

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Sergio Eduardo Contreras Espinoza

**Modelo em Espaço de Estado para Séries
Temporais com Distribuição Poisson
Bivariada - Uma aplicação da Metodologia
Durbin-Koopman**

Tese de Doutorado

Tese apresentada ao Programa de Pós-graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da PUC-Rio como parte dos requisitos parciais para obtenção do título de Doutor em Engenharia Elétrica

Orientador: Prof. Cristiano Augusto Coelho Fernandes

Rio de Janeiro
Março de 2004



Sergio Eduardo Contreras Espinoza

**Modelo em Espaço de Estado para Séries
Temporais com Distribuição Poisson
Bivariada - Uma aplicação da Metodologia
Durbin-Koopman**

Tese de Doutorado apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Dr. Cristiano Augusto Coelho Fernandes

Orientador

Departamento de Engenharia Elétrica — PUC-Rio

Dr. Antonio Carlos Ponce de Leon

UERJ

Dr. Eduardo Lima Campos

ENCE-IBGE

Dr. Luiz Koodi Hotta

UNICAMP

Dr. Edison Huarsaya Tito

Petrobras

Prof. José Eugenio Leal

Coordenador Setorial do Centro Técnico Científico

PUC-Rio

Rio de Janeiro, 30 de Março de 2004

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Sergio Eduardo Contreras Espinoza

Professor do Departamento de Matemática da Facultad de Ciencias da Universidad del Bío-Bío, Concepción-Chile. Em 1992 obteve o grau de Licenciatura em Matemática e o título de Engenheiro-Matemático, e em 1997 o grau de Mestre em Estatística. Todos os títulos foram obtidos na Universidad de Concepción-Chile. No seu constante desejo de aprimoramento veio para o Rio de Janeiro em 2000, afastando-se de sua atividade profissional, onde iniciou seus estudos de doutorado no Departamento de Engenharia Elétrica da PUC-Rio.

Ficha Catalográfica

Contreras Espinoza, Sergio Eduardo

Modelo em espaço de estado para séries temporais com distribuição Poisson bivariada: uma aplicação da metodologia Durbin-Koopman. / Sergio Eduardo Contreras Espinoza; orientador: Cristiano Augusto Coelho Fernandes. — Rio de Janeiro : PUC, Departamento de Engenharia Elétrica, 2004.

82 f. : il. ; 30 cm

Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica.

Inclui referências bibliográficas.

1. Engenharia elétrica – Teses. 2. Distribuição poisson bivariada. 3. Filtro de Kalman. 4. Simulação de Monte Carlo. 5. Amostragem por importância. 6. Variáveis antitéticas. I. Fernandes, Cristiano Augusto Coelho. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

Agradecimentos

Meu agradecimento a Deus pelas oportunidades que me tem concedido.

À minha esposa Paulina e a meus filhos, Daniela e Felipe por seu infinito amor, apoio, força, compreensão nos momentos de ausência, o estímulo constante, suas existências inspiradoras e pelo prazer reconfortante de nossa convivência.

Ao meu orientador, Professor Cristiano Fernandes pelo incentivo, perseverança, competência indiscutível, disponibilidade e confiança depositada.

Aos meus pais que me mostraram o caminho a seguir na vida, e porque tudo que eu faça ou seja é fruto primordialmente do exemplo de dedicação, trabalho e perseverança de suas vidas e de seu amor incondicional.

Ao professor Reinaldo Castro Souza pelo seu apoio e orientação nos meus primeiros anos no Brasil.

Aos meus amigos André Gustavo, Francisco Carlos, Tomás e Zanini pelas incomensuráveis ajudas prestadas e companhias em muitos momentos em minha vida no Brasil, e sobretudo por aguentar meu mau-humor e “Portuñol”.

Aos meus amigos e colegas: Alberto, Ana Paula, Cetale, Dan, Edison, Evandro, Felipe, Isabela, Juan, Joares, Leonardo Xavier, Lúcio, Macrini, Mayte, Rui, Savano e Yván, pelo apoio e companheirismo.

Ao Adrian pela infinita ajuda e paciência nas inúmeras discussões do tema, sem as quais o presente trabalho não tinha chegado ao seu fim.

A todos os funcionários do Departamento de Engenharia Elétrica pela gentileza e apoio no “dia a dia” da PUC, em especial a Alcina, Márcia, Ana, Danilo, Evandro, Esnarde, Luiz Fernando e Manuel.

Ao Roberto, Francisco e Katia por guiar nossos primeiros passos no Brasil.

À comunidade Peruana da PUC-Rio, por me terem aceito como um deles e permitido-me que aprendesse um pouco da sua rica cultura e tradições.

Aos meus amigos, companheiros de “pelada” e do tradicional Rainha do Mar: Joel, Presvítero, David, Juan José, Hugo, Frank, Daniel e tantos outros que não é possível mencionar, por me permitirem que continuasse com meus velhos vícios.

Ao Brasil pela sua gente e belezas naturais.

Aos meus colegas do Departamento de Matemática de la Universidad del Bío-Bío pelo apoio e compreensão.

Finalmente minha gratidão ao CNPq, OEA, Universidad del Bío-Bío e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Caminante no hay camino
se hace camino al andar.
Al andar se hace camino,
y al volver la vista atrás
Se ve la senda que nunca
se ha de volver a pisar.
Caminante no hay camino,
sino estelas en la mar.

Antonio Machado

Resumo

Contreras Espinoza, Sergio Eduardo; Fernandes, Cristiano (Orientador). **Modelo em Espaço de Estado para Séries Temporais com Distribuição Poisson Bivariada - Uma aplicação da Metodologia Durbin-Koopman** Rio de Janeiro, 2004. 82p. Tese de Doutorado — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Nesta tese, consideramos um modelo de espaço de estado bivariado para dados de contagem. A abordagem usada para resolver integrais não-analíticas que se apresentam no modelo é uma natural extensão da metodologia proposta por Durbin & Koopman-(DK), no sentido de que o Modelo Gaussiano Aproximador deve possuir algumas matrizes de covariâncias diagonais. Esta modificação traz a vantagem de viabilizar o uso do tratamento univariado para séries multivariadas com as recursões de Kalman, o qual, como se sabe, é mais eficiente do que o tratamento usual e facilita o uso de inicializações exatas destas mesmas recursões. O vetor de estado do modelo proposto é definido usando-se abordagem estrutural, onde os elementos do vetor de estado têm interpretação direta como tendência e sazonalidade. Apresentamos exemplos simulados e reais para ilustrar o modelo.

Palavras-chaves

Amostragem por Importância, Distribuição Poisson Bivariada, Filtro de Kalman, Simulação de Monte Carlo, Variáveis Antitéticas, Verossimilhança.

Abstract

Contreras Espinoza, Sergio Eduardo; Fernandes, Cristiano (Adviser). **State Space Model for Time Series with Bivariate Poisson Distribution—An application of Durbin-Koopman methodology**. Rio de Janeiro, 2004. 82p. PhD. Thesis — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

In this thesis we consider a state space model for bivariate observations of count data. The approach used to solve the non analytical integrals that appears as the solution of the resulting non-Gaussian filter is a natural extension of the methodology advocated by Durbin and Koopman (DK).

In our approach the approximated Gaussian Model (AGM), has a diagonal Covariance matrix, while in the original DK, this is a full matrix. This modification make it possible to use univariate Kalman recourses to construct the AGM, resulting in a computationally more efficient solution for the estimation of a Bivariate Poisson model. This also facilitates the use of exact initialization of those recursions.

The state vector is specified using the structural approach, where the state elements are components which have direct interpretation, such as trend and seasonals. In our bivariate set up the dependence between the bivariate vector of time series is accomplished by use of common components which drive both series. We present both simulation and real life examples illustrating the use of our model.

Keywords

Importance Sampling, Poisson Bivariate Distribution, Kalman Filter, Monte Carlo Simulation, Antithetic Variable, Likelihood.

Conteúdo

1	Introdução	11
2	Modelo de Espaço de Estado Linear Gaussiano	15
2.1	Introdução	15
2.1.1	O Filtro de Kalman	19
2.2	Modelos Estruturais para Séries Temporais	22
2.3	Regressão com Coeficientes Variantes no Tempo	32
2.4	Variáveis Explicativas	32
3	Modelo Não-Gaussiano: Abordagem Modificada de Durbin–Koopman	34
3.1	Introdução	34
3.2	O Modelo Não-Gaussiano	34
3.3	Estimação dos Hiperparâmetros	36
3.3.1	Amostragem por Importância	36
3.3.2	Cálculo da Verossimilhança	37
3.3.3	Seleção da Densidade de Importância	39
3.3.4	Verossimilhança por Monte Carlo	41
3.3.5	Variáveis Antitéticas	42
3.3.6	Estimação dos Hiperparâmetros Passo-a-Passo	43
3.4	Alisamento	46
4	Modelo Poisson bivariado	48
4.1	Introdução	48
4.2	Modelo Poisson Bivariado	48
4.3	Modelo de Espaço Estado Poisson Bivariado	50
4.4	Verossimilhança por Monte Carlo	54
4.5	Seleção do Modelo Gaussiano Aproximador	54
4.6	Estimação dos Hiperparâmetros e do Vetor de Estado	57
4.7	Simulação	59
5	Aplicação	63
6	Conclusões e Possíveis Extensões	68
	Referências Bibliográficas	69
A	Apêndice	73
A.1	Modelo Poisson Bivariado proposto por Fernandes	73
A.2	Programa utilizado na tese	75

Lista de Figuras

4.1	Simulação do modelo de nível local Poisson bivariado para $\lambda = 1$, $\sigma_{\eta}^2 = 0.1$ e $b = 0.9$.	52
4.2	Simulação do Modelo Nível e Sazonalidade Locais para $\lambda = 1$, $\sigma_{\eta}^2 = 0.04$, $\sigma_{\omega}^2 = 0.02$, $s = 4$ $a = 0.8$ e $b = .9$.	53
4.3	Estado estimado para o modelo de nível local Poisson bivariado para $\lambda = 1$, $\sigma_{\eta}^2 = 0.1$, $b = 0.9$ e $N = 50$.	61
4.4	Estado Estimado modelo de nível local Poisson bivariado para $\lambda = 1$, $\sigma_{\eta}^2 = 0.05$, $b = 0.9$ e $N = 70$.	62
5.1	y_1 : Número de crianças falecidas com menos de 27 dias de vida.	63
5.2	y_2 : Número de crianças falecidas com idade entre 28 dias e 11 meses e 29 dias.	64
5.3	Número de crianças falecidas com menos de 11 meses e 29 dias.	64
5.4	Média condicional para y_1 usando o modelo bivariado.	66
5.5	Média condicional para y_2 usando o modelo bivariado.	66
5.6	Média condicional para y_1 usando o modelo univariado.	67
5.7	Média condicional para y_2 usando o modelo univariado.	67

Lista de Tabelas

2.1	Dimensões do modelo de espaço de estado.	16
4.1	Resultados da Simulação para $\sigma_{\eta}^2 = 0.1$, $b = 0.9$ e $\lambda = 1$.	60
5.1	Resultado da estimação para σ_{η} para mortalidade infantil.	65