

**Processos de Cox com  
intensidade difusiva afim**

Alan De Genaro Dario

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

Programa: Pós-Graduação em Estatística

Orientador: Prof. Dr. Adilson Simonis

São Paulo, 24 de agosto de 2011

# Processos de Cox com intensidade difusiva afim

Esta é a versão depositada  
por Alan De Genaro Dario.

Banca Examinadora:

- Prof. Dr. Adilson Simonis (orientador) - IME-USP.
- Prof. Dr. Marco Avellaneda - Courant Institute of Mathematical Sciences - NYU.
- Prof. Dr. Jorge Zubelli - IMPA.
- Prof. Dr. Cristiano Fernandes - PUC-Rio.
- Prof. Dr. Julio Stern - IME-USP.

## Agradecimentos

Ao iniciar os agradecimento é certo que muitos foram aqueles que contribuíram direta ou indiretamente na concretização desse sonho. Em primeiro lugar eu gostaria de agradecer o Prof. Adilson por ter acreditado em mim mesmo antes do meu ingresso no Programa de Pós-graduação em Estatística do IME. O Prof. Adilson não hesitou em aceitar-me como aluno, mesmo sabendo que eu teria que conciliar a vida acadêmica com a profissional. O Prof. Adilson não foi apenas meu orientador, mas sim o condutor dentro do IME, sugerindo disciplinas que permitiram preencher algumas lacunas existentes na minha formação. Agradeço aos Profs. Júlio Stern e Walter Mascarenhas, membros da minha banca de qualificação, pelas sugestões que ajudaram a aperfeiçoar o trabalho. Também seria impossível escrever esta tese sem o apoio que a BM&FBOVESPA me proporcionou, em especial na pessoa do Luis Vicente, meu gestor imediato e amigo, que me liberou durante o período em que tive que cumprir os créditos-aula. Agradeço também aos Profs. Jorge Zubelli e Cristiano Fernandez pela oportunidade de apresentar versões preliminares desta tese nos seminários por eles organizados. Não poderia deixar de agradecer o *staff* do Nonno Ruggero pela competência em nos servir durante nossos almoços de trabalho. Gostaria de agradecer meus pais, meus sogros, que me apoiaram sempre que precisei. Por último, o agradecimento mais importante é para minha mulher, Deborah, que entendeu e ajudou nos momentos mais difíceis durante o curso, mas especialmente pelo fato dela ter sonhando um sonho meu.



## Resumo

Esta Tese explora o Processo de Cox quando sua intensidade pertence a uma família de difusões afim. A forma da função densidade de Probabilidade do Processo de Cox é obtida quando a intensidade é descrita por uma difusão afim  $d$ -dimensional arbitrária. Analisa-se também o acoplamento e convergência para o Processo de Cox com intensidade afim. Para ilustrar assume-se que a intensidade do Processo é governada por uma difusão de Feller e resultados mais detalhados são obtidos. Adicionalmente, os parâmetros da intensidade do Processo são estimados por meio do Filtro de Kalman conjugado com o estimador de Quase-Máxima Verossimilhança.

**Palavras-chave:** Processo de Cox, Difusão Afim, Filtro de Kalman, Livro de Ofertas.



# Abstract

This Thesis deals with the Cox Process when its intensity belongs to a family of affine diffusions. The form of the probability density function of the Cox process is obtained when the density is described by an arbitrary  $d$ -dimensional affine diffusion. Coupling and convergence results are also addressed for a general Cox Process with affine intensity. We adopted the Feller diffusion for driving the underlying intensity of the Cox Process to illustrate our results. Additionally the parameters of the underlying intensity processes are estimated by means of the Kalman Filter in conjunction with Quasi-Maximum Likelihood estimation.

**Keywords:** Cox Process, affine diffusion, Kalman Filter, Order Book.





# Sumário

<b>Lista de Abreviaturas</b>	<b>ix</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Resultados Anteriores . . . . .	1
1.2 Aplicações do Processo de Poisson Duplamente Estocástico em finanças . . . . .	2
1.3 Objetivos da Tese . . . . .	3
1.4 Estrutura da tese . . . . .	4
<b>2 Processo de Poisson Duplamente Estocástico - Modelo Geral</b>	<b>5</b>
2.1 Introdução . . . . .	5
2.2 Construção do Processo Duplamente Estocástico . . . . .	5
2.3 Existência, unicidade e admissibilidade de $(X_t)_{t \geq 0}$ . . . . .	8
2.3.1 Existência . . . . .	8
2.3.2 Unicidade . . . . .	10
2.3.3 Admissibilidade . . . . .	13
2.4 Transformada de Laplace para o processo $(\Lambda_t)_{t \geq 0}$ . . . . .	14
2.5 Distribuição do Processo de Cox $(N_t)_{t \geq 0}$ . . . . .	17
2.5.1 Momentos do processo $(N_t)_{t \geq 0}$ . . . . .	19
2.5.2 O processo Compensado $(M_t)_{t \geq 0}$ . . . . .	19
2.6 Estacionariedade do processo $(N_t)_{t \geq 0}$ . . . . .	20
2.6.1 Estacionariedade do processo $(\lambda_t)_{t \geq 0}$ . . . . .	23
2.7 Distribuição Estacionária do processo $(N_t)_{t \geq 0}$ . . . . .	23
<b>3 Processo de Poisson Duplamente Estocástico - Intensidade de Feller</b>	<b>25</b>
3.1 Introdução . . . . .	25
3.2 Construção do Processo Duplamente Estocástico . . . . .	25
3.3 Transformada de Laplace para o processo $(\Lambda_t)_{t \geq 0}$ . . . . .	26
3.4 Distribuição do Processo de Cox $(N_t)_{t \geq 0}$ . . . . .	28
3.4.1 Momentos do processo $(N_t)_{t \geq 0}$ . . . . .	29
3.5 Distribuição estacionária do processo $(N_t)_{t \geq 0}$ . . . . .	30

<b>4</b>	<b>Acoplamento para Processos de Poisson Duplamente Estocásticos</b>	<b>39</b>
4.1	Definições . . . . .	39
4.2	Resultados e Teoremas . . . . .	40
<b>5</b>	<b>Estimação de Processos de Cox com intensidade afim</b>	<b>45</b>
5.1	Introdução . . . . .	45
5.2	Revisão dos modelos de estimação . . . . .	46
5.3	Filtro de Kalman . . . . .	47
5.3.1	Derivação da matriz de covariância <i>a posteriori</i> . . . . .	49
5.3.2	Derivação do ganho de Kalman . . . . .	50
5.3.3	Simplificação da fórmula <i>a posteriori</i> da covariância . . . . .	51
5.4	Filtro de Kalman para processos de Cox com intensidade de Feller . . . . .	52
5.5	Estimação do PPDE para dados de alta-frequência . . . . .	53
5.5.1	Descrição da amostra . . . . .	54
5.5.2	Resultados da Estimação . . . . .	54
5.6	Simulações . . . . .	57
<b>6</b>	<b>Conclusões</b>	<b>61</b>
6.1	Considerações Finais . . . . .	61
6.2	Sugestões para Pesquisas Futuras . . . . .	61
<b>A</b>	<b>Códigos desenvolvidos</b>	<b>63</b>
	<b>Referências Bibliográficas</b>	<b>69</b>

## Lista de Abreviaturas

ACD	Duração Autorregressiva Condicional ( <i>Autoregressive Conditional Duration</i> )
ATS	Estrutura a Termo Afim ( <i>Affine Term Structure- ATS</i> )
EDE	Equação Diferencial Estocástica ( <i>Stochastic Differential Equation</i> )
EDO	Equação Diferencial Ordinária ( <i>Ordinary Differential Equation</i> )
EDP	Equação Diferencial Parcial ( <i>Partial Differential Equation</i> ).
EFP	Equação de Fokker-Plank ( <i>Fokker-Plank Equation</i> ).
$E_i$	Função Integral Exponencial ( <i>Exponential Integral</i> )
FDP	Função Distribuição de Probabilidade ( <i>Probability Distribution Function</i> )
i.i.d.	Independente e Identicamente Distribuído ( <i>Independent and identically-distributed</i> )
MLE	Estimador de Máxima Verossimilhança ( <i>Maximum Likelihood Estimator</i> )
PPDE	Processo de Poisson Duplamente Estocástico ( <i>Doubly Stochastic Poisson Processes</i> )
QME	Estimador de Quase-Máxima Verossimilhança ( <i>Quasi Maximum Likelihood Estimator</i> )



# Capítulo 1

## Introdução

Os Processos de Poisson Duplamente Estocásticos, (**PPDE**), foram introduzidos por Cox (1955) ao permitir que a intensidade do processo de Poisson pudesse ser descrito por uma variável aleatória positiva, e não apenas por um fator determinístico. O objetivo de tal generalização foi permitir que a dinâmica de um processo exógeno ao modelo influenciasse as taxas de transição do processo pontual de interesse.

Em um primeiro momento e, sem recorrer ao formalismo necessário que será desenvolvido no próximo capítulo, um processo pontual  $N_t$  pode ser interpretado como a representação do ato de contar um fenômeno com o passar do tempo.

Os processos pontuais possuem aplicações em diferentes áreas do conhecimento. De maneira muito resumida pode-se apresentar as três principais áreas: bioestatística, finanças e teoria da confiabilidade. Em bioestatística eles são o arcabouço teórico para estudar eventos recorrentes (*recurrent events*), tais como Gail *et al.* (1980), que estudou a quantidade de tumores em ratos ao longo de um período. Em finanças, Lando (1998) foi o pioneiro em utilizar processo pontuais para descrever a ocorrência de eventos de crédito. Na teoria de confiabilidade, o trabalho de Dalal e McIntosh (1994) desenvolveu uma estimativa para determinar o tempo de parada ótimo para testar e validar (*debugging*) um software.

Nas seções a seguir será realizada uma breve revisão dos principais trabalhos sobre esse assunto, bem como a caracterização do problema, os objetivos e a estrutura desta tese.

### 1.1 Resultados Anteriores

O trabalho seminal de Cox (1955) introduziu o Processo de Poisson Duplamente Estocástico, **PPDE**, também conhecido por Processo de Cox.

A principal obra nesta área é o trabalho de Grandell (1976), que apresenta as principais propriedades do **PPDE** a partir da construção clássica em probabilidade. De maneira complementar, os trabalhos de Brémaud (1972) e Daley e Vere-Jones (1988) constroem o processo segundo ideias da Teoria de Martingais. A principal semelhança entre estas obras é o foco em derivar as propriedades

gerais do **PPDE**, porém sem explorar a forma funcional da intensidade do processo e, portanto com reduzido número de aplicações.

O estudo do **PPDE** tomou nova direção quando se passou a especificar a forma funcional para a intensidade do processo, e conseqüentemente foi possível obter expressões analíticas para a função densidade de probabilidade e para os momentos do processo. Dentre eles, pode-se citar Bouzas *et al.* (2002) que utiliza a distribuição Normal truncada para descrever a intensidade do processo e Bouzas *et al.* (2006), que generaliza a intensidade para o caso do oscilador harmônico. A contribuição destes trabalhos foi a obtenção de expressões analíticas em forma fechada para a função densidade do processo de Cox, bem como para os seus momentos. Entretanto, em ambos os casos, os autores utilizam construções nas quais a restrição de não-negatividade para a medida de intensidade é violada. Para contornar tal limitação os autores determinam uma região no espaço paramétrico onde a probabilidade de realizações de valores negativos para a intensidade é reduzida.

Com o intuito de garantir que a restrição de não-negatividade para a medida de intensidade seja preservada, Basu e Dassios (2002) e Kozachenko e Pogorilyak (2008) sugerem a adoção de um modelo Lognormal para a intensidade do processo. Uma formulação que incorpora a intensidade em uma formulação dinâmica é desenvolvida em Dassios e Jang (2008), que utilizaram a forma funcional de um processo do tipo *Shot Noise*, o qual, a despeito da garantia da não-negatividade da medida de intensidade, não se trata de um processo difusivo.

Por sua vez Wei *et al.* (2002), ao assumir que a intensidade é governada por um processo de Feller unidimensional, obteve a Função Densidade de Probabilidade para o Processo Pontual  $N_t$ . O processo de Feller, introduzido por Feller (1951), foi então consagrado na literatura de finanças após o trabalho de Cox *et al.* (1985). Dentre as propriedades, tem-se que o processo está bem definido apenas no  $\mathbb{R}_+$ , garantindo que a condição de não-negatividade seja automaticamente atendida.

Neste trabalho assume-se que a intensidade é governada por um processo difusivo afim<sup>1</sup>, conforme formalizado por Duffie e Kan (1996), *d-dimensional* flexível o suficiente para incorporar o processo de Feller em uma ou  $d$  dimensões. Desta forma, os trabalhos de Bouzas *et al.* (2002), Wei *et al.* (2002), Basu e Dassios (2002) e Kozachenko e Pogorilyak (2008) podem ser visto com casos particulares do modelo ora proposto.

## 1.2 Aplicações do Processo de Poisson Duplamente Estocástico em finanças

A utilização de processos pontuais em finanças, em especial dos processos de Poisson Duplamente Estocásticos, **PPDE**, teve um enorme desenvolvimento no final dos anos 90 com o desenvolvimento dos modelos para gerenciar e apreçar o risco de crédito. Diferente dos trabalhos de

---

<sup>1</sup>Mais formalmente, uma função  $F : R^d \rightarrow R$ , é denominada afim, se existe  $a \in R$  e  $b \in R^d$  tal que  $F(x) = a + b^\top x$  para todo  $x \in R^d$ . Observe que  $b^\top$  denota o transposto do vetor *d-dimensional*.

Grandell (1976), Brémaud (1972) e Daley e Vere-Jones (1988), os trabalhos de Duffie e Singleton (1999) e Duffie *et al.* (2003) formalizam a construção da probabilidade do primeiro salto do processo pontual  $N_t$ , em uma estrutura afim. O tempo até o primeiro salto do processo  $N_t$ , representa, em um contexto de risco de crédito, o tempo até a falência (*Default*) de uma empresa e/ou país. Assim, uma vez atingido o estado absorvente, não se fazia necessário o estudo da dinâmica e das propriedades do **PPDE** posteriores ao primeiro salto.

Recentemente uma nova área de aplicações dos processos pontuais em finanças surgiu com o uso destes modelos para descrever o processo de chegada de ordens de compra e venda de um ativo em uma ambiente de negociação eletrônica. Nestes modelos, o processo de chegada de ordens evolui ao longo do tempo e o intuito é caracterizar a dinâmica do processo e obter expressões tratáveis analiticamente para descrever a probabilidade de uma ordem enviada, segundo uma configuração de mercado, seja executada antes que o preço se altere, conforme descrito em Cont *et al.* (2010).

Assim, como nos modelos onde a probabilidade de uma empresa/país não honrar suas dívidas pode depender de variáveis de estado exógenas ao modelo, o número de ofertas de compra e venda que chegam no livro de ordens pode depender, por exemplo, do nível de aversão ao risco dos investidores em um determinado dia, ou da divulgação de uma informação, por exemplo, a divulgação de uma nova tecnologia que impacta um determinado setor da economia. Por esta razão a construção de um modelo analiticamente tratável que incorpore endogenamente o comportamento estocástico da intensidade de um Processo Pontual é uma contribuição à literatura.

### 1.3 Objetivos da Tese

Diante do exposto e da caracterização do problema, o objetivo principal desta tese é estudar o Processo de Poisson Duplamente Estocástico quando a intensidade pertence a uma família *d-dimensional* de difusões afim. Neste sentido, pretende-se emprestar todo o arcabouço teórico dos modelos afim desenvolvido ao longo de décadas para modelar estruturas a termo (*Affine Term Structure- ATS*) da taxa de juros e aplicá-los nos processos pontuais para estudar sua dinâmica com o passar do tempo. Desta forma, tem-se como a primeira contribuição deste trabalho o Teorema 1, que descreve a forma da Função Densidade de Probabilidade para o Processo de Poisson Duplamente Estocásticos quando a intensidade pertence a uma família *d-dimensional* de difusões afim. Em um caso particular, quando a intensidade é governada por uma difusão de Feller unidimensional, os resultados são comparáveis ao encontrado em Wei *et al.* (2002). A segunda contribuição é o Teorema 2 que prova a velocidade de convergência do processo de Cox construído no Capítulo 3. A terceira contribuição desta tese é o Teorema 3 onde demonstra-se como é possível acoplar dois Processos de Poisson Duplamente Estocásticos quando as condições para tanto são construídas nas equações diferenciais estocásticas que governam a dinâmica das intensidades dos processos.

Além disso, sugere-se aqui, de forma pioneira, a estimação do processo de Cox com intensidade

afim por meio do Filtro de Kalman conjugado com o método de Quase-Máxima Verossimilhança. A razão para esta proposta advém do fato que o Filtro de Kalman é um algoritmo recursivo sendo computacionalmente eficiente em problemas com elevado número de observações, conforme descrito na seção 5.5. Além disso, como se verá, o método proposto se comporta bem em um exercício de Simulações de Monte Carlo.

#### 1.4 Estrutura da tese

Além desta introdução a tese é formada por 5 Capítulos e um Apêndice. No Capítulo 2 o processo de Cox com intensidade afim em sua forma geral é construído e os principais resultados apresentados. No Capítulo 3 adota-se o processo de Feller para descrever a intensidade do Processo de Cox, sendo possível assim obter resultados mais detalhados para o processo pontual. No Capítulo 4 tem-se a construção das condições para o acoplamento dos processos de Cox com intensidade afim. No Capítulo 5 apresenta-se uma metodologia, baseada no Filtro de Kalman, para estimar os parâmetros do processo de Cox com intensidade de Feller e seus resultados são avaliados por meio de uma Simulação de Monte Carlo. O Capítulo 6 apresenta as considerações finais, bem como as sugestões para pesquisas futuras. O Apêndice contém os códigos desenvolvidos em *Matlab*<sup>®</sup> utilizados no Capítulo 5.

Com o intuito de facilitar a leitura, adota-se ao longo desta tese a seguinte nomenclatura: quando for feito uso de Teoremas já conhecidos estes serão denominados **Resultados**, quando for necessário reunir diferentes resultados para apresentar uma nova contribuição estes serão denominados por **Proposição** e apenas as contribuições mais significativas desta tese serão denominadas de **Teoremas**.



## Capítulo 2

# Processo de Poisson Duplamente Estocástico - Modelo Geral

### 2.1 Introdução

O Processo de Poisson Duplamente Estocástico ou Processo de Cox é uma possível generalização do Processo de Poisson no qual a intensidade é definida por um processo estocástico de tal maneira que condicionado a uma particular realização da intensidade  $\lambda_t(\omega)$ , o salto do processo pontual torna-se um processo de Poisson não-homogêneo com intensidade  $\lambda_t(\omega)$ . Na sequência<sup>2</sup> as bases para analisar as propriedades do processo de Cox são construídas.

### 2.2 Construção do Processo Duplamente Estocástico

Considere o espaço de probabilidade filtrado  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ , onde  $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$  é uma filtração que contém os conjunto de medida  $\mathbb{P}$ -nula contínua à direita.

**Definição 1** *Define-se,  $(X_t)_{t \geq 0}$ , com  $X_t : \mathcal{D} \times \mathbb{R}_+ \rightarrow \mathbb{R}^d$ , a **variável de estado**, como a solução da seguinte Equação Diferencial Estocástica (EDE):*

$$dX_t = \mu(X, t)dt + \sigma(X, t)dW_t \quad (2.1)$$

sendo:  $\mu(X, t) : \mathcal{D} \times \mathbb{R}_+ \rightarrow \mathbb{R}^d$ , o coeficiente de *drift*;  $\sigma(X, t) : \mathcal{D} \times \mathbb{R}_+ \rightarrow \mathbb{R}^{d \times d}$ , o coeficiente de difusão do processo e,  $W_t$  o Movimento Browniano Padrão *d-dimensional* em  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ . Onde  $\mathcal{D} \subset \mathbb{R}^d$  será definido na seção 2.3.3.

**Definição 2 (Duffie e Kan (1996))** *Define-se  $(X_t)_{t \geq 0}$  como um **processo afim**, se (2.1) atender simultaneamente:*

1. Coeficiente de *drift*

$$\mu(X, t) = \mathcal{K}(\Theta - X_t) \quad (2.2)$$

Para  $\Theta \in \mathbb{R}^d$  e  $\mathcal{K} \in \mathbb{R}^{d \times d}$

---

<sup>2</sup>Para não utilizar notações em excesso, toma-se a seguinte simplificação:  $\lambda_t := \lambda_t(\omega)$

2. Matriz de Covariância

$$\sigma(X, t) = \Sigma\sqrt{\sigma} \quad (2.3)$$

Onde  $\Sigma$  é uma matriz não-singular em  $\mathbb{R}^{d \times d}$  e  $\sigma$  é a matriz diagonal em  $\mathbb{R}^{d \times d}$ , tendo o  $i$ -ésimo elemento em sua diagonal dado por:

$$\sigma_i = a_i + b_i X_t, \quad i = 1, \dots, d, \quad (2.4)$$

com  $a \in \mathbb{R}$  e  $b \in \mathbb{R}^d$ .

Desta forma, a variável de estado em sua forma afim, segundo Duffie e Kan (1996), será escrita como:

$$dX_t = \mathcal{K}(\Theta - X_t)dt + \Sigma\sqrt{\sigma}dW_t \quad (2.5)$$

onde  $W_t$  é Movimento Browniano Padrão  $d$ -dimensional em  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ .

Com esta parametrização é possível representar qualquer difusão afim ao se definir os parâmetros  $\{\mathcal{K}, \Theta, a, b\}$  apropriadamente. Para ilustrar este argumento, tomam-se algumas EDE usualmente utilizadas em finanças e as representamos segundo a forma descrita em (2.5):

Para o caso univariado ( $d = 1$ ), têm-se:

- Ornstein-Ulhenbeck (Vasicek)

$$dX_t = \kappa(\theta - X_t)dt + \sigma dW_t$$

- Feller (Raiz quadrada, Cox-Ingersoll-Ross)

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t \quad (2.6)$$

- Movimento Geometrico Browniano

$$X_t := \ln S_t \quad dX_t = \left( \mu - \frac{1}{2}\sigma^2 \right) dt + \sigma dW_t$$

Analogamente, têm-se as seguintes EDE's afins no caso multivariado:

- Heston

$$dX_t = \left( \begin{bmatrix} \mu \\ \kappa\theta \end{bmatrix} + \begin{bmatrix} 0 & -\frac{1}{2} \\ 0 & -\kappa \end{bmatrix} X_t \right) dt + \begin{bmatrix} 1 & 0 \\ \rho\sigma & \sqrt{1-\rho^2}\sigma \end{bmatrix} \left( \sqrt{\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} X_t} \right) dW_t$$

- CIR Multivariado

$$dX_t = \left( \begin{bmatrix} \kappa_1 \theta_1 \\ \kappa_2 \theta_2 \end{bmatrix} + \begin{bmatrix} -\kappa_1 & 0 \\ 0 & -\kappa_2 \end{bmatrix} X_t \right) dt + \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \left( \sqrt{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} X_t \right) dW_t$$

As quais descrevem um processo bi-dimensional afim  $X = (V, Y)$  com espaço de estado  $\mathcal{D} = \mathbb{R}_+ \times \mathbb{R}_+$ .

**Definição 3** *Define-se como **tempo de parada** a variável aleatória estendida não-negativa  $\tau : \Omega \rightarrow \mathbb{R}^+ \cup \{\infty\}$ , definida em  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ , tal que  $\{\tau \leq t\} \in \mathcal{G}_t \quad \forall t \geq 0$ .*

**Definição 4** *Define-se a **intensidade** de um processo como:*

$$\lambda_t = \rho_0 + \rho_1 \cdot X_t \quad (2.7)$$

$$= \rho_0 + \sum_{i=1}^d \rho_{i,1} X_{i,t} \quad (2.8)$$

Para  $\rho_0 \in \mathbb{R}$  e  $\rho_1 \in \mathbb{R}^d$ . Com  $X_t$  descrito na Definição 2.

**Definição 5** *Dado o espaço  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$  e sendo  $\mathbb{F}$  uma sub-filtração de  $\mathbb{G}$ , define-se a **intensidade acumulada**  $(\Lambda_t)_{t \geq 0}$ , como um processo crescente  $\mathcal{F}$ -mensurável, absolutamente contínuo, tal que  $\mathbb{P}$ -quase certamente:*

$$\Lambda_0 = 0 \quad , \quad \Lambda_\infty = \infty \quad \text{e}$$

$$\Lambda_{0,t} = \int_0^t \lambda_u du \quad \text{com} \quad t \geq 0$$

Onde  $\lambda_u$  é definida de acordo com a equação (2.7).

**Definição 6** *Dada uma sequência não-decrescente de tempos de parada  $\{\tau_i, i \in \mathbb{N}\}$ , define-se o **Processo Pontual**  $(N_t)_{t \geq 0}$ , como um processo contínuo à direita (càdlàg) construído por:*

$$N_t := \sum_i \mathbb{1}_{\{\tau_i \leq t\}}$$

A construção da variável  $\tau$  como sendo um tempo de parada requer a correta escolha da filtração. Desta forma, neste momento se expandirá  $\mathbb{F}$  para uma filtração maior  $\mathbb{G}$ , onde na seção 2.5.2, justifica-se sua necessidade com a observação 1. Dentre as inúmeras possibilidades de expandir  $\mathbb{F}$ , a estratégia aqui adotada será expandir  $\mathbb{F}$  o suficiente para que  $\tau$  seja um tempo de parada. Desta forma, a nova filtração  $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ , pode ser construída como:

$$\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$$

onde<sup>3</sup>

$$\mathcal{H}_t = \sigma\{N_s : 0 \leq s \leq t\}$$

**Definição 7** Considere  $(N_t)_{t \geq 0}$  um processo estocástico construído em  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ , e toma-se  $\mathbb{F}$  uma sub-filtração de  $\mathbb{G}$ , então define-se  $N_t$  como um **Processo de Poisson Duplamente Estocástico** com relação à  $\Lambda_t$ , se  $\lambda_t$  é  $\mathbb{F}$ -mensurável e, se para qualquer  $0 \leq t \leq T$  e todo  $k = 0, 1, \dots$ , for satisfeito que:

$$P(N_T - N_t = k | \mathcal{G}_t \vee \mathcal{F}_\infty) = \frac{1}{k!} (\Lambda_{t,T})^k \exp(-\Lambda_{t,T}) \quad (2.9)$$

onde  $\mathcal{F}_\infty = \sigma(\mathcal{F}_u : u \in \mathbb{R}_+)$

Em particular, tem-se que:

$$P(N_T - N_t = k | \mathcal{G}_t \vee \mathcal{F}_\infty) = P(N_T - N_t = k | \mathcal{F}_\infty) \quad (2.10)$$

Ou seja, condicionado à  $\sigma$ -álgebra  $\mathcal{F}_\infty$  o incremento  $N_T - N_t$  é independente da  $\sigma$ -álgebra  $\mathcal{G}_t$ .

Desta forma, a probabilidade de nenhuma ocorrência no intervalo  $[t, T]$  para o processo de Poisson não-homogêneo  $(N_t)_{t \geq 0}$  com função intensidade  $(\lambda_t)_{t \geq 0}$  é dada por:

$$P(N_T = 0 | \mathcal{F}_\infty) = \exp\left(-\int_t^T \lambda_u du\right) \quad (2.11)$$

Nas seções seguintes a equação (2.11) será fundamental para a obtenção da função densidade de probabilidade dos Processos de Poisson Duplamente Estocásticos com intensidade afim.

## 2.3 Existência, unicidade e admissibilidade de $(X_t)_{t \geq 0}$

### 2.3.1 Existência

Para definir a existência da solução forte<sup>4</sup> da equação diferencial estocástica (2.5) se faz necessário que o seguinte resultado seja satisfeito:

**Proposição 1 (Ikeda e Watanabe (1981))** Ao se tomar as seguintes definições de norma para vetores  $a \in \mathbb{R}^d$  e matrizes  $(d \times d)$   $B = [b_{ij}]$ :

$$\|a\|^2 = \sum_{i=1}^d a_i^2 \quad e \quad \|b\|^2 = \sum_{j=1}^d \sum_{i=1}^d b_{ij}^2$$

tem-se que a condição de crescimento linear deve satisfazer:

<sup>3</sup>  $A \vee B :=$  máximo entre A e B.

<sup>4</sup>  $X_t$  é definido como a **solução forte** da EDE (2.1) se para todo  $t > 0$ ,  $(X_t)_{t \geq 0}$  é uma função  $F(t, (W_s, s \leq t))$  e as integrais estocásticas  $\int_0^t \mu(X_s) ds$  e  $\int_0^t \sigma(X_s) dW_s$  existem para a equação diferencial estocástica (2.1) e a solução integral  $X_t = X_0 + \int_0^t \mu(X_s) ds + \int_0^t \sigma(X_s) dW_s$  é satisfeita.

$$\forall X \in \mathbb{R}^d \exists c \in \mathbb{R} : \|\mathcal{K}(\Theta - X)\|^2 + \|\Sigma\sqrt{\sigma}\|^2 \leq c(1 + \|X\|) \quad (2.12)$$

### Prova da Proposição 1

Para o *drift* a demonstração se dá trivialmente graças a estrutura afim:

$$\begin{aligned} \|\mathcal{K}(\Theta - X)\|^2 &\leq (\|\mathcal{K}\Theta\| + \|\mathcal{K}X\|)^2 \leq \\ &\leq (\|\mathcal{K}\Theta\| + \|\mathcal{K}\| \|X\|)^2 = \|\mathcal{K}\Theta\|^2 + \|\mathcal{K}\|^2 \|X\|^2 + 2\|\mathcal{K}\| \|\Theta\| \|X\| \end{aligned}$$

No caso do coeficiente de difusão, tem-se, devido as propriedades da norma:

$$\begin{aligned} \|\Sigma\sqrt{\sigma}\|^2 &\leq \|\Sigma\|^2 \|\sqrt{\sigma}\|^2 = \|\Sigma\|^2 \sum_{i=1}^d (a_i - b_i X)^2 = \\ \|\Sigma\|^2 \sum_{i=1}^d \left( a_i^2 + (b_i X)^2 - 2a_i b_i X \right) &\leq \|\Sigma\|^2 \sum_{i=1}^d \left( a_i^2 + \|b_i\|^2 \|X\|^2 - 2a_i b_i X \right) \leq \\ &\leq \|\Sigma\|^2 \sum_{i=1}^d \left( a_i^2 + \|b_i\|^2 \|X\|^2 + 2\|a_i\| \|b_i\| \|X\| \right) \end{aligned}$$

Então, definindo

$$\bar{a} = \max_j |a_j| \quad \text{e} \quad \bar{b} = \max_j |b_j| \quad ,$$

a condição (2.12) torna-se

$$\begin{aligned} \|\mathcal{K}(\Theta - X)\|^2 + \|\Sigma\sqrt{\sigma}\|^2 &\leq \|\mathcal{K}\Theta\|^2 + \|\mathcal{K}\|^2 \|X\|^2 + 2\|\mathcal{K}\| \|\Theta\| \|X\| + \\ &+ \|\Sigma\|^2 \sum_{i=1}^d \left( a_i^2 + \|b_i\|^2 \|X\|^2 + 2\|a_i\| \|b_i\| \|X\| \right) \leq \\ \|\mathcal{K}\Theta\|^2 + \|\mathcal{K}\|^2 \|X\|^2 + 2\|\mathcal{K}\| \|\Theta\| \|X\| &+ \|\Sigma\|^2 d \left( \bar{a}^2 + \bar{b}^2 \|X\|^2 + 2\bar{a}\bar{b} \|X\| \right) = \\ = \|X\|^2 \left( \|\mathcal{K}\| + \|\Sigma\| d \bar{b}^2 \right) + \|X\| \left( 2\|\mathcal{K}\| \|\Theta\| + \|\Sigma\|^2 2d\bar{a}\bar{b} \right) &+ \\ + \|\mathcal{K}\Theta\| + \|\Sigma\|^2 d \bar{a}^2 &\leq c \left( 1 + \|X\| \right) \end{aligned}$$

onde  $c$  é uma constante definida apropriadamente na sequência. Assim, reagrupando os termos, se obtém:

$$\begin{aligned} \|X\|^2 \left( \|\mathcal{K}\| + \|\Sigma\| d \bar{b}^2 - c \right) + \|X\| \left( 2\|\mathcal{K}\| \|\Theta\| + \|\Sigma\|^2 2d\bar{a}\bar{b} \right) &+ \\ + \left( \|\mathcal{K}\Theta\| + \|\Sigma\|^2 d \bar{a}^2 - c \right) &\leq \end{aligned} \quad (2.13)$$

Temos uma equação polinomial de grau 2 em  $\|X\|$  negativa para  $\forall \|X\|$  desde que as seguintes condições se verifiquem:

$$\begin{cases} \left( \|\mathcal{K}\| + \|\Sigma\|d\bar{b}^2 - c \right) < 0 \\ \left( \|\mathcal{K}\Theta\| + \|\Sigma\|^2d\bar{a}^2 - c \right) < 0 \\ \left( 2\|\mathcal{K}\|^2\|\Theta\| + \|\Sigma\|^22d\bar{a}\bar{b} \right) < 4\left( \|\mathcal{K}\Theta\| + \|\Sigma\|^2d\bar{a}^2 - c \right)\left( \|\mathcal{K}\| + \|\Sigma\|d\bar{b}^2 - c \right) \end{cases} \quad (2.14)$$

É possível encontrar um  $c$  grande o suficiente tal que o sistema (2.14) seja satisfeito e portanto existe uma escolha apropriada de  $c$  tal que (2.13) também seja satisfeito. Finalmente, ter-se-á que (2.12) será satisfeita.

Soma-se ao resultado acima o fato de  $\mu(X, t)$   $\sigma(X, t)$  serem funções da classe  $C^0$  e, portanto garante-se a existência da equação diferencial estocástica (2.5).

■

### 2.3.2 Unicidade

A prova usual para unicidade de uma solução forte, conforme Karatzas e Shreve (1991), impõem que os coeficientes  $\mu(X, t)$  e  $\sigma(X, t)$  satisfaçam a condição de serem globalmente Lipschitzianos. Entretanto para alguns dos processos afim que serão abordados nesta tese a condição de globalmente Lipschitz não se verifica. Por exemplo, o processo de Feller unidimensional (2.6) só existe em  $\mathbb{R}_+$  e o coeficiente de difusão não é globalmente Lipschitziano, pois a derivada de  $\sigma(X_t) := \sigma\sqrt{X_t}$  não é limitada na vizinhança de zero. No caso de difusões unidimensionais pode-se garantir a existência e unicidade da solução a partir de condições menos restritivas sobre os coeficientes da difusão  $\mu(X, t)$  e  $\sigma(X, t)$ , conforme o resultado a seguir:

**Resultado 1 (Yamada-Watanabe I)** *Suponha que  $\mu(X, t)$  seja Lipschitz e que  $\sigma(X, t)$  satisfaça a condição de Hölder de ordem  $\alpha$ ,  $\alpha \geq 1/2$ , ou seja, existe uma constante  $K$  tal que:*

$$|\sigma(x) - \sigma(y)| < K |x - y|^\alpha$$

*Então existe uma única solução forte para a difusão.*

Para o processo de Feller unidimensional, tem-se que o coeficiente de difusão é  $\sigma(X, t) := \sigma\sqrt{X_t}$  e ao tomarmos  $\alpha = 1/2$ , percebe-se que a condição é suficiente para a unicidade de uma solução forte da equação diferencial estocástica (2.5), requeridas no Resultado 1.

Entretanto, no caso geral a unicidade de uma solução para uma EDE multidimensional é obtida por meio do seguinte Resultado:

**Resultado 2 (Yamada-Watanabe II)** *Considere a seguinte equação diferencial estocástica  $d$ -dimensional:*

$$dX_t = \mu(X, t)dt + \sigma(X, t)dW_t, \quad X, W \in \mathbb{R}^d \quad (2.15)$$

e suponha que  $X_0 \geq 0$ . Então as condições suficientes para unicidade<sup>5</sup> (pathwise) são dadas por:

Existem duas funções positivas e crescentes  $\rho, h : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  tais que:

- $h$  é (fracamente) côncava e o drift satisfaz:

$$|\mu_i(u) - \mu_i(v)| \leq h(\|u - v\|), \quad \forall u, v \in \mathbb{R}^d, \quad i = 1, \dots, d \quad (2.16)$$

e

$$\int_0^\epsilon h^{-1}(y)dy = +\infty, \quad \forall \epsilon > 0;$$

- O coeficiente de difusão tem a seguinte forma:

$$\sigma(X) = \begin{bmatrix} \sigma_1(X_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_d(X_d) \end{bmatrix} \quad (2.17)$$

com os elementos da diagonal  $\sigma_i : \mathbb{R} \rightarrow [0, +\infty], i = 1, \dots, d$ , satisfazendo:

$$|\sigma_i(\eta) - \sigma_i(\xi)| \leq \rho(\|\eta - \xi\|), \quad \forall \eta, \xi \in \mathbb{R}, \quad i = 1, \dots, d \quad (2.18)$$

e

$$\int_0^\epsilon \rho^{-2}(y)dy = +\infty, \quad \text{com algum } \epsilon > 0.$$

Uma vez estabelecidas as condições suficientes para a unicidade de uma EDE  $d$ -dimensional, tem-se que que na representação afim dada pela equação (2.5) o coeficiente de difusão deverá satisfazer:

$$\Sigma \in \mathcal{D}, \quad b \in \mathcal{D}^+, \quad a_i = 0 \quad \forall i. \quad (2.19)$$

Assim, para se obter a unicidade da solução, dado que o coeficiente de difusão  $\Sigma\sqrt{\sigma}$  não é globalmente Lipschitziano, este deve ser uma matriz diagonal.

Entretanto, como forma de flexibilizar as restrições sobre o coeficiente de difusão  $\Sigma\sqrt{\sigma}$ , dadas em (2.19), Dai e Singleton (2000), constroem as difusões afim em uma estrutura mais flexível denominada Representação Canônica. Para tanto, seja  $m = \text{posto}(\sigma)$ , que determinará o número de variáveis de estado que estão na matriz  $\sigma$ . Assim, qualquer equação diferencial estocástica  $d$ -dimensional afim, poderá ser classificada em uma das  $d + 1$  sub-famílias geradas a partir de  $m$ ,

<sup>5</sup>Porém de acordo com Ikeda e Watanabe (1981) pag. 52, se para qualquer medida de probabilidade de Borel  $\mathbb{P}$  em  $\mathbb{R}^d$ , uma solução  $Z_t$  para 2.15 existe tal que a distribuição do valor inicial  $Z_0$  é  $\mathbb{P}$ , e a unicidade pathwise é assegurada, então  $Z_t$  é uma solução forte de (2.15).

pois este poderá variar de 0 a  $d$ . Com isso a Representação Canônica  $\mathbb{A}_m(d)$ , definirá o modelo  $d$ -dimensional indexado pelo parâmetro  $m$ .

**Definição 8** *Seja  $m = \text{posto}(\sigma)$ , então para cada  $m$ , particiona-se o vetor de estados  $X_t$  como  $X' = (X_t^B, X_t^D)$  onde  $X_t^B$  é um vetor em  $\mathbb{R}^m$  e  $X_t^D$  é um vetor em  $\mathbb{R}^{d-m}$ . A **Representação Canônica**  $\mathbb{A}_m(d)$ , é definida como um caso especial da equação (2.5) quando<sup>6</sup>:*

$$\mathcal{K} = \begin{bmatrix} \mathcal{K}_{m \times m}^{BB} & \mathcal{K}_{m \times (d-m)}^{BD} \\ \mathcal{K}_{(d-m) \times m}^{DB} & \mathcal{K}_{(d-m) \times (d-m)}^{DD} \end{bmatrix}, \quad (2.20)$$

$$\Theta = \begin{bmatrix} \Theta^B \\ \Theta^D \end{bmatrix}, \quad \Theta^B \in \mathbb{R}^m, \quad \Theta^D \in \mathbb{R}^{d-m}, \quad (2.21)$$

$$\Sigma = \begin{bmatrix} \Sigma_{m \times m}^{BB} & \Sigma_{m \times (d-m)}^{BD} \\ \Sigma_{(d-m) \times m}^{DB} & \Sigma_{(d-m) \times (d-m)}^{DD} \end{bmatrix}, \quad (2.22)$$

$$a = \begin{bmatrix} a^B \\ a^D \end{bmatrix}, \quad a^B \in \mathbb{R}^m, \quad a^D \in \mathbb{R}^{d-m}, \quad (2.23)$$

$$b_i = \begin{bmatrix} b_i^B & b_i^D \end{bmatrix}, \quad b_i^B \in \mathbb{R}^m, \quad b_i^D \in \mathbb{R}^{d-m}, \quad (2.24)$$

$$\mathcal{B} = \begin{bmatrix} b_1^B & b_1^D \\ \vdots & \vdots \\ b_d^B & b_d^D \end{bmatrix} = \begin{bmatrix} \mathcal{B}_{m \times m}^{BB} & \mathcal{B}_{m \times (d-m)}^{BD} \\ \mathcal{B}_{(d-m) \times m}^{DB} & \mathcal{B}_{(d-m) \times (d-m)}^{DD} \end{bmatrix}, \quad (2.25)$$

E os coeficientes acima devem satisfazer:

1. Da equação (2.19) têm-se:

$$a^B = 0, \quad \mathcal{B}^{BB} \in \mathcal{D}^+, \quad \Sigma^{BB} \in \mathcal{D}^+ \quad (2.26)$$

2. O Processo  $X^B$  é autônomo em relação à  $X^D$

$$\mathcal{K}^{BD} = 0, \quad \Sigma^{BD} = 0, \quad \mathcal{B}^{BD} = 0, \quad (2.27)$$

3. A dependência exclusiva da volatilidade dos fatores de  $X^D$  em  $X^B$  implica

$$\mathcal{B}^{DD} = 0 \quad (2.28)$$

---

<sup>6</sup>De acordo com Dai e Singleton (2000) algumas transformações invariantes preservam a admissibilidade e identificação do modelo. Para detalhes sobre quais transformações invariante possuem esta propriedade vide anexo A de Dai e Singleton (2000).



A aparente complexidade introduzida com a Representação Canônica, na verdade permitirá provar a existência e unicidade para processos mais gerais do que os que satisfazem as condições expressas em (2.19), uma vez que, com esta representação, a condição (2.26) garante a existência e unicidade de  $X^B$ . Adicionalmente, a volatilidade de  $X^D$  condicionada em  $X^B$  é conhecida, logo garante-se a unicidade de  $X^D$ . Finalmente, pela condição (2.27) sabe-se que o processo  $X^D$ , quando  $X^D \neq 0$ , possui distribuição Gaussiana e independente de  $X^B$ .

### 2.3.3 Admissibilidade

O fato de se assumir que a variável de estado  $(X_t)_{t \geq 0}$  é descrita por uma equação diferencial estocástica afim em sua representação canônica  $\mathbb{A}_m(d)$ , que engloba uma série de processos, impõem restrições conjuntas sobre o espaço de estado  $\mathcal{D}$  e os parâmetros  $\{\mathcal{K}, \Theta, \Sigma, \sigma\}$ , para que o processo  $(X_t)_{t \geq 0}$  uma vez definido em  $\mathcal{D}$  permaneça em  $\mathcal{D}$ . Ao mesmo tempo, dado um processo definido em  $\mathcal{D}$ , tem-se que  $\Sigma\sigma(x)[\Sigma\sigma(x)]^\top$  deve definir uma matriz semi-positiva definida representando a matriz de variância do processo  $(X_t)_{t \geq 0}$ .

**Definição 9** *Uma equação diferencial estocástica afim do tipo (2.5) é classificada como **admissível** se o processo de variância (2.4) é quase-certamente não negativo  $\forall t$ .*

Para atender a Definição 9, proposta por Dai e Singleton (2000), o espaço de estado  $\mathcal{D} \subset \mathbb{R}^d$  pode ser construído como:

$$\mathcal{D} = \left\{ x \in \mathbb{R}^d : a_i + b_i \cdot x \geq 0, \quad i \in \{1, \dots, d\} \right\} \quad (2.29)$$

No caso do processo de Feller unidimensional, como se verá na seção 3.5, o processo é admissível desde que a  $\kappa\theta > \frac{1}{2}\sigma^2$  (condição de Feller) seja atendida. Neste caso, garante-se que o *drift* impedirá que o processo se torne negativo quando este se aproximar de sua fronteira,  $\partial\mathcal{D} = \{x \in \bar{\mathcal{D}} : \sigma(x) = 0\}$

Duffie e Kan (1996) estendem a condição de Feller para o caso multidimensional, inclusive para os casos onde  $\langle X_{i,t}, X_{j,t} \rangle \neq 0$  para  $i \neq j$ , impedindo que a correlação entre os processos viole a restrição de negatividade quando uma das variáveis de estado se aproxima de  $\partial\mathcal{D}$ . Assim a condição de Duffie e Kan (1996) é definida como:

**Condição A (Duffie e Kan (1996)).** *Para todo  $i$ :*

1.  $\forall x$  tal que  $\sigma_i(x) = 0$ ,  $b_i^\top(\mathcal{K}(\Theta - X_t) > \frac{1}{2}b_i^\top \Sigma \Sigma^\top b_i$
2.  $\forall j$ , se  $(b_i^\top \Sigma)_j \neq 0$ , então  $\sigma_i = k\sigma_j$  para  $k > 0$

Adicionalmente e diferente de Duffie e Kan (1996), no presente caso deve-se impor a condição de não-negatividade em  $(\lambda_t)_{t \geq 0}$  seja atendida. Neste sentido, um segundo conjunto de restrições

conjuntas sobre o espaço de estado  $\mathcal{D}$  e sobre os parâmetros  $\{\mathcal{K}, \Theta, \Sigma, \sigma\}$  deve ser imposto para que a intensidade seja uma medida positiva:

$$\mathcal{D} = \left\{ x \in \mathbb{R}^d : \lambda_t(x) \geq 0 \right\} \quad (2.30)$$

Neste sentido, a forma mais forte de garantir que (2.29) e (2.30) sejam atendidas simultaneamente é definir  $\mathcal{D} = \mathbb{R}_+^d$ . Entretanto, por se tratar de uma restrição muito forte no espaço de estado e que excluiria uma grande classe de processos difusivos afim, busca-se, alternativamente, relaxar esta condição graças à representação canônica  $\mathbb{A}_m(d)$ , uma vez que, por hipótese,  $X^B \in \mathcal{D}^+$ . Desta forma, a possibilidade de  $(\lambda_t)_{t \geq 0}$  assumir valores negativo advém do fato de  $X^D$  ser Gaussiano com média  $\mu^D$  e desvio-padrão  $\sigma^D$ . Neste sentido, ao impor a condição:

$$\mu^D > 0, \quad (2.31)$$

pode-se estabelecer a seguinte Proposição:

**Proposição 2** *Se  $\gamma^* \mu^D > \sigma^D$ , então o processo  $X^D$  é quase-certamente não-negativo,  $X^D \in \mathcal{D}^+$ .*

**Prova da Proposição 2**

Pela Representação Canônica o processo  $X^D$  é Gaussiano com média  $\mu^D$  e desvio-padrão  $\sigma^D$ , logo pode-se calcular a probabilidade deste ser quase-certamente positivo, ou seja:

$$\mathbb{P}(X^D > 0) = \Phi\left(\frac{\mu^D}{\sigma^D}\right) \quad (2.32)$$

onde  $\Phi(\cdot)$  é a Função densidade acumulada da distribuição Normal Padrão.

