

**Análise bayesiana  
do modelo fatorial dinâmico  
para um vetor de séries temporais  
usando distribuições elípticas**

LÍVIA COSTA BORGES

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

Programa: **Estatística**

Orientadora: **Profa. Dra. Lúcia Pereira Barroso**

São Paulo, agosto de 2008.

**Análise bayesiana  
do modelo fatorial dinâmico  
para um vetor de séries temporais  
usando distribuições elípticas**

Este exemplar corresponde à redação final  
da tese devidamente corrigida e  
defendida por Livia Costa Borges  
e aprovada pela comissão julgadora.

Banca Examinadora:

Profa. Dra. Lúcia Pereira Barroso (presidente) - IME/USP

Prof. Dr. Pedro A. Morettin - IME/USP

Prof. Dr. Daniel Furtado Ferreira - DEX/UFLA

Prof. Dr. Marinho Gomes de Andrade Filho - ICMC/USP

Profa. Dra. Thelma Sáfyadi - DEX/UFLA

*Aos meus pais, Albertino e Sara  
e às minhas irmãs Andreísa e Valéria.*

# Agradecimentos

Desejo expressar meus mais sinceros agradecimentos à minha família e aos meus amigos queridos por me acompanharem e apoiarem, sempre. Obrigado por serem tão pacientes comigo e por suportarem a minha ausência várias vezes durante este período de dedicação quase exclusiva aos estudos. Aos meus pais e irmãs agradeço pelo estímulo e apoio incondicional, desde o primeiro instante. Aos meus sobrinhos Gabriel e Vinícius, agradeço pelo fato de existirem em minha vida. Mesmo sem entender suportaram minha ausência.

À Prof. Lúcia Pereira Barroso, pela orientação, paciência, críticas e sugestões feitas durante a orientação.

À minha amiga Tatiana Terabayashi Melhado, pela amizade e disponibilidade desde o nosso primeiro contato. O seu apoio foi fundamental durante esta jornada.

Minha gratidão a todo o pessoal do Departamento de Ciências Exatas da Universidade Federal de Lavras - DEX/UFLA, especialmente aos professores Thelma Sáfydi pela co-orientação e Daniel Furtado Ferreira, por apostar e acreditar em minha capacidade e pela disponibilidade em ajudar sempre. Ao Professor Daniel Furtado dedico esta vitória. Serei eternamente grata aos meus colegas de mestrado no DEX, Marcelo Cirillo e Denys Nogueira, pelo apoio.

A Juan Carlos Ruilova, pela atenção especial concedida na etapa final do meu trabalho.

Aos meus amigos do Citibank, Carlos Auricchio, Ricardo Almeida, Jack, Fer e Paulo Sirimarco pela força e pelos momentos legais proporcionados.

Obrigado, mil vezes obrigado, a meus mentores e anjos da guarda por me mostrarem a luz, me acompanharem e me orientarem em minha jornada. Um agradecimento especial ao meu terapeuta Dr. Mateus, que me ajudou a crescer muito.

À Capes, pela bolsa de doutorado concedida.

# Resumo

A análise fatorial é uma importante ferramenta estatística que tem amplas aplicações práticas e explica a correlação entre um grande número de variáveis observáveis em termos de um pequeno número de variáveis não observáveis, conhecidas como variáveis latentes. A proposta deste trabalho é fazer a análise Bayesiana, que incorpora à análise o conhecimento que se tenha sobre os parâmetros antes da coleta dos dados, do modelo fatorial dinâmico na classe de modelos elípticos multivariados, assumindo que a um vetor de  $q$  séries temporais pode-se ajustar um modelo fatorial com  $k < q$  fatores mais um ruído branco, e que a parte latente segue um modelo vetorial auto-regressivo. A classe de modelos elípticos citada acima é rica em distribuições simétricas com caudas mais pesadas que as da distribuição normal, característica importante na análise de séries financeiras. Essa classe inclui as distribuições  $t$  de Student, exponencial potência, normal contaminada, entre outras. A inferência sobre os parâmetros foi feita utilizando métodos de Monte Carlo via Cadeias de Markov, com os algoritmos Metropolis-Hastings e Griddy-Gibbs, através da obtenção das distribuições a posteriori dos parâmetros e dos fatores. A determinação da convergência do processo foi feita por técnicas gráficas e pelos métodos de Geweke (1992), de Heidelberger e Welch (1983) e Half-Width. O método foi ilustrado usando dados reais e simulados.

# Abstract

The factor analysis is an important statistical tool that has wide practical applications and it explains the correlation among a large number of observable variables in terms of a small number of unobservable variables, known as latent variables. The proposal of this work is the Bayesian analysis, which incorporates the information we have concerning the parameters before collecting data into the analysis of a dynamical factor model in the class of multivariate elliptical models, where the factors follow a multivariate autoregressive model, assuming that a vector of  $q$  time series can be adjusted with  $k < q$  factors and a white noise. The class of elliptical models is rich in symmetrical distributions with heavier tails than the normal distribution, which is an important characteristic in financial series analysis. This class includes t-Student, power exponential, contaminated normal and other distributions. The parameters inference was made through Monte Carlo Markov Chain methods, with Metropolis-Hastings and Griddy-Gibbs algorithms, by obtaining the parameters and factors posteriori distributions. The convergence process was made through graphical technics and by Geweke (1992) and by Heidelberger and Welch (1983) and Half-Width methods. The method was illustrated using simulated and real data.

# Índice

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Distribuições Elípticas</b>	<b>5</b>
2.1	Definições e Exemplos . . . . .	7
2.2	Propriedades da Distribuição Elíptica . . . . .	12
2.3	Distribuição t multivariada . . . . .	14
<b>3</b>	<b>Análise Bayesiana</b>	<b>16</b>
3.1	Inferência Bayesiana . . . . .	16
3.2	Métodos de Monte Carlo via Cadeias de Markov . . . . .	18
3.2.1	O Algoritmo Metropolis-Hastings . . . . .	20
3.2.2	O Amostrador Griddy-Gibbs . . . . .	25
3.2.2.1	O Algoritmo Griddy-Gibbs . . . . .	25
3.2.3	Procedimentos para Monitoração da Convergência . . . . .	27
<b>4</b>	<b>Estimação Bayesiana do Modelo Fatorial Proposto</b>	<b>32</b>
4.1	Modelo Fatorial Dinâmico . . . . .	33
4.2	Distribuições a Posteriori . . . . .	35
4.3	Simulação Estocástica via Cadeias de Markov . . . . .	38
4.3.1	Simulação via Algoritmo Metropolis-Hastings . . . . .	38
4.3.2	Simulação via Amostrador Griddy-Gibbs . . . . .	38
4.4	Monitoração da Convergência . . . . .	39

4.4.1	Convergência do Algoritmo Metropolis-Hastings . . . . .	39
4.4.2	Convergência do Amostrador Griddy-Gibbs . . . . .	40
4.4.3	Simulação via Máxima Verossimilhança . . . . .	44
<b>5</b>	<b>Aplicação</b>	<b>48</b>
<b>6</b>	<b>Conclusões</b>	<b>58</b>
<b>A</b>	<b>Programa R - simulação e análise via Metropolis-Hastings.</b>	<b>60</b>
<b>B</b>	<b>Programa S-Plus - simulação e análise via Griddy-Gibbs.</b>	<b>68</b>
<b>C</b>	<b>Séries financeiras utilizadas na aplicação</b>	<b>80</b>
	<b>Bibliografia</b>	<b>84</b>



# Capítulo 1

## Introdução

Inovações metodológicas e aplicações da análise fatorial, ou seja, modelos com estruturas latentes, têm se desenvolvido rapidamente nos últimos anos, devido ao incremento de ferramentas computacionais apropriadas.

A análise fatorial tem como objetivo principal descrever as relações entre um conjunto de variáveis observadas e outras não observáveis, conhecidas como variáveis latentes ou fatores.

O estudo de variáveis latentes teve início no começo do Século XX, com a criação da análise fatorial na área de Psicometria por Spearman (1904), para analisar escores individuais em testes mentais. Esse primeiro trabalho envolveu o uso de somente uma variável latente, a habilidade geral ou inteligência. Thurstone (1944) estendeu essa teoria para modelos com mais variáveis, introduzindo a análise fatorial múltipla, também na área de Psicologia. Mais detalhes a respeito podem ser encontrados em Thurstone (1947).

Somente após 1960 é que o assunto teve tratamento estatístico e um dos trabalhos dessa época é o livro de Lawley e Maxwell (1971).

Ainda nos anos 60 teve início o desenvolvimento de *covariance structure analysis* com o trabalho de Jöreskog (1966), sendo esta técnica conhecida como *modelagem de equações estruturais*. Esse modelo nasceu na área de Ciências Sociais na busca de relações de causalidade e seus resultados são baseados na comparação da matriz de covariância estimada

pelos dados livres de estrutura com a matriz de covariância estimada dado o modelo. Vários softwares como o AMOS, o PRELIS, o CALIS e o LISREL foram desenvolvidos para o ajuste desse modelo, sendo que o último tornou-se um sinônimo do mesmo.

Atualmente, análise fatorial tem sido aplicada em um grande número de situações, por exemplo, para analisar quantidades econômicas, testes de atitudes e comportamentos, medidas físicas, etc.

Peña e Box (1987) propuseram um método para identificar fatores comuns em um vetor de séries temporais. Eles assumem que um vetor de séries temporais observado de dimensão  $q$  pode ser escrito em função de uma matriz de parâmetros desconhecidos com dimensão  $(q \times k)$ , um vetor não observável  $k$ -dimensional e um vetor de erros  $q$ -dimensional (ruído branco), com matriz de covariância  $\Sigma_\epsilon$  de posto completo. Esse modelo é relevante se e somente se  $k < q$  ou se  $k = q$  e  $\Sigma_\epsilon = \mathbf{0}$ . No primeiro caso ocorre redução da dimensionalidade sem perda de informação e no segundo caso, em que  $k = q$ , um problema interessante é encontrar uma transformação linear das séries que permitam uma representação simples do sistema. Os autores mostraram que uma representação simples usando um pequeno número de fatores foi adequada para o modelo de séries temporais e que, diferentemente da análise fatorial para variáveis observadas em unidades independentes, as restrições colocadas pela presença de fatores comuns na correlação, autocorrelação parcial e matrizes paramétricas tornam a identificação do número de fatores e a estimativa da matriz de cargas fatoriais relativamente fáceis.

Diebold e Nerlove (1989) utilizaram um modelo multivariado para um vetor de séries temporais em que a parte latente segue um modelo ARCH (autoregressive conditional heterocedastic) para estudar padrões de volatilidade temporal na taxa de câmbio. Primeiramente, estimaram o modelo univariado e em seguida usaram as informações obtidas para especificar o modelo multivariado. A principal finalidade da aproximação multivariada nesse trabalho foi explorar a estrutura fatorial, que facilita a estimação através da redução do número de parâmetros a serem estimados.

Uma estrutura latente nos modelos de séries temporais foi proposta por Harvey e Chung (2000), aplicada à série de desemprego no Reino Unido. Esse trabalho teve abor-

## Gracias por visitar este Libro Electrónico

Puedes leer la versión completa de este libro electrónico en diferentes formatos:

- HTML(Gratis / Disponible a todos los usuarios)
- PDF / TXT(Disponible a miembros V.I.P. Los miembros con una membresía básica pueden acceder hasta 5 libros electrónicos en formato PDF/TXT durante el mes.)
- Epub y Mobipocket (Exclusivos para miembros V.I.P.)

Para descargar este libro completo, tan solo seleccione el formato deseado, abajo:

