**Filtro de Kalman Ensemble:
Uma análise da estimação conjunta
dos estados e dos parâmetros**

Rafael Oliveira Silva

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

Programa: Estatística

Orientador: Prof^a. Dr^a. Márcia D'Elia Branco

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CNPq

São Paulo, abril de 2019

Filtro de Kalman Ensemble: Uma análise da estimação conjunta dos estados e dos parâmetros

Filtro de Kalman Ensemble: Uma análise da estimação conjunta dos estados e dos parâmetros

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 08/04/2019. Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof^a. Dr^a. Márcia D'Elia Branco (orientadora) - IME-USP
- Prof^a. Dr^a. Glauro da Conceição Franco - ICEx-UFMG
- Prof. Dr. Rafael Bráz Azevedo Farias - DEMA-UFC

Agradecimentos

Ao Instituto de Matemática e Estatística da Universidade de São Paulo e seus membros do Programa de Pós-graduação em Estatística.

À Professora Dr^a. Márcia D'Elia Branco, pela paciência e compreensão em me orientar.

Aos membros da banca examinadora, pelo interesse e disponibilidade.

Aos Professores Dr. José Ailton Alencar Andrade, Dr. João Maurício Araujo Mota, Dr. Juvêncio Santos Nobre, Dr. Rafael Bráz Azevedo Farias, Dr. Luis Gustavo B. Pinho e Dr. Pushpa Narayan Rathie, pela dedicação ao ensino de estatística e por me motivarem ao longo do curso de graduação em Estatística na Universidade Federal do Ceará.

Aos meus amigos Alessandra (irmã Toinha), Kevin, Rosália, Nádia, Jonatas, Estefanía, Andrés, Shu, Renato, Tyler, Anderson, Robert, Fernando, Luiz, Thaís, Raphael, Camila, pelas conversas e pelos momentos de distração.

À minha família e minha namorada Beatriz por me apoiarem em todos os momentos.

Resumo

SILVA, R. O. **Filtro de Kalman Ensemble: Uma análise da estimação conjunta dos estados e dos parâmetros**. 2019. 103 f. Dissertação (Mestrado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2019.

O Filtro de Kalman Ensemble (EnKF) é um algoritmo de Monte Carlo sequencial para inferência em modelos de espaço de estados lineares e não lineares. Este filtro combinado com alguns outros métodos propaga a distribuição *a posteriori* conjunta dos estados e dos parâmetros ao longo do tempo. Existem poucos trabalhos que consideram o problema da estimação simultânea dos estados e parâmetros, e os métodos existentes possuem limitações. Nesta dissertação analisamos a eficiência desses métodos por meio de estudos de simulação em modelos de espaço de estados lineares e não lineares. O problema de estimação não linear aqui tratado refere-se ao modelo de produção excedente logístico, para o qual o EnKF pode ser considerado uma possível alternativa aos algoritmos MCMC. Os resultados da simulação revelam que a acurácia das estimativas aumenta quando a série temporal cresce, mas alguns parâmetros apresentam problemas na estimação.

Palavras-chave: Filtro de Kalman Ensemble, estimação conjunta, modelo de produção excedente logístico.

Abstract

SILVA, R. O. **Ensemble Kalman Filter: An analysis of the joint estimation of states and parameters**. 2019. 103 f. Dissertação (Mestrado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2019.

The Ensemble Kalman Filter (EnKF) is a sequential Monte Carlo algorithm for inference in linear and nonlinear state-space models. This filter combined with some other methods propagates the joint posterior distribution of states and parameters over time. There are fewer papers that consider the problem of simultaneous state-parameter estimation and existing methods have limitations. The purpose of this dissertation is to analyze the efficiency of these methods by means of simulation studies in linear and nonlinear state-space models. The nonlinear estimation problem addressed here refers to the logistic surplus-production model, for which the EnKF can be considered as a possible alternative to MCMC algorithms. The simulation results reveal that the accuracy of the estimates increases when the time series grows, but some parameters present problems in the estimation.

Keywords: Ensemble Kalman Filter, joint estimation, logistic surplus-production model.

Sumário

Lista de Figuras	vi
Lista de Tabelas	ix
1 Introdução	2
1.1 Objetivos e Organização do Trabalho	3
2 Modelos de Espaço de Estados	4
2.1 Modelo de Espaço de Estados Linear Normal	5
2.2 Modelos de Espaço de Estados Não Lineares Normais	8
2.3 Modelos de Espaço de Estados Não Lineares e Não Normais	9
3 Modelos de Produção Excedente	10
3.1 Modelos de Produção	10
3.1.1 Modelo de Produção Logístico	11
3.1.2 Outros Modelos de Produção	13
3.2 Modelos de Produção Excedente vistos como Modelos de Espaço de Estados	14
4 Inferência em Modelos de Espaço de Estados via KF e EnKF	15
4.1 Filtro de Kalman (KF)	16
4.2 Filtro de Kalman Ensemble (EnKF)	18
4.2.1 Estimação dos Estados	19
4.2.2 Estimação dos parâmetros	20
5 Simulação I - Modelo de Espaço de Estados Linear Normal	25
5.1 Estimação dos parâmetros σ^2 e τ^2	25
5.2 Estimação do parâmetro σ^2 com $\tau^2 = \kappa\sigma^2$	30
5.3 Estimação dos parâmetros α e σ^2	31
5.4 Estimação dos parâmetros q , α e σ^2	39
6 Simulação II - Modelo de Produção Excedente	45
6.1 Análise de Sensibilidade	46

6.2	Estimação dos Estados	50
6.3	Análise de dados reais	54
7	Conclusão	58
A	Apêndice	60
A.1	Filtro de Kalman Ensemble	60
B	Apêndice	62
B.1	Efeito do tamanho da amostra ensemble nas estimativas dos parâmetros .	62
B.2	Comparando as estimativas do EnKF com as estimativas obtidas via o pacote DLM	63
C	Apêndice	65
C.1	Figuras do Capítulo 5	65
C.1.1	Figuras da seção 5.1	65
C.1.2	Figuras da seção 5.3	67
C.2	Figuras do Capítulo 6	71
D	Códigos R	75
D.1	Códigos R do capítulo 5	75
D.1.1	Seção 5.3 - Estimação de α e σ^2	75
D.1.2	Seção 5.4 - Estimação de q , α e σ^2	77
D.2	Códigos R do capítulo 6	79
D.2.1	Seção 6.2 - Estimação de r , K e σ^2	79
D.2.2	Seção 6.3 - Estimação de r , K e σ^2	82
D.3	Códigos do Apêndice B.2	83
	Referências Bibliográficas	88

Lista de Figuras

2.1	Representação gráfica do modelo de espaço de estados.	5
3.1	Função de produção do modelo de Schaefer.	11
3.2	Biomassa da população quando $x_0 < K$, $x_0 = K$, e $x_0 > K$	12
5.1	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 50 - (σ^2 e τ^2).	28
5.2	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 100 - (σ^2 e τ^2).	28
5.3	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 100 - Caso I.	36
5.4	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 100 - Caso II.	36
5.5	Estimativa do viés e do desvio absoluto de α para diferentes comprimentos da série de observações considerando os casos I e II.	38
5.6	Estimativa do viés e do desvio absoluto de σ^2 para diferentes comprimentos da série de observações considerando os casos I e II.	38
5.7	Distribuição de <i>filtragem</i> do estado X_{100} para o parâmetro q assumindo os valores 0.5, 1, 5 e 9.	39
5.8	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 100 - Caso I.	43
5.9	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 100 - Caso II.	44
6.1	Estimativas das distribuições preditivas e a <i>posteriori</i> de cada estado para um modelo não linear e uma série de observações de comprimento igual a 100 com $\sigma^2 \sim GI(0.1, 0.1)$	51
6.2	Estimativas das distribuições preditivas e a <i>posteriori</i> de cada estado para um modelo não linear e uma série de observações de comprimento igual a 100 com $\sigma \sim HC(25)$	51
6.3	Dados de Camarão do Chile.	54
6.4	Estimativas das distribuições preditivas das observações.	57

B.1	Na gráfico do lado esquerdo apresentamos as estimativas do parâmetro α obtidas via EnKF e MCMC. Do lado direito apresentamos as estimativas de σ^2 obtidas pelos os dois métodos.	64
C.1	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 30 - (σ^2 e τ^2)	65
C.2	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 250 - (σ^2 e τ^2)	66
C.3	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 500 - (σ^2 e τ^2).	66
C.4	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 30 - Caso I.	67
C.5	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 50 - Caso I.	67
C.6	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 250 - Caso I.	68
C.7	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 500 - Caso I.	68
C.8	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 30 - Caso II.	69
C.9	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 50 - Caso II.	69
C.10	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 250 - Caso II.	70
C.11	Estimativa da distribuição preditiva e a <i>posteriori</i> de cada estado para uma série de observações de comprimento igual a 500 - Caso II.	70
C.12	Estimativas das distribuições preditivas e a <i>posteriori</i> de cada estado para um modelo não linear e uma série de observações de comprimento igual a 30 com $\sigma^2 \sim GI(0.1, 0.1)$	71
C.13	Estimativas das distribuições preditivas e a <i>posteriori</i> de cada estado para um modelo não linear e uma série de observações de comprimento igual a 50 com $\sigma^2 \sim GI(0.1, 0.1)$	71
C.14	Estimativas das distribuições preditivas e a <i>posteriori</i> de cada estado para um modelo não linear e uma série de observações de comprimento igual a 250 com $\sigma^2 \sim GI(0.1, 0.1)$	72
C.15	Estimativas das distribuições preditivas e a <i>posteriori</i> de cada estado para um modelo não linear e uma série de observações de comprimento igual a 500 com $\sigma^2 \sim GI(0.1, 0.1)$	72

C.16 Estimativas das distribuições preditivas e a <i>posteriori</i> de cada estado para um modelo não linear e uma série de observações de comprimento igual a 30 com $\sigma \sim HC(25)$	73
C.17 Estimativas das distribuições preditivas e a <i>posteriori</i> de cada estado para um modelo não linear e uma série de observações de comprimento igual a 50 com $\sigma \sim HC(25)$	73
C.18 Estimativas das distribuições preditivas e a <i>posteriori</i> de cada estado para um modelo não linear e uma série de observações de comprimento igual a 250 com $\sigma \sim HC(25)$	74
C.19 Estimativas das distribuições preditivas e a <i>posteriori</i> de cada estado para um modelo não linear e uma série de observações de comprimento igual a 500 com $\sigma \sim HC(25)$	74

Lista de Tabelas

5.1	Análise de sensibilidade de τ^2	26
5.2	Medidas de acurácia na análise de sensibilidade de τ^2	26
5.3	Análise de sensibilidade de σ^2	26
5.4	Medidas de acurácia na análise de sensibilidade de σ^2	27
5.5	Taxa de cobertura do intervalo interquartil da distribuição de filtragem. . . .	27
5.6	Estimativas dos parâmetros σ^2 e τ^2 para os diferentes comprimentos da série de observações.	29
5.7	Medidas de acurácia das estimativas dos parâmetros σ^2 e τ^2 para os dife- rentes comprimentos da série de observações.	29
5.8	Análise de sensibilidade de σ^2 com $\sigma^2 = 9$	30
5.9	Medidas de acurácia na análise de sensibilidade de σ^2 com $\sigma^2 = 9$	30
5.10	Análise de sensibilidade de σ^2 com $\alpha = 1$ e $\sigma^2 = 4$	32
5.11	Medidas de acurácia na análise de sensibilidade de σ^2 com $\alpha = 1$ e $\sigma^2 = 4$	32
5.12	Análise de sensibilidade de σ^2 com $\alpha = 1$ e $\sigma^2 = 1$	32
5.13	Medidas de acurácia na análise de sensibilidade de σ^2 com $\alpha = 1$ e $\sigma^2 = 1$	33
5.14	Análise de sensibilidade de α com $\alpha = 1$ e $\sigma^2 = 4$	33
5.15	Medidas de acurácia na análise de sensibilidade de α com $\alpha = 1$ e $\sigma^2 = 4$	34
5.16	Análise de sensibilidade de α com $\alpha = 1$ e $\sigma^2 = 1$	34
5.17	Medidas de acurácia na análise de sensibilidade de α com $\alpha = 1$ e $\sigma^2 = 1$	35
5.18	Taxa de cobertura do intervalo interquartil da distribuição de filtragem. . . .	35
5.19	Estimativas dos parâmetros α e σ^2 para diferentes comprimentos da série de observações.	37
5.20	Análise de sensibilidade de q	40
5.21	Medidas de acurácia na análise de sensibilidade de q	40
5.22	Estimativas dos parâmetros q , α e σ^2 para diferentes comprimentos da sé- rie de observações.	41
5.23	Medidas de acurácia das estimativas dos parâmetros q , α e σ^2 para dife- rentes comprimentos da série de observações.	42
5.24	Taxa de cobertura média dos intervalos interquartil da distribuição de fil- tragem.	43

6.1	Análise de sensibilidade do parâmetro r .	46
6.2	Medidas de acurácia na análise de sensibilidade do parâmetro r .	47
6.3	Análise de sensibilidade do parâmetro K .	48
6.4	Medidas de acurácia na análise de sensibilidade do parâmetro K .	48
6.5	Análise de sensibilidade σ^2 .	49
6.6	Medidas de acurácia na análise de sensibilidade do parâmetro σ^2 .	50
6.7	Taxa de cobertura média dos intervalos interquartil das distribuições de <i>filtragem</i> dos estados.	50
6.8	Estimativas dos parâmetros para os diferentes comprimentos da série de observações.	52
6.9	Análise da acurácia das estimativas dos parâmetros para diferentes tamanhos do conjunto de dados.	53
6.10	Estatísticas descritivas dos dados de Camarão do Chile.	54
6.11	Estimativas dos parâmetros r, K, q e σ^2 via EnKF	55
6.12	Estimativas dos parâmetros r, K e σ^2 via EnKF	55
6.13	Estimativas dos parâmetros r, K, q, σ^2 e τ^2 via MCMC.	55
6.14	Estimativas B_{MSY}, H_{MSY} e MSY via EnKF - (q desconhecido)	56
6.15	Estimativas das quantidades B_{MSY}, H_{MSY} e MSY via EnKF - (q conhecido)	56
6.16	Estimativas dos parâmetros r, K, q, σ^2 e τ^2 via MCMC.	56
B.1	Estimativas dos parâmetros α e σ^2 para diferentes tamanhos da amostra ensemble.	62
B.2	Análise da acurácia das estimativas dos parâmetros para diferentes tamanhos da amostra ensemble.	63
B.3	Tempo computacional em segundos para obter as estimativas em um série de comprimento N .	64

Capítulo 1

Introdução

Os modelos de séries temporais abordados nesta dissertação são conhecidos como modelos dinâmicos ou modelos de espaço de estados. Esses modelos são geralmente representados por um sistema que descreve a dependência probabilística entre os estados e os dados observados. Na literatura de modelos dinâmicos, os estados são variáveis latentes, isto é, quantidades não observáveis que podem ser mensuradas indiretamente por meio de variáveis observadas relacionadas a elas.

No processo de inferência dos estados existem várias abordagens que podem ser encontradas na literatura, uma delas é o método de *filtragem* que é um procedimento recursivo que permite estimar as variáveis de estado a partir de novas observações. Um dos mais populares algoritmos de inferência sequencial dos estados é o Filtro de Kalman (KF, do inglês *Kalman Filter*). Este filtro foi proposto por Kalman (1960) para inferência dos estados em modelos dinâmicos lineares normais e pode ser considerado como um propagador da distribuição probabilística dos estados no tempo, já que possibilita uma maneira de obter recursivamente a distribuição exata de predição e de *filtragem* dos estados.

Para modelos normais e não lineares, a obtenção de um método de *filtragem* exato é geralmente inviável, pois não é possível obter as distribuições de *filtragem* e de predição dos estados analiticamente. Deste modo, é necessário o uso de aproximações. Uma das primeiras aproximações do Filtro de Kalman para modelos não lineares normais foi o Filtro de Kalman Extendido (EKF, de *Extended Kalman Filter*) proposto por Maybeck (1979). Este filtro lineariza as equações não lineares dos estados e das observações para a utilização do Filtro de Kalman padrão. A linearização feita no EKF leva a atualizações precárias das estimativas a medida em que o grau de não linearidade aumenta. Uma outra extensão do KF é o Filtro de Kalman *Ensemble* (EnKF, de *Ensemble Kalman Filter*) proposto por Evensen (1994), que é um método de Monte Carlo sequencial que usa uma amostra ou conjunto (*ensemble*, do inglês) como uma representação da distribuição dos estados ao invés da distribuição explícita como no KF.

O EnKF é construído com base em um modelo de espaço de estados linear gaussiano, mas segundo Katzfuss *et al.* (2016) é robusto a desvios dessas suposições. Este filtro lida com a não linearidade de uma maneira mais flexível do que métodos de linearização analítica, como o filtro de Kalman Extendido, e permite fazer inferência em modelos de espaço de estados com dimensão na ordem de milhões, enquanto que abordagens padrões, como os métodos de Monte Carlo via Cadeias de Markov (MCMC), apresentam limitações em termos de custo computacional em tais problemas. Por estas características o EnKF é amplamente utilizado em aplicações de assimilação de dados

com dimensão na ordem de dezenas de milhões e estudado principalmente na área de Geofísica. Algumas referências nessa área são: Evensen e Van Leeuwen (1996), Houtekamer e Mitchell (1998), Houtekamer e Zhang (2016), Mitchell e Houtekamer (2000), Wikle e Berliner (2007), Evensen (1997), Stroud e Bengtsson (2007), Houtekamer e Mitchell (1998) e Evensen (1994, 2009b). Katzfuss *et al.* (2016) faz uma breve revisão do método com a intenção de divulgar a metodologia para a comunidade estatística.

Contudo, embora muitos trabalhos tenham sido feitos para melhorar as estimativas dos estados via EnKF, existem poucos trabalhos com foco na estimação conjunta dos estados e dos parâmetros dos modelos. Algumas alternativas para resolver o problema de estimação conjunta foram propostas por Anderson (2001) e Stroud *et al.* (2016) e, em muitos casos, podem resolver os problemas de estimação conjunta de forma eficiente. Neste trabalho iremos investigar a eficiência dessas alternativas para modelos univariados lineares e não lineares.

1.1 Objetivos e Organização do Trabalho

O presente trabalho tem o objetivo de estudar o Filtro de Kalman *Ensemble* (EnKF) com relação a acurácia das estimativas dos estados e dos parâmetros em modelos univariados lineares e não lineares por meio de simulações. Nas simulações verificamos o efeito do tamanho da amostra *ensemble* nas estimativas dos estados, o efeito do comprimento da série de observações nas estimativas *a posteriori* e a influência da informação *a priori* na distribuição *a posteriori* dos parâmetros.

Esta dissertação está organizada da seguinte forma: No Capítulo 2, apresentamos uma breve revisão dos modelos de espaço de estados.

No Capítulo 3, introduzimos um modelo de espaço de estados não linear denominado modelo de produção excedente, também são apresentadas as principais funções de produção e suas medidas para a biomassa e captura associadas a produção máxima sustentável de uma população animal.

No Capítulo 4, apresentamos os conceitos do Filtro de Kalman e Filtro de Kalman *Ensemble* para a estimação dos estados e também alguns métodos que, em conjunto com o EnKF, estimam os parâmetros.

No Capítulo 5, consideramos o EnKF em conjunto com alguns métodos de estimação de parâmetros para fazer inferências em um modelo de espaço de estados linear gaussiano. Nas simulações analisamos o efeito do comprimento da série temporal das observações e a influência da distribuição *a priori* na distribuição *a posteriori* dos parâmetros e dos estados.

No Capítulo 6, consideramos o EnKF em conjunto com alguns métodos de estimação de parâmetros para fazer inferências em um modelo não linear gaussiano representado pelo modelo de produção logístico. Nas simulações analisamos o efeito do comprimento da série temporal das observações e a influência da distribuição *a priori* na distribuição *a posteriori* dos parâmetros e dos estados.

Finalmente, no Capítulo 7 discutimos algumas conclusões e extensões para modelos não gaussianos.

Capítulo 2

Modelos de Espaço de Estados

Modelos de espaço de estados oferecem uma forma unificada de modelar muitos tipos de séries temporais e outros dados, pois são uma rica e flexível família de modelos que inclui, como casos especiais, muitos dos modelos clássicos usados para a análise de séries temporais (Brockwell e Davis, 2016). Além disso, a propriedade Markoviana atribuída a variável de estado permite a aplicação de algoritmos recursivos de estimação e previsão por meio do Filtro de Kalman. A flexibilidade dessa modelagem faz com que esses modelos sejam utilizados em diversas áreas.

Um modelo de espaço de estados é composto por uma série temporal $\mathbb{R}^m$ –dimensional $(\mathbf{x}_t : t = 0, 1, \dots)$ e uma série temporal $\mathbb{R}^r$ –dimensional $(\mathbf{y}_t : t = 1, 2, \dots)$. A primeira série temporal é formada por variáveis que não são diretamente observadas, chamadas de estados, e a segunda série temporal é formada por um conjunto de observações $(\mathbf{y}_t)_{t \geq 1}$, que são assumidas independentes uma da outra quando condicionadas ao processo $(\mathbf{x}_t)_{t \geq 0}$. A estrutura de um modelo de espaço de estados é formada por:

- i) **Distribuição Inicial**, que descreve o conhecimento *a priori* do estado inicial $\mathbf{x}_0$ no tempo inicial $t = 0$ baseado em um conjunto de informações iniciais $\mathbf{D}_0$ e é denotada por $p(\mathbf{x}_0|\mathbf{D}_0)$.
- ii) **Modelo de Transição**, que descreve a dinâmica do sistema e suas incertezas com um processo de Markov, em que a distribuição de $\mathbf{x}_t$ dado $\mathbf{x}_{t-1}, \dots, \mathbf{x}_1$ é a mesma que a distribuição de $\mathbf{x}_t$ dado $\mathbf{x}_{t-1}$, isto é, o estado $\mathbf{x}_{t-1}$ resume todas as informações probabilísticas relevantes do passado. Tal propriedade é o aspecto crucial da modelagem de espaço de estados, pois não há a necessidade do armazenamento e processamento do restante do histórico dos estados ou das observações. Em termos de uma distribuição de probabilidade o modelo de transição é representado por $p(\mathbf{x}_t|\mathbf{x}_{t-1})$.
- iii) **Modelo das Observações**, que descreve como a observação $\mathbf{y}_t$ depende do estado corrente $\mathbf{x}_t$. Essa dependência é modelada especificando a distribuição de probabilidade condicional da observação, $\mathbf{y}_t$, dado o estados $\mathbf{x}_t$, que é denotado por $p(\mathbf{y}_t|\mathbf{x}_t)$.

Assim, um modelo de espaço de estados pode ser probabilisticamente escrito da seguinte forma,

Distribuição Inicial	$\mathbf{x}_0 \sim p(\mathbf{x}_0 \mathbf{D}_0)$
Distribuição de Transição	$\mathbf{x}_t \mathbf{x}_{t-1} \sim p(\mathbf{x}_t \mathbf{x}_{t-1})$
Distribuição da Observação	$\mathbf{y}_t \mathbf{x}_t \sim p(\mathbf{y}_t \mathbf{x}_t)$

Na Figura 2.1 é apresentada a representação gráfica do modelo

Figura 2.1: Representação gráfica do modelo de espaço de estados.

As distribuições que compõem o modelo de espaço de estados podem depender de uma estrutura de parâmetros desconhecidos $\psi \in \mathbb{R}^k$ invariante no tempo. E assim, denotaremos a distribuição inicial, a distribuição de transição e a distribuição das observações por, $p(\mathbf{x}_0 | \psi)$, $p(\mathbf{x}_t | \mathbf{x}_{t-1}, \psi)$ e $p(\mathbf{y}_t | \mathbf{x}_t, \psi)$, respectivamente. Deste modo, considerando um conjunto de estados $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$ e um conjunto de observações $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$, temos que a distribuição conjunta dos estados e das observações é dada por

$$P(\mathbf{x}_{0:n}, \mathbf{y}_{1:n} | \psi) = \left(\prod_{t=1}^n p(\mathbf{y}_t | \mathbf{x}_t, \psi) p(\mathbf{x}_t | \mathbf{x}_{t-1}, \psi) \right) p(\mathbf{x}_0 | \psi), \quad (2.1)$$

em que $\mathbf{x}_{i:j}$ e $\mathbf{y}_{i:j}$ denotam um conjunto de estados e observações que estão entre i e j , ou seja, $\{\mathbf{x}_i, \dots, \mathbf{x}_j\}$ e $\{\mathbf{y}_i, \dots, \mathbf{y}_j\}$. Mediante condicionamentos e marginalizações da distribuição (2.1) podemos obter as distribuições de interesse para inferência.

2.1 Modelo de Espaço de Estados Linear Normal

O modelo de espaço de estados linear gaussiano, também conhecido como modelo dinâmico linear normal, é um caso especial importante de modelos de espaço de estados. Neste modelo, atribui-se aos erros distribuições gaussianas e assume-se que a relação dos componentes latentes é dada de forma linear em ambas as equações do

modelo. Uma descrição detalhada desse modelo pode ser vista em [Brockwell e Davis \(2016\)](#), [West e Harrison \(2006\)](#) e também em [Campagnoli et al. \(2009\)](#).

Um modelo de espaço de estados linear para uma série temporal multivariada $\{y_t, t = 1, 2, \dots, n\}$ consiste de duas equações. A primeira equação é denominada equação de estados (equação de transição ou equação do processo) e relaciona o estado no tempo t com o estado no tempo $t - 1$. Assim,

$$\mathbf{x}_t = \mathbf{M}_t \mathbf{x}_{t-1} + \mathbf{w}_t, \quad t = 1, 2, \dots \quad (2.2)$$

em que $\mathbf{w}_t \sim \mathcal{N}_m(\mathbf{0}, \mathbf{Q}_t)$, $(\mathbf{Q}_t)_{t \geq 1}$ e $(\mathbf{M}_t)_{t \geq 1}$ são duas sequências de matrizes conhecidas de dimensão $m \times m$.

A segunda equação é conhecida como equação de observação, ela relaciona o vetor de observações r -dimensional ao vetor de estados m -dimensional por meio de uma função linear adicionada de um ruído. A equação de observação é dada por,

$$\mathbf{y}_t = \mathbf{H}_t' \mathbf{x}_t + \mathbf{v}_t, \quad t = 1, 2, \dots, n \quad (2.3)$$

em que $\mathbf{v}_t \sim \mathcal{N}_r(\mathbf{0}, \mathbf{R}_t)$, $(\mathbf{R}_t)_{t \geq 1}$ é uma sequência de matrizes conhecidas de dimensão $r \times r$ e $(\mathbf{H}_t)_{t \geq 1}$ é uma sequência de matrizes conhecidas de dimensão $m \times r$. Para completar a especificação do modelo, assume-se que o estado inicial está condicionado a alguma informação inicial $\mathbf{D}_0$. Assim,

$$\mathbf{x}_0 | \mathbf{D}_0 \sim \mathcal{N}_m(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$

em que $\boldsymbol{\mu}_0$ e $\boldsymbol{\Sigma}_0$ representam essa informação inicial. As duas sequências de erros $(\mathbf{v}_t)_{t \geq 1}$ e $(\mathbf{w}_t)_{t \geq 1}$ são internamente e mutuamente independentes, e independentes de $\mathbf{x}_0 | \mathbf{D}_0$.

As equações apresentadas do modelo de espaço de estados linear podem ser especificadas na forma de distribuições de probabilidade condicionais. Um modelo linear de espaço de estados pode ser especificado pelas densidades de transição e observação,

$$\mathbf{x}_t | \mathbf{x}_{t-1} \sim \mathcal{N}_m(\mathbf{M}_t \mathbf{x}_{t-1}, \mathbf{Q}_t)$$

e

$$\mathbf{y}_t | \mathbf{x}_t \sim \mathcal{N}_r(\mathbf{H}_t' \mathbf{x}_t, \mathbf{R}_t),$$

respectivamente.

Casos Especiais do Modelo Linear Normal

Modelo de nível local

O modelo de nível local ou modelo passeio aleatório mais ruído é o modelo mais simples entre os modelos de espaço de estados e é apropriado para séries temporais que não possuem variação de tendência ou sazonalidade. Assim, para uma série temporal univariada $(y_t : t = 1, 2, \dots, n)$, o modelo de nível local é definido pela equação de observação

$$y_t = x_t + v_t, \quad v_t \sim N(0, \sigma_v^2)$$

e por uma equação de estados que é um simples passeio aleatório

$$x_t = x_{t-1} + w_t, \quad w_t \sim N(0, \sigma_w^2).$$

As sequências de erros $(v_t)_{t \geq 1}$ e $(w_t)_{t \geq 1}$ são internamente e mutuamente independentes e independentes de $x_0 | D_0 \sim N(\mu_0, \sigma_{w0}^2)$. Comparando com a estrutura geral apresentada nas equações (2.2) e (2.3), esse modelo é caracterizado por $m = r = 1$, $\mathbf{H}_t = \mathbf{M}_t = 1$, $\mathbf{R} = \sigma_v^2$ e $\mathbf{Q} = \sigma_w^2$.

Modelo de crescimento linear

O modelo de nível local ao incorporar localmente uma tendência linear passa a ser chamado de modelo de crescimento linear ou modelo de tendência linear. Nesse modelo, a equação de observação é a mesma do modelo de nível local, mas a equação de estados é modificada com a inclusão de uma tendência que pode variar ao longo do tempo, como segue:

$$\begin{aligned} y_t &= x_{t,1} + v_t, & v_t &\sim N(0, \sigma_v^2) \\ x_{t,1} &= x_{t-1,1} + x_{t-1,2} + w_{t,1}, & w_{t,1} &\sim N(0, \sigma_{w1}^2) \\ x_{t,2} &= x_{t-1,2} + w_{t,2}, & w_{t,2} &\sim N(0, \sigma_{w2}^2) \end{aligned}$$

em que os erros v_t , $w_{t,1}$ e $w_{t,2}$ são não correlacionados. Podemos expressar o modelo de tendência linear na forma introduzida nas equações (2.2) e (2.3),

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}_t + v_t$$

e

$$\mathbf{x}_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x}_{t-1} + \mathbf{w}_t.$$

Podemos identificar as seguintes estruturas do modelo,

$$\mathbf{x}_t = \begin{bmatrix} x_{t,1} \\ x_{t,2} \end{bmatrix}, \quad \mathbf{w}_t = \begin{bmatrix} w_{t,1} \\ w_{t,2} \end{bmatrix}, \quad \mathbf{M}_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{H}_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{Q}_t = \begin{bmatrix} \sigma_{w1}^2 & 0 \\ 0 & \sigma_{w2}^2 \end{bmatrix} \text{ e } \mathbf{R}_t = \sigma_v^2.$$

Note que, as matrizes $\mathbf{H}_t$, $\mathbf{M}_t$ e as matrizes de covariâncias $\mathbf{R}_t$ e $\mathbf{Q}_t$ são constantes ao longo do tempo, e neste caso, o modelo é dito invariante no tempo.

2.2 Modelos de Espaço de Estados Não Lineares Normais

Os modelos de espaço de estados não lineares gaussianos são mais flexíveis que os modelos de espaço de estados lineares gaussianos, pois neles a equação de estados descreve a evolução do vetor de estados m -dimensional, $\mathbf{x}_t$, por meio de uma função de transição não linear $\mathcal{M}_t(\cdot)$, como segue,

$$\mathbf{x}_t = \mathcal{M}_t(\mathbf{x}_{t-1}, \mathbf{w}_t), \quad t = 1, 2, \dots$$

em que $\mathbf{w}_t \sim \mathcal{N}_m(\mathbf{0}, \mathbf{Q}_t)$ e $(\mathbf{Q}_t)_{t \geq 1}$ é uma sequência de matrizes conhecidas de dimensão $m \times m$. Na equação de observação, o vetor r -dimensional de observações, $\mathbf{y}_t$, está relacionado ao vetor de estados, $\mathbf{x}_t$, por meio de uma função determinística $h_t(\cdot)$, chamada de função de observação. Essa relação é representada por:

$$\mathbf{y}_t = h_t(\mathbf{x}_t, \mathbf{v}_t) \quad t = 1, 2, \dots, n$$

em que $\mathbf{v}_t \sim \mathcal{N}_r(\mathbf{0}, \mathbf{R}_t)$ e $(\mathbf{R}_t)_{t \geq 1}$ é uma sequência de matrizes conhecidas de dimensão $r \times r$. O modelo fica completo com a especificação do estado inicial, que está condicionado a alguma informação inicial D_0 ,

$$\mathbf{x}_0 | D_0 \sim \mathcal{N}_m(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$

em que $\boldsymbol{\mu}_0$ e $\boldsymbol{\Sigma}_0$ representam essa informação inicial. As duas sequências de erros $(\mathbf{v}_t)_{t \geq 1}$ e $(\mathbf{w}_t)_{t \geq 1}$ são internamente e mutuamente independentes. Além disso, assume-se que $\mathbf{x}_0 | D_0$ é independente das duas sequências de erros.

As funções vetoriais $h_t(\cdot)$ e $\mathcal{M}_t(\cdot)$ são conhecidas e podem variar no tempo. Além disso, podem depender de uma estrutura de parâmetros desconhecidos que não variam no tempo. Quanto aos erros, eles podem ser inseridos no modelo na forma aditiva ou multiplicativa. O caso aditivo é o mais comum nas aplicações e a forma multiplicativa pode ser reformulada na forma aditiva usando transformação de variável.

Os modelos de espaço de estados não lineares aditivos podem ser especificados na forma de distribuições de probabilidade condicionais chamadas de densidade de transição e densidade de observação,

$$\mathbf{x}_t | \mathbf{x}_{t-1} \sim \mathcal{N}_m(\mathcal{M}_t(\mathbf{x}_{t-1}), \mathbf{Q}_t)$$

e

$$\mathbf{y}_t | \mathbf{x}_t \sim \mathcal{N}_r(h_t(\mathbf{x}_t), \mathbf{R}_t),$$

respectivamente.

2.3 Modelos de Espaço de Estados Não Lineares e Não Normais

Com o objetivo de dar uma maior flexibilidade ao modelo de espaço de estados, outras distribuições podem ser atribuídas aos erros do processo e das observações. A estrutura do modelo de espaço de estados não linear e não gaussiano é dada por

$$\mathbf{x}_t = \mathcal{M}_t(\mathbf{x}_{t-1}, \mathbf{w}_t), \quad t = 1, 2, \dots \quad (2.4)$$

$$\mathbf{y}_t = h_t(\mathbf{x}_t, \mathbf{v}_t), \quad t = 1, 2, \dots, n \quad (2.5)$$

$$\mathbf{x}_0 | D_0 \sim p_m(\cdot). \quad (2.6)$$

A equação de estados (2.4) descreve a evolução do vetor de estados m -dimensional $\mathbf{x}_t$ por meio de uma função de transição $\mathcal{M}_t(\cdot)$ perturbada por uma sequência de erros $(\mathbf{w}_t)_{t \geq 1}$ não gaussianos. A equação de observação (2.5) relaciona o vetor de observações r -dimensional, $\mathbf{y}_t$, ao vetor de estados não observados, $\mathbf{x}_t$, por meio da função de observação $h_t(\cdot)$. Essa relação pode ser perturbada por uma sequência de erros $(\mathbf{v}_t)_{t \geq 1}$ não gaussianos. O estado inicial (2.6) está condicionado a alguma informação inicial D_0 , possui uma distribuição multivariada não gaussiana $p_m(\cdot)$. Além disso, assume-se que as duas sequências de erros $(\mathbf{v}_t)_{t \geq 1}$ e $(\mathbf{w}_t)_{t \geq 1}$ são internamente e mutuamente independentes, sendo também independentes da informação inicial $\mathbf{x}_0 | D_0$. Diversos autores tem trabalhado com essa descrição mais geral de modelos de espaço de estados, entre eles, [Harvey \(1990\)](#), [Gordon *et al.* \(1993\)](#), [Kitagawa \(1996\)](#), [Carlin *et al.* \(1992\)](#) e [Tanizaki \(2001\)](#).

No próximo capítulo introduzimos um modelo de espaço de estados não linear denominado modelo de produção excedente e também será apresentado as principais funções de produção e suas medidas para a biomassa e captura associadas a produção máxima sustentável.

Capítulo 3

Modelos de Produção Excedente

Modelos de produção excedente são modelos que buscam fornecer recomendações para a exploração sustentável de algum recurso biológico e são muito utilizados na área pesqueira. A produção excedente é definida como a quantidade de biomassa produzida além da necessária para a reposição no meio ambiente. Em muitos casos, a biomassa de um sistema biológico não pode ser observada diretamente. Entretanto, a biomassa pode ser mensurada indiretamente por meio de quantidades observáveis, tais como, captura e índice de abundância.

No contexto de modelos de espaço de estados, a biomassa representa um estado da população no tempo, geralmente medido em anos, e a dinâmica dos estados é definida por uma função determinística chamada de função de crescimento ou função de produção. Algumas dessas funções são usadas nos modelos de produção propostos por [Schaefer \(1954\)](#), [Pella e Tomlinson \(1969\)](#) e [Fox \(1970\)](#). Com esses modelos, busca-se estimar algumas quantidades como: a produção máxima sustentável (MSY, do inglês *Maximum Sustainable Yield*), a biomassa associada a produção máxima sustentável, captura associada a produção máxima sustentável e outras medidas, que podem dar base científica para as recomendações de exploração sustentável do meio ambiente.

3.1 Modelos de Produção

Os modelos de produção buscam descrever a dinâmica da biomassa (x_t) com base em uma equação sequencial em tempo discreto. Nessa equação, a biomassa no tempo $t+1$ depende da biomassa no tempo t , da biomassa produzida entre o tempo t e o tempo $t+1$, e também da biomassa que é capturada nesse intervalo de tempo. Essa equação é expressa da seguinte forma:

$$x_{t+1} = x_t + G(x_t) - H_t, \quad x_t \geq 0 \text{ e } H_t \geq 0,$$

em que, H_t representa a captura total observada ao longo do tempo t e $G(x_t) = \frac{dx_t}{dt}$ é chamada de função de crescimento ou função de produção. A função de crescimento é uma equação diferencial ordinária (ODE) que representa a taxa de variação da biomassa em função do tempo em um estoque ([Pitcher e Hart, 1982](#)). Os modelos de produção excedente variam de acordo com a escolha da função de produção. Esta pode ser escolhida com base em suposições sobre o comportamento do crescimento da biomassa

da população em estudo. Um dos principais modelos utilizados é o modelo proposto por [Schaefer \(1954\)](#), o qual detalharemos na próxima seção.

3.1.1 Modelo de Produção Logístico

O modelo de produção logístico é também conhecido como modelo de crescimento de [Schaefer \(1954\)](#) e é dado por:

$$x_{t+1} = x_t + rx_t \left(1 - \frac{x_t}{K}\right) - H_t, \quad r > 0, K > 0, \text{ para } t = 0, 1, 2, \dots$$

O modelo de produção logístico tem uma função de crescimento que possui dois parâmetros que são: a taxa intrínseca de crescimento, r , e a capacidade de carga que o ambiente pode suportar, K . Além disso, a quantidade H_t representa a captura total durante o ano t .

A função de crescimento na equação (3.1) é uma função quadrática da biomassa e seu gráfico é apresentado na Figura 3.1.

$$G(x_t) = \frac{dx_t}{dt} = rx_t \left(1 - \frac{x_t}{K}\right). \quad (3.1)$$

Figura 3.1: Função de produção do modelo de Schaefer.

Note que a equação (3.1) possui duas raízes, ou seja, dois pontos em que a taxa de crescimento da população é nula. Estes pontos são chamados de pontos de equilíbrio e ocorrem quando, em determinado tempo, a biomassa é igual a zero ($x_t = 0$) ou quando a biomassa da população é igual a carga máxima que o ambiente pode suportar ($x_t = K$). A partir da equação (3.1) podemos calcular as principais quantidades de interesse e, para isso, precisamos diferenciá-la com respeito a x_t ,

$$\frac{dG(x_t)}{dx_t} = r - \frac{2rx_t}{K}$$

igualar a zero e isolar x_t . Deste modo, podemos obter a *Biomassa associada à produção máxima sustentável* (B_{MSY}):

$$B_{MSY} = \frac{K}{2}. \quad (3.2)$$

Assim, a produção máxima sustentável é obtida quando o tamanho do estoque é exatamente igual a metade da capacidade máxima que o ambiente pode suportar.

Substituindo o resultado (3.2) na equação (3.1) obtemos a *Produção Máxima Sustentável*, que denotaremos por MSY (MSY do inglês *Maximum Sustainable Yield*),

$$MSY = \frac{rK}{4}.$$

Dividindo MSY por B_{MSY} obtemos a *taxa de captura associada a produção máxima sustentável* (H_{MSY}),

$$H_{MSY} = \frac{r}{2}.$$

Para calcular a biomassa da população em determinado tempo t precisamos resolver a equação diferencial (3.1) da seguinte forma:

$$\int_{x_0}^{x_t} \frac{K}{x_t(K - x_t)} dx_t = \int_0^t r dt.$$

A solução final é dada por,

$$x_t = \frac{K}{1 + \frac{K-x_0}{x_0} e^{-rt}} \quad , \quad (3.3)$$

em que x_0 é a biomassa no estado inicial. A função (3.3) é conhecida como função logística e seu gráfico está representado na Figura 3.2.

Figura 3.2: Biomassa da população quando $x_0 < K$, $x_0 = K$, e $x_0 > K$

3.1.2 Outros Modelos de Produção

Na literatura existem outras alternativas ao modelo de produção logístico e uma delas é o modelo de Fox (1970), com função de produção,

$$G(x_t) = rx_t \left(1 - \log \left(\frac{x_t}{K} \right) \right),$$

denominada função de produção de Gompertz. Esta função de produção é assimétrica à direita e, deste modo, o máximo crescimento sustentável da população ocorre abaixo da metade da capacidade de carga que o ambiente pode suportar, ou seja, $MSY < \frac{K}{2}$. As medidas de interesse usando o modelo de Fox são dadas por:

i) *Produção Máxima Sustentável*

$$MSY = \frac{r \times K}{e \log(K)}.$$

ii) *Biomassa associada a produção máxima sustentável*

$$B_{MSY} = \frac{K}{e}.$$

iii) *Taxa de captura associada a produção máxima sustentável*

$$H_{MSY} = \frac{r}{\log(K)}.$$

Uma outra alternativa mais geral é o modelo de produção de Pella e Tomlinson (1969). Esse modelo usa a seguinte função de produção:

$$G(x_t) = r \frac{x_t}{\phi} \left(1 - \frac{x_t}{K} \right)^\phi, \quad \text{para } \phi > 0. \quad (3.4)$$

Note que, se $\phi = 1$, a função de produção (3.4) se reduz a função de produção de Verhulst, (3.1), e no limite, quando $\phi \rightarrow 0$, possui como caso especial a função de produção de Gompertz. As medidas de interesse usando o modelo de Pella e Tomlinson são dadas por:

i) *Produção Máxima Sustentável*

$$MSY = r \times K \times (1 + \phi)^{-\frac{1+\phi}{\phi}}.$$

ii) *Biomassa associada à produção máxima sustentável*

$$B_{MSY} = K \times (1 + \phi)^{-\frac{1}{\phi}}.$$

iii) *Taxa de captura associada à produção máxima sustentável*

$$H_{MSY} = \frac{r}{\phi + 1}.$$

3.2 Modelos de Produção Excedente vistos como Modelos de Espaço de Estados

Os modelos de produção excedente abordados aqui necessitam de duas séries temporais que são: série temporal da captura, $H_{t=1:N}$, e a série temporal dos índices de abundância (ou índice relativo de biomassa), $y_{t=1:N}$. Os índices de abundância, na prática, são frequentemente referidos como captura por unidade de esforço (CPUE) e são geralmente assumidos proporcionais a biomassa no tempo corrente multiplicado por um parâmetro chamado de coeficiente de capturabilidade, $q > 0$, que é considerado constante no tempo, como segue,

$$y_t = qx_t, \quad \forall t \in \{1, 2, \dots, N\}.$$

Note que, cada índice de abundância em cada tempo t depende da biomassa que é comumente não observada, mas pode ser estimada. Na equação de observação, pode-se assumir que os erros das observações, v_t , são aditivos ou multiplicativos. Geralmente assume-se que $v_t \sim N(0, \sigma_v^2)$. Em Montenegro e Branco (2016) são atribuídas algumas distribuições pertencentes à classe de misturas de escala da distribuição normal assimétrica (SMSN, scale mixture of skew-normal, em inglês) para os erros das observações.

A equação de estados é não linear e é adicionada de erros, w_t , que podem ser aditivos,

$$x_t = x_{t-1} + G(x_{t-1}) - H_{t-1} + w_t$$

ou multiplicativos,

$$x_t = \left(x_{t-1} + G(x_{t-1}) - H_{t-1} \right) e^{w_t}.$$

Os erros dos estados são considerados seguindo $N(0, \sigma_w^2)$ e o tempo inicial, $t = 0$, é usualmente assumido $x_0 \sim N(\mu_0, \sigma_{w0}^2)$, em que μ_0 e σ_{w0}^2 representam algum conhecimento inicial.

Montenegro e Branco (2016) utilizam os métodos MCMC (Monte Carlo via Cadeias de Markov) para estimar os parâmetros do modelo de espaço de estados logístico, cujo os códigos foram implementados no OpenBugs. No entanto, segundo os autores, devido a complexidade de implementação dos modelos esta metodologia pode ser pouco atrativa aos usuários da área de ciência de pesca. Desse modo, o EnKF parece ser uma alternativa a ser investigada, já que este método é de fácil implementação e fornece estimativas para os parâmetros e estados em modelos de espaço de estados lineares e não lineares, segundo Katzfuss *et al.* (2016).

No próximo capítulo apresentaremos o Filtro de Kalman e o Filtro de Kalman Ensemble, utilizados para fazer inferência em modelos de espaço de estados.

Capítulo 4

Inferência em Modelos de Espaço de Estados via KF e EnKF

Um importante problema em modelagem com espaço de estados é a estimação e predição dos estados, que usualmente não podem ser observados diretamente. Para a estimação dos estados, várias abordagens podem ser encontradas na literatura. Uma abordagem usual é o método de *filtragem*, um procedimento recursivo para a estimação dos estados a partir de novas observações.

Um dos mais populares algoritmos de *filtragem* é o Filtro de Kalman (Kalman, 1960). Utilizamos a notação KF, do inglês *Kalman Filter*. Este filtro é construído para o modelo de espaço de estados linear gaussiano, e portanto, sua aplicação está limitada a fenômenos com pouca complexidade. Algumas de suas variações, tais como o Filtro de Kalman Extendido (Maybeck, 1979) e o Filtro de Kalman Ensemble (Evensen, 1994) possuem maior aplicabilidade, pois permitem fazer inferências em modelos mais flexíveis. Por exemplo, não lineares e não gaussianos. Estes algoritmos utilizam ideias similares ao KF, no entanto, usam aproximações diferentes para as distribuições a *posteriori*, uma vez que os sistemas não lineares geralmente produzem algoritmos intratáveis analiticamente.

De modo geral, nos métodos de *filtragem* a estimação do estado $\mathbf{x}_t$ é feita calculando a distribuição de *filtragem* no tempo t , que é definida como a distribuição a *posteriori* de $\mathbf{x}_t$ condicionado a todas as observações $\mathbf{y}_{1:t} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t\}$ obtidas até o tempo t , isto é, $p(\mathbf{x}_t | \mathbf{y}_{1:t})$. A obtenção da densidade de *filtragem* no tempo t é composta por dois passos, que são: predição e atualização. O passo de predição consiste na obtenção das distribuições preditivas um passo a frente dos estados e das observações e, o passo de atualização consiste na obtenção da distribuição de *filtragem*.

No passo de predição é necessário o conhecimento da distribuição de *filtragem* no tempo $t - 1$, ou seja $p(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1})$, para calcular a distribuição preditiva um passo à frente dos estados,

$$p(\mathbf{x}_t | \mathbf{y}_{1:t-1}) = \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1}) d\mathbf{x}_{t-1}. \quad (4.1)$$

Usando (4.1) calculamos a densidade preditiva um passo à frente das observações,

$$p(\mathbf{y}_t | \mathbf{y}_{1:t-1}) = \int p(\mathbf{y}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{y}_{1:t-1}) d\mathbf{x}_t. \quad (4.2)$$

A atualização é obtida incorporando a observação no tempo t , em termos da verossimilhança $p(\mathbf{y}_t|\mathbf{x}_t)$, ao conhecimento *a priori* de $\mathbf{x}_t$ dado por (4.1) por meio do teorema de Bayes. Portanto, a distribuição de *filtragem* no tempo t é dada por

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})}. \quad (4.3)$$

Essa densidade possui toda a informação sobre $\mathbf{x}_t$ que está contida nas observações $\mathbf{y}_{1:t}$ e na distribuição *a priori* de $\mathbf{x}_t$.

Deste modo, a inferência sobre os estados pode ser feita recursivamente por meio das equações de predição (4.1) e (4.2), e da equação de *filtragem* ou atualização (4.3). Essas equações, alternadamente, estimam as densidades do estado corrente e estado predito um passo a frente, ambas condicionadas nas observações. Essa forma de estimar os estados é chamada de Filtragem Bayesiana ou Estimção Recursiva Bayesiana (Ho e Lee, 1964).

As integrais em (4.1) e (4.2) podem ser intratáveis analiticamente, o que torna muito difícil calcular a densidade de *filtragem*. Uma exceção é o modelo de espaço de estados linear gaussiano, neste caso, obtém-se o Filtro de Kalman.

4.1 Filtro de Kalman (KF)

O Filtro de Kalman (Kalman, 1960) é um algoritmo que permite fazer inferência em um modelo de espaço de estados linear gaussiano. Em essência, esse algoritmo fornece uma maneira de calcular recursivamente os momentos das distribuições de predição e de filtragem em dois passos: predição e atualização. A forma recursiva é facilitada pelos resultados da distribuição normal multivariada, em que, para um vetor conjunto gaussiano $(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_t, \mathbf{y}_1, \dots, \mathbf{y}_t)$, as distribuições marginais e condicionais são também gaussianas em qualquer tempo $t = 1, 2, \dots, N$, podendo ser completamente determinadas por suas médias e variâncias.

Assim, considerando o modelo de espaço de estados linear gaussiano introduzido no Capítulo 2 nas equações (2.2) e (2.3) e usando as propriedades da distribuição normal multivariada podemos obter o passo de predição calculando os momentos da distribuição preditiva um passo à frente dos estados que é

$$\mathbf{x}_t|\mathbf{y}_{1:t-1} \sim \mathcal{N}_m(\tilde{\boldsymbol{\mu}}_t = \mathbf{M}_t\hat{\boldsymbol{\mu}}_{t-1}, \tilde{\boldsymbol{\Sigma}}_t = \mathbf{M}_t\hat{\boldsymbol{\Sigma}}_{t-1}\mathbf{M}_t' + \mathbf{Q}_t). \quad (4.4)$$

O vetor de médias e a matriz de covariâncias podem ser calculados usando as propriedades condicionais da média e da variância, como segue,

$$\tilde{\boldsymbol{\mu}}_t = E[\mathbf{x}_t|\mathbf{y}_{1:t-1}] = E[E[\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_{1:t-1}]|\mathbf{y}_{1:t-1}] = E[\mathbf{M}_t\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1}] = \mathbf{M}_t\hat{\boldsymbol{\mu}}_{t-1} \quad (4.5)$$

e

$$\begin{aligned} \tilde{\boldsymbol{\Sigma}}_t &= Var[\mathbf{x}_t|\mathbf{y}_{1:t-1}] = Var[E[\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_{1:t-1}]|\mathbf{y}_{1:t-1}] + E[Var[\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_{1:t-1}]|\mathbf{y}_{1:t-1}] \\ &= Var[\mathbf{M}_t\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1}] + E[\mathbf{Q}_t|\mathbf{y}_{1:t-1}] \\ &= \mathbf{M}_t\hat{\boldsymbol{\Sigma}}_{t-1}\mathbf{M}_t' + \mathbf{Q}_t. \end{aligned} \quad (4.6)$$

em que $\hat{\boldsymbol{\mu}}_{t-1}$ e $\hat{\boldsymbol{\Sigma}}_{t-1}$ representam os momentos da densidade de *filtragem* no tempo $t-1$. Também podemos calcular os momentos da distribuição preditiva um passo à frente das observações da seguinte forma,

$$E[\mathbf{y}_t | \mathbf{y}_{1:t-1}] = E[E[\mathbf{y}_t | \mathbf{x}_t, \mathbf{y}_{1:t-1}] | \mathbf{y}_{1:t-1}] = E[\mathbf{H}_t \mathbf{x}_t | \mathbf{y}_{1:t-1}] = \mathbf{H}_t \tilde{\boldsymbol{\mu}}_t$$

e

$$\begin{aligned} Var[\mathbf{y}_t | \mathbf{y}_{1:t-1}] &= Var[E[\mathbf{y}_t | \mathbf{x}_t, \mathbf{y}_{1:t-1}] | \mathbf{y}_{1:t-1}] + E[Var[\mathbf{y}_t | \mathbf{x}_t, \mathbf{y}_{1:t-1}] | \mathbf{y}_{1:t-1}] \\ &= Var[\mathbf{H}_t \mathbf{x}_t | \mathbf{y}_{1:t-1}] + E[\mathbf{R}_t | \mathbf{y}_{1:t-1}] \\ &= \mathbf{H}_t \tilde{\boldsymbol{\Sigma}}_t \mathbf{H}_t' + \mathbf{R}_t. \end{aligned}$$

Resultando em,

$$\mathbf{y}_t | \mathbf{y}_{1:t-1} \sim \mathcal{N}_r(\mathbf{H}_t \tilde{\boldsymbol{\mu}}_t, \mathbf{H}_t \tilde{\boldsymbol{\Sigma}}_t \mathbf{H}_t' + \mathbf{R}_t). \quad (4.7)$$

Já o passo de atualização envolve o cálculo dos momentos da distribuição de *filtragem* no tempo t , a qual podemos derivar a partir da distribuição conjunta de $\mathbf{x}_t$ e $\mathbf{y}_t$ dado $\mathbf{y}_{1:t-1}$, sendo dada pela seguinte distribuição normal multivariada,

$$\begin{pmatrix} \mathbf{x}_t \\ \mathbf{y}_t \end{pmatrix} \Big| \mathbf{y}_{1:t-1} \sim \mathcal{N}_{r+m} \left(\begin{pmatrix} \tilde{\boldsymbol{\mu}}_t \\ \mathbf{H}_t \tilde{\boldsymbol{\mu}}_t \end{pmatrix}, \begin{pmatrix} \tilde{\boldsymbol{\Sigma}}_t & \tilde{\boldsymbol{\Sigma}}_t \mathbf{H}_t' \\ \mathbf{H}_t \tilde{\boldsymbol{\Sigma}}_t & \mathbf{H}_t \tilde{\boldsymbol{\Sigma}}_t \mathbf{H}_t' + \mathbf{R}_t \end{pmatrix} \right).$$

Usando a propriedade de condicionamento da distribuição normal multivariada (Mardia *et al.*), a densidade preditiva (4.4) é então atualizada tendo em conta a nova observação $\mathbf{y}_t$, resultando na densidade de *filtragem* no tempo t ,

$$\mathbf{x}_t | \mathbf{y}_{1:t} \sim \mathcal{N}_m(\hat{\boldsymbol{\mu}}_t, \hat{\boldsymbol{\Sigma}}_t) \quad (4.8)$$

cujos os momentos são,

$$\hat{\boldsymbol{\mu}}_t = \tilde{\boldsymbol{\mu}}_t + \mathbf{K}_t(\mathbf{y}_t - \mathbf{H}_t \tilde{\boldsymbol{\mu}}_t) \quad (4.9)$$

$$\hat{\boldsymbol{\Sigma}}_t = (\mathbf{I}_m - \mathbf{K}_t \mathbf{H}_t) \tilde{\boldsymbol{\Sigma}}_t \quad (4.10)$$

e $\mathbf{K}_t = \tilde{\boldsymbol{\Sigma}}_t \mathbf{H}_t' (\mathbf{H}_t \tilde{\boldsymbol{\Sigma}}_t \mathbf{H}_t' + \mathbf{R}_t)^{-1}$ é uma matriz $m \times r$ chamada de ganho de Kalman.

As equações (4.9) e (4.10) são chamadas de equações de atualização e são usadas em conjunto com as equações de predição (4.5) e (4.6) em um processo recursivo. Em qualquer tempo t , elas fornecem o valor que contém toda informação necessária para fazer boas previsões. Entretanto, para modelos de espaço de estados com grandes dimensões o custo computacional pode ser alto, já que é necessário inverter uma matriz de alta dimensão no ganho de Kalman.

Com base em um conhecimento prévio, os valores iniciais de $\hat{\boldsymbol{\mu}}_0$ e $\hat{\boldsymbol{\Sigma}}_0$ são especificados e a partir deles as equações podem ser atualizadas sucessivamente calculando o vetor de estados com cada nova observação. No *Algoritmo 1* apresentamos uma forma geral de fazer inferência exata em modelos de espaço de estados lineares normais.

Algoritmo 1: FILTRO DE KALMAN

- 1 **Passo inicial:** Inicialização do processo recursivo com $\hat{\mu}_0$ e $\hat{\Sigma}_0$ fixados.
- 2 **Passo de predição:** Calcule o vetor de médias e a matriz de covariâncias da distribuição predita dos estados $\mathbf{x}_t | \mathbf{y}_{1:t-1} \sim \mathcal{N}_m(\tilde{\mu}_t, \tilde{\Sigma}_t)$ fazendo,

$$\begin{aligned}\tilde{\mu}_t &= \mathbf{M}_t \hat{\mu}_{t-1} \\ \tilde{\Sigma}_t &= \mathbf{M}_t \hat{\Sigma}_{t-1} \mathbf{M}_t' + \mathbf{Q}_t\end{aligned}$$

- 3 **Passo de atualização:** Calcule o vetor de médias e a matriz de covariâncias da distribuição de *filtragem* exata $\mathbf{x}_t | \mathbf{y}_{1:t} \sim \mathcal{N}_m(\hat{\mu}_t, \hat{\Sigma}_t)$ fazendo,

$$\begin{aligned}\hat{\mu}_t &= \tilde{\mu}_t + \mathbf{K}_t(\mathbf{y}_t - \mathbf{H}_t \tilde{\mu}_t) \\ \hat{\Sigma}_t &= (\mathbf{I}_m - \mathbf{K}_t \mathbf{H}_t) \tilde{\Sigma}_t\end{aligned}$$

em que $\mathbf{K}_t = \tilde{\Sigma}_t \mathbf{H}_t' (\mathbf{H}_t \tilde{\Sigma}_t \mathbf{H}_t' + \mathbf{R}_t)^{-1}$.

- 4 Repetir os passos de predição e atualização para $t = 1, 2, \dots, N$.
-

Quando o modelo de espaço de estados é não linear, geralmente, não é possível produzir algoritmos analiticamente tratáveis e alguma forma de aproximação deve ser usada. Exemplos dessas aproximações são: o Filtro de Kalman Extendido e o Filtro de Kalman *Ensemble*. O Filtro de Kalman Extendido foi uma das primeiras generalizações do Filtro de Kalman para modelos não lineares. Este método lineariza as equações não lineares dos estados e das observações para a utilização do Filtro de Kalman padrão. Essa linearização leva a atualizações precárias das estimativas à medida em que o grau de não linearidade aumenta. No Filtro de Kalman *Ensemble* não é necessário fazer essa linearização, e usa-se outro tipo de aproximação. A formulação teórica do Filtro de Kalman *Ensemble* será descrita a seguir.

4.2 Filtro de Kalman Ensemble (EnKF)

O Filtro de Kalman Ensemble é um método de Monte Carlo sequencial que fornece uma alternativa ao filtro de Kalman tradicional para a estimação dos estados. A principal diferença entre os dois métodos é que o KF considera a distribuição exata dos estados (condicional aos parâmetros) e o EnKF usa uma amostra ou conjunto (*ensemble*, do inglês) para representar a distribuição dos estados. Essa amostra é armazenada, atualizada e propagada no tempo a cada nova observação.

A representação amostral é uma forma de redução de dimensão, o que leva a viabilidade computacional mesmo para modelos de espaço de estado com dimensão na ordem de milhões. Sendo, por isso, amplamente utilizado em aplicações de assimilação de dados com dezenas de milhões de dimensões. Algumas referências do uso desse algoritmo [Evensen e Van Leeuwen \(1996\)](#), [Houtekamer e Mitchell \(1998\)](#); [Mitchell e Houtekamer \(2000\)](#), [Evensen \(1994, 1997, 2009b\)](#).

4.2.1 Estimação dos Estados

A estimação dos estados via EnKF consiste de um passo de predição e um de atualização para todo tempo t , de forma similar ao Filtro de Kalman. Vamos derivar o algoritmo considerando o modelo de espaço de estados linear gaussiano introduzido no Capítulo 2 nas equações (2.2) e (2.3).

O algoritmo é iniciado gerando uma amostra de tamanho L da distribuição de *filtragem* $\hat{\mathbf{x}}_0^{(i)} \sim \mathcal{N}_m(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0)$ no tempo $t = 0$, em que $\hat{\boldsymbol{\mu}}_0$ e $\hat{\boldsymbol{\Sigma}}_0$ são especificados de acordo com um conhecimento prévio. De (4.8), temos que no tempo $t - 1$, a distribuição de *filtragem* será $\hat{\mathbf{x}}_{t-1}^{(i)} \sim \mathcal{N}_m(\hat{\boldsymbol{\mu}}_{t-1}, \hat{\boldsymbol{\Sigma}}_{t-1})$. Assim, a partir de uma amostra $\hat{\mathbf{x}}_{t-1}^{(1)}, \dots, \hat{\mathbf{x}}_{t-1}^{(L)}$ simulada da distribuição de *filtragem* no tempo $t - 1$ podemos obter uma amostra da distribuição de predição (4.4) pela simples aplicação dos membros da amostra da distribuição de *filtragem* na equação de evolução (2.2), ou seja,

$$\tilde{\mathbf{x}}_t^{(i)} = \mathbf{M}_t \hat{\mathbf{x}}_{t-1}^{(i)} + \mathbf{w}_t^{(i)}, \quad \mathbf{w}_t^{(i)} \sim \mathcal{N}_m(\mathbf{0}, \mathbf{Q}_t) \quad i = 1, 2, \dots, L.$$

O passo de atualização é feito com base no conjunto predito, $\tilde{\mathbf{x}}_t^{(1)}, \dots, \tilde{\mathbf{x}}_t^{(L)}$ e na observação real no tempo t , $\mathbf{y}_t$, como segue,

$$\hat{\mathbf{x}}_t^{(i)} = \tilde{\mathbf{x}}_t^{(i)} + \hat{\mathbf{K}}_t(\mathbf{y}_t - \tilde{\mathbf{y}}_t^{(i)})$$

em que $\tilde{\mathbf{y}}_t^{(i)} = \mathbf{H}_t \tilde{\mathbf{x}}_t^{(i)} - \mathbf{v}^{(i)}$, $\mathbf{v}^{(i)} \sim \mathcal{N}_r(\mathbf{0}, \mathbf{R}_t)$ e $\hat{\mathbf{K}}_t$ é uma aproximação do ganho de Kalman em (4.9). A expressão do ganho de Kalman amostral é dada por

$$\hat{\mathbf{K}}_t = \tilde{\mathbf{P}}_t \mathbf{H}_t' (\mathbf{H}_t \tilde{\mathbf{P}}_t \mathbf{H}_t' + \mathbf{R}_t)^{-1},$$

em que $\tilde{\mathbf{P}}_t$ é a matriz de covariâncias amostral de $\tilde{\mathbf{x}}_t^{(1)}, \dots, \tilde{\mathbf{x}}_t^{(L)}$. Pode-se mostrar que $\hat{\mathbf{x}}_t^{(i)}$ segue $\mathcal{N}_m(\hat{\boldsymbol{\mu}}_t, \hat{\boldsymbol{\Sigma}}_t)$. As demonstrações são apresentadas em A.1.

No Algoritmo 2 apresentamos a estimação sequencial dos estados via EnKF,

Algoritmo 2: ENKF ESTOCÁSTICO

1 **Passo inicial:** Gere uma amostra inicial $\hat{\mathbf{x}}_0^{(1)}, \hat{\mathbf{x}}_0^{(2)}, \dots, \hat{\mathbf{x}}_0^{(L)}$ de $\hat{\mathbf{x}}_0^{(i)} \sim \mathcal{N}_m(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0)$.

2 **Passo de predição:** Gere uma amostra da distribuição preditiva

$\mathbf{x}_t | \mathbf{y}_{1:t-1} \sim \mathcal{N}_m(\tilde{\boldsymbol{\mu}}_t, \tilde{\boldsymbol{\Sigma}}_t)$ usando a expressão,

$$\tilde{\mathbf{x}}_t^{(i)} = \mathbf{M}_t \hat{\mathbf{x}}_{t-1}^{(i)} + \mathbf{w}_t^{(i)} \quad \text{com } \mathbf{w}_t^{(i)} \sim \mathcal{N}_m(\mathbf{0}, \mathbf{Q}_t).$$

3 **Passo de atualização:** Gere uma amostra da distribuição de *filtragem*

$\mathbf{x}_t | \mathbf{y}_{1:t} \sim \mathcal{N}_m(\hat{\boldsymbol{\mu}}_t, \hat{\boldsymbol{\Sigma}}_t)$ usando a expressão,

$$\hat{\mathbf{x}}_t^{(i)} = \tilde{\mathbf{x}}_t^{(i)} + \hat{\mathbf{K}}_t(\mathbf{y}_t - \tilde{\mathbf{y}}_t^{(i)})$$

em que $\hat{\mathbf{K}}_t = \tilde{\mathbf{P}}_t \mathbf{H}_t' (\mathbf{H}_t \tilde{\mathbf{P}}_t \mathbf{H}_t' + \mathbf{R}_t)^{-1}$.

4 Repetir os passos de predição e atualização para $t = 1, 2, \dots, N$.

O conjunto inicial $\hat{\mathbf{x}}_0^{(1)}, \hat{\mathbf{x}}_0^{(2)}, \dots, \hat{\mathbf{x}}_0^{(L)}$ deve ser idealmente escolhido com base em algum conhecimento prévio, no entanto, uma má especificação do conjunto inicial geralmente não influencia muito as estimativas ao longo do tempo (Evensen, 2003). Contudo, ao usar um conjunto inicial não informativo pode ser necessário fazer o tamanho da amostra ensemble (L) maior que 100 (Evensen, 2009a), uma vez que, do teorema central do limite, a acurácia das estimativas aumenta quando $L \rightarrow \infty$, já que o EnKF converge para o KF. A demonstração dessa convergência assintótica pode ser vista em Butala *et al.* (2008).

A inferência em um modelo não linear aditivo é feita trocando os operadores lineares $\mathbf{M}_t \hat{\mathbf{x}}_{t-1}^{(i)}$ e $\mathbf{H}_t \hat{\mathbf{x}}_t^{(i)}$ pelos operadores não lineares $\mathcal{M}_t(\hat{\mathbf{x}}_{t-1}^{(i)})$ e $h_t(\hat{\mathbf{x}}_t^{(i)})$ no Algoritmo 2. Segundo Evensen (1994), o EnKF é robusto a desvios da suposição de linearidade, de modo a fornecer inferências acuradas para muitos modelos não lineares. Além disso, segundo o autor, o algoritmo pode ser aplicado, com cautela, em modelos não gaussianos.

Alguns modelos de espaço de estados contêm parâmetros desconhecidos que precisam ser estimados em conjunto com os estados. O problema de estimação desses parâmetros tem sido abordado por alguns poucos autores como, por exemplo, Anderson (2001) e Stroud *et al.* (2016). Apresentaremos a seguir os métodos propostos por esses autores para a estimação dos parâmetros.

4.2.2 Estimação dos parâmetros

Considere o seguinte modelo de espaço de estados gaussiano,

$$\begin{aligned} \mathbf{x}_t &= \mathcal{M}_t(\mathbf{x}_{t-1}, \boldsymbol{\gamma}) + \mathbf{w}_t, & \mathbf{w}_t &\sim \mathcal{N}_m(\mathbf{0}, \mathbf{Q}_t(\boldsymbol{\theta})) \\ \mathbf{y}_t &= \mathbf{H}_t(\boldsymbol{\theta})\mathbf{x}_t + \mathbf{v}_t, & \mathbf{v}_t &\sim \mathcal{N}_r(\mathbf{0}, \mathbf{R}_t(\boldsymbol{\theta})) \end{aligned}$$

com $\mathbf{y}_t$ representando o vetor de observações de dimensão $r \times 1$ e $\mathbf{x}_t$ um vetor de estados de dimensão $m \times 1$ para $t = 1, 2, \dots$, em que $\mathbf{H}_t, \mathbf{R}_t, \mathbf{Q}_t$ são matrizes que podem depender de $\boldsymbol{\theta}$, um vetor de parâmetros desconhecidos e invariante no tempo e $\mathcal{M}_t(\cdot)$ é a evolução dos estados que pode depender de um outro vetor de parâmetros desconhecidos $\boldsymbol{\gamma}$ que também é invariante no tempo. O modelo fica totalmente especificado com a atribuição de distribuições *a priori* para o estado inicial $\mathbf{x}_0 | \boldsymbol{\theta}, \boldsymbol{\gamma}$ e para os vetores de parâmetros $\boldsymbol{\theta}$ e $\boldsymbol{\gamma}$.

Embora muitos trabalhos tenham sido feitos para melhorar a estimativa dos estados via EnKF, existem poucos trabalhos com foco na estimação dos parâmetros do modelo. A seguir, apresentaremos alguns métodos de estimação que podem ser combinados para a estimação dos dois vetores de parâmetros.

EnKF Aumentado

O EnKF Aumentado (Anderson, 2001) busca estimar simultaneamente os estados e os parâmetros $\boldsymbol{\gamma}$ contidos na função de evolução. Este método é definido como a simples inclusão desse vetor de parâmetros no vetor de estados. O método é construído com base na suposição de linearidade das equações das observações e dos estados e na suposição de normalidade para a distribuição *a priori* do vetor $\boldsymbol{\gamma}$, para assim, poder obter o EnKF padrão para o vetor aumentado $\mathbf{z}_t = [\mathbf{x}_t', \boldsymbol{\gamma}']'$. Ao adicionar a estrutura de parâmetros $\boldsymbol{\gamma}$ no vetor de estados a atualização será dada de forma linear.

O novo sistema de espaço de estados aumentado é dado por:

$$\begin{aligned} \mathbf{z}_t &= \mathcal{M}_{t-1}^*(\mathbf{z}_{t-1}) + \mathbf{w}_t^* \\ \mathbf{y}_t &= \mathbf{H}_t^* \mathbf{z}_t + \mathbf{v}_t \end{aligned}$$

em que $\mathcal{M}_{t-1}^*(\mathbf{z}_{t-1}) = \begin{bmatrix} \mathcal{M}_{t-1}(\mathbf{z}_{t-1}) \\ \gamma \end{bmatrix}$, $\mathbf{w}_t^* = [\mathbf{w}_t \ 0]'$, $\mathbf{H}_t^* = [\mathbf{H}_t \ 0]$, com 0 representando uma matriz de zeros com dimensões apropriadas. Pode-se verificar que o EnKF Aumentado pode ser derivado de forma similar ao EnKF padrão, usando a distribuição predita dos estados e das observações um passo a frente e a atualização do estado predito. Assim, no EnKF Aumentado essas distribuições são dadas por:

$$p(\mathbf{z}_t | \mathbf{y}_{1:t-1}) = \int p(\mathbf{x}_t | \mathbf{x}_{t-1}, \gamma) p(\mathbf{z}_{t-1} | \mathbf{y}_{1:t-1}) d\mathbf{x}_{t-1} \quad (4.11)$$

$$p(\mathbf{y}_t | \mathbf{y}_{1:t-1}) = \int p(\mathbf{y}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{y}_{1:t-1}) d\mathbf{x}_t \quad (4.12)$$

$$p(\mathbf{z}_t | \mathbf{y}_{1:t}) = \frac{p(\mathbf{z}_t, \mathbf{y}_t | \mathbf{y}_{1:t-1})}{p(\mathbf{y}_t | \mathbf{y}_{1:t-1})}. \quad (4.13)$$

A equação (4.11) refere-se a distribuição de predição um passo a frente do vetor $\mathbf{z}_t$, a equação (4.12) refere-se a distribuição de predição um passo a frente das observações e a equação (4.13) é a atualização *a posteriori* da equação (4.11). Essa atualização de $\mathbf{z}_t$, como no EnKF padrão, é baseada na observação corrente $\mathbf{y}_t$. As equações de atualização são construídas com base na suposição de linearidade da função de transição $\mathcal{M}(\cdot)$. As expressões das equações de atualização e predição do vetor $\mathbf{z}_t$ serão similares ao do vetor $\mathbf{x}_t$, porém com a presença da estrutura de parâmetros γ . Assim, é possível mostrar que a matriz de covariâncias da distribuição preditiva um passo a frente de $\mathbf{z}_t$ é dada por,

$$\tilde{\Sigma}_t^* = \begin{bmatrix} \tilde{\Sigma}_{\mathbf{x}_t \mathbf{x}_t} & \tilde{\Sigma}'_{\gamma \mathbf{x}_t} \\ \tilde{\Sigma}_{\gamma \mathbf{x}_t} & \tilde{\Sigma}_{\gamma \gamma} \end{bmatrix}$$

que é composta por quatro matrizes de covariâncias e covariâncias cruzadas de $\mathbf{x}_t$ e γ . O Ganho de Kalman, $\mathbf{K}_t^*$, agora é dado por,

$$\begin{aligned} \mathbf{K}_t^* &= \tilde{\Sigma}_t^* \mathbf{H}_t'^* (\mathbf{H}_t^* \tilde{\Sigma}_t^* \mathbf{H}_t'^* + \mathbf{R}_t)^{-1} \\ &= \begin{bmatrix} \tilde{\Sigma}_{\mathbf{x}_t \mathbf{x}_t} & \tilde{\Sigma}'_{\gamma \mathbf{x}_t} \\ \tilde{\Sigma}_{\gamma \mathbf{x}_t} & \tilde{\Sigma}_{\gamma \gamma} \end{bmatrix} \begin{bmatrix} \mathbf{H}_t \\ \mathbf{0} \end{bmatrix} \left(\begin{bmatrix} \mathbf{H}_t & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\Sigma}_{\mathbf{x}_t \mathbf{x}_t} & \tilde{\Sigma}'_{\gamma \mathbf{x}_t} \\ \tilde{\Sigma}_{\gamma \mathbf{x}_t} & \tilde{\Sigma}_{\gamma \gamma} \end{bmatrix} \begin{bmatrix} \mathbf{H}_t \\ \mathbf{0} \end{bmatrix} + \mathbf{R}_t \right)^{-1} \\ &= \begin{bmatrix} \tilde{\Sigma}_{\mathbf{x}_t \mathbf{x}_t} \mathbf{H}_t' (\mathbf{H}_t \tilde{\Sigma}_{\mathbf{x}_t \mathbf{x}_t} \mathbf{H}_t' + \mathbf{R}_t)^{-1} \\ \tilde{\Sigma}_{\gamma \mathbf{x}_t} \mathbf{H}_t' (\mathbf{H}_t \tilde{\Sigma}_{\mathbf{x}_t \mathbf{x}_t} \mathbf{H}_t' + \mathbf{R}_t)^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{K}_{\mathbf{x}_t} \\ \mathbf{K}_{\gamma} \end{bmatrix} \end{aligned}$$

Desses resultados seguem as equações para a atualização dos estados e dos parâmetros,

$$\begin{bmatrix} \hat{\boldsymbol{\mu}}_t \\ \hat{\gamma}_t \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{\mu}}_t \\ \tilde{\gamma}_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{K}_{\mathbf{x}_t} (\mathbf{y}_t - \mathbf{H}_t \tilde{\boldsymbol{\mu}}_t) \\ \mathbf{K}_{\gamma} (\mathbf{y}_t - \mathbf{H}_t \tilde{\boldsymbol{\mu}}_t) \end{bmatrix}.$$

O vetor de parâmetros γ é mantido invariante, mas aqui será indexado por t para indicar a atualização.

No *Algoritmo 3* apresentamos os passos para a estimação EnKF Aumentado.

Algoritmo 3: ESTIMANDO OS ESTADOS E O CONJUNTO DE PARÂMETROS γ .

- 1 **Passo de inicial:** O algoritmo é inicializado com a amostra *a priori* $\gamma^{(i)} \sim p(\gamma)$ e $\hat{\mathbf{x}}_0^{(i)} \sim \mathcal{N}_m(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0)$ para $i = 1, \dots, L$.
- 2 **Passo de predição:** O vetor de parâmetros $\gamma_{t-1}^{(i)}$ é mantido invariante, enquanto os estados $\tilde{\mathbf{x}}_t^{(i)}$ são obtidos por,

$$\tilde{\mathbf{x}}_t^{(i)} = \mathcal{M}_t\left(\hat{\mathbf{x}}_{t-1}^{(i)}, \gamma_{t-1}^{(i)}\right) + \mathbf{w}_t^{(i)}; \quad \mathbf{w}_t^{(i)} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_t)$$

- 3 **Passo de atualização:** Para uma nova observação, todos os membros $\mathbf{x}_{t-1}^{(i)}$ e $\gamma_{t-1}^{(i)}$, são atualizados com,

$$\hat{\mathbf{x}}_t^{(i)} = \tilde{\mathbf{x}}_t^{(i)} + \tilde{\mathbf{P}}_{\tilde{\mathbf{x}}_t, \tilde{\mathbf{y}}_t} \tilde{\mathbf{P}}_{\tilde{\mathbf{y}}_t, \tilde{\mathbf{y}}_t}^{-1} \left(\mathbf{y}_t - \tilde{\mathbf{y}}_t^{(i)} \right) \quad \text{e} \quad \gamma_t^{(i)} = \gamma_{t-1}^{(i)} + \tilde{\mathbf{P}}_{\gamma_{t-1}, \tilde{\mathbf{y}}_t} \tilde{\mathbf{P}}_{\tilde{\mathbf{y}}_t, \tilde{\mathbf{y}}_t}^{-1} \left(\mathbf{y}_t - \tilde{\mathbf{y}}_t^{(i)} \right),$$

em que $\tilde{\mathbf{y}}_t^{(i)} = \mathbf{H}_t \tilde{\mathbf{x}}_t^{(i)} + \mathbf{v}_t^{(i)}$, $\mathbf{v}_t^{(i)} \sim \mathcal{N}_r(\mathbf{0}, \mathbf{R}_t)$.

- 4 Repetir os passos de predição e atualização para $t = 1, 2, \dots, N$.
-

em que $\tilde{\mathbf{P}}_{\tilde{\mathbf{y}}_t, \tilde{\mathbf{y}}_t}$ é a matriz de covariâncias amostral das observações preditas, $\tilde{\mathbf{P}}_{\tilde{\mathbf{x}}_t, \tilde{\mathbf{y}}_t}$ é matriz de covariâncias cruzada amostral das observações preditas e dos estados preditos, e $\tilde{\mathbf{P}}_{\gamma_{t-1}, \tilde{\mathbf{y}}_t}$ a covariâncias cruzada amostral da estrutura de parâmetros e das observações preditas.

Estimação dos estados e do vetor de parâmetros θ

O método de estimação introduzido por [Stroud et al. \(2016\)](#) também é baseado na linearidade e normalidade do modelo de espaço de estados e estima o vetor de parâmetros θ em conjunto com os estados utilizando-se da decomposição da seguinte distribuição conjunta,

$$p(\mathbf{x}_t, \theta | \mathbf{y}_{1:t}) = p(\mathbf{x}_t | \theta, \mathbf{y}_{1:t}) p(\theta | \mathbf{y}_{1:t}).$$

Em que, $p(\mathbf{x}_t | \theta, \mathbf{y}_{1:t})$ é a distribuição de *filtragem* e $p(\theta | \mathbf{y}_{1:t})$ é a densidade marginal *a posteriori* dos parâmetros, que pode ser reescrita de forma recursiva como

$$p(\theta | \mathbf{y}_{1:t}) \propto p(\mathbf{y}_t | \theta, \mathbf{y}_{1:t-1}) p(\theta | \mathbf{y}_{1:t-1}).$$

Em que, o primeiro termo do produto representa a verossimilhança predita um passo a frente e o segundo termo representa a distribuição *a posteriori* de θ no tempo $t - 1$. Esta forma recursiva produz uma maneira de propagar a informação do parâmetro no tempo.

A função de verossimilhança predita um passo à frente $p(\mathbf{y}_t | \theta, \mathbf{y}_{1:t-1})$ em um modelo linear normal é apresentada em (4.7). No entanto, [Stroud et al. \(2016\)](#) propõe uma aproximação usando o vetor de médias e a matriz de covariâncias da representação amostral

para que a evolução da verossimilhança de um modelo de alta dimensão seja computacionalmente viável. Essa aproximação é dada por,

$$\hat{p}(\mathbf{y}_t | \boldsymbol{\theta}, \mathbf{y}_{1:t-1}) \propto |\hat{\mathbf{B}}_t(\boldsymbol{\theta})|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} \hat{\mathbf{e}}_t(\boldsymbol{\theta})' \hat{\mathbf{B}}_t(\boldsymbol{\theta})^{-1} \hat{\mathbf{e}}_t(\boldsymbol{\theta}) \right\} \quad (4.14)$$

em que $\hat{\mathbf{e}}_t(\boldsymbol{\theta}) = \mathbf{y}_t - \mathbf{H}_t(\boldsymbol{\theta})\tilde{\mathbf{a}}_t$ e $\hat{\mathbf{B}}_t(\boldsymbol{\theta})$ representa a matriz de covariâncias definida por,

$$\hat{\mathbf{B}}_t(\boldsymbol{\theta}) = \mathbf{H}_t(\boldsymbol{\theta})\tilde{\mathbf{P}}_t^f(\boldsymbol{\theta})\mathbf{H}_t(\boldsymbol{\theta})' + \mathbf{R}(\boldsymbol{\theta})$$

com $\tilde{\mathbf{P}}_t^f(\boldsymbol{\theta}) = \tilde{\mathbf{P}}_t + \mathbf{Q}_t(\boldsymbol{\theta})$. As quantidades $\tilde{\mathbf{a}}_t$ e $\tilde{\mathbf{P}}_t$ representam, respectivamente, o vetor de médias e a matriz de covariâncias amostral do conjunto gerado por $\tilde{\mathbf{x}}_t^{(pi)} = \mathcal{M}(\tilde{\mathbf{x}}_{t-1}^{(i)})$, sem a presença do ruído.

Geralmente a distribuição a *posteriori* $p(\boldsymbol{\theta} | \mathbf{y}_{1:t-1})$ não terá forma fechada, sendo assim necessário o uso de uma aproximação. As aproximações propostas são:

- i) **Aproximação pela distribuição normal:** Nesta abordagem a distribuição a *posteriori* é aproximada por uma distribuição normal com média $\mathbf{m}_t$ e matriz de covariâncias $\mathbf{C}_t$, como segue,

$$\hat{p}(\boldsymbol{\theta} | \mathbf{y}_{1:t}) \propto \exp \left\{ -\frac{1}{2} (\boldsymbol{\theta} - \mathbf{m}_t)' \mathbf{C}_t^{-1} (\boldsymbol{\theta} - \mathbf{m}_t) \right\}.$$

Assim, a distribuição a *posteriori* aproximada é proporcional ao produto da verossimilhança aproximada e da distribuição *priori* aproximada, que é,

$$\hat{p}(\boldsymbol{\theta} | \mathbf{y}_{1:t}) \propto \hat{p}(\mathbf{y}_t | \boldsymbol{\theta}, \mathbf{y}_{1:t}) \hat{p}(\boldsymbol{\theta} | \mathbf{y}_{1:t-1})$$

em que a distribuição $\hat{p}(\boldsymbol{\theta} | \mathbf{y}_{1:t-1})$ é normal com média $\mathbf{m}_{t-1}$ e matriz de covariâncias $\mathbf{C}_{t-1}$. A aproximação a *posteriori* é dada pela maximização de $\ell(\boldsymbol{\theta}) = \log(\hat{p}(\mathbf{y}_t | \boldsymbol{\theta}, \mathbf{y}_{1:t}) \hat{p}(\boldsymbol{\theta} | \mathbf{y}_{1:t-1}))$ de modo a obter,

$$\mathbf{m}_t = \arg \max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}) \quad \text{e} \quad \mathbf{C}_t = - \left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right]_{\boldsymbol{\theta}=\mathbf{m}_t}^{-1}.$$

O ótimo global só é garantido quando $\ell(\boldsymbol{\theta})$ é côncava e o custo computacional tem um crescimento cúbico com o aumento do número de parâmetros a serem estimados.

- ii) **Aproximação via amostragem por importância:** Nesta abordagem a distribuição dos parâmetros é especificada por um conjunto de pontos $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_D\}$ que estão associados a um conjunto de pesos $\{\pi_{t,1}, \pi_{t,2}, \dots, \pi_{t,D}\}$, cujo a soma $\sum_{d=1}^D \pi_{t,d} = 1$. A distribuição discreta é definida por, $\hat{p}(\boldsymbol{\theta} | \mathbf{y}_{1:t}) = \sum_{d=1}^D \delta_{\boldsymbol{\theta}_d}(\boldsymbol{\theta}) \pi_{t,d}$ em que $\delta(\cdot)$ é a função delta de Kronecker. A distribuição a *posteriori* é obtida com a atualização dos pesos da seguinte forma,

$$\pi_{t,d} \propto \hat{p}(\mathbf{y}_t | \boldsymbol{\theta}_d, \mathbf{y}_{1:t-1}) \pi_{t-1,d}$$

para $d = 1, \dots, D$ e os novos pesos $\pi_{t,d}$ são normalizados de forma a somar 1. Esta forma de aproximação da distribuição a *posteriori* possui um custo computacional

que cresce exponencialmente à medida que o número de parâmetros aumenta. Uma outra limitação é devido ao conjunto de valores que é fixado *a priori* e eventualmente a distribuição *a posteriori* pode se concentrar em um ponto desse conjunto de valores.

Após a atualização dos parâmetros, é então simulado o conjunto predito $\tilde{\mathbf{x}}_t^{(i)} = \tilde{\mathbf{x}}_t^{(pi)} + \mathbf{w}_t^{(i)}$ em que $\mathbf{w}_t^{(i)} \sim \mathcal{N}_m(\mathbf{0}, \mathbf{Q}_t(\boldsymbol{\theta}^{(i)}))$ e as observações $\tilde{\mathbf{y}}_t^{(i)} = \mathbf{H}_t(\boldsymbol{\theta}^{(i)})\tilde{\mathbf{x}}_t^{(i)} + \mathbf{v}_t^{(i)}$, com $\mathbf{v}_t^{(i)} \sim \mathcal{N}_r(\mathbf{0}, \mathbf{R}_t(\boldsymbol{\theta}^{(i)}))$, $i = 1, 2, \dots, L$. O conjunto *a posteriori* dos estados no tempo t é obtido baseado nos valores da amostra *a posteriori* de $\boldsymbol{\theta}$. Assim, temos que,

$$\hat{\mathbf{x}}_t^{(i)} = \tilde{\mathbf{x}}_t^{(i)} + \hat{\mathbf{K}}_t(\boldsymbol{\theta}^{(i)})(\mathbf{y}_t - \tilde{\mathbf{y}}_t^{(i)})$$

em que a estimativa do ganho de Kalman é dada por $\hat{\mathbf{K}}_t(\boldsymbol{\theta}^{(i)}) = \mathbf{H}_t(\boldsymbol{\theta}^{(i)})\hat{\mathbf{P}}_t^f(\boldsymbol{\theta}^{(i)})\hat{\mathbf{B}}_t(\boldsymbol{\theta}^{(i)})^{-1}$.

No *Algoritmo 4* apresentamos os passos para a estimação.

Algoritmo 4: ESTIMANDO OS ESTADOS E O CONJUNTO DE PARÂMETROS $\boldsymbol{\theta}$.

- 1 O algoritmo é inicializado com a amostra *a priori* $\boldsymbol{\theta}^{(i)} \sim p(\boldsymbol{\theta})$ e $\hat{\mathbf{x}}_0^{(i)} \sim \mathcal{N}_m(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0)$ para $i = 1, \dots, L$.
Em cada ciclo $t = 1, 2, \dots$ proceder da seguinte forma:
- 2 Calcule o valor predito: $\tilde{\mathbf{x}}_t^{(pi)} = \mathcal{M}(\hat{\mathbf{x}}_{t-1}^{(i)}) \quad i = 1, \dots, L$
- 3 Aproxime a função de verossimilhança usando:

$$\hat{p}(\mathbf{y}_t | \boldsymbol{\theta}, \mathbf{y}_{1:t-1}) \propto |\hat{\mathbf{B}}_t(\boldsymbol{\theta})|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} \hat{\mathbf{e}}_t(\boldsymbol{\theta})' \hat{\mathbf{B}}_t(\boldsymbol{\theta})^{-1} \hat{\mathbf{e}}_t(\boldsymbol{\theta}) \right\}$$

- 4 Atualize a distribuição dos parâmetros usando amostragem por importância ou aproximação normal.

$$\hat{p}(\boldsymbol{\theta} | \mathbf{y}_{1:t}) \propto \hat{p}(\mathbf{y}_t | \boldsymbol{\theta}, \mathbf{y}_{1:t-1}) \hat{p}(\boldsymbol{\theta} | \mathbf{y}_{1:t-1})$$

- 5 Gere uma amostra da *posteriori* atualizada: $\boldsymbol{\theta}^{(i)} \sim \hat{p}(\boldsymbol{\theta} | \mathbf{y}_{1:t})$, $i = 1, \dots, L$.
- 6 Gere a amostra predita ensemble:

$$\tilde{\mathbf{x}}_t^{(i)} = \tilde{\mathbf{x}}_t^{(pi)} + \mathbf{w}_t^{(i)}, \quad \text{em que } \mathbf{w}_t^{(i)} \sim \mathcal{N}_m(\mathbf{0}, \mathbf{Q}_t(\boldsymbol{\theta}^{(i)})) \quad i = 1, \dots, L.$$

- 7 Gere a distribuição de *filtragem* ensemble usando: $\hat{\mathbf{x}}_t^{(i)} = \tilde{\mathbf{x}}_t^{(i)} + \hat{\mathbf{K}}_t(\boldsymbol{\theta}^{(i)})(\mathbf{y}_t - \tilde{\mathbf{y}}_t^{(i)})$
-

O *Algoritmo 4* tem seu melhor desempenho quando o comprimento do vetor de parâmetros $\boldsymbol{\theta}$ é pequeno. Assim, essa condição limita o uso desse algoritmo a modelos de baixa dimensão com estados e parâmetros desconhecidos. Já o EnKF Aumentado, como apresentado por [Stroud e Bengtsson \(2007\)](#), também possui uma limitação, pois ele não estima bem parâmetros que apresentam baixa correlação linear com o vetor de estados.

Nos próximos capítulos verificaremos por meio de estudos de simulação a eficiência desses métodos no ajuste de modelos de espaço de estados univariados lineares e não lineares.

Capítulo 5

Simulação I - Modelo de Espaço de Estados Linear Normal

Neste capítulo realizamos um estudo de simulação para avaliar a eficiência do filtro de Kalman Ensemble (EnKF) com relação a acurácia das estimativas dos estados e dos parâmetros em modelos univariados lineares. Nas simulações iremos verificar o efeito do tamanho da amostra ensemble nas estimativas dos estados, o efeito do comprimento da série de observações e a influência da informação *a priori* na distribuição *a posteriori* dos parâmetros. Os algoritmos desenvolvidos pelo autor foram implementados no software R (R Core Team, 2019) e os códigos são apresentados no Apêndice D.

5.1 Estimação dos parâmetros σ^2 e τ^2

Para analisar a estimação de σ^2 e τ^2 consideramos o modelo de nível local,

$$y_t = x_t + v_t \quad (5.1)$$

$$x_t = x_{t-1} + w_t \quad (5.2)$$

em que $v_t \sim N(0, \sigma^2)$ e $w_t \sim N(0, \tau^2)$ para $t = 1, 2, \dots, N$.

A estimação dos parâmetros σ^2 e τ^2 é feita usando o *Algoritmo 4*, que utiliza a seguinte função de verossimilhança predita um passo à frente para obter as estimativas *a posteriori*,

$$\hat{p}(y_t | \mathbf{y}_{1:t-1}, \sigma^2, \tau^2) \propto \exp \left\{ -\frac{1}{2} \left(\frac{(y_t - \tilde{a}_t)^2}{\tilde{p}_t^2 + \tau^2 + \sigma^2} \right) \right\} \quad (5.3)$$

em que $\tilde{a}_t$ e $\tilde{p}_t^2$ representam a versão univariada de $\tilde{\mathbf{a}}_t$ e $\tilde{\mathbf{P}}_t$ em (4.14). Note que $\tilde{p}_t^2$ depende implicitamente de τ^2 e σ^2 .

Nas Tabelas 5.1 e 5.2 apresentamos uma análise de sensibilidade para τ^2 , verificando o efeito das distribuições *a priori* em suas estimativas. Na simulação executamos 1000 réplicas e assumimos que $x_0 = 20$, $\sigma^2 = 9$ e $\tau^2 = 4$. Já no processo de inferência do modelo consideramos o tamanho da amostra ensemble $L = 200$, o comprimento da série de observações $N = 100$, uma amostra inicial de $\hat{x}_0^{(i)} | \tau^2 \sim N(20, \tau^2)$ e o método de amostragem por importância para estimar σ^2 e τ^2 com a grade de valores pertencendo ao intervalo $(0, 20]$. Nesta análise de sensibilidade fixamos uma distribuição *a priori* Gama

Inversa (GI) para σ^2 dada por $\sigma^2 \sim GI(12.125, 100.125)$ ($E(\sigma^2) = 9$ e $var(\sigma^2) = 8$) e atribuímos algumas distribuições *a priori* bem informativas Gama Inversa para τ^2 , em todas a média é igual a 4 e as variâncias são iguais a 2, 1, 0.5 e 0.2, respectivamente.

Tabela 5.1: *Análise de sensibilidade de τ^2 .*

$\tau^2 \sim$	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
$GI(10, 36)$	$\sigma^2 = 9$	7.662	1.221	7.542	5.660	10.320
	$\tau^2 = 4$	1.975	0.314	1.936	1.483	2.682
$GI(18, 68)$	$\sigma^2 = 9$	7.647	1.213	7.531	5.643	10.278
	$\tau^2 = 4$	1.985	0.248	1.960	1.578	2.526
$GI(34, 132)$	$\sigma^2 = 9$	7.735	1.241	7.614	5.701	10.436
	$\tau^2 = 4$	1.988	0.222	1.968	1.618	2.466
$GI(82, 324)$	$\sigma^2 = 9$	7.623	1.209	7.510	5.630	10.261
	$\tau^2 = 4$	1.992	0.182	1.979	1.681	2.376

Tabela 5.2: *Medidas de acurácia na análise de sensibilidade de τ^2 .*

$\tau^2 \sim$	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$GI(10, 36)$	$\sigma^2 = 9$	-1.338	1.338	4.660	100%
	$\tau^2 = 4$	-2.025	2.025	1.199	0%
$GI(18, 68)$	$\sigma^2 = 9$	-1.353	1.353	4.634	100%
	$\tau^2 = 4$	-2.015	2.015	0.947	0%
$GI(34, 132)$	$\sigma^2 = 9$	-1.265	1.265	4.734	100%
	$\tau^2 = 4$	-2.012	2.012	0.848	0%
$GI(82, 324)$	$\sigma^2 = 9$	-1.377	1.377	4.631	100%
	$\tau^2 = 4$	-2.008	2.008	0.695	0%

Nas Tabelas 5.3 e 5.4 faremos a análise de sensibilidade para σ^2 , neste segundo caso fixamos *a priori* $\tau^2 \sim GI(10, 36)$ e atribuímos distribuições Gama Inversa para σ^2 , em toda a média é igual a 9 e as variâncias iguais a 8, 4, 2 e 1, respectivamente.

Tabela 5.3: *Análise de sensibilidade de σ^2 .*

$\sigma^2 \sim$	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
$GI(12.125, 100.125)$	$\sigma^2 = 9$	7.662	1.221	7.542	5.660	10.320
	$\tau^2 = 4$	1.975	0.314	1.936	1.483	2.682
$GI(22.25, 191.25)$	$\sigma^2 = 9$	7.889	1.112	7.789	6.033	10.276
	$\tau^2 = 4$	1.992	0.182	1.978	1.680	2.375
$GI(42.5, 373.5)$	$\sigma^2 = 9$	8.214	0.979	8.143	6.537	10.292
	$\tau^2 = 4$	1.992	0.182	1.979	1.680	2.376
$GI(83, 738)$	$\sigma^2 = 9$	8.397	0.883	8.334	6.872	10.252
	$\tau^2 = 4$	1.992	0.182	1.979	1.680	2.376

Tabela 5.4: Medidas de acurácia na análise de sensibilidade de σ^2 .

$\sigma^2 \sim$	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$GI(12.125, 100.125)$	$\sigma^2 = 9$	-1.338	1.338	4.660	100%
	$\tau^2 = 4$	-2.025	2.025	1.199	0%
$GI(22.25, 191.25)$	$\sigma^2 = 9$	-1.111	1.111	4.242	100%
	$\tau^2 = 4$	-2.008	2.008	0.695	0%
$GI(42.5, 373.5)$	$\sigma^2 = 9$	-0.786	0.786	3.754	100%
	$\tau^2 = 4$	-2.008	2.008	0.696	0%
$GI(83, 738)$	$\sigma^2 = 9$	-0.603	0.603	3.381	100%
	$\tau^2 = 4$	-2.008	2.008	0.696	0%

Nas análises de sensibilidade apresentadas nas Tabelas 5.1, 5.2, 5.3 e 5.4, notamos que o *Algoritmo 4* possui problemas na estimação do parâmetro τ^2 . Mesmo assumindo distribuições *a priori* informativas obtivemos um viés alto, estimativas distantes dos verdadeiros valores dos parâmetros, intervalos de credibilidade não contendo o verdadeiro valor de τ^2 e por consequência obtivemos uma taxa de cobertura nula. As estimativas de σ^2 e τ^2 mostram-se robustas à escolha da distribuição *a priori* de τ^2 e apenas as estimativas do parâmetro σ^2 mostram-se sensíveis à escolha da distribuição *a priori* de σ^2 .

Para verificar o efeito da estimação dos parâmetros σ^2 e τ^2 na estimativa dos estados, calculamos a taxa de cobertura para o intervalo interquartil da distribuição de *filtragem* com os seguintes comprimentos da série de observações $N = 30, 50, 100, 250$ e 500 . As taxas estão apresentadas na Tabela 5.5. No processo de inferência assumimos distribuições *a priori* bem informativas Gama Inversa para $\sigma^2 \sim GI(12.125, 100.125)$, $\tau^2 \sim GI(10, 36)$ e $x_0^{(i)} | \tau^2 \sim N(20, \tau^2)$.

Tabela 5.5: Taxa de cobertura do intervalo interquartil da distribuição de *filtragem*.

N	30	50	100	250	500
porcentagem	47.6%	48.8%	48%	47.4%	45%

A taxa de cobertura do intervalo interquartil apresentou-se um pouco abaixo do valor nominal de 50%. Esta baixa porcentagem das estimativas dos estados pertencendo ao intervalo interquartil pode estar associado ao fato dos parâmetros σ^2 e τ^2 estarem sendo mal estimados pelo *Algoritmo 4* e isso interfere na estimativa dos estados.

Nas Figuras 5.1 e 5.2 apresentamos a estimativa da distribuição preditiva e de *filtragem* de cada estado para o comprimento da série de observações igual a $N = 50$ e 100 . Os gráficos de $N = 30, 250$ e 500 são apresentados no Apêndice C. Cada box-plot representa a distribuição média de cada estado a respeito das 1000 réplicas.

Figura 5.1: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 50 - (σ^2 e τ^2).

Figura 5.2: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 100 - (σ^2 e τ^2).

Nas Figuras 5.1 e 5.2 notamos que as distribuições preditivas e de filtragem dos estados foram bem estimadas para o comprimento da série de observações igual a 50 e 100. Portanto, não aparentam estar sendo muito afetadas pela má estimação dos parâmetros de escala.

Nas Tabelas 5.6 e 5.7 verificaremos o comportamento das estimativas dos parâmetros σ^2 e τ^2 com relação ao comprimento da série de observações.

Tabela 5.6: Estimativas dos parâmetros σ^2 e τ^2 para os diferentes comprimentos da série de observações.

N	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
30	$\sigma^2 = 9$	8.504	1.490	8.344	6.109	11.782
	$\tau^2 = 4$	1.975	0.313	1.937	1.484	2.677
50	$\sigma^2 = 9$	8.218	1.394	8.071	5.955	11.273
	$\tau^2 = 4$	1.975	0.313	1.938	1.485	2.677
100	$\sigma^2 = 9$	7.662	1.221	7.542	5.660	10.320
	$\tau^2 = 4$	1.975	0.314	1.936	1.483	2.682
250	$\sigma^2 = 9$	6.498	0.903	6.423	4.975	8.420
	$\tau^2 = 4$	1.975	0.313	1.936	1.485	2.678
500	$\sigma^2 = 9$	5.475	0.667	5.428	4.328	6.873
	$\tau^2 = 4$	1.976	0.313	1.937	1.483	2.679

Tabela 5.7: Medidas de acurácia das estimativas dos parâmetros σ^2 e τ^2 para os diferentes comprimentos da série de observações.

	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
30	$\sigma^2 = 9$	-0.496	0.496	5.673	100%
	$\tau^2 = 4$	-2.025	2.025	1.193	0%
50	$\sigma^2 = 9$	-0.782	0.782	5.318	100%
	$\tau^2 = 4$	-2.025	2.025	1.193	0%
100	$\sigma^2 = 9$	-1.338	1.338	4.660	100%
	$\tau^2 = 4$	-2.025	2.025	1.199	0%
250	$\sigma^2 = 9$	-2.502	2.502	3.445	1.7%
	$\tau^2 = 4$	-2.025	2.025	1.193	0 %
500	$\sigma^2 = 9$	-3.525	3.525	2.545	0 %
	$\tau^2 = 4$	-2.024	2.024	1.195	0 %

Como vimos nas Tabelas 5.1, 5.2, 5.3, 5.4, 5.6 e 5.7, o Algoritmo 4 não consegue estimar bem os parâmetros σ^2 e τ^2 mesmo aumentando o tamanho do comprimento da série de observações e assumindo distribuições *a priori* muito informativas para ambos parâmetros. Também observamos que em todas as situações simuladas o viés nas estimativas são altos e a taxa de cobertura do parâmetro τ^2 é nula. Além disso, as estimativas de σ^2 e τ^2 foram distantes dos verdadeiros valores dos parâmetros e robustas à escolha da distribuição *a priori*. Dessa forma, o Algoritmo 4 apresenta problemas para estimar o modelo de espaço de estados linear normal quando as duas variâncias são desconhecidas e isso necessita de uma investigação futura. Assim, na próxima seção faremos a suposição de proporcionalidade das variâncias $\frac{\tau^2}{\sigma^2} = \kappa$ com o valor de κ predefinido e consideraremos somente o parâmetro σ^2 desconhecido. Esta suposição pode ser vista em West e Harrison (2006) e Campagnoli *et al.* (2009).

5.2 Estimação do parâmetro σ^2 com $\tau^2 = \kappa\sigma^2$

Considerando que o modelo de nível local em (5.1) e (5.2) possui somente a variância das observações (σ^2) desconhecida e a proporcionalidade das variâncias $\frac{\tau^2}{\sigma^2} = \kappa$ com o valor de κ predefinido, é possível obter analiticamente a distribuição *a posteriori* dos estados e do parâmetro de escala usando a conjugação ao atribuir a distribuição Gama Inversa para σ^2 , conforme abordado em West e Harrison (2006) e Campagnoli *et al.* (2009). Segundo West e Harrison (2006), a variância das observações σ^2 é muitas vezes desconhecida e grande em relação a variância dos estados τ^2 e segundo Campagnoli *et al.* (2009) a magnitude relativa(κ) de τ^2 e σ^2 é um fator que afeta o ganho de Kalman, que por sua vez, determina como as distribuições de predição e de filtragem serão sensíveis as informações das observações.

Agora verificaremos se o *Algoritmo 4* consegue estimar os estados e o parâmetro σ^2 do modelo. Nas simulações executamos 1000 réplicas e assumimos que $x_0 = 20$, $\sigma^2 = 9$ e $\kappa = \frac{4}{9}$ para assim termos um modelo similar ao modelo anterior. No processo de inferência do modelo consideramos o tamanho da amostra ensemble $L = 200$, o comprimento da série de observações $N = 100$, uma amostra inicial de $\hat{x}_0^{(i)} | \sigma^2 \sim N(20, \tau^2 = \kappa\sigma^2)$ e o método de amostragem por importância para estimar σ^2 com a grade de valores pertencendo ao intervalo (0,20]. Faremos a análise de sensibilidade para σ^2 e atribuímos as mesmas distribuições Gama Inversa para da situação anterior, em que a média é igual a 9 e as variâncias iguais a 8, 4, 2 e 1, respectivamente. Também atribuímos *a priori* não informativa $\sigma^2 \sim GI(0.1, 0.1)$.

Os resultados das simulações estão apresentados nas Tabelas 5.8 e 5.9.

Tabela 5.8: Análise de sensibilidade de σ^2 com $\sigma^2 = 9$.

$\sigma^2 \sim$	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
$GI(0.1, 0.1)$	$\sigma^2 = 9$	8.133	2.647	7.909	3.670	13.804
$GI(12.125, 100.125)$	$\sigma^2 = 9$	9.965	2.183	9.745	6.386	14.732
$GI(22.25, 191.25)$	$\sigma^2 = 9$	9.594	1.751	9.424	6.718	13.401
$GI(42.5, 375.5)$	$\sigma^2 = 9$	9.347	1.323	9.236	7.117	12.160
$GI(83, 738)$	$\sigma^2 = 9$	9.243	0.981	9.181	7.544	11.297

Tabela 5.9: Medidas de acurácia na análise de sensibilidade de σ^2 com $\sigma^2 = 9$.

$\sigma^2 \sim$	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$GI(0.1, 0.1)$	$\sigma^2 = 9$	-0.867	1.576	10.134	97.9%
$GI(12.125, 100.125)$	$\sigma^2 = 9$	0.965	1.162	8.346	99.8%
$GI(22.25, 191.25)$	$\sigma^2 = 9$	0.594	0.763	6.683	100%
$GI(42.5, 375.5)$	$\sigma^2 = 9$	0.347	0.457	5.044	100%
$GI(83, 738)$	$\sigma^2 = 9$	0.243	0.296	3.753	100%

Nas Tabelas 5.8 e 5.9 observando os intervalos de credibilidade notamos que o *Algoritmo 4* conseguiu estimar razoavelmente bem o parâmetro σ^2 mesmo assumindo uma distribuição a *priori* não informativa $\sigma^2 \sim GI(0.1, 0.1)$. Quanto a sensibilidade, o parâmetro σ^2 mostrou-se sensível à escolha da distribuição a *priori*, visto que a amplitude, o desvio-padrão, o desvio absoluto e o viés diminuíram com a atribuição de *prioris* mais informativas. A taxa de cobertura para o parâmetro σ^2 é alta em todas as situações simuladas.

Na próxima seção também faremos a suposição de proporcionalidade das variâncias $\frac{\tau^2}{\sigma^2} = \kappa$ com o valor de κ predefinido e consideraremos um modelo um pouco mais geral, com os parâmetros α e σ^2 desconhecidos.

5.3 Estimação dos parâmetros α e σ^2

Consideramos inicialmente o modelo linear gaussiano, para o qual o filtro de Kalman foi inicialmente proposto, isto é,

$$y_t = x_t + v_t, \quad (5.4)$$

$$x_t = \alpha x_{t-1} + w_t \quad \alpha \in \mathbb{R} \quad (5.5)$$

em que $v_t \sim N(0, \sigma^2)$ e $w_t \sim N(0, \tau^2)$ para $t = 1, 2, \dots, N$. A partir deste modelo analisamos as estimativas a *posteriori* dos estados e dos parâmetros.

Assumindo que os parâmetros α e σ^2 do modelo apresentado em (5.4) e (5.5) são desconhecidos e as variâncias dos estados e das observações são proporcionais $\frac{\tau^2}{\sigma^2} = \kappa$, faremos agora uma análise de sensibilidade para verificar a influência das distribuições a *priori* nos valores estimados.

No processo de inferência foram utilizados os *Algoritmos 3* e *4* com o tamanho da amostra ensemble $L = 200$, uma série temporal de comprimento $N = 100$ e mil réplicas. No *Algoritmo 4* usamos a amostragem por importância para estimar σ^2 com intervalo de variação $(0, 20]$ e em todos os casos usamos uma amostra ensemble inicial não informativa $\hat{x}_0^{(i)} \sim N(0, \tau^2 = \kappa\sigma^2)$, em que assumiremos $\kappa = 0.25$, ou seja, a variância dos estados representa um quarto da variância das observações.

Análise de sensibilidade - σ^2

Na análise de sensibilidade do parâmetro σ^2 fixamos a *priori* vaga $\alpha \sim N(0, 10^2)$ e assumimos dois tipos de distribuições a *priori* para σ^2 . Primeiro, consideramos a classe conjugada Gama Inversa(GI). Como alternativa, foi considerada uma *priori* Half-Cauchy(HC) para σ . Esta distribuição de cauda pesada tem um parâmetro de escala que, à medida que aumenta torna a distribuição menos informativa, podendo assim ser uma alternativa para a *priori* Gama Inversa (Gelman *et al.*, 2006). Em Hall (2008) é recomendado que esse parâmetro de escala seja igual a 25.

Para verificar a influência de cada distribuição a *priori* em σ^2 primeiro assumimos que $x_0 = 20$, $\alpha = 1$ e $\sigma^2 = 4$ e atribuímos algumas distribuições a *priori* conjugadas Gama Inversa(GI) para σ^2 com médias igual a 4 e variâncias 8 e 2 (com exceção da *priori* vaga $GI(0.1, 0.1)$). Também atribuímos a *priori* HC(25) para σ . Em cada uma

das situações calculamos as médias das estimativas dos parâmetros *a posteriori* e mais algumas outras quantidades. Os resultados são apresentados nas Tabelas 5.10 e 5.11.

Tabela 5.10: *Análise de sensibilidade de σ^2 com $\alpha = 1$ e $\sigma^2 = 4$.*

$\sigma^2 \sim$	Parâmetros	Média	Desvio padrão	Mediana	Q2.5%	Q97.5%
$GI(0.1, 0.1)$	$\alpha = 1$	1.001	0.005	1.001	0.991	1.010
	$\sigma^2 = 4$	4.959	1.089	4.853	3.167	7.328
$GI(4, 12)$	$\alpha = 1$	0.993	0.006	0.993	0.982	1.005
	$\sigma^2 = 4$	3.623	0.944	3.523	2.090	5.690
$GI(10, 36)$	$\alpha = 1$	0.994	0.005	0.994	0.984	1.005
	$\sigma^2 = 4$	3.325	0.701	3.252	2.187	4.865
$\sigma \sim$						
$HC(25)$	$\alpha = 1$	0.996	0.005	0.996	0.987	1.006
	$\sigma^2 = 4$	4.478	0.938	4.388	2.945	6.527

Tabela 5.11: *Medidas de acurácia na análise de sensibilidade de σ^2 com $\alpha = 1$ e $\sigma^2 = 4$.*

$\sigma^2 \sim$	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$GI(0.1, 0.1)$	$\alpha = 1$	0.001	0.002	0.018	98.0%
	$\sigma^2 = 4$	0.959	1.013	4.161	93.3%
$GI(4, 12)$	$\alpha = 1$	-0.007	0.007	0.023	94.4%
	$\sigma^2 = 4$	-0.377	0.591	3.600	98.0%
$GI(10, 36)$	$\alpha = 1$	-0.006	0.006	0.021	98.2%
	$\sigma^2 = 4$	-0.675	0.699	2.677	94.0%
$\sigma \sim$					
$HC(25)$	$\alpha = 1$	-0.004	0.004	0.019	99.2%
	$\sigma^2 = 4$	0.478	0.632	3.581	97.4%

Nesta segunda situação faremos $x_0 = 20$, $\alpha = 1$ e $\sigma^2 = 1$ e atribuímos algumas distribuições *a priori* conjugadas Gama Inversa(GI) para σ^2 com médias igual a 1 e variâncias 8 e 2 (com exceção da *priori* vaga $GI(0.1, 0.1)$). Neste caso, também atribuímos *a priori* $HC(25)$ para σ . Os resultados são apresentados nas Tabelas 5.12 e 5.13.

Tabela 5.12: *Análise de sensibilidade de σ^2 com $\alpha = 1$ e $\sigma^2 = 1$.*

σ^2	Parâmetros	Média	Desvio padrão	Mediana	Q2.5%	Q97.5%
$GI(0.1, 0.1)$	$\alpha = 1$	1.003	0.002	1.003	0.998	1.007
	$\sigma^2 = 1$	1.218	0.267	1.194	0.780	1.796
$GI(2.125, 1.125)$	$\alpha = 1$	1.003	0.002	1.003	0.998	1.007
	$\sigma^2 = 1$	1.230	0.280	1.203	0.770	1.839
$GI(2.5, 1.5)$	$\alpha = 1$	1.002	0.003	1.002	0.997	1.007
	$\sigma^2 = 1$	1.093	0.249	1.067	0.688	1.639
$\sigma \sim$						
$HC(25)$	$\alpha = 1$	0.997	0.003	0.997	0.992	1.003
	$\sigma^2 = 1$	0.948	0.208	0.928	0.613	1.398

Tabela 5.13: Medidas de acurácia na análise de sensibilidade de σ^2 com $\alpha = 1$ e $\sigma^2 = 1$.

	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$\sigma^2 \sim GI(0.1, 0.1)$	$\alpha = 1$	0.003	0.003	0.008	93.0%
	$\sigma^2 = 1$	0.218	0.234	1.016	94.4%
$\sigma^2 \sim GI(2.125, 1.125)$	$\alpha = 1$	0.003	0.003	0.009	93.4%
	$\sigma^2 = 1$	0.230	0.246	1.069	94.0%
$\sigma^2 \sim GI(2.5, 1.5)$	$\alpha = 1$	0.002	0.002	0.010	100.0%
	$\sigma^2 = 1$	0.093	0.161	0.951	97.8%
$\sigma \sim$					
$HC(25)$	$\alpha = 1$	-0.003	0.003	0.010	96.5%
	$\sigma^2 = 1$	-0.052	0.128	0.785	98.1%

Nas Tabelas 5.10, 5.11, 5.12 e 5.13 podemos notar que, observando a taxa de cobertura, o método conseguiu recuperar os verdadeiros valores dos parâmetros, no entanto, as estimativas de σ^2 apresentaram-se viesadas. Esse viés pode ser reduzido com o uso de algumas prioris, tal como observado na priori $\sigma \sim HC(25)$. Quanto a sensibilidade, σ^2 mostrou-se sensível à escolha das *prioris*. Já as estimativas do parâmetro α mostraram-se robustas em relação à escolha das *prioris* de σ^2 .

Análise de sensibilidade α

Na análise de influência da distribuição a *priori* de α fixamos $\sigma^2 \sim GI(0.1, 0.1)$ e atribuímos distribuições normais para α , algumas centramos no zero e alteramos a variância e em outras fixamos a variância e alteramos a média. Neste primeiro caso, assumimos que $x_0 = 20$, $\alpha = 1$ e $\sigma^2 = 4$. Os resultados obtidos são apresentados nas Tabelas 5.14 e 5.15.

Tabela 5.14: Análise de sensibilidade de α com $\alpha = 1$ e $\sigma^2 = 4$.

$\alpha \sim$	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
$N(0, 10^2)$	$\alpha = 1$	1.001	0.005	1.001	0.991	1.010
	$\sigma^2 = 4$	4.959	1.089	4.853	3.167	7.328
$N(0, 2^2)$	$\alpha = 1$	0.997	0.007	0.997	0.984	1.010
	$\sigma^2 = 4$	4.811	1.049	4.710	3.080	7.087
$N(0, 0.5^2)$	$\alpha = 1$	1.004	0.003	1.004	0.998	1.010
	$\sigma^2 = 4$	4.961	1.098	4.858	3.146	7.339
$N(1, 2^2)$	$\alpha = 1$	0.993	0.006	0.993	0.982	1.004
	$\sigma^2 = 4$	3.644	1.088	3.540	1.842	5.996
$N(2, 2^2)$	$\alpha = 1$	0.996	0.005	0.996	0.987	1.005
	$\sigma^2 = 4$	3.149	1.000	3.052	1.492	5.316
$N(3, 2^2)$	$\alpha = 1$	0.999	0.004	0.999	0.991	1.006
	$\sigma^2 = 4$	3.304	1.035	3.207	1.592	5.542

Tabela 5.15: Medidas de acurácia na análise de sensibilidade de α com $\alpha = 1$ e $\sigma^2 = 4$.

$\alpha \sim$	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$N(0, 10^2)$	$\alpha = 1$	0.001	0.002	0.018	98.0%
	$\sigma^2 = 4$	0.959	1.013	4.161	93.3%
$N(0, 2^2)$	$\alpha = 1$	-0.003	0.004	0.027	99.4%
	$\sigma^2 = 4$	0.811	0.896	4.007	95.4%
$N(0, 0.5^2)$	$\alpha = 1$	0.004	0.004	0.012	85.1%
	$\sigma^2 = 4$	0.961	1.013	4.193	93.5%
$N(1, 2^2)$	$\alpha = 1$	-0.007	0.007	0.022	86.8%
	$\sigma^2 = 4$	-0.356	0.658	4.154	97.1%
$N(2, 2^2)$	$\alpha = 1$	-0.004	0.004	0.017	95.2%
	$\sigma^2 = 4$	-0.851	0.948	3.824	90.2%
$N(3, 2^2)$	$\alpha = 1$	-0.001	0.002	0.015	99.2%
	$\sigma^2 = 4$	-0.696	0.823	3.950	93.7%

Agora assumimos que $x_0 = 20$, $\alpha = 1$ e $\sigma^2 = 1$ e atribuímos a mesmas distribuições a priori para α apresentadas no caso anterior. Também assumimos que $\sigma^2 \sim GI(0.1, 0.1)$. Os resultados obtidos são apresentados nas Tabelas 5.16 e 5.17.

Tabela 5.16: Análise de sensibilidade de α com $\alpha = 1$ e $\sigma^2 = 1$.

$\alpha \sim$	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
$N(0, 10^2)$	$\alpha = 1$	1.003	0.002	1.003	0.998	1.007
	$\sigma^2 = 1$	1.218	0.267	1.194	0.780	1.796
$N(0, 2^2)$	$\alpha = 1$	0.998	0.003	0.998	0.992	1.003
	$\sigma^2 = 1$	1.168	0.253	1.144	0.755	1.715
$N(0, 0.5^2)$	$\alpha = 1$	0.998	0.003	0.998	0.993	1.003
	$\sigma^2 = 1$	1.228	0.261	1.203	0.800	1.793
$N(1, 2^2)$	$\alpha = 1$	1.003	0.002	1.003	0.998	1.007
	$\sigma^2 = 1$	1.111	0.241	1.088	0.717	1.636
$N(2, 2^2)$	$\alpha = 1$	0.996	0.003	0.996	0.989	1.002
	$\sigma^2 = 1$	1.214	0.263	1.188	0.783	1.783
$N(3, 2^2)$	$\alpha = 1$	1.000	0.002	1.000	0.996	1.005
	$\sigma^2 = 1$	1.183	0.251	1.159	0.771	1.727

Nas Tabelas 5.14, 5.15, 5.16 e 5.17 notamos que o método conseguiu estimar razoavelmente bem os parâmetros em todas as escolhas de *prioris* atribuídas a α , no entanto, o parâmetro σ^2 apresentou estimativas viesadas e um pouco sensíveis as *prioris* atribuídas a α . O menor viés de σ^2 foi obtido com $N(1, 2^2)$. Já o parâmetro α possui boas estimativas em todas as situações apresentadas e não é sensível à escolha das distribuições a priori, visto que, o conhecimento a priori não influenciou muito as estimativas a posteriori.

Tabela 5.17: Medidas de acurácia na análise de sensibilidade de α com $\alpha = 1$ e $\sigma^2 = 1$.

$\alpha \sim$	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$N(0, 10^2)$	$\alpha = 1$	0.003	0.003	0.008	93.0%
	$\sigma^2 = 1$	0.218	0.234	1.016	94.4%
$N(0, 2^2)$	$\alpha = 1$	-0.002	0.002	0.011	99.6%
	$\sigma^2 = 1$	0.168	0.200	0.960	94.8%
$N(0, 0.5^2)$	$\alpha = 1$	-0.002	0.002	0.010	99.3%
	$\sigma^2 = 1$	0.228	0.241	0.993	91.3%
$N(1, 2^2)$	$\alpha = 1$	0.003	0.003	0.009	97.0%
	$\sigma^2 = 1$	0.111	0.164	0.919	97.5%
$N(2, 2^2)$	$\alpha = 1$	-0.004	0.004	0.013	92.8%
	$\sigma^2 = 1$	0.214	0.231	1.001	92.9%
$N(3, 2^2)$	$\alpha = 1$	0.000	0.001	0.009	100%
	$\sigma^2 = 1$	0.183	0.206	0.956	93.8%

Estimação dos estados

A análise das estimativas dos estados e dos parâmetros com relação ao comprimento da série temporal das observações será feita com os seguintes valores: $x_0 = 20$, $\alpha = 1$, $\sigma^2 = 4$, $\kappa = 0.25$. Nas simulações executamos 1000 réplicas para cada um dos seguintes comprimentos da série de observações $N = 30, 50, 100, 250$ e 500 . Em cada um desses comprimentos consideramos que $\hat{x}_0^{(i)} | \sigma^2 \sim N(0, \tau^2 = \kappa \sigma^2)$.

Na estimativa da distribuição de *filtragem* calculamos a taxa de cobertura do intervalo de credibilidade interquartil para cada comprimento da série de observações considerando dois casos para distribuições *a priori*. No Caso I assumimos que $\alpha \sim N(0, 10^2)$ e $\sigma^2 \sim GI(0.1, 0.1)$ e no Caso II assumimos que $\alpha \sim N(1, 2^2)$ e $\sigma \sim HC(25)$. Os resultados são apresentados na Tabela 5.18.

Tabela 5.18: Taxa de cobertura do intervalo interquartil da distribuição de filtragem.

Casos	N = 30	N = 50	N = 100	N = 250	N = 500
Caso I	47%	49.9%	51.8%	51.7%	52.9%
Caso II	60.3%	55.4%	51.9%	51.3%	50.1%

A taxa de cobertura tende a ficar próxima de 50% para os comprimentos da série temporal das observações maiores que 30 e como podemos perceber, as taxas foram geralmente superiores à 50% para ambos os casos.

Nas Figuras 5.3 e 5.4 apresentamos a estimativa da distribuição preditiva e de *filtragem* de cada estado para o comprimento da série de observações igual a $N = 100$ considerando os casos I e II, os demais gráficos são apresentados no Apêndice C. Cada box-plot representa a distribuição média de cada estado a respeito das 1000 réplicas.

Figura 5.3: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 100 - Caso I.

Figura 5.4: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 100 - Caso II.

Nas Figuras 5.3 e 5.4 podemos observar que as estimativas das distribuições preditivas e a *posteriori* dos estados iniciais obtidas no Caso II possuem uma dispersão maior do que as estimativas obtidas por Caso I.

Na Tabela 5.19 apresentamos as estimativas dos parâmetros obtidas com diferentes comprimentos da série de observações para os casos I e II:

Tabela 5.19: Estimativas dos parâmetros α e σ^2 para diferentes comprimentos da série de observações.

	N	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
Caso I	30	$\alpha = 1$	0.988	0.011	0.988	0.966	1.009
		$\sigma^2 = 4$	5.439	2.394	5.037	2.013	11.131
	50	$\alpha = 1$	0.992	0.009	0.992	0.975	1.008
		$\sigma^2 = 4$	4.873	1.561	4.663	2.494	8.430
	100	$\alpha = 1$	1.001	0.005	1.001	0.991	1.010
		$\sigma^2 = 4$	4.959	1.089	4.853	3.167	7.328
	250	$\alpha = 1$	0.999	0.003	0.999	0.994	1.004
		$\sigma^2 = 4$	4.581	0.601	4.545	3.530	5.827
	500	$\alpha = 1$	1.001	0.002	1.001	0.998	1.005
		$\sigma^2 = 4$	4.686	0.419	4.668	3.930	5.536
Caso II	30	$\alpha = 1$	0.991	0.018	0.991	0.956	1.026
		$\sigma^2 = 4$	4.777	2.472	4.299	1.478	10.718
	50	$\alpha = 1$	0.994	0.010	0.994	0.976	1.013
		$\sigma^2 = 4$	4.403	1.473	4.193	2.188	7.758
	100	$\alpha = 1$	1.000	0.004	1.000	0.992	1.009
		$\sigma^2 = 4$	4.181	0.869	4.099	2.759	6.073
	250	$\alpha = 1$	1.002	0.002	1.002	0.998	1.007
		$\sigma^2 = 4$	3.850	0.475	3.822	3.027	4.828
	500	$\alpha = 1$	0.996	0.003	0.996	0.991	1.001
		$\sigma^2 = 4$	4.100	0.346	4.089	3.486	4.792

Na Tabela 5.19 notamos que as estimativas tendem a ficar mais acuradas com o crescimento do comprimento da série de observações, uma vez que o desvio padrão, o desvio absoluto e a amplitude média do intervalo de credibilidade empírico dos parâmetros, diminuem. Notamos também que geralmente as estimativas de σ^2 obtidas no Caso II foram um pouco melhores que as do Caso I. Já as estimativas de α foram similares na maioria dos casos.

Nas Figuras 5.5 e 5.6 apresentamos as estimativa do viés e do desvio absoluto dos parâmetros para diferentes comprimentos da série de observações considerando os casos I e II.

Figura 5.5: Estimativa do viés e do desvio absoluto de α para diferentes comprimentos da série de observações considerando os casos I e II.

Figura 5.6: Estimativa do viés e do desvio absoluto de σ^2 para diferentes comprimentos da série de observações considerando os casos I e II.

Nas Figuras 5.5 e 5.6 observamos que o viés e o desvio absoluto α são similares. Por outro lado, as estimativas do viés e do desvio absoluto de σ^2 são menores no Caso II, indicando a preferência pelo uso dessas distribuições *a priori*.

5.4 Estimação dos parâmetros q , α e σ^2

Consideramos agora um modelo linear gaussiano um pouco mais geral,

$$y_t = qx_t + v_t, \quad q \in \mathbb{R} \quad (5.6)$$

$$x_t = \alpha x_{t-1} + w_t \quad \alpha \in \mathbb{R} \quad (5.7)$$

em que $v_t \sim N(0, \sigma^2)$ e $w_t \sim N(0, \tau^2)$ para $t = 1, 2, \dots, N$. A distribuição de *filtragem* desse modelo é $x_t|y_{1:t} \sim N(\hat{\mu}_t, \hat{\sigma}_t^2)$ com

$$\begin{aligned} \hat{\mu}_t &= \tilde{\mu}_t + K_t(y_t - q\tilde{\mu}_t) \\ \hat{\sigma}_t^2 &= (1 - qK_t)\tilde{\sigma}_t^2 \end{aligned}$$

em que $\tilde{\mu}_t = \alpha\hat{\mu}_{t-1}$ e $\tilde{\sigma}_t^2 = \alpha^2\hat{\sigma}_{t-1}^2 + \tau^2$ representam, respectivamente, a média e a variância da distribuição predita dos estados um passo a frente e $K_t = q\tilde{\sigma}_t^2(q^2\tilde{\sigma}_t^2 + \sigma^2)^{-1}$ é o ganho de Kalman. Assim, o parâmetro q influencia tanto a média quando a variância da distribuição de *filtragem*.

Para ilustrar essa influência considere que $\alpha = 1$, $\tau^2 = 1$, $\sigma^2 = 4$ e $x_0 = 20$, são conhecidos. Alterando o valor de q verificamos o seguinte efeito na dispersão e na média da distribuição de *filtragem* exata:

Figura 5.7: Distribuição de filtragem do estado X_{100} para o parâmetro q assumindo os valores 0.5, 1, 5 e 9.

Note que, à medida em que aumentamos os valores de q a distribuição de *filtragem* tende a ficar mais concentrada e a média tende a se deslocar.

Agora faremos uma análise de sensibilidade do parâmetro q e para isso consideramos os casos I e II para as distribuições a priori de α e σ^2 apresentadas na seção anterior. No processo de inferência foram utilizados os *Algoritmos 3* e *4* com o tamanho da amostra ensemble $L = 200$, uma série temporal de comprimento $N = 100$ e mil réplicas. No *Algoritmo 4* usamos a amostragem por importância para estimar q e σ^2 com $q \in (-5, 5)$ e $\sigma^2 \in (0, 20]$, em todos os casos uma amostra ensemble inicial não informativa $\hat{x}_0^{(i)} \sim N(0, \tau^2 = \kappa\sigma^2)$, em que $\kappa = 0.25$. Os resultados da análise de sensibilidade são apresentados nas Tabelas 5.20 e 5.21.

Tabela 5.20: *Análise de sensibilidade de q .*

Caso I	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
$q \sim N(1, 10^2)$	$q = 1$	0.000	0.003	0.000	-0.003	0.003
	$\alpha = 1$	0.814	0.582	0.841	-0.437	1.944
	$\sigma^2 = 4$	5.840	3.136	4.973	2.221	13.818
$q \sim N(1, 5^2)$	$q = 1$	0.002	0.005	0.003	-0.003	0.007
	$\alpha = 1$	0.824	0.564	0.855	-0.409	1.931
	$\sigma^2 = 4$	5.915	3.231	5.020	2.133	13.983
$q \sim N(1, 1^2)$	$q = 1$	0.004	0.003	0.004	0.001	0.007
	$\alpha = 1$	0.824	0.569	0.854	-0.399	1.929
	$\sigma^2 = 4$	5.924	3.213	5.022	2.194	14.025
Caso II						
$q \sim N(1, 10^2)$	$q = 1$	0.000	0.000	0.000	0.000	0.000
	$\alpha = 1$	0.993	0.002	0.994	0.990	0.997
	$\sigma^2 = 4$	13.890	2.990	13.553	9.121	20.514
$q \sim N(1, 5^2)$	$q = 1$	0.000	0.000	0.000	0.000	0.000
	$\alpha = 1$	1.007	0.001	1.007	1.005	1.009
	$\sigma^2 = 4$	16.985	3.617	16.580	11.212	24.998
$q \sim N(1, 1^2)$	$q = 1$	0.008	0.000	0.008	0.008	0.008
	$\alpha = 1$	1.003	0.001	1.003	1.001	1.004
	$\sigma^2 = 4$	10.297	2.277	10.046	6.683	15.316

Tabela 5.21: *Medidas de acurácia na análise de sensibilidade de q .*

Caso I	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$q \sim N(1, 10^2)$	$q = 1$	-1.000	1.000	0.006	0.0%
	$\alpha = 1$	-0.186	0.197	2.381	69.3%
	$\sigma^2 = 4$	1.840	2.498	11.597	80.7%
$q \sim N(1, 5^2)$	$q = 1$	-0.998	0.998	0.010	0.0%
	$\alpha = 1$	-0.176	0.189	2.340	59.4%
	$\sigma^2 = 4$	1.915	2.422	11.850	84.0%
$q \sim N(1, 1^2)$	$q = 1$	-0.996	0.996	0.006	0.0
	$\alpha = 1$	-0.176	0.189	2.328	58.5%
	$\sigma^2 = 4$	1.924	2.460	11.832	81.7%
Caso II					
$q \sim N(1, 10^2)$	$q = 1$	-1.000	1.000	0.000	0.0%
	$\alpha = 1$	-0.007	0.007	0.007	3.3%
	$\sigma^2 = 4$	9.890	9.890	11.393	0.03%
$q \sim N(1, 5^2)$	$q = 1$	-1.000	1.000	0.000	0.0%
	$\alpha = 1$	0.007	0.007	0.004	11.2%
	$\sigma^2 = 4$	12.985	12.985	13.786	0.3%
$q \sim N(1, 1^2)$	$q = 1$	-0.992	0.992	0.000	0.0%
	$\alpha = 1$	0.003	0.003	0.003	9.8%
	$\sigma^2 = 4$	6.297	6.297	8.632	9.3%

Podemos notar nas Tabelas 5.20 e 5.21 que, quando o parâmetro q é desconhecido, as estimativas dos outros parâmetros são afetadas, como pode ser observado principalmente nas taxas de cobertura. Mesmo assumindo distribuições *a priori* bem informativas para q , as estimativas tenderam a se concentrar em torno do zero. Esse problema pode estar associado a falta de identificabilidade do modelo, pelo fato da quantidade qX_t envolver o produto de duas quantidades desconhecidas. Este tópico sugere investigação futura que está fora do propósito dessa dissertação.

Na análise das estimativas dos estados e dos parâmetros com relação ao comprimento da série temporal das observações consideramos $N = 30, 50, 100, 250$ e 500 , e em cada um desses comprimentos assumimos que $\hat{x}_0^{(i)}|\sigma^2 \sim N(0, \tau^2 = 0.25\sigma^2)$. Também consideramos dois casos para as distribuições *a priori*, que são: Caso I assumimos que $q \sim N(1, 5^2)$, $\alpha \sim N(0, 10^2)$ e $\sigma^2 \sim GI(0.1, 0.1)$ e no Caso II assumimos que $q \sim N(1, 5^2)$, $\alpha \sim N(1, 2^2)$ e $\sigma \sim HC(25)$. Os resultados são apresentados nas Tabelas 5.22 e 5.23.

Tabela 5.22: Estimativas dos parâmetros q, α e σ^2 para diferentes comprimentos da série de observações.

Caso I	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
$N = 30$	$q = 1$	-0.000	0.025	-0.000	-0.025	0.025
	$\alpha = 1$	0.020	7.469	0.046	-15.330	15.146
	$\sigma^2 = 4$	3.563	4.680	1.375	0.110	16.818
$N = 50$	$q = 1$	0.000	0.025	0.001	-0.025	0.025
	$\alpha = 1$	0.045	5.514	0.089	-11.513	11.464
	$\sigma^2 = 4$	3.506	4.648	1.343	0.110	16.765
$N = 100$	$q = 1$	0.002	0.005	0.003	-0.003	0.007
	$\alpha = 1$	0.824	0.564	0.855	-0.409	1.931
	$\sigma^2 = 4$	5.915	3.231	5.020	2.133	13.983
$N = 250$	$q = 1$	0.002	0.000	0.002	0.002	0.002
	$\alpha = 1$	1.002	0.000	1.002	1.001	1.003
	$\sigma^2 = 4$	14.922	1.510	14.883	12.167	17.871
$N = 500$	$q = 1$	0.002	0.000	0.002	0.002	0.002
	$\alpha = 1$	1.001	0.000	1.001	1.000	1.001
	$\sigma^2 = 4$	19.975	0.055	20.000	19.826	20.000
Caso II						
$N = 30$	$q = 1$	0.000	0.000	0.000	0.000	0.001
	$\alpha = 1$	0.975	0.113	0.982	0.764	1.143
	$\sigma^2 = 4$	80.616	38.616	72.934	29.127	173.369
$N = 50$	$q = 1$	0.001	0.000	0.001	0.001	0.001
	$\alpha = 1$	0.977	0.028	0.979	0.932	1.015
	$\sigma^2 = 4$	31.930	11.962	29.874	15.072	59.893
$N = 100$	$q = 1$	0.000	0.000	0.000	0.000	0.000
	$\alpha = 1$	1.007	0.001	1.007	1.005	1.009
	$\sigma^2 = 4$	16.985	3.617	16.580	11.212	24.998
$N = 250$	$q = 1$	0.001	0.000	0.001	0.001	0.001
	$\alpha = 1$	1.004	0.001	1.004	1.003	1.005
	$\sigma^2 = 4$	27.901	2.760	27.719	23.126	33.645
$N = 500$	$q = 1$	0.000	0.000	0.000	0.000	0.000
	$\alpha = 1$	1.005	0.000	1.005	1.005	1.005
	$\sigma^2 = 4$	87.518	5.780	87.269	77.103	99.205

Tabela 5.23: Medidas de acurácia das estimativas dos parâmetros q , α e σ^2 para diferentes comprimentos da série de observações.

Caso I	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
N = 30	$q = 1$	-1.000	1.000	0.050	0.0 %
	$\alpha = 1$	-0.980	1.005	30.476	100%
	$\sigma^2 = 4$	-0.437	0.468	16.707	100%
N = 50	$q = 1$	-1.000	1.000	0.050	0.0%
	$\alpha = 1$	-0.955	0.962	22.977	99.9%
	$\sigma^2 = 4$	-0.494	0.515	16.655	100%
N = 100	$q = 1$	-0.998	0.998	0.010	0.0%
	$\alpha = 1$	-0.176	0.189	2.340	59.4%
	$\sigma^2 = 4$	1.915	2.422	11.850	84.0%
N = 250	$q = 1$	-0.998	0.998	0.000	0.0%
	$\alpha = 1$	0.002	0.002	0.001	5.2%
	$\sigma^2 = 4$	10.922	10.922	5.704	0.0%
N = 500	$q = 1$	-0.998	0.998	0.000	0.0%
	$\alpha = 1$	0.001	0.001	0.000	2.6%
	$\sigma^2 = 4$	15.975	15.975	1.740	0.0%
Caso II					
N = 30	$q = 1$	-1.000	1.000	0.000	0.0%
	$\alpha = 1$	-0.025	0.041	0.378	85.5%
	$\sigma^2 = 4$	76.616	76.731	144.242	39.5%
N = 50	$q = 1$	-0.999	0.999	0.000	0.0%
	$\alpha = 1$	-0.023	0.023	0.083	56.7%
	$\sigma^2 = 4$	27.930	27.942	44.821	25.2%
N = 100	$q = 1$	-1.000	1.000	0.000	0.0%
	$\alpha = 1$	0.007	0.007	0.004	11.2%
	$\sigma^2 = 4$	12.985	12.985	13.786	0.3%
N = 250	$q = 1$	-0.999	0.999	0.000	0.0%
	$\alpha = 1$	0.004	0.004	0.002	13.6%
	$\sigma^2 = 4$	23.901	23.901	10.519	0.0%
N = 500	$q = 1$	-1.000	1.000	0.000	0.0%
	$\alpha = 1$	0.005	0.005	0.000	0.0%
	$\sigma^2 = 4$	83.518	83.518	22.102	0.0%

Uma observação interessante nas Tabelas 5.22 e 5.23 é que, à medida em que aumentamos o comprimento da série temporal das observações, as estimativas dos parâmetros q e α tendem a convergir para valores constantes que são diferentes dos verdadeiros valores dos parâmetros. Observando as quantidades obtidas, percebemos que quando o parâmetro q é desconhecido o método não consegue recuperar bem os verdadeiros valores dos parâmetros.

As estimativas dos estados são também extremamente afetadas e isso pode ser observado na taxa de cobertura média dos intervalos interquartil da distribuição de *filtragem*. As taxas estão apresentadas na Tabela 5.24.

Tabela 5.24: Taxa de cobertura média dos intervalos interquartil da distribuição de *filtragem*.

	N = 30	N = 50	N = 100	N = 250	N = 500
Caso I	100 %	99.8%	77.9%	32.4%	15.8%
Caso II	38.6%	22.7%	11.7%	3.8 %	3.0%

As taxas de cobertura do intervalo interquartil da distribuição de *filtragem* apresentadas na Tabela 5.24 são um indicativo de que o parâmetro q está sendo mal estimado pelo *Algoritmo 4* e isso interfere na estimativa dos estados e dos demais parâmetros. É importante mencionar que o problema da falta de identificabilidade pode ser a causa desse resultado, porque diferentes valores de q e X_t são capazes de reproduzir o mesmo modelo de espaço de estados, um vez que o *Algoritmo 4* está estimando o produto de duas quantidades desconhecidas, qX_t .

Nas Figuras (5.8) e (5.9) apresentamos a estimativa da distribuição preditiva e de *filtragem* de cada estado para o comprimento da série de observações igual a $N = 100$ considerando os casos I e II.

Figura 5.8: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 100 - Caso I.

Figura 5.9: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 100 - Caso II.

Nas Figuras 5.8 e 5.9 notamos que os estados não estão sendo bem estimados em ambos os casos. As estimativas das distribuições apresentaram grande dispersão, o que tornou irrelevante as informações das observações no passo de atualização. Dessa forma, o algoritmo obtém estimativas das distribuições de predição e de filtragem com muita incerteza sobre os verdadeiros valores dos estados.

Das simulações apresentadas neste capítulo notamos que para fazer a estimação dos parâmetros do modelo de espaço de estados linear normal devemos considerar o parâmetro q conhecido e as variâncias proporcionais. Com essas condições verificamos que o método consegue estimar razoavelmente bem os parâmetros α e σ^2 . No entanto, o viés de σ^2 não tende a zero, mesmo para *prioris* bem informativas ou um comprimento da série temporal consideravelmente grande. Esse viés pode ser reduzido com o uso de algumas *prioris*, tal como observado na *priori* $\sigma \sim HC(25)$. Já o parâmetro α é bem estimado, visto que suas estimativas apresentaram um pequeno viés e uma alta taxa de cobertura em todos os casos analisados.

No Apêndice B são apresentados simulações para verificar o efeito do tamanho da amostra *ensemble* L nas estimativas e também uma comparação com as estimativas obtidas pelo pacote DLM do software R. No próximo capítulo verificaremos a eficiência desses métodos no ajuste do modelo não linear denominado modelo de produção logístico.

Capítulo 6

Simulação II - Modelo de Produção Excedente

Neste capítulo realizamos um estudo de simulação para avaliar a eficiência do filtro de Kalman Ensemble (EnKF) com relação a acurácia das estimativas dos estados e dos parâmetros em um modelo não linear gaussiano representado pelo modelo de produção excedente logístico. Ao analisar esse modelo também estamos verificando se o EnKF pode ser uma possível alternativa aos algoritmos MCMC no ajuste desse modelo. Nas simulações iremos verificar o efeito do tamanho da amostra ensemble nas estimativas dos estados, o efeito do comprimento da série de observações e a influência da informação *a priori* na distribuição *a posteriori* dos parâmetros.

Os modelos de produção excedente são representados por um sistema que descreve, de forma integrada, a evolução temporal do estoque de pesca e a sua relação com os dados de captura por unidade de esforço (CPUE), como segue:

$$\begin{aligned}y_t &= qx_t + v_t \\x_t &= x_{t-1} + rx_{t-1}\left(1 - \frac{x_{t-1}}{K}\right) + w_t, \quad r > 0 \text{ e } K > 0,\end{aligned}$$

em que $v_t \sim N(0, \sigma^2)$ e $w_t \sim N(0, \tau^2 = \kappa\sigma^2)$.

No processo de inferência utilizamos o filtro de Kalman Ensemble (EnKF) em conjunto com os métodos apresentados na seção 4.2.2 para fazer a estimação da biomassa $x_{t=0:N}$, da taxa intrínseca de crescimento r , da capacidade de carga que o ambiente pode suportar K e de σ^2 . Em contrapartida, como vimos na seção 5.4, devemos considerar o parâmetro q fixo e as variâncias proporcionais.

Faremos agora uma análise de sensibilidade para verificar a influência das distribuições *a priori* nos valores estimados *a posteriori*. Assim, escolhemos algumas distribuições *a priori* para r , K e σ^2 e em todos os casos usamos uma amostra ensemble inicial $x_0^{(i)} \sim N(0, \tau^2)$, em que $\tau^2 = \kappa\sigma^2$. Na estimação, usamos conjuntamente os Algoritmos 3 e 4 com o tamanho da amostra ensemble $L = 200$, uma série temporal de comprimento $N = 100$ e 1000 réplicas. No Algoritmo 4 usamos a amostragem por importância para estimar σ com intervalo de variação $(0, 100]$. Os algoritmos foram implementados no software R e os códigos são apresentados no Apêndice D.

Os dados foram gerados assumindo que: $x_0 = 10$, $r = 0.42$, $K = 46$, $\sigma = 37$, $q = 9$ e $\tau^2 = 0.005\sigma^2$. Estes valores foram escolhidos com base nas estimativas obtidas por Montenegro e Branco (2016) em uma análise de dados reais. Neste mesmo estudo foram especificadas distribuições *a priori* informativas para os parâmetros r e K baseadas em Montenegro *et al.* (2009), em que r possui Log-normal com média 0.33 e variância 0.007 e K também possui Log-normal com média 33 e variância 68.06. Estas distribuições serão utilizadas em nosso estudo de sensibilidade dos parâmetros.

Na literatura de modelos de produção excedente é comum atribuir a distribuição Log-normal (ln) para os parâmetros r e K . Deste modo, na análise de sensibilidade também estaremos testando a robustez do Algoritmo 3 quanto a não normalidade das *prioris* de tais parâmetros, uma vez que este algoritmo é construído com base na suposição de normalidade das distribuições *a priori* dos parâmetros contidos na função de evolução. Já o parâmetro σ^2 geralmente é assumido uma *priori* não informativa Gama Inversa, mas aqui também analisaremos a *priori* Half-Cauchy como uma alternativa.

6.1 Análise de Sensibilidade

Na análise de sensibilidade do parâmetro r consideramos dois cenários, no primeiro fixamos a média e alteramos a variância e no segundo fixamos a variância e alteramos a média. Em todos os cenários fixamos as distribuições *a priori* não informativas para K e σ^2 , em que K segue Log-normal com média 33 e variância 1000 e $\sigma^2 \sim GI(0.1, 0.1)$. Nas Tabelas 6.1 e 6.2 são apresentadas as seguintes quantidades *a posteriori* dos parâmetros: média, mediana, desvio padrão, viés, amplitude do intervalo empírico com 95% de credibilidade, desvio absoluto e a taxa de cobertura (proporção de vezes que o verdadeiro valor do parâmetro está contido nos intervalos de credibilidade).

Tabela 6.1: Análise de sensibilidade do parâmetro r .

$r \sim ln$	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
$E(r) = 0.33$	$r = 0.42$	0.552	0.169	0.534	0.191	0.837
$Var(r) = 0.5$	$K = 46$	46.906	1.965	46.701	43.038	49.853
	$\sigma = 37$	35.961	3.281	35.785	30.166	42.724
$E(r) = 0.33$	$r = 0.42$	0.316	0.141	0.309	0.202	0.742
$Var(r) = 0.01$	$K = 46$	47.099	2.060	46.929	43.910	51.298
	$\sigma = 37$	37.826	3.397	37.139	31.297	44.340
$E(r) = 0.33$	$r = 0.42$	0.314	0.063	0.310	0.206	0.448
$Var(r) = 0.007$	$K = 46$	46.705	1.166	46.668	44.578	49.045
	$\sigma = 37$	35.919	3.222	35.757	30.188	42.533
$E(r) = 0.42$	$r = 0.42$	0.489	0.156	0.474	0.241	0.827
$Var(r) = 0.5$	$K = 46$	45.331	1.178	45.321	43.112	47.622
	$\sigma = 37$	37.912	3.273	37.443	31.797	44.333
$E(r) = 1$	$r = 0.42$	0.656	0.197	0.640	0.324	1.087
$Var(r) = 0.5$	$K = 46$	45.903	1.516	45.887	43.053	48.845
	$\sigma = 37$	40.838	3.672	40.664	34.300	48.343

Tabela 6.2: Medidas de acurácia na análise de sensibilidade do parâmetro r .

$r \sim \ln$	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$E(r) = 0.33$	$r = 0.42$	0.132	0.166	0.646	85.3%
$Var(r) = 0.5$	$K = 46$	0.906	0.868	6.815	94.1%
	$\sigma = 37$	-1.039	2.660	12.558	93.2%
$E(r) = 0.33$	$r = 0.42$	-0.104	0.122	0.539	89.9%
$Var(r) = 0.1$	$K = 46$	1.099	1.218	7.387	97.0%
	$\sigma = 37$	0.826	2.167	12.043	97.3%
$E(r) = 0.33$	$r = 0.42$	-0.106	0.106	0.242	74.9%
$Var(r) = 0.007$	$K = 46$	0.705	0.813	4.467	97.3%
	$\sigma = 37$	-1.081	2.372	12.345	96.1%
$E(r) = 0.42$	$r = 0.42$	0.069	0.141	0.586	86.6%
$Var(r) = 0.5$	$K = 46$	-0.669	0.819	4.511	90.1%
	$\sigma = 37$	0.912	2.418	12.536	94.5%
$E(r) = 1$	$r = 0.42$	0.236	0.265	0.763	75.6%
$Var(r) = 0.5$	$K = 46$	-0.097	1.533	5.792	89.9%
	$\sigma = 37$	3.838	4.425	14.043	83.0%

Nas Tabelas 6.1 e 6.2 podemos notar que, de modo geral, o método conseguiu recuperar os verdadeiros valores dos parâmetros em todas as distribuições atribuídas a r . No entanto, suas estimativas apresentaram-se enviesadas. Quanto a sensibilidade, r , K e σ^2 mostraram-se sensíveis à escolha da distribuição *a priori*, visto que houve uma redução do desvio padrão, do desvio absoluto e da amplitude do intervalo de credibilidade com a suposição de *prioris* mais informativas.

Os verdadeiros valores dos parâmetros pertenceram ao intervalo empírico com 95% de credibilidade, mas suas taxas de cobertura mantiveram-se abaixo de 95% na maioria das situações simuladas. O melhor ajuste das situações analisadas ocorreu quando foi atribuída a r a distribuição Log-normal com média 0.33 e variância 0.007 e o pior ajuste ocorreu com r assumindo Log-normal com média 1 e variância 0.5.

Na análise de sensibilidade do parâmetro K fixamos as *prioris* para r e σ^2 , em que r segue Log-normal com média 1 e variância 0.5 e $\sigma^2 \sim GI(0.1, 0.1)$. Nas Tabelas 6.3 e 6.4 são apresentadas as estimativas para cada distribuição *a priori* de K obtidas considerando dois cenários, no primeiro fixamos a média e alteramos a variância e no segundo fixamos a variância e alteramos a média, como segue:

Tabela 6.3: *Análise de sensibilidade do parâmetro K .*

$K \sim \ln$	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
$E(K) = 33$	$r = 0.42$	0.656	0.197	0.640	0.324	1.087
$Var(K) = 1000$	$K = 46$	45.903	1.516	45.887	43.053	48.845
	$\sigma = 37$	40.838	3.672	40.664	34.300	48.343
$E(K) = 33$	$r = 0.42$	0.329	0.136	0.315	0.121	0.626
$Var(K) = 100$	$K = 46$	45.357	1.501	45.341	42.552	48.276
	$\sigma = 37$	39.413	3.578	39.217	33.050	46.793
$E(K) = 33$	$r = 0.42$	0.523	0.164	0.509	0.257	0.883
$Var(K) = 68.06$	$K = 46$	44.899	0.993	44.902	42.994	46.777
	$\sigma = 37$	38.622	3.290	38.453	32.788	45.373
$E(K) = 46$	$r = 0.42$	0.476	0.118	0.466	0.279	0.731
$Var(K) = 100$	$K = 46$	45.877	0.892	45.841	44.313	47.612
	$\sigma = 37$	37.087	3.106	36.920	31.618	43.531
$E(K) = 60$	$r = 0.42$	0.500	0.081	0.495	0.361	0.670
$Var(K) = 100$	$K = 46$	46.594	0.894	46.561	44.997	48.360
	$\sigma = 37$	39.021	3.297	38.838	33.202	45.839

Tabela 6.4: *Medidas de acurácia na análise de sensibilidade do parâmetro K .*

$K \sim \ln$	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$E(K) = 33$	$r = 0.42$	0.236	0.265	0.763	75.6%
$Var(K) = 1000$	$K = 46$	-0.097	1.533	5.792	89.9%
	$\sigma = 37$	3.838	4.425	14.043	83.0%
$E(K) = 33$	$r = 0.42$	-0.091	0.137	0.505	84.9%
$Var(K) = 100$	$K = 46$	-0.643	1.451	5.723	90.9%
	$\sigma = 37$	2.413	3.711	13.743	89.6%
$E(K) = 33$	$r = 0.42$	0.103	0.158	0.626	86.2%
$Var(K) = 68.06$	$K = 46$	-1.101	1.157	3.784	80.1%
	$\sigma = 37$	1.622	3.050	12.585	90.4%
$E(K) = 46$	$r = 0.42$	0.056	0.108	0.452	87.4%
$Var(K) = 100$	$K = 46$	-0.123	0.604	3.299	93.0%
	$\sigma = 37$	0.087	2.770	11.913	92.5%
$E(K) = 60$	$r = 0.42$	0.080	0.111	0.310	72.6%
$Var(K) = 100$	$K = 46$	0.594	0.693	3.363	94.7%
	$\sigma = 37$	2.021	2.715	12.637	93.8%

Nas Tabelas 6.3 e 6.4 são apresentadas as estimativas *a posteriori* obtidas para cada uma das distribuições *a priori* atribuídas ao parâmetro K . Apesar das estimativas serem enviesadas os verdadeiros valores dos parâmetros pertenceram ao intervalo empírico com 95% de credibilidade, mas a taxa de cobertura manteve-se abaixo de 95% em todas as situações simuladas. Quanto a sensibilidade, os parâmetros mostraram-se sensíveis à escolha das *prioris* de K , visto que, na maioria dos casos, houve uma redução do desvio absoluto, do desvio padrão e da amplitude do intervalo de credibilidade com a suposição de *prioris* mais informativas. Já o viés das estimativas, em grande parte, não variou conforme as diferentes propostas de distribuições *a priori*.

Na análise de sensibilidade do parâmetro σ^2 fixamos as *prioris* para r e K , em que r segue Log-normal com média 1 e variância 0.5 e K segue Log-normal com média 33 e variância 1000. Nas Tabelas 6.5 e 6.6 são apresentadas estimativas de σ^2 para algumas distribuições *a priori* conjugadas Gama Inversa(GI) com médias igual a 36 e variâncias 100 e 10 (com exceção da *priori* vaga $GI(0.1, 0.1)$), respectivamente. Neste caso também atribuímos a *priori* $HC(25)$ para σ .

Tabela 6.5: Análise de sensibilidade σ^2 .

$\sigma^2 \sim$	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
$GI(0.1, 0.1)$	$r = 0.42$	0.656	0.197	0.640	0.324	1.087
	$K = 46$	45.903	1.516	45.887	43.053	48.845
	$\sigma = 37$	40.838	3.672	40.664	34.300	48.343
$E(\sigma^2) = 37^2$ $Var(\sigma^2) = 100$	$r = 0.42$	0.536	0.190	0.520	0.218	0.946
	$K = 46$	45.819	0.819	45.815	44.267	47.404
	$\sigma = 37$	37.002	0.135	37.000	36.746	37.265
$E(\sigma^2) = 37^2$ $Var(\sigma^2) = 10$	$r = 0.42$	0.498	0.208	0.480	0.154	0.961
	$K = 46$	45.819	0.924	45.781	44.179	47.662
	$\sigma = 37$	37.000	0.043	36.999	36.917	37.082
$HC(25)$	$r = 0.42$	0.459	0.175	0.443	0.176	0.844
	$K = 46$	45.573	0.946	45.576	43.775	47.347
	$\sigma = 37$	39.298	3.429	39.105	33.247	46.417

Nas Tabelas 6.5 e 6.6, podemos notar que todas as distribuições atribuídas ao parâmetro de escala conseguiram recuperar razoavelmente bem os verdadeiros valores dos parâmetros, e à medida que usamos distribuições *a priori* mais informativas, o viés, o desvio padrão, o desvio absoluto e a amplitude do intervalo de credibilidade empírico de σ tenderam a diminuir consideravelmente. Também notamos que as estimativas de σ obtidas pela *priori* não informativa $\sigma \sim HC(25)$ foram melhores que as estimativas obtidas com $\sigma^2 \sim GI(0.1, 0.1)$. Quanto a sensibilidade, os parâmetros mostraram-se sensíveis à escolha das distribuições *a priori*.

Tabela 6.6: Medidas de acurácia na análise de sensibilidade do parâmetro σ^2 .

$\sigma^2 \sim$	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$GI(0.1, 0.1)$	$r = 0.42$	0.236	0.265	0.763	75.6%
	$K = 46$	-0.097	1.533	5.792	89.9%
	$\sigma = 37$	3.838	4.425	14.043	83.0%
$E(\sigma^2) = 37^2$ $Var(\sigma^2) = 100$	$r = 0.42$	0.116	0.171	0.729	89.0%
	$K = 46$	-0.181	0.821	3.137	88.0%
	$\sigma = 37$	0.002	0.009	0.519	100.0%
$E(\sigma^2) = 37^2$ $Var(\sigma^2) = 10$	$r = 0.42$	0.078	0.161	0.807	95.0%
	$K = 46$	-0.181	0.532	3.484	99.0%
	$\sigma = 37$	-0.000	0.002	0.165	100.0%
$\sigma \sim$					
$HC(25)$	$r = 0.42$	0.039	0.162	0.668	87.3%
	$K = 46$	-0.427	1.075	3.572	80.2%
	$\sigma = 37$	2.298	3.207	13.171	88.9%

6.2 Estimação dos Estados

A análise das estimativas dos estados e dos parâmetros com relação ao comprimento da série temporal das observações será feita considerando os seguintes comprimentos $N = 30, 50, 100, 250$ e 500 , em cada um desses comprimentos atribuímos as *prioris* informativas usadas por Montenegro e Branco (2016) para os parâmetros r e K , e verificamos o desempenho de $\sigma \sim HC(25)$ como uma alternativa a $\sigma^2 \sim GI(0.1, 0.1)$.

Na estimativa da distribuição de *filtragem* calculamos a taxa de cobertura do intervalo de credibilidade interquartil para cada comprimento da série de observações considerando as *prioris* $\sigma^2 \sim GI(0.1, 0.1)$ e $\sigma \sim HC(25)$. Os resultados são apresentados na Tabela 6.7.

Tabela 6.7: Taxa de cobertura média dos intervalos interquartil das distribuições de filtragem dos estados.

	N = 30	N = 50	N = 100	N = 250	N = 500
$\sigma^2 \sim GI(0.1, 0.1)$	45.2%	49.1%	49.9%	51.4 %	50.1%
$\sigma \sim HC(25)$	50.2%	49.5%	51.0%	49.7%	49.9%

A taxa de cobertura do intervalo interquartil da distribuição de *filtragem* começou baixa para o comprimento da série temporal das observações igual a 30 e tendeu a se aproximar de 50% à medida em que comprimento da série temporal aumentou quando utilizamos a *priori* $\sigma^2 \sim GI(0.1, 0.1)$ e quando utilizamos a *priori* $\sigma \sim HC(25)$ em todos os comprimentos obtivemos taxas próximas a 50%.

A não linearidade do modelo afeta de certa forma o EnKF na estimativa dos estados e isso pode ser percebido nas Figuras 6.1 e 6.2, em que apresentamos a estimativa da distribuição preditiva e de *filtragem* de cada estado para o comprimento da série de observações igual $N = 100$. Cada box-plot representa a distribuição média de cada estado a respeito das 1000 réplicas. Os gráficos de $N = 30, 250$ e 500 são apresentados no Apêndice C.

Figura 6.1: Estimativas das distribuições preditivas e a posteriori de cada estado para um modelo não linear e uma série de observações de comprimento igual a 100 com $\sigma^2 \sim GI(0.1, 0.1)$.

Figura 6.2: Estimativas das distribuições preditivas e a posteriori de cada estado para um modelo não linear e uma série de observações de comprimento igual a 100 com $\sigma \sim HC(25)$.

Nas Figuras 6.1 e 6.2 observamos que as estimativas obtidas pelo EnKF tendem a não estimar bem os estados que se distanciam de forma abrupta do estado do tempo anterior.

Nas Tabelas 6.8 e 6.9 apresentamos as estimativas *a posteriori* dos parâmetros e algumas outras medidas obtidas com diferentes comprimentos da série de observações para $\sigma^2 \sim GI(0.1, 0.1)$ e $\sigma \sim HC(25)$:

Tabela 6.8: Estimativas dos parâmetros para os diferentes comprimentos da série de observações.

	N	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
$\sigma^2 \sim GI(0.1, 0.1)$	30	$r = 0.42$	0.334	0.072	0.327	0.214	0.488
		$K = 46$	44.731	2.103	44.664	40.878	49.013
		$\sigma = 37$	43.531	7.386	42.816	31.329	59.684
	50	$r = 0.42$	0.347	0.069	0.341	0.230	0.494
		$K = 46$	44.496	1.548	44.441	41.658	47.636
		$\sigma = 37$	39.006	5.079	38.646	30.286	49.706
	100	$r = 0.42$	0.324	0.062	0.320	0.216	0.453
		$K = 46$	45.399	1.032	45.381	43.460	47.439
		$\sigma = 37$	38.012	3.269	37.844	32.217	44.728
	250	$r = 0.42$	0.387	0.052	0.385	0.294	0.494
		$K = 46$	46.149	0.504	46.143	45.192	47.131
		$\sigma = 37$	36.095	1.867	36.036	32.689	39.890
	500	$r = 0.42$	0.377	0.038	0.377	0.307	0.455
		$K = 46$	46.257	0.292	46.257	45.702	46.815
		$\sigma = 37$	36.792	1.304	36.758	34.403	39.394
$\sigma \sim HC(25)$	30	$r = 0.42$	0.360	0.074	0.354	0.237	0.521
		$K = 46$	45.720	1.994	45.655	42.066	49.782
		$\sigma = 37$	39.219	6.935	38.610	27.629	54.162
	50	$r = 0.42$	0.351	0.070	0.346	0.232	0.502
		$K = 46$	45.019	1.432	44.969	42.401	47.922
		$\sigma = 37$	37.533	4.741	37.207	29.380	47.580
	100	$r = 0.42$	0.372	0.066	0.368	0.258	0.513
		$K = 46$	44.800	0.861	44.795	43.167	46.470
		$\sigma = 37$	36.002	3.060	35.841	30.568	42.329
	250	$r = 0.42$	0.387	0.052	0.385	0.295	0.493
		$K = 46$	45.627	0.482	45.626	44.705	46.558
		$\sigma = 37$	36.455	1.858	36.397	33.075	40.192
	500	$r = 0.42$	0.384	0.038	0.383	0.314	0.460
		$K = 46$	45.827	0.296	45.827	45.260	46.397
		$\sigma = 37$	36.746	1.299	36.715	34.348	39.323

Nas Tabelas 6.8 e 6.9 são apresentadas as estimativas *a posteriori* obtidas para cada comprimento da série de observações (N). Nas Tabelas, notamos que as estimativas tendem a ficar mais acuradas quando aumentamos o valor de N , já que o desvio padrão, o viés, o desvio absoluto e a amplitude do intervalo de credibilidade empírico de 95% diminuem. No entanto, o viés do parâmetro de escala não decresceu para próximo de zero quando aumentamos o comprimento da série de observações. Já com relação as estimativas dos parâmetros obtidas usando *a priori* $\sigma \sim HC(25)$ observamos que foram um pouco melhores que as estimativas obtidas por $\sigma^2 \sim GI(0.1, 0.1)$ na maioria dos comprimentos analisados.

Tabela 6.9: Análise da acurácia das estimativas dos parâmetros para diferentes tamanhos do conjunto de dados.

		Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
$\sigma^2 \sim GI(0.1, 0.1)$	30	$r = 0.42$	-0.086	0.087	0.274	93.2%
		$K = 46$	-1.269	1.403	8.135	98.8%
		$\sigma = 37$	6.531	7.147	28.355	89.1%
	50	$r = 0.42$	-0.073	0.074	0.264	94.4%
		$K = 46$	-1.504	1.526	5.978	96.3%
		$\sigma = 37$	2.006	3.421	19.420	98.2%
	100	$r = 0.42$	-0.096	0.096	0.237	76.7%
		$K = 46$	-0.601	0.695	3.979	98.7%
		$\sigma = 37$	1.012	2.255	12.511	97.1%
	250	$r = 0.42$	-0.033	0.044	0.200	94.5%
		$K = 46$	0.149	0.332	1.939	97.5%
		$\sigma = 37$	-0.905	1.471	7.201	93.0%
	500	$r = 0.42$	-0.043	0.049	0.148	76.0%
		$K = 46$	0.257	0.322	1.113	87.0%
		$\sigma = 37$	-0.208	0.849	4.992	99.0%
$\sigma \sim HC(25)$	30	$r = 0.42$	-0.060	0.060	0.284	99.9%
		$K = 46$	-0.280	0.792	7.715	100.0%
		$\sigma = 37$	2.219	4.655	26.533	97.3%
	50	$r = 0.42$	-0.069	0.069	0.271	98.3%
		$K = 46$	-0.981	1.077	5.521	98.9%
		$\sigma = 37$	0.533	3.090	18.200	98.3%
	100	$r = 0.42$	-0.048	0.051	0.255	99.0%
		$K = 46$	-1.200	1.203	3.303	78.6%
		$\sigma = 37$	-0.998	2.282	11.760	96.2%
	250	$r = 0.42$	-0.033	0.043	0.199	94.7%
		$K = 46$	-0.373	0.430	1.853	93.3%
		$\sigma = 37$	-0.545	1.321	7.117	96.2%
	500	$r = 0.42$	-0.036	0.049	0.146	77.4%
		$K = 46$	-0.173	0.289	1.136	87.9%
		$\sigma = 37$	-0.254	0.925	4.975	96.4%

O EnKF em conjunto os métodos representados pelos *Algoritmos 3 e 4* apresentaram uma performance razoável na estimação dos estados e dos parâmetros no modelo de produção excedente logístico, desde que seja assumida a proporcionalidade das variâncias e o parâmetro q conhecido. Na prática, a suposição do parâmetro q conhecido torna inviável o uso desse método em aplicações reais.

Ao analisar as estimativas dos parâmetros K e r notamos que o *Algoritmo 3* mostrou-se robusto quanto a não normalidade das *prioris* desses parâmetros, uma vez que este algoritmo é construído com base na suposição de normalidade das distribuições a *priori* dos parâmetros contidos na função de evolução. Os parâmetros K e r podem ser considerados os principais parâmetros do modelo, pois a partir de suas estimativas são obtidas medidas para uma exploração sustentável de algum recurso biológico.

6.3 Análise de dados reais

Os dados são formados por duas séries temporais que são: série de capturas (em toneladas) anuais de camarão (*Heterocarpus reedi*) e a série temporal dos rendimentos médios de pesca, cujo a unidade de medida é quilogramas por hora de arrasto. Estas séries foram coletadas entre 1981 e 2013 ($N = 32$) na região centro sul do Chile. As duas séries são apresentadas na Figura 6.3.

Figura 6.3: *Dados de Camarão do Chile.*

Na Tabela 6.10 apresentamos algumas estatísticas descritivas das duas séries temporais.

Tabela 6.10: *Estatísticas descritivas dos dados de Camarão do Chile.*

	Min.	1ºQ.	Mediana	Média	3ºQ.	Max.	Desvio-padrão
Rendimento de Pesca (Kg Hora de arrasto)	61.00	106.00	129.50	177.94	246.50	383.00	101.48
Captura (Mil toneladas)	1.10	2.42	3.07	3.26	4.16	6.26	1.39

Na Tabela 6.10 notamos que o intervalo de variação do rendimento médio de pesca é de 61 quilogramas por horas de arrasto até 383 quilogramas por horas de arrasto, já o espaço de variação da captura foi de 1.1 mil toneladas até 6.26 mil toneladas. O coeficiente de variação das séries foram 57% e 42.6%, respectivamente.

Na modelagem dos dados consideramos a abordagem de espaço de estados no modelo de crescimento logístico para obter conhecimento sobre a biomassa anual da população de camarões. No nosso modelo a série temporal do rendimento médio de pesca (Y_t) é a variável resposta e a série de captura (H_t) será considerada constante. Com isso consideramos o seguinte modelo apresentado na seção 3.2,

$$\begin{aligned}
 y_t &= qx_t + v_t, & q > 0 \\
 x_t &= x_{t-1} + rx_{t-1} \left(1 - \frac{x_{t-1}}{K}\right) - H_t + w_t, & r > 0 \text{ e } K > 0,
 \end{aligned}$$

em que assumimos $v_t \sim N(0, \sigma^2)$ e $w_t \sim N(0, \tau^2 = \kappa\sigma^2)$. Aqui faremos apenas a suposição de proporcionalidade das variâncias com $k = 0.005$, valor calculado com

a razão das estimativas das variâncias das observações e dos estados obtidas via MCMC por Montenegro e Branco (2016), cujo as distribuições *a priori* utilizadas foram: $q \sim IG(0.1, 0.1)$, $\sigma^2 \sim IG(0.1, 0.1)$, $\tau^2 \sim IG(0.1, 0.1)$, r possui distribuição Log-normal com média 0.33 e variância 0.007 e K possui distribuição Log-normal com média 33 e variância 68.06. Utilizaremos estas mesmas distribuições *a priori* para obter as estimativas via EnK com o objetivo comparar com as estimativas obtidas via MCMC por Montenegro e Branco (2016).

Na Tabela 6.11 são apresentadas as estimativas dos parâmetros q , K , r e σ^2 obtidas usando os Algoritmos 3 e 4.

Tabela 6.11: Estimativas dos parâmetros r , K , q e σ^2 via EnKF

Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
r	0.001	0.001	0.001	0.001	0.002
K	184.039	15.474	184.362	151.357	217.595
q	0.512	0.000	0.512	0.512	0.512
σ	9.970	0.027	9.975	9.904	10.000

Na Tabela 6.12 consideramos o parâmetro q conhecido e obtemos as estimativas dos parâmetros K , r e σ^2 .

Tabela 6.12: Estimativas dos parâmetros r , K e σ^2 via EnKF

Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
r	0.416	0.048	0.413	0.320	0.508
K	47.137	3.848	46.615	41.106	55.830
σ	31.041	5.077	30.917	21.939	40.758

Na Tabela 6.13 apresentamos as estimativas obtidas por Montenegro e Branco (2016) não considerando a suposição de proporcionalidade das variâncias.

Tabela 6.13: Estimativas dos parâmetros r , K , q , σ^2 e τ^2 via MCMC.

Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
r	0.42	0.07	0.42	0.28	0.58
K	46.58	8.19	45.79	32.61	65.14
q	9.11	1.77	9.00	5.93	13.04
τ	2.57	0.93	2.40	1.27	4.85
σ	37.57	5.63	36.97	28.33	50.35

Fonte: Montenegro e Branco (2016)

Algumas medidas consideradas importantes para a administração de recursos pesqueiros são: a Produção Máxima Sustentável (*Maximum Sustainable Yield*, MSY) e medidas associadas tais como a Biomassa (B_{MSY}) e a Taxa de captura (H_{MSY}). Estas quantidades são funções dos parâmetros K e r . Na Tabela 6.14 apresentamos as estimativas dessas quantidades considerando o parâmetro q desconhecido.

Tabela 6.14: Estimativas B_{MSY} , H_{MSY} e MSY via EnKF - (q desconhecido)

Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
B_{MSY}	92.020	7.737	92.181	75.678	108.797
H_{MSY}	0.001	0.001	0.001	0.001	0.001
MSY	0.060	0.064	0.048	0.040	0.096

Na Tabela 6.15 consideramos o parâmetro q conhecido e obtemos as estimativas dos parâmetros das quantidades B_{MSY} , H_{MSY} e MSY .

Tabela 6.15: Estimativas das quantidades B_{MSY} , H_{MSY} e MSY via EnKF - (q conhecido)

Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
B_{MSY}	23.723	1.926	23.517	20.620	27.215
H_{MSY}	0.202	0.021	0.201	0.165	0.245
MSY	4.779	0.452	4.763	4.013	5.630

Na Tabela 6.16 apresentamos as estimativas obtidas via MCMC não considerando a suposição de proporcionalidade das variâncias.

Tabela 6.16: Estimativas dos parâmetros r , K , q , σ^2 e τ^2 via MCMC.

Quantidade	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
B_{MSY}	23.29	4.09	22.90	16.45	32.63
H_{MSY}	0.21	0.03	0.21	0.14	0.29
MSY	4.84	0.68	4.84	3.47	6.22

Fonte: Montenegro e Branco (2016)

Apesar dos modelos serem diferentes devido a suposição de proporcionalidade assumida no ajuste do EnKF, estamos mais interessados em comparar as estimativas dos principais parâmetros e quantidades obtidas pelo modelo que são r , K , B_{MSY} , H_{MSY} e MSY . Comparando as estimativas nas Tabelas 6.11, 6.12, 6.13, 6.14, 6.15 e 6.16 notamos que as estimativas obtidas via EnKF são extremamente afetadas quando o parâmetro q é desconhecido, no entanto, ao considerá-lo conhecido obtemos estimativas semelhantes as obtidas via MCMC, mas a variabilidade das estimativas foram subestimadas.

Na Figura 6.4 são apresentadas as estimativas das distribuições preditivas das observações considerando o parâmetro q desconhecido e conhecido.

Figura 6.4: Estimativas das distribuições preditivas das observações.

Com o parâmetro q desconhecido temos uma ideia de que o parâmetro q e os estados estão sendo mal estimados por meio das distribuições preditivas das observações. Dessa forma, a aplicação do EnKF é inviável no ajuste do modelo de crescimento logístico.

Capítulo 7

Conclusão

Nesta dissertação estudamos o filtro de Kalman Ensemble (EnKF), que é uma ferramenta para inferência dos estados em modelos de espaço de estados de alta dimensão. Este filtro em conjunto com os métodos propostos por [Anderson \(2001\)](#) e [Stroud et al. \(2016\)](#) pode estimar simultaneamente os estados e os parâmetros de um modelo. Esse processo de estimação é feito de forma sequencial, de modo que a cada nova observação a densidade *a posteriori* dos parâmetros é atualizada enquanto que a distribuição de predição e de *filtragem* do vetor de estados dado os parâmetros são obtidas pelo EnKF.

Nos resultados obtidos notamos que para fazer a estimação dos estados e dos parâmetros devemos considerar o parâmetro q conhecido e as variâncias proporcionais $\frac{\tau^2}{\sigma^2} = \kappa$ com o valor de κ predefinido e σ^2 desconhecido. Com essas condições verificamos que o método consegue estimar razoavelmente bem os parâmetros e os estados dos modelos. Observamos também que as estimativas obtidas *a posteriori* tendem a ser mais acuradas com o crescimento do comprimento da série de observações.

A escolha de fixar o parâmetro q se dá pelo fato de que as estimativas dos estados e dos demais parâmetros são extremamente afetadas quando este é desconhecido. A causa desse resultado pode estar relacionado ao problema da falta de identificabilidade, o que requer uma investigação futura.

No modelo linear, foram obtidas estimativas razoavelmente boas para os parâmetros α e σ^2 . No entanto, σ^2 apresentou estimativas enviesadas. Esse viés pode ser reduzido com o uso de algumas *prioris*, tal como foi observado na priori $\sigma \sim HC(25)$. Quanto a sensibilidade, α mostrou-se robusto à escolha das *prioris* e as estimativas do parâmetro σ^2 mostraram-se sensíveis. Já no modelo não linear, os algoritmos também estimaram razoavelmente bem os parâmetros K , r e σ^2 , apesar de as estimativas serem enviesadas, com r apresentando grande viés. Na análise de sensibilidade, os parâmetros mostraram-se sensíveis à escolha das *prioris*.

Ao estimar os parâmetros K e r o [Algoritmo 3](#) mostrou-se robusto quanto a não normalidade das *prioris* dos parâmetros r e K , uma vez que este algoritmo é construído com base na suposição de normalidade das distribuições *a priori* dos parâmetros contidos na função de evolução e geralmente na literatura é comum atribuir a distribuição Log-normal a estes parâmetros.

Na estimação dos estados no modelo linear gaussiano o EnKF conseguiu estimar bem os estados mesmo não tendo um conjunto inicial dos estados bem informativo. Enquanto que, no modelo não linear gaussiano as simulações mostraram que o EnKF possui uma performance razoável na estimação dos estados no modelo de produção excedente logístico. Observamos também que no modelo linear as estimativas das distribuições preditivas e a *posteriori* dos estados iniciais obtidas por $\sigma \sim HC(25)$ possuem uma dispersão maior do que as estimativas obtidas por $\sigma^2 \sim GI(0.1, 0.1)$.

O EnKF não consegue estimar o modelo de produção excedente logístico quando todos os parâmetros são desconhecidos e deste modo não pode ser empregado como uma alternativa aos métodos MCMC. Por outro lado, o EnKF consegue estimar distribuições a posteriori complicadas de se obter analiticamente em muitas aplicações, especialmente em casos que exigem atualizações frequentes das distribuições a posteriori à medida que novos dados são observados. Já os métodos MCMC são de uso limitado para inferência sequencial, porque sempre que uma nova observação se torna disponível, uma cadeia de Markov totalmente nova tem que ser simulada. Nesse sentido, ao contrário do filtro de Kalman Ensemble, o MCMC não é indicado para um uso sequencial.

Contudo, no filtro de Kalman Ensemble existem muitos problemas e questões pendentes que podem ser objeto de pesquisas futuras. Um dos principais problemas é o da estimação conjunta dos estados e dos parâmetros, que é pouco explorado na literatura e os métodos existentes apresentam problemas na estimação de alguns parâmetros dos modelos, como apresentado aqui. Uma outra possibilidade de pesquisa futura é a construção de um filtro *ensemble* com base em distribuições assimétricas, tais como as pertencentes à classe de mistura de escala da distribuição normal assimétrica. A construção de filtros assimétricos abriria espaço para a investigação de problemas associados a inferência em modelos de espaço de estados não lineares e não gaussianos.

Apêndice A

Apêndice

A.1 Filtro de Kalman Ensemble

A equação de predição do Filtro Kalman Ensemble é dada por

$$\tilde{\mathbf{x}}_t^{(i)} = \mathbf{M}_t \hat{\mathbf{x}}_{t-1}^{(i)} + \mathbf{w}_t^{(i)}, \quad \mathbf{w}_t^{(i)} \sim \mathcal{N}_m(\mathbf{0}, \mathbf{Q}_t) \quad i = 1, 2, \dots, L.$$

Sabendo que $\hat{\mathbf{x}}_{t-1}^{(i)} \sim N(\hat{\boldsymbol{\mu}}_{t-1}, \hat{\boldsymbol{\Sigma}}_{t-1})$, então $\tilde{\mathbf{x}}_{t-1}^{(i)}$ possui distribuição normal com os seguintes momentos,

$$\begin{aligned} E(\tilde{\mathbf{x}}_t^{(i)}) &= E(\mathbf{M}_t \hat{\mathbf{x}}_{t-1}^{(i)} + \mathbf{w}_t^{(i)}) \\ &= \mathbf{M}_t \hat{\boldsymbol{\mu}}_{t-1} \\ &= \tilde{\boldsymbol{\mu}}_t \end{aligned}$$

e

$$\begin{aligned} Var(\tilde{\mathbf{x}}_t^{(i)}) &= Var(\mathbf{M}_t \hat{\mathbf{x}}_{t-1}^{(i)} + \mathbf{w}_t^{(i)}) \\ &= \mathbf{M}_t \hat{\boldsymbol{\Sigma}}_{t-1} \mathbf{M}_t' + \mathbf{Q}_t \\ &= \tilde{\boldsymbol{\Sigma}}_t. \end{aligned}$$

Na equação de atualização

$$\hat{\mathbf{x}}_t^{(i)} = \tilde{\mathbf{x}}_t^{(i)} + \hat{\mathbf{K}}_t(\mathbf{y}_t - \tilde{\mathbf{y}}_t^{(i)})$$

em que $\hat{\mathbf{K}}_t = \tilde{\boldsymbol{\Sigma}}_t \mathbf{H}' (\mathbf{H} \tilde{\boldsymbol{\Sigma}}_t \mathbf{H}' + \mathbf{R}_t)^{-1}$, $\tilde{\mathbf{y}}_t^{(i)} = \mathbf{H}_t \tilde{\mathbf{x}}_t^{(i)} + \mathbf{v}^{(i)}$, e $\mathbf{v}^{(i)} \sim \mathcal{N}_r(\mathbf{0}, \mathbf{R}_t)$ teremos que $\hat{\mathbf{x}}_t^{(i)}$ possui distribuição normal com os seguintes momentos,

$$\begin{aligned} E(\hat{\mathbf{x}}_t^{(i)}) &= E(\tilde{\mathbf{x}}_t^{(i)} + \hat{\mathbf{K}}_t(\mathbf{y}_t - \tilde{\mathbf{y}}_t^{(i)})) \\ &= \tilde{\boldsymbol{\mu}}_t + \hat{\mathbf{K}}_t(\mathbf{y}_t - \mathbf{H}_t \tilde{\boldsymbol{\mu}}_t) \\ &= \hat{\boldsymbol{\mu}}_t \end{aligned}$$

e

$$\begin{aligned}
Var(\hat{\mathbf{x}}_t^{(i)}) &= Var(\tilde{\mathbf{x}}_t^{(i)} + \hat{\mathbf{K}}_t(\mathbf{y}_t - \tilde{\mathbf{y}}_t^{(i)})) \\
&= Var((\mathbf{I} - \hat{\mathbf{K}}_t\mathbf{H}_t)\tilde{\mathbf{x}}_t^{(i)} - \hat{\mathbf{K}}_t\mathbf{v}_t) \\
&= (\mathbf{I} - \hat{\mathbf{K}}_t\mathbf{H}_t)Var(\tilde{\mathbf{x}}_t^{(i)})(\mathbf{I} - \hat{\mathbf{K}}_t\mathbf{H}_t)' + \hat{\mathbf{K}}_tVar(\mathbf{v}_t)\hat{\mathbf{K}}_t' \\
&= (\mathbf{I} - \hat{\mathbf{K}}_t\mathbf{H}_t)\tilde{\Sigma}_t(\mathbf{I} - \hat{\mathbf{K}}_t\mathbf{H}_t)' + \hat{\mathbf{K}}_t\mathbf{R}_t\hat{\mathbf{K}}_t' \\
&= (\mathbf{I} - \hat{\mathbf{K}}_t\mathbf{H}_t)(\tilde{\Sigma}_t - \tilde{\Sigma}_t\mathbf{H}_t'\hat{\mathbf{K}}_t') + \hat{\mathbf{K}}_t\mathbf{R}_t\hat{\mathbf{K}}_t' \\
&= \tilde{\Sigma}_t - 2\hat{\mathbf{K}}_t\mathbf{H}_t\tilde{\Sigma}_t + \hat{\mathbf{K}}_t(\mathbf{H}_t\tilde{\Sigma}_t\mathbf{H}_t' + \mathbf{R}_t)\hat{\mathbf{K}}_t' \\
&= \tilde{\Sigma}_t - 2\hat{\mathbf{K}}_t\mathbf{H}_t\tilde{\Sigma}_t + \hat{\mathbf{K}}_t(\mathbf{H}_t\tilde{\Sigma}_t\mathbf{H}_t' + \mathbf{R}_t)(\tilde{\Sigma}_t\mathbf{H}'(\mathbf{H}\tilde{\Sigma}_t\mathbf{H}' + \mathbf{R}_t)^{-1})' \\
&= \tilde{\Sigma}_t - 2\hat{\mathbf{K}}_t\mathbf{H}_t\tilde{\Sigma}_t + \hat{\mathbf{K}}_t(\mathbf{H}_t\tilde{\Sigma}_t\mathbf{H}_t' + \mathbf{R}_t)((\mathbf{H}\tilde{\Sigma}_t\mathbf{H}_t' + \mathbf{R}_t)^{-1}\mathbf{H}_t\tilde{\Sigma}_t) \\
&= \tilde{\Sigma}_t - 2\hat{\mathbf{K}}_t\mathbf{H}_t\tilde{\Sigma}_t + \hat{\mathbf{K}}_t\mathbf{H}_t\tilde{\Sigma}_t \\
&= (\mathbf{I} - \hat{\mathbf{K}}_t\mathbf{H}_t)\tilde{\Sigma}_t.
\end{aligned}$$

Apêndice B

Apêndice

B.1 Efeito do tamanho da amostra ensemble nas estimativas dos parâmetros

Para verificar o efeito do tamanho da amostra *ensemble* nas estimativas dos parâmetros executamos 1000 simulações do modelo apresentado na seção 5.3. No processo de inferência foram utilizados os Algoritmos 3 e 4 com o tamanho da amostra *ensemble* igual a $L = 30, 50, 100, 200, 500$ e 1000, uma série temporal de comprimento $N = 100$ e mil réplicas. No Algoritmo 4 usamos a amostragem por importância para estimar σ^2 com intervalo de variação $(0, 20]$ e em todos os casos usamos uma amostra *ensemble* inicial não informativa $\hat{x}_0^{(i)} \sim N(0, \tau^2 = \kappa\sigma^2)$, em que assumiremos $\kappa = 0.25$. Também assumimos que o parâmetro σ^2 possui distribuição $GI(0.1, 0.1)$ e α tem distribuição $N(0, 10)$. Os resultados são apresentados nas Tabelas B.1 e B.2.

Tabela B.1: Estimativas dos parâmetros α e σ^2 para diferentes tamanhos da amostra ensemble.

L	Parâmetros	Média	Desvio-padrão	Mediana	Q2.5%	Q97.5%
30	$\alpha = 1$	0.994	0.005	0.994	0.986	1.002
	$\sigma^2 = 4$	5.000	1.029	4.900	3.442	7.019
50	$\alpha = 1$	0.994	0.005	0.994	0.985	1.004
	$\sigma^2 = 4$	4.926	1.026	4.829	3.313	7.023
100	$\alpha = 1$	0.994	0.006	0.994	0.983	1.006
	$\sigma^2 = 4$	4.829	1.023	4.734	3.176	7.011
200	$\alpha = 1$	0.994	0.006	0.994	0.982	1.006
	$\sigma^2 = 4$	4.804	1.015	4.707	3.124	7.004
500	$\alpha = 1$	0.994	0.006	0.994	0.981	1.006
	$\sigma^2 = 4$	4.779	1.015	4.681	3.086	7.014
1000	$\alpha = 1$	0.994	0.007	0.994	0.981	1.007
	$\sigma^2 = 4$	4.761	1.014	4.665	3.065	7.005

Tabela B.2: Análise da acurácia das estimativas dos parâmetros para diferentes tamanhos da amostra *ensemble*.

	Parâmetros	Viés	Desvio Absoluto	Amplitude do ICr	Taxa de Cobertura
30	$\alpha = 1$	-0.006	0.007	0.016	62.0%
	$\sigma^2 = 4$	1.000	1.063	3.578	82.4%
50	$\alpha = 1$	-0.006	0.006	0.019	77.1%
	$\sigma^2 = 4$	0.926	0.986	3.710	87.5%
100	$\alpha = 1$	-0.006	0.006	0.023	93.9%
	$\sigma^2 = 4$	0.829	0.896	3.835	93.5%
200	$\alpha = 1$	-0.006	0.006	0.024	98.0%
	$\sigma^2 = 4$	0.804	0.881	3.879	94.6%
500	$\alpha = 1$	-0.006	0.006	0.025	99.9%
	$\sigma^2 = 4$	0.779	0.853	3.927	96.3%
1000	$\alpha = 1$	-0.006	0.006	0.026	99.9%
	$\sigma^2 = 4$	0.761	0.831	3.940	96.2%

Com a análise das tabelas percebemos que à medida que aumentamos o tamanho da amostra *ensemble* as estimativas tendem a melhorar, uma vez que o viés e o desvio absoluto diminuem. No entanto, para amostras *ensemble* de tamanho abaixo de 100 as taxas de cobertura ficam abaixo de 95% e para amostras *ensemble* de tamanho acima de 100 as taxas começam a ficar superiores a 95%. Evensen (2009a) sugere usar, caso necessário, o tamanho da amostra *ensemble* maior que 100 quando o conjunto inicial é não informativo, já que o EnKF converge para o KF em um modelo de espaço de estados linear normal. Assim, um tamanho da amostra *ensemble* igual a 200 foi considerado nas simulações deste trabalho.

B.2 Comparando as estimativas do EnKF com as estimativas obtidas via o pacote DLM

Para comparar as estimativas dos parâmetros obtidas via EnKF com as estimativas obtidas via o pacote DLM (Petrís, 2010) do software R consideramos o modelo apresentado na seção 5.3 com os seguintes comprimentos da série de observações $N = 30, 50, 100, 250$ e 500 . Em cada um desses comprimentos consideramos $x_0 = 20$, $\alpha = 1$, $\sigma^2 = 4$ e $\kappa = 0.25$.

No processo de inferência via EnKF foram utilizados os Algoritmos 3 e 4 com o tamanho da amostra *ensemble* igual a $L = 200$ e no Algoritmo 4 usamos a amostragem por importância para estimar σ^2 com intervalo de variação $(0, 20]$. Já o pacote DLM fornece algumas funções para facilitar a obtenção da posteriori dos parâmetros via MCMC. Assim, para obter as estimativas executamos 150 iterações e descartamos as 50 primeiras. Este número de iterações é razoável para a convergência da cadeia. Em ambos os métodos utilizamos as seguintes distribuições a priori: $\hat{x}_0^{(i)} \sim N(0, \tau^2 = 0.25\sigma^2)$, $\sigma^2 \sim GI(0.1, 0.1)$ e $\alpha \sim N(0, 10)$.

Figura B.1: Na gráfico do lado esquerdo apresentamos as estimativas do parâmetro α obtidas via EnKF e MCMC. Do lado direito apresentamos as estimativas de σ^2 obtidas pelos os dois métodos.

Na Figura B.1 notamos que ambos os métodos estão obtendo estimativas razoáveis das distribuições a posteriori dos parâmetros α e σ^2 . Dessa forma, em modelos de espaço de estados lineares o EnKF pode ser considerado uma alternativa ao MCMC. A diferença nas estimativas a posteriori obtidas pelos métodos ocorre principalmente nas distribuições a posteriori do parâmetro α quando o comprimento da série é pequeno.

Na Tabela B.3 comparamos o tempo computacional para a execução dos códigos de ambos os métodos.

Tabela B.3: Tempo computacional em segundos para obter as estimativas em um série de comprimento N .

N	30	50	100	250	500
EnKF	0.039	0.411	0.462	0.659	1.122
MCMC	0.433	0.635	1.104	2.348	4.532

Na Tabela B.3 notamos que o tempo de execução do código usando o EnKF é bem inferior ao tempo de execução usando o pacote DLM. Os códigos são apresentados em D.3.

Apêndice C

Apêndice

C.1 Figuras do Capítulo 5

C.1.1 Figuras da seção 5.1

Conforme indicado na seção 5.1, apresentamos nas Figuras C.1, C.2 e C.3 a estimativa da distribuição preditiva e de *filtragem* de cada estado para o comprimento da série de observações igual a $N = 30, 250$ e 500 .

Figura C.1: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 30 - (σ^2 e τ^2).

Figura C.2: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 250 -(σ^2 e τ^2).

Figura C.3: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 500 -(σ^2 e τ^2).

C.1.2 Figuras da seção 5.3

Conforme indicado na seção 5.3, apresentamos nas Figuras C.4, C.5, C.6, C.7, C.8, C.9, C.10 e C.11 a estimativa da distribuição preditiva e de *filtragem* de cada estado para o comprimento da série de observações igual a $N = 30, 50, 250$ e 500 .

Caso I

Figura C.4: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 30 - Caso I.

Figura C.5: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 50 - Caso I.

Figura C.6: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 250 - Caso I.

Figura C.7: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 500 - Caso I.

Caso II

Figura C.8: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 30 - Caso II.

Figura C.9: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 50 - Caso II.

Figura C.10: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 250 - Caso II.

Figura C.11: Estimativa da distribuição preditiva e a posteriori de cada estado para uma série de observações de comprimento igual a 500 - Caso II.

C.2 Figuras do Capítulo 6

Figura C.12: Estimativas das distribuições preditivas e a posteriori de cada estado para um modelo não linear e uma série de observações de comprimento igual a 30 com $\sigma^2 \sim GI(0.1, 0.1)$.

Figura C.13: Estimativas das distribuições preditivas e a posteriori de cada estado para um modelo não linear e uma série de observações de comprimento igual a 50 com $\sigma^2 \sim GI(0.1, 0.1)$.

Figura C.14: Estimativas das distribuições preditivas e a posteriori de cada estado para um modelo não linear e uma série de observações de comprimento igual a 250 com $\sigma^2 \sim GI(0.1, 0.1)$.

Figura C.15: Estimativas das distribuições preditivas e a posteriori de cada estado para um modelo não linear e uma série de observações de comprimento igual a 500 com $\sigma^2 \sim GI(0.1, 0.1)$.

Figura C.16: Estimativas das distribuições preditivas e a posteriori de cada estado para um modelo não linear e uma série de observações de comprimento igual a 30 com $\sigma \sim HC(25)$.

Figura C.17: Estimativas das distribuições preditivas e a posteriori de cada estado para um modelo não linear e uma série de observações de comprimento igual a 50 com $\sigma \sim HC(25)$.

Figura C.18: Estimativas das distribuições preditivas e a posteriori de cada estado para um modelo não linear e uma série de observações de comprimento igual a 250 com $\sigma \sim HC(25)$.

Figura C.19: Estimativas das distribuições preditivas e a posteriori de cada estado para um modelo não linear e uma série de observações de comprimento igual a 500 com $\sigma \sim HC(25)$.

Apêndice D

Códigos R

D.1 Códigos R do capítulo 5

D.1.1 Seção 5.3 - Estimação de α e σ^2 .

```
require(invgamma, ggplot2, xtable, gridExtra)
Re <- 1000
N <- 100
L <- 200
X <- numeric(N)

X[1] <- 20
q <- 1
alpha <- 1
kappa <- 0.25
sigma <- 2
tau <- sqrt(kappa*sigma^2)

for(i in 2:N){
X[i] <- alpha*X[i-1] + rnorm(1,0,tau)
}

theta <- seq(0.01,20,len = 2*L) # Espaço de variação de sigma^2

X.tilde.resume <- X.hat.resume <- alpha_resume <- Sig_resume <- list()
alpha.resume <- SigM.resume <- matrix(0,ncol = 5, nrow = Re)
alpha.resume2 <- SigM.resume2 <- matrix(0,ncol = 3, nrow = Re)
porcent1 <- porcent2 <- porcent3 <- matrix(0,ncol = 1, nrow = Re)

for(j in 1:Re){

Y <- q * X + rnorm(N,0,sigma)

X.tilde <- X.hat <- alpha.Matr <- MatSig <- matrix(0,ncol = L,nrow = N+1)
```



```

Yi <- c(0,Y)

#priors
alpha.Matr[1,] <- rnorm(L, 0,10)
tau0 <- sqrt(kappa*rinvgamma(L, shape = 0.1, rate = 0.1))
X.hat[1,] <- sapply( tau0, FUN = function(o) rnorm(1,0,o))
log.theta <- dinvgamma(theta,shape = 0.1, rate = 0.1, log = T)

#EnKF
for( i in 2:(N+1)){
  X.tilde[i,] <- alpha.Matr[i-1,]*X.hat[i-1,]

  log.theta <- dnorm(Yi[i],q*mean(X.tilde[i,]), sqrt(q^2*(var(X.tilde[i,]) +
    kappa*theta ) + theta ) , log = T) + log.theta
  wi <- exp(log.theta - max(log.theta))
  wi <- wi/sum(wi)

  MatSig[i,] <- sample(theta, size = L, replace = T, prob = wi)
  Tau <- sqrt(kappa*MatSig[i,])

  X.tilde[i,] <- X.tilde[i,] + sapply(Tau, function(o) rnorm(1,0,o))

  gain.alpha <- q*cov(X.tilde[i,],alpha.Matr[i-1,])/(q^2*var(X.tilde[i,]) + MatSig[i,])
  gain.Kalman <- q*var(X.tilde[i,])/(q^2*var(X.tilde[i,]) + MatSig[i,])

  yobs <- q*X.tilde[i,] - sapply(MatSig[i,], function(o) rnorm(1,0,sqrt(o)))
  alpha.Matr[i,] <- alpha.Matr[i-1,] + gain.alpha*(Yi[i] - yobs )
  X.hat[i,] <- X.tilde[i,] + gain.Kalman*(Yi[i] - yobs )

}

SigM <- MatSig[-1,]
alpha.Ma <- alpha.Matr[-1,]
X..hat <- X.hat[-1,]
X..tilde <- X.tilde[-1,]

# Medidas Resumo

X.hat.resume[[j]] <- t(apply(X..hat, 1, FUN = function(z)
  c(mean(z),sd(z),min(z),quantile(z,c(0.25,0.5,0.75)),max(z))))
X.tilde.resume[[j]] <- t(apply(X..tilde, 1, FUN = function(z)
  c(mean(z),sd(z),min(z),quantile(z,c(0.25,0.5,0.75)),max(z))))
alpha_resume[[j]] <- t(apply( alpha.Ma, 1, FUN = function(z)
  c(mean(z),quantile(z,c(0.025,0.975)))))
Sig_resume[[j]] <- t(apply(SigM, 1, FUN = function(z)
  c(mean(z),quantile(z,c(0.025,0.975)))))

```

```

porcent1[j] <- sum(X.hat.resume[[j]][,4] <= X & X <= X.hat.resume[[j]][,6])/N

SigM.resume[j,] <- c(mean(SigM[N,]),sd(SigM[N,]),quantile(SigM[N,],
c(0.5,0.025,0.975)))
alpha.resume[j,] <- c(mean(alpha.Ma[N,]),sd(alpha.Ma[N,]),quantile(alpha.Ma[N,],
c(0.5,0.025,0.975)))

SigM.resume2[j,] <- c(mean(SigM[N,]) - sigma^2, abs(mean(SigM[N,]) - sigma^2),
quantile(SigM[N,],0.975) - quantile(SigM[N,],0.025))
alpha.resume2[j,] <- c(mean(alpha.Ma[N,]) - alpha, abs(mean(alpