiana dos modelos ARMA e TAR, respectivamente. No segundo, estendemos o trabalho de Chen and Lee(1995) para o modelo TAR, considerando o modelo TARMA rearranjado. A inferência para os parâmetros foi obtida através do Amostrador de Gibbs e do Algoritmo de Metropolis-Hastings.

Um outro modelo aqui desenvolvido envolvendo uma variável limiar e por nós considerado é o modelo de Médias Móveis Assimétricas (“Asymmetric Moving Average”, ASMA). Para o modelo ASMA, a assimetria percebida é uma reação diferente para o retardo da inovação dependendo do sinal desta, ou seja, poderíamos dizer que a série temporal reage com uma de duas regras diferentes conforme a inovação seja positiva ou negativa.

Os modelos lineares com coeficientes aleatórios têm sido aplicados em diversas áreas. Vamos considerar também o modelo autoregressivo com coeficientes aleatórios (“Random Coefficients Autoregressive”, RCA), como definido por Diaz(1990). A inferência para os parâmetros e a previsão 1-passo à frente é feita através do Amostrador de Gibbs.

O Capítulo 2 é destinado às ferramentas utilizadas neste trabalho, ou seja, apresentamos a inferência Bayesiana assim como alguns métodos de simulação e controle de convergência. No Capítulo 3 descrevemos os modelos de Séries Temporais que serão analisados. Os Capítulos 4 e 5 destinam-se às análises dos modelos TARMA e TARMA Rearranjado, respectivamente. No Capítulo 6 analisamos o modelo ASMA e a análise do modelo RCA é encontrada no Capítulo 7. Finalizamos cada Capítulo com aplicações onde os dados são simulados ou reais.

No Apêndice 1 apresentamos as distribuições univariadas e multivariadas que serão utilizadas neste trabalho e no Apêndice 2 podem ser encontradas as séries de dados reais utilizadas nas aplicações.

Para as aplicações utilizamos o “software” MATLAB versão 4.2.

Capítulo 2

Análise Bayesiana

2.1 Inferência Bayesiana

O problema fundamental da Estatística é fazer inferência. A inferência Bayesiana trabalha na presença de observações y cujos valores inicialmente incertos são descritos através de uma distribuição de probabilidades com densidade ou função de probabilidade $f(y|\theta)$.

O indexador θ é uma quantidade de interesse que se supõe possuir um sentido bastante preciso dentro do problema em estudo. Além disso, o pesquisador sempre possui uma informação prévia sobre esta quantidade θ . É recomendável que este conhecimento seja incorporado à análise.

A teoria Bayesiana julga que a descrição da incerteza deve ser feita por meio de probabilidades e por essa razão descreve probabilisticamente toda quantidade desconhecida.

Vemos, então, que enquanto a metodologia frequentista não tolera informação alguma que não seja julgada proveniente da observação de um ex-

perimento “repetível”, o Bayesiano incorpora a opinião sobre θ à análise através de uma densidade $P(\theta)$, denominada densidade *a priori*.

Podemos dizer que a diferença formal entre a inferência Bayesiana e a frequentista é que para a primeira os parâmetros θ são variáveis aleatórias e, portanto, possuem uma distribuição, enquanto que para a segunda os parâmetros são valores fixos e não é permitido tratá-los como aleatórios ou associar a eles probabilidades.

No que segue, o valor da distribuição condicional $f(y|\theta)$, será considerado como função de θ , dada a amostra y efetivamente observada, representada por $\mathcal{L}(\theta|y)$ e denominada *função de verossimilhança*. A densidade *a priori* $P(\theta)$ possui este nome por ser a distribuição de probabilidade de θ antes que se observe Y . O Cálculo de Probabilidades impõe que qualquer inferência sobre θ seja baseada na sua distribuição de probabilidade depois que os dados forem observados. Esta distribuição, $P(\theta|y)$, é denominada distribuição *a posteriori* e é obtida através do Teorema de Bayes. Isto é,

$$P(\theta|y) = \frac{\mathcal{L}(\theta|y)P(\theta)}{\int \mathcal{L}(\theta|y)P(\theta)d\theta}.$$

Como o denominador da expressão acima não depende de θ , o Teorema de Bayes pode ser reescrito como

$$P(\theta|y) \propto \mathcal{L}(\theta|y)P(\theta), \quad (2.1)$$

onde o símbolo $\propto$ significa proporcionalidade. Podemos pensar no Teorema de Bayes como um mecanismo de atualização da opinião do estatístico sobre θ .

A expressão das distribuições a priori e a atribuição de valores aos hiperparâmetros, isto é, aos parâmetros destas prioris é um problema pertinente na abordagem Bayesiana. Eventualmente, ocorrem situações onde exista muito pouca ou nenhuma informação disponível a priori, ou mesmo que se deseje deliberadamente omitir totalmente a opinião do cientista. Nestes casos, a distribuição a priori considerada é do tipo não-informativa. A ignorância ou não informação sobre θ , sugere que todos os possíveis valores de θ são olhados como iguais.

Existe entre os estatísticos muita controvérsia sobre a especificação deste tipo de priori, parte da divergência devendo-se ao fato que muitas vezes esta especificação leva a distribuições impróprias, opondo-se ao cálculo de probabilidades. A priori de Jeffreys dada por

$$P(\theta) \propto |I(\theta)|^{1/2}, \quad (2.2)$$

onde $I(\theta) = E(-\frac{\partial^2 \log \mathcal{L}(\theta|y)}{\partial \theta^2} | \theta)$ é a medida de informação esperada de Fisher, é, entre as prioris não informativas, uma das mais usadas.

Um outro elemento importante na inferência Bayesiana é a distribuição *preditiva*. Usualmente estamos interessados em fazer inferência sobre alguma observação futura, w , utilizando as observações passadas y . Tal inferência deve ser baseada na distribuição a posteriori de w , $P(w|y)$. Observemos que $P(w|y)$ pode ser obtida da posteriori conjunta de w e θ por integração, ou seja,

$$P(w|y) = \int P(w, \theta|y) d\theta = \int P(w|\theta, y) P(\theta|y) d\theta, \quad (2.3)$$

onde $P(w|\theta, y)$ em muitos modelos se reduz a $P(w|\theta)$. Conforme veremos posteriormente, na maioria das vezes não conseguimos obter analiticamente

da distribuição a posteriori conjunta as distribuições marginais de interesse. Este problema tem sido solucionado com os métodos de simulação. A seguir descreveremos os que serão utilizados neste trabalho.

2.2 Amostrador de Gibbs

O Amostrador de Gibbs é uma técnica para gerar amostras aleatórias de uma distribuição (marginal) sem que se conheça a sua densidade. Embora o Amostrador de Gibbs(A.G.) tenha sido muito aplicado em modelos Bayesianos, ele também é extremamente útil para se amostrar a função de verossimilhança nos métodos frequentistas.

O A.G. ganhou popularidade nos últimos anos a partir do trabalho de Geman e Geman(1984), no contexto de processamento de imagens. Entretanto, o método já havia sido descrito por Metropolis et al(1953) e posteriormente, por Hastings(1970). Recentemente, Gelfand e Smith(1990) despertaram novos interesses na aplicação do método ao mostrarem sua aplicação a uma grande variedade de problemas da estatística convencional.

Em modelos complicados, raramente se consegue obter amostras diretamente das distribuições a posteriori. A idéia do método de Monte Carlo via Cadeias de Markov(MCMC) é simular um passeio aleatório no espaço do parâmetro θ , o qual converge para uma distribuição estacionária, que é a distribuição a posteriori $P(\theta|Y)$, onde Y é o vetor de observações.

O A.G. é um particular algoritmo MCMC e tem sido extremamente útil

na resolução de problemas multidimensionais e é definido em termos de subvetores de θ .

Pode-se pensar no A.G. como uma implementação prática do fato que o conhecimento das distribuições condicionais é suficiente para determinar a distribuição conjunta (se ela existir). Consideremos, por exemplo, o caso bivariado. Sejam X e Y variáveis aleatórias e suponhamos conhecidas as densidades condicionais $X|Y$ e $Y|X$. Observando que a distribuição conjunta (desconhecida) pode ser dada por $f_{X,Y}(x,y) = f_{X|Y}(x|y)f_Y(y)$ obtemos a distribuição marginal de X , ou seja,

$$\begin{aligned} f_X(x) &= \int f_{X,Y}(x,y)dy = \int f_{X|Y}(x|y)f_Y(y)dy \\ &= \int f_{X|Y}(x|y) \int f_{Y|X}(y|t)f_X(t)dt dy \\ &= \int h(x,t)f_X(t)dt, \end{aligned}$$

onde $h(x,t) = \int f_{X|Y}(x|y)f_{Y|X}(y|t)dy$. Maiores detalhes podem ser encontrados em Casella e George(1992).

Suponhamos que o vetor de parâmetros θ seja dividido em k subvetores, $(\theta_1, \theta_2, \dots, \theta_k)'$ e que as distribuições condicionais de cada parâmetro θ_i dado todos os outros, sejam conhecidas. Essas distribuições são denotadas por $f_1(\theta_1|\theta_2, \dots, \theta_k, Y), f_2(\theta_2|\theta_1, \theta_3, \dots, \theta_k, Y), \dots, f_k(\theta_k|\theta_1, \dots, \theta_{k-1}, Y)$.

Dados os valores iniciais $\theta^0 = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)})'$ retiramos $\theta_1^{(1)}$ de $f_1(\theta_1|\theta_2^{(0)}, \theta_3^{(0)}, \dots, \theta_k^{(0)}, Y)$, $\theta_2^{(1)}$ de $f_2(\theta_2|\theta_1^{(1)}, \theta_3^{(0)}, \dots, \theta_k^{(0)}, Y)$,

$\theta_k^{(1)}$ de $f_k(\theta_k|\theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_{k-1}^{(1)}, Y)$,

obtendo-se na primeira iteração $\theta^1 = (\theta_1^{(1)}, \dots, \theta_k^{(1)})'$. Após um número grande de iterações, digamos M , obtém-se $\theta^M = (\theta_1^{(M)}, \dots, \theta_k^{(M)})'$. As distribuições marginais a posteriori desejadas podem ser aproximadas pelas distribuições marginais dos N valores $(\theta_1^{(i)}, \dots, \theta_k^{(i)})_{i=M+1}^{M+N}$.

Um problema encontrado é a escolha apropriada dos valores de M e N . Uma solução geral para a escolha de N é monitorarmos a convergência da sequência de Gibbs sob algum aspecto. Cowles e Carlin(1996) apresentam um estudo comparativo entre vários métodos para monitoração da convergência, alguns com uma única cadeia, outros com cadeias múltiplas e a monitoração é feita com métodos gráficos, quantitativos ou qualitativos. Através de dois modelos simples mostram que todos os métodos podem falhar em detectar a convergência.

Para a monitoração da convergência veremos dois critérios simples.

Monitoração Informal da Convergência

Gelfand e Smith(1990) sugeriram técnicas gráficas para a verificação da convergência.

Após um número suficientemente grande N de iterações em m cadeias paralelas, forma-se uma amostra de θ e pode-se construir um histograma de qualquer uma de suas componentes. O mesmo procedimento pode ser repetido após $N + k$ iterações. Não havendo diferença perceptível a olho nu entre

os gráficos obtidos após N e $N+k$ iterações. concluimos pela convergência da cadeia. O valor de k não pode ser muito pequeno, pois a correlação inerente à cadeia de Markov estará exercendo sua influência e não se poderá dizer que a similaridade é devida à convergência ou à correlação do processo. Valores de k muito grandes são desnecessários, pois se há suspeita de convergência após N iterações, não há necessidade de ir muito além na cadeia apenas para uma verificação. Valores para k entre 10 e 50 são apropriados.

A monitoração gráfica deve ser usada com cautela pois as técnicas gráficas podem ser ilusórias, indicando uma constância que pode não ser tão evidente sob outra escala. Além disso, muitas cadeias podem exibir um comportamento similar ao de convergência, sem que esta tenha sido atingida.

Verificação Formal da Convergência

Uma técnica simples é proposta por Gelman e Rubin(1992), a qual consiste em considerarmos várias cadeias em paralelo partindo de valores iniciais distintos e compararmos, para cada parâmetro (escalar) de interesse, a variabilidade dentro e entre as cadeias amostradas. Denotemos por θ cada um desses parâmetros. Uma vez que a estacionaridade tenha sido atingida, digamos na i -ésima iteração, consideramos as realizações $\theta_j, \theta_{j+h}, \dots, \theta_{j+Nh}, j \geq i$ como uma amostra aleatória da distribuição desejada.

A convergência é monitorada através do fator $\hat{R}$, introduzido por Gelman e Rubin(1992). Tomemos $n \geq 2$ cadeias de comprimento N . No caso de multimodalidade é recomendável inicializar pelo menos uma cadeia a partir de cada moda. Calculamos a variabilidade *dentro* das cadeias, W , e a média

B das N variâncias *entre* as cadeias,

$$W = \frac{1}{n} \sum_{j=1}^n s_j^2,$$

$$B = \frac{N}{n-1} \sum_{j=1}^n (\bar{\theta}_{\cdot j} - \bar{\theta}_{\cdot\cdot})^2$$

onde $\bar{\theta}_{\cdot j} = \frac{1}{N} \sum_{i=1}^N \theta_{ij}$ e $s_j^2 = \frac{1}{N-1} \sum_{i=1}^N (\theta_{ij} - \bar{\theta}_{\cdot j})^2$.

Sob convergência, todos os nN valores são gerados da posteriori e a variância da distribuição pode ser estimada de forma não-viciada por

$$\hat{\sigma}^2 = \frac{N-1}{N} W + \frac{1}{N} B.$$

Definimos $\sqrt{\hat{R}} = \sqrt{\frac{\hat{\sigma}^2}{W}}$. Se o valor de $\sqrt{\hat{R}}$ é grande, devemos considerar mais iterações para obtermos uma melhor estimativa dos parâmetros. Quando $\sqrt{\hat{R}} \approx 1$ a convergência ocorre e a amostra selecionada das iterações é i.i.d.

2.3 Algoritmo de Metropolis-Hastings

O Amostrador de Gibbs seleciona amostras aleatórias de todas as densidades condicionais obtidas da distribuição a posteriori conjunta $P(\theta|Y)$. Se estivermos considerando distribuições a priori conjugadas, as distribuições condicionais são, em geral, da forma Gama, Normal, Beta, Poisson, etc e as amostras geradas dessas distribuições são facilmente obtidas, pois estas distribuições estão disponíveis em muitos “software”.

Quando estas distribuições não são identificáveis ou se a geração não-iterativa for muito complicada ou custosa, usamos o algoritmo de Metropolis-Hastings(M.H.).

Seja $P(\theta|Y)$ uma distribuição não identificada. O algoritmo de M.H. cria sequências de pontos aleatórios $(\theta^1, \theta^2, \dots)$ cuja distribuição converge para $P(\theta|Y)$. Cada sequência pode ser considerada como um passeio aleatório com distribuição estacionária $P(\theta|Y)$. Denotemos $P(\theta|Y)$ por $P(\theta)$.

O algoritmo de M.H. é especificado por uma densidade $q(\theta, \beta)$, denominada núcleo de transição. Esta densidade é selecionada de uma família de distribuições que requer a especificação de algum “tuning” do parâmetro, tal como escala ou posição. O núcleo de transição, $q(\theta, \beta)$ representa a probabilidade de movimento de θ para β . Geralmente o processo se move de θ para β com mais frequência do que de β para θ , contradizendo a condição de reversibilidade

$$P(\theta)q(\theta, \beta) = P(\beta)q(\beta, \theta).$$

Este fato pode ser corrigido, se diminuirmos o número de movimentos de θ para β , e para isto introduzimos uma probabilidade de movimento p .

O algoritmo de M.H. é como segue:

- (1) Inicialize com um valor(arbitrário) $\theta^{(0)}$, $j = 0$,
- (2) Gere β de $q(\theta^{(j)}, \cdot)$ e u de uma Uniforme(0,1),
- (3) Seja $p = \min\{1, \frac{P(\beta)q(\theta^{(j)}, \beta)}{P(\theta^{(j)})q(\beta, \theta^{(j)})}\}$,

(4) Se $u \leq p$; faça $\theta^{(j+1)} = \beta$, senão faça $\theta^{(j+1)} = \theta^{(j)}$,

(5) Repita (2) e (3) até que a distribuição estacionária tenha sido obtida.

O núcleo de transição q apenas define uma proposta de movimento, a qual deve (ou não) ser confirmada por p .

Para que o algoritmo de M.H. seja implementado necessitamos de uma densidade q adequada. A escolha dessas densidades é objeto de estudos recentes. Embora estes estudos estejam longe de estarem completos, já se conhece o suficiente para a condução de simulações.

Uma família de densidades aparece no trabalho de Metropolis et al.(1953) e é dada por $q(\theta, \beta) = q_1(\beta - \theta)$, onde $q_1(\cdot)$ é uma densidade multivariada. O candidato β é obtido de $\beta = \theta + \alpha$, onde α é uma variável aleatória incremento com distribuição q_1 . Sendo β igual ao valor atual θ mais um ruído, esta cadeia é denominada passeio aleatório. A densidade q_1 pode ser normal multivariada ou t-multivariada, com a escolha adequada dos parâmetros.

Uma segunda família é dada por $q(\theta, \beta) = q_2(\beta)$. O candidato β é obtido independente do valor atual θ . Para q_2 podemos escolher uma densidade normal multivariada ou uma t-multivariada sem que a escolha dos parâmetros seja necessária.

Uma terceira escolha, quando possível, é explorar a forma de $P(\cdot)$ para obter a densidade $q(\cdot)$ (veja Chib e Greenberg(1994)). Por exemplo, se $P(\theta)$ for dada por $P(\theta) \propto \Psi(\theta)h(\theta)$, onde $h(\theta)$ é uma densidade que pode ser amostrada e $\Psi(\theta)$ é uniformemente limitada, fazemos $q(\theta, \beta) = h(\theta)$.

Outras possíveis famílias podem ser encontradas em Chib e Greenberg(1995).

Um outro fator crítico na eficiência do algoritmo é a escolha do fator escala ou dispersão da densidade q . O fator escala da densidade q afeta o comportamento da cadeia em pelo menos duas dimensões, uma é a taxa de aceitação (percentagem de vezes que um movimento é aceito) e o outro é a região do espaço amostral que é coberta pela cadeia. Para vermos isto, consideremos uma situação na qual a cadeia convergiu e a densidade está sendo amostrada em torno da moda. Se a dispersão da densidade q é muito grande, alguns pontos de β estarão longe dos valores atuais θ e por isto terão baixa probabilidade de serem aceitos. Por outro lado se o fator escala for muito pequeno a cadeia não cobrirá o suporte da densidade e regiões de baixa probabilidade serão amostradas(“undersampling”).

Quando a densidade q escolhida é o passeio aleatório q_1 , Roberts, Gelman e Gilks(1994) mostram que considerando-se $P(\cdot)$ e $q_1(\cdot)$ normais, o fator escala escolhido para q_1 deve ser tal que a taxa de aceitação seja aproximadamente 45% nos problemas unidimensionais, aproximadamente 23% para dimensões excessivamente grandes e que, para dimensões menores que 6 a taxa de aceitação deve estar em torno de 25%.

Se a densidade q_2 é considerada, Chib e Geweke(trabalho em andamento) mostram que é importante que as caudas da densidade q_1 dominem as da densidade $P(\cdot)$.

As técnicas de verificação de convergência descritas para o A.G. também podem ser utilizadas aqui.

O algoritmo de M.H. pode ser executado dentro do ciclo do A.G., ou seja, para aquelas distribuições não identificadas ou difíceis de serem geradas, aplicamos o algoritmo de M.H. para obtermos, por exemplo, $\theta_i^{(j)}$ e a seguir

continuarmos o ciclo do A.G.

Maiores detalhes sobre o algoritmo de M.H. podem ser encontrados em Casella e George(1992), Chib e Greenberg(1995) e Gamerman(1996).

Capítulo 3

Alguns Modelos de Séries Temporais

3.1 Introdução

Nos últimos anos muito se tem estudado sobre os modelos não lineares, principalmente com o progresso da área computacional, facilitando a análise desses modelos.

A idéia de modelar uma série temporal discreta usando modelos lineares por partes, foi enunciada primeiramente por Tong(1977). Uma apresentação compreensível, juntamente com inúmeras aplicações e discussões foi considerada por Tong e Lim (1980).

O uso de variáveis e valores limiares nos permite modelar uma série temporal em modelos lineares por partes. Entre os modelos com limiares vamos descrever os modelos TARMA e ASMA.

Além dos modelos com limiares vamos descrever o modelo RCA. Os mo-

delos lineares com coeficientes aleatórios têm sido aplicados em várias áreas. Por exemplo, os modelos de regressão com coeficientes aleatórios são encontrados em economia e sociologia, em aplicações onde é razoável supor que os parâmetros mudem com o tempo.

3.2 Modelo TARMA

Seja $\{Y_t, t = 0, \pm 1, \dots\}$ um processo estocástico discreto satisfazendo:

$$Y_t = \phi_{l0} + \sum_{i=1}^{p_l} \phi_{li} Y_{t-i} + a_t^{(l)} + \sum_{j=1}^{q_l} \theta_{lj} a_{t-j}^{(l)}, \quad \text{se } Y_{t-d} \in R_l \quad (3.1)$$

para $l = 1, \dots, k$, com

i) $\{a_t^{(l)}\}$ é uma sequência de variáveis aleatórias i.i.d. $\sim N(0, \tau_l^{-1})$, onde a precisão $\tau_l > 0, \tau_l^{-1} = \sigma_l^2 = \text{var}(a_t^{(l)})$ é desconhecida, para cada l ;

ii) R_l é um intervalo da reta $\mathbb{R}$ tal que $R_i \cap R_j = \emptyset$, se $i \neq j$ e $\cup_{l=1}^k R_l = \mathbb{R}$, isto é, $\{R_1, \dots, R_k\}$ formam uma partição da reta com $k-1$ variáveis limiares;

iii) d é um parâmetro de retardo e assumiremos $d \in \{1, \dots, d_0\}$.

Então (3.1) nos diz que Y_t segue k diferentes processos ARMA com diferentes ordens (p_l, q_l) , $l = 1, \dots, k$, dependendo do valor da variável limiar Y_{t-d} . Nós denotaremos tal processo por $TARMA(k; p_1, \dots, p_k, q_1, \dots, q_k)$, um

processo ARMA com k regimes. Neste trabalho consideramos um modelo TARMA com 2 regimes dado por:

$$Y_t = \begin{cases} \phi_{10} + \sum_{i=1}^{p_1} \phi_{1i} Y_{t-i} + a_t^{(1)} + \sum_{j=1}^{q_1} \theta_{1j} a_{t-j}^{(1)}, & \text{se } Y_{t-d} \leq r \\ \phi_{20} + \sum_{i=1}^{p_2} \phi_{2i} Y_{t-i} + a_t^{(2)} + \sum_{j=1}^{q_2} \theta_{2j} a_{t-j}^{(2)}, & \text{se } Y_{t-d} > r \end{cases} \quad (3.2)$$

onde r é o parâmetro de limiar.

Neste caso os parâmetros do modelo são $(\tau_1, \tau_2, \gamma_1, \gamma_2, r, d)$ onde $\gamma_i = (\phi_{i0}, \phi_{i1}, \dots, \phi_{ip_i}, \theta_{i1}, \dots, \theta_{iq_i})$, $i = 1, 2$. Assumiremos também que as ordens p_1, p_2, q_1, q_2 , são conhecidas. Caso contrário, podem ser determinadas através de algum critério de informação, como AIC ou BIC.

3.3 Modelo ASMA

A noção de séries temporais assimétricas, além do interesse econômico, tem importantes implicações para a análise de séries temporais por apresentarem propriedades de retardo assimétrico não encontradas em modelos como AR(p), MA(q), ARMA(p,q) e BL(p,q,P,Q) (bilinear).

Se a inovação for uma sequência de variáveis aleatórias com função de distribuição assimétrica (como por exemplo lognormal), os modelos acima podem descrever séries temporais com distribuições marginais ou condicionais assimétricas. Mas este não é o fenômeno assimétrico observado por Wecker(1981) (e inicialmente por Stigler e Kindahl(1970)). A assimetria percebida é uma reação diferente para o retardo da inovação dependendo do sinal desta. Ou seja, poderíamos dizer que a série temporal reage com uma de duas regras diferentes conforme a inovação seja positiva ou negativa.

O modelo de médias móveis assimétricas de ordem q , denotado por $ASMA(q)$, é dado por

$$Y_t = a_t + \sum_{i=1}^q \theta_i^- a_{t-i} + \sum_{i=1}^q (\theta_i^+ - \theta_i^-) \mathbf{I}(a_{t-i} \geq 0) a_{t-i}, \quad (3.3)$$

onde $\{a_t\}$ é sequência i.i.d. $\sim (0, \tau^{-1})$, $\tau > 0$.

Para $q = 1$, o modelo $ASMA(1)$ pode ser escrito como

$$Y_t = a_t + \theta_1^+ a_{t-1}^+ + \theta_1^- a_{t-1}^-$$

onde $a_t^+ = \max\{0, a_t\}$ e $a_t^- = \min\{0, a_t\}$.

Ou ainda, o modelo $ASMA(1)$ pode ser dado em regimes, onde a_{t-1} é a variável limiar, i.e.

$$Y_t = \begin{cases} a_t + \theta_1^+ a_{t-1}, & \text{se } a_{t-1} \geq 0, \\ a_t + \theta_1^- a_{t-1}, & \text{se } a_{t-1} < 0. \end{cases} \quad (3.4)$$

Por simplicidade vamos supor $\{a_t\}$ i.i.d. $\sim N(0, \tau^{-1})$.

Observamos que enquanto para o modelo $MA(1)$, a média de Y_t é zero, para o modelo (3.4) ela é função dos parâmetros θ_1^+ e θ_1^- e em geral é diferente de zero.

De Wecker(1981) temos para o modelo $ASMA(1)$:

$$\begin{aligned} E(Y_t) &= \mu = \frac{\theta_1^+ - \theta_1^-}{\sqrt{2\pi}} \tau^{-1/2}, \\ Var(Y_t) &= \gamma_0 = (1 + \frac{\theta_1^{+2} + \theta_1^{-2}}{2}) \tau^{-1} - \mu^2, \\ \gamma_1 &= (\frac{\theta_1^+ + \theta_1^-}{2}) \tau^{-1} \end{aligned}$$

e

$$\gamma_k = 0; k \geq 2.$$

Quando o modelo é $MA(1)$ temos $\theta_1^+ = \theta_1^-$ e, portanto, $\mu = 0$, $Var(Y_t) = \tau^{-1}(1 + \theta^2)$, $\gamma_1 = \theta\tau^{-1}$ e $\gamma_k = 0$ para $k > 1$.

Assim temos que os modelos simétricos não são diferenciados dos assimétricos se olharmos para a autocovariância, já que para $MA(q)$ e $ASMA(q)$, $\gamma_k = 0$ para $k > q$. Observemos ainda que se os parâmetros do modelo $ASMA(1)$ são tais que $\theta_1^+ = -\theta_1^-$, então é impossível distinguirmos, baseados apenas na autocovariância, o modelo $ASMA(1)$ de uma sequência i.i.d. Maiores detalhes assim como um teste para assimetria poderão ser encontrados em Wecker(1981).

Temos também que o modelo (3.4) é invertível se $|\theta_1^+| < 1$ e $|\theta_1^-| < 1$. Condições similares para modelos mais complicados, como $ASMA(q)$, não têm sido obtidas na prática. Uma regra prática pode ser encontrada em Tong(1990).

O modelo $ASMA(q)$ pode ser representado em 2^q regimes. Por exemplo, para o modelo $ASMA(2)$ temos:

$$Y_t = \begin{cases} a_t + \theta_1^+ a_{t-1} + \theta_2^+ a_{t-2}, & \text{se } a_{t-1} \geq 0 \text{ e } a_{t-2} \geq 0 \\ a_t + \theta_1^+ a_{t-1} + \theta_2^- a_{t-2}, & \text{se } a_{t-1} \geq 0 \text{ e } a_{t-2} < 0 \\ a_t + \theta_1^- a_{t-1} + \theta_2^+ a_{t-2}, & \text{se } a_{t-1} < 0 \text{ e } a_{t-2} \geq 0 \\ a_t + \theta_1^- a_{t-1} + \theta_2^- a_{t-2}, & \text{se } a_{t-1} < 0 \text{ e } a_{t-2} < 0 \end{cases} \quad (3.5)$$

De Wecker(1981) temos ainda que, para o modelo $ASMA(2)$,

$$\mu = \frac{(\theta_1^+ + \theta_2^+) - (\theta_1^- + \theta_2^-)}{\sqrt{2\pi}} \tau^{-1/2},$$

$$\gamma_0 = \frac{1 + (\theta_1^+)^2 + (\theta_2^+)^2}{2} \tau^{-1} + \frac{1 + (\theta_1^-)^2 + (\theta_2^-)^2}{2} \tau^{-1} - \mu^2,$$

$$\gamma_k = 0; k > 2.$$

3.4 Modelo RCA

Uma série temporal $\{Y_t\}$ segue um modelo autoregressivo com coeficientes aleatórios de ordem k , RCA(k), se Y_t satisfizer a equação da forma

$$Y_t = \sum_{i=1}^k (\beta_i + B_i(t)) Y_{t-i} + e_t, \quad (3.6)$$

onde

- (1) $\{e_t\}$ é uma sequência de v.a. i.i.d. com média zero e precisão τ ,
- (2) os β_i são constantes, $i = 1, \dots, k$,
- (3) fazendo $B(t) = (B_k(t), \dots, B_1(t))$, $\{B(t)\}$ é uma sequência de vetores independentes $(1 \times k)$ com média zero e $E(B'(t)B(t)) = C$,
- (4) $B(t)$ e e_t são independentes.

Consideremos o caso particular, RCA(1), onde o parâmetro autoregressivo θ_t segue um modelo da forma $\theta_t = \mu + \epsilon_t$. Esse modelo foi analisado por Diaz(1990). Em seu trabalho, Diaz tem como objetivo determinar a densidade preditiva de observações futuras. Neste trabalho vamos fazer inferência sobre

os parâmetros através do Amostrador de Gibbs e fazer previsão 1-passo à frente.

O modelo RCA(1) considerado é dado por

$$\begin{aligned} Y_t &= \theta_t Y_{t-1} + e_t, \\ \theta_t &= \mu + \epsilon_t, \end{aligned} \tag{3.7}$$

onde e_t e ϵ_t são independentes e normalmente distribuídos com médias iguais a zero e precisões τ e $\xi\tau$, respectivamente, $\tau > 0, \xi > 0$. Então $e_t \sim N(0, \tau^{-1})$ e $\epsilon_t \sim N(0, (\xi\tau)^{-1})$. Vamos assumir e_t independente de ϵ_t e o parâmetro μ real.

Capítulo 4

Modelo TARMA

Este Capítulo é dedicado à análise do modelo TARMA descrito em (3.1). Vamos considerar uma priori própria para os parâmetros (seção 4.1) e a priori imprópria de Jeffreys (seção 4.2). Na seção 4.3 vamos determinar a distribuição preditiva para previsão 1-passo à frente. Finalmente na seção 4.4 faremos uma aplicação considerando a série “Canadian Lynx”.

Consideremos o modelo $TARMA(2; p_1, p_2, q_1, q_2)$ definido em (3.2). Vamos supor inicialmente que $\tau = \tau_1 = \tau_2$ seja comum aos dois regimes, $d = 1$, $\gamma = (\gamma_1, \gamma_2)'$ e r é independente de γ e τ .

4.1 Análise considerando uma priori própria

Seja $P(\gamma, \tau, r)$ uma priori própria dada por

$$P(\gamma, \tau, r) = P_1(\gamma|\tau)P_2(\tau)P_3(r) , \quad (4.1)$$

onde

$$\begin{aligned}\gamma|\tau &\sim \text{Normal}(\mu, \tau^{-1}\mathbf{Q}^{-1}), \\ \tau &\sim \text{Gama}(\alpha, \beta), \\ r &\sim \text{Uniforme}\end{aligned}$$

e a distribuição de r é uniforme entre o mínimo e o máximo dos valores amostrais observados.

A priori (4.1) pode ser escrita como

$$P(\gamma, \tau, r) \propto \tau^{\alpha-1} \exp\left\{-\frac{\tau}{2}\{2\beta + (\gamma - \mu)' \mathbf{Q}(\gamma - \mu)\}\right\}. \quad (4.2)$$

Dada uma amostra $Y = \{Y_1, \dots, Y_n\}$, sejam $p = \max(p_1, p_2)$ e $q = \max(q_1, q_2)$. Condicionando nas p primeiras observações e assumindo $a_1 = a_2 = \dots = a_p = a_0 = \dots = a_{p-q-1} = 0$, com $q > p + 1$, a função de verossimilhança condicional pode ser aproximada por

$$\mathcal{L}(\gamma, \tau, r|D) \propto \tau^{(n-p)/2} \exp\left\{-\frac{\tau}{2} \sum_{t=p+1}^n a_t^2\right\}, \quad (4.3)$$

onde $D = \{Y_{(1)}, \dots, Y_{(n)}\}$ e $Y_{(i)}$ é a i -ésima estatística de ordem. Observemos que $r \in D_r = \{Y_{(2)}, \dots, Y_{(n-1)}\}$. Se $r = Y_{(1)}$ ou $r = Y_{(n)}$, teríamos um único regime. Portanto, vamos supor que r seja uniforme sobre $\{Y_{(2)}, \dots, Y_{(n-1)}\}$.

A verossimilhança não é linear em γ , já que o erro $\{a_t\}$ em (4.3) não é uma função linear dos parâmetros. A equação (4.3) pode ser feita aproximadamente linear se usarmos os resíduos

$$\hat{a}_t = \begin{cases} Y_t - \hat{\phi}_{10} - \sum_{i=1}^{p_1} \hat{\phi}_{1i} Y_{t-i} - \sum_{j=1}^{q_1} \hat{\theta}_{1j} \hat{a}_{t-j}, & \text{se } Y_{t-1} \leq r; \\ Y_t - \hat{\phi}_{20} - \sum_{i=1}^{p_2} \hat{\phi}_{2i} Y_{t-i} - \sum_{j=1}^{q_2} \hat{\theta}_{2j} \hat{a}_{t-j}, & \text{se } Y_{t-1} > r. \end{cases} \quad (4.4)$$

onde $\hat{\phi}_i$ e $\hat{\theta}_i$ são os estimadores de mínimos quadrados de ϕ_i e θ_i , respectivamente, e são obtidos minimizando a soma de quadrados

$$S(\phi, \theta) = \sum_{t=p+1}^n a_t^2 \quad (4.5)$$

em relação a ϕ_i e θ_i . Maiores detalhes podem ser encontrados em Harvey(1981). A função de verossimilhança aproximada é dada por

$$\begin{aligned} \mathcal{L}^*(\gamma, \tau, r|D) \propto \tau^{(n-p)/2} \exp\left\{-\frac{\tau}{2}\left[\sum_1 (Y_t - (\phi_{10} + \sum_{i=1}^{p_1} \phi_{1i} Y_{t-i} + \sum_{j=1}^{q_1} \theta_{1j} \hat{a}_{t-j}))^2 \right. \right. \\ \left. \left. + \sum_2 (Y_t - (\phi_{20} + \sum_{i=1}^{p_2} \phi_{2i} Y_{t-i} + \sum_{j=1}^{q_2} \theta_{2j} \hat{a}_{t-j}))^2\right]\right\}, \end{aligned} \quad (4.6)$$

onde $\sum_1$ e $\sum_2$ são somas sobre $\{t = p+1, \dots, n, Y_{t-1} \leq r\}$ e $\{t = p+1, \dots, n, Y_{t-1} > r\}$, respectivamente.

Temos então o seguinte teorema:

Teorema 4.1 Se a priori (4.2) for combinada com a função de verossimilhança (4.6) obtemos a distribuição a posteriori normal-gama

$$\begin{aligned} P(\gamma, \tau, r|D) \propto \tau^{(n-p+2\alpha)/2-1} \exp\left\{-\frac{\tau}{2}\{F(r) + [\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]' \right. \\ \left. \times (\mathbf{Q} + \mathbf{A}(r))[\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]\right\}, \end{aligned} \quad (4.7)$$

onde

$$\begin{aligned} F(r) &= \{2\beta + h(r) + \mu' \mathbf{Q} \mu + \hat{\gamma}(r)' \mathbf{A}(r) \hat{\gamma}(r) \\ &\quad - (\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))' (\mathbf{Q} + \mathbf{A}(r))^{-1} (\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))\}, \\ h(r) &= \sum_{t=p+1}^n Y_t^2 - \mathbf{B}'(r) \mathbf{A}^{-1}(r) \mathbf{B}(r), \\ \mathbf{B}(r) &= (\mathbf{B}_1(r), \mathbf{B}_2(r))', \text{ um vetor } (m \times 1), \end{aligned}$$

$\mathbf{B}_i(r) = (\mathbf{B}_{1i}, \mathbf{B}_{2i})'$. um vetor $(p_i + q_i + 1) \times 1$, $i = 1, 2$,

$$\mathbf{B}_{1i} = \begin{pmatrix} \sum_i Y_t \\ \sum_i Y_t Y_{t-1} \\ \vdots \\ \sum_i Y_t Y_{t-p_i} \end{pmatrix} , \quad \mathbf{B}_{2i} = \begin{pmatrix} \sum_i Y_t \hat{a}_{t-1} \\ \vdots \\ \sum_i Y_t \hat{a}_{t-q_i} \end{pmatrix} , \quad i = 1, 2,$$

$$\mathbf{A}(r) = \text{diag}\{\mathbf{A}_1(r), \mathbf{A}_2(r)\},$$

$$\mathbf{A}_i(r) = \begin{pmatrix} \mathbf{A}_{11}^{(i)} & \mathbf{A}_{12}^{(i)} \\ \mathbf{A}_{21}^{(i)} & \mathbf{A}_{22}^{(i)} \end{pmatrix} , \quad \text{simétrica} , \quad (p_i + q_i + 1) \times (p_i + q_i + 1) , i = 1, 2,$$

$$\mathbf{A}_{11}^{(i)} = \begin{pmatrix} \sum_i 1 & \sum_i Y_{t-1} & \dots & \sum_i Y_{t-p_i} \\ \sum_i Y_{t-1} & \sum_i Y_{t-1}^2 & \dots & \sum_i Y_{t-1} Y_{t-p_i} \\ \vdots & \vdots & \dots & \vdots \\ \sum_i Y_{t-p_i} & \sum_i Y_{t-1} Y_{t-p_i} & \dots & \sum_i Y_{t-p_i}^2 \end{pmatrix} ,$$

$$\mathbf{A}_{12}^{(i)} = \begin{pmatrix} \sum_i \hat{a}_{t-1} & \sum_i \hat{a}_{t-2} & \dots & \sum_i \hat{a}_{t-q_i} \\ \sum_i \hat{a}_{t-1} Y_{t-1} & \sum_i \hat{a}_{t-2} Y_{t-1} & \dots & \sum_i \hat{a}_{t-q_i} Y_{t-1} \\ \vdots & \vdots & \dots & \vdots \\ \sum_i \hat{a}_{t-1} Y_{t-p_i} & \sum_i \hat{a}_{t-2} Y_{t-p_i} & \dots & \sum_i \hat{a}_{t-q_i} Y_{t-p_i} \end{pmatrix} ,$$

$$\mathbf{A}_{22}^{(i)} : (q_i \times q_i) \text{matriz cujo } (j,k)\text{-ésimo elemento é } \sum_i \hat{a}_{t-j} \hat{a}_{t-k}$$

e finalmente

$$\hat{\gamma}(r) = (\hat{\gamma}_1(r), \hat{\gamma}_2(r))' = \mathbf{A}^{-1}(r) \mathbf{B}(r) .$$

Demonstração:

Combinando (4.2) com (4.6) temos

$$P(\gamma, \tau, r | D) \propto \tau^{\frac{n-p+2\alpha}{2}-1} \exp\left\{-\frac{\tau}{2}[2\beta + (\gamma - \mu)' \mathbf{Q}(\gamma - \mu)\right]$$

$$\begin{aligned}
& + \sum_1 (Y_t - (\phi_{10} + \sum_{i=1}^{p_1} \phi_{1i} Y_{t-i} + \sum_{j=1}^{q_1} \theta_{1j} \hat{a}_{t-j}))^2 \\
& + \sum_2 (Y_t - (\phi_{20} + \sum_{i=1}^{p_2} \phi_{2i} Y_{t-i} + \sum_{j=1}^{q_2} \theta_{2j} \hat{a}_{t-j}))^2 \}
\end{aligned}$$

Sejam $\phi_l = (\phi_{l0}, \phi_{l1}, \dots, \phi_{lp_l})'$, $\theta_l = (\theta_{l1}, \theta_{l2}, \dots, \theta_{lq_l})'$, $Y_l = (1, Y_{t-1}, \dots, Y_{t-p_l})'$ e $\hat{a}_l = (\hat{a}_{t-1}, \hat{a}_{t-2}, \dots, \hat{a}_{t-q_l})'$.

Resolvendo a expressão

$$\sum_l (Y_t - (\phi_{l0} + \sum_{i=1}^{p_l} \phi_{li} Y_{t-i} + \sum_{j=1}^{q_l} \theta_{lj} \hat{a}_{t-j}))^2$$

para $l = 1, 2$, obtemos

$$\begin{aligned}
& \sum_l (Y_t - \phi_l' Y_l - \theta_l' \hat{a}_l)(Y_t - Y_l' \phi_l - \hat{a}_l' \theta_l) \\
& = \sum_l Y_t^2 - \mathbf{B}_{1l}' \phi_l - \mathbf{B}_{2l}' \theta_l - \phi_l' \mathbf{B}_{1l} - \theta_l' \mathbf{B}_{2l} \\
& \quad + \phi_l' \mathbf{A}_{11}^{(l)} \phi_l + \phi_l' \mathbf{A}_{12}^{(l)} \theta_l + \theta_l' \mathbf{A}_{21}^{(l)} \phi_l + \theta_l' \mathbf{A}_{22}^{(l)} \theta_l \\
& = \sum_l Y_t^2 + \gamma_l' \begin{pmatrix} \mathbf{A}_{11}^{(l)} & \mathbf{A}_{12}^{(l)} \\ \mathbf{A}_{21}^{(l)} & \mathbf{A}_{22}^{(l)} \end{pmatrix} \gamma_l - \gamma_l' \begin{pmatrix} \mathbf{B}_{1l} \\ \mathbf{B}_{2l} \end{pmatrix} - (\mathbf{B}_{1l}', \mathbf{B}_{2l}') \gamma \\
& = \sum_l Y_t^2 + \gamma_l' \mathbf{A}_l(r) \gamma_l - \gamma_l' \mathbf{B}_l(r) - \mathbf{B}_l(r)' \gamma.
\end{aligned}$$

Fazendo

$$\mathbf{A}(r) = \begin{pmatrix} \mathbf{A}_1(r) & 0 \\ 0 & \mathbf{A}_2(r) \end{pmatrix} \text{ e } \hat{\gamma}(r) = \mathbf{A}^{-1}(r) \mathbf{B}(r),$$

obtemos que a posteriori é dada por

$$P(\gamma, \tau, r|D) \propto \tau^{(n-p-2\alpha)/2-1} \exp\left\{\frac{-\tau}{2}\left[2\beta + \sum_{t=p+1}^n Y_t^2\right.\right.$$

$$\left. + (\gamma - \mu)' \mathbf{Q}(\gamma - \mu) + (\gamma - \hat{\gamma}(r))' \mathbf{A}(r)(\gamma - \hat{\gamma}(r))\right.$$

$$\left. - \mathbf{B}'(r) \mathbf{A}^{-1}(r) \mathbf{B}(r)\right]\}$$

$$\propto \tau^{(n-p-2\alpha)/2-1} \exp\left\{\frac{-\tau}{2}\left[2\beta + \sum_{t=p+1}^n Y_t^2\right.\right.$$

$$\left. + (\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)))'(\mathbf{Q} + \mathbf{A}(r))(\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)))\right.$$

$$\left. - \mathbf{B}'(r) \mathbf{A}^{-1}(r) \mathbf{B}(r) - (\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))'(\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))\right.$$

$$\left. + \mu' \mathbf{Q}\mu + \hat{\gamma}'(r) \mathbf{A}(r) \hat{\gamma}(r) - \mathbf{B}'(r) \mathbf{A}^{-1}(r) \mathbf{B}(r)\right]\}.$$

Portanto com a notação do Teorema temos

$$P(\gamma, \tau, r|D) \propto \tau^{(n-p-2\alpha)/2-1} \exp\left\{\frac{-\tau}{2}\{F(r)\right.$$

$$\left. -[\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]'(\mathbf{Q} + \mathbf{A}(r))[\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]\right\}\} \square$$

Para obtermos as distribuições marginais a posteriori de γ e τ vamos primeiramente determinar as marginais de γ e τ condicionadas à variável limiar r .

A distribuição a posteriori conjunta de γ e τ , condicionada a r , $P(\gamma, \tau|r, D)$, obtida de (4.7), é dada por

$$\begin{aligned} P(\gamma, \tau|r, D) &\propto |\tau(\mathbf{Q} + \mathbf{A}(r))|^{-\frac{1}{2}} \\ &\exp\left\{\frac{-\tau}{2}([\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]'(\mathbf{Q} + \mathbf{A}(r))\right. \\ &\quad \times [\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]\} \\ &\quad \times |\mathbf{Q} + \mathbf{A}(r)|^{-\frac{1}{2}} \tau^{(n-m-p+2\alpha)/2-1} \exp\left\{\frac{-\tau}{2}F(r)\right\}. \end{aligned} \quad (4.8)$$

Corolário 4.1 As distribuições marginais condicionais são obtidas da equação (4.8) e são dadas por

$$\gamma|\tau, r, D \sim \text{Normal}((\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)), \tau^{-1}(\mathbf{Q} + \mathbf{A}(r))^{-1}), \quad (4.9)$$

$$\tau|r, D \sim \text{Gama}\left(\frac{\nu}{2}, \frac{F(r)}{2}\right), \nu = n - m - p + 2\alpha. \quad (4.10)$$

Além disso, integrando (4.8) em relação a τ obtemos

$$\gamma|r, D \sim t_m(\nu, (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)), \mathbf{W}(r)), \quad (4.11)$$

onde $\mathbf{W}(r) = \nu F^{-1}(r)(\mathbf{Q} + \mathbf{A}(r))$ e t_m é a distribuição t-multivariada de dimensão m .

Demonstração:

As equações (4.9) e (4.10) são imediatas de (4.8). Para a equação (4.11) observemos que (4.8) pode ser reescrita por

$$\begin{aligned} P(\gamma, \tau|r, D) &\propto \tau^{(n-p+2\alpha)/2-1} \exp\left\{\frac{-\tau}{2}\left[1\right.\right. \\ &\quad \left.+[\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]F^{-1}(r)(\mathbf{Q} + \mathbf{A}(r))\right. \\ &\quad \left.\left.\times[\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]\right]\right\}. \end{aligned}$$

Completando a expressão acima de modo a obtermos uma densidade Gama para τ e integrando em relação a τ obtemos

$$\begin{aligned} P(\gamma|r, D) &\propto \{1 + [\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]F^{-1}(r)(\mathbf{Q} + \mathbf{A}(r)) \\ &\quad \times[\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]\}^{-(n-p+2\alpha)/2} \\ &\propto \left\{1 + \frac{1}{\nu}[\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]F^{-1}(r)(\mathbf{Q} + \mathbf{A}(r))\right. \\ &\quad \left.\times[\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]\}^{-(\nu+m)/2}, \end{aligned}$$

onde $\nu = n - p - m + 2\alpha$. $\square$

Para encontrarmos as distribuições exatas de γ e τ necessitamos da distribuição a posteriori de r , a qual é obtida integrando (4.8) em relação a γ e τ . Isto é, como

$$P(\tau, r|d) \propto |\mathbf{Q} + \mathbf{A}(r)|^{\frac{-1}{2}} \tau^{(n-m-p+2\alpha)/2-1} \exp\left\{\frac{-\tau}{2} F(r)\right\},$$

temos

$$P(r|D) \propto |\mathbf{Q} + \mathbf{A}(r)|^{\frac{-1}{2}} \{F(r)\}^{\frac{-\nu}{2}}, r \in D_r \quad (4.12)$$

onde a constante normalizadora de (4.12) e os momentos da distribuição podem ser encontrados numericamente.

Portanto, as marginais exatas de γ e τ são

$$P(\tau|D) = \sum_{r \in D_r} P(\tau|r, D)P(r|D), \quad (4.13)$$

e

$$P(\gamma|D) = \sum_{r \in D_r} P(\gamma|r, D)P(r|D), \quad (4.14)$$

respectivamente. Temos que (4.13) é uma mistura de distribuições Gama, com

$$E(\tau|D) = E_{r|D}(E(\tau|r, D)) = \sum_{r \in D_r} \frac{\nu}{F(r)} P(r|D),$$

e

$$\begin{aligned} Var(\tau|D) &= E_{r|D}(Var(\tau|r, D)) + Var_{r|D}(E(\tau|r, D)) \\ &= \sum_{r \in D_r} \frac{2\nu}{[F(r)]^2} P(r|D) + \sum_{r \in D_r} [E(\tau|r, D) - E(\tau|D)]^2 P(r|D). \end{aligned}$$

Analogamente, (4.14) é uma mistura de distribuições t-multivariadas, com médias e variâncias dadas respectivamente por

$$E(\gamma|D) = \sum_{r \in D_r} (\mathbf{Q} + \mathbf{A}(r))^{-1} (\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) P(r|D)$$

e

$$\begin{aligned} Var(\gamma|D) = & \sum_{r \in D_r} \frac{F(r)(\mathbf{Q} + \mathbf{A}(r))^{-1}}{\nu - 2} P(r|D) \\ & + \sum_{r \in D_r} [(\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) - E(\gamma|D)][(\mathbf{Q} + \mathbf{A}(r))^{-1} \\ & \times (\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) - E(\gamma|D)]]' P(r|D) . \end{aligned}$$

Observemos que se a variável limiar não existe, então os regimes são iguais. Uma maneira simples de verificarmos a existência da variável limiar é supondo $p = p_1 = p_2, q = q_1 = q_2$ e definindo $\eta = M\hat{\gamma}(r)$, onde $M = (\mathbf{I}_{p+q+1}, -\mathbf{I}_{p+q+1})$. Observemos que $\eta = \hat{\gamma}_1(r) - \hat{\gamma}_2(r)$ e, portanto, se $\eta = 0$ temos que os regimes são iguais e a variável limiar não existe.

4.2 Análise a posteriori com priori imprópria de Jeffreys

Se temos pouca ou nenhuma informação sobre os parâmetros podemos usar a priori de Jeffreys

$$P(\gamma, \tau, r) \propto \tau^{-1}. \quad (4.15)$$

Para esta priori, a posteriori ainda será uma normal-gama. Os resultados obtidos na seção 4.2 são modificados fazendo-se $\beta \rightarrow 0$, $\mathbf{Q} \rightarrow \mathbf{0}$ e $\alpha \rightarrow -m/2$ na distribuição a posteriori conjunta.

Corolário 4.2 Se $n > p + 2m$, então as distribuições marginais são dadas por

$$\gamma|\tau, r, D \sim \text{Normal}(\hat{\gamma}(r), \tau^{-1}\mathbf{A}(r)^{-1}) , \quad (4.16)$$

$$\tau|r, D \sim \text{Gama}\left(\frac{\nu'}{2}, \frac{h(r)}{2}\right) , \nu' = n - p - 2m , \quad (4.17)$$

$$\gamma|r, D \sim t_m(\nu', \hat{\gamma}(r), \nu'h^{-1}(r)\mathbf{A}(r)) . \quad (4.18)$$

A demonstração do Corolário 4.2 é análoga à do Corolário 4.1.

A distribuição de r , obtida de (4.12), é dada por

$$P(r|D) \propto |\mathbf{A}(r)|^{-1/2}[h(r)]^{-\nu'/2} , r \in D_r . \quad (4.19)$$

As distribuições marginais a posteriori para τ e γ são como em (4.13) e (4.14), respectivamente, com ν e $F(r)$ substituídos por ν' e $h(r)$ nas expressões para $E(\tau|D)$ e $Var(\tau|D)$, respectivamente. Para a distribuição de $\gamma|D$, temos

$$E(\gamma|D) = \sum_{r \in D_r} \hat{\gamma}(r)P(r|D) \quad (4.20)$$

e

$$\begin{aligned} Var(\gamma|D) &= \sum_{r \in D_r} \frac{h(r)\mathbf{A}(r)^{-1}}{\nu' - 2} P(r|D) \\ &+ \sum_{r \in D_r} [(\hat{\gamma}(r) - E(\gamma|D))(\hat{\gamma}(r) - E(\gamma|D))'] P(r|D) . \end{aligned} \quad (4.21)$$

4.3 Distribuição preditiva

Como visto no Capítulo 2, a distribuição preditiva Bayesiana é a densidade condicional de observações futuras dadas observações passadas. Vamos

determinar a densidade preditiva para previsão 1-passo à frente, Y_{n+1} , dado que temos as observações $Y_1, \dots, Y_n$.

Observemos que

$$Y_{n+1} = \begin{cases} \phi_{10} + \sum_{i=1}^{p_1} \phi_{1i} Y_{n+1-i} + a_{n+1} + \sum_{j=1}^{q_1} \theta_{1j} a_{n+1-j}, & \text{se } Y_n \leq r \\ \phi_{20} + \sum_{i=1}^{p_2} \phi_{2i} Y_{n+1-i} + a_{n+1} + \sum_{j=1}^{q_2} \theta_{2j} a_{n+1-j}, & \text{se } Y_n > r \end{cases} \quad (4.22)$$

e, portanto

$$a_{n+1} = \begin{cases} Y_{n+1} - \mathbf{z}_1' \gamma_1, & \text{se } Y_n \leq r \\ Y_{n+1} - \mathbf{z}_2' \gamma_2, & \text{se } Y_n > r \end{cases}$$

onde

$$\mathbf{z}_i = (1, Y_n, Y_{n-1}, \dots, Y_{n+1-p_i}, a_n, \dots, a_{n+1-q_i})', i = 1, 2.$$

Suponhamos que $Y_n \leq r$ (Se $Y_n > r$, o procedimento é análogo). Então temos que

$$Y_{n+1} = \mathbf{z}_1' \gamma_1 + a_{n+1}.$$

A função de verossimilhança aproximada de Y_{n+1} , condicionada a todos os parâmetros e observações passadas é dada por

$$\mathcal{L}^*(Y_{n+1}|\gamma, \tau, r, D) \propto \tau^{1/2} \exp\left\{-\frac{\tau}{2}(Y_{n+1} - \mathbf{z}_1' \gamma_1)^2\right\}. \quad (4.23)$$

A densidade conjunta de Y_{n+1}, γ e τ dados r e D , obtida de (4.23) e (4.8) é dada por

$$P(Y_{n+1}, \gamma, \tau|r, D) \propto \tau^{(n-p+2\alpha+1)/2-1} \exp\left\{-\frac{\tau}{2}[F(r) + (Y_{n+1} - \mathbf{z}_1' \gamma_1)^2]\right\}$$

$$+(\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)))'(\mathbf{Q} + \mathbf{A}(r))$$

$$\times(\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)))\}]. \quad (4.24)$$

A densidade preditiva de Y_{n+1} , dados r e D , é obtida de (4.24), após completarmos o quadrado em γ e Y_{n+1} e integrarmos em relação a γ e τ .

Teorema 4.2 A densidade preditiva de Y_{n+1} , dados r e D , é uma distribuição t-univariada com ν graus de liberdade, valor esperado

$$\hat{Y}(r) = E(Y_{n+1}|r, D) = \mathbf{z}'_1 \mathbf{I}_1(r) (\mathbf{Q} + \mathbf{A}(r))^{-1} (\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) \quad (4.25)$$

e variância

$$Var(Y_{n+1}|r, D) = \mathbf{z}'_1 \mathbf{I}_1(r) \frac{F(r)(\mathbf{Q} + \mathbf{A}(r))^{-1}}{\nu - 2} \mathbf{I}'_1(r) \mathbf{z}_1 + E(\tau^{-1}|r, D) \quad (4.26)$$

onde

$$\mathbf{I}_1(r) = (\mathbf{I}_{p_1+q_1+1}, \mathbf{0}_{(p_1+q_1+1) \times (p_2+q_2+1)}).$$

Demonstração:

Observemos que $\gamma_1 = \mathbf{I}_1(r)\gamma$ e, portanto,

$$\begin{aligned} P(Y_{n+1}, \gamma, \tau|r, D) &\propto \tau^{(n-p+2\alpha+1)/2-1} \exp\left\{\frac{-\tau}{2}[F(r) + (Y_{n+1} - \mathbf{z}'_1 \mathbf{I}_1(r)\gamma)^2 \right. \\ &\quad \left. + [\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]'(\mathbf{Q} + \mathbf{A}(r))[\gamma - (\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))]\right\} \end{aligned}$$

$$\alpha \quad \tau^{(n-p+2\alpha+1)/2-1} \exp\left\{\frac{-\tau}{2}[F(r) + Y_{n+1}^2 - Y_{n+1} \mathbf{z}_1' \mathbf{I}_1(r) \gamma\right.$$

$$\left. - \gamma_1' \mathbf{I}_1'(r) \mathbf{z}_1 Y_{n+1} + \gamma' \mathbf{I}_1'(r) \mathbf{z}_1 \mathbf{z}_1' \mathbf{I}_1(r) \gamma + \gamma' (\mathbf{Q} + \mathbf{A}(r)) \gamma\right.$$

$$\left. - \gamma' (\mathbf{Q} \mu + \mathbf{A}(r) \hat{\gamma}(r)) - (\mathbf{Q} \mu + \mathbf{A}(r) \hat{\gamma}(r))' \gamma\right.$$

$$\left. + (\mathbf{Q} \mu + \mathbf{A}(r) \hat{\gamma}(r))' (\mathbf{Q} + \mathbf{A}(r))^{-1} (\mathbf{Q} \mu + \mathbf{A}(r) \hat{\gamma}(r))\right\}$$

$$\alpha \quad \tau^{(n-p+2\alpha+1)/2-1} \exp\left\{\frac{-\tau}{2}[F(r) + Y_{n+1}^2\right.$$

$$\left. + \gamma' (\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r) \mathbf{z}_1 \mathbf{z}_1' \mathbf{I}_1(r))\right) \gamma - (\mathbf{Q} \mu + \mathbf{A}(r) \hat{\gamma}(r))' + Y_{n+1} \mathbf{z}_1' \mathbf{I}_1(r) \gamma$$

$$\left. - \gamma' (\mathbf{Q} \mu + \mathbf{A}(r) \hat{\gamma}(r)) + \mathbf{I}_1'(r) \mathbf{z}_1 Y_{n+1}\right)$$

$$\left. + ((\mathbf{Q} \mu + \mathbf{A}(r) \hat{\gamma}(r)))' (\mathbf{Q} + \mathbf{A}(r))^{-1} ((\mathbf{Q} \mu + \mathbf{A}(r) \hat{\gamma}(r)))\right\}$$

$$\alpha \quad \tau^{(n-p+2\alpha+1)/2-1} \exp\left\{\frac{-\tau}{2}[F(r) + Y_{n+1}^2\right.$$

$$\left. + \{\gamma - [(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r) \mathbf{z}_1 \mathbf{z}_1' \mathbf{I}_1(r))]\}^{-1} ((\mathbf{Q} \mu + \mathbf{A}(r) \hat{\gamma}(r)) + \mathbf{I}_1'(r) \mathbf{z}_1 Y_{n+1})\}'\right.$$

$$\times(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r)\mathbf{z}_1\mathbf{z}_1'\mathbf{I}_1(r))$$

$$\times\{\gamma - [(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r)\mathbf{z}_1\mathbf{z}_1'\mathbf{I}_1(r))]\}^{-1}((\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) + \mathbf{I}'_1(r)\mathbf{z}_1Y_{n+1})\}$$

$$-((\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) + \mathbf{I}'_1(r)\mathbf{z}_1Y_{n+1})'[(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r)\mathbf{z}_1\mathbf{z}_1'\mathbf{I}_1(r))]\}^{-1}$$

$$\times((\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) + \mathbf{I}'_1(r)\mathbf{z}_1Y_{n+1}))\}$$

$$\alpha \quad \tau^{(n-m-p+2\alpha+1)/2-1}|\tau(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r)\mathbf{z}_1\mathbf{z}_1'\mathbf{I}_1(r))|\}^{\frac{1}{2}}$$

$$\times|\tau(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r)\mathbf{z}_1\mathbf{z}_1'\mathbf{I}_1(r))|\}^{\frac{-1}{2}}\exp\{\frac{-\tau}{2}$$

$$\times[\gamma - [(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r)\mathbf{z}_1\mathbf{z}_1'\mathbf{I}_1(r))]]^{-1}((\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) + \mathbf{I}'_1(r)\mathbf{z}_1Y_{n+1})]'$$

$$\times(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r)\mathbf{z}_1\mathbf{z}_1'\mathbf{I}_1(r))$$

$$\times[\gamma - [(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r)\mathbf{z}_1\mathbf{z}_1'\mathbf{I}_1(r))]]^{-1}((\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) + \mathbf{I}'_1(r)\mathbf{z}_1Y_{n+1})]$$

$$\times\exp\{\frac{-\tau}{2}[F(r) + Y_{n+1}^2 - ((\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) + \mathbf{I}'_1(r)\mathbf{z}_1Y_{n+1})']$$

$$\times [(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r)\mathbf{z}_1\mathbf{z}_1'\mathbf{I}_1(r)))]^{-1}((\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) + \mathbf{I}'_1(r)\mathbf{z}_1Y_{n+1}))].$$

Integrando em relação a γ obtemos

$$P(Y_{n+1}, \tau|r, D) \propto \tau^{(n-m-p+2\alpha+1)/2-1} |(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r)\mathbf{z}_1\mathbf{z}_1'\mathbf{I}_1(r))|^{-\frac{1}{2}}$$

$$\times \exp\left\{\frac{-\tau}{2}[F(r) + Y_{n+1}^2 - ((\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) + \mathbf{I}'_1(r)\mathbf{z}_1Y_{n+1})']\right\}$$

$$\times [(\mathbf{Q} + \mathbf{A}(r)) + (\mathbf{I}'(r)\mathbf{z}_1\mathbf{z}_1'\mathbf{I}_1(r)))]^{-1}((\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r)) + \mathbf{I}'_1(r)\mathbf{z}_1Y_{n+1})).$$

Completando o quadrado em Y_{n+1} e integrando em relação a τ obtemos a distribuição de $Y_{n+1}|r, D$. $\square$

Corolário 4.3 A densidade preditiva de Y_{n+1} é uma mistura de distribuições t-univariadas, dada por

$$P(Y_{n+1}|D) = \sum_{r \in D_r} P(Y_{n+1}|r, D)P(r|D), \quad (4.27)$$

com valores esperados

$$E(Y_{n+1}|D) = \sum_{r \in D_r} \mathbf{z}_1'\mathbf{I}_1(r)(\mathbf{Q} + \mathbf{A}(r))^{-1}(\mathbf{Q}\mu + \mathbf{A}(r)\hat{\gamma}(r))P(r|D) \quad (4.28)$$

e variâncias

$$Var(Y_{n+1}|D) = \sum_{r \in D_r} Var(Y_{n+1}|r, D) + \sum_{r \in D_r} (\hat{Y}(r) - E(Y_{n+1}|D))^2 P(r|D). \quad (4.29)$$

Corolário 4.4 Se combinarmos (4.23) com a priori de Jeffreys (4.15), os momentos da densidade preditiva de Y_{n+1} são dados por

$$E(Y_{n+1}|D) = \sum_{r \in D_r} \mathbf{z}'_1 \hat{\gamma}_1(r) P(r|D) \quad (4.30)$$

e

$$Var(Y_{n+1}|D) = \sum_{r \in D_r} Var(Y_{n+1}|r, D) + \sum_{r \in D_r} (\hat{Y}(r) - E(Y_{n+1}|D))^2 P(r|D) , \quad (4.31)$$

onde

$$Var(Y_{n+1}|r, D) = \mathbf{z}'_1 \mathbf{I}_1(r) \frac{h(r) \mathbf{A}(r)^{-1}}{\nu' - 2} \mathbf{I}_1(r)' \mathbf{z}_1 + E(\tau^{-1}|r, D) . \quad (4.32)$$

As demonstrações para os Corolários 4.3 e 4.4 são imediatas.

4.4 Aplicações

Seja $\{Z_t, t = 1, \dots, 101\}$ a série “Canadian Lynx”, considerada por Tong(1983), para o período 1821-1922 e seja $Y_t = \log_{10} Z_t$. Vamos considerar as $n = 100$ primeiras observações desta série.

Ordenando os valores de Y_t obtemos D_r . Os valores limiares considerados são 2,58, 2,77 e 3,01, os quais correspondem aos quantis 32, 43 e 55. Para o parâmetro de retardo consideramos $d = 2$.

Para cada valor limiar consideramos dois regimes e para cada regime ajustamos um modelo ARMA. As ordens dos modelos foram obtidas usando-

se o critério AICC (Brockwell e Davis, 1991a), de modo que para cada r temos um modelo $TARMA(2; p_1, p_2, q_1, q_2)$.

Consideremos a priori de Jeffreys, às distribuições marginais condicionadas a r são obtidas do Corolário 4.2.

A Tabela I,(a)-(c) mostra os resultados obtidos para cada valor de r . Observemos que embora tenhamos considerado τ comum aos regimes, esta suposição pode não ser válida. Uma possibilidade é considerar que temos duas sequências de ruídos branco, com precisões τ_1 e τ_2 , independentes, e supor que cada τ_i tem uma distribuição a priori Gama com r independente de γ, τ_1 e τ_2 .

A previsão 1-passo à frente é obtida do Teorema 4.2. Vamos calcular Y_{101} .

Como $Y_{99} = 1, 90$, temos que Y_{101} pertence ao primeiro regime para todos os modelos. Portanto,

$$Y_{n+1} = \mathbf{z}_1' \gamma_1 + a_{n+1}$$

onde $\mathbf{z}_1 = (1, Y_n, \dots, Y_{n+1-p_1}, a_n, \dots, a_{n+1-q_1})'$.

Considerando a priori de Jeffreys, do Teorema 4.2 temos

$$\hat{Y}(r) = E(Y_{n+1}|r, D) = \mathbf{z}_1' \hat{\gamma}_1(r),$$

$$Var(Y_{n+1}|r, D) = \mathbf{z}_1' \mathbf{I}_1(r) \frac{\mathbf{A}(r)^{-1}}{\nu' - 2} \mathbf{I}_1'(r) \mathbf{z}_1 + E(\tau^{-1}|r, D).$$

Tabela I

parâmetros	médias	d.p.
ϕ_{10}	3,762	0,089
ϕ_{11}	-0,418	0,033
θ_{11}	0,516	0,192
ϕ_{20}	0,702	0,059
ϕ_{21}	0,741	0,019
θ_{21}	-0,108	0,140
θ_{22}	-0,560	0,141
θ_{23}	-0,577	0,140
σ^2	0,222	0,032

(a) modelo TARMA(2;1,1,1,3), $r=2,58$

parâmetros	médias	d.p.
ϕ_{10}	1,807	0,064
ϕ_{11}	0,332	0,023
θ_{11}	1,027	0,134
θ_{12}	-0,167	0,137
ϕ_{20}	-0,235	0,053
ϕ_{21}	1,059	0,017
θ_{21}	-0,667	0,113
θ_{22}	-0,715	0,114
σ^2	0,154	0,032

(b) modelo TARMA(2;1,1,2,2), $r=2,77$

parâmetros	médias	d.p.
ϕ_{10}	0,710	0,079
ϕ_{11}	0,733	0,028
θ_{11}	-0,363	0,162
θ_{12}	-0,550	0,165
ϕ_{20}	2,918	0,086
ϕ_{21}	0,022	0,029
θ_{21}	0,105	0,186
θ_{22}	-0,556	0,183
θ_{23}	-0,261	0,183
σ^2	0,304	0,045

(c) modelo TARMA(2;1,1,2,3), $r=3,01$

A Tabela II mostra as previsões 1-passo à frente para os diferentes valores de r .

Tabela II-Previsão 1-passo à frente

r	$\hat{Y}(r)$	$Var(Y_{101} r, D)$
2,58	2,555	0,039
2,77	1,913	0,154
3,01	2,681	0,345

O valor real de Y_{101} é 2,36, de modo que analisando em termos do vício e da variância, o modelo para $r = 2,58$ nos dá uma melhor previsão 1-passo à frente, para os valores limiares considerados.

Considerações

1 Observemos que algoritmos de simulação não foram necessários, já que as distribuições a posteriori são conhecidas. Apesar disso, para uma aplicação prática, o trabalho seria excessivo, já que para cada valor de r teríamos que ajustar 2 modelos, uma vez que as ordens p_l e q_l mudam de regime para regime. Se o valor limiar, r , for conhecido, a análise usando este método é simples, já que necessitaríamos ajustar apenas 2 modelos ARMA, um para cada regime e a inferência sobre os parâmetros seria obtida diretamente das distribuições a posteriori.

No caso de não conhecermos o valor limiar, ou termos vários limiares, o problema pode ser resolvido se utilizarmos o método de autoregressão rearranjada proposto por Tsay(1989), o qual será apresentado no Capítulo 5.

2 A generalização para o modelo TARMA com k regimes é feita sem grandes dificuldades. A única modificação é a escolha da distribuição a priori para as variáveis limiares. Para ilustrarmos, vamos considerar o caso onde $k = 3$.

Teremos então três processos ARMA, com ordens $(p_i, q_i), i = 1, 2, 3$ e dois valores limiares r_1 e r_2 , tais que $Y_{t-d} \leq r_1$ para o primeiro regime, $r_1 < Y_{t-d} \leq r_2$ para o segundo regime e $Y_{t-d} > r_2$ para o terceiro, onde $-\infty < r_1 < r_2 < \infty$.

Dada a amostra $Y_1, \dots, Y_n$, sejam $D = \{Y_{(1)}, \dots, Y_{(n)}\}$. Como $r_1 \in$

$(Y_{(1)}, Y_{(n)})$ e $r_2 \in (Y_{(1)}, Y_{(n)})$, com a condição $r_1 < r_2$, temos então que $r = (r_1, r_2)'$ é uniforme na região triangular $\{Y_{(1)} < r_1 < Y_{(n)}, Y_{(1)} < r_2 < Y_{(n)}, r_1 < r_2\}$ e, portanto, a distribuição a priori de r , $P(r)$, é tal que

$$\int_{Y_{(1)}}^{Y_{(n)}} \int_0^{r_2} P(r) dr_1 dr_2 = 1$$

De maneira análoga para k regimes, consideramos os $k-1$ valores limiares $r_1, r_2, \dots, r_{k-1}$, com a condição $r_1 < r_2 < \dots < r_{k-1}$.

3 Testes para não-linearidade e existência dos valores limiares são encontrados em Tong(1990).

Capítulo 5

Modelo TARMA Rearranjado

Neste Capítulo vamos analisar o modelo TARMA considerando uma extensão da auto-regressão rearranjada, método proposto por Tsay(1989). A seção 5.1 descreve o método e apresenta o modelo TARMA rearranjado. A análise Bayesiana é encontrada na seção 5.2, a qual é feita através do Amostrador de Gibbs e do Algoritmo de Metropolis-Hastings. Para finalizar apresentamos um exemplo numérico onde os dados são simulados.

5.1 Descrição do Modelo

Uma extensão do conceito de auto-regressão rearranjada é útil para o modelo ARMA porque separa efetivamente os regimes, ou seja, os dados são

agrupados de modo que as observações em cada grupo sigam o mesmo modelo ARMA. Observa-se que para esta separação não é necessário conhecermos o valor exato do limiar r . Apenas o número de observações em cada regime depende de r .

Um processo ARMA(p, q) com n observações pode ser visto como

$$Y_t = (1, Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}, a_{t-1}, a_{t-2}, \dots, a_{t-q})\gamma + a_t$$

$t = 1, \dots, n$, onde $\{a_t\}$ é um ruído e $\gamma = (\phi_0, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)'$ é o vetor de coeficientes de dimensão $(p + q + 1)$.

A variável limiar Y_{t-d} pode assumir os valores $\{Y_{p+1-d}, Y_{p+2-d}, \dots, Y_{n-d}\}$ para $t = p + 1, \dots, n$. Comparamos essas observações com o valor limiar r e associamos o índice π_i às observações em ordem de ocorrência. Assim, temos que as s observações do primeiro regime são Y_{π_i+d} para $i = 1, \dots, s$ e, para $i > s$ temos as observações do segundo regime.

O modelo (3.2) pode ser reescrito por:

$$Y_{\pi_i+d} = \begin{cases} \phi_{10} + \sum_{j=1}^{p_1} \phi_{1j} Y_{\pi_i+d-j} + a_{\pi_i+d}^{(1)} + \sum_{k=1}^{q_1} \theta_{1k} a_{\pi_i+d-k}^{(1)}, & \text{se } i \leq s \\ \phi_{20} + \sum_{j=1}^{p_2} \phi_{2j} Y_{\pi_i+d-j} + a_{\pi_i+d}^{(2)} + \sum_{k=1}^{q_2} \theta_{2k} a_{\pi_i+d-k}^{(2)}, & \text{se } i > s \end{cases} \quad (5.1)$$

onde s satisfaz $Y_{\pi_s} \leq r < Y_{\pi_{s+1}}$.

O modelo (5.1) será denominado “Modelo Rearranjado”. Observemos que apenas rearranjamos os valores de $(Y_t, 1, Y_{t-1}, \dots, Y_{t-p_i}, a_{t-1}^{(i)}, \dots, a_{t-q_i}^{(i)})$, $i = 1, 2$ baseando-nos no valor da variável limiar.

Temos um modelo com s observações no primeiro regime e $(n - p - s)$ observações no segundo regime. Ou ainda, as observações do primeiro regime

ordenadas em ordem de ocorrência são $\{Y_{\pi_1+d}, \dots, Y_{\pi_s+d}\}$ e as do segundo regime são $\{Y_{\pi_{s+1}+d}, \dots, Y_{\pi_{n-p}+d}\}$.

Sejam $Y_1^* = (Y_{\pi_1+d}, \dots, Y_{\pi_s+d})'$ e $Y_2^* = (Y_{\pi_{s+1}+d}, \dots, Y_{\pi_{n-p}+d})'$.

Façamos ainda

$$X_1^* = (x_{1,\pi_1+d}, x_{1,\pi_2+d}, \dots, x_{1,\pi_s+d}), X_2^* = (x_{2,\pi_{s+1}+d}, \dots, x_{2,\pi_{n-p}+d}),$$

onde

$$x_{1,k} = (1, Y_{\pi_k+d-1}, Y_{\pi_k+d-2}, \dots, Y_{\pi_k+d-p_1}, a_{\pi_k+d-1}^{(1)}, \dots, a_{\pi_k+d-q_1}^{(1)})',$$

$k = 1, \dots, s$ e

$$x_{2,k} = (1, Y_{\pi_k+d-1}, Y_{\pi_k+d-2}, \dots, Y_{\pi_k+p_2}, a_{\pi_k+d-1}^{(2)}, \dots, a_{\pi_k+d-q_2}^{(2)})',$$

$k = s+1, \dots, n-p$.

Então,

$$(X_1^*)' = \begin{pmatrix} x_{1,\pi_1+d} \\ x_{1,\pi_2+d} \\ \vdots \\ x_{1,\pi_s+d} \end{pmatrix} = \begin{pmatrix} 1 & Y_{\pi_1+d-1} \dots Y_{\pi_1+d-p_1} & a_{\pi_1+d-1}^{(1)} \dots a_{\pi_1+d-q_1}^{(1)} \\ 1 & Y_{\pi_2+d-1} \dots Y_{\pi_2+d-p_1} & a_{\pi_2+d-1}^{(1)} \dots a_{\pi_2+d-q_1}^{(1)} \\ & & \vdots \\ 1 & Y_{\pi_s+d-1} \dots Y_{\pi_s+d-p_1} & a_{\pi_s+d-1}^{(1)} \dots a_{\pi_s+d-q_1}^{(1)} \end{pmatrix}$$

é matriz $(s \times (1 + p_1 + q_1))$. De maneira análoga, definimos a matriz $(X_2^*)'$, a qual é $((n-p-s) \times (1 + p_2 + q_2))$.

5.2 Análise a posteriori

Os parâmetros para o modelo (5.1) são $\tau_1, \tau_2, \gamma_1, \gamma_2, r$ e d onde $\gamma_i = (\phi_{i0}, \phi_{i1}, \dots, \phi_{ip_i}, \theta_{i1}, \theta_{i2}, \dots, \theta_{iq_i})'$.

Suporemos que as distribuições a priori são independentes e dadas por:

- 1) γ_1 e γ_2 são independentes $\sim N(\gamma_{0i}, \mathbf{Q}_i^{-1})$, $i = 1, 2$,
- 2) τ_1 e τ_2 são independentes $\sim \text{Gama}(\alpha_i/2, \beta_i/2)$, $i = 1, 2$,
- 3) $r \sim \text{Uniforme}(a, b)$, onde a e b são como no Capítulo 4,
- 4) $d \sim \text{Uniforme} \{1, 2, \dots, d_0\}$.

Os hiperparâmetros $\gamma_{0i}, \mathbf{Q}_i, \alpha_i, \beta_i, d_0, i = 1, 2$ são supostos conhecidos.

A distribuição a priori conjunta é dada por

$$P(\gamma_1, \gamma_2, \tau_1, \tau_2, r, d) \propto (\tau_1^{(\alpha_1/2)-1} \tau_2^{(\alpha_2/2)-1}) \times \exp\left\{\frac{-1}{2} \sum_{i=1}^2 (\beta_i \tau_i + (\gamma_i - \gamma_{0i})' \mathbf{Q}_i (\gamma_i - \gamma_{0i}))\right\}. \quad (5.2)$$

A função de verossimilhança é dada aproximadamente por

$$\mathcal{L}(\gamma_1, \gamma_2, \tau_1, \tau_2, r, d|Y) \propto (\tau_1^{s/2}) (\tau_2^{(n-p-s)/2}) \times$$

$$\exp\left\{\frac{-\tau_1}{2} \sum_{i=1}^s (Y_{\pi_i+d} - \phi_{10} - \sum_{j=1}^{p_1} \phi_{1j} Y_{\pi_i+d-j} - \sum_{k=1}^{q_1} \theta_{1k} \hat{a}_{\pi_i+d-k}^{(1)})^2\right\}$$

$$\frac{-\tau_2}{2} \sum_{i=s+1}^{n-p-s} (Y_{\pi_i+d} - \phi_{20} - \sum_{j=1}^{p_2} \phi_{2j} Y_{\pi_i+d-j} - \sum_{k=1}^{q_2} \theta_{2k} \hat{a}_{\pi_i+d-k}^{(2)})^2 \} \quad (5.3)$$

onde $\hat{a}_t^{(i)}$ é o estimador do resíduo $a_t^{(i)}$.

Substituindo $a_t^{(i)}$ por $\hat{a}_t^{(i)}$ em $x_{i,t}$ $i = 1, 2$, a função de verossimilhança pode ser reescrita como

$$\mathcal{L}(\gamma_1, \gamma_2, \tau_1, \tau_2, r, d|Y) \propto \tau_1^{n_1/2} \tau_2^{n_2/2} \exp\left\{\sum_{i=1}^2 \frac{-\tau_i}{2} (Y_i^* - X_i^{*'} \gamma_i)' (Y_i^* - X_i^{*'} \gamma_i)\right\} \quad (5.4)$$

onde n_1 , e n_2 são os números de observações do primeiro e segundo regimes, respectivamente.

Teorema 5.1 A distribuição a posteriori obtida de (5.2) e (5.4) é normal-gama e é dada por

$$\begin{aligned} P(\gamma_1, \gamma_2, \tau_1, \tau_2, r, d|Y) &\propto (\tau_1)^{(n_1+\alpha_1)/2-1} (\tau_2)^{(n_2+\alpha_2)/2-1} \times \\ &\exp\left\{\frac{-1}{2} \sum_{i=1}^2 [\tau_i(\beta_i + Y_i^{*'} Y_i^*) - (\tau_i X_i^{*'} Y_i^* + \mathbf{Q}_i \gamma_{0i})' (\tau_i X_i^{*'} X_i^{*'} + \mathbf{Q}_i)^{-1} (\tau_i X_i^{*'} Y_i^* + \mathbf{Q}_i \gamma_{0i}) \right. \\ &\quad \left. + [\gamma_i - (\tau_i X_i^{*'} X_i^{*'} + \mathbf{Q}_i)^{-1} (\tau_i X_i^{*'} Y_i^* + \mathbf{Q}_i \gamma_{0i})]' (\tau_i X_i^{*'} X_i^{*'} + \mathbf{Q}_i) \right. \\ &\quad \left. \times [\gamma_i - (\tau_i X_i^{*'} X_i^{*'} + \mathbf{Q}_i)^{-1} (\tau_i X_i^{*'} Y_i^* + \mathbf{Q}_i \gamma_{0i})]]\right\} \end{aligned} \quad (5.5)$$

Demonstração: A demonstração é análoga à do Teorema 4.1, modificando-se o fato de termos precisões τ_1 e τ_2 . $\square$

Para fazermos inferência sobre os parâmetros precisamos determinar as distribuições marginais a posteriori. Para cada parâmetro, a distribuição marginal é obtida integrando (5.5) em relação aos outros parâmetros. Neste caso em particular, estas distribuições não são facilmente obtidas. Este problema pode ser resolvido através do Amostrador de Gibbs (definido no Capítulo 3). Para aplicarmos o Amostrador de Gibbs e posteriormente estimarmos os parâmetros precisamos das distribuições a posteriori de cada parâmetro condicionado a todos os outros. Estas serão dadas pelo Corolário 5.1, cuja demonstração assim como as dos próximos Corolários serão omitidas por serem similares às do capítulo anterior.

Corolário 5.1 As distribuições a posteriori de cada parâmetro, condicionado a todos os outros, são dadas por

$$P(\gamma_i | \tau_1, \tau_2, r, d, Y) \propto \exp\{(-1/2)[\gamma_i - (\tau_i X_i^* X_i^{*'} + Q_i)^{-1}(\tau_i X_i^* Y_i^* + Q_i \gamma_{0i})]'$$

$$\times (\tau_i X_i^* X_i^{*'} + Q_i)[\gamma_i - (\tau_i X_i^* X_i^{*'} + Q_i)^{-1}(\tau_i X_i^* Y_i^* + Q_i \gamma_{0i})]\},$$

isto é,

$$\gamma_i | \tau_1, \tau_2, r, d, Y \sim N_{p_i+q_i+1}((\tau_i X_i^* X_i^{*'} + Q_i)^{-1}(\tau_i X_i^* Y_i^* + Q_i \gamma_{0i}), (\tau_i X_i^* X_i^{*'} + Q_i)^{-1}),$$

$i = 1, 2,$

$$P(\tau_i|\gamma_1, \gamma_2, r, d, Y) \propto (\tau_i)^{(n_i+\alpha_i)/2-1} \exp\left\{\frac{-\tau_i}{2}[\beta_i + (Y_i^* - X_i^{*'}\gamma_i)'(Y_i^* - X_i^{*'}\gamma_i)]\right\},$$

isto é,

$$\tau_i|\gamma_1, \gamma_2, r, d, Y \sim \text{Gama}((n_i+\alpha_i)/2, (\beta_i + (Y_i^* - X_i^{*'}\gamma_i)'(Y_i^* - X_i^{*'}\gamma_i))/2), i = 1, 2.$$

Corolário 5.2 A função de probabilidade a posteriori condicional de r é dada por

$$P(r|\gamma_1, \gamma_2, \tau_1, \tau_2, d, Y) \propto \tau_1^{n_1/2} \tau_2^{n_2/2} \exp\left\{(-1/2) \sum_{i=1}^2 \tau_i (Y_i^* - X_i^{*'}\gamma_i)'(Y_i^* - X_i^{*'}\gamma_i)\right\}$$

sendo n_1 e n_2 funções de r .

Corolário 5.3 A função de probabilidade a posteriori condicional de d é uma multinomial com probabilidade

$$p(d|\gamma_1, \gamma_2, \tau_1, \tau_2, r, Y) = \frac{\mathcal{L}(\gamma_1, \gamma_2, \tau_1, \tau_2, r, d|Y)}{\sum_{d=1}^{d_0} \mathcal{L}(\gamma_1, \gamma_2, \tau_1, \tau_2, r, d|Y)}$$

onde $\mathcal{L}(\gamma_1, \gamma_2, \tau_1, \tau_2, r, d|Y)$ é a função de verossimilhança dada em (5.4).

Observemos que para aplicarmos o Amostrador de Gibbs seria necessário conhecermos a distribuição de r dada no Corolário 5.2. Uma solução é aplicarmos o Algoritmo de Metropolis-Hastings(definido em 2.3), juntamente com o Amostrador de Gibbs.

5.3 Previsão

Nesta seção vamos encontrar a distribuição preditiva para previsão 1-passo à frente.

Sejam $\{Y_1, Y_2, Y_3, \dots, Y_n\}$ as n observações conhecidas e seja Y_{n+1} a observação a ser prevista.

O primeiro passo é determinarmos a qual regime Y_{n+1} pertence, ou seja, teremos Y_{n+1} pertencente ao primeiro regime se $Y_{n+1-d} \leq r$ e Y_{n+1} pertencente ao segundo caso contrário. Vamos supor, sem perda de generalidade, que $Y_{n+1-d} \leq r$.

Pela própria construção do modelo TARMA Rearranjado temos que a última observação do primeiro regime é Y_{π_s-d} e, portanto,

$$Y_{n+1} = \phi_{10} + \sum_{j=1}^{p_1} \phi_{1j} Y_{\pi_s+d-j} + a_{\pi_s+d}^{(1)} + \sum_{k=1}^{q_1} \theta_{1k} a_{\pi_s+d-k}^{(1)}.$$

A densidade de Y_{n+1} é dada por

$$P(Y_{n+1} | \gamma_1, \gamma_2, \tau_1, \tau_2, r, d) \propto \tau_1^{1/2} \exp\left\{-\frac{\tau_1}{2}(Y_{n+1} - (1, Y_{\pi_s+d-1}, Y_{\pi_s+d-2}, \dots, Y_{\pi_s+d-p_1}, a_{\pi_s+d-1}^{(1)}, a_{\pi_s+d-2}^{(1)}, \dots, a_{\pi_s+d-q_1}^{(1)})\gamma_1)^2\right\}$$

onde o vetor $(1, Y_{\pi_s+d-1}, Y_{\pi_s+d-2}, \dots, Y_{\pi_s+d-p_1}, a_{\pi_s+d-1}^{(1)}, a_{\pi_s+d-2}^{(1)}, \dots, a_{\pi_s+d-q_1}^{(1)})$ é a s -ésima linha da matriz $(X_1^*)'$. Portanto, considerando $I = (0, 0, 0, \dots, 0, 1)'$ um vetor $(p_1 + q_1 + 1) \times 1$ cujos todos os elementos são nulos com exceção do último, podemos reescrever a densidade de Y_{n+1} como

$$P(Y_{n+1}|\gamma_1, \gamma_2, \tau_1, \tau_2, r, d) \propto \tau_1^{1/2} \exp\left\{-\frac{\tau_1}{2}(Y_{n+1} - I'(X_1^*)'\gamma_1)^2\right\}. \quad (5.6)$$

A distribuição conjunta de $Y_{n+1}, \gamma_1, \gamma_2, \tau_1, \tau_2, r, d$ e Y é obtida das equações (5.5) e (5.6).

A inferência para Y_{n+1} será feita através do A.G. e para isto necessitamos da distribuição de Y_{n+1} dados todos os outros parâmetros e observações passadas. Temos

$$Y_{n+1}|\gamma_1, \gamma_2, \tau_1, \tau_2, r, d, Y \propto N(I'(X_1^*)'\gamma_1, \tau_1^{-1}).$$

5.4 Aplicação

Vamos ilustrar o método através de um exemplo onde os dados serão simulados. Geramos 200 observações para o modelo TARMA(2;1,1,1,1), dado por:

$$Y_t = \begin{cases} \phi_{11}Y_{t-1} + a_t^{(1)} + \theta_{11}a_{t-1}^{(1)}, & \text{se } Y_{t-d} \leq r \\ \phi_{21}Y_{t-1} + a_t^{(2)} + \theta_{21}a_{t-1}^{(2)}, & \text{se } Y_{t-d} > r \end{cases} \quad (5.7)$$

onde $\{a_t^{(i)}\}$ é uma sequência de variáveis aleatórias i.i.d. $\sim N(0, \tau_i^{-1}), i = 1, 2$ e $\{a_t^{(1)}\}$ independente de $\{a_t^{(2)}\}$.

Os valores escolhidos para os parâmetros são $\phi_{11} = 0,8, \theta_{11} = -0,5, \tau_1 = 0,5, \phi_{21} = -0,3, \theta_{21} = 0,7, \tau_2 = 1,0, r = 0,4$ e $d = 1$.

Para os hiperparâmetros consideramos $\gamma_{0i} = (0, 0, 0, 0)'$, $\alpha_i = 2, 0$, $\beta_i = 1$,
 $Q_i = \begin{pmatrix} 0, 1 & 0, 0 \\ 0, 0 & 0, 2 \end{pmatrix}$, $i = 1, 2$.

Suponhamos, também, que $a = \rho_{25}$ e $b = \rho_{75}$, onde ρ_k é o k-ésimo percentil para os dados.

Os resíduos foram estimados do modelo AR(2) ajustado para as observações e foram obtidos recursivamente de $\hat{a}_t = Y_t - \hat{\phi}_1 Y_{t-1} - \hat{\phi}_2 Y_{t-2}$, onde $\hat{\phi}_i$ é o estimador de mínimos quadrados de ϕ_i , $i = 1, 2$.

O Algoritmo de M.H. e o Amostrador de Gibbs foram usados para 5 cadeias com 5 valores iniciais distintos. Para o M.H. consideramos inicialmente a distribuição normal como núcleo de transição. Como a convergência da cadeia estava difícil de ocorrer decidimos por uma mudança da função núcleo e para tanto consideramos a distribuição logística. Optamos por uma taxa de aceitação em torno de 50% uma vez que os estudos para os quais já se conhece a taxa de aceitação não inclui a distribuição logística.

Iniciamos com 1000 iterações por cadeia, num total de 5000 observações. As Figuras 1-a e 1-b mostram gráficos e histogramas para a variável r , considerando diferentes números de iterações. Os histogramas apresentados na Figura 1-b indicam uma possível convergência das cadeias.

As inferências para os parâmetros e para previsão 1-passo à frente foram feitas considerando-se as 5000 iterações. Com a finalidade de diminuirmos o efeito dos valores iniciais desprezamos as 400 primeiras observações e tomamos das restantes uma a cada 15, para que os dados fossem aproximadamente independentes. Os resultados obtidos são apresentados na Tabela III-a. Para o parâmetro d , obtivemos 1 para a mediana e 0 para a medida de dispersão.

Observemos que o fator $\sqrt{\hat{R}} \approx 1$ para todos os parâmetros, isto é, a convergência ocorre e as amostras selecionadas são i.i.d.

Os resultados mostrados na Tabela III-a nos dizem que o M.H. e Amostrador de Gibbs produziram bons resultados. Os estimadores são robustos, uma vez que os mesmos resultados foram obtidos para outros valores iniciais.

Como pudemos comprovar a convergência através dos dois métodos escolhidos, poderíamos finalizar o exemplo neste ponto. Apesar disso vamos considerar 5000 iterações para cada cadeia num total de 25000. Com o objetivo de diminuirmos a influência dos valores iniciais desprezamos os 3500 primeiros valores gerados para cada parâmetro e consideramos os últimos 1500. Tomamos um valor a cada 15, desta forma temos 100 observações em cada cadeia num total de 500 para inferência. Os resultados obtidos são apresentados na Tabela III-b. A Figura 2 apresenta vários histogramas para a variável r para valores das iterações entre 24000 e 25000. Podemos comprovar a semelhança entre as Figuras 1-b e 2, comprovando assim que a convergência já havia ocorrido para as primeiras 5000 iterações.

Observemos que as variâncias estimadas com o método Bayesiano são grandes, o que é devido ao fato da metodologia Bayesiana incorporar a variação a posteriori dos parâmetros. Na Figura 3 apresentamos os gráficos para os dados gerados e para os parâmetros assim como seus respectivos histogramas.

Tabela III-a

Resultado da simulação com 200 observações obtidas das 5000 iterações,
sendo 1000 em cada cadeia.

parâmetros	valor real	média	d.p.	$\sqrt{\hat{R}}$
ϕ_{11}	0,8	0,776	0,246	1,0147
θ_{11}	-0,5	-0,422	0,263	0,9954
ϕ_{21}	-0,3	-0,276	0,216	1,0000
θ_{21}	0,7	0,738	0,241	1,0001
τ_1	0,5	0,431	0,066	1,0006
τ_2	1,0	0,891	0,208	1,0236
r	0,4	0,443	0,066	1,0101
Y_{n+1}	0,9067	0,9297	1,1224	1,0045

Tabela III-b

Resultado da simulação com 500 observações obtidas das 25000 iterações,
sendo 5000 em cada cadeia.

parâmetros	valor real	média	d.p.	$\sqrt{\hat{R}}$
ϕ_{11}	0,8	0,754	0,236	1,0019
θ_{11}	-0,5	-0,458	0,254	1,0028
ϕ_{21}	-0,3	-0,274	0,238	1,0029
θ_{21}	0,7	0,731	0,260	0,9985
τ_1	0,5	0,432	0,060	1,0053
τ_2	1,0	0,885	0,216	1,0011
r	0,4	0,448	0,063	1,0065
Y_{n+1}	0,9067	0,8673	1,1406	0,9998

Figura 1-a.

Gráficos de r para diferentes números de iterações

Figura 1-b.
Histogramas de r para diferentes números de iterações

Figura 2.
Histogramas de r para diferentes números de iterações

Figura 3. (a) Observações TARMA(2;1,1,1)
(b)-(h) Amostrador de Gibbs para os parâmetros

Capítulo 6

Modelo ASMA

Neste Capítulo vamos considerar os Modelos *ASMA*(1) e *ASMA*(2), para os quais a análise Bayesiana será feita considerando-se uma priori própria e a priori de Jeffreys(seção 6.1). A distribuição preditiva para previsão 1-passo à frente será encontrada na seção 6.2. A seção 6.3 é destinada a aplicações do método, para isto consideramos dados simulados e reais.

6.1 Análise do Modelo ASMA

Para os modelos *ASMA*(1) e *ASMA*(2), definidos em (3.4) e (3.5), respectivamente, vamos supor $\{a_t\}$ sequência i.i.d. $\sim N(0, \tau^{-1})$, onde $\tau^{-1} = \text{Var}(a_t) > 0$ é desconhecida. Os parâmetros para o modelo (3.4) são $\gamma = (\theta_1^+, \theta_1^-)'$ e τ e para o modelo (3.5) temos $\gamma = (\theta_1^+, \theta_1^-, \theta_2^+, \theta_2^-)'$ e τ .

Consideremos uma amostra com n observações $Y = \{Y_1, Y_2, \dots, Y_n\}$.

A verossimilhança para o modelo (3.4) é dada por:

$$\begin{aligned}\mathcal{L}(\theta_1^+, \theta_1^-, \tau|Y) &= \prod_{t=1}^n p(Y_t|\hat{a}_t, a_0, \theta_1^+, \theta_1^-, \tau) \\ &= \frac{\tau^{n/2}}{\sqrt{2\pi}^n} \exp\left\{-\frac{\tau}{2}\left[\sum_1 (Y_t - \theta_1^+ \hat{a}_{t-1})^2 + \sum_2 (Y_t - \theta_1^- \hat{a}_{t-1})^2\right]\right\}\end{aligned}$$

onde $\sum_1$ é somatória em $\{t = 1, \dots, n; a_{t-1} \geq 0\}$, $\sum_2$ é somatória em $\{t = 1, \dots, n; a_{t-1} < 0\}$, e os resíduos $\hat{a}_t$ são calculados recursivamente dado que $a_0 = 0$.

Desenvolvendo as expressões quadráticas dentro da exponencial obtemos

$$\begin{aligned}& \sum_1 (Y_t - \theta_1^+ \hat{a}_{t-1})^2 + \sum_2 (Y_t - \theta_1^- \hat{a}_{t-1})^2 \\ &= \sum_1 Y_t^2 + \sum_2 Y_t^2 - (\theta_1^+, \theta_1^-) \begin{pmatrix} \sum_1 Y_t \hat{a}_{t-1} \\ \sum_2 Y_t \hat{a}_{t-1} \end{pmatrix} - \left(\sum_1 Y_t \hat{a}_{t-1} \sum_2 Y_t \hat{a}_{t-1}\right) \begin{pmatrix} \theta_1^+ \\ \theta_1^- \end{pmatrix} \\ & \quad + (\theta_1^+, \theta_1^-) \begin{pmatrix} \sum_1 \hat{a}_{t-1}^2 & 0 \\ 0 & \sum_2 \hat{a}_{t-1}^2 \end{pmatrix} \begin{pmatrix} \theta_1^+ \\ \theta_1^- \end{pmatrix} \\ &= \sum_{t=1}^n Y_t^2 - \gamma' \mathbf{B} - \mathbf{B}' \gamma + \gamma' \mathbf{A} \gamma \\ &= \sum_{t=1}^n Y_t^2 + (\gamma - \mathbf{A}^{-1} \mathbf{B})' \mathbf{A} (\gamma - \mathbf{A}^{-1} \mathbf{B}) - \mathbf{B}' \mathbf{A}^{-1} \mathbf{B},\end{aligned}$$

onde

$$\mathbf{B} = \begin{pmatrix} \sum_1 Y_t \hat{a}_{t-1} \\ \sum_2 Y_t \hat{a}_{t-1} \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} \sum_1 \hat{a}_{t-1}^2 & 0 \\ 0 & \sum_2 \hat{a}_{t-1}^2 \end{pmatrix}.$$

Portanto, a função de verossimilhança pode ser reescrita como

$$\mathcal{L}(\theta_1^+, \theta_1^-, \tau|Y) \propto \tau^{n/2} \exp\left\{-\frac{\tau}{2}\left[\sum_{t=1}^n Y_t^2 - \mathbf{B}' \mathbf{A}^{-1} \mathbf{B} + (\gamma - \mathbf{A}^{-1} \mathbf{B})' \mathbf{A} (\gamma - \mathbf{A}^{-1} \mathbf{B})\right]\right\}. \quad (6.1)$$

Para o modelo(3.5), a função de verossimilhança é encontrada de maneira análoga:

$$\begin{aligned}\mathcal{L}(\gamma, \tau|Y) &= \prod_{t=1}^n p(Y_t|\hat{a}_{t-1}, a_0, a_{-1}, \gamma, \tau) \\ &\propto \tau^{n/2} \exp\left\{-\frac{\tau}{2}\left[\sum_1 (Y_t - \theta_1^+ \hat{a}_{t-1} - \theta_2^+ \hat{a}_{t-2})^2 + \sum_2 (Y_t - \theta_1^+ \hat{a}_{t-1} - \theta_2^- \hat{a}_{t-2})^2\right.\right. \\ &\quad \left.+\sum_3 (Y_t - \theta_1^- \hat{a}_{t-1} - \theta_2^+ \hat{a}_{t-2})^2 + \sum_4 (Y_t - \theta_1^- \hat{a}_{t-1} - \theta_2^- \hat{a}_{t-2})^2\right]\} \\ &\propto \tau^{n/2} \exp\left\{-\frac{\tau}{2}\left[\sum_{t=1}^n Y_t^2 - \mathbf{B}'\mathbf{A}^{-1}\mathbf{B} + (\gamma - \mathbf{A}^{-1}\mathbf{B})'\mathbf{A}(\gamma - \mathbf{A}^{-1}\mathbf{B})\right]\right\}, \quad (6.2)\end{aligned}$$

supondo $a_0 = a_{-1} = 0$ e, onde $\sum_1$ é a somatória em $\{t = 1, \dots, n; a_{t-1} \geq 0, a_{t-2} \geq 0\}$, $\sum_2$ é a somatória em $\{t = 1, \dots, n; a_{t-1} \geq 0, a_{t-2} < 0\}$, $\sum_3$ é a somatória em $\{t = 1, \dots, n; a_{t-1} < 0, a_{t-2} \geq 0\}$, $\sum_4$ é a somatória em $\{t = 1, \dots, n; a_{t-1} < 0, a_{t-2} < 0\}$,

$$\mathbf{A} = \begin{pmatrix} \sum_1 \hat{a}_{t-1}^2 + \sum_2 \hat{a}_{t-2}^2 & 0 & 2\sum_1 \hat{a}_{t-1}\hat{a}_{t-2} & 2\sum_2 \hat{a}_{t-1}\hat{a}_{t-2} \\ 0 & \sum_3 \hat{a}_{t-1}^2 + \sum_4 \hat{a}_{t-2}^2 & 2\sum_3 \hat{a}_{t-1}\hat{a}_{t-2} & 2\sum_4 \hat{a}_{t-1}\hat{a}_{t-2} \\ 0 & 0 & \sum_1 \hat{a}_{t-2}^2 + \sum_3 \hat{a}_{t-2}^2 & 0 \\ 0 & 0 & 0 & \sum_2 \hat{a}_{t-2}^2 + \sum_4 \hat{a}_{t-2}^2 \end{pmatrix}$$

e

$$\mathbf{B} = \begin{pmatrix} \sum_1 Y_t \hat{a}_{t-1} + \sum_2 Y_t \hat{a}_{t-1} \\ \sum_3 Y_t \hat{a}_{t-1} + \sum_4 Y_t \hat{a}_{t-1} \\ \sum_1 Y_t \hat{a}_{t-2} + \sum_3 Y_t \hat{a}_{t-2} \\ \sum_2 Y_t \hat{a}_{t-2} + \sum_4 Y_t \hat{a}_{t-2} \end{pmatrix}.$$

Consideremos uma distribuição a priori própria para os parâmetros γ e τ , de modo que

$$\gamma|\tau \sim \text{Normal}(\gamma_0, \tau^{-1}\mathbf{Q}^{-1})$$

e

$$\tau \sim \text{Gama}(\alpha/2, \beta/2).$$

A distribuição a priori é dada por:

$$P(\gamma, \tau) \propto \tau^{\alpha/2-1} \exp\left\{-\frac{\tau}{2}[\beta + (\gamma - \gamma_0)' \mathbf{Q}(\gamma - \gamma_0)]\right\}. \quad (6.3)$$

Teorema 6.1 De (6.1) ou (6.2) e (6.3), temos que a distribuição a posteriori é uma normal-gama dada por

$$\begin{aligned} P(\gamma, \tau|Y) \propto & \tau^{(n+\alpha)/2-1} \exp\left\{-\frac{\tau}{2}\left[\beta + \sum_{t=1}^n Y_t^2 - \right. \right. \\ & (\mathbf{B} + \mathbf{Q}\gamma_0)'(\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q}\gamma_0) + \gamma_0' \mathbf{Q}\gamma_0 + \\ & \left. \left. + (\gamma - (\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q}\gamma_0))'(\mathbf{A} + \mathbf{Q})(\gamma - (\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q}\gamma_0))\right]\right\}. \quad (6.4) \end{aligned}$$

Demonstração: Será omitida por seguir exatamente os mesmos passos da demonstração do Teorema 4.1.

Corolário 6.1 As distribuições marginais a posteriori obtidas por integração de (6.4) são:

$$P(\tau|Y) \propto \tau^{(n+\alpha-2q)/2-1} \exp\left\{-\frac{\tau}{2}\left[\beta + \sum_{t=1}^n Y_t^2 + \gamma_0' \mathbf{Q}\gamma_0\right]\right\}$$

$$-(\mathbf{B} + \mathbf{Q}\gamma_0)'(\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q}\gamma_0)]\}$$

isto é,

$$\tau|Y \sim \text{Gama}\left(\frac{n + \alpha - 2^q}{2}, \frac{\beta + \sum_{t=1}^n Y_t^2 + \gamma_0' \mathbf{Q} \gamma_0 - (\mathbf{B} + \mathbf{Q}\gamma_0)'(\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q}\gamma_0)}{2}\right). \quad (6.5)$$

$$P(\gamma|Y) \propto \{1 + (\gamma - (\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q}\gamma_0))' \frac{\mathbf{W}}{\nu} (\gamma - (\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q}\gamma_0))\}^{-\frac{\nu+2^q}{2}},$$

onde

$$\mathbf{W} = \nu[\beta + \sum_{t=1}^n Y_t^2 + \gamma_0' \mathbf{Q} \gamma_0 - (\mathbf{B} + \mathbf{Q}\gamma_0)'(\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q}\gamma_0)]^{-1}(\mathbf{A} + \mathbf{Q})$$

e

$$\nu = n + \alpha - 2^q,$$

isto é,

$$\gamma|Y \sim t_{2^q}(\nu, (\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q}\gamma_0), \mathbf{W}). \quad (6.6)$$

Teorema 6.2 Usando-se a priori de Jeffreys , $P(\gamma, \tau) \propto \tau^{-1}$, temos que a distribuição a posteriori é normal-gama e é dada por

$$P(\gamma, \tau|Y) \propto \tau^{(n-2^q)/2-1} \exp\left\{-\frac{\tau}{2}\left[\sum_{t=1}^n Y_t^2 - \mathbf{B}'\mathbf{A}^{-1}\mathbf{B} + (\gamma - \mathbf{A}^{-1}\mathbf{B})'\mathbf{A}(\gamma - \mathbf{A}^{-1}\mathbf{B})\right]\right\}. \quad (6.7)$$

Corolário 6.2 As distribuições marginais a posteriori obtidas de (6.7) são:

$$P(\tau|Y) \propto \tau^{(n-2^{q+1})/2-1} \exp\left\{-\frac{\tau}{2}\left[\sum_{t=1}^n Y_t^2 - \mathbf{B}'\mathbf{A}^{-1}\mathbf{B}\right]\right\}$$

isto é,

$$\tau|Y \sim \text{Gama}\left(\frac{n-2^{q+1}}{2}, \frac{\sum_{t=1}^n Y_t^2 - \mathbf{B}'\mathbf{A}^{-1}\mathbf{B}}{2}\right). \quad (6.8)$$

$$P(\gamma|Y) \propto (1 + (\gamma - \mathbf{A}^{-1}\mathbf{B})' \frac{\mathbf{W}}{\nu'} (\gamma - \mathbf{A}^{-1}\mathbf{B}))^{-(\nu'+2^q)/2},$$

com

$$\mathbf{W} = \nu' \left(\sum_{t=1}^n Y_t^2 - \mathbf{B}'\mathbf{A}\mathbf{B} \right)^{-1} \mathbf{A} \text{ e } \nu' = n - 2^{2q+1},$$

isto é,

$$\gamma|Y \sim t_{2^q}(\nu', \mathbf{A}^{-1}\mathbf{B}, \mathbf{W}), \quad (6.9)$$

onde q é a ordem do modelo ASMA.

Observemos que para se obter (6.8) e (6.9) basta fazermos $\mathbf{Q} \rightarrow \mathbf{0}, \beta \rightarrow 0$ e $\alpha \rightarrow -2^q$ na distribuição conjunta a posteriori (6.4).

6.2 Distribuição preditiva

Vamos determinar a densidade preditiva para previsão um passo à frente, Y_{n+1} , dado que temos as observações $Y_1, \dots, Y_n$.

Observemos que para o modelo ASMA(1) temos

$$Y_{n+1} = \begin{cases} a_{n+1} + \theta_1^+ a_n, & \text{se } a_n \geq 0, \\ a_{n+1} + \theta_1^- a_n, & \text{se } a_n < 0, \end{cases} \quad (6.10)$$

e, portanto,

$$a_{n+1} = \begin{cases} Y_{n+1} - I_1' \gamma, & \text{se } a_n \geq 0, \\ Y_{n+1} - I_2' \gamma, & \text{se } a_n < 0, \end{cases}$$

onde $I_1 = (a_n, 0)'$, $I_2 = (0, a_n)'$ e $\gamma = (\theta_1^+, \theta_1^-)'$.

Suponhamos que $a_n \geq 0$. (Se $a_n < 0$, o procedimento é análogo.)

A função de verossimilhança aproximada de Y_{n+1} condicionada às observações passadas é dada por

$$\mathcal{L}(Y_{n+1}|\tau, \gamma, Y) \propto \tau^{1/2} \exp\left\{-\frac{\tau}{2}(Y_{n+1} - I_1' \gamma)^2\right\}. \quad (6.11)$$

Combinando (6.11) com (6.4) obtemos

$$\begin{aligned} P(\gamma, \tau, Y_{n+1}|Y) &\propto \tau^{(n+\alpha+1)/2-1} \exp\left\{-\frac{\tau}{2}[(Y_{n+1} - I_1' \gamma)^2 + \beta + \sum_{t=1}^n Y_t^2 \right. \\ &\quad \left. + \gamma_0' \mathbf{Q} \gamma_0 - (\mathbf{B} + \mathbf{Q} \gamma_0)'(\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q} \gamma_0) \right. \\ &\quad \left. + (\gamma - (\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q} \gamma_0))'(\mathbf{A} + \mathbf{Q})(\gamma - (\mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q} \gamma_0))\right\}. \end{aligned} \quad (6.12)$$

A densidade preditiva de Y_{n+1} , dado Y é obtida de (6.12) após completarmos o quadrado em γ e Y_{n+1} e integrarmos em relação a γ e τ .

Teorema 6.3 A densidade preditiva de Y_{n+1} dado Y é uma t-univariada com $\nu_p = n + \alpha - 2^q$ graus de liberdade (neste caso $q = 1$), valor esperado

$$E(Y_{n+1}|Y) = (1 - I_1'(I_1 I_1' + \mathbf{A} + \mathbf{Q})^{-1} I_1)^{-1} (I_1'(I_1 I_1' + \mathbf{A} + \mathbf{Q})^{-1})(\mathbf{B} + \mathbf{Q} \gamma_0)$$

e matriz de precisão

$$\begin{aligned} \mathbf{W} &= \nu_p [\beta + \sum_{t=1}^n Y_t^2 + \gamma_0' \mathbf{Q} \gamma_0 - (\mathbf{B} + \mathbf{Q} \gamma_0)'(I_1 I_1' + \mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q} \gamma_0) \\ &\quad - (I_1'(I_1 I_1' + \mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q} \gamma_0))'(1 - I_1'(I_1 I_1' + \mathbf{A} + \mathbf{Q})^{-1} I_1)^{-1} \end{aligned}$$

$$\times (I_1'(I_1 I_1' + \mathbf{A} + \mathbf{Q})^{-1}(\mathbf{B} + \mathbf{Q}\gamma_0))].$$

Teorema 6.4 Se estivermos considerando a priori de Jeffreys, a densidade preditiva de Y_{n+1} dado Y é uma t-univariada com $\nu_p = n - 2^{q+1}$, valor esperado

$$E(Y_{n+1}) = (1 - I_1'(I_1 I_1' + \mathbf{A})^{-1} I_1)^{-1} (I_1'(I_1 I_1' + \mathbf{A})^{-1} \mathbf{B})$$

e precisão

$$\begin{aligned} \mathbf{W} = & \nu_p \left[\sum_{t=1}^n Y_t^2 - \mathbf{B}'(I_1 I_1' + \mathbf{A})^{-1} \mathbf{B} \right. \\ & \left. + (I_1'(I_1 I_1' + \mathbf{A})^{-1} \mathbf{B})'(1 - I_1'(I_1 I_1' + \mathbf{A})^{-1} I_1)^{-1} (I_1'(I_1 I_1' + \mathbf{A})^{-1} \mathbf{B}) \right]. \end{aligned}$$

Para fazermos previsão para o modelo ASMA(2) observemos que

$$Y_{n+1} = \begin{cases} a_{n+1} + \theta_1^+ a_n + \theta_2^+ a_{n-1}, & \text{se } a_n \geq 0 \text{ e } a_{n-1} \geq 0 \\ a_{n+1} + \theta_1^+ a_n + \theta_2^- a_{n-1}, & \text{se } a_n \geq 0 \text{ e } a_{n-1} < 0 \\ a_{n+1} + \theta_1^- a_n + \theta_2^+ a_{n-1}, & \text{se } a_n < 0 \text{ e } a_{n-1} \geq 0 \\ a_{n+1} + \theta_1^- a_n + \theta_2^- a_{n-1}, & \text{se } a_n < 0 \text{ e } a_{n-1} < 0 \end{cases}$$

ou ainda,

$$Y_{n+1} = \begin{cases} a_{n+1} + I_1' \gamma, & \text{se } a_n \geq 0 \text{ e } a_{n-1} \geq 0 \\ a_{n+1} + I_2' \gamma, & \text{se } a_n \geq 0 \text{ e } a_{n-1} < 0 \\ a_{n+1} + I_3' \gamma, & \text{se } a_n < 0 \text{ e } a_{n-1} \geq 0 \\ a_{n+1} + I_4' \gamma, & \text{se } a_n < 0 \text{ e } a_{n-1} < 0 \end{cases},$$

onde $I_1 = (a_n, 0, a_{n-1}, 0)'$, $I_2 = (a_n, 0, 0, a_{n-1})'$, $I_3 = (0, a_n, a_{n-1}, 0)'$, $I_4 = (0, a_n, 0, a_{n-1})'$ e $\gamma = (\theta_1^+, \theta_1^-, \theta_2^+, \theta_2^-)'$.

Neste caso, o procedimento para encontrarmos a preditiva é análogo ao usado para o modelo ASMA(1).

6.3 Aplicações

Exemplo 1: Análise para ASMA(1)

Geramos 3001 observações para o modelo ASMA(1) considerando $\theta_1^+ = 0,5, \theta_1^- = -0,4$ e $\tau = 1$.

As 3000 primeiras observações foram usadas para a inferência dos parâmetros e a última para compararmos com a previsão 1-passo à frente.

Considerando a priori de Jeffreys, $p(\gamma, \tau) \propto \tau^{-1}$, onde $\gamma = (\theta_1^+, \theta_1^-)'$, do Corolário (6.2) obtemos:

$$\tau \sim \text{Gama}\left(\frac{2996}{2}, \frac{2999,8}{2}\right)$$

e

$$\gamma \sim t_2(2992, \mathbf{M}, \mathbf{W}),$$

onde

$$\mathbf{M} = \begin{pmatrix} 0,4894 \\ -0,4061 \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} 1470,3 & 0 \\ 0 & 1522,7 \end{pmatrix}.$$

A Tabela IV mostra os valores médios e os desvios padrão para os parâmetros.

A previsão 1-passo à frente, $\hat{Y}_{3001}$, foi calculada usando-se o Teorema 6.3.

O resultado se encontra na Tabela V.

Tabela IV-Média e desvio padrão para os parâmetros

parâmetro	média	d.p.
τ	0,9987	0,0258
θ_1^+	0,4894	0,0261
θ_1^-	-0,4061	0,0256

Tabela V-Previsão 1-passo à frente

	Valor Real	Previsão	d.p.
Y_{3001}	0,1596	0,1217	0,0183

Exemplo 2: Análise para ASMA(2)

Geramos 3001 observações para o modelo ASMA(2), supondo que $\theta_1^+ = 0,5$; $\theta_1^- = 0,6$; $\theta_2^+ = -0,4$; $\theta_2^- = -0,5$ e $\tau = 1$.

Para a análise consideramos a priori de Jeffreys e as 3000 primeiras observações.

Do Corolário 6.2 obtemos

$$\tau \sim \text{Gama}(1496; 1496, 9)$$

e

$$\gamma = (\theta_1^+, \theta_1^-, \theta_2^+, \theta_2^-)' \sim t_4(2968, \mathbf{D}, \mathbf{W}),$$

onde

$$\mathbf{D} = \begin{pmatrix} 0,4828 \\ 0,6002 \\ -0,4621 \\ -0,4939 \end{pmatrix}, \quad \mathbf{W} = 10^3 \times \begin{pmatrix} 1,4903 & 0 & 0,9268 & -0,9775 \\ 0 & 1,4691 & -0,9573 & 0,9589 \\ 0 & 0 & 1,4606 & 0 \\ 0 & 0 & 0 & 1,5135 \end{pmatrix}.$$

A Tabela VI mostra os valores médios e os desvios padrões para os parâmetros, e a Tabela VII apresenta a previsão 1-passo à frente.

Tabela VI- Média e desvio padrão para os parâmetros

parâmetros	médias	d.p.
θ_1^+	0,4828	0,0259
θ_1^-	0,6002	0,0261
θ_2^+	-0,4621	0,0262
θ_2^-	-0,4939	0,0257
τ	0,9994	0,0258

Tabela VII-Previsão 1-passo à frente

	Valor Real	Previsão	d.p.
Y_{3001}	-0,5017	-0,5976	0,0183

Exemplo 3:

As aplicações para dados reais foram feitas com as séries

BLS Tin Plate

NBER Tin Plate

BLS Regular Gas

NBER Regular Gas

BLS CSS+S-Hot Rolled

NBER CSS+S-Hot Rolled,

obtidas de Stigler e Kindahl(1970). As séries Bureau of Labor Statistics (BLS) são séries de preços cotados e as séries National Bureau of Economic Research (NBER) são séries de preços de transação. Estas séries podem ser encontradas no Apêndice 2.

Segundo Stigler e Kindahl as séries BLS são assimétricas e as NBER são simétricas. Assim como Wecker(1981), fazemos uma primeira diferença nos dados e ajustamos para cada série um modelo simétrico MA(1) e um assimétrico ASMA(1). Os coeficientes dos modelos MA(1) foram estimados usando-se o pacote ITSM(Brockwell e Davis,1991b) e os coeficientes dos modelos ASMA(1) foram obtidos do Corolário 6.2 onde consideramos a priori de Jeffreys. Embora tenhamos ajustados estes modelos para os dados não significa serem estes os modelos mais adequados. Os resultados que obtivemos estão na tabela VIII-a e os encontrados por Wecker estão na tabela VIII-b.

Tabela VIII-a

Séries	Simétrico	Assimétrico	
	θ	θ_1^+	θ_1^-
CSS+S Hot Rolled			
BLS	0,0418	0,0199	-0,1952
NBER	-0,052	-0,0733	0,0768
Tin Plate			
BLS	-0,0161	0	-0,8864
NBER	0,0961	0,1361	-0,0253
Regular Gas			
BLS	0,2066	0,4723	0,1771
NBER	0,2509	0,0224	0,1754

Tabela VIII-b
Resultados de Wecker

Séries	Simétrico	Assimétrico	
	θ	θ_1^+	θ_1^-
CSS+S Hot Rolled			
BLS	0,05	0	-0.95
NBER	-0,05	-0,05	0,15
Tin Plate			
BLS	0	0	-1.10
NBER	0,10	0,20	-0,05
Regular Gas			
BLS	0,20	0,45	-0.55
NBER	0,25	0	0,40

Capítulo 7

Modelo RCA

Este Capítulo é destinado a análise Bayesiana do modelo Autoregressivo com coeficientes aleatórios normalmente distribuídos.

Vamos considerar o modelo RCA(1) definido por Diaz(1990). O trabalho de Diaz tem como objetivo fazer previsão 1-passo à frente. A nossa análise é feita através do Amostrador de Gibbs e além da inferência sobre os parâmetros faremos previsão 1-passo à frente.

Para finalizar apresentamos um exemplo numérico, onde os dados são simulados.

7.1 Análise a posteriori

A função de verossimilhança para o modelo (3.7) é dada por

$$\begin{aligned}\mathcal{L}(\theta, \tau, \xi, \mu|Y) &\propto \tau^{n/2} \exp\left\{-\frac{\tau}{2} \left[\sum_{t=1}^n (Y_t - \theta_t Y_{t-1})^2\right]\right\} (\tau\xi)^{n/2} \exp\left\{-\frac{\tau\xi}{2} \left[\sum_{t=1}^n (\theta_t - \mu)^2\right]\right\} \\ &\propto \tau^n \xi^{n/2} \exp\left\{-\frac{\tau}{2} \left[\sum_{t=1}^n (Y_t - \theta_t Y_{t-1})^2\right] + \xi \sum_{t=1}^n (\theta_t - \mu)^2\right\},\end{aligned}\quad (7.1)$$

onde $\theta = (\theta_1, \theta_2, \dots, \theta_n)'$, $Y = \{Y_1, Y_2, \dots, Y_n\}$ é o vetor de todas as observações e Y_0 é suposto fixado.

Para os parâmetros μ, τ e ξ consideremos as prioris independentes $\tau \sim \text{Gama}(\alpha, \beta)$, $\mu \sim \text{Uniforme}$ e para ξ a priori imprópria de Jeffreys, ou seja,

$$P(\tau) \propto \tau^{\alpha-1} \exp\{-\tau\beta\}, \quad P(\xi) \propto \frac{1}{\xi}.$$

A priori conjunta é dada por

$$P(\mu, \tau, \xi) \propto \xi^{-1} \tau^{\alpha-1} \exp\{-\beta\tau\}. \quad (7.2)$$

Teorema 7.1 A distribuição a posteriori conjunta obtida das equações (7.1) e (7.2) é dada por

$$P(\theta, \tau, \xi, \mu|Y) \propto \tau^{n+\alpha-1} \xi^{n/2-1} \exp\left\{-\frac{\tau}{2} \left[2\beta + \sum_{t=1}^n (Y_t - \theta_t Y_{t-1})^2 + \xi \sum_{t=1}^n (\theta_t - \mu)^2\right]\right\}. \quad (7.3)$$

As inferências sobre os parâmetros serão feitas com o uso do Amostrador de Gibbs, e para isto necessitamos das distribuições a posteriori condicionais completas.

Teorema 7.2 As distribuições condicionais completas são dadas por:

$$\tau|\theta, \xi, \mu, Y \sim \text{Gama}(n + \alpha, \frac{1}{2}(\sum_{t=1}^n (Y_t - \theta_t Y_{t-1})^2 + \xi \sum_{t=1}^n (\theta_t - \mu)^2 + 2\beta)), \quad (7.4)$$

$$\xi|\tau, \theta, \xi, \mu, Y \sim \text{Gama}(\frac{n}{2}, \frac{1}{2}\tau(\sum_{t=1}^n (\theta_t - \mu)^2)), \quad (7.5)$$

$$\mu|\theta, \tau, \xi, Y \sim N(\frac{1}{n} \sum_{t=1}^n \theta_t, \frac{(\tau\xi)^{-1}}{n}), \quad (7.6)$$

$$\theta_t|\mu, \tau, \xi, Y \sim N(\frac{(Y_t Y_{t-1} + \xi\mu)}{(Y_{t-1}^2 + \xi)}, (\tau(Y_{t-1}^2 + \xi))^{-1}), \quad (7.7)$$

Demonstração:.

As distribuições de $\tau|\theta, \mu, \xi, Y$ e $\xi|\theta, \mu, \tau, Y$ são imediatas, já que

$$P(\tau|\theta, \mu, \xi, Y) \propto \tau^{n+\alpha-1} \exp\{-\frac{\tau}{2}[\sum_{t=1}^n (Y_t - \theta_t Y_{t-1})^2 + \xi \sum_{t=1}^n (\theta_t - \mu)^2 + 2\beta]\}$$

e

$$P(\xi|\theta, \mu, \tau, Y) \propto \xi^{n/2-1} \exp\{-\frac{\tau\xi}{2}[\sum_{t=1}^n (\theta_t - \mu)^2]\}.$$

Para $\mu|\theta, \tau, \xi, Y$ observemos que

$$\begin{aligned}
P(\mu|\theta, \tau, \xi, Y) &\propto \exp\left\{-\frac{\tau\xi}{2}\left[\sum_{t=1}^n(\theta_t - \mu)^2\right]\right\} \\
&\propto \exp\left\{-\frac{\tau\xi}{2}\left[\sum_{t=1}^n \theta_t^2 - 2\mu \sum_{t=1}^n \theta_t + n\mu^2\right]\right\} \\
&\propto \exp\left\{-\frac{\tau\xi}{2}\left[\left(\mu^2 - \frac{2\mu \sum_{t=1}^n \theta_t}{n}\right)n\right]\right\} \\
&\propto \exp\left\{-\frac{\tau\xi n}{2}\left(\mu - \frac{\sum_{t=1}^n \theta_t}{n}\right)^2 - \left(\frac{\sum_{t=1}^n \theta_t}{n}\right)^2\right\} \\
&\propto \exp\left\{-\frac{\tau\xi n}{2}\left(\mu - \frac{\sum_{t=1}^n \theta_t}{n}\right)^2\right\}.
\end{aligned}$$

Para o parâmetro θ temos que

$$\begin{aligned}
P(\theta|\mu, \tau, \xi, Y) &\propto \exp\left\{-\frac{\tau}{2}\left[\sum_{t=1}^n[(Y_t - \theta_t Y_{t-1})^2 + \xi(\theta_t - \mu)^2]\right]\right\} \\
&\propto \exp\left\{-\frac{\tau}{2}\left[\sum_{t=1}^n Y_t^2 - 2\sum_{t=1}^n Y_t Y_{t-1} \theta_t + \sum_{t=1}^n Y_{t-1}^2 \theta_t^2 + \xi \sum_{t=1}^n \theta_t^2 - 2\xi\mu \sum_{t=1}^n \theta_t + n\xi\mu^2\right]\right\} \\
&\propto \exp\left\{-\frac{\tau}{2}\left[\sum_{t=1}^n[(Y_{t-1}^2 + \xi)\theta_t^2 - 2(Y_t Y_{t-1} + \xi\mu)\theta_t]\right]\right\} \\
&\propto \left\{-\frac{\tau}{2}\left[\sum_{t=1}^n\left(\theta_t^2 - 2\frac{(Y_t Y_{t-1} + \xi\mu)}{(Y_{t-1}^2 + \xi)}\theta_t\right)(Y_{t-1}^2 + \xi)\right]\right\}
\end{aligned}$$

$$\propto \exp\left\{-\frac{\tau}{2}\left[\sum_{t=1}^n\left(\theta_t - \frac{(Y_t Y_{t-1} + \xi\mu)}{(Y_{t-1}^2 + \xi)}\right)^2(Y_{t-1}^2 + \xi) - \sum_{t=1}^n \frac{(Y_t Y_{t-1} + \xi\mu)^2}{(Y_{t-1}^2 + \xi)}\right]\right\}$$

$$\propto \exp\left\{-\frac{\tau}{2}\left[\sum_{t=1}^n\left(\theta_t - \frac{(Y_t Y_{t-1} + \xi\mu)}{(Y_{t-1}^2 + \xi)}\right)^2(Y_{t-1}^2 + \xi)\right]\right\}$$

$$\propto \prod_{t=1}^n \exp\left\{-\frac{\tau(Y_{t-1}^2 + \xi)}{2}\left(\theta_t - \frac{(Y_t Y_{t-1} + \xi\mu)}{(Y_{t-1}^2 + \xi)}\right)^2\right\}.$$

Observemos que $\theta_t|\mu$ são condicionalmente independentes e, portanto, para $t = 1, \dots, n$, temos

$$P(\theta_t|\mu, \tau, \xi, Y) \propto \exp\left\{-\frac{\tau(Y_{t-1}^2 + \xi)}{2}\left(\theta_t - \frac{(Y_t Y_{t-1} + \xi\mu)}{(Y_{t-1}^2 + \xi)}\right)^2\right\} \square$$

7.2 Densidade preditiva

Nesta seção vamos obter a densidade de Y_{n+1} condicionada a todos os parâmetros e observações passadas de modo a utilizarmos o Amostrador de Gibbs e fazermos previsão 1-passo à frente.

A densidade conjunta de Y_{n+1} e θ_{n+1} dados θ, τ, ξ e μ é dada por

$$P(Y_{n+1}, \theta_{n+1}|\theta, \tau, \xi, \mu) \propto \tau \xi^{1/2} \exp\left\{\frac{-\tau}{2}[(Y_{n+1} - \theta_{n+1} Y_n)^2 + \xi(\theta_{n+1} - \mu)^2]\right\}. \quad (7.8)$$

Combinando (7.3) com (7.8) obtemos a densidade conjunta

$$P(Y_{n+1}, \theta_{n+1}, \theta, \tau, \xi, \mu|Y) \propto \tau^{n+\alpha} \xi^{\frac{(n+1)}{2}-1} \exp\left\{\frac{-\tau}{2}\left[\sum_{t=1}^{n+1}(Y_t - \theta_t Y_{t-1})^2\right]\right\}$$

$$+ \xi \sum_{t=1}^{n+1} (\theta_t - \mu)^2 + 2\beta\}.$$

A posteriori de Y_{n+1} , dados $\theta_{n+1}, \theta, \tau, \xi, \mu$ e Y é dada por

$$\begin{aligned} P(Y_{n+1} | \theta, \theta_{n+1}, \tau, \xi, \mu, Y) &\propto \exp\left\{\frac{-\tau}{2} \left[\sum_{t=1}^{n+1} (Y_t - \theta_t Y_{t-1})^2\right]\right\} \\ &\propto \exp\left\{\frac{-\tau}{2} [(Y_{n+1} - \theta_{n+1} Y_n)^2 + \sum_{t=1}^n (Y_t - \theta_t Y_{t-1})^2]\right\} \end{aligned}$$

$$\propto \exp\left\{\frac{-\tau}{2} [(Y_{n+1} - \theta_{n+1} Y_n)^2]\right\}.$$

Portanto, temos que

$$Y_{n+1} | \theta, \theta_{n+1}, \tau, \xi, \mu, Y \sim N(\theta_{n+1} Y_n, \tau^{-1}).$$

7.3 Aplicação

Para esta aplicação vamos simular as observações e através do Amostrador de Gibbs, estimar os valores de τ, μ, ξ e fazer previsão 1-passo à frente. As distribuições envolvidas no processo são as do Teorema 7.2 e a equação (3.8). Portanto, para n observações, o Amostrador de Gibbs será aplicado para as distribuições:

$$\tau | \theta, \xi, \mu, Y,$$

$$\xi | \tau, \mu, \theta, Y,$$

$$\begin{aligned}
&\mu|\tau, \xi, \mu, \theta, Y, \\
&\theta_1|\tau, \xi, \mu, Y, \\
&\theta_2|\tau, \xi, \mu, Y, \\
&\vdots \\
&\theta_n|\tau, \xi, \mu, Y, \\
&\theta_{n+1}|\tau, \xi, \mu, \theta, Y, \\
&Y_{n+1}|\theta, \theta_{n+1}, \tau, \xi, \mu, Y.
\end{aligned}$$

Geramos 500 observações para o modelo (3.7) e para os parâmetros escolhemos os valores $\tau = 5, \mu = 0,7$ e $\xi = 0,2$. Das 500 observações desprezamos as 200 primeiras e trabalhamos com as 300 últimas.

Os valores iniciais para a primeira iteração foram $\tau^{(0)} = 3, \xi^{(0)} = 0,1, \mu^{(0)} = 0,5, Y_{n+1}^{(0)} = -1$ e $\theta_t^{(0)} = j$, onde $t = 1, \dots, 301$ e j inicia em -0,52 e segue uma p.a. com razão 0,1.

Tabela IX- Média, desvio padrão e previsão 1-passo à frente

Parâmetros	Valor real	Valor estimado	d.p.
τ	5	4,84	0,86
ξ	0,2	0,27	0,06
μ	0,7	0,78	0,07
Y_{301}	-0,88	0,11	0,59

Fizemos 1000 iterações, desprezamos as 100 primeiras e para as restantes tomamos uma a cada 15, de modo a obtermos valores aproximadamente independentes.

Os valores médios obtidos das iterações para os parâmetros e para previsão 1-passo à frente estão na tabela IX.

Podemos observar que os valores estimados através do A.G., com exceção da previsão, foram muito bons.

Na Figura 4 apresentamos os gráficos para os parâmetros e para previsão 1-passo à frente assim como seus respectivos histogramas.

Figura 4

Discussões e Conclusões

O objetivo deste trabalho foi a análise Bayesiana de alguns modelos não lineares de Séries Temporais. Os modelos analisados foram TARMA, ASMA e RCA.

O modelo TARMA ainda não havia sido analisado seja com a metodologia clássica ou Bayesiana. Talvez isto se deva ao fato de alguns pesquisadores não serem favoráveis ao uso de uma aproximação de distribuições normal-gama para a função de verossimilhança. Atualmente com a crescente utilização do Amostrador de Gibbs e do Algoritmo de Metropolis-Hastings esta aproximação pode ser evitada.

Para o modelo ASMA também não havia na literatura uma análise Bayesiana. Os parâmetros do modelo foram estimados por Wecker minimizando a soma de quadrados dos resíduos.

A nossa contribuição para o modelo RCA foi a análise através do Amostrador de Gibbs, embora fosse possível encontrarmos na literatura uma análise Bayesiana para o modelo, esta envolvia distribuições condicionais e a sugestão que métodos numéricos fossem utilizados para a análise.

A literatura para análise Bayesiana de modelos de Séries Temporais ainda é muito limitada apesar do crescente interesse. Trabalhos futuros envolvem a análise de outros modelos de Séries Temporais e Econometria. Vários modelos não lineares podem ser encontrados em Tong(1990) e Granger e Terasvirta(1993). Estamos particularmente interessados nos modelos TARMASO, ARCH e GARCH.

O primeiro é uma extensão do modelo TARMA e podemos dizer que um sistema (Y_t, Z_t) é denominado "Open loop threshold autoregressive moving average", TARMASO, com Y_t sendo os dados observáveis de saída e Z_t os dados observáveis de entrada, se

$$Y_t = \phi_{l0} + \sum_{i=1}^{p_l} \phi_{li} Y_{t-i} + a_t^{(l)} + \sum_{j=1}^{q_l} \theta_{lj} a_{t-j}^{(l)} + \sum_{m=1}^{m_l} \beta_{lm} Z_{t-m}, \quad \text{se } Z_{t-d} \in R_l$$

para $l = 1, \dots, k$. A análise envolve o Amostrador de Gibbs e já se encontra em fase adiantada.

Bibliografia

- Brockwell,P.J. and Davis,R.A.(1991a). *Time Series: Theory and Methods*. New York: Springer-Verlag.
- Brockwell,P.J. and Davis,R.A.(1991b). *ITSM: An Interactive Time Series Modelling Package for the PC*. New York: Springer-Verlag.
- Broemeling,L.D.(1985). *Bayesian analysis of linear models* New York: Marcel Decker.
- Broemeling,L.D. and Cook,P.(1992). Bayesian analysis of threshold autoregressions. *Communications in Statistics-Theory and Methods*,21(9),2459-2482.
- Broemeling,L.D. and Shaarawy,S.(1984). Bayesian inferences and forecasts with moving average processes. *Communications in Statistics-Theory and Methods*, 13(5), 1871-1888.
- Broemeling,L.D. and Shaarawy,S.(1986). *A Bayesian analysis of time series. In Bayesian inference and decision techniques* (P.Goel and A. Zellner, eds.). Elsevier, New York.
- Broemeling,L.D. and Shaarawy,S.(1988). Time series: A Bayesian analysis in the time domain. In *Bayesian analysis of time series and dynamic models*, J.C. Spall, Ed., Marcel Dekker,NY, 1-21.
- Casella,G. and George,E.(1992). Explaining the Gibbs Sampler. *The American Statistician*,46, 167-174.
- Chen,C.W.S. and Lee,J.C.(1995). Bayesian inference of threshold autoregres-

- sive models. *J. Time Series Analysis*, 16(5), 483-392.
- Chib, S. and Greenberg, E. (1994). Bayes inference for regression models with ARMA(p,q) errors. *Journal of Econometrics*, 64, 183-206.
- Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4), 327-335.
- Cowles, M.K. and Carlin, B.P. (1996). Markov chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association*, 91(434).
- Diaz, J. (1990). Bayesian forecasting for AR(1) models with normal coefficients. *Communications in Statistics- Theory and Methods*, 19(6), 2229-2246.
- Gamerman, D. (1996) Simulação Estocástica via Cadeia de Markov. *XII SINAPE*
- Gelfand, A.E. and Smith, A.F.M. (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85, 398-409.
- Gelman, A. and Rubin, D.B. (1992). Inference from iterative simulation using multiple sequence. *Statistical Science*, 7(4), 457-511.
- Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.
- Granger, C.W.J. and Terasvirta, T. (1993). *Modelling Nonlinear Economic*

- Relationships*. Oxford University Press.
- Harvey,A.C.(1981). *Time Series Models*.New York: Wiley.
- Hastings ,W.K.(1970) Monte Carlo sampling methods using Markov chains and their applications.*Biometrika*, 57, 97-109.
- Jeffreys,H.(1939).*Theory of Probability*. Oxford at the Clarendon Press.
- Metropolis,N., Rosenbluth,A.W., Rosenbluth,M.N., Teller,A.H. and Teller,E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics.*, 21, 1087-1092.
- Monahan,J.F.(1983). Fully Bayesian analysis of ARMA time series models. *Journal of Econometrics*, 21, 307-331.
- Newbold,P.(1973). Bayesian estimation of Box-Jenkins transfer function noise models. *Journal of the Royal Statistical Society, B*, 323-336.
- Pole,A., West,M. and Harrison,J. (1994). *Applied Bayesian Forecasting and Time Series Analysis*. New York: Chapman and Hall.
- Roberts,G.O., Gelman,A. and Gilks,W.R.(1994). Weak convergence and optimal scaling of random walk Metropolis Algorithms. *Technical Report, University of Cambridge*.
- Shaarawy,S. and Broemeling,L.D.(1985). Inferences and prediction with ARMA processes. *Communications in Statistics-Theory and Methods*, 14(10), 2029-2040.
- Stigler,G.J. and Kindahl,J.K.(1970).*The Behavior of Industrial Prices*.New York:National Bureau of Economic Research.
- Tong,H.(1977). Discussion of a paper by A.J. Lawrance and N.T. Kottegoda.

Journal of the Royal Statistical Society A, 140, 34-35.

Tong,H.(1983). *Threshold Models in Non-linear Time Series*. New York: Springer Verlag.

Tong,H.(1990). *Non-linear Time Series: A Dynamical Approach*. Oxford: Oxford University Press.

Tong,H. and Lim,K.S.(1980). Threshold autoregression, limit cycles and cyclical data (with discussion). *Journal of the Royal Statistical Society B*, 42, 245-292.

Tsay,R.S.(1989). Testing and modeling threshold autoregressive process. *Journal of the American Statistical Association* , 84, 231-240.

Wecker,W.E.(1981). Asymmetric Time Series. *Journal of the American Statistical Association*,76(373),16-21.

West,M. and Harrison,J.(1989). *Bayesian forecasting and dynamic models*. New York : Springer.

Zellner,A.(1971). *An Introduction to Bayesian Inference in Econometrics*. New York: John Wiley and Sons. Inc.

Apêndice 1

Este Apêndice é destinado a apresentação das distribuições univariadas e multivariadas utilizadas neste trabalho. As formas aqui apresentadas são as mais convenientes para a Análise Bayesiana.

I- Distribuições Univariadas

1- Distribuição Gama

Uma quantidade aleatória contínua x tem uma distribuição *Gama* com parâmetros α e β ($\alpha > 0, \beta > 0$) se sua função de densidade $Gama(x|\alpha, \beta)$ é

$$Gama(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{(\alpha-1)} \exp\{-\beta x\}, x > 0.$$

Tem-se $E(x) = \frac{\alpha}{\beta}$ e $Var(x) = \frac{\alpha}{\beta^2}$.

2-Distribuição Gama-Inversa

Uma quantidade aleatória contínua x tem uma distribuição *Gama – Inversa* com parâmetros α e β ($\alpha > 0, \beta > 0$) se sua função de densidade $GI(x|\alpha, \beta)$ é

$$GI(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} \exp\left\{-\frac{\beta}{x}\right\}, x > 0.$$

Tem-se $E(x) = \frac{\beta}{\alpha-1}, \alpha > 1$ e $Var(x) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}, \alpha > 2$.

O termo *Gama – Inversa* vem do fato que se $y \sim Gama(\alpha, \beta)$ então $x = y^{-1} \sim GI(\alpha, \beta)$.

3- Distribuição Normal

Uma quantidade aleatória contínua x tem uma distribuição *Normal* com parâmetros μ e σ^2 ($\mu \in \mathbb{R}, \sigma^2 > 0$) se sua função de densidade $N(x|\mu, \sigma^2)$ é

$$N(x|\mu, \sigma^2) = \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\}, x \in \mathbb{R}.$$

Tem-se $E(x) = \mu$ e $Var(x) = \sigma^2$. Se $\tau^{-1} = \sigma^2$, dizemos que τ representa a precisão da distribuição e a função de densidade é

$$N(x|\mu, \tau^{-1}) = \frac{\tau^{1/2}}{\sqrt{2\pi}} \exp\left\{-\frac{\tau}{2}(x - \mu)^2\right\}, \tau > 0, x \in \mathbb{R}.$$

4-Distribuição Logística

Uma quantidade aleatória contínua x tem uma distribuição *Logística* com parâmetros α e β , ($\alpha \in \mathbb{R}, \beta > 0$) se sua função de densidade $Lo(x|\alpha, \beta)$ é

$$Lo(x|\alpha, \beta) = \beta^{-1} \frac{\exp\left\{-\left(\frac{x-\alpha}{\beta}\right)\right\}}{\left[1 + \exp\left\{-\left(\frac{x-\alpha}{\beta}\right)\right\}\right]^2}, x \in \mathbb{R}.$$

A distribuição *Logística* torna-se mais simples se escrita em termos de sua função de distribuição

$$F_x(x) = \left[1 + \exp\left\{-\left(\frac{x - \alpha}{\beta}\right)\right\}\right]^{-1}.$$

Tem-se $E(x) = \alpha$ e $Var(x) = \frac{\beta^2 \pi^2}{3}$.

5-Distribuição de Student(t)

Uma quantidade aleatória contínua x tem uma distribuição de *Student* com parâmetros μ e λ, α ($\mu \in \mathbb{R}, \lambda > 0, \alpha > 0$) se sua função de densidade $t(x|\alpha, \mu, \lambda)$ é

$$t(x|\alpha, \mu, \lambda) = \frac{\Gamma((\alpha + 1)/2)}{\Gamma(\alpha/2)\Gamma(1/2)} \left(\frac{\lambda}{\alpha}\right)^{1/2} \left\{1 + \frac{\lambda}{\alpha}(x - \mu)^2\right\}^{-\frac{\alpha+1}{2}}, x \in \mathbb{R}.$$

Tem-se $E(x) = \mu$, se $\alpha > 0$ e $Var(x) = \frac{1}{\lambda} \frac{\alpha}{\alpha-2}$, se $\alpha > 2$.

II- Distribuições Multivariadas

6-Distribuição Multinomial

Um vetor aleatório discreto $\mathbf{x} = (x_1, \dots, x_k)'$ tem uma distribuição *Multinomial* de dimensão k , com parâmetros $\theta = (\theta_1, \dots, \theta_k)'$ e n ($0 < \theta_i < 1, \sum_i \theta_i < 1, n = 1, 2, \dots$) se sua função de probabilidade $Mu(\mathbf{x}|\theta, n)$ para $x_i = 0, 1, 2, \dots$, com $\sum_{i=1}^k x_i \leq n$ é

$$Mu(\mathbf{x}|\theta, n) = \frac{n!}{\prod_{i=1}^k x_i! (n - \sum_{i=1}^k x_i)!} \prod_{i=1}^k \theta_i^{x_i} (1 - \sum_{i=1}^k \theta_i)^{(n - \sum_{i=1}^k x_i)}.$$

O vetor de média e a matriz de covariância são dados por $E(x_i) = n\theta_i, Var(x_i) = n\theta_i(1 - \theta_i), Cov(x_i, x_j) = -n\theta_i\theta_j$.

7-Distribuição Normal-Gama

Um vetor aleatório contínuo (x, y) tem uma distribuição *Normal – Gama* com parâmetros $\mu, \sigma^2, \alpha, \beta (\mu \in \mathbb{R}, \sigma^2 > 0, \alpha > 0, \beta > 0)$ se sua densidade $NG(x, y|\mu, \sigma^2, \alpha, \beta)$ é

$$NG(x, y|\mu, \sigma^2, \alpha, \beta) = N(x|\mu, y\sigma^2)Gama(y|\alpha, \beta), x \in \mathbb{R}, y > 0,$$

onde as densidades Normal e Gama já foram definidas.

8-Distribuição Normal Multivariada

Um vetor aleatório contínuo $\mathbf{x} = (x_1, \dots, x_k)'$ tem uma distribuição *Normal Multivariada* de dimensão k , com parâmetros $\mu = (\mu_1, \dots, \mu_k)'$, λ , onde $(\mu \in \mathbb{R}^k)$ e λ é uma matriz $k \times k$ simétrica, positiva-definida, se sua densidade $N(\mathbf{x}|\mu, \lambda)$ é

$$N_k(\mathbf{x}|\mu, \lambda) = \frac{|\lambda|^{1/2}}{(2\pi)^{k/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \mu)' \lambda (\mathbf{x} - \mu)\right\}, \mathbf{x} \in \mathbb{R}^k.$$

No caso geral, $E(x_i) = \mu_i$ e com $\Sigma = \lambda^{-1}$ de elementos σ_{ij} , $Var(x_i) = \sigma_{ii}$ e $Cov(x_i, x_j) = \sigma_{ij}$ tal que $\mathbf{V}(\mathbf{x}) = \lambda^{-1}$. O parâmetro μ é denominado vetor de médias e λ é a matriz de precisão (inversa da matriz de covariância).

9-Distribuição de Student Multivariada

Um vetor aleatório contínuo $\mathbf{x} = (x_1, \dots, x_k)'$ tem uma distribuição de *Student Multivariada* de dimensão k , com parâmetros $\mu = (\mu_1, \dots, \mu_k)'$, $\lambda, \alpha, (\mu \in$

$\mathbb{R}^k$, λ matriz simétrica $k \times k$ positiva definida, $\alpha > 0$) se sua densidade é

$$t_k(\mathbf{x}|\alpha, \mu, \lambda) = \frac{\Gamma((\alpha + k)/2)}{\Gamma(\alpha/2)(\alpha\pi)^{k/2}} |\lambda|^{1/2} \left\{ 1 + \frac{1}{\alpha} (\mathbf{x} - \mu)' \lambda (\mathbf{x} - \mu) \right\}^{-\frac{(\alpha+k)}{2}}, \mathbf{x} \in \mathbb{R}^k.$$

Tem-se $E(\mathbf{x}) = \mu$ e $\mathbf{V}(\mathbf{x}) = \lambda^{-1}(\frac{\alpha}{\alpha-2})$. Embora não seja exatamente a inversa da matriz de covariância, o parâmetro λ é frequentemente tratado como a “matriz de precisão” da distribuição.

Em particular, a densidade apropriada para cada subvetor de $\mathbf{x}$ é Student(multivariada), de dimensão apropriada, com média e matriz de precisão dados pelo subvetor de μ e submatriz de λ correspondentes.

Apêndice 2

Tabela 2-A. Canadian Lynx

269	523	151	377	236	358	469	59	758	1388	229	1000
321	98	45	225	245	784	736	188	1307	2713	399	1590
585	184	68	360	552	1594	2042	377	3465	3800	1132	2657
871	279	213	731	1623	1676	2811	1292	6991	3091	2432	3396
1475	409	546	1638	3311	2251	4431	4031	6313	2985	3574	
2821	2285	1033	2725	6721	1426	2511	3495	3794	3790	2935	
3928	2685	2129	2871	4254	756	389	587	1836	674	1537	
5943	3409	2536	2119	687	299	73	105	345	81	529	
4950	1824	957	684	255	201	39	153	382	80	485	
2577	409	361	299	473	229	49	387	808	108	662	

Tabela 2-B. Carbon Steel, Sheet and Strip-Hot Rolled

	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966
BLS										
January	91.49	95.33	96.93	96.93	96.93	96.93	96.93	100.00	100.00	99.99
February	91.49	95.33	96.93	96.93	96.93	96.93	96.93	100.00	100.00	99.99
March	91.49	94.53	96.93	96.93	96.93	96.93	96.93	100.00	100.00	99.99
April	91.49	94.53	96.93	96.93	96.93	96.93	97.49	100.00	100.00	99.99
May	91.48	94.53	96.93	96.93	96.93	96.93	100.00	100.00	100.00	99.99
June	91.48	94.25	96.93	96.93	96.93	96.93	100.00	100.00	100.00	99.99
July	95.33	94.25	96.93	96.93	96.93	96.93	100.00	100.00	100.00	99.99
August	95.33	96.93	96.93	96.93	96.93	96.93	100.00	100.00	99.99	102.33
September	95.33	96.93	96.93	96.93	96.93	96.93	100.00	100.00	99.99	102.33
October	95.33	96.93	96.93	96.93	96.93	96.93	100.00	100.00	99.99	102.33
November	95.33	96.93	96.93	96.93	96.93	96.93	100.00	100.00	99.99	102.33
December	95.33	96.93	96.93	96.93	96.93	96.93	100.00	100.00	99.99	102.33
NBER										
January	88.40	92.29	96.67	96.94	96.93	96.90	96.80	99.77	99.15	98.99
February	88.63	92.82	96.68	96.94	96.90	96.91	96.76	99.89	99.21	98.99
March	88.87	93.37	96.69	96.94	96.89	96.94	96.73	100.01	99.27	98.98
April	89.10	93.92	96.69	96.94	96.89	96.94	96.83	100.13	99.27	98.98
May	89.54	93.92	96.78	96.94	96.83	96.96	98.61	100.13	99.56	99.00
June	89.97	93.92	96.86	96.94	96.78	96.96	99.17	100.13	99.85	99.06
July	93.84	93.92	96.86	96.94	96.73	96.96	99.45	100.13	100.15	99.12
August	93.82	99.74	96.83	96.94	96.70	97.02	99.51	100.13	100.13	99.41
September	93.81	98.25	96.81	96.94	96.49	97.09	99.58	100.13	100.12	99.69
October	93.79	96.79	96.87	96.95	96.50	97.14	99.65	100.13	100.10	99.96
November	93.89	96.75	96.94	96.94	96.62	96.97	99.69	99.85	99.64	99.96
December	94.00	96.71	96.94	96.94	96.74	96.81	99.73	99.57	99.18	99.96

Tabela 2-C. Tinplate

	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966
BLS										
January	92.29	96.34	100.01	100.01	100.01	100.01	100.00	100.00	100.00	102.57
February	92.29	96.34	100.01	100.01	100.01	100.01	100.00	100.00	100.00	102.57
March	92.29	96.34	100.01	100.01	100.01	100.01	100.00	100.00	100.00	102.57
April	92.29	96.34	100.01	100.01	100.01	100.01	100.00	100.00	100.00	102.57
May	96.34	96.34	100.01	100.01	100.01	100.01	100.00	100.00	100.00	102.57
June	96.34	96.34	100.01	100.01	100.01	100.01	100.00	100.00	100.00	102.57
July	96.34	96.34	100.01	100.01	100.01	100.01	100.00	100.00	100.00	102.57
August	96.34	96.34	100.01	100.01	100.01	100.01	100.00	100.00	100.00	102.57
September	96.34	96.34	100.01	100.01	100.01	100.01	100.00	100.00	100.00	102.57
October	96.34	96.34	100.01	100.01	100.01	100.01	100.00	100.00	100.00	102.57
November	96.34	100.01	100.01	100.01	100.01	100.01	100.00	100.00	99.78	102.57
December	96.34	100.01	100.01	100.01	100.01	100.01	100.00	100.00	102.57	102.57
NBER										
January	96.10	97.38	100.07	105.58	100.75	100.49	99.53	100.61	100.16	102.84
February	96.10	97.81	100.15	101.27	100.80	100.56	99.61	100.49	100.29	103.44
March	96.01	97.80	100.28	101.32	100.89	100.47	99.75	100.39	100.48	103.30
April	95.98	97.78	100.40	101.45	100.90	100.46	99.89	100.29	100.67	103.16
May	96.88	97.76	100.53	101.58	100.91	100.44	100.89	100.19	100.85	103.02
June	96.58	97.77	100.40	101.56	100.94	100.35	100.95	100.06	100.69	103.26
July	97.51	97.77	100.28	101.12	100.96	100.25	101.00	99.93	101.25	103.50
August	97.18	98.53	100.15	101.01	100.98	100.16	101.06	99.12	101.09	103.74
September	97.06	98.41	99.87	101.00	100.75	99.91	101.00	99.37	100.63	103.57
October	96.83	98.59	100.44	100.90	100.51	99.61	100.93	99.63	100.17	103.82
November	96.61	99.60	104.81	100.81	100.28	99.36	100.87	99.89	99.71	104.07
December	96.95	99.99	105.19	100.69	100.42	99.45	100.74	100.02	101.75	104.60

Tabela 2-D. Regular Gasoline

	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966
BLS										
January	111.50	110.57	105.69	99.70	111.22	105.51	103.57	102.65	100.53	106.15
February	116.56	106.24	106.24	101.36	110.57	101.45	101.36	101.64	98.69	104.95
March	116.56	104.49	107.90	104.77	110.39	98.96	103.57	99.15	99.33	104.12
April	116.56	103.66	108.64	105.51	107.16	107.07	104.68	97.30	100.81	105.32
May	116.01	102.46	107.90	101.91	105.41	106.80	108.09	100.44	103.11	107.81
June	115.09	103.66	104.86	105.69	108.82	107.44	109.84	100.90	104.31	111.22
July	113.06	107.26	109.95	108.27	109.28	106.89	108.82	101.36	104.31	110.67
August	113.06	109.84	108.36	110.85	107.62	105.51	104.03	99.52	104.31	111.86
September	114.08	109.84	107.81	111.22	105.04	109.19	104.31	95.92	103.75	111.86
October	112.88	107.44	105.32	111.22	101.10	108.36	103.39	99.52	103.75	111.86
November	110.57	105.87	103.29	110.85	103.85	106.98	99.33	100.81	106.15	111.86
December	110.57	105.23	100.71	110.85	105.51	105.51	103.39	100.81	106.15	109.01
NBER										
January	107.91	109.10	102.67	103.93	106.77	103.66	102.54	101.40	99.55	98.45
February	109.05	108.17	102.68	104.35	106.82	103.18	102.37	101.27	98.28	98.46
March	109.14	108.03	103.11	104.55	106.75	102.87	102.54	100.83	98.24	98.52
April	109.01	106.97	103.56	104.47	106.41	103.82	102.57	100.36	98.29	98.32
May	109.49	105.31	103.17	104.96	106.16	103.91	103.38	100.38	98.35	98.68
June	110.12	104.66	103.27	104.59	105.57	103.76	103.40	99.87	98.17	99.36
July	110.05	104.14	103.59	105.06	105.30	103.86	102.69	99.55	98.19	99.58
August	110.30	103.87	104.18	105.21	104.67	103.64	102.44	99.39	97.98	99.62
September	110.00	103.87	104.24	105.45	104.44	103.90	102.39	99.39	97.97	99.76
October	109.86	103.56	103.82	106.04	103.90	103.60	102.12	99.29	98.05	99.94
November	109.61	102.81	103.79	106.11	103.90	103.16	101.48	99.06	98.13	100.49
December	109.33	102.96	104.52	106.64	103.88	102.80	101.96	99.19	98.20	100.95