Graças a condição (2.31), pode-se impor que  $\mathbb{P}(X^D > 0) = 1$ , logo o problema descrito em (2.32), passa a ser escrito como:

$$\gamma^* = \inf_{\gamma}(\gamma : \Phi(\gamma) = 1) \quad (2.33)$$

onde:  $\gamma = \frac{\mu^D}{\sigma^D}$

Logo ao impor que  $\gamma^* \mu^D > \sigma^D$  a condição de não-negatividade é satisfeita quase-certamente. ■

**2.4 Transformada de Laplace para o processo  $(\Lambda_t)_{t \geq 0}$**

A transformada de Laplace para integrais de processos estocásticos existe, e pode ser descrita numa forma fechada analítica, para um número bastante restrito de processos Markovianos.

Neste sentido, o trabalho de Albanese e Lawi (2004) apresenta os critérios para definir quais processos difusivos podem ter, de maneira analítica, a transformada de Laplace da integral deste processo. Para tanto, toma-se  $(X_t)_{t \geq 0}$  um processo difusivo definido no espaço de probabilidade filtrado  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$  e considere a transformada de Laplace  $L_t(X, t, \bar{\mu})$ , definida como:

$$L_t(X, t, \bar{\mu}) := \mathbb{E}_t \left[ f(X, T) \exp \left( -\bar{\mu} \int_t^T \phi(X, s) ds \right) \middle| \mathcal{G}_t \right] \quad (2.34)$$

Onde  $f \in C^{2,1}$  e  $\phi \in C^1$  é limitada e  $\bar{\mu} \geq 0$ .

Então tem-se que:

**Resultado 3 (Albanese e Lawi (2004))** *A classe de processos difusivos que possuem a transformada de Laplace para sua integral é dada por:*

$$dX_t = 2 \frac{h'(X_t)}{h(X_t)} \frac{A(X_t)^2}{R(X_t)} dt + \frac{\sqrt{2A(X_t)}}{\sqrt{R(X_t)}} dW_t \quad (2.35)$$

Com as seguintes restrições:

1. Três polinômios do segundo grau em  $x$ :  $A(x), R(x), Q(x)$ . Sendo que  $A(x)$  pertence ao conjunto  $\{1, x, x(1-x), x^2+1\}$  e  $R(X_t) \geq 0$ ;
2. A função  $h(x)$  é uma combinação linear de funções hipergeométricas confluentes<sup>7</sup> do tipo  ${}_1F_1$  se  $A(x) \in \{1, x\}$  e Gaussianas do tipo  ${}_2F_1$  caso contrário.

Então a transformada de Laplace  $L_t(X, t, \bar{\mu})$  é especificada por:

$$\phi(x) = \frac{Q(x, \bar{\mu})}{\bar{\mu} R(x)} \quad (2.36)$$

Uma vez garantida a existência da transformada  $L_t(X, t, \bar{\mu})$ , pode-se enunciar o seguinte Resultado, o qual determinará a forma funcional da transformada e será útil ao longo deste trabalho.

Para as difusões afins de que trata esta tese pode-se enunciar o seguinte resultado:

---

<sup>7</sup>Uma função hipergeométrica em sua forma geral pode ser descrita por:

$${}_pF_q(\alpha_1, \dots, \alpha_p; \gamma_1, \dots, \gamma_q; z)$$

Para  $p \leq q+1$ ,  $\gamma_j \in \mathbb{C} \setminus \mathbb{Z}_+$  e são representados pela seguinte expansão de Taylor ao redor de  $z=0$

$${}_pF_q(\alpha_1, \dots, \alpha_p; \gamma_1, \dots, \gamma_q; z) = \sum_0^{\infty} \frac{(\alpha_1)_n \dots (\alpha_p)_n}{(\gamma_1)_n \dots (\gamma_q)_n} \frac{z^n}{n!}$$

Uma função confluyente é uma forma degenerada de uma função hipergeométrica quando duas das três singularidades regulares se agrupam em uma singularidade irregular.

**Resultado 4 (Feynman-Kač)** Se  $(X_t)_{t \geq 0}$  é um processo difusivo que satisfaz a EDE (2.1) e seja  $\mathcal{A}$  seu gerador infinitesimal. Assuma que  $F, V \in C^{2,1}(\mathcal{D} \times [t, T])$  e  $\bar{\mu} \geq 0$ . Então:

$$F(X, t, \bar{\mu}) = \mathbb{E}_t \left[ f(X, T) \exp \left( -\bar{\mu} \int_t^T V(X, s) ds \right) \middle| \mathcal{G}_t \right] \quad (2.37)$$

é solução da seguinte Equação Diferencial Parcial (EDP):

$$\begin{cases} \frac{\partial F}{\partial t} = \mathcal{A}F(X, t, \bar{\mu}) - V(X, t)F(X, t, \bar{\mu}) \\ F(x, 0, \bar{\mu}) = f(x) \quad x \in \mathcal{D} \end{cases} \quad (2.38)$$

e

$$\mathcal{A}F(X, t) = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial X} \mu(X, t) + \frac{1}{2} \text{Tr} \left[ \sigma(X, t) \sigma(X, t)^\top \frac{\partial^2 F(X, t)}{\partial X \partial X} \right] \quad (2.39)$$

Além disso<sup>8</sup>, se  $U(X, t, \bar{\mu}) \in C^{2,1}$  é uma solução limitada da equação (2.38), então  $U(X, t, \bar{\mu}) = F(X, t, \bar{\mu})$  sendo expresso por (2.37).

Com o Resultado 4 pode-se enunciar a seguinte proposição:

**Proposição 3** Se  $(X_t)_{t \geq 0}$  é um processo difusivo afim  $d$ -dimensional, conforme descrito em (2.5), e atende as condições de admissibilidade (seção 2.3.3), então a solução de (2.37) é da forma:

$$F(X, t) = e^{\alpha(t) + \beta(t) \cdot X}, \quad (2.40)$$

onde os coeficientes  $\alpha(t)$  e  $\beta(t)$  são determinísticos<sup>9</sup> e satisfazem:

$$\beta'(t) = \rho_1 - \mathcal{K}^\top \beta(t) - \frac{1}{2} \beta^\top b \beta \quad (2.41)$$

$$\alpha'(t) = \rho_0 - \mathcal{K} \Theta \beta(t) - \frac{1}{2} \beta^\top a \beta \quad (2.42)$$

Com condição de contorno  $\alpha(0) = 0$  e  $\beta(0) = \bar{\mu}$ . Tendo  $a \in \mathbb{R}$  e  $b \in \mathbb{R}^d$ .

### Prova da Proposição 3

Assumindo que os processos  $(X_t)_{t \geq 0}$  e  $(\lambda_t)_{t \geq 0}$  são, respectivamente, afins conforme a Definição 2 e a equação (2.7) então a equação (2.38) torna-se:

$$0 = -(\rho_0 + \rho_1 \cdot X)F(X, t) + F_t(X, t) + F_X(X, t)(\mathcal{K}(\Theta - X_t)) + \frac{1}{2} \sum_{i,j} \frac{\partial F(X, t)}{\partial X_i \partial X_j} (a_{ij} + b_{ij} \cdot X) \quad (2.43)$$

<sup>8</sup>Onde  $Tr$  denota o traço de uma matriz.

<sup>9</sup>Onde  $'$  denota a derivada em relação à  $t$ .

Substituindo  $F(X, t) = e^{\alpha(t) + \beta(t) \cdot X}$  na EDP acima, e agrupando os termos em  $X$  se obtém:

$$u(\cdot)X + v(\cdot) = 0$$

onde

$$u(\cdot) = -\beta'(t) + \rho_1 - \mathcal{K}^\top \beta(t) - \frac{1}{2} \beta(t)^\top b \beta(t) \quad (2.44)$$

$$v(\cdot) = \alpha'(t) + \rho_0 - \mathcal{K} \Theta \beta(t) - \frac{1}{2} \beta(t)^\top a \beta(t) \quad (2.45)$$

com  $a \in \mathbb{R}$  e  $b \in \mathbb{R}^d$ .

Utilizando a técnica de separação de variáveis tem-se que  $\alpha$  e  $\beta$  satisfazem uma EDO de Riccati com condição de contorno  $\alpha(0) = 0$  e  $\beta(0) = \bar{\mu}$ , a qual surge da condição de contorno  $f(x) = e^{\bar{\mu}x}$  para todo  $x$ .

■

O sistema de Riccati acima possui solução analítica em forma fechada em muitos casos, sendo que, mesmo nos casos onde esta não está disponível é possível obter  $\alpha(t)$  e  $\beta(t)$  numericamente por meio do método explícito de Runge-Kutta<sup>10</sup>, por exemplo.

## 2.5 Distribuição do Processo de Cox $(N_t)_{t \geq 0}$

Sabe-se que, para o processo de Poisson não-homogêneo  $(N_t)_{t \geq 0}$  com intensidade  $(\lambda_t)_{t \geq 0}$ , a probabilidade do número de pontos ocorrendo no intervalo  $[t, T]$  é dada por:

$$P(N_T - N_t = k) = \frac{1}{k!} \mathbb{E} \left[ \left( \int_t^T \lambda_u du \right)^k \exp \left( - \int_t^T \lambda_u du \right) \right] \quad (2.46)$$

para  $k = 0, 1, 2, \dots$

Como um passo intermediário anunciam-se os seguintes resultados:

**Definição 10** *Define-se  $G_X(\bar{\mu})$  a função geradora de momentos para uma variável aleatória absolutamente contínua  $X$  por:*

$$G_X(\bar{\mu}) = \mathbb{E} \left( e^{\langle \bar{\mu}, X \rangle} \right)$$

onde  $\bar{\mu} \in \mathbb{R}$  e  $\langle \cdot, \cdot \rangle$  é o produto escalar, com a seguinte propriedade:

<sup>10</sup>No *Matlab*<sup>®</sup>, por exemplo, este algoritmo é implementado por meio da rotina `ode45`

**Lema 1** Se  $G_X(\bar{\mu})$  é a função geradora de momentos de  $X$ , então:

$$G_X^k(\bar{\mu}) = \mathbb{E} \left( X^k e^{\bar{\mu}X} \right), \quad (2.47)$$

onde  $G_X^k(\bar{\mu})$  é a  $k$ -ésima derivada da Função Geradora de Momentos de  $X$ .

**Prova do Lema 1**

Toma-se  $n = 1$ , então:

$$G_X^1(\bar{\mu}) = \frac{d}{d\bar{\mu}} \int_{-\infty}^{\infty} e^{\bar{\mu}X} f_X(X) dX \quad (2.48)$$

$$= \int_{-\infty}^{\infty} \left( \frac{d}{d\bar{\mu}} e^{\bar{\mu}X} \right) f_X(X) dX \quad (2.49)$$

$$= \int_{-\infty}^{\infty} X e^{\bar{\mu}x} f_X(X) dX \quad (2.50)$$

$$= \mathbb{E} (X e^{\bar{\mu}X}) \quad (2.51)$$

Para  $n > 1$  a dedução é análoga. ■

Desta forma, pode-se enunciar o seguinte Teorema:

**Teorema 1** Ao se escolher um elemento qualquer da família de difusões afim

$$dX_t = \mathcal{K}(\Theta - X_t)dt + \Sigma \sqrt{a + bX_t} dW_t \quad (2.52)$$

para descrever a intensidade do Processo de Poisson Duplamente estocástico  $(N_s)_{s \geq 0}$ , a forma da função distribuição de probabilidade no intervalo  $[t, T]$  é dada por:

$$P(N_T - N_t = k) = \frac{1}{k!} \mathbb{E} \left[ (\Lambda_t)^k e^{-\Lambda_t} \right] = \frac{1}{k!} G_{\Lambda_t}^k(1), \quad (2.53)$$

onde  $G_{\Lambda_t}^k(\bar{\mu})$  é a  $k$ -ésima derivada da Função Geradora de Momentos de  $\Lambda_t$  a qual, para  $\bar{\mu} = 1$ , possui forma:

$$F(X, t) = e^{\alpha(t) + \beta(t) \cdot X} \quad (2.54)$$

com os coeficientes  $\alpha(t)$  e  $\beta(t)$  determinísticos e satisfazem:

$$\beta'(t) = \rho_1 - \mathcal{K}^\top \beta(t) - \frac{1}{2} \beta^\top b \beta \quad (2.55)$$

$$\alpha'(t) = \rho_0 - \mathcal{K} \Theta \beta(t) - \frac{1}{2} \beta^\top a \beta \quad (2.56)$$

com  $a \in \mathbb{R}$  e  $b \in \mathbb{R}^d$ .

### Prova do Teorema 1

Com o uso dos resultados anteriores a prova é direta. Para tanto, ao tomar  $k = 0$  na equação (2.46) tem-se que a prova é obtida a partir do uso da Proposição 3. A generalização para  $k > 0$  é obtida ao se combinar a solução de  $k = 0$  com o Lema 1. ■

#### 2.5.1 Momentos do processo $(N_t)_{t \geq 0}$

Uma das propriedades do Processo de Poisson Duplamente Estocástico, **PPDE**, é que condicionado a uma particular realização da intensidade  $(\lambda_t)_{t \geq 0}$ , o salto do processo pontual torna-se um processo de Poisson não-homogêneo com intensidade  $\lambda_t$ . Sendo assim, pode-se enunciar que:

**Propriedade 1** *O PPDE tem as seguintes propriedades condicionais:*

$$1. - \quad \mathbb{E}(N_t | \mathcal{F}_t) = \Lambda_t \quad (2.57)$$

$$2. - \quad \text{Var}(N_t | \mathcal{F}_t) = \Lambda_t \quad (2.58)$$

Onde  $\Lambda_t$  é o processo de intensidade acumulado.

Desta forma, a expressão não condicional para o primeiro momento é dada por:

$$\mathbb{E}(N_t) = \mathbb{E}(\Lambda_t) \quad (2.59)$$

Enquanto que, o segundo momento do processo  $(N_t)_{t \geq 0}$  utiliza, adicionalmente, a fórmula da decomposição de variância e as propriedades (2.57) e (2.58), assim:

$$\text{Var}(N_t) = \mathbb{E}(\text{Var}(N_t | \mathcal{F}_t)) + \text{Var}(\mathbb{E}(N_t | \mathcal{F}_t)) \quad (2.60)$$

$$\text{Var}(N_t) = \mathbb{E}(\Lambda_t) + \text{Var}(\Lambda_t) \quad (2.61)$$

Desta forma, os momentos do processo  $(N_t)_{t \geq 0}$  são obtidos a partir dos momentos do processo acumulado  $(\Lambda_t)_{t \geq 0}$ .

#### 2.5.2 O processo Compensado $(M_t)_{t \geq 0}$

Sabe-se que o processo de contagem  $(N_t)_{t \geq 0}$  é um submartingal<sup>11</sup> adaptado a filtração  $\mathcal{G}_t$  e assim será possível utilizar do seguinte resultado.

---

<sup>11</sup>Um processo càdlàg  $\mathcal{G}_t$ -adaptado  $(N_t)_{t \geq 0}$  é um submartingale se  $E[|N_t|] < \infty$  para cada  $t$  e se  $s < t$  implica que  $E[N_t | \mathcal{G}_s] \geq N_s$ .

**Resultado 5 (Decomposição de Doob-Meyer)** *Seja  $(N_t)_{t \geq 0}$  um submartingal não negativo contínuo à direita com respeito a  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbf{P})$  e sendo  $\mathbb{F}$  uma sub-filtração de  $\mathbb{G}$ . Então existe quase-certamente um Martingal  $(M_t)_{t \geq 0}$  contínuo à direita e uma sequência contínua à direita previsível e crescente  $\Lambda_t$ , tal que  $\mathbb{E}(\Lambda_t) < \infty$  e*

$$N_t = M_t + \Lambda_t \quad \forall t \geq 0 \quad (2.62)$$

Pode-se apresentar o seguinte resultado, diretamente relacionado ao Resultado 5:

**Proposição 4** *Dado o Processo de Poisson Duplamente Estocástico  $(N_t)_{t \geq 0}$  com intensidade  $\Lambda_t$ , então o processo compensado  $M_t = N_t - \Lambda_t$  é um  $\mathcal{G}_t$ -Martingal.*

**Prova da Proposição 4** Toma-se  $0 \leq t < T$  e a partir da Definição 7 e da Propriedade 1, pode-se escrever que

$$\begin{aligned} \mathbb{E}(M_T - M_t | \mathcal{G}_t) &= \mathbb{E}(N_T - N_t | \mathcal{G}_t) - \mathbb{E}(\Lambda_T - \Lambda_t | \mathcal{G}_t) \\ &= \mathbb{E}(\mathbb{E}(N_T - N_t | \mathcal{G}_t \vee \mathcal{F}_\infty) | \mathcal{G}_t) - \mathbb{E}(\Lambda_T - \Lambda_t | \mathcal{G}_t) \\ &= \mathbb{E}(\mathbb{E}(N_T - N_t | \mathcal{F}_\infty) | \mathcal{G}_t) - \mathbb{E}(\Lambda_T - \Lambda_t | \mathcal{G}_t) \\ &= \mathbb{E}(\Lambda_T - \Lambda_t | \mathcal{G}_t) - \mathbb{E}(\Lambda_T - \Lambda_t | \mathcal{G}_t) \\ &= 0 \end{aligned}$$

■

Com a Proposição 4 e o Resultado 5, pode-se justificar, conforme apontado na seção 2.2, a necessidade de expandir a filtração do tempo de parada  $\tau$ . Para tanto tem-se a seguinte observação:

**Observação 1** *O tempo de parada  $\tau$  não é um  $\mathbb{F}$ -tempo de parada quando  $\mathcal{F}_t := \sigma\{W_s : 0 \leq s \leq t\}$*

### Prova da Observação 1

Assuma por absurdo que  $\tau$  é  $\mathbb{F}$ -tempo de parada quando  $\mathcal{F}_t := \sigma\{W_s : 0 \leq s \leq t\}$ , então pela Proposição 4 existirá um processo compensado (um  $\mathbb{F}$ -Martingale)  $(M_t)_{t \geq 0}$  que poderá ser representado como a integral estocástica com respeito ao Movimento Browniano  $W$ . O que é uma contradição pois  $M_t$  deve saltar em  $\tau$ .

■

## 2.6 Estacionariedade do processo $(N_t)_{t \geq 0}$

Um elemento importante no estudo das propriedades assintóticas do processo  $(N_t)_{t \geq 0}$  ou na estimação de seus parâmetros, como será tratado no Capítulo 5, é garantir a estacionariedade do processo  $(N_t)_{t \geq 0}$ . Neste sentido, esta seção apresenta os requisitos para a estacionariedade do



PPDE quando sua intensidade é estocástica na forma afim, como descrito na equação (2.7).

Para tanto, inicia-se com a seguinte definição:

**Definição 11** *O processo  $(N_t)_{t \geq 0}$  é estacionário se, para cada  $A_1, \dots, A_n \in \mathcal{F}$  e  $h \in \mathbb{R}$*

$$(N_{A_1+h}, \dots, N_{A_n+h}) \stackrel{\mathcal{D}}{=} (N_{A_1}, \dots, N_{A_n})$$

Ou seja os incrementos de  $(N_t)_{t \geq 0}$  são invariantes em distribuição em relação à quaisquer translações  $h \in \mathbb{R}$ . A definição acima pode ser escrita com a seguinte notação compacta:

$$\theta_h N \stackrel{\mathcal{D}}{=} N, \quad h \in \mathbb{R}$$

Onde:

$$\theta_h N_A := N_{A+h}$$

e  $\stackrel{\mathcal{D}}{=}$  denota igualdade em distribuição.

Na sequência apresenta-se o funcional de Laplace que será necessário para a caracterização da estacionariedade do processo  $(N_t)_{t \geq 0}$ :

**Resultado 6** *Se o processo de Poisson  $(N_t)_{t \geq 0}$  em  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$  possui intensidade  $\lambda_t$  e  $f : \Omega \rightarrow \mathbb{R}_+$ , então o funcional de Laplace é dado por:*

$$\mathbb{E} \left[ e^{-Nf} \right] = \exp \left[ - \int_{\Omega} (1 - e^{-f(x)}) \lambda(dx) \right] \quad (2.63)$$

### Prova do Resultado 6

Considere primeiro a classe das funções simples  $f(x) = \sum_{i=1}^k a_i \mathbb{I}_{\{x \in B_i\}}$  para a sequência não negativa  $a_1, \dots, a_k$  e para os conjuntos disjuntos mensuráveis  $B_1, \dots, B_k \in \Omega$ , tais que  $\bigcup B_i = \Omega$ . Então:

$$Nf = \sum_{i=1}^k a_i \int_{\Omega} \mathbb{I}_{\{x \in B_i\}} N(dx) = \sum_{i=1}^k a_i N(B_i) \quad (2.64)$$

A partir do resultado acima e da independência de  $N(B_i)$  se obtém:

$$\mathbb{E} \left[ e^{-Nf} \right] = \prod_{i=1}^k \mathbb{E} \left[ e^{-a_i N(B_i)} \right] = \exp \left[ - \sum_{i=1}^k \mu(B_i) (1 - e^{-a_i}) \right] \quad (2.65)$$

$$= \exp \left[ - \int_{\Omega} (1 - e^{f(x)}) \mu(dx) \right] \quad (2.66)$$

Sendo  $f : \Omega \rightarrow \mathbb{R}_+$  contínua, existe uma sequência  $f_n$  tal que  $f_n \uparrow f$ , logo pelo Teorema da

Convergência Monótona, têm-se:

$$\mathbb{E} \left[ e^{-Nf} \right] = \lim_{n \rightarrow \infty} \mathbb{E} \left[ e^{-Nf_n} \right] = \lim_{n \rightarrow \infty} \exp \left[ - \int_{\Omega} (1 - e^{f_n(x)}) \mu(dx) \right] \quad (2.67)$$

$$= \exp \left[ - \int_{\Omega} (1 - e^{f(x)}) \mu(dx) \right] \quad (2.68)$$

■

Finalmente, pode-se definir as condições para caracterizar a distribuição de equilíbrio do processo  $(N_t)_{t \geq 0}$ :

**Proposição 5** *O PPDE  $(N_t)_{t \geq 0}$  é estacionário se a intensidade  $(\lambda_t)_{t \geq 0}$  é estacionária*

**Prova da Proposição 5**

A transformada de Laplace de um processo pontual  $(N_t)_{t \geq 0}$  é definida por

$$L_N(f) := \mathbb{E} \left\{ \exp \left[ - \int_{\Omega} f(t) N(dt) \right] \right\} \quad (2.69)$$

Adicionalmente, para um processo pontual  $(N_t)_{t \geq 0}$ , com intensidade  $(\lambda_t)_{t \geq 0}$  sua transformada de Laplace é dada por:

$$L_N(f) = \mathbb{E} \left\{ \exp \left[ - \int_{\Omega} (1 - e^{-f(t)}) \lambda(t)(dt) \right] \right\} \quad (2.70)$$

Desta forma, tomando um  $h \in \mathbb{R}$  e (2.69) têm-se:

$$\begin{aligned} L_{\theta_h N} &= \mathbb{E} \left\{ \exp \left[ - \int_{\Omega} f(t) N(dt + h) \right] \right\} \\ &= \mathbb{E} \left\{ \exp \left[ - \int_{\Omega} f(t - h) N(dt) \right] \right\} \\ &= \mathbb{E} \left\{ \exp \left[ - \int_{\Omega} (1 - e^{-f(t-h)}) \lambda(t)(dt) \right] \right\} \\ &= \mathbb{E} \left\{ \exp \left[ - \int_{\Omega} (1 - e^{-f(t)}) \theta_h \lambda(t)(dt) \right] \right\} \end{aligned} \quad (2.71)$$

Portanto, se  $(\lambda_t)_{t \geq 0}$  é estacionário, tal que  $\theta_h \lambda \stackrel{D}{=} \lambda$ , então usando este fato em (2.71) tem-se que:

$$L_{\theta_h N}(f) = L_N(f) \quad (2.72)$$

Portanto, quando existir, a transformada de Laplace é única e prova-se que o PPDE  $(N_t)_{t \geq 0}$  é estacionário se sua intensidade  $(\lambda_t)_{t \geq 0}$  é estacionária.

■

Portanto, a partir da Proposição 5, define-se a estacionariedade do processo  $(N_t)_{t \geq 0}$  a partir do comportamento de  $(\lambda_t)_{t \geq 0}$ .

### 2.6.1 Estacionariedade do processo $(\lambda_t)_{t \geq 0}$

Conforme descrito na seção anterior a caracterização da estacionariedade do processo pontual  $(N_t)_{t \geq 0}$ , reside no comportamento de sua intensidade,  $(\lambda_t)_{t \geq 0}$ . Uma vez que a intensidade foi definida pela equação (2.7):

$$\lambda_t = \rho_0 + \sum_{i=1}^d \rho_{i,1} X_{i,t} \quad (2.73)$$

$$= \rho_0 + \rho_1 \cdot X_t \quad (2.74)$$

se faz necessário estudar a estacionariedade da variável de estado  $(X_t)_{t \geq 0}$ . Neste sentido, Dai e Singleton (2000) adicionam às condições de admissibilidade descritas na seção 2.3.3 que para a equação:

$$dX_t = \mathcal{K}(\Theta - X_t)dt + \Sigma\sqrt{\sigma}dW_t$$

os auto-valores de  $\mathcal{K}$  sejam todos positivos.

## 2.7 Distribuição Estacionária do processo $(N_t)_{t \geq 0}$

Quando o processo  $(N_t)_{t \geq 0}$  for estacionário, busca-se, adicionalmente, determinar de maneira analítica a forma da distribuição estacionária, ou seja, o comportamento de  $(N_t)_{t \geq 0}$  quando  $t \rightarrow \infty$ . Desta forma, pode-se introduzir a seguinte Definição:

**Definição 12** *Suponha que  $\mu, \mu_1, \dots$  são medidas localmente finitas<sup>12</sup> em  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ . Então  $\mu_n$  converge vagamente para  $\mu$ , com notação  $\mu_n \xrightarrow{v} \mu$ , se*

$$\mu_n f \rightarrow \mu f, \text{ quando } n \rightarrow \infty, \text{ para cada } f \in C_K^+(\Omega)$$

onde:  $C_K^+(\Omega)$  é a classe das funções contínuas com suporte compacto  $f : \Omega \rightarrow \mathbb{R}_+$

A partir da Definição 12 pode-se apresentar o resultado bem conhecido de convergência para processos pontuais:

**Resultado 7 (Kallenberg (1986))** *Para a sequência de processos pontuais  $N_1, N_2, \dots$  em  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$  os seguintes resultados são equivalentes quando  $t \rightarrow \infty$ :*

1.  $N_t \xrightarrow{v} N_\infty$

---

<sup>12</sup>Seja  $(\Omega, T)$  uma espaço topológico de Hausdorff e  $F$  uma  $\sigma$ -álgebra em  $\Omega$  que contém a topologia  $T$ . Uma medida é **localmente finita** se  $\forall p \in \Omega, \exists N_p \in T$  tal que  $p \in N_p$  e  $|\mu(N_p)| < +\infty$ .

2.  $N_t f \xrightarrow{v} N_\infty f$ , para cada  $f \in C_K^+(\Omega)$
3.  $\mathbb{E}(e^{N_t f}) \xrightarrow{v} \mathbb{E}(e^{N_\infty f})$ , para cada  $f \in C_K^+(\Omega)$

Fazendo uso do Resultado 7, tem-se:

**Resultado 8** Para cada  $t \geq 1$ , suponha que  $(N_t)_{t \geq 0}$  é um processo de Poisson em  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$  com intensidade  $(\lambda_t)_{t \geq 0}$ . Se  $\lambda_t \xrightarrow{v} \lambda$  e  $\lambda$  é localmente finita, então  $N_t \xrightarrow{v} N_\infty$ , onde  $N_\infty$  é um processo de Poisson com intensidade  $\lambda$

### Prova do Resultado 8

A partir do Resultado 6, sabe-se que  $\mathbb{E}[e^{N_t f}] = e^{\lambda_t h}$ , para  $f \in C_K^+(\Omega)$ , onde  $h(t) = (1 - e^{-f(t)})$ . Adicionalmente, pelo Resultado 7, se tem  $\lambda_t h \xrightarrow{v} \lambda h$ , e então:

$$\mathbb{E}[e^{N_t f}] = e^{\lambda_t h} \xrightarrow{v} e^{\lambda h} = \mathbb{E}[e^{N_\infty f}] \quad (2.75)$$

Segue que  $N_t \xrightarrow{v} N_\infty$  pelo Resultado 7. ■

Neste capítulo foram derivadas algumas propriedades gerais do processo  $(N_t)_{t \geq 0}$ , quando a intensidade  $(\lambda_t)_{t \geq 0}$  segue um processo afim  $d$ -dimensional. No Capítulo 3 uma forma específica da intensidade será assumida e resultados mais detalhados serão obtidos.

## Capítulo 3

# Processo de Poisson Duplamente Estocástico - Intensidade de Feller

### 3.1 Introdução

O objetivo deste Capítulo é utilizar os resultados obtidos no Capítulo 2 quando se impõem uma forma específica para os parâmetros  $\mathcal{K}(\Theta - X_t)$  e  $\Sigma\sqrt{\sigma}$ . A parametrização escolhida é a difusão de Feller unidimensional devido a sua facilidade em obter soluções analíticas.

### 3.2 Construção do Processo Duplamente Estocástico

Considere o espaço de probabilidade filtrado  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ , onde  $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$  é uma filtração que contém os conjunto de medida  $\mathbb{P}$ -nula contínua à direita, como definido no Capítulo 2.

**Definição 13** *Define-se  $X_t : \Omega \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , a **difusão de Feller unidimensional**, como:*

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t \quad (3.1)$$

onde  $\kappa > 0$ ,  $\theta > 0$  e  $\sigma > 0$  são constantes e  $W_t$  é o Movimento Browniano Padrão unidimensional.

A partir da definição da variável de estado conforme uma difusão de Feller unidimensional, a intensidade do processo pontual  $(\lambda_t)_{t \geq 0}$ , é obtida ao estabelecer  $\rho_0 = 0$  e  $\rho_1 = 1$  na definição (2.7). A dinâmica de  $(\lambda_t)_{t \geq 0}$  é obtida a partir do uso do Lema de Itô e de (3.1), desta forma, a intensidade do processo de Poisson Duplamente estocástico é:

$$d\lambda_t = \kappa(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t \quad (3.2)$$

onde  $\kappa > 0$ ,  $\theta > 0$  e  $\sigma > 0$  são constantes e  $W_t$  é o Movimento Browniano Padrão unidimensional.

De acordo com a representação canônica descrita na seção 2.3.1 este processo é da classe  $\mathbb{A}_1(1)$ , logo garante-se automaticamente que  $(\lambda_t)_{t \geq 0} \geq 0$ .

### 3.3 Transformada de Laplace para o processo $(\Lambda_t)_{t \geq 0}$

Com base no Resultado 3, deseja-se avaliar se a transformada de Laplace do processo de intensidade acumulado  $\Lambda_t := \int_0^t \lambda_u du$  existe. Para tanto,

**Proposição 6** *A transformada de Laplace  $L_t(\bar{\mu}, X_t)$  para o processo difusivo de Feller unidimensional existe.*

#### Prova da Proposição 6

Sabe-se do Resultado 3, (Albanese e Lawi (2004)), que a transformada de Laplace da integral de um processo existe, se é possível representá-lo conforme a seguinte EDE:

$$dX_t = 2 \frac{h'(X_t)}{h(X_t)} \frac{A(X_t)^2}{R(X_t)} dt + \frac{\sqrt{2}A(X_t)}{\sqrt{R(X_t)}} dW_t \quad (3.3)$$

Assim, assumamos as seguintes formas funcionais para os polinômios:

$$A(x) = x, \quad R(x) = \frac{2x}{\sigma^2}, \quad h(x) = x^{a/\sigma^2} e^{-\frac{b}{\sigma^2}x} \quad (3.4)$$

Substituindo-os em (3.3), e realizando a mudança de variável  $x = \lambda$ , obtém-se

$$d\lambda_t = \kappa(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t$$

■

Uma vez garantida a existência da transformada  $L_t(\bar{\mu}, X_t)$ , pode-se utilizar o resultado do Teorema 1 para determinar explicitamente os parâmetros  $\alpha$  e  $\beta$ .

**Proposição 7** *A Transformada de Laplace,  $L_t(\bar{\mu}, \lambda_t)$ , para o processo difusivo de Feller, é dada por:*

$$L_t(\bar{\mu}, \lambda_t) := \mathbb{E} \left( e^{-\bar{\mu} \int_t^T \lambda_u du} \right) = e^{\alpha(t,T) - \beta(t,T)\lambda_t}, \quad (3.5)$$

onde

$$\alpha(t, T) = \frac{2\kappa\theta}{\sigma^2} \ln \left( \frac{2\gamma (e^{(\gamma+\kappa)/2})}{(\gamma + \kappa) (e^{-\gamma(T-t)} - 1) + 2\gamma} \right), \quad (3.6)$$

$$\beta(t, T) = \left[ \frac{2\mu (e^{-\gamma(T-t)} - 1)}{(\gamma + \kappa) (e^{-\gamma(T-t)} - 1) + 2\gamma} \right] \quad (3.7)$$

$$\text{e } \gamma = \sqrt{\kappa^2 + 2\sigma^2\bar{\mu}} \quad (3.8)$$

**Prova da Proposição 7** Como a intensidade do processo de Poisson é descrita por (3.2) o Resultado 4 permite obter a Transformada de Laplace, como a solução da seguinte EDP:

$$\frac{\partial F}{\partial t} + \kappa(\theta - \lambda) \frac{\partial F}{\partial \lambda} + \frac{1}{2} \sigma^2 \lambda \frac{\partial^2 F}{\partial \lambda^2} - \lambda \bar{\mu} F = 0, \quad (3.9)$$

com condição de contorno  $F(T, \lambda, \bar{\mu}) = 1$

Conforme enunciado na Proposição 3 seja a solução da EDP da seguinte forma:

$$F(t, T, \lambda) = \alpha(t, T) e^{-\beta(t, T) \lambda_t} \quad (3.10)$$

Segue que,

$$\begin{aligned} F_\lambda &= -\alpha \beta e^{-\beta \lambda} \\ F_{\lambda \lambda} &= \alpha \beta^2 e^{-\beta \lambda} \\ F_t &= \alpha_t e^{-\beta \lambda} - \alpha \beta_t \lambda e^{-\beta \lambda} \end{aligned}$$

Substituindo os termos acima na equação (3.9):

$$\lambda \left( \frac{\sigma^2}{2} \alpha \beta^2 - \alpha \beta_t + \alpha \beta - \alpha \right) = \kappa \theta \alpha \beta - \alpha_t \quad (3.11)$$

Ao igualar os dois lados da equação acima obtém-se:

$$\begin{cases} \frac{\sigma^2}{2} \beta^2 - \beta_t + \beta - 1 = 0 \\ \alpha_t - \kappa \theta \alpha \beta = 0 \end{cases} \quad (3.12)$$

A primeira equação do sistema (3.12) é a equação de Riccati com solução  $\beta(t, T) = v(t, T)/u(t, T)$  onde  $v(t, T)$  e  $u(t, T)$  são soluções para o seguinte sistema de equações<sup>13</sup>:

$$\begin{cases} \frac{\sigma^2}{2} v(t, T) + u'(t, T) = 0 \\ u(t, T) + v'(t, T) - \kappa v(t, T) = 0 \end{cases} \quad (3.13)$$

Defina  $\Delta = T - t$ , então  $\frac{\partial}{\partial t} = -\frac{d}{d\Delta}$  e o sistema acima pode ser reescrito como:

$$\begin{cases} \frac{\sigma^2}{2} v(\Delta) - u'(\Delta) = 0 \\ u(\Delta) - v'(\Delta) - \kappa v(\Delta) = 0 \end{cases} \quad (3.14)$$

Da segunda equação de (3.14) tem-se

$$\begin{cases} u(\Delta) = + v'(\Delta) + \kappa v(\Delta) \\ u'(\Delta) = + v''(\Delta) + \kappa v'(\Delta) \end{cases} \quad (3.15)$$

<sup>13</sup> $u$  e  $v$  são funções de  $t$  e  $T$ , mas fixa-se  $T$ , então  $v'(t, T)$  denota a derivada com respeito a  $t$  e  $F_x$  denota a derivada de  $F$  em relação a  $x$ .

Substituindo na primeira equação de (3.14) e reescrevendo em termo do operador defasagem  $\mathcal{Q}$  se obtém a forma quadrática:

$$\left( \mathcal{Q}^2 + \kappa \mathcal{Q} - \frac{\sigma^2}{2} \right) v(\Delta) = 0 \quad (3.16)$$

As raízes da equação (3.17) acrescida da condição de contorno  $\beta(T, T) = 0 = v(0)/u(0)$ , resultam na seguinte solução:

$$v(\Delta) = e^{0.5(\gamma-\kappa)\Delta} - e^{0.5(-\gamma-\kappa)\Delta} \quad (3.17)$$

Derivando (3.17) e substituindo em (3.15) obtém-se:

$$u(\Delta) = 0.5(\gamma - \kappa)e^{0.5(\gamma-\kappa)\Delta} - 0.5(-\gamma + \kappa)e^{0.5(-\gamma-\kappa)\Delta} \quad (3.18)$$

Uma vez que  $\Delta = T - t$  a solução da equação de Riccati é obtida a partir das equações (3.17) e (3.18):

$$\begin{aligned} \beta(t, T) &= v(\Delta)/u(\Delta) \\ \beta(t, T) &= \left[ \frac{2(e^{-\gamma(T-t)} - 1)}{(\gamma + \kappa)(e^{-\gamma(T-t)} - 1) + 2\gamma} \right] \end{aligned} \quad (3.19)$$

Ao tomar  $T$  fixo tem-se que  $\alpha(t, T)$  na equação (3.12) é função apenas de  $t$ , logo:

$$\begin{aligned} \frac{\partial \alpha}{\partial t} &= \kappa \theta \alpha \beta \\ \alpha(t, T) &= \exp \left( -\kappa \theta \int_t^T \beta(s, T) ds \right) \end{aligned}$$

Substituindo o valor de  $\beta(t, T)$ , conforme determinado na equação (3.19), obtém-se:

$$\alpha(t, T) = \frac{2\kappa\theta}{\sigma^2} \ln \left( \frac{2\gamma (e^{(\gamma+\kappa)/2})}{(\gamma + \kappa)(e^{-\gamma(T-t)} - 1) + 2\gamma} \right) \quad (3.20)$$

■

### 3.4 Distribuição do Processo de Cox $(N_t)_{t \geq 0}$

Sabe-se que, para o processo de Poisson não-homogêneo  $(N_t)_{t \geq 0}$  com função intensidade  $(\lambda_t)_{t \geq 0}$ , a probabilidade do número de pontos ocorrendo no intervalo  $[t, T]$  é dada por:

$$P(N_t - N_T = k) = \frac{1}{k!} \mathbb{E} \left[ \left( \int_t^T \lambda_u du \right)^k \exp \left( - \int_t^T \lambda_u du \right) \right] \quad (3.21)$$

para  $k = 0, 1, 2, \dots$



Utilizando a equação (3.21) e o Teorema 1 pode-se enunciar a seguinte Proposição:

**Proposição 8** *A função distribuição de probabilidade, (FDP), para o processo de Poisson não-homogêneo  $(N_t)_{t \geq 0}$  no intervalo  $[t, T]$ , quando a função intensidade  $(\lambda_t)_{t \geq 0}$  é descrita pelo processo*

$$d\lambda_t = \kappa(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t \quad (3.22)$$

é dada por:

$$P(N_T = k) = \frac{1}{k!} \mathbb{E} \left[ (\Lambda_t)^k e^{-\Lambda_t} \right] = \frac{1}{k!} G_{\Lambda_t}^k(1) \quad (3.23)$$

$$G_{\Lambda_t}(1) = \mathbb{E} \left( e^{-\int_t^T \lambda_u du} \right) = e^{\alpha(t,T) - \beta(t,T)\lambda_t} \quad (3.24)$$

e

$$\alpha(t, T) = \frac{2\kappa\theta}{\sigma^2} \ln \left( \frac{2\gamma (e^{(\gamma+\kappa)/2})}{(\gamma + \kappa) (e^{-\gamma(T-t)} - 1) + 2\gamma} \right) \quad (3.25)$$

$$\beta(t, T) = \left[ \frac{2\mu (e^{-\gamma(T-t)} - 1)}{(\gamma + \kappa) (e^{-\gamma(T-t)} - 1) + 2\gamma} \right] \quad (3.26)$$

$$\text{onde } \gamma = \sqrt{\kappa^2 + 2\sigma^2} \quad (3.27)$$

### Prova da Proposição 8

Para  $k = 0$  na equação (3.21) a solução foi obtida com a transformada de Laplace do processo de intensidade acumulado derivada na Proposição 7. A generalização para  $k > 0$  é obtido ao se combinar a solução de  $k = 0$  com o Lema 1. ■

#### 3.4.1 Momentos do processo $(N_t)_{t \geq 0}$

Uma das propriedades do Processo de Poisson Duplamente Estocástico  $(N_t)_{t \geq 0}$ , é que condicionado a uma particular realização da intensidade  $\lambda_t$ , o salto do processo pontual torna-se um processo de Poisson não-homogêneo com intensidade  $\lambda_t$ . Sendo assim, pode-se enunciar que:

**Proposição 9** *Os dois primeiros momentos do processo  $(N_t)_{t \geq 0}$  com intensidade estocástica descrita por (3.2) são:*

1.  $E(N_t) = \theta t + \frac{1-e^{-\kappa t}}{\kappa}(\lambda_0 - \theta)$

2.  $\text{Var}(N_t) = \frac{2\theta t}{\kappa} \left[ (e^{-\kappa t} + 1)(\lambda_0 - \theta) - 2(\theta + \lambda_0) \right] + \frac{\sigma^2}{\kappa^3} \left( \frac{\theta e^{-\kappa t}}{2} + \frac{4e^{-\kappa} - 5}{2} - \lambda_0 e^{-2\kappa t} \right) + \frac{\sigma^2 t}{\kappa} \left( \frac{3\theta - 2\lambda_0}{\kappa} \right)$

**Prova da Proposição 9**

A prova do item 1 é obtida com uso da transformada de Laplace de  $\Lambda_t$  avaliada em:  $-\frac{\partial}{\partial \bar{\mu}} \mathbb{E} \left( e^{-\bar{\mu} \int_0^t \lambda_s ds} \right) \Big|_{\bar{\mu}=0}$ .

Para demonstrar o item 2 faz-se uso do seguinte resultado:

$$\text{Var}(\Lambda_t) = \frac{\partial^2}{\partial \bar{\mu}^2} \mathbb{E} \left( e^{-\bar{\mu} \int_0^t \lambda_s ds} \right) \Big|_{\bar{\mu}=0} - \left( \frac{\partial}{\partial \bar{\mu}} \mathbb{E} \left( e^{-\bar{\mu} \int_0^t \lambda_s ds} \right) \Big|_{\bar{\mu}=0} \right)^2 \quad (3.28)$$

Substituindo a expressão acima em  $\text{Var}(N_t) = \mathbb{E}(\Lambda_t) + \text{Var}(\Lambda_t)$  obtém-se o resultado desejado. ■

**3.5 Distribuição estacionária do processo  $(N_t)_{t \geq 0}$** 

Nesta seção estuda-se a estacionariedade do processo de Cox quando sua intensidade é estocástica na forma (3.2). Para o processo de Feller univariado as condições requeridas no Resultado 1 são atendidas ao tomarmos  $\alpha = 1/2$ , pois o coeficiente de difusão é  $\sigma(X_t) := \sigma \sqrt{X_t}$ . Portanto, têm-se que as condições suficientes para a existência e unicidade de uma solução forte para a equação diferencial estocástica (3.2), conforme exposto no Resultado 1, são atendidas.

A prova da estacionariedade para a difusão de Feller unidimensional é feita a partir de uma construção probabilística, conforme Karlin e Taylor (1981). Sendo assim, tem-se que a probabilidade da difusão (3.2) partindo de um ponto arbitrário  $x$  atinja  $b$  antes de tocar em  $a$ , ou seja,  $\mathbb{P}_x(T_b < T_a)$ , onde  $T_j := \inf\{t > 0 : X_t \in j\}$  será importante para avaliar a recorrência do processo. Além disso, o comportamento do processo de Feller em sua fronteira pode ser analisado de maneira simples.

A probabilidade  $\mathbb{P}_x(T_b < T_a)$  será calculada com a ajuda da função  $S(\lambda)$  a qual soluciona a seguinte Equação Diferencial Parcial:

$$\frac{1}{2} \sigma^2(\lambda) S''(\lambda) + \mu(\lambda) S'(\lambda) = 0 \quad \text{ou na forma compacta} \quad \mathcal{L}S = 0 \quad (3.29)$$

Qualquer função que resolva a equação (3.29) é denominada função harmônica de  $\mathcal{L}$ . Apenas soluções positivas são de interesse e ao se excluir as soluções constantes, tem-se que toda função  $\mathcal{L}$ -harmônica que resolve (3.29) é dada por:

$$S(\lambda) := \int^\lambda s(\eta) d\eta \quad (3.30)$$

e denominada função medida de escala, e

$$s(\lambda) := \exp \left( - \int^\lambda \frac{2\mu(\eta)}{\sigma^2(\eta)} d\eta \right) \quad (3.31)$$

define a função densidade de escala.

Especificamente, no tratamento das propriedades do processo  $d\lambda_t = \kappa(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t$ , têm-se que a *drift* e a difusão são respectivamente,  $\mu(\lambda) = a\sigma^2 - b\sigma^2\lambda$  e  $\sigma(\lambda) = \sigma\sqrt{\lambda}$ , com a nova parametrização  $a = \kappa\theta/\sigma^2$  e  $b = \kappa/\sigma^2$ .

Sendo assim, para determinar se o processo  $\lambda_t$  é estacionário, considere os seguintes Lemas:

**Lema 2** *O comportamento da função medida de escala  $S(x)$  na fronteiras  $[l, r]$  do processo  $(\lambda_t)_{t \geq 0}$  é:*

$$S(0, \alpha] = -\infty \quad (3.32)$$

$$S[\alpha, \infty) = \infty \quad (3.33)$$

### Prova do Lema 2

Com a nova parametrização do processo de Feller, tem-se que a densidade de escala é dada por:

$$s(\lambda) = e^{\lambda 2b} \lambda^{-2a} \quad (3.34)$$

Uma condição usual para garantir que o processo  $(\lambda_t)_{t \geq 0}$  seja recorrente, é diferenciar os casos  $a > 1/2$  e  $a < 1/2$ . Tomando  $a = 1$  tem-se que:

$$S(t) := \int_0^\alpha s(\eta) d\eta \quad (3.35)$$

$$= \int_0^\alpha \left( \frac{e^{2b\eta}}{\eta^2} \right) d\eta \quad (3.36)$$

Integrando por partes tem-se:

$$S(t) = \frac{e^{2b\eta}}{\eta} - \underbrace{2b \int_0^\alpha \left( \frac{e^{2b\eta}}{\eta} \right) d\eta}_1 \quad (3.37)$$

O termo 1 na equação (3.37), pode ser identificado como a função integral exponencial<sup>14</sup> (*Exponential Integral*)  $Ei(x)$ . Têm-se que:  $Ei(0+) = -\infty$  e  $Ei(\infty) = \infty$ , desta forma:

$$\lim_{\alpha \rightarrow 0+} S(0, \alpha] = -\infty$$

<sup>14</sup>A função integral exponencial é definida como:

$$Ei(x) = \int_{-\infty}^x \frac{e^t}{t} dt$$

Ao tomar outro valor para  $a > 1/2$ , por exemplo,  $a = 2$ , percebe-se que a mesma estrutura polinomial é obtida,

$$8b^3 \left( -1/24 \frac{e^{2b\alpha}}{b^3\alpha^3} - 1/24 \frac{e^{2b\alpha}}{b^2\alpha^2} - 1/12 \frac{e^{2b\alpha}}{b\alpha} - 1/6 Ei(1, -2b\alpha) \right)$$

E portanto o comportamento do limite não se altera.

De maneira análoga, devido ao comportamento da função integral exponencial, pode-se mostrar que

$$\lim_{\alpha \rightarrow \infty} S[\alpha, \infty] = \infty$$

■

**Lema 3** *O comportamento da função medida de velocidade  $M(x)$  na fronteiras  $[l, r]$  do processo  $(\lambda_t)_{t \geq 0}$  é:*

$$M(0, \alpha] < \infty \quad (3.38)$$

$$M[\alpha, \infty] < \infty \quad (3.39)$$

### Prova do Lema 3

A função medida de velocidade é definida por:

$$M(\alpha, \beta] := \int_{\alpha}^{\beta} m(\lambda) d\lambda \quad (3.40)$$

onde:

$$m(\lambda) := \frac{1}{\sigma^2(\lambda)s(\lambda)} \quad (3.41)$$

é a função densidade de velocidade.

**Caso A**  $M(0, \alpha] < \infty$

Para  $a > 0$  tem-se ao substituir os valores de  $\sigma^2(\lambda)$  e  $s(\lambda)$  em (3.40):

$$M(0, \alpha] = \frac{1}{\sigma^2} \int_0^{\alpha} \frac{1}{\lambda^{1-2a}} e^{-2b\lambda} d\lambda < \infty \quad (3.42)$$

Para obter o resultado acima, integra-se por partes o lado direito de (3.42) e o fato do termo  $e^{-2b\lambda}$  ser dominante e, convergir a zero.

**Caso B**  $M[\alpha, \infty] < \infty$

Novamente têm-se ao substituir os valores de  $\sigma^2(\lambda)$  e  $s(\lambda)$  em (3.40):

$$M[\alpha, \infty] = \frac{1}{\sigma^2} \int_{\alpha}^{\infty} \frac{1}{\lambda^{1-2a}} e^{-2b\lambda} d\lambda < \infty \quad (3.43)$$

O argumento é análogo ao utilizado no caso A. ■

Assim, pode-se enunciar o seguinte Resultado:

**Resultado 9** *Seja  $(X_t)_{t \geq 0}$  uma difusão com gerador  $\mathcal{L}$  e  $\sigma(x) > 0$  em  $[a, b]$  contínuo. Para  $X_0 = x$  e  $a < x < b$ , então*

$$\mathbb{P}_x(T_b < T_a) = \frac{S(x) - S(a)}{S(b) - S(a)} \quad (3.44)$$

onde  $S(x)$  é definida pela equação (3.30).

### Prova do Resultado 9

Utilizando a Fórmula de Dynkin<sup>15</sup> tem-se que:

$$\mathbb{E}_x \left[ S(X_{t \wedge \tau}) - \int_0^{t \wedge \tau} \mathcal{L}S(X_s) ds \right] = S(x) \quad (3.45)$$

onde  $\tau := T_a \wedge T_b$

Mas  $\mathcal{L}S = 0$ , portanto

$$\mathbb{E}_x [S(X_{t \wedge \tau})] = S(x)$$

Tomando o limite  $t \rightarrow \infty$ , o Teorema da Convergência Dominada garante que:

$$\mathbb{E}_x S(X_\tau) = S(X_0) = S(x)$$

Reescrevendo a expressão acima como:

$$S(x) = \mathbb{E}_x S(X_\tau) \quad (3.46)$$

$$= S(b)\mathbb{P}_x(T_b < T_a) + S(a)(1 - \mathbb{P}_x(T_b < T_a)), \quad (3.47)$$

e isolando os termos o resultado é obtido. ■

De posse do comportamento da função medida de escala  $S(x)$ , dado pelo Lema 2, pode-se enunciar a seguinte proposição:

**Proposição 10** *O processo  $(\lambda_t)_{t \geq 0}$ , parametrizado como  $d\lambda_t = \kappa(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t$ , é recorrente.*

<sup>15</sup>O valor esperado de uma função de uma difusão avaliado em um tempo de parada é uma das inúmeras contribuições de Eugene Dynkin apresentados em Dynkin, E. B. *Markov processes*. Vols. I, II. Academic Press Inc., Publishers, New York; Springer-Verlag, Berlin-Göttingen-Heidelberg 1965 Vol. I: xii+365 pp.; Vol. II: viii+274 pp. Traduzido do original Die Grundlehren der Mathematischen Wissenschaften, Bände 121, 122 por J. Fabius, V. Greenberg, A. Maitra, G. Majone.

**Prova da Proposição 10** Primeiro mostra-se que o processo não atinge o ponto zero a despeito, possivelmente, de um conjunto  $B$ , tal que  $\mu(B) = 0$ . Para tanto se utiliza do Resultado 9, e obtém-se que  $\mathbb{P}_x(T_\alpha < T_b)$  pode ser suficientemente pequena ao tomar  $a \rightarrow \alpha$ . Portanto, a fronteira  $\alpha$  não é atingida antes de um  $b$  arbitrário. Analogamente para a outra fronteira, i.e,  $\infty$ . Como o processo não explode, tem-se que o  $T_\infty = \lim_{b \rightarrow \infty} T_b = \infty$  e, para  $y < x$  arbitrários, tem-se:

$$\mathbb{P}_x(T_y < \infty) = \lim_{b \rightarrow \infty} \mathbb{P}_x(T_y < b) = \lim_{b \rightarrow \infty} \frac{S(x) - S(b)}{S(y) - S(b)} \quad (3.48)$$

De posse novamente do Lema 2 sabe-se que  $S(\infty) = \lim_{b \rightarrow \infty} S(b) = \infty$ . Utilizando a regra de l'Hôpital na equação (3.48) tem-se que  $\mathbb{P}_x(T_y < \infty) = 1$  para  $y$  arbitrário, provando a recorrência do processo  $(\lambda_t)_{t \geq 0}$ . ■

Desta forma, a partir do Lema 3 e da Proposição 10 pode-se enunciar que:

**Proposição 11** *O processo  $d\lambda_t = \kappa(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t$  :*

1. *É estacionário;*
2. *Sua distribuição estacionária  $\Psi(x)$  é dada por:*

$$\Psi(x) = \frac{C}{\sigma^2(x)s(x)}$$

**Prova da Proposição 11**

Seja  $p(s, t, x, y)$  a densidade de transição do processo  $(\lambda_t)_{t \geq 0}$ , por simplicidade, assumamos que  $\lambda_t = c \in \mathbb{R}_+$  no instante 0. Assim, a densidade de transição poderá ter uma notação mais simples,  $p(t, y)$ . Da equação *Forward* de Kolmogorov (ou Fokker-Plank, EFP), tem-se:

$$\frac{\partial p}{\partial t}(t, y) = \frac{1}{2} \frac{\partial^2}{\partial y^2} \left[ \sigma^2(y)p(t, y) \right] - \frac{\partial}{\partial y} \left[ \mu(y)p(t, y) \right] \quad (3.49)$$

A distribuição estacionária,  $\Psi(x)$  quando existir deve, necessariamente, satisfazer:

$$\Psi(y) = \int \Psi(x)p(t, x, y)dx \quad \text{para todo } t > 0$$

Se o processo alcançou o regime estacionário, sua distribuição de probabilidade será independente de  $t$ , logo  $\frac{\partial p}{\partial t}(t, y) = 0$  e, ao combinar com (3.49), tem-se<sup>16</sup>:

$$0 = \frac{1}{2} \frac{\partial^2}{\partial y^2} \left[ \sigma^2(y)\Psi(y) \right] - \frac{\partial}{\partial y} \left[ \mu(y)\Psi(y) \right] \quad (3.50)$$

---

<sup>16</sup>Uma prova rigorosa para esta passagem pode ser encontrada na página 181 e 219 de Pinsky (1995)

Integrando a equação (3.50) tem-se:

$$\frac{d}{dy} \left[ \frac{\sigma^2(y)\Psi(y)}{2} \right] - \mu(y)\Psi(y) = \frac{1}{2}C_1 \quad (3.51)$$

Onde  $C_1$  é uma constante arbitrária. Multiplicando pelo fator integrante:

$$s(y) = \exp \left( - \int_y \frac{2\mu(\nu)}{\sigma^2(\nu)} d\nu \right)$$

pode-se escrever (3.51) em uma forma mais compacta:

$$\frac{d}{dy} \left[ s(y)\sigma^2(y)\Psi(y) \right] = C_1 s(y)$$

Uma segunda integração, com  $S(x) = \int_x s(y)dy$ , gera:

$$\Psi(x) = C_1 \frac{S(x)}{s(x)\sigma^2(x)} + C_2 \frac{1}{s(x)\sigma^2(x)} \quad (3.52)$$

$$= m(x)[C_1 S(x) + C_2] \quad (3.53)$$

onde  $m(x)$  foi definido pela equação (3.41), e as constantes  $C_1$  e  $C_2$  são determinada tal que  $\Psi(x)$  seja uma função densidade de probabilidade. Do Lema 2, sabe-se que o comportamento da função  $S(x)$  implica que  $C_1 = 0$  para manter  $\Psi(x)$  positivo em  $(l, r)$ , assim  $C_2$  é escolhido para que  $\int_l^r \Psi(x)dx = 1$ , portanto:

$$\Psi(x) = \frac{1}{M} \frac{1}{\sigma^2(x)s(x)} dx \quad \forall x \in (l, r)$$

Logo, pelo resultado do Lema 3, tem-se que:

$$M = \int_l^r \frac{1}{\sigma^2(x)s(x)} dx < \infty$$

Segue que  $\Psi(x)$  existe e portanto o processo é estacionário. ■

Finalmente, ao combinar a Proposição 5 com a Proposição 11 obtém-se:

**Proposição 12** *O processo  $(N_t)_{t \geq 0}$  com intensidade da forma  $d\lambda_t = \kappa(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t$  é estacionário.*

**Prova da Proposição 12** O resultado é direto ao combinar as Proposição 5 com a Proposição 11. ■

Finalmente, pode-se analisar o comportamento dos processos  $(\lambda_t)_{t \geq 0}$  e  $(N_t)_{t \geq 0}$  quando  $t \rightarrow \infty$ . Desta forma, inicia-se com a determinação da distribuição de  $\lambda_t$ . E para tanto, enuncia-se a seguinte Proposição:

**Proposição 13** *A distribuição estacionária do processo  $(\lambda_t)_{t \geq 0}$  é a distribuição Gama, com parâmetros  $\alpha = \frac{2\kappa\theta}{\sigma^2}$  e  $\beta = \frac{2\kappa}{\sigma^2}$*

**Prova da Proposição 13**

A partir do resultado da Proposição 11, encontrou-se que a distribuição estacionária tem o seguinte formato:

$$\Psi(x) = \frac{C}{\sigma^2(x)s(x)}$$

Logo, substituindo a forma de  $s(x)$  conforme descrito na equação (3.34), deve-se determinar a constante normalizadora  $C$ . Então

$$\begin{aligned}\Psi(\lambda) &= C e^{\left(\frac{2\kappa}{\sigma^2}\right)\lambda^{\frac{2\theta\kappa}{\sigma^2}-1}} \\ \int_0^\infty \left( C e^{-\left(\frac{2\kappa}{\sigma^2}\right)\lambda^{\frac{2\theta\kappa}{\sigma^2}-1}} \right) d\lambda &= 1\end{aligned}$$

Realizando a mudança de variável  $t = \frac{2\kappa\lambda}{\sigma^2}x$ , tem-se:

$$\int_0^\infty \left[ e^{-t} \left( t \frac{\sigma^2}{2\kappa} \right)^{\alpha-1} \frac{\sigma^2}{2\kappa} \right] dt = \frac{1}{C}$$

e simplificando,

$$\left[ \frac{\sigma^2}{2\kappa} \right]^\alpha \underbrace{\int_0^\infty e^{-t} t^{\alpha-1} dt}_{\Gamma(\alpha)} = \frac{1}{C}$$

onde  $\Gamma(\alpha)$  é a função Gama.

Assim a distribuição estacionária é dada por:

$$\Psi(\lambda) = \beta^\alpha \frac{1}{\Gamma(\alpha)} e^{-\beta\lambda} \lambda^{\alpha-1}$$

Ou seja,  $\Psi(\lambda) \sim Gama(\alpha, \beta)$

■

A partir da Proposição 12, foi possível definir as condições para assegurar quando o processo  $(N_t)_{t \geq 0}$  é estacionário. Adicionalmente, deseja-se determinar de maneira analítica o formato da distribuição de  $(N_t)_{t \geq 0}$  quando  $t \rightarrow \infty$ .

**Proposição 14** *A distribuição estacionária de  $N_t$  para  $t \rightarrow \infty$ , representada por  $N_t^\pi$ , quando  $\lambda_t$  é descrita por  $d\lambda_t = \kappa(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t$  é a distribuição binomial negativa.*

**Prova da Proposição 14** Para tanto faz-se uso da Proposição 13 e do Resultado 7:



$$\begin{aligned}
\mathbb{P}(N_t = n) &= \int_0^\infty \mathbb{P}(N_t = n | \Lambda = \lambda) \Psi(\lambda) d\lambda \\
&= \int_0^\infty \frac{(\lambda t)^k e^{-\lambda t}}{k!} \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta \lambda} \lambda^{\alpha-1} d\lambda \\
&= \frac{t^k}{k!} \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha + k)}{(\beta + t)^{\alpha+k}} \int_0^\infty \frac{(\beta + t)^{\alpha+k}}{\Gamma(\alpha + k)} e^{-\lambda(\beta+t)} \lambda^{k+\alpha-1} d\lambda
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(N_t = n) &= \frac{(\alpha + k - 1)!}{(\alpha - 1)! \alpha!} \left( \frac{t}{\beta + t} \right)^k \left( \frac{\beta}{\beta + t} \right)^\alpha \\
&= \text{Neg}(\alpha, p) \quad \text{com} \quad p = \frac{\beta}{\beta + t}
\end{aligned}$$

Portanto  $N_t^\pi \sim \text{Neg}(\alpha, p)$

■

Uma vez determinada a distribuição estacionária do processo  $(N_t)_{t \geq 0}$  busca-se encontrar qual é a velocidade com que este processo converge para a distribuição estacionária  $N_t^\pi$ . Desta forma, pode-se enunciar o seguinte resultado:

**Teorema 2** *O processo  $(N_t)_{t \geq 0}$  converge para sua distribuição estacionária,  $N_t^\pi$ , exponencialmente a taxa  $2\kappa$ .*

**Prova do Teorema 2** Sabe-se que a densidade de transição do processo  $(\lambda_t)_{t \geq 0}$  quando este segue uma difusão de Feller é a distribuição Qui-quadrado não centralizada. Na Proposição 13 mostrou-se que o processo  $(\lambda_t)_{t \geq 0}$  converge para a distribuição Gama. Têm-se, conforme Karlin e Taylor (1981), que ambas as distribuições possuem representação espectral dada por:

$$p(t, x, y) = (2\kappa)^{2\kappa\theta} y^{2\kappa\theta-1} e^{2\kappa\theta y} \sum_{n=1}^{\infty} e^{-2\kappa t} L_n^{2\kappa-1} \left( \frac{\kappa x}{\sigma^2/2} \right) L_n^{2\kappa-1} \left( \frac{\kappa y}{\sigma^2/2} \right) \frac{\Gamma(n+1)}{\Gamma(n+\kappa\theta)} \quad (3.54)$$

onde  $L_n^{2\kappa-1}(\cdot)$  é o polinômio de Laguerre com parâmetro  $2\kappa-1$ , com a propriedade  $L_0^{2\kappa-1}(\cdot) = 1$  e  $\Gamma$  é a função gama.

No caso da distribuição estacionária  $p(y)$ , tem-se que a distribuição Gama é obtida, segundo a representação espectral, quando  $n = 0$  na equação (3.54).

Para calcular a velocidade de convergência de  $(\lambda_t)_{t \geq 0}$ , tem-se que:

$$|p(t, x, y) - p(y)| = \mathcal{O}(e^{-2\kappa t}), \forall x > 0 \quad \text{quando} \quad t \rightarrow \infty \quad (3.55)$$

Pois na representação espectral para  $p(t, x, y)$  e  $p(y)$  tem-se que  $e^{-2\kappa t}$  é o único termo em  $t$  para  $\forall n$ .

Finalmente tomando a versão não-condicional do PPDE:

$$P(N_T - N_t = k) = \int \left[ \frac{1}{k!} (\Lambda_{t,T})^k \exp(-\Lambda_{t,T}) p(t, x, y) | \mathcal{F}_t \right] d\mathbb{P} \quad (3.56)$$

Segue que,

$$|N_t - N_t^\pi| = \mathcal{O}(e^{-2\kappa t}) \quad \text{quando } t \rightarrow \infty \text{ e} \quad (3.57)$$

Ter-se-á a convergência do processo  $(N_t)_{t \geq 0}$  para  $N_t^\pi$ .

■

Neste Capítulo foram derivadas, a partir do resultados gerais apresentados no Capítulo 2, algumas propriedades do processo  $N_t$ , quando se especifica a forma funcional da intensidade do processo evoluindo segundo um processo difusivo do tipo Feller. A difusão de Feller é particularmente útil pois o processo está bem definido apenas no  $\mathbb{R}_+$ , garantindo que a condição de não-negatividade para a medida aleatória seja automaticamente atendida. Adicionalmente, ao se adotar a intensidade de Feller é possível obter funções tratáveis analiticamente que serão úteis na etapa de estimação dos parâmetros do Processo de Poisson Duplamente Estocástico, conforme descrito no Capítulo 5.

## Capítulo 4

# Acoplamento para Processos de Poisson Duplamente Estocásticos

### 4.1 Definições

Para provar os resultados deste capítulo novas definições serão introduzidos. Desta forma,

**Definição 14** *Seja  $\mathcal{X}$  o espaço de estado e  $\mathcal{B}(\mathcal{X})$  uma  $\sigma$ -álgebra em  $\mathcal{X}$ . Uma medida de Borel em  $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$  é uma medida localmente finita<sup>17</sup>.*

**Definição 15** *Uma medida aleatória  $\Lambda$  é um mapeamento mensurável do espaço de probabilidade  $(\Omega, \mathcal{G}, \mathbb{P})$  em  $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ , onde  $\mathcal{M}$  é dotado da topologia vaga e formado por todas as medidas de Borel e  $\mathcal{B}(\mathcal{M})$  é a  $\sigma$ -álgebra gerada pelos conjuntos aberto em  $\mathcal{M}$ .*

**Definição 16** *Seja  $\mathcal{N} \in \mathcal{B}(\mathcal{M})$  o conjunto de todos os valores inteiros ou infinito de  $\mathcal{M}$ .*

Adota-se na sequência que os elementos em  $\mathcal{M}$  e  $\mathcal{N}$  são representados por  $\varsigma$  e  $\nu$ , respectivamente.

**Definição 17** *A distribuição de  $\Lambda$  é a medida de probabilidade  $\Pi$  em  $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$  induzida por  $\Lambda$ .*

**Definição 18** *Uma medida aleatória com distribuição  $\Pi$  é denominada de **processo pontual** se  $\Pi(\mathcal{N}) = 1$ .*

**Definição 19** *Um Processo Pontual  $N_t$  com distribuição  $\int_{\mathcal{M}} \Pi_{\varsigma} \Pi(d\varsigma)$  para alguma medida de probabilidade em  $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$  é chamado de **Processo de Poisson Duplamente estocástico**.*

Neste caso, tem-se que para qualquer  $B$  limitado pode-se escrever:

$$\mathbb{P}(N_t(B) = k) = \mathbb{P}(\nu \in \mathcal{N}; \nu(B) = k) \quad (4.1)$$

$$= \int_{\mathcal{M}} \frac{\varsigma(B)^k}{k!} e^{-\varsigma(B)} \Pi(d\varsigma) \quad (4.2)$$

$$= \mathbb{E} \left( \frac{\Lambda(B)^k}{k!} e^{-\Lambda(B)} \right) \quad (4.3)$$

---

<sup>17</sup>Definição pode ser encontrada na nota 12.

Seja  $\mathcal{N} \times \mathcal{M}$  o produto de  $\mathcal{N}$  e  $\mathcal{M}$  dotado da topologia do produto, e seja  $\mathcal{B}(\mathcal{N}) \times \mathcal{B}(\mathcal{M})$  a  $\sigma$ -álgebra gerada por todos os retângulos  $B_{\mathcal{N}} \times B_{\mathcal{M}}$ . Portanto  $(N_t, \Lambda)$  é um mapeamento mensurável do espaço de probabilidade  $(\Omega, \mathcal{G}, \mathbb{P})$  em  $(\mathcal{N} \times \mathcal{M}, \mathcal{B}(\mathcal{N}) \times \mathcal{B}(\mathcal{M}))$  com distribuição dada por:

$$\mathbb{P}(N_t \in B_{\mathcal{N}}, \Pi \in B_{\mathcal{M}}) = \int_{B_{\mathcal{M}}} \Pi_{\zeta}(B_{\mathcal{N}}) \Pi(d\zeta) \quad (4.4)$$

Para todo  $B_{\mathcal{N}} \in \mathcal{B}(\mathcal{N})$  e  $B_{\mathcal{M}} \in \mathcal{B}(\mathcal{M})$ .

Esta construção do Processo de Poisson Duplamente Estocástico será conveniente para provar o Teorema 3. Pois, ao assumir que  $N_t$  e  $\Lambda$  estão definidos no mesmo espaço de probabilidade é possível considerar que primeiro se obtém uma realização  $\zeta$  de  $\Lambda$ , e então na sequência é gerado o processo  $N_t$  com medida de intensidade  $\zeta$ .

## 4.2 Resultados e Teoremas

Considere o espaço de probabilidade filtrado  $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ , onde estão definidos o movimento Browniano<sup>18</sup>,  $W_t(\zeta)$ , e o processo Pontual  $N_t$ .

Tome duas EDE's:

$$dY_t = \mu_1(X_t, t)dt + \sigma_1(X_t)dW_t(\zeta) \quad (4.5)$$

$$dX_t = \mu_2(Y_t, t)dt + \sigma_2(Y_t)dW_t(\zeta) \quad (4.6)$$

onde:  $\mu_i(\cdot, t)$  e  $\sigma_i(\cdot)$  para  $i = 1, 2$  satisfazem as condições de regularidade tais que as equações acima possuem uma única solução forte<sup>19</sup>.

Sejam as intensidade  $\lambda_t^X = f(X, t)$  e  $\lambda_t^Y = g(Y, t)$  para as funções mensuráveis  $f : x \mapsto x$  e  $g : y \mapsto y$ , onde  $X_t$  e  $Y_t$  resolvem as equações (4.5) e (4.6). Associadas à estas medidas são construídos dois processos de Poisson Duplamente Estocásticos independentes,  $N_t^X$  e  $N_t^Y$ , com intensidades  $\lambda_t^X$  e  $\lambda_t^Y$ , respectivamente.

**Teorema 3** *Se  $\lambda_t^X \geq \lambda_t^Y$ , então  $N_t^X \geq N_t^Y$  quase-certamente, com intensidades  $\lambda_t^X$  e  $\lambda_t^Y$ , respectivamente.*

### Prova do Teorema 3

Para provar o resultado acima será utilizada a técnica de acoplamento, para tanto tome  $B_{\mathcal{M}} \in \mathcal{B}(\mathcal{M})$  arbitrário, conforme a construção empregada para se obter a equação (4.4).

Na sequência defina  $\delta = \lambda_t^X - \lambda_t^Y \geq 0$ . Defina também o par  $(N_t^{X,1}, N_t^{Y,1})$  onde  $(N_t^{X,1} = N_t^{Y,1} + N_t^{\delta})$  com intensidades  $\lambda_t^Y$  e  $\lambda_t^{\delta}$ , respectivamente. Então:

<sup>18</sup>A notação  $W_t(\zeta)$  é apenas para enfatizar a dependência do Movimento Browniano à realização do evento  $\zeta$

<sup>19</sup>Definição pode ser encontrada na nota 4.

$$N_t^{Y,1} \stackrel{\mathcal{D}}{=} N_t^Y \quad \text{e} \quad N_t^{X,1} \stackrel{\mathcal{D}}{=} N_t^X \quad (4.7)$$

A igualdade acima deve-se à propriedade de sobreposição para Processos Duplamente Estocásticos.

Adicionalmente, tem-se que para uma função mensurável  $G_t$ :

$$\mathbb{P}^Y(N_t^Y \geq G_t, \forall t \leq u \leq T) = \mathbb{P}^{Y,1}(N_t^{Y,1} \geq G_t, \forall t \leq u \leq T) \quad (4.8)$$

$$\leq \mathbb{P}^{X,1}(N_t^{X,1} \geq G_t, \forall t \leq u \leq T) \quad (4.9)$$

$$= \mathbb{P}^X(N_t^X \geq G_t, \forall t \leq u \leq T) \quad (4.10)$$

Uma vez que  $B_{\mathcal{M}} \in \mathcal{B}(\mathcal{M})$  é arbitrário, o resultado é obtido. ■

**Corolário 1** *Sejam  $\lambda_t^{1,n}$  e  $\lambda_t^{2,n}$  duas seqüências monótonas que convergem para  $\lambda_t^{1,N}$  e  $\lambda_t^{2,N}$ , respectivamente. Se  $\lambda_t^{1,n} \geq \lambda_t^{2,n}$  para  $\forall n \leq N$ , então  $N_t^{1,n} \geq N_t^{2,n}$  para  $\forall n \leq N$ , com intensidades  $\lambda_t^{1,n}$  e  $\lambda_t^{2,n}$ , respectivamente.*

**Prova do Corolário 1** Tome  $n$  arbitrário, tal que  $\lambda_t^{1,n} \geq \lambda_t^{2,n}$  então, pelo Teorema 3 tem-se que  $N_t^{1,n} \geq N_t^{2,n}$  para  $n$  arbitrário. Adicionalmente, uma vez que os processos  $\lambda_t^{i,N}$  ( $i = 1, 2$ ) convergem monotonamente, sabe-se, pelo Resultado 8, que se  $\lambda_t^{1,n} \xrightarrow{v} \lambda^{1,N}$  e  $\lambda^{1,N}$  é localmente finita, então  $N_t \xrightarrow{v} N_t^1$ , onde  $N_t^1$  é um processo de Poisson com intensidade  $\lambda^{1,N}$ . Repetindo o mesmo argumento para  $\lambda_t^{2,n}$  ter-se-á que  $N_t^2$  é um processo de Poisson com intensidade  $\lambda^{2,N}$ . Pelo fato que  $\lambda^{1,N} \geq \lambda^{2,N}$ , utilizando novamente o Teorema 3 o resultado é obtido. ■

Na seqüência são apresentadas as condições suficientes para que se verifique  $\lambda_t^X \geq \lambda_t^Y$  quase-certamente.

**Resultado 10 (Ferreyra e Sundar (2006))** *Para o caso de difusões unidimensionais,  $Y_t$  e  $X_t$ , que satisfazem as condições:*

**H1.**  $\mu_1$  e  $\mu_2$  são funções contínuas em  $\mathcal{D}$ , mensuráveis em  $\mathcal{D} \times \mathbb{R}_+$  tais que exista uma constante  $K$  tal que:

$$|\mu_1(x, t) - \mu_1(y, t)| + |\mu_2(x, t) - \mu_2(y, t)| \leq K|x - y| \quad (4.11)$$

e

$$|\mu_1(x, t)| + |\mu_2(x, t)| \leq K\sqrt{(1 + x^2)} \quad (4.12)$$

Para todo  $t \geq 0$ ;  $x, y \in \mathcal{D}$

**H2.** As funções  $\sigma_1$  e  $\sigma_2$  são estritamente positivas,  $C^1$ , com primeira derivada limitada, e  $\bar{b}(t, a) \leq \bar{c}(t, b)$ , para todo  $t \geq 0$ ;  $a < b \in \mathcal{D}$

onde

$$\bar{b} = \frac{\mu_1}{\sigma_1} - \frac{Z^2 \sigma_1'}{2} \quad e \quad \bar{c} = \frac{\mu_2}{\sigma_2} - \frac{Z^2 \sigma_2'}{2}$$

Para  $Z_t$  um processo contínuo unidimensional e mensurável.

**H3.**

$$\infty > \int_{\xi}^x \frac{du}{\sigma_1(u)} \geq \int_{\eta}^x \frac{du}{\sigma_2(u)} > -\infty \quad (4.13)$$

Para todo  $x \in \mathcal{D}$

então as soluções  $Y_t, X_t$  das equações (4.5) e (4.6):

$$Y_t = \xi + \int_0^t Z_s \mu_1(Y_s, s) ds + \int_0^t \sigma_1(Y_s) dW_s(\zeta) \quad (4.14)$$

$$X_t = \eta + \int_0^t Z_s \mu_2(X_s, s) ds + \int_0^t \sigma_2(X_s) dW_s(\zeta) \quad (4.15)$$

satisfazem:

$$\mathbb{P}(Y_t \leq X_t : t \geq 0) = 1 \quad (4.16)$$

■

O Resultado 10 estabelece as condições para a dominação estocástica de difusões gerais. No caso específico das difusões afim unidimensionais definidas no Capítulo 2, pode-se enunciar a seguinte Proposição:

**Proposição 15** Para difusões afim unidimensionais na forma:

$$Y_t = \xi + \int_0^t \mathcal{K}_1(\Theta_1 - Y_s) ds + \int_0^t \Sigma_1 \sqrt{a_1 + b_1 Y_s} dW_s \quad (4.17)$$

$$X_t = \eta + \int_0^t \mathcal{K}_2(\Theta_2 - X_s) ds + \int_0^t \Sigma_2 \sqrt{a_2 + b_2 X_s} dW_s \quad (4.18)$$

as condições suficientes para:

$$\mathbb{P}(Y_t \leq X_t : t \geq 0) = 1 \quad (4.19)$$

são:

**C1.**  $0 < b_1 < b_2$

**C2.**  $x_0 = \xi = \eta = \frac{a_1 - a_2}{b_2 - b_1}$

**C3.**  $x_0 > A$ , onde  $A = (\frac{-b_1}{a_1}, \frac{-b_2}{a_2})^+$

**C4.**  $\mathcal{K}_2 \Theta_2 > \mathcal{K}_1 \Theta_1$

**C5.**  $b_1 \Sigma_1^2 > b_2 \Sigma_2^2$

**C6.**  $\mathcal{K}_1 c > \mathcal{K}_2 d$  para  $c < d \in \mathcal{D}$

**C7.**  $\Sigma_1(a_1 - 1) > \Sigma_2(a_2 - 1)$

### Prova da Proposição 15

Se todas as condições forem atendidas simultaneamente, tem-se as condições suficientes para a dominação estocástica,  $\mathbb{P}(Y_t \leq X_t : t \geq 0) = 1$ . As condições C1, C2 e C3 satisfazem H3 no intervalo  $(A, \infty)$ . As condições C4, C5, C6 e C7 são necessárias para atender H2. A prova da existência e unicidade da solução forte para (4.17) e (4.18) foram apresentadas na seção 2.3.1.

■





## Capítulo 5

# Estimação de Processos de Cox com intensidade afim

### 5.1 Introdução

A estimação de processos pontuais é um tópico de pesquisa bastante ativo, sendo provavelmente a econometria financeira a área que mais explorou este assunto<sup>20</sup>. O principal objetivo na econometria financeira dos processos pontuais é, a partir do fluxo de chegada de ofertas de compra e venda em um ambiente de negociação eletrônica, modelar a dinâmica dos negócios, tanto no que diz respeito ao volume ou preço negociado.

Embora existam diversos modelos estatísticos para modelar processos pontuais o método mais comum envolve estimar o tempo de espera entre eventos sucessivos. Neste sentido Engle e Russell (1998) e Engle (2000), com a classe de modelos *autoregressive conditional duration* (ACD), foram os precursores ao modelar a duração entre eventos financeiros capaz de incorporar alguns dos principais fatos estilizados<sup>21</sup> presentes nas séries de ordens enviadas e negócios realizados. De fato, como se observa, as propriedades presentes na dinâmica da duração de eventos financeiros são muito similares às encontradas no estudo da volatilidade diária de um ativo financeiro.

Os modelos ACD, a despeito de sua simplicidade e possibilidade de emprestar os resultados desenvolvidos para os modelos da família ARCH/GARCH, o que os tornaram muito atrativos do ponto de vista prático, possuem limitações que acabaram por estimular soluções alternativas à análise da duração dos eventos financeiros.

Uma maneira alternativa à duração, no estudo de eventos financeiros, é modelar de maneira direta a intensidade do processo pontual e desta forma ser capaz de incorporar a ocorrência dos eventos a qualquer instante de tempo, ou seja, construir o modelo em tempo contínuo, bem como atribuir um comportamento estocástico à intensidade, ausente na família de modelos ACD's, se

---

<sup>20</sup>Grande parte da disseminação de novas técnicas pode ser atribuída à disponibilidade crescente de dados em alta-frequência, os quais, atualmente, estão disponíveis conforme as ofertas são submetidas, em sua maioria no intervalo de milissegundos.

<sup>21</sup>Os principais fatos estilizados dos dados de negociação de ativos financeiros são: espaçamento irregular das observações no tempo, o movimento de preços é discreto, sazonalidade *intraday* e agrupamento de pontos no tempo (*clustered over time*) implicando que a duração possuirá autocorrelação positiva e elevada persistência.

aproximando conceitualmente da definição do Processo de Poisson Duplamente Estocástico desenvolvido no Capítulo 2.

A contrapartida a tornar a intensidade estocástica é que esta passa a ser um processo estocástico não observável, desta forma os Processos de Poisson Duplamente Estocásticos pertencem a classe de modelos com variáveis latentes dinâmicos uma vez que apenas as realizações  $(N_t)_{t \geq 0}$  são observáveis enquanto que a intensidade do processo,  $(\lambda_t)_{t \geq 0}$ , é latente.

De maneira geral, a função de verossimilhança do processo  $(N_t)_{t \geq 0}$ , dado um vetor de parâmetros  $\theta$  e  $(\lambda_t)_{t \geq 0}$ , pode ser descrita por:

$$\mathcal{L}(N|\theta) = \prod_{t=1}^T f(N_t|\lambda_t, \theta) f(\lambda_t, \theta) \quad (5.1)$$

## 5.2 Revisão dos modelos de estimação

Nos casos onde a equação de medida é uma função linear do processo latente,  $(\lambda_t)_{t \geq 0}$ , e a distribuição dos erros é Gaussiana, o filtro de Kalman pode ser utilizado em conjunto com o Estimador de Quase-Máxima Verossimilhança (QME) para avaliar (5.1).

Alternativamente, os filtros de partícula são, em geral, utilizados em modelos de espaço de estado não linear com o objetivo de atualizar, de maneira sequencial, a distribuição da variável latente a partir do conhecimento de variáveis observáveis. Em sua grande maioria, os filtros de partícula, diferente do filtro de Kalman, não possuem solução analítica em forma fechada e nestes casos as soluções são aproximadas, obtidas por meio de processos de amostragens, denominadas partículas.

Uma alternativa, assim como Harvey e Fernandes (1989), é utilizar uma distribuição conjugada, Gama, para a intensidade e assim obter a distribuição binomial negativa como solução de (5.1). A partir das propriedades derivadas para o PPDE apresentadas no Capítulo 2, especificamente na Proposição 14, a estimação do modelo proposto em Harvey e Fernandes (1989) é um caso particular de  $\mathcal{L}(\theta)$  para o processo  $(N_t)_{t \geq 0}$  quando  $t \rightarrow \infty$ .

Uma possível alternativa seria a adoção de técnicas<sup>22</sup> Bayesianas para a solução do problema em (5.1), neste caso o algoritmo MCMC seria a solução para extrair a variável latente  $\lambda_t$ , bem como a distribuição a posteriori do vetor de parâmetros  $\theta$ .

A alternativa escolhida nesta tese para estimar o vetor  $\theta$  será o filtro de Kalman combinado com o estimador de Quase-Máxima Verossimilhança. A razão para esta escolha advém do fato de

<sup>22</sup>Neste sentido, o trabalho de Johannes e Polson (2009) é uma referência bastante completa sobre a estimação de modelos em tempo contínuo para séries financeiras.

que a equação de medida é log-linear em relação a intensidade, fazendo uso portanto das condições de otimalidade do filtro de Kalman. Uma segunda razão decorre do fato que o filtro de Kalman é recursivo sendo computacionalmente eficiente em problemas com elevado número de observações, que será o caso, conforme descrito na seção 5.5.

Nas seções seguintes os passos para a estimação do processo de Cox segundo a metodologia proposta são apresentados. Assim, na seção 5.3 apresenta-se o filtro de Kalman. Na seção 5.4, ao combinar os resultados precedentes ter-se-á o procedimento completo de estimação para o vetor  $\theta$  do processo  $(N_t)_{t \geq 0}$ . Finalmente, na seção 5.6 recorre-se a um exercício de simulação de Monte Carlo para avaliar a qualidade do método de estimação ora proposto.

### 5.3 Filtro de Kalman

A técnica de Filtro de Kalman foi popularizada em finanças, especialmente na literatura de estruturas a termo de taxas de juros, após o trabalho de Duan e Simonato (1999). O uso desta estratégia é especialmente útil em situações como a desenvolvida nesta tese, onde a variável de estado não é observável. A utilização do Filtro de Kalman para extrair o componente não-observável tem como fundamento o Teorema 1 que determina a relação entre o componente observável (a ocorrência dos processos pontuais) e a variável latente. Com o intuito de formalizar a construção do filtro de Kalman faz-se necessário a definição da equação de medida que irá determinar o relacionamento entre os componentes observáveis e não-observáveis, bem como da equação de transição que descreve a dinâmica da variável latente. Desta forma, as equações de medida e transição representam a formulação do problema em uma forma de espaço de estado. Por sua vez, o filtro de Kalman faz uso desta formulação de espaço de estado para recursivamente inferir os valores da variável latente, condicionado nas realizações do processo pontual de interesse. Finalmente, tendo como ferramenta a inferência recursiva fornecida pelo filtro de Kalman e do método de Quase-Máxima Verossimilhança pode-se estimar os parâmetros que descrevem a dinâmica da intensidade do processo pontual,  $(\lambda_t)_{t \geq 0}$ .

O uso do filtro de Kalman necessita que o processo de interesse seja reescrito na forma de espaço de estado, para cada instante de tempo  $k$ . Assim, adotam-se as seguintes notações  $F_k$ , a matriz de transição;  $H_k$ , a matriz de medida;  $Q_k$ , a matriz de variância e covariância do erro;  $R_k$ , a matriz de covariância para o erro da mensuração.

Desta forma, o filtro de Kalman necessita que a dinâmica da variável de estado evolua segundo:

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{w}_k$$

onde:

$w_k$  é erro construído a partir de uma distribuição Normal multivariada com matriz de covariância  $Q_k$ .

$$\mathbf{w}_k \sim N(\boldsymbol{\mu}_k, \mathbf{Q}_k)$$

Em um instante de tempo  $k$ , o mapeamento de processo latente,  $x_k$ , em uma observação,  $z_k$ , é feito por:

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k$$

Onde  $v_k$  é o erro de mensuração definido como um ruído branco Gaussiano com matriz de covariância diagonal  $R_k$ .

$$\mathbf{v}_k \sim N(0, \mathbf{R}_k)$$

O fato do filtro de Kalman ser um algoritmo recursivo é particularmente útil quando se trata de modelos onde a dimensão é elevada, uma vez que apenas a última inferência da variável de estado e a observação corrente do sistema são necessários para estimar o estado atual do processo. Na derivação dos resultados a seguir adota-se a seguinte notação  $\hat{\mathbf{x}}_{n|m}$  representa a estimativa de  $\mathbf{x}$  no instante  $n$  dada as observações até o instante  $m$ , inclusive.

Formalmente o filtro é construído com o uso de duas variáveis:

- $\hat{\mathbf{x}}_{k|k}$ , a estimativa *a posteriori* do estado no instante  $k$ , dadas as observações até  $k$ , inclusive;
- $\mathbf{P}_{k|k}$ , a matriz de variância e covariância *a posteriori*.

Teoricamente pode-se escrever o filtro de Kalman em apenas uma única etapa, entretanto, tradicionalmente, este é descrito em duas fases distintas: previsão e atualização. A fase de previsão utiliza a estimativa do estado no instante anterior para prever o estado atual, sendo frequentemente esta previsão chamada de *a priori* pois não incorpora as observações do instante atual.

Por sua vez, na fase de atualização a estimativa *a priori* é combinada com a última informação disponível, dando origem a previsão *posteriori* do estado.

### Previsão

As previsões *a priori* são dadas por:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{u}_k$$

A matriz de variância e covariância *a priori*:

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{Q}_k$$

**Atualização**

Erro de mensuração:

$$\tilde{\mathbf{y}}_k = \mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}$$

Covariância do erro de mensuração:

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k$$

Ganho ótimo de Kalman:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1}$$

Previsões *a posteriori* são dadas por:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}_k$$

A matriz de variância e covariância *a posteriori*:

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1}$$

É importante ressaltar que as fórmulas para a atualização do sistema e da matriz de covariância são válidas apenas quando são realizadas através do ganho ótimo de Kalman. Adicionalmente, assumindo que o modelo proposto reflita o processo gerador de dados e, os valores iniciais  $\hat{\mathbf{x}}_{0|0}$  e  $\mathbf{P}_{0|0}$  refletem a distribuição inicial dos estados, então as seguintes relações de invariância são preservadas:

- $E[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}] = E[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}] = 0$
- $E[\tilde{\mathbf{y}}_k] = 0$

Onde  $E[\xi]$  é o valor esperado de  $\xi$ , e para a matriz de variância e covariância têm-se:

- $\mathbf{P}_{k|k} = \text{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})$
- $\mathbf{P}_{k|k-1} = \text{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})$
- $\mathbf{S}_k = \text{cov}(\tilde{\mathbf{y}}_k)$

**5.3.1 Derivação da matriz de covariância *a posteriori***

Partindo da matriz de variância e covariância invariante para o erro, definida acima,  $\mathbf{P}_{k|k}$ , tem-se:

$$\mathbf{P}_{k|k} = \text{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})$$

Ao substituir as definições de  $\hat{\mathbf{x}}_{k|k}$

$$\mathbf{P}_{k|k} = \text{cov}(\mathbf{x}_k - (\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}_k))$$

Igualmente para  $\tilde{\mathbf{y}}_k$

$$\mathbf{P}_{k|k} = \text{cov}(\mathbf{x}_k - (\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1})))$$

e por último de  $\mathbf{z}_k$

$$\mathbf{P}_{k|k} = \text{cov}(\mathbf{x}_k - (\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1})))$$

tem-se ao agrupar o vetor de erros:

$$\mathbf{P}_{k|k} = \text{cov}((I - \mathbf{K}_k \mathbf{H}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) - \mathbf{K}_k \mathbf{v}_k)$$

Uma vez que o erro de mensuração  $v_k$  não é correlacionado com outros termos, tem-se:

$$\mathbf{P}_{k|k} = \text{cov}((I - \mathbf{K}_k \mathbf{H}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})) + \text{cov}(\mathbf{K}_k \mathbf{v}_k)$$

Com uso das propriedades da matriz de covariância:

$$\mathbf{P}_{k|k} = (I - \mathbf{K}_k \mathbf{H}_k) \text{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) (I - \mathbf{K}_k \mathbf{H}_k)^\top + \mathbf{K}_k \text{cov}(\mathbf{v}_k) \mathbf{K}_k^\top$$

Finalmente, tomando a invariância de  $P_{k|k-1}$  e a definição de  $R_k$  se obtém:

$$\mathbf{P}_{k|k} = (I - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} (I - \mathbf{K}_k \mathbf{H}_k)^\top + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^\top \quad (5.2)$$

Esta fórmula, também conhecida como forma de Joseph (introduzida em Bucy e Joseph (1968)) para atualização da matriz de covariância, é válida para qualquer valor de  $K_k$ . Caso se utilize o valor ótimo de  $K_k$ , pode-se simplificar, conforme descrito na seção 5.3.2.

### 5.3.2 Derivação do ganho de Kalman

O filtro de Kalman pertence a classe dos estimadores que minimizam o erro quadrático médio da estimativa *a posteriori* do estado do sistema.

Desta forma, pode-se formalizar o objetivo do filtro de Kalman como obter a solução do seguinte problema de otimização:

$$\text{argmin} \quad \mathbb{E}[|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}|^2]$$

Adicionalmente, a solução deste problema pode ser vista como equivalente ao problema de minimizar o traço da matriz de variância e covariância *a posteriori*  $\mathbf{P}_{k|k}$ . Desta forma, ao expandir

os termos da equação acima e agrupar os termos, obtém-se:

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{K}_k^T + \mathbf{K}_k (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k) \mathbf{K}_k^T \quad (5.3)$$

$$= \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{K}_k^T + \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T \quad (5.4)$$

O traço é minimizado quando:

$$\frac{\partial \text{tr}(\mathbf{P}_{k|k})}{\partial \mathbf{K}_k} = -2(\mathbf{H}_k \mathbf{P}_{k|k-1})^T + 2\mathbf{K}_k \mathbf{S}_k = 0$$

Resolvendo para  $\mathbf{K}_k$  se obtém o ganho ótimo de Kalman:

$$\mathbf{K}_k \mathbf{S}_k = (\mathbf{H}_k \mathbf{P}_{k|k-1})^T = \mathbf{P}_{k|k-1} \mathbf{H}_k^T$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1}$$

### 5.3.3 Simplificação da fórmula *a posteriori* da covariância

Quando se faz uso do ganho ótimo de Kalman, a fórmula da covariância do erro (5.2) pode ser simplificada. Para tanto, multiplicam-se os dois lados do ganho de Kalman por  $\mathbf{S}_k \mathbf{K}_k^T$ :

$$\mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{K}_k^T$$

Tomando a fórmula (5.2):

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{K}_k^T + \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T$$

verifica-se que o último termo se cancela, produzindo:

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k|k-1} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1}$$

O grande apelo desta fórmula é seu baixo custo computacional, característica esta vital em problemas de grande dimensão, como será colocado na seção 5.5.

Uma vez determinada a forma como a variável de estado será obtida o próximo passo é a estimação dos parâmetros que descrevem a dinâmica da intensidade do processo. De acordo com Harvey (1989) o filtro de Kalman oferece todas as informações necessárias para calcular a função de Quase-Máxima Verossimilhança, assim:

$$\log f(\mathbf{y}_t | \mathbf{x}_t; \boldsymbol{\theta}) = -\frac{1}{2} \log 2\pi(T - K) - \frac{1}{2} \sum_{t=K+1}^T \log |S_t| - \frac{1}{2} \sum_{t=K+1}^T \tilde{\mathbf{y}}_t' S_t^{-1} \tilde{\mathbf{y}}_t \quad (5.5)$$

Onde  $T$  é o tamanho da amostra;  $K$  é a dimensão da equação que descreve a intensidade do processo.

#### 5.4 Filtro de Kalman para processos de Cox com intensidade de Feller

Com base nos resultados introduzidos nas seções anteriores sobre o filtro de Kalman, o objetivo desta seção é representar o processo de Cox  $(N_t)_{t \geq 0}$ , com intensidade de Feller  $(\lambda_t)_{t \geq 0}$ , em uma forma de espaço de estado que permitirá aplicar o filtro de Kalman conjugado com o método de Quase-Máxima Verossimilhança na estimação dos parâmetros.

Na seção 2.3.1 a equação diferencial estocástica que governa a intensidade foi reescrita segundo a representação canônica introduzida por Dai e Singleton (2000). Esta representação foi útil para provar a existência e unicidade da EDE em um contexto  $d$ -dimensional. Entretanto, de acordo com Dai e Singleton (2000), um segundo ganho de utilizar a representação canônica  $\mathbb{A}_m(d)$  é que esta também é *maximal* no sentido que se impõem restrições mínimas na etapa de identificação dos parâmetros a serem estimados.

Para construir a equação de medida, parte-se dos resultados obtidos na seção 3.4:

$$P(N_{t-T} = k) = \frac{1}{k!} \mathbb{E} \left[ (\Lambda_s)^k e^{-\Lambda_s} \right] = \frac{1}{k!} G_{\Lambda_s}^k(1) \quad (5.6)$$

$$G_{\Lambda_s}(1) = \mathbb{E} \left( e^{-\int_t^T \lambda_s ds} \right) = e^{\alpha(t,T) - \beta(t,T)\lambda_s} \quad (5.7)$$

e

$$\alpha(t, T) = \frac{2\kappa\theta}{\sigma^2} \ln \left( \frac{2\gamma (e^{(\gamma+\kappa)/2})}{(\gamma + \kappa) (e^{-\gamma(T-t)} - 1) + 2\gamma} \right) \quad (5.8)$$

$$\beta(t, T) = \left[ \frac{2\mu (e^{-\gamma(T-t)} - 1)}{(\gamma + \kappa) (e^{-\gamma(T-t)} - 1) + 2\gamma} \right] \quad (5.9)$$

$$\text{onde } \gamma = \sqrt{\kappa^2 + 2\sigma^2} \quad (5.10)$$

Para apresentar o estimador dos parâmetros do processo latente, considere a probabilidade  $P(N_{t,T} = 0)$  do processo de Cox com intensidade de Feller no intervalo  $(t - T)$ . Desta forma, as equações acima podem ser reescritas como:

$$P(N_{t,T} = 0) = \alpha(t, T) e^{-\beta(t,T)\lambda_s} \quad (5.11)$$

sendo que  $\alpha(t, T)$  e  $\beta(t, T)$  não se alteram.

É possível linearizar (5.11) e obter:



$$\ln P(N_{t,T} = 0) = \ln \alpha(t, T) - \beta(t, T)\lambda_s \quad (5.12)$$

Que passa a ser uma equação log-linear em  $\lambda_t$  e portanto a equação de medida pode ser definida como:

$$\ln P(N_{t,T} = 0) = \ln \alpha(t, T) - \beta(t, T)\lambda_s + \chi_s \quad (5.13)$$

A equação (5.13) inclui um termo errático  $\chi_s$  enquanto que em (5.11) este termo está ausente. Este fato pode parecer inconsistente com o processo de Cox, entretanto a inclusão do termo errático (5.13) justifica-se em um contexto de estimação, pois assume-se que a intensidade do processo pontual é um processo de Feller, entretanto caso o verdadeiro processo que governa a intensidade não seja o de Feller, a equação (5.11) estará mal especificada e os resultados serão sistematicamente diferente dos observados. Alternativamente, caso o modelo esteja corretamente especificado o termo errático não será autocorrelacionado e terá média zero.

Adicionalmente, sabe-se que a densidade de transição do processo de Feller,  $p(t, x, y) = \mathbb{P}(X_t \in dy | X_{t-1} = dx)$ , é a distribuição qui-quadrado não centralizada. Entretanto, no processo de estimação do componente não-observável por meio do filtro de Kalman, conjugado com o estimador de Quase-Máxima Verossimilhança se adotará a substituição da densidade de transição exata pela densidade normal:

$$\lambda_s | \lambda_{s-1} \sim N(\mu_s, Q_s)$$

onde  $\mu_s$  e  $Q_s$  são definidos de tal forma que os momentos da distribuição aproximada sejam exatamente os mesmos da distribuição exata. Assim, tem-se que:

$$\mu_s = \theta[1 - \exp(-\kappa)] + \exp(-\kappa)\lambda_{s-1} \quad (5.14)$$

e  $Q_s$  uma matriz diagonal com elementos:

$$Q_s = \sigma^2 \frac{1 - \exp(-\kappa)}{\theta} \left( \frac{\theta}{2} [1 - \exp(-\kappa)] + \exp(-\kappa)\lambda_{s-1} \right) \quad (5.15)$$

## 5.5 Estimação do PPDE para dados de alta-frequência

A partir da formulação descrita na seção anterior, deseja-se modelar o processo de chegada de ofertas em um ambiente de negociação eletrônica como um PPDE e para tanto devem-se estimar<sup>23</sup> os parâmetros do processo de Cox  $(N_t)_{t \geq 0}$ , com intensidade de Feller,  $(\lambda_t)_{t \geq 0}$ .

---

<sup>23</sup>As rotinas para estimar os parâmetros foram desenvolvidas pelo autor em *MATLAB*<sup>®</sup> e encontram-se no apêndice A

### 5.5.1 Descrição da amostra

A amostra é formada por todas as ordens submetidas ao livro de ofertas do contrato futuro de Reais/Dólar (código de negociação FUT DOL) negociados na BM&FBOVESPA no mês de outubro de 2009. Por se tratar de um sistema de negociação totalmente eletrônico o momento exato (*timestamp*) em que as ordens são submetidas são armazenados com precisão de milissegundos. Adicionalmente, além do momento exato em que a ordem é submetida a amostra possui a quantidade ofertada, preço e lado (oferta de compra ou oferta de venda). A negociação deste contrato ocorre de maneira contínua das 09:00 as 18:00.

A variável que se deseja modelar (variável observável) é a frequência de ordens de venda<sup>24</sup> que chegam ao livro de ofertas em um determinado intervalo de 1 minuto ao longo do mês de outubro. Como colocado, as ordens enviadas possuem *timestamp* em milissegundos, porém como estas ordens trafegam em diferentes sistemas e redes de comunicação o intervalo entre duas ordens consecutivas não é inferior a 10 milissegundos. Sendo assim, seria possível, em um intervalo de 1 segundo, que 100 ordens de venda ingressassem no livro de oferta, logo em um intervalo de 60 segundos seria possível que 6.000 ofertas fossem enviadas. Desta forma, a variável de interesse será contruída como o número<sup>25</sup> de ofertas enviadas em relação ao total de ofertas que poderiam ingressar no sistema, assim para cada um dos dias da amostra:

$$y_{(t-s=60 \text{ seg})} = \frac{\#(\text{ofertas de venda})}{6.000} \quad (5.16)$$

Uma vez calculado o número de ordens que chegam por minuto para cada um dos dias da amostra, calcula-se a média ao longo dos dias e se obtém o número médio de ordens que são submetidas ao ambiente de negociação em intervalos de 60 segundos, totalizando portanto 535 observações.

### 5.5.2 Resultados da Estimação

Na sequência os parâmetros do Processo de Poisson Duplamente Estocástico quando a intensidade é um processo de Feller são estimados. Os valores estimados  $\hat{\boldsymbol{\theta}} = (\theta, \kappa, \sigma)$  encontram-se na tabela 5.1, os quais, segundo Hamilton (1994), têm o seguinte comportamento:

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0) \stackrel{\mathcal{D}}{=} N\left(\mathbf{0}, [\mathcal{I}_{2D} \mathcal{I}_{OP}^{-1} \mathcal{I}_{2D}]^{-1}\right) \quad (5.17)$$

Onde<sup>26</sup>:

$$\mathcal{I}_{2D} = -\frac{1}{T} \mathbb{E} \left( \sum_{t=1}^T \frac{\partial^2 \log f(\mathbf{y}_t | \mathbf{x}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \bigg|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0} \right) \quad (5.18)$$

<sup>24</sup>Poderia, sem prejuízo dos resultados, escolher as ofertas de compra ou outra data.

<sup>25</sup>Onde # na equação (5.16) é a medida de contagem no intervalo ( $t - s = 60 \text{ seg}$ )

<sup>26</sup> $\stackrel{\mathcal{D}}{=}$  denota igualdade em distribuição.

$$\mathcal{I}_{OP} = \text{plim} \frac{1}{T} \left( \sum_{t=1}^T \frac{\partial \log f(\mathbf{y}_t | \mathbf{x}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \log f(\mathbf{y}_t | \mathbf{x}_t; \boldsymbol{\theta})'}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right) \quad (5.19)$$

As matrizes de informação  $\mathcal{I}_{2D}$  e  $\mathcal{I}_{OP}$  foram calculadas numericamente a partir da equação (5.5).

	$\theta$	$\kappa$	$\sigma$	Erro Padrão $\chi$
Estimativa	0.065	0.0043	0.00267	0.0010
Erro Padrão	2.6E-05	1.7E-05	7.47E-08	2.68E-09

**Tabela 5.1:** Parâmetros estimados para as ordens de Venda para o contrato futuro de BRL/USD: (ticker DOL FUT - vencimento NOV09)

Como se verifica a partir da tabela acima, todos os parâmetros estimados são estatisticamente significantes a 1%, e a estimativa para o erro de mensuração  $\chi$  é duas vezes menor que a variância do processo,  $\sigma$ , corroborando a qualidade do modelo proposto para descrever o processo de ordens de venda.

A seguir é apresentado de maneira gráfica o ajuste do modelo proposto para estimar a probabilidade de ocorrência de ordens de venda que ingressam no livro de negociação e os valores efetivamente verificados.

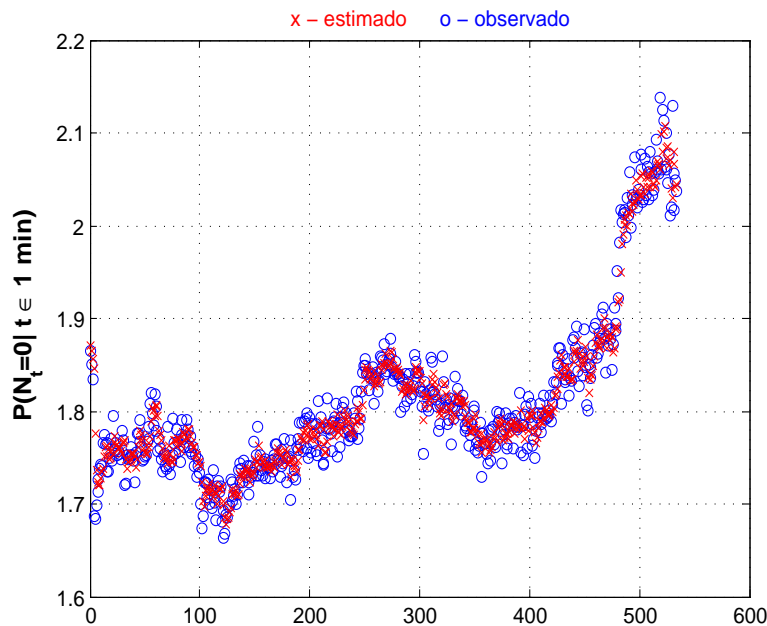


Figura 5.1: Ajuste do modelo em relação aos dados observados: ordens de venda para o contrato futuro de R\$/DOL (ticker DOL FUT - vencimento NOV09)

Percebe-se que o modelo é flexível o suficiente para acompanhar as diferentes mudanças no número de ordens que são submetidas no ambiente de negociação.

Na sequência os resíduos do modelo estimado são analisados para avaliar a hipótese deste ser um ruído branco. Inicialmente, na figura 5.2 mostra-se o correlograma dos resíduos até a vigésima defasagem. Percebe-se que a correlação entre as diferentes defasagens não são significativas a 5% conforme as barras horizontais que denotam o intervalo de confiança para esta estimativa.

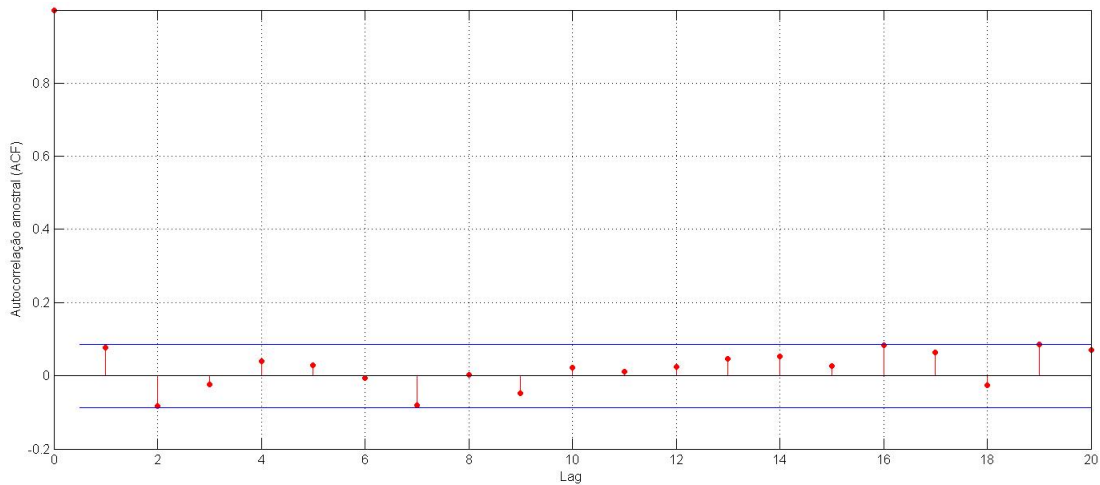


Figura 5.2: Correlograma dos resíduos

Adicionalmente, na figura 5.3 exibem-se a função de autocorrelação parcial e novamente percebe-se que os valores das estimativas encontram-se dentro do intervalo de confiança de 95%, a despeito da segunda defasagem onde a estimativa é igual ao intervalo de confiança.

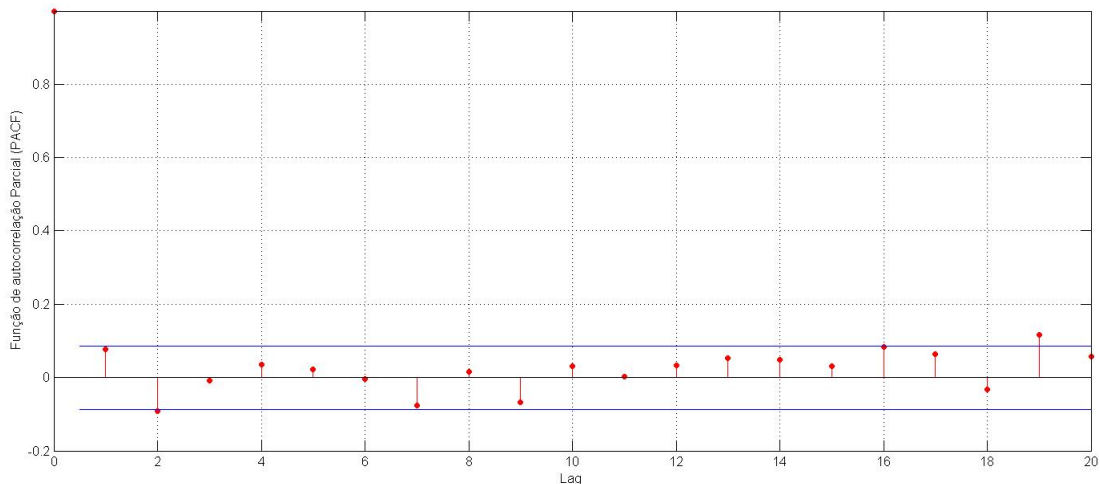


Figura 5.3: Correlograma parcial dos resíduos

Além da análise das funções de autocorrelação (ACF) e autocorrelação Parcial (PACF), realiza-se o teste conjunto de Ljung-Box para avaliar a autocorrelação dos resíduos até a  $l$ -ésima defasagem.

A estatística do teste de Ljung-Box é dada por:

$$Q(L) = N(N + 2) \sum_{k=1}^L \frac{r_k^2}{(N - k)} \quad (5.20)$$

Onde  $N$  é o tamanho da amostra,  $L$  é  $l$ -ésima defasagem e  $r_k^2$  é o quadrado da autocorrelação amostral da  $k$ -ésima defasagem.

A estatística do teste possui distribuição Qui-quadrado com  $L$  graus de liberdade. Assim, o teste para diferentes valores de  $L$  encontra-se na tabela a seguir.

$L$	Q(L)	p-valor
5	8.36	0.13
10	13.47	0.19
15	17	0.31

**Tabela 5.2:** Teste de Ljung-Box para correlação dos resíduos

Da análise da tabela acima verifica-se que existem evidências que os resíduos não são correlacionados até a 15<sup>a</sup> defasagem, corroborando o bom ajuste do modelo estimado.

## 5.6 Simulações

Para avaliar a qualidade do método de estimação que combina o Filtro de Kalman e o método de Quase-Máxima Verossimilhança para inferir os parâmetros do processo de intensidade se recorre ao uso de um conjunto de Simulações de Monte Carlo.

Para tanto, se constroem 200 séries, cada uma com 500 observações, de acordo com parâmetros previamente definidos  $\hat{\theta} = (\theta, \kappa, \sigma)$  e, na sequência se aplica o filtro de Kalman conjugado com o estimador de QML. Ao final será possível comparar a distribuição empírica dos parâmetros estimados conforme sugerido e comparar estes valores com os parâmetros iniciais usados para gerar as séries. A construção da variável de estado,  $(\lambda_t)_{t \geq 0}$ , nas simulações se valerá do fato que a densidade de transição do processo de Feller possui fórmula fechada, assim as realizações de  $(\lambda_t)_{t \geq 0}$ , para  $t - s > 0$ , são construídas de acordo com:

$$\lambda_t = c_s \times f_\lambda(d, l)$$

onde<sup>27</sup>:

$$c_s = \frac{\sigma^2(1 - e^{-\kappa(t-s)})}{4\kappa}$$

<sup>27</sup>Onde  $f_\lambda(d, l)$  é a densidade de probabilidade da distribuição Qui-Quadrado com  $d$  graus de liberdade e parâmetro de deslocamento  $l$

$$d = \frac{4\kappa\theta}{\sigma^2}$$

$$l = \frac{\lambda_s e^{-\kappa(t-s)}}{c_s}$$

Uma vez geradas as realizações da variável de estado  $(\lambda_t)_{t \geq 0}$  para cada instante de tempo  $t$ , a equação (5.11) faz o mapeamento entre a variável simulada e a probabilidade de ocorrência de nenhum evento no intervalo  $t - s$ .

	$\theta$	$\kappa$	$\sigma$
Original	0.04	0.2	0.05
Média das estimações	0.04	0.3698	0.0448
EQM	1.5303e-005	0.37	0.0023
SE	0.0039	0.52	0.0001

**Tabela 5.3:** Resultados da Simulação de Monte Carlo para estimação dos parâmetros

Nas figuras 5.4, 5.5 e 5.6, a barra vertical denota o parâmetro gerador dos dados e, os eixos, os valores estimados e verdadeiro para os parâmetros do modelo.

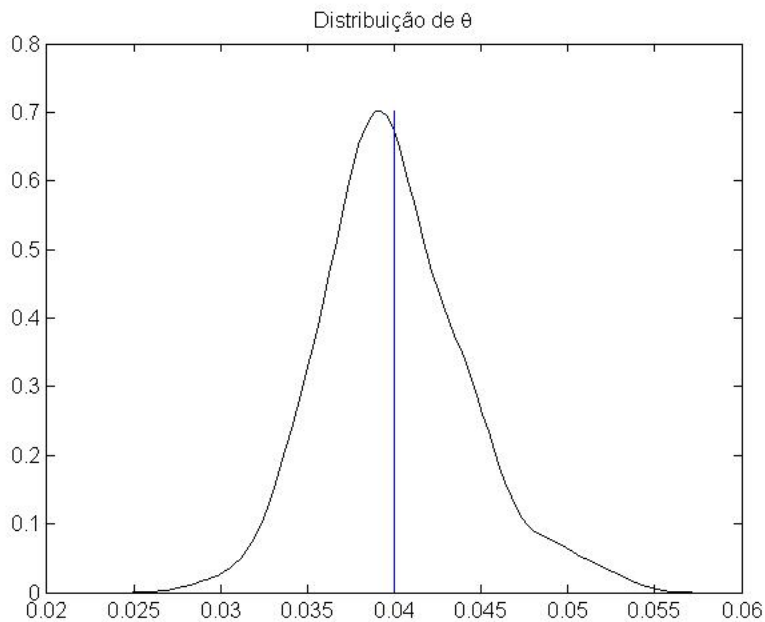
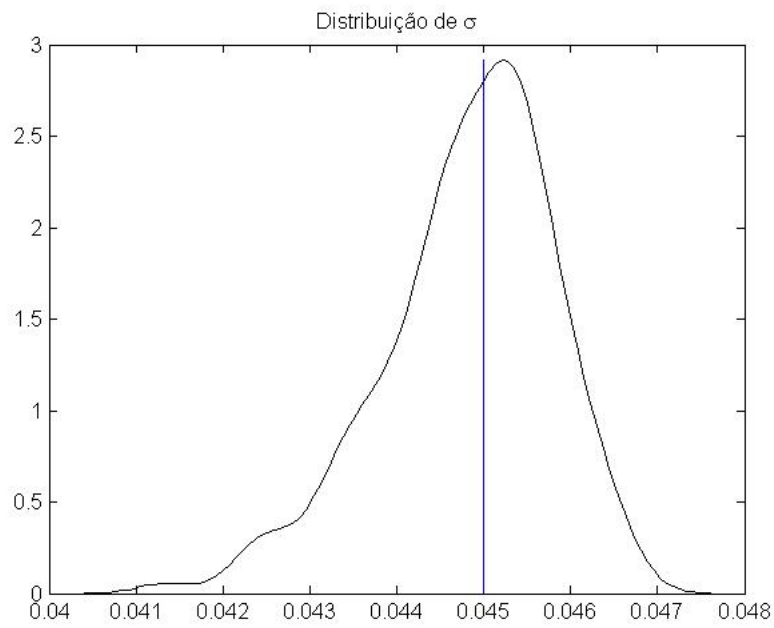
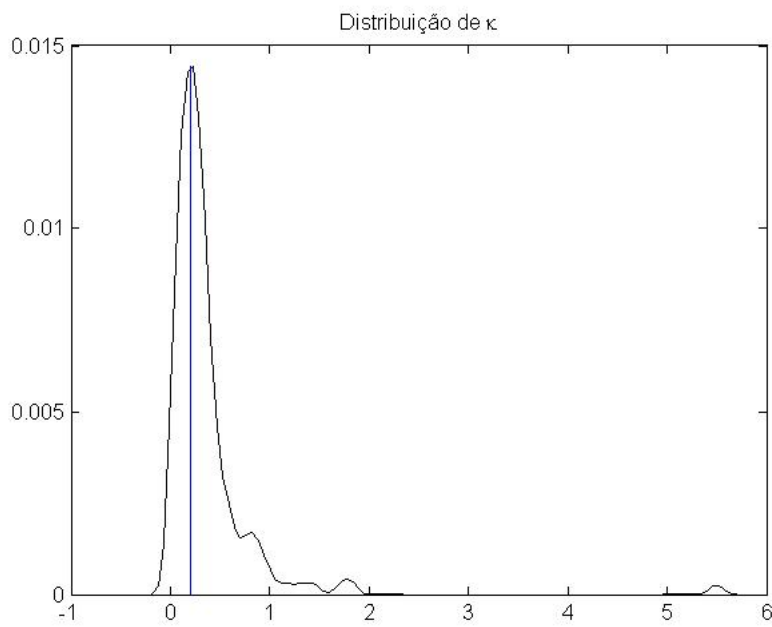


Figura 5.4: Distribuição empírica do estimador de  $\theta$

Figura 5.5: Distribuição empírica do estimador de  $\sigma$ Figura 5.6: Distribuição empírica do estimador de  $\kappa$

Da análise dos resultados da simulação de Monte Carlo, contidos na tabela 5.3 e nas figuras 5.4, 5.5 e 5.6, percebe-se que o critério de estimação para o processo pontual  $(N_t)_{t \geq 0}$  por meio do filtro de Kalman conjugado com o estimador de Quase-Máxima Verossimilhança comporta-se adequadamente.

O viés do estimador é, em média, baixo e a distribuição empírica (*finite sample*) dos parâmetros aproxima-se da esperada, conforme estabelecido na equação (5.17). Estes resultados quando comparados com os obtidos por Chen e Scott (2003) e Duan e Simonato (1999), que aplicaram o método ora proposto para estimar os parâmetros que descrevem a estrutura temporal da taxa de juros, são muito similares. Nos trabalhos mencionados, assim como neste, o estimador para amostras finitas de  $\theta$  e  $\sigma$  comporta-se como sua contraparte assintótica, com baixo viés. Por sua vez, o parâmetro de reversão à média  $\kappa$  apresenta o maior EQM entre os parâmetros estimados.

Se faz necessário destacar que, a despeito dos bons resultados encontrados ao se utilizar o Filtro de Kalman conjugado com o estimador de Quase-Máxima Verossimilhança quando aplicado a dados de negociação, o intuito do presente capítulo foi apresentar a viabilidade de implementar um método de estimação para os parâmetros do processo de Cox quando a sua intensidade possui uma forma difusiva afim. O entendimento de outras questões relacionadas à estimação não cobertas neste capítulo, tais como exogeneidade, *overfitting*, estabilidade dos parâmetros, bem como sua performance relativa em relação a outros métodos de estimação não foram tratado por estarem fora do escopo desta tese.



## Capítulo 6

### Conclusões

#### 6.1 Considerações Finais

Esta Tese explora o Processo de Cox quando sua intensidade pertence a uma família de difusões afim. A forma da função densidade de Probabilidade do Processo de Cox é obtida quando a intensidade é descrita por uma difusão afim  $d$ -dimensional arbitrária. Analisa-se também o acoplamento e convergência para o Processo de Cox com intensidade afim. Com o intuito de se obter resultados mais específicos assume-se que a intensidade do processo é uma difusão de Feller unidimensional e resultados mais detalhados são obtidos. Adicionalmente, os parâmetros da intensidade do Processo, quando esta é definida como uma difusão de Feller unidimensional, são estimados por meio do Filtro de Kalman conjugado com o estimador de Quase-Máxima Verossimilhança. Por meio de um exercício de Monte Carlo verifica-se que método de estimação proposto comporta-se adequadamente.

#### 6.2 Sugestões para Pesquisas Futuras

Entende-se como uma extensão natural deste trabalho a generalização do processo pontual  $N_t$ , para o caso  $n$ -dimensional. Neste caso, obter-se-á uma cadeia de Markov não-homogênea onde, no caso geral, o gerador infinitesimal não possuirá a propriedade de comutatividade. Sendo assim, é factível pesquisar a estrutura mínima que se deve impor no gerador infinitesimal para que seja possível obter analiticamente as probabilidades de transição da cadeia.

Uma segunda possibilidade, independente de ser possível determinar analiticamente a probabilidade de transição do processo, consiste em determinar o comportamento assintótico da cadeia, tal como sua distribuição estacionária, caso esta exista, e neste caso sua velocidade de convergência.

Uma outra linha de pesquisa seria focar exclusivamente no método de estimação proposto. Questões como a possível existência de vieses de discretização, correção do viés do estimador no caso de amostras finitas, bem como avaliar a estabilidade dos parâmetros estimados parecem ser um objeto de estudo bastante promissor.



## Apêndice A

### Códigos desenvolvidos

```
function [x, sumll,exitflag,se,test1,test2] = orders_KF(Y);
% Função para estimar os parâmetros via Filtro de Kalman e QME
global test1 test2
%  $S(t+1) = \mu + F S(t) + \text{erro}(Q)$ 
%  $Y(t) = A + H S(t) + \text{erro}(R)$ 
[nrow, ncol] = size(Y); tau = [1 1 1 1];

options = optimset('fmincon');
options = optimset('TolFun',1e-8,'Algorithm','interior-point');
para0 = [0.05, 0.1, 0.1, 0.01*rand(1,ncol).*ones(1,ncol)];
[x,fval,exitflag,output,lambda,grad,hessian] = fmincon(@loglik,...
para0,[],[],[],[],[0.0001,0.0001,0.0001,
0.00001*ones(1,ncol)],[ones(1,length(para0))],[],options,Y,
tau,... nrow, ncol);

sumll = fval; % valor da função de MLE
se=transpose(diag(inv(hessian))); % erro padrão calculado numericamente

%Gráfico com os valores previstos e realizados
figure
obs=1:1:nrow;
plot(obs, log(-Y*100),'o')
hold on
plot(obs, log(-test1*100), 'x','Color','red')
grid on
title(['\color{red}x - estimado \color{blue}o - observado'])
ylabel('P(N_t=0| t \in 1 min)','fontsize',12,'fontweight','b')
end
```

```

function sumll = loglik(para,Y, tau, nrow, ncol)
% função de Máxima Verossimilhança para estimar os parâmetros
global test

sigmai = para(4:end); R = eye(ncol);
theta = para(1); kappa = para(2); sigma = para(3);

for i = 1:ncol
    R(i,i) = sigmai(i)^2;
end

dt = 1/1; initx = theta; initV = sigma^2*theta/(2*kappa);

% Equação de transição

F = exp(-kappa*dt);

% Equação de Medida
A = zeros(1, ncol); H = A; for i = 1:ncol
    AffineGamma = sqrt((kappa+0)^2+2*sigma^2);
    AffineBeta = 2*(exp(AffineGamma*tau(i))-1)/((AffineGamma+kappa+0)*...
        (exp(AffineGamma*tau(i))-1)+2*AffineGamma);
    AffineAlpha = 2*kappa*theta/(sigma^2)*log(2*AffineGamma*exp((AffineGamma+kappa+0)*...
        tau(i)/2)/((AffineGamma+kappa+0)*...
        (exp(AffineGamma*tau(i))-1)+2*AffineGamma));
    A(i) = -AffineAlpha/tau(i);
    H(i) = AffineBeta/tau(i);
end

%Recursividades
AdjS = initx; VarS = initV;
ll = zeros(nrow,1); %log-likelihood
for i = 1:nrow
    mu = theta*(1-exp(-kappa*dt));
    PredS = mu+F*AdjS; %Valores previstos para S e Y
    Q = theta*sigma*sigma*(1-exp(-kappa*dt))^2/(2*kappa)...
        +sigma*sigma/kappa*(exp(-kappa*dt)-exp(-2*kappa*dt))*AdjS;
    VarS = F*VarS*F'+Q;
end

```

```
PredY = A+H*PredS;
PredError = Y(i,:)-PredY;
VarY = H'*VarS*H+R;
InvVarY = inv(VarY);
DetY = det(VarY);
%Atualizacao
KalmanGain = VarS*H*InvVarY;
AdjS = PredS+KalmanGain*PredError';
VarS = VarS*(1-KalmanGain*H');
ll(i) = -(ncol/2)*log(2*pi)-0.5*log(DetY)-0.5*PredError*InvVarY*PredError';

Z(i,1) = PredY;
Z2(i,1) = AdjS;

end

sumll = -sum(ll);
test_1(Z);
test_2(Z2);
end
```

```

%script para realizar as simulações de Monte Carlo
clear all close all %%
%EDE a ser simulada
% $$$dr = a(b-r)dt + \sigma\sqrt{r}dW$$$

t = 0:1/500:1;
a = 0.2;          % parametro de reversao a media
b = 0.04;        % media de longo prazo
s = 0.05 ;       % volatilidade
r0 = 0.04;       % valor inicial
MC=10;           % Numero de replicacoes da simulacao de Monte Carlo
%%
n = length(t); r = nan(n,1);

for j=1:MC
r(:,j) = cirpath(t,a,b,s,r0); % simulando a intensidade
    kappa=a;
    theta=b;
    sigma=s;
    tau=1;
    disp('iteration #')
    j
for i=1:n
    AffineGamma = sqrt((kappa+0)^2+2*sigma^2);
    AffineBeta = 2*(exp(AffineGamma*tau)-1)/((AffineGamma+kappa+0)*...
    (exp(AffineGamma*tau)-1)+2*AffineGamma);
    AffineAlpha = 2*kappa*theta/(sigma^2)*...
    log(2*AffineGamma*exp((AffineGamma+kappa+0)*tau/2)/((AffineGamma+kappa+0)*...
    (exp(AffineGamma*tau)-1)+2*AffineGamma));
    A = -AffineAlpha/tau;
    H = AffineBeta/tau;
    Rt(i,j)=A+H*r(i,j);
end
[x, sumll,exitflag,se,test1,test2] = orders_KF(Rt(:,j));
PAR(j,:) = x;
end
%distribuicao da media de longo prazo
figure [f1,xi1] = ksdensity(PAR(:,1)); plot(xi1,f1/MC,'k-') hold
on q1=ones(1,length(f1))*b; plot(q1,f1/MC) title('Distribuição de

```

```

\theta') EQM(1,2)=sum((PAR(:,1)-b).^2)/MC SE(1,2)=std(PAR(:,1));
%=====
%distribuicao do parametro de reversao a media
figure [f2,xi2] = ksdensity(-log(1-PAR(:,2)./PAR(:,1)));
plot(xi2,f2/MC,'k-') hold on q2=ones(1,length(f2))*a;
plot(q2,f2/MC) title('Distribuição de \kappa')
EQM(1,1)=sum((-log(1-PAR(:,2)./PAR(:,1))-b).^2)/MC
SE(1,1)=std(-log(1-PAR(:,2)./PAR(:,1)));
%=====
%distribuicao da volatilidade
figure [f3,xi3] = ksdensity(sqrt(PAR(:,3))); plot(xi3,f3/MC,'k-')
hold on q3=ones(1,length(f3))*0.045; plot(q3,f3/MC)
title('Distribuição de \sigma') EQM(1,3)=sum((PAR(:,3)-s).^2)/MC;
SE(1,3)=std(PAR(:,3));

```

```
function[r] = cirpath(t,a,b,s,r0)
% funcao para simular o processo de Feller
% INPUTS : t      - tempo, n*1 vector
%          a,b,s - parametros do modelo, > 0
%          r0    - valor inicial, cir(t(1))

r = [r0; nan*dt];
v = s^2; d = 4*a*b/v;
e = exp(-a*dt);
c = v.*(1-e)/(4*a);
for i = 1:(n-1)
    l = r(i)*e(i)/c(i);
    r(i+1) = c(i)*ncx2rnd(d,l);
end
```



## Referências Bibliográficas

- Albanese e Lawi(2004)** C. Albanese e S. Lawi. Laplace transforms for integrals of markov processes. *Markov Processes and Related Fields*, 11:677–724. 15, 26
- Basu e Dassios(2002)** S. Basu e A. Dassios. A cox process with log-normal intensity. *Insurance: Mathematics and Economics*, 31:297–302. 2
- Bouzas et al.(2002)** P.R. Bouzas, M.J. Valderrama, e A.M. Aguilera. Forecasting a class of doubly poisson processes. *Statistical Papers*, 43:507–523. 2
- Bouzas et al.(2006)** P.R. Bouzas, M.J. Valderrama, e A.M. Aguilera. On the characteristic functional of a doubly stochastic poisson process: Application to a narrow-band process. *Applied Mathematical Modelling*, 30:1021–1032. 2
- Brémaud(1972)** P. Brémaud. *Point Processes and Queues: Martingale Dynamics*. Springer-Verlag, New York. 1, 3
- Bucy e Joseph(1968)** R. S. Bucy e P. D. Joseph. *Filtering for Stochastic Processes, with Applications to Guidance*. Wiley, New York. 50
- Chen e Scott(2003)** R.-R. Chen e L. Scott. Multi-factor cox-ingersoll-ross models of the term structure: Estimates and tests from a kalman filter model. *The Journal of Real Estate Finance and Economics*, 27:143–172. 60
- Cont et al.(2010)** R. Cont, S. Stoikov, e R. Talreja. A stochastic model for order book dynamics. *Operations Research*, no Prelo. 3
- Cox(1955)** D. R. Cox. Some statistical methods connected with series of events. *Journal of Royal Statistical Society B*, 17:129–164. 1
- Cox et al.(1985)** J. Cox, J. Ingersoll, e S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–408. 2
- Dai e Singleton(2000)** Q. Dai e K. Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, LV(5):1943–1978. 11, 12, 13, 23, 52
- Dalal e McIntosh(1994)** S. Dalal e A. McIntosh. When to stop testing for large software systems with changing code. *IEEE Trans. Software Eng*, 20:318–323. 1
- Daley e Vere-Jones(1988)** D.J. Daley e D. Vere-Jones. *An Introduction to Theory of Point Processes*. Springer-Verlag, New York. 1, 3
- Dassios e Jang(2008)** A. Dassios e J. Jang. The distribution of the interval between events of a cox process with shot noise intensity. *Journal of Applied Mathematics and Stochastic Analysis*. 2

- Duan e Simonato(1999)** J. Duan e J. Simonato. Estimating and testing exponential-affine term structure models by kalman filter. *Review of Quantitative Finance and Accounting*, 13:111–135. 47, 60
- Duffie e Kan(1996)** D. Duffie e R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379–406. 2, 5, 6, 13
- Duffie e Singleton(1999)** D. Duffie e K. Singleton. Modeling term structures defaultable bonds. *Review of Financial Studies*, 12:687–720. 3
- Duffie et al.(2003)** D. Duffie, D. Filipovic, e W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13:984–1053. 3
- Engle(2000)** R. F. Engle. The econometrics of ultra-high-frequency data. *Econometrica*, 68:1–22. 45
- Engle e Russell(1998)** R. F. Engle e J. R. Russell. Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, 66:1127–1162. 45
- Feller(1951)** W. Feller. Two singular diffusion problems. *Annals of Mathematics*, 54:173–182. 2
- Ferreira e Sundar(2006)** G. Ferreira e P. Sundar. Comparison of solutions of stochastic equations and applications. *Stochastic Analysis and Applications*, 18:211–229. 41
- Gail et al.(1980)** M. Gail, T. Santner, e C. Brown. An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics*, 36:255–266. 1
- Grandell(1976)** J. Grandell. *Doubly Stochastic Process*. Springer-Verlan, New York, 1ª edição. 1, 3
- Hamilton(1994)** J.D. Hamilton. *Time Series Analysis*. Princeton University Press, Princeton. 54
- Harvey e Fernandes(1989)** A. C. Harvey e C. Fernandes. Time series models for count or qualitative observations. *Journal of Business & Economic Statistics*, 7(4):407–417. 46
- Harvey(1989)** A.C. Harvey. *Forecasting, structural time series models and the Kalman Filter*. Cambridge University Press, Cambridge. 51
- Ikeda e Watanabe(1981)** N. Ikeda e S. Watanabe. *Stochastic Differential Equations and Diffusions Processes*. North-Holland. 8, 11
- Johannes e Polson(2009)** Michael Johannes e Nicholas Polson. MCMC methods for financial econometrics. Em Y. Ait-Sahalia e L.P. Hansen, editors, *Handbook of Financial Econometrics*, volume 2, páginas 1–72. North-Holland. 46
- Kallenberg(1986)** O. Kallenberg. *Random Measures*. Academic Press, London, 4ª edição. 23
- Karatzas e Shreve(1991)** I. Karatzas e S. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlan, New York, 2ª edição. 10
- Karlin e Taylor(1981)** S. Karlin e H. Taylor. *A Second Course in Stochastic Process*. Academic Press, New York. 30, 37

- Kozachenko e Pogorilyak(2008)** Yu. V. Kozachenko e O. O. Pogorilyak. A method of modelling log gaussian cox process. *Theory of Probability and Mathematical Statistics*, 77:91–105. 2
- Lando(1998)** D. Lando. On cox processes and credit risky securities. *Review of Derivatives Research*, 2:99–120. 1
- Pinsky(1995)** R. Pinsky. *Positive Harmonic Functions and Diffusions*. Cambridge University Press. 34
- Wei et al.(2002)** G. Wei, P. Clifford, e J. Feng. Population death sequences and cox processes driven by interacting feller diffusions. *Journal of Physics A: Mathematical and General*, 35:931. 2, 3