

Universidade Federal do Rio de Janeiro  
Instituto de Matemática  
Departamento de Métodos Estatísticos



**Aplicações de Modelos Markovianos  
Ocultos em Modelos Dinâmicos  
Não-Lineares e Não Gaussianos**

**Gabriella Pires Pacca**

**Orientador: Prof. Carlos Abanto Valle**

2015

# Aplicações de Modelos Markovianos Ocultos em Modelos Dinâmicos Não Lineares e Não Gaussianos

Gabriella Pires Pacca

Projeto de Graduação apresentado ao curso de Ciências Atuariais do Instituto de Matemática da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários para a obtenção do grau de Bacharel em Ciências Atuariais.

Examinado por:

---

Prof. Carlos Abanto Valle  
Dr. Sc. - IM - UFRJ - Orientador.

---

Prof. Vinicius Pinheiro Israel  
Dr. Sc. - IM - UFRJ.

---

Dmitri Santos  
Bacharel em Ciências Atuariais.

Rio de Janeiro, RJ - Brasil

2015

*À minha família e meus grandes amigos.*

*“Todas as vitórias ocultam uma abdicação.”*

Simone de Beauvoir.

# Agradecimentos

Agradeço primeiramente à minha família, por todo amor, carinho e por me ensinar sempre a importância de se dedicar aos estudos. Além, é claro, de todo apoio e confiança que colocaram em minhas escolhas pessoais e profissionais, que me permitiram chegar até aqui.

Agradeço também a minha mãe Fátima Cristina (*in memoriam*), que infelizmente me deixou precocemente, mas me deu a vida e muito carinho enquanto pôde.

Agradeço a meus demais familiares, que contribuíram de modo direto ou indireto para que eu chegasse até aqui, com destaque para minha tia Celina. Agradeço por todo apoio e incentivo que me deu, das mais diferentes formas.

Agradeço em especial ao professor Carlos Abanto, por toda a boa vontade e paciência ao orientar esse nosso projeto, enriquecendo-o com todo o seu conhecimento.

Agradeço ao corpo docente do IM - UFRJ, por todo o conhecimento que me transmitiram, com bastante empenho, ao longo das diversas disciplinas que fiz na graduação. Gostaria de destacar o professor Rafael Souza dos Santos, que felizmente tive a chance de ter como monitor.

Por fim, agradeço ainda ao professor Vinicius Israel e ao atuário Dmitri Santos por terem aceitado fazer parte da banca avaliadora desse projeto.

# Resumo

A modelagem de séries temporais, inferência e previsão, baseadas em modelos dinâmicos, é uma das mais importantes áreas que surgiram na estatística, visto que muitos dos problemas práticos envolvendo estatística podem ser colocados nesta estrutura. O objetivo desta monografia é aplicar a metodologia de aproximação a verossimilhança através de modelos markovianos ocultos em modelos dinâmicos não lineares e não Gaussianos.

Para ilustrar essa aplicação, são apresentados estudos de simulação e de casos para dois tipos de modelos, que são: modelos de volatilidade estocástica e modelos de dados binários. Ao longo desta monografia, encontram-se também um resumo de cadeia de Markov, alguns conceitos importantes de modelos dinâmicos e por fim conclui-se com algumas sugestões para trabalhos futuros.

*Palavras-Chaves:* modelos markovianos ocultos, séries temporais, integração numérica, dados binários, volatilidade estocástica.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Preliminares</b>	<b>3</b>
2.1	Cadeias de Markov . . . . .	3
2.2	Modelos Markovianos Ocultos . . . . .	5
2.3	Distribuição Generalizada de Valores Extremos . . . . .	6
<b>3</b>	<b>Modelos Dinâmicos</b>	<b>8</b>
3.1	Modelo Dinâmico Geral . . . . .	9
3.2	Modelo Dinâmico Linear . . . . .	10
<b>4</b>	<b>Apresentação da Metodologia</b>	<b>12</b>
4.1	Aproximação da Verossimilhança . . . . .	12
<b>5</b>	<b>Aplicações - Modelo de Volatilidade Estocástica</b>	<b>15</b>
5.1	Estudo de Simulações - Modelo de Volatilidade Estocástica (SV) . . . . .	16
5.2	Aplicações em Dados Reais - Modelo de Volatilidade Estocástica (SV) . . . . .	19
<b>6</b>	<b>Aplicações - Modelo para Dados Binários</b>	<b>22</b>
6.1	Estudo de Simulações - Modelo para Dados Binários . . . . .	23
6.1.1	Modelo 1: Função de Ligação T-Student . . . . .	23
6.1.2	Modelo 2: Função de Ligação Probit . . . . .	25
6.1.3	Modelo 3: Função de Ligação GEV . . . . .	27
6.2	Aplicações em Dados Reais - Modelo de Dados Binários . . . . .	29
<b>7</b>	<b>Conclusão</b>	<b>32</b>
<b>A</b>	<b>Programação no R</b>	<b>33</b>
A.1	Aproximação da Verossimilhança . . . . .	33

A.2 Reconstrução da Volatilidade . . . . .	34
<b>Referências Bibliográficas</b>	<b>36</b>



# Lista de Figuras

5.1	Modelo de Volatilidade Estocástica (SV) - Observações Simuladas . . . . .	16
5.2	Modelo de Volatilidade Estocástica (SV) - Estados Simulados . . . . .	17
5.3	Retorno Ibovespa . . . . .	19
5.4	Reconstrução da Volatilidade . . . . .	21

# Lista de Tabelas

5.1	Estudo de Simulações - Modelo de Volatilidade Estocástica (SV) ( $T = 2.000$ ), com $\phi = 0,98$ , $\sigma = 0,2$ , $\beta = 0,05$ . . . . .	17
5.2	Estudo de Simulações - Modelo de Volatilidade Estocástica (SV) ( $T = 2.000$ e $m = 50$ ) e com $\phi = 0,98$ , $\sigma = 0,15$ , $\beta = 0,02$ . . . . .	18
5.3	Retornos Ibovespa . . . . .	20
5.4	IBOVESPA (19/04/2006/ a 16/05/2014) com $T = 2.000$ . . . . .	20
6.1	Estudo de simulações: $g_{max} = -g_{min} = 2$ - Função de Ligação T-Student. . . . .	24
6.2	Estudo de simulações: $g_{max} = -g_{min} = 4$ - Função de Ligação T-Student. . . . .	25
6.3	Estudo de simulações: $g_{max} = -g_{min} = 2$ - Função de Ligação Probit. . . . .	26
6.4	Estudo de simulações: $g_{max} = -g_{min} = 3$ - Função de Ligação Probit. . . . .	26
6.5	Estudo de simulações: $g_{max} = -g_{min} = 4$ - Função de Ligação Probit. . . . .	27
6.6	Estudo de simulações: $g_{max} = -g_{min} = 2$ - Função de Ligação GEV. . . . .	28
6.7	Estudo de simulações: $g_{max} = -g_{min} = 3$ - Função de Ligação GEV. . . . .	28
6.8	Estudo de simulações: $g_{max} = -g_{min} = 4$ - Função de Ligação GEV. . . . .	29
6.9	Aplicações em Dados Reais - Função de Ligação Probit. . . . .	30
6.10	Aplicações em Dados Reais - Função de Ligação GEV. . . . .	30

# Capítulo 1

## Introdução

Os Modelos Dinâmicos, também conhecidos como modelos de espaço de estado, são formulados para permitir alterações nos valores de estados com o passar do tempo. Há bastante interesse no estudo da modelagem, inferência e previsão para esses modelos, pois eles formam uma classe muito flexível, conseguindo ser aplicados em diversas situações reais. Por conta disso, essa área se tornou uma das mais importantes que surgiram na estatística durante o último século.

O filtro de Kalman pode ser utilizado para fazer previsões dos estados e das observações dos modelos dinâmicos lineares e/ou Gaussiano. No entanto, para a classe dos modelos dinâmicos não lineares e/ou não Gaussiano, devido à maior complexidade da sua função de verossimilhança, torna-se necessário a utilização de outras técnicas para fazer inferência.

Avanços recentes em computação estocástica aumentaram muito o potencial de utilização dos modelos dinâmicos nas mais diversas áreas. Métodos como Monte Carlo via Cadeia de Markov (MCMC) e Monte Carlo sequencial (SMC) foram desenvolvidos e estão bem documentados.

No entanto, a inferência através dos métodos citados acima, depende fortemente da especificação do modelo. A finalidade desta monografia é apresentar uma metodologia de aproximação da função de verossimilhança usando modelos markovianos ocultos (*Hidden Markov Models* - HMM), tornando assim a inferência e previsão dos estados, do ponto de vista da estatística clássica, mais rápida e com baixo custo computacional.

As principais contribuições desse trabalho incluem: a aplicação da metodologia pro-

posta em modelos de volatilidade estocástica e em modelos dinâmicos com observações binárias. Em ambos, estudam-se as propriedades dos estimadores de máxima verossimilhança através de um estudo de simulações.

A metodologia foi implementada computacionalmente usando o software R que está disponível de forma gratuita, além de possuir ampla documentação e fornecer uma ampla variedade de rotinas para a modelagem estatística, as quais podem ser estendidas pelos usuários.

A estrutura da presente monografia está descrita a seguir. O Capítulo 2 introduz a parte conceitual necessária para a melhor compreensão e contextualização do tema proposto. São apresentados conceitos introdutórios de cadeia de Markov, modelos markovianos ocultos, e além disso, também é apresentado um breve resumo da função GEV (*Generalized Extreme Value*).

O Capítulo 3, fornece uma descrição dos modelos dinâmicos e as técnicas de estimação dos estados e observações.

O Capítulo 4 apresenta a metodologia de aproximação de verossimilhança através de um modelo markoviano oculto, bem como suas vantagens e desvantagens.

O Capítulo 5 apresenta o modelo de volatilidade estocástica (SV) usando a aproximação da verossimilhança via modelos markovianos ocultos. Um estudo de simulação é apresentado, bem como uma aplicação da metodologia proposta é ilustrada com os retornos do Índice da Bolsa de Valores de São Paulo.

No Capítulo 6 descreve-se um modelo dinâmico para observações binárias. Diferentes funções de ligação são introduzidas. Um estudo de simulação é implementado e uma aplicação com dados reais é fornecida.

No Capítulo 7 são apresentadas algumas considerações finais, os pontos mais relevantes do trabalho e, além disso, são apresentadas algumas considerações para trabalhos futuros. Por fim, apresenta-se um apêndice com a programação, de forma genérica, utilizada na implementação da metodologia no Software R.

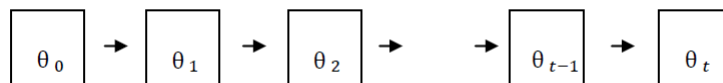
# Capítulo 2

## Preliminares

Este capítulo conduz a parte conceitual necessária para a melhor compreensão e contextualização do tema proposto. São apresentados conceitos introdutórios de um processo de Markov, entre eles cadeia de Markov e modelos markovianos ocultos.

### 2.1 Cadeias de Markov

“o futuro independe do passado, dado o presente”



Uma cadeia de Markov é um processo estocástico em que qualquer probabilidade associada ao futuro do processo pode ser descrita apenas conhecendo o valor que esse processo assume no presente, podendo assim abdicar de toda informação do passado. Ou seja, informalmente, o futuro desses processos só dependem do presente. A sua probabilidade de transição, ou seja, a probabilidade do processo transitar de um estado  $i$  para o  $j$  em instantes de tempo consecutivos é dado por:

$$P_{ij} = P(i, j) = P(\theta_t = j \mid \theta_{t-1} = i_{t-1}, \theta_{t-2} = i_{t-2}, \dots, \theta_0 = i_0) = P(\theta_t = j \mid \theta_{t-1} = i_{t-1})$$

As probabilidades de transição satisfazem as seguintes propriedades:

- (1)  $0 \leq p(i, j) \leq 1$ ,      para  $1 \leq i, j \leq n$   
(2)  $\sum_{j=1}^n p(i, j) = 1$ ,      para  $1 \leq i \leq n$

Estas probabilidades de transição, por conseguinte, podem ser expostas por uma matriz  $\mathbf{P}$  de probabilidade de transição, na qual a soma de qualquer uma das linhas tem valor 1.

Também conhecida como matriz estocástica, sua diagonal principal representa as probabilidade de um estado permanecer no mesmo lugar, e as demais entradas representam a probabilidade de transitar de um estado  $i$  para  $j$ .

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1j} & \cdots \\ p_{21} & p_{22} & \cdots & p_{2j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \\ p_{i1} & p_{i2} & \cdots & p_{ij} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \end{pmatrix}$$

Outros conceitos importantes de uma cadeia de Markov são a **irredutibilidade** e o **período**. A irredutibilidade garante que a cadeia possa visitar todos os estados, a qualquer tempo, com probabilidade positiva para qualquer que seja o ponto inicial, em outras palavras, uma cadeia de Markov é irredutível se, e somente se, ela possui uma única classe de comunicação. Analiticamente, tem-se:

$$P(\theta_t = j \mid \theta_0 = i) > 0, \quad \text{para } 1 \leq i, j \leq n$$

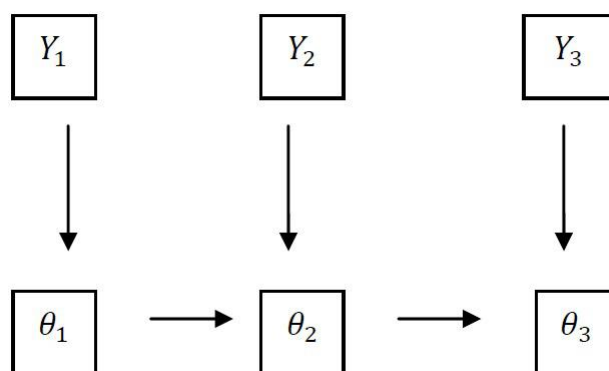
O período de uma cadeia de Markov irredutível é o período comum de seus estados. Uma cadeia é chamada de aperiódica se o seu período é igual a 1, ou seja, quando não há nenhuma repartição do espaço amostral. Logo, a aperiodicidade de uma cadeia de Markov assegura que a cadeia não se mova de modo cíclico entre seus estados.

Maiores detalhes sobre o cadeia de Marlov podem ser encontrados em G. F. LAWLER [16]

## 2.2 Modelos Markovianos Ocultos

Em uma cadeia de Markov é considerado um estado correspondente a uma observação, fazendo com que seja um modelo muito restrito para aplicar em certos problemas. Os modelos Markovianos ocultos (HMM) são uma extensão da cadeia de Markov, pois nesses modelos temos um processo estocástico de duas camadas. Um HMM é composto por dois componentes. O primeiro, um não observável (estado), é projetado para dar conta de correlação da série. O segundo é um processo dependente do estado (observação).

O processo pode ser representado na figura abaixo:



Essa classe de modelo, fornece artifícios flexíveis para modelar problemas de séries temporais, onde as observações dependem de uma sequência de estados subjacentes correlacionados. Eles tem sido aplicados com sucesso a uma grande variedade de tipos de séries de temporais, como por exemplo:

- Variáveis Discretas;
- Variáveis Contínuas;
- Séries Multivariadas;

Originalmente, os modelos Markovianos ocultos foram desenvolvidos no campo do reconhecimento da fala (Rabiner 1989); a sequência de estados subjacente é, então, dada pela sequência falada de fonemas, enquanto as observações são essencialmente dadas pelas transformadas de Fourier dos sons gravados. O reconhecimento de voz tenta decodificar

a sequência falada das observações com ruídos. A metodologia foi introduzida nos anos 60 e o fato de sua estrutura matemática poder ser aplicada em inúmeros problemas, e por apresentar resultados consistentes, fez com que sua aplicação se tornasse mais constante.

Esse modelo vem sendo frequentemente aplicado, pois, além de ser simples para estimar os parâmetros e fazer inferência, ele é rico o bastante para ser aplicado em dados reais.

Considere o modelo a seguir:

$$\mathbf{y}_t = a(\boldsymbol{\theta}_t, \epsilon_t), \quad (2.1)$$

$$\boldsymbol{\theta}_t = b(\boldsymbol{\theta}_{t-1}, \eta_t), \quad (2.2)$$

onde  $\mathbf{y}_t$  e o vetor de observações e o  $\boldsymbol{\theta}_t$  e o vetor de estados não observáveis.

Respeitando a estrutura do modelo, a função conjunta pode ser escrita da seguinte forma:

$$p(y_1, \dots, y_T, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_T) = p(\boldsymbol{\theta}_1) \cdot p(y_1 | \boldsymbol{\theta}_1) \cdot \prod_{t=2}^T p(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1}) \cdot p(y_t | \boldsymbol{\theta}_t),$$

onde  $T$  representa o número de observações.

## 2.3 Distribuição Generalizada de Valores Extremos

Os modelos de ligação GEV (*Generalized Extreme Value*) são baseados na distribuição GEV. Sua função de distribuição acumulada é dada por:

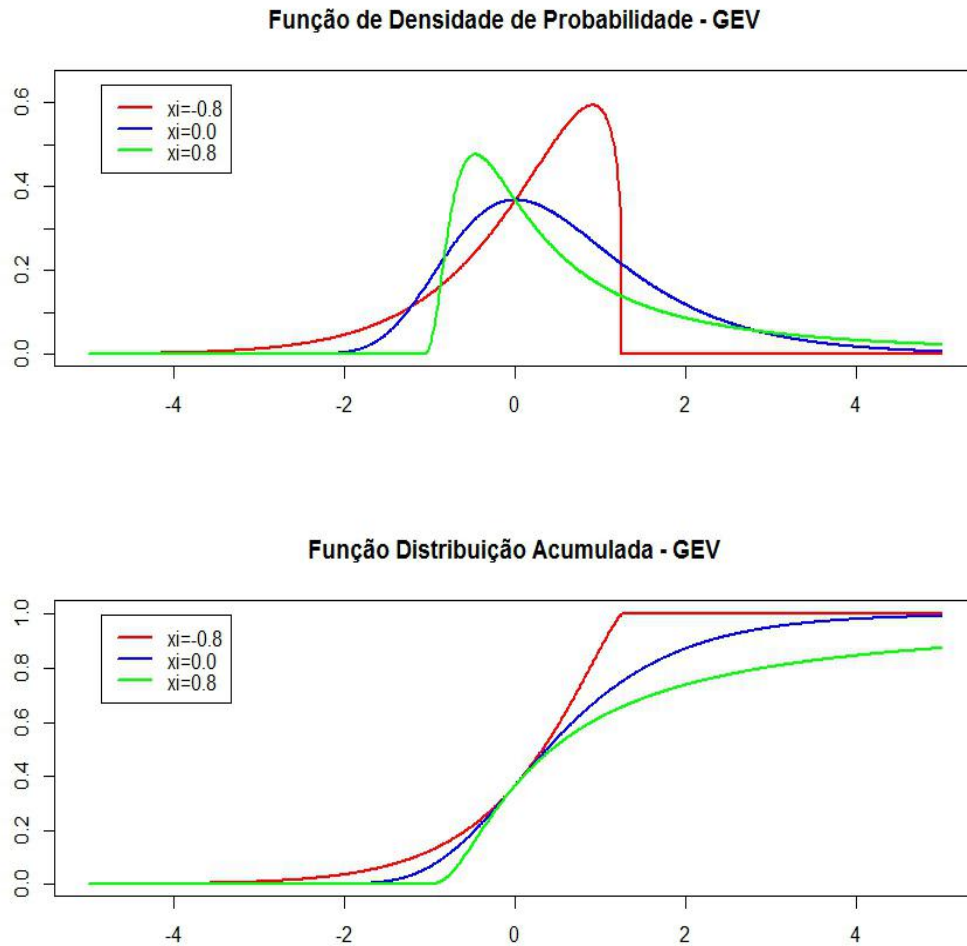
$$G(x) = \exp \left[ - \left\{ 1 + \xi \frac{x-\mu}{\sigma} \right\}^{-\frac{1}{\xi}} \right],$$

onde  $\mu \in \mathbb{R}$  é o parâmetro de localização,  $\sigma \in \mathbb{R}$  é o parâmetro de escala,  $\xi \in \mathbb{R}$  é o parâmetro de forma (assimetria) e  $x = \max(x, 0)$ . A distribuição do modelo acima é chamada de distribuição generalizada de valores extremos. Para facilitar, chamaremos de GEV.

A sua importância como uma função de ligação resulta do fato de que o parâmetro de forma  $\xi$  controla o comportamento da cauda da distribuição.



A figura abaixo fornece uma comparação da função densidade de probabilidade bem como da função distribuição acumulada da classe GEV com diferentes valores para  $\xi$  para mostrar a flexibilidade de tal distribuição. Olhando o gráfico da função de distribuição acumulada fica claro como o valor do parâmetro de forma altera a assimetria, o mesmo acontece com a velocidade que se aproxima de 1 e 0.



Obs.:  $\xi = \xi$

# Capítulo 3

## Modelos Dinâmicos

O modelo de espaço de estado, também conhecido como modelagem dinâmica, teve início na década de 60, primeiramente com Kalman(1960) [3] e posteriormente pela solução encontrada por Kalman e Bucy (1960) [4], nomeada Filtro de Kalman. Entretanto, estas descobertas só foram utilizadas por estatísticos cerca de dez anos depois e, desde então, o interesse e o uso dessa classe de modelos na análise de séries temporais vem crescendo de forma considerável.

Modelos dinâmicos proporcionam a modelagem de uma série temporal utilizando a combinação de inúmeras componentes, tais como: tendência, sazonalidade e componente regressiva. Além disso, tem uma estrutura probabilística robusta que evolui ao longo do tempo, garantindo aos modelos flexibilidade e aplicabilidade em uma ampla variedade de problemas nas mais diversas áreas. Os modelos de espaço de estados consideram uma série afetada pelo tempo, através de deformações dinâmicas e aleatórias.

Os problemas de estimação e inferência são resolvidos por algoritmos recursivos, calculando as distribuições condicionais das quantidades de interesse dadas as informações observadas. Neste sentido, estes modelos são naturalmente tratados seguindo o paradigma bayesiano.

Esses modelos podem ser usados em séries temporais univariadas ou multivariadas, também na presença de mudanças estruturais não estacionárias e padrões irregulares. De fato, uma análise de série temporal depende da possibilidade de encontrar uma regularidade no comportamento do problema estudado, tornando assim possível a previsão de um comportamento futuro.

Os modelos dinâmicos são descritos por dois processos: o processo de estados latentes  $\{\boldsymbol{\theta}_t\}$  e o processo observacional  $\{\mathbf{y}_t\}$ , onde  $t$  é um indicador temporal. Com a evolução do tempo, toda a informação relevante para prever o futuro é atualizada e pode ser usada na revisão e crítica do modelo.

Supondo que o tempo inicial seja  $t = 0$  e que  $\boldsymbol{\theta}_0$  represente a informação relevante e disponível sobre o modelo, a qual será usada pelo estatístico para fazer as previsões iniciais do futuro. De um modo semelhante, suponha que para qualquer tempo  $t > 0$ , a informação disponível e relevante seja denotada por  $\boldsymbol{\theta}_t$ . Qualquer afirmação sobre o futuro será condicionada nesta informação. Uma vez que  $\mathbf{y}_t$  foi observado no tempo  $t$ , define-se  $\{D_t\} = \{y_t, D_{t-1}\}$ .

### 3.1 Modelo Dinâmico Geral

Considere o modelo abaixo,

$$\begin{aligned} \mathbf{y}_t &= a(\boldsymbol{\theta}_t, \epsilon_t) && \text{para } t = 1, \dots, n, \\ \boldsymbol{\theta}_t &= b(\boldsymbol{\theta}_{t-1}, \eta_t) && \text{para } t = 1, \dots, n, \end{aligned}$$

onde  $\theta_t$  {para  $t = 1, \dots, T$ } representa o processo de estados latentes,  $y_t$  {para  $t = 1, \dots, T$ } é a variável resposta.  $\epsilon_t$  e  $\eta_t$  são perturbações aleatórias independentes e identicamente distribuídas.  $a(.,.)$  e  $b(.,.)$  são funções não lineares conhecidas.

A verossimilhança de um modelo dinâmico (MD) é dada por uma integral múltipla de ordem grande que em geral não pode ser resolvida de forma direta. A estimação dos parâmetros em um modelo linear e Gaussiano pode ser resolvido pelo filtro de Kalman. Por outro lado, o caso não linear e não Gaussiano é mais complexo. Possíveis métodos de estimação dos parâmetros do modelo dinâmico não linear e não Gaussiano são:

- Filtro de Kalman Estendido
- Método dos Momentos Generalizados
- Método de Monte Carlo
- Integração Numérica

## 3.2 Modelo Dinâmico Linear

Os modelos de espaço de estados lineares e Gaussianos, ou modelos dinâmicos lineares (MLD) são especificados por uma distribuição a priori Normal para o vetor de estados no instante zero.

$$\boldsymbol{\theta}_0 \sim \mathcal{N}_p(\mathbf{m}_0, \mathbf{C}_0)$$

Para cada  $t \geq 1$ , o modelo dinâmico linear é definido por:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{F}'_t \boldsymbol{\theta}_t + \mathbf{v}_t, & \mathbf{v}_t &\sim \mathcal{N}(\mathbf{0}, \mathbf{V}_t), \\ \boldsymbol{\theta}_t &= \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, & \boldsymbol{\omega}_t &\sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t), \end{aligned}$$

- $\boldsymbol{\theta}_t$ : é um vetor  $p$ -dimensional de parâmetros de estados do modelo dinâmico
- $\mathbf{F}'_t$  é a matriz de regressão  $n \times p$ , cujos elementos são conhecidos
- $\mathbf{G}_t$  é uma matriz  $p \times p$  conhecida, que descreve a evolução dos parâmetros de estado no tempo
- $\mathbf{V}_t$  é a matriz de covariância  $n \times n$  associada ao erro observacional  $\mathbf{v}_t$
- $\mathbf{W}_t$  é a matriz de covariância  $p \times p$  associada ao erro da evolução dos parâmetros de estado  $\boldsymbol{\omega}_t$

As equações podem ser escritas da seguinte forma:

$$\begin{aligned} \mathbf{y}_t \mid \boldsymbol{\theta}_t &\sim \mathcal{N}(\mathbf{F}'_t \boldsymbol{\theta}_t, \mathbf{V}_t) \\ \boldsymbol{\theta}_t \mid \boldsymbol{\theta}_{t-1} &\sim \mathcal{N}(\mathbf{G}_t \boldsymbol{\theta}_{t-1}, \mathbf{W}_t) \end{aligned}$$

A verossimilhança de um modelo dinâmico é dada por uma integral múltipla de ordem grande que em geral não pode ser resolvida de forma direta. Um modelo dinâmico linear e gaussiano pode ser resolvido pelo Filtro de Kalman. O Filtro de Kalman é um

método para avaliar a distribuição a posteriori  $(\theta_t | y_t)$  baseado na distribuição a priori de  $(\theta_t | y_{t-1})$ . Sua função é atualizar de forma recursiva a distribuição dos estados quando uma nova observação está disponível.

É complicado garantir em situação real que as premissas ocorram. Porém, a normalidade pode ser justificada por argumentos que utilizem o Teorema do Limite Central. Ainda assim, existem extensões importantes que permitem a modelagem de outliers usando distribuições de caudas pesadas, comumente utilizadas no tratamento de séries temporais discretas. O preço pago por não obter normalidade é a complexidade computacional.

# Capítulo 4

## Apresentação da Metodologia

O presente capítulo apresenta a metodologia de aproximação de verossimilhança através de um modelo markoviano oculto, bem como suas vantagens e desvantagens.

### 4.1 Aproximação da Verossimilhança

Como já destacado no capítulo anterior, a verossimilhança de um modelo dinâmico é dada por uma integral múltipla de ordem grande que em geral não pode ser resolvida de forma direta. Para solucionar este problema é apresentado abaixo uma forma de aproximação da função de verossimilhança.

Assumindo que  $y_t$  e  $\theta_t$  são escalares, a função de verossimilhança do modelo representado pelas equações (2.1) e (2.2) pode ser descrita da seguinte forma:

$$\begin{aligned}\mathcal{L} &= \int \dots \int f(\mathbf{y}, \boldsymbol{\theta}) d\boldsymbol{\theta} \\ &= \int \dots \int f(\theta_1) f(y_1 | \theta_1) \prod_{t=2}^T f(\theta_t | \theta_{t-1}) f(y_t | \theta_t) d\theta_T \dots d\theta_1 \\ &\approx \sum_{i_1=1}^m \dots \sum_{i_T=1}^m P(\theta_1 \in B_{i_1}) f(y_1 | \theta_1 = b_{i_1}^*) \\ &\quad \times \prod_{t=2}^T P(\theta_t \in B_{i_t} | \theta_{t-1} = b_{i_{t-1}}^*) f(y_t | \theta_t = b_{i_t}^*)\end{aligned}\tag{4.1}$$

onde  $f$  é utilizado para representar uma densidade e  $T$  o número de observações.

Pode-se notar que esta aproximação da integração numérica foi ligeiramente modificada. Mais precisamente, usou-se uma aproximação do tipo:

$$\int_a^b g(x)h(x)dx \approx \int_a^b g(x)dxh\left(\frac{a+b}{2}\right)$$

em vez de

$$\int_a^b g(x)h(x)dx \approx g\left(\frac{a+b}{2}\right)h\left(\frac{a+b}{2}\right).$$

Essa aproximação da verossimilhança de fato corresponde à de um modelo markoviano oculto adequadamente estruturado. Para isso, considerou-se os pontos médios  $b_i^*$ ,  $i = 1, \dots, m$ , como possíveis valores de uma cadeia de Markov  $h_t$  com  $m$  estados, com matriz de probabilidade de transição  $\Gamma = \gamma_{ij}$ , onde:

$$\gamma(i, j) = P(\theta_t \in B_j \mid \theta_{t-1} = b_i^*),$$

e com distribuição inicial  $\delta$ , onde  $\delta_i := P(\theta_0 \in B_i)$ . Dessa forma, as probabilidades de transição  $\gamma_{ij}$  para  $i, j = 1, \dots, m$  são determinadas pela equações de estado do modelo dinâmico.

Por exemplo, se o processo de estado é um autoregressivo de ordem 1 (AR(1)) com parâmetros  $\phi$  e  $\sigma$ , então,

$$\gamma_{ij} = \Phi\left(\frac{(b_j - \phi b_i^*)}{\sigma}\right) - \Phi\left(\frac{(b_{j-1} - \phi b_i^*)}{\sigma}\right)$$

onde  $\Phi(\bullet)$  representa a função de distribuição acumulada da normal padrão. Além disso, definimos:

$$P(y_t) := \text{diag}(f(y_t|\theta_t = b_1^*), \dots, f(y_t|\theta_t = b_m^*)),$$

onde  $f(y_t|\theta_t = b_i^*)$  é determinada pela equação de observações do modelo dinâmico. Com isso, pode-se reescrever a aproximação da verossimilhança utilizando notação de modelo markoviano oculto, como segue abaixo:

$$\mathcal{L} \approx \delta P(y_1) \Gamma P(y_2) \Gamma \dots \Gamma P(y_T) \mathbf{1}^\top,$$

Em resumo, a verossimilhança do modelo dinâmico pode ser aproximada utilizando integração numérica, e essa aproximação é representada por um modelo markoviano oculto

determinado pela cadeia de Markov  $h_t$  e com função de probabilidade associada ao estado dependente com densidade  $f(y_t | \theta_t = b_i^*) =: f(y_t | h_t = b_i^*)$ .

A aproximação do modelo markoviano oculto será estacionário se  $\delta$  for a distribuição estacionária implícita na matriz de transição  $\Gamma$ , ou seja, se  $\delta\Gamma = \delta$ , sujeita à restrição  $\sum_i \delta_i = 1$ .

A escolha de  $m$  tem uma forte influência sobre a precisão da aproximação. A precisão melhora à medida que  $m$  aumenta, mas com isso o tamanho das matrizes na equação  $\mathcal{L}$  também aumentam, o que retarda a avaliação da verossimilhança.

Enquanto  $m$  necessita ser grande o suficiente para fornecer uma boa aproximação, o número de parâmetros do modelo não depende de  $m$ ; as entradas da matriz  $\Gamma(m \times m)$  dependem apenas da equação de estado do modelo dinâmico.

Enquanto a ideia da aproximação da verossimilhança não é original, a questão do modelo markoviano oculto ainda não foi discutida na literatura. Além disso, a aproximação da verossimilhança é modificada se comparada com os métodos citados na seção 3.1. Uma vantagem do método de estimação proposto por meio do modelo markoviano oculto é que existem fórmulas explícitas simples para os resíduos e para previsão das distribuições de um HMM. Além disso, as estimativas do processo latente podem ser obtidas utilizando o Algoritmo de Viterbi.

A formulação de um modelo markoviano oculto o torna particularmente simples de considerar uma variedade de modelos dinâmicos não padrões que são fáceis de implementar. O método discutido é particularmente fácil de aplicar em caso de espaços de estado unidimensionais. No entanto, ele sofre a chamada “maldição de dimensionalidade”, e por isso os modelos com espaços de estados com mais do que duas, ou talvez três dimensões, podem precisar recorrer ao método MCMC.



# Capítulo 5

## Aplicações - Modelo de Volatilidade Estocástica

Os modelos ARCH (autoregressive conditional heteroscedasticity) introduzidos por Engle (1982) [5] e a generalização GARCH (generalized autoregressive conditional heteroscedasticity) proposta por Bollerslev (1986) [6] são amplamente aplicados para modelar a volatilidade de séries financeiras (Taylor 1982 [1], Bollerslev et al. 1992 [7], Engle 2001 [10]).

O uso de modelos de volatilidade estocástica (Stochastic volatility ou SV) tem sido uma alternativa satisfatória para analisar séries temporais financeiras em comparação com os modelos usuais tipo GARCH (Kim Shephard 1998 [8]).

Os modelos SV são mais flexíveis para modelar séries financeiras, pois assumem dois processos: um para as observações e outro para as volatilidades latentes.

Usualmente, essa classe de modelos é representada da seguinte forma:

$$y_t = \beta \exp \left\{ \frac{\theta_t}{2} \right\} \varepsilon_t, \quad (5.1)$$

$$\theta_t = \phi \theta_{t-1} + \sigma_\eta \eta_t, \quad (5.2)$$

onde  $y_t$  e  $\theta_t$  são respectivamente os retornos compostos e o logaritmo da volatilidade no tempo  $t$ . Assume-se inovações  $\varepsilon_t$  e  $\eta_t$  mutuamente independentes e normalmente distribuídas com média zero e variância unitária.

## 5.1 Estudo de Simulações - Modelo de Volatilidade Estocástica (SV)

Para o modelo de volatilidade estocástica descrito nas equações (5.1) e (5.2), foram simuladas  $T = 2.000$  observações no R.

As figuras 5.1 e 5.2 representam as observações e os estados simulados respectivamente.

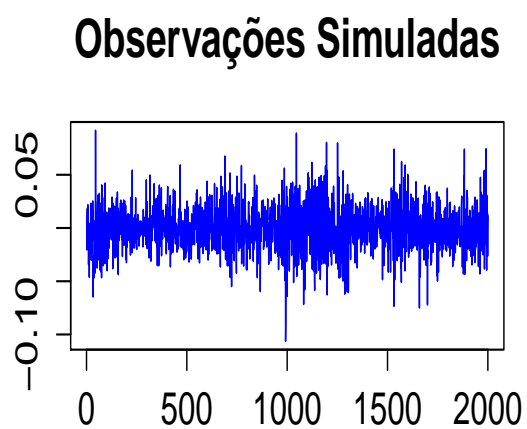


Figura 5.1: Modelo de Volatilidade Estocástica (SV) - Observações Simuladas

## Estados Simulados

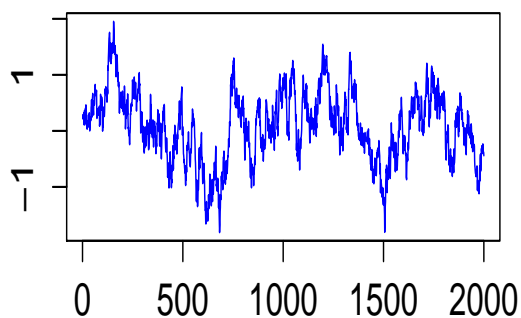


Figura 5.2: Modelo de Volatilidade Estocástica (SV) - Estados Simulados

A tabela 5.1 apresenta os resultados obtidos através da metodologia descrita no capítulo anterior. Pode-se perceber que a escolha do  $m$  requer atenção, já que se for escolhido um  $m$  pequeno, a estimação não trará um resultado tão próximo do valor verdadeiro, e se o  $m$  for grande demais, o custo computacional será maior. Logo, o  $m$  deve ser escolhido de forma com que haja um equilíbrio entre custo computacional e um resultado consistente. No caso apresentado abaixo, deve-se trabalhar com  $m=50$ .

Tabela 5.1: Estudo de Simulações - Modelo de Volatilidade Estocástica (SV) ( $T = 2.000$ ), com  $\phi = 0,98$ ,  $\sigma = 0,2$ ,  $\beta = 0,05$

$m$	$\log \mathcal{L}$	tempo (seg)	$\hat{\phi}$	$\hat{\sigma}$	$\hat{\beta}$
50	15.095,6	35	0,982 (0,975 ; 0,990)	0,198 (0,189 ; 0,219)	0,05 (0,045 ; 0,058)
100	15.095,6	85	0,982 (0,975 ; 0,990)	0,202 (0,184 ; 0,221)	0,05 (0,045 ; 0,056)
200	15.095,6	256	0,982 (0,975 ; 0,990)	0,202 (0,184 ; 0,221)	0,05 (0,045 ; 0,056)

A tabela 5.2 mostra algumas medidas que comprovam que os parâmetros foram bem estimados. A primeira e a segunda coluna, indicam respectivamente os parâmetros e os valores verdadeiros. As duas colunas posteriores, apresentam respectivamente o viés

médio e o erro quadrático médio (EQM) com valores bem pequenos, podendo-se concluir que o ajuste do modelo está consistente. Por fim, as últimas duas colunas indicam o intervalo de confiança e o seu respectivo nível de confiança.

Tabela 5.2: Estudo de Simulações - Modelo de Volatilidade Estocástica (SV) ( $T = 2.000$  e  $m = 50$ ) e com  $\phi = 0,98$ ,  $\sigma = 0,15$ ,  $\beta = 0,02$

Parâmetro	Valor Verdadeiro	Viés Médio	Erro Quadrático Médio	IC	%
$\phi$	0,98	0,0039	0,00017	(0,9536 ; 0,9986)	93,00
$\sigma$	0,15	-0,0060	0,00073	(0,0942 ; 0,2177)	99,00
$\beta$	0,02	-0,0007	0,00004	(0,0331 ; 0,0482)	84,00

## 5.2 Aplicações em Dados Reais - Modelo de Volatilidade Estocástica (SV)

Para ilustrar uma aplicação da metodologia em um modelo de volatilidade estocástica, considerou-se uma série de retornos diários do Ibovespa.

O Ibovespa é um índice de cerca de 50 ações que são negociadas na Bolsa de Valores de São Paulo (Bovespa). O índice é composto por uma carteira teórica com as ações que representavam 80% do volume negociado nos últimos 12 meses e que foram negociados em pelo menos 80% dos dias de negociação. Ele é revisado trimestralmente.

Extraídas do site “finance.yahoo.com”, a série é composta por  $T = 2.000$  observações, de forma a contemplar os preços de fechamento de 19/04/2006 a 16/05/2014. A partir do preços de fechamento  $p_t$ , o retorno foi calculado da seguinte forma:

$$y_t = \log \frac{p_t}{p_{t-1}}$$

A figura 5.3 apresenta os retornos da Ibovespa entre 19/04/2006 e 16/05/2014.

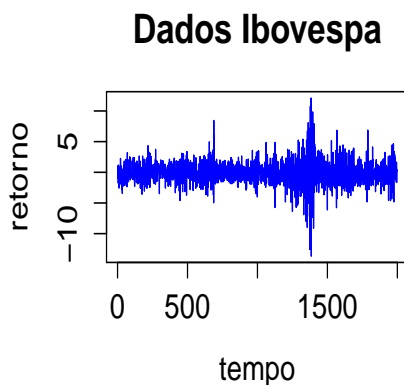


Figura 5.3: Retorno Ibovespa

Tabela 5.3: Retornos Ibovespa

Média	Desvio Padrão	Curtose
$1.92 \times 10^{-17}$	1.8529	6.3243

A partir da aplicação da metodologia no conjunto de dados descrito anteriormente, foram obtidos os resultados expostos na tabela apresentada abaixo. Nela, podemos ver que os parâmetros estimados apresentam resultados bem próximo para os dois  $m$ 's escolhidos. Assim, pode-se concluir que os resultados estão consistentes. Como foi dito anteriormente, uma das vantagens da metodologia em questão é o custo computacional. Nesse exemplo, obtivemos os resultados, em média com 348 segundos.

Tabela 5.4: IBOVESPA (19/04/2006/ a 16/05/2014) com  $T = 2.000$ 

$m$	$g_{max} = -g_{min}$	$\hat{\phi}$	$\hat{\sigma}$	$\hat{\beta}$
100	3	0,9816 (0,9693 ; 0,9877)	0,1436 (0,1042 ; 0,1629)	1,5414 (1,3075 ; 1,6561)
	4	0,9805 (0,9677 ; 0,9933)	0,1465 (0,1051 ; 0,1879)	1,4975 (1,2467 ; 1,7482)
200	3	0,9816 (0,9693 ; 0,9940)	0,1444 (0,1052 ; 0,1836)	1,5426 (1,3082 ; 1,7772)
	4	0,9805 (0,9677 ; 0,9933)	0,1480 (0,1069 ; 0,1891)	1,4986 (1,2478 ; 1,7494)

O  $\phi$  está perto de 0,98 em todos os casos. Este fato indica uma alta persistência da volatilidade.

A figura 5.4 apresenta o gráfico da reconstrução da volatilidade. Observando a linha vermelha, pode-se perceber que ela acompanha a série de dados. Para fazer a reconstrução, foi utilizado o Algoritmo de Viterbi.

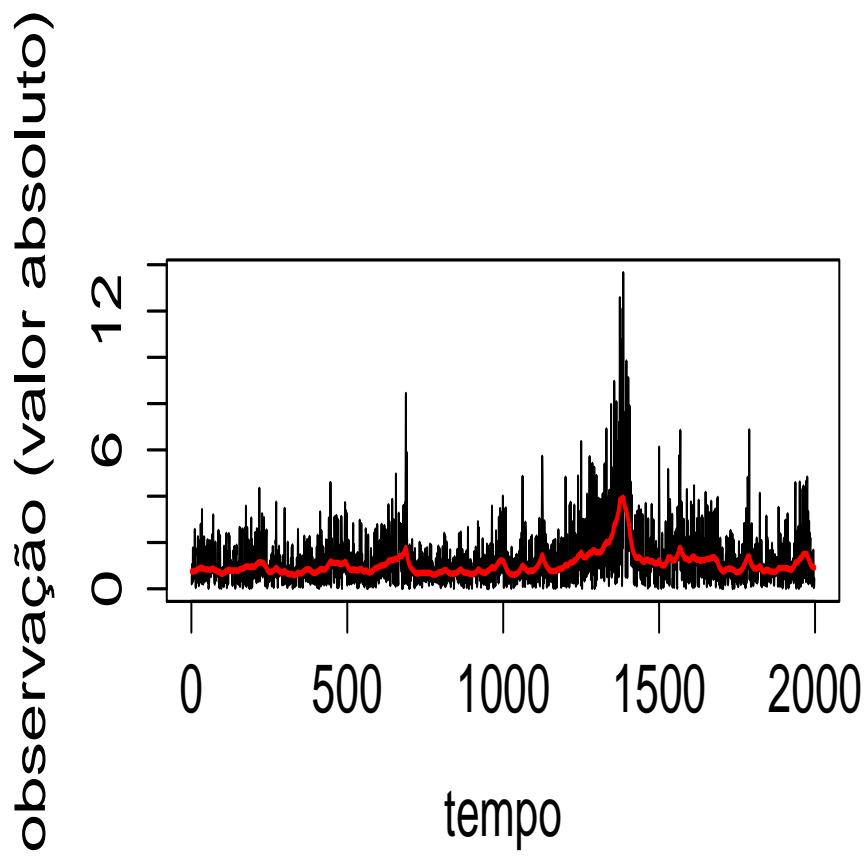


Figura 5.4: Reconstrução da Volatilidade

# Capítulo 6

## Aplicações - Modelo para Dados Binários

Em muitas áreas de aplicações o pesquisador pode encontrar observações que assumem valores 0 e 1. Observações binárias são coletadas juntamente com informações de covariáveis, podem ser contínuas, discretas ou até mesmo categóricas.

Séries temporais binárias podem ser descritas por modelos lineares generalizados. No entanto, se existir correlação na série, estes modelos podem não ser adequados, e algumas aproximações podem ser feitas. Os modelos lineares generalizados dinâmicos podem contornar o problema.

Uma questão crítica na modelagem de observações binárias é a escolha da função de ligação. No contexto de problemas de regressão binária, funções de ligações simétricas tem sido muito utilizadas. No entanto, quando a probabilidade de uma resposta se aproximar a zero é muito maior do que a probabilidade da resposta se aproximar de um ou vice-versa, a função de ligação simétrica se torna inadequadas. Para lidar com esse problema, funções de ligação assimétricas devem ser utilizadas, como por exemplo a função de ligação skew-t generalizada, skew-probit e a função de ligação GEV.



## 6.1 Estudo de Simulações - Modelo para Dados Binários

Essa seção é composta por estudos de simulações para três diferentes tipos de função de ligação.

### 6.1.1 Modelo 1: Função de Ligação T-Student

Considere uma série temporal  $y_t$ , para  $t = 1, \dots, T$ , em que cada termo pode assumir valores 1 ou 0 com probabilidade de sucesso  $\pi_t$ . Assume-se que  $\pi_t$  está relacionada com um vetor de covariáveis  $x_t$  e um vetor  $q$ -dimensional de estados latentes  $\theta_t$ .

$$y_t \sim \text{Bern}(\pi_t) \quad t = 1, \dots, T \quad (6.1)$$

$$\pi_t = H(\mathbf{x}'_t \beta + \theta_t) \quad (6.2)$$

$$\theta_t = \phi \theta_{t-1} + \sigma \eta_t \quad \eta_t \sim N_q(\mathbf{0}, \mathbf{W}_t) \quad (6.3)$$

Na configuração definida acima, o processo observacional  $Y_t$  é descrito pelas equações (6.2) e (6.3), onde  $\pi_t = P(Y_t = 1 \mid \theta_t, x_t)$  é a probabilidade condicional de sucesso. O processo latente  $\theta_t$  é descrito por um processo markoviano de primeira ordem na equação (6.3), onde  $\mathbf{W}_t$  é a matriz de variância-covariância do erro  $\eta_t$ . *Bern* e  $N_q$  indicam respectivamente a distribuição de Bernoulli e distribuição Normal multivariada de dimensão  $q$ . Na terminologia de modelos dinâmicos generalizados,  $F$  é a função de ligação inversa. Este modelo refere-se a  $H$  como função de ligação da T-student.

Para o modelo descrito acima, foram simuladas  $T=400$  observações no  $R$ . Nas tabelas 6.1 e 6.2 encontram-se os resultados para  $g_{max} = -g_{min} = 2$  e  $g_{max} = -g_{min} = 4$ .

Tabela 6.1: Estudo de simulações:  $g_{max} = -g_{min} = 2$  - Função de Ligação T-Student.

Parâmetros	Valor Verdadeiro	Média	EQM	VMR	VMRA
$m = 50$					
$\phi$	0.9800	0.9431	0.0014	0.0376	0.0386
$\sigma$	0.1500	0.2324	0.0068	-0.5490	0.6985
$\beta_1$	-0.1000	-0.0066	0.0087	0.9337	0.9835
$\beta_2$	0.5000	0.5360	0.0013	-0.0719	0.1801
$\beta_3$	-0.4500	-0.7955	0.1194	-0.7678	0.7678
$\nu$	5.0000	9.3084	18.5621	-0.8617	1.1632
$m = 100$					
$\phi$	0.9800	0.9404	0.0016	0.0404	0.0410
$\sigma$	0.1500	0.2333	0.0069	-0.5550	0.6843
$\beta_1$	-0.1000	-0.0030	0.0094	0.9701	0.9947
$\beta_2$	0.5000	0.5267	0.0007	-0.0534	0.1770
$\beta_3$	-0.4500	-0.7908	0.1162	-0.7574	0.7574
$\nu$	5.0000	9.2390	17.9695	-0.8478	1.1312
$m = 200$					
$\phi$	0.9800	0.9460	0.0012	0.0347	0,0355
$\sigma$	0.1500	0.2294	0.0063	-0.5295	0.6870
$\beta_1$	-0.1000	-0.0037	0.0093	0.9628	1.0068
$\beta_2$	0.5000	0.5306	0.0009	-0.0611	0.1823
$\beta_3$	-0.4500	-0.7950	0.1190	-0.7667	0.7667
$\nu$	5.0000	9.3148	18.6173	-0.8630	1.1755

Tabela 6.2: Estudo de simulações:  $g_{max} = -g_{min} = 4$  - Função de Ligação T-Student.

Parâmetros	Valor Verdadeiro	Média	EQM	VMR	VMRA
$m = 50$					
$\phi$	0.9800	0.9326	0.0022	0.0484	0.0490
$\sigma$	0.1500	0.2540	0.0108	-0.6936	0.8307
$\beta_1$	-0.1000	0.0036	0.0107	1.0362	1.0367
$\beta_2$	0.5000	0.5536	0.0029	-0.1072	0.2170
$\beta_3$	-0.4500	-0.8285	0.1432	-0.8411	0.8411
$\nu$	5.0000	9.6204	21.3485	-0.9241	1.2578
$m = 100$					
$\phi$	0.9800	0.9298	0.0025	0.0512	0.0517
$\sigma$	0.1500	0.2601	0.0121	-0.7339	0.8498
$\beta_1$	-0.1000	0.0004	0.0101	1.0036	1.0036
$\beta_2$	0.5000	0.5468	0.0022	-0.0937	0.2125
$\beta_3$	-0.4500	-0.8172	0.1349	-0.8161	0.8161
$\nu$	5.0000	10.1608	26.6339	-1.0322	1.3379
$m = 200$					
$\phi$	0.9800	0.9309	0.0024	0.0501	0.0508
$\sigma$	0.1500	0.2541	0.0108	-0.6938	0.8291
$\beta_1$	-0.1000	0.0009	0.0102	1.0091	1.0091
$\beta_2$	0.5000	0.5420	0.0018	-0.0841	0.2108
$\beta_3$	-0.4500	-0.8160	0.1339	-0.8133	0.8133
$\nu$	5.0000	10.2043	27.0845	-1.0409	1.3414

### 6.1.2 Modelo 2: Função de Ligação Probit

Considere uma série temporal  $y_t$ , para  $t = 1, \dots, T$ , em que cada termo pode assumir valores 1 ou 0 com probabilidade de sucesso  $\pi_t$ . Assume-se que  $\pi_t$  está relacionada com um vetor de covariáveis  $x_t$  e um vetor  $q$ -dimensional de estados latentes  $\theta_t$ .

$$y_t \sim \text{Bern}(\pi_t) \quad t = 1, \dots, T \quad (6.4)$$

$$\pi_t = F(\theta_t) \quad (6.5)$$

$$\theta_t = \phi\theta_{t-1} + \sigma\eta_t \quad \eta_t \sim N_q(\mathbf{0}, \mathbf{W}_t) \quad (6.6)$$

Na configuração definida acima, o processo observacional  $Y_t$  é descrito pelas equações (6.5) e (6.6), onde  $\pi_t$  é a probabilidade condicional de sucesso. O processo latente  $\theta_t$  é descrito por um processo markoviano de primeira ordem na equação (6.6), onde  $\mathbf{W}_t$  é a matriz de variância-covariância do erro  $\eta_t$ .  $\text{Bern}$  e  $N_q$  indicam respectivamente a distribuição de Bernoulli e distribuição Normal multivariada de dimensão  $q$ . Na terminologia de modelos dinâmicos generalizados,  $F$  é a função de ligação inversa. Para facilitar a

exposição, refere-se a F como função de ligação probit.

Para o modelo descrito acima, foram simuladas T=400 observações no R. Nas tabelas 6.3 e 6.4 encontram-se os resultados para  $g_{max} = -g_{min} = 2$ ,  $g_{max} = -g_{min} = 3$  e  $g_{max} = -g_{min} = 4$ .

Tabela 6.3: Estudo de simulações:  $g_{max} = -g_{min} = 2$  - Função de Ligação Probit.

Parâmetros	Valor Verdadeiro	Média	EQM	VMR	VMRA
$m = 100$					
$\phi$	0.9500	0.9182	0.0010	0.0335	0.0645
$\sigma$	0.2000	0.2531	0.0028	-0.2656	0.7715
$m = 200$					
$\phi$	0.9500	0.9194	0.0009	0.0322	0.0641
$\sigma$	0.2000	0.2521	0.0027	-0.2603	0.7687
$m = 300$					
$\phi$	0.9500	0.9194	0.0009	0.0322	0.0641
$\sigma$	0.2000	0.2521	0.0027	-0.2603	0.7689

Tabela 6.4: Estudo de simulações:  $g_{max} = -g_{min} = 3$  - Função de Ligação Probit.

Parâmetros	Valor Verdadeiro	Média	EQM	VMR	VMRA
$m = 100$					
$\phi$	0.9500	0.9066	0.0019	0.0457	0.0754
$\sigma$	0.2000	0.2762	0.0058	-0.3808	0.9074
$m = 200$					
$\phi$	0.9500	0.9067	0.0019	0.0456	0.0759
$\sigma$	0.2000	0.2767	0.0059	-0.3834	0.9062
$m = 300$					
$\phi$	0.9500	0.9067	0.0019	0.0456	0.0759
$\sigma$	0.2000	0.2767	0.0059	-0.3836	0.9064

Tabela 6.5: Estudo de simulações:  $g_{max} = -g_{min} = 4$  - Função de Ligação Probit.

Parâmetros	Valor Verdadeiro	Média	EQM	VMR	VMRA
$m = 100$					
$\phi$	0.9500	0.9022	0.0023	0.0503	0.0795
$\sigma$	0.2000	0.2899	0.0081	-0.4494	0.9742
$m = 200$					
$\phi$	0.9500	0.9027	0.0022	0.0498	0.0795
$\sigma$	0.2000	0.2901	0.0081	-0.4505	0.9762
$m = 300$					
$\phi$	0.9500	0.9028	0.0022	0.0497	0.0800
$\sigma$	0.2000	0.2905	0.0082	-0.4523	0.9750

### 6.1.3 Modelo 3: Função de Ligação GEV

Considere uma série temporal  $y_t, t = 1, \dots, T$ , em que cada termo pode assumir valores 1 ou 0 com probabilidade de sucesso  $\pi_t$ . Assume-se que  $\pi_t$  está relacionada com um vetor de covariáveis  $x_t$  e um vetor  $q$ -dimensional de estados latentes  $\theta_t$ .

$$y_t \sim \text{Bern}(\pi_t) \quad t = 1, \dots, T \quad (6.7)$$

$$\pi_t = 1 - p\text{GEV}(-\theta_t) \quad (6.8)$$

$$\theta_t = \phi\theta_{t-1} + \sigma\eta_t \quad \eta_t \sim N_q(\mathbf{0}, \mathbf{W}_t) \quad (6.9)$$

Na configuração definida acima, o processo observacional  $Y_t$  é descrito pelas equações (6.8) e (6.9), onde  $\pi_t$  é a probabilidade condicional de sucesso. O processo latente  $\theta_t$  é descrito por um processo markoviano de primeira ordem na equação (6.9), onde  $\mathbf{W}_t$  é a matriz de variância-covariância do erro  $\eta_t$ .  $\text{Bern}$  e  $N_q$  indicam respectivamente a distribuição de Bernoulli e distribuição Normal multivariada de dimensão  $q$ .

Para o modelo descrito acima, foram simuladas  $T=400$  observações no R. Nas tabelas 6.6 e 6.7 encontram-se os resultados para  $g_{max} = -g_{min} = 2$ ,  $g_{max} = -g_{min} = 3$  e  $g_{max} = -g_{min} = 4$ .

Tabela 6.6: Estudo de simulações:  $g_{max} = -g_{min} = 2$  - Função de Ligação GEV.

Parâmetros	Valor Verdadeiro	Média	EQM	VMR	VMRA
$m = 100$					
$\xi$	-0.0300	0.5261	-0.8261	2.7537	13.8618
$\phi$	0.9500	0.8580	0.0920	0.0968	0.1277
$\sigma$	0.2000	0.2385	-0.0385	-0.1925	0.7796
$m = 200$					
$\xi$	-0.0300	1.3317	-1.6317	5.4391	17.0894
$\phi$	0.9500	0.8502	0.0998	0.1051	0.1350
$\sigma$	0.2000	0.2383	-0.0383	-0.1914	0.7907
$m = 300$					
$\xi$	-0.0300	0.2136	-0.5136	1.7119	21.0312
$\phi$	0.9500	0.8431	0.1069	0.1125	0.1425
$\sigma$	0.2000	0.2387	-0.0387	-0.1935	0.7975

Tabela 6.7: Estudo de simulações:  $g_{max} = -g_{min} = 3$  - Função de Ligação GEV.

Parâmetros	Valor Verdadeiro	Média	EQM	VMR	VMRA
$m = 100$					
$\xi$	-0.0300	0.5343	-0.8343	2.7808	13.38660
$\phi$	0.9500	0.8562	0.0938	0.0987	0.1286
$\sigma$	0.2000	0.2730	-0.0730	-0.3648	0.9851
$m = 200$					
$\xi$	-0.0300	0.8113	-1.1113	3.7140	15.4225
$\phi$	0.9500	0.8536	0.0964	0.1015	0.1310
$\sigma$	0.2000	0.2825	-0.0825	-0.4125	1.0030
$m = 300$					
$\xi$	-0.0300	1.4362	-1.7362	5.7873	17.7906
$\phi$	0.9500	0.8400	0.1100	0.1158	0.1447
$\sigma$	0.2000	0.2897	-0.0897	-0.4484	1.0445

Tabela 6.8: Estudo de simulações:  $g_{max} = -g_{min} = 4$  - Função de Ligação GEV.

Parâmetros	Valor Verdadeiro	Média	EQM	VMR	VMRA
$m = 100$					
$\xi$	-0.0300	1.1859	1.4859	4.9530	13.3878
$\phi$	0.9500	0.8420	0.1080	0.1137	0.1430
$\sigma$	0.2000	0.3215	-0.1215	-0.6073	1.2058
$m = 200$					
$\xi$	-0.0300	0.6532	-0.9532	3.1673	15.0608
$\phi$	0.9500	0.8460	0.1040	0.1091	0.1389
$\sigma$	0.2000	0.3202	-0.1202	-0.5982	1.2051
$m = 300$					
$\xi$	-0.0300	1.1248	-1.4248	4.7493	17.3538
$\phi$	0.9500	0.8419	0.1081	0.1138	0.1425
$\sigma$	0.2000	0.3222	-0.1222	-0.6109	1.2133

## 6.2 Aplicações em Dados Reais - Modelo de Dados Binários

Para ilustrar a técnica aplicada às respostas binários, considerou-se respostas de um macaco que executa paradigma de atenção descrito em Abanto, Dey e Jiang [12].

Estabeleceu-se para esta experiência dados comportamentais compostos por uma série temporal de observações binários, onde a observação 1 corresponde a recompensa a ser entregue e um 0 correspondente a recompensa não ser entregue em cada ensaio.

O objetivo da experiência é determinar se, uma vez que o desempenho diminuiu com o aumento da fadiga, se uma estimulação cerebral profunda (DBS - Deep Brain Stimulation) permite que o animal possa recuperar seu nível de desempenho pré-fadiga. Nesta experiência, o macaco performou 1250 vezes, sendo 741 (ou 59,28%) respostas 1 corretas. Para este conjunto de dados foi possível ajustar o modelo de espaço de estado binário com duas funções de ligação padrão.

A primeira função de ligação utilizada para ajustar o modelo foi a função de ligação Probit, que pode ser representada pelas equações (6.5) e (6.6).

A segunda e última função de ligação utilizada para ajustar o modelo foi a função de ligação GEV, que pode ser representada pelas equações (6.8) e (6.9).

Tabela 6.9: Aplicações em Dados Reais - Função de Ligação Probit.

$g_{max} = -g_{min}$	$\hat{\phi}$	$\hat{\sigma}$	tempo (segundos)
$m = 100$			
2	0.9978	0.0676	4
3	0.9975	0.0777	4
4	0.9975	0.0769	5
$m = 200$			
2	0.9978	0.0683	14
3	0.9975	0.0793	20
4	0.9975	0.0795	14
$m = 300$			
2	0.9978	0.0684	29
3	0.9975	0.0795	34
4	0.9975	0.0800	31

Tabela 6.10: Aplicações em Dados Reais - Função de Ligação GEV.

$g_{max} = -g_{min}$	$\hat{\phi}$	$\hat{\sigma}$	$\hat{\xi}$	tempo(segundos)
$m = 100$				
2	0.9969	0.0458	-7.0159	12
3	0.9968	0.0429	-6.4674	15
4	0.9972	0.0469	-8.1715	10
$m = 200$				
2	0.9970	0.0479	-7.6045	37
3	0.9970	0.0465	-7.2980	42
4	0.9969	0.0493	-7.0094	41
$m = 300$				
2	0.9971	0.0487	-7.8207	57
3	0.9970	0.0476	-7.5898	59
4	0.9971	0.0493	-8.1445	61



Observando as tabelas acima, pode-se perceber que para os dois modelos considerados, o valor estimado para  $\phi$  está bem próximo de 1. Dessa forma, mostrando a maior persistência dos parâmetros autoregressivos para as variáveis estados.

Os resultados apresentados na tabela 6.10 indicam que há uma assimetria negativa da função, já que o parâmetro  $\xi$  apresenta valores negativos afastados de zero.

Tendo como base de comparação os resultados obtidos através de aplicação de MCMC em Abanto, Dey e Jiang [15], pode-se concluir que os resultados da estimação via HMM são consistentes e satisfatórios.

# Capítulo 7

## Conclusão

A metodologia de aproximação de verossimilhança via modelos markovianos ocultos apresentada neste trabalho, fornece um artifício conveniente e flexível para a implementação de diversas classes de modelos dinâmicos, podendo ser um modelo linear e Gaussiano ou não. Apesar de fornecer bons resultados, o esforço computacional envolvido na modelagem é modesto. Além disso, o algoritmo de Viterbi pode ser aplicado para a fazer a reconstrução da volatilidade.

Foram propostos dois modelos para aplicação da metodologia. O primeiro foi o de volatilidade estocástica que tem sido amplamente utilizado para modelar séries temporais financeira pois é muito flexível, já que possui dois processo, um para as observações e outro para os estados. Para ilustrar seu funcionamento foram feitas aplicações tanto para observações simuladas quanto para dados reais. Dessas aplicações pudemos concluir que os parâmetros foram bem estimados e com baixo custo computacional.

O segundo, é um modelo para observações binárias. Foram feitas aplicações para observações simuladas e para dados reais, e também pudemos concluir que os parâmentros forma bem estimados e com baixo custo computacional.

Por fim, como proposta para trabalhos futuros, sugere-se a aplicação da metodologia do ponto de vista bayesiano. Assim, tornando possível fazer a comparação dos resultados obtidos através outras aproximação de enfoque bayesiano.

# Apêndice A

## Programação no R

### A.1 Aproximação da Verossimilhança

#### 1. Preparando os parâmetros

```
pn2pw <- function(beta,phi,sigma){
  lbeta<- log(beta)
  lphi <- log((1+phi)/(1-phi))
  lsigma <- log(sigma)
  parvect <- c(lbeta,lphi,lsigma)
  return(parvect)
}

pw2pn <- function(parvect){
  beta <- exp(parvect[1])
  phi <- (exp(parvect[2])-1)/(exp(parvect[2])+1)
  sigma <- exp(parvect[3])
  return(list(beta=beta,phi=phi,sigma=sigma))
}
```

#### 2. Aproximando a Verossimilhança

```
mllk2 <-function(parvect,y,m){
  p <- pw2pn(parvect)
```

```

K= m+1
b=seq(-gmax,gmax,length=K)
bs=(b[-1]+b[-K])*0.5
sey= p$beta*exp(bs/2)
Gamma=matrix(0,m,m)
for (i in 1:m){
Gamma[i,]=diff(pnorm(b,p$phi*bs[i],p$sigma))}
Gamma= Gamma/apply(Gamma,1,sum)
foo= solve(t(diag(m)-Gamma+1),rep(1,m))
llk=0
for (t in 1:length(y)){
foo=foo%*%Gamma*dnorm(y[t],0,sey)
sumfoo= sum(foo)
llk=llk+log(sumfoo)
foo=foo/sumfoo
}
return(-llk)
}

```

## A.2 Reconstrução da Volatilidade

```

viterbi <- function(x,m,gbmax,mod){
if (model=="SV0") mod$nu <- 10^10
K <- m+1
gb <- seq(-gbmax,gbmax,length=K)
g <- (gb[-1]+gb[-K])*0.5
beg <- mod$mu*exp(g/2)
gamma <- matrix(0,m,m)
E <- mod$phi*g
intlen <- gb[2]-gb[1]
for (i in 1:m){
goo <- dnorm(g,E[i],mod$sigma)*intlen
gamma[i,] <- goo/sum(goo)
}
}

```

```

delta <- dnorm(g,0,mod$sigma/sqrt(1-mod$phi^2))*intlen
n <- length(x)
xi <- matrix(0,n,m)
foo <- delta*1/beg*dt(x[1]/beg,mod$nu)
xi[1,] <- foo/sum(foo)
for (i in 2:n){
  foo <- apply(xi[i-1,]*gamma,2,max)*1/beg*dt(x[i]/beg,mod$nu)
  xi[i,] <- foo/sum(foo)
}
iv <- numeric(n)
iv[n] <-which.max(xi[n,])
for (i in (n-1):1){
  iv[i] <- which.max(gamma[,iv[i+1]]*xi[i,])
}
return ( iv)
}

```

# Referências Bibliográficas

- [1] S. TAYLOR. “Financial returns modeled by the product of two stochastic processes, a study of daily sugar prices”, *Time Series Analysis: Theory and Practice*, Amsterdam-Holland (1982).
- [2] G. PETRIS, S. PETRONE, P. CAMPAGNOLI. *Dynamic Linear Models with R*, Springer (2009).
- [3] R. KALMAN. “A new approach to linear filtering and prediction problems”, *Journal of Basic Engineering* (1960), 82 , 35-45.
- [4] R. KALMAN, R. BUCY. “New results in linear filtering and prediction theory”, *Journal of Basic Engineering* (1960), 83D, 95108
- [5] R. ENGLE. “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation”, *Econometric* (1982), 50, 987-1007.
- [6] T. BOLLERSLEV. “Generalized autoregressive conditional heterosdasticity”, *Journal of Econometrics*(1986), 31, 307-327.
- [7] T. BOLLERSLEV, R. CHOU, K. KRONER. “Arch modeling in finance: a review of the theory and empirical evidence”, *Journal of Econometrics* (1992), 52, 559.
- [8] S. KIM, N. SHERPHARD. “Stochastic volatility: likelihood inference and comparison with arch models”, *Review of Economic Studies* (1998), 65, 361393.
- [9] R. MEYER, T. YU. “Bugs for a bayesian analysis of stochastic volatility models”, *Econometrics Journal* 3, (2000), 198215.
- [10] R. ENGLE. “Garch 101: The use of arch/garch models in applied econometrics”, *Journal of Economic Perspectives* (2001), 15, 157168.

- [11] C. ABANTO, H. MIGON, H. LOPES. “Bayesian modeling of financial returns: A relationship between volatility and trading volume”, *Applied Stochastic Models in Business and Industry* (2010), 26, 172-193.
- [12] R. LANGROCK. “Some applications of nonlinear and non-Gaussian state-space modeling by means of hidden Markov models”, *Journal of Application Statistics* (2011), 38, 2955-2970.
- [13] R. LANGROCK, I. MACDONALDS, W. ZUCCHINI. “Some nonstandard stochastics volatility models and their estimation using structured hidden Markov models”, *Journal of Empirical Finance* (2012), 19, 147-161.
- [14] C. ABANTO, D. DEY. “State Space mixed models for binary responses with scale mixture of normal distribution links”, *Computational Statistics and Data Analysis* (2014), 71, 274-287.
- [15] C. ABANTO, D. DEY, X.JIANG. “Binary state mixed models with flexible link function: a case study on deep brain stimulation on attention reaction time”, *Statistics and Its Inference* (2015), 8, 187-194.
- [16] G. F. LAWLER. *Introduction to Stochastic Processes*, Chapman Hall/CRC Probability Series (2000).
- [17] C. ABANTO. “Métodos de Simulação Estocástica em Modelos Dinâmicos não Lineares: Uma Aplicação em Modelos de Volatilidade”, *Tese de Doutorado do Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro*(2005).
- [18] R. LANGROCK. “On some special-purpose hidden Markov models”, *Dissertation presented for the degree of Doctor of Philosophy at the Faculty of Economics and Business Administration of the Georg-August-Universität Göttingen*
- [19] Finance Yahoo, Market Data. Disponível em: <http://finance.yahoo.com/market-overview/>.