1
Introdução

A abordagem em espaço de estado proporciona uma metodologia unificada para estudar uma ampla variedade de problemas em séries temporais. Por exemplo, por meio desta abordagem é possível modelar o comportamento das diferentes componentes de uma série (tendência, sazonalidade e ciclo) separadamente, e combinar estes sub-modelos em um modelo único para a série de interesse. Neste contexto, os modelos em espaço de estado são denominados de *Modelos Estruturais*. (Ver por exemplo, Harvey [17] e Durbin & Koopman [7] (DK)).

Os modelos em espaço de estado são formados por dois tipos de variáveis: as variáveis não-observáveis do estado, que determinam o movimento do sistema no tempo, e as observações da série.

Os modelos em espaço de estado multivariados lineares gaussianos encontram-se amplamente discutidos na literatura (ver por exemplo, Durbin & Koopman [7] e Anderson & Moore [1]). Todavia, existem situações de interesse prático em que estes modelos não conseguem uma representação apropriada do comportamento dos dados. Por exemplo, se em uma determinada região têm-se as séries do número diário de pessoas mortas em acidentes de trânsito e do número diário de pessoas feridas em acidentes de trânsito, e estes números são pequenos, a modelagem adequada sugere uma distribuição Poisson bivariada, a qual se espera proporcione um modelo mais apropriado para os dados em relação a um modelo baseado na distribuição normal bivariada.

Em algumas situações de interesse, é possível que um modelo linear gaussiano não consiga representar o comportamento dos dados de forma satisfatória, e assim deve-se considerar um modelo não-gaussiano.

A implementação de modelos em espaço de estado em ambientes não–gaussianos requer soluções de integrais decorrentes do cálculo das densidades dos modelos, as quais necessitam de métodos numéricos e/ou aproximações analíticas. Somente um pequeno número de modelos não–gaussiano simples têm recursões de filtragem conjugadas, por exemplo os apresentados em

Harvey & Fernandes [18].

West *et al* [34] introduziram os modelos em espaço de estado para a família exponencial por meio de uma abordagem Bayesiana, que foi extendida por West & Harrison [35]. A família exponencial também foi considerada por Fahrmeir [11].

Tanizaki [33] estudou e desenvolveu alguns filtros não-lineares e/ou não-gaussianos.

Técnicas de simulação baseadas em *Markov Chain Monte Carlo* (MCMC) para modelos em espaço de estado não-gaussianos foram desenvolvidas por Carlin et al. [2], Carter & Kohn [3], Shephard [31], Shephard & Pitt [32], Cargnoni [4] e Gamerman [14] [15].

Um número considerável de algoritmos baseados no filtro de partículas (*particle filtering*), conhecido também como Método de Monte Carlo Sequencial (*Sequential Monte Carlo*) têm sido considerado recentemente na literatura. Ver por exemplo Pitt & Shephard [28] e Doucet *et al* [6].

Durbin & Koopman [8] consideraram um caso especial de modelo onde a distribuição das observações, condicional ao estado, é não-gaussianas e o estado é gaussiano. O objetivo é calcular estimadores por máxima verossimilhança dos hiperparâmetros deste modelo por simulação.

Durbin & Koopman [9] estudaram estes modelos sob a perspectiva clássica e Bayesiana. O tratamento foi baseado em simulação de Monte Carlo. Em seu livro, Durbin & Koopman [7] descrevem uma metodologia baseada em técnicas de simulação para estudar modelos em espaço de estado não-lineares/não-gaussianos, esta metodologia é fundamental no presente trabalho.

A diferença fundamental entre a formulação linear gaussiana e a formulação geral para os modelos não-lineares/não-gaussianos está nas possíveis generalizações das equações das observações e do estado. Para a equação do estado no caso linear pode-se considerar, por exemplo, uma densidade com caudas pesadas como a distribuição t. No caso da distribuição condicional das observações dado o vetor de estado, também no caso linear, podem ser consideradas distribuições diferentes da gaussiana, por exemplo, distribuições da família exponencial. Assim é possível modelar dados de contagem através de distribuições do tipo Poisson, Binomial e Multinomial.

A abordagem de Durbin & Koopman ([7], [8], [9]) é uma das metodologias possíveis para estimar modelos não-lineares/não-gaussianos, e está baseada em simulação de Monte Carlo utilizando amostragem por importância (*importance sampling*), técnica descrita em Ripley [29], variáveis antitéticas (ver Ripley, [29]) e simulação suavizada (*simulation smoothing*) que é um algoritmo para obter amostras do vetor de estado e das perturbações condicional a todas as observações disponíveis. O algoritmo para simulação suavizada pode ser estudado em De Jong & Shephard [21] ou em Durbin & Koopman [10].

Os métodos considerados por Durbin & Koopman obtém estimadores computacionalmente eficientes de médias e variâncias de funções arbitrárias do vetor de estado condicionais as observações. Estas técnicas consideram geralmente o problema em que as observações são univariadas e o vetor de estado multivariado. Entretanto, a literatura é bastante escassa para o caso em que as observações são multivariadas e provenientes de distribuições não-gaussianas. Neste contexto, Fernandes [12] considera um modelo Poisson bivariado.

Este trabalho pode ser visto como uma extensão da metodologia de Durbin & Koopman, no sentido de que algumas matrizes de covariâncias serão diagonais, visando à uma maior eficiência computacional e à possibilidade de que se efetuem iniciliazações exatas do processo de estimação do estado. Da mesma forma que na metologia de DK original, será possível considerar observações provenientes da distribuição Poisson bivariada.

Assim sendo, as contribuições deste trabalho são as seguintes:

1. Especificação de um modelo em espaço de estado com distribuição Poisson bivariada;

2. Uma modificação da metologia de DK que se caracteriza por ser mais eficiente.

3. Adaptação desta nova metodologia desenvolvida para estimação do modelo proposto em (1), estabelecendo expressões para derivadas analíticas necessárias para as implemetações computacionais e discussões sobre a identificação da função de verossimilhança associada;

4. Estudo por simulação para investigar a adequabilidade do modelo e da metodologia de estimação proposta;

5. Aplicação do modelo proposto em dados reais.

O capítulo 2 apresenta um breve resumo do modelo de espaço de estado linear gaussiano com a intenção de apresentar muito sucintamente o modelo. Este capítulo é muito sucinto em relação a tudo que poderia ser escrito a respeito, a principal razão é porque existem muitos livros que consideram este tópico, entre eles cabe destacar o livro de Harvey [17].

No capítulo 3, discutem-se os pontos principais da metologia modificada de Durbin & Koopman no tratamento dos modelos estruturais não-gaussianos; descrevem-se, também, as idéias gerais das técnicas de simulações que a metodologia usa, a saber: técnicas de amostragem por importância e variáveis antitéticas. O tratamento completo deste tema pode ser encontrado em Durbin & Koopman [7], [9].

No capítulo 4, apresenta-se o modelo em espaço de estado com a distribuição Poisson bivariada proposto nesta tese. Apresenta-se um método de estimação dos hiperparâmetros e das funções do vetor de estado baseado na metodologia discutida no capítulo 3, além de um estudo por simulação para verificação da eficiência do modelo na estimação dos hiperparâmetros.

No capítulo 5 descreve-se uma aplicação em dados reais, e no capítulo 6 são feitas as conclusões finais e a descrição de possíveis extensões de nosso trabalho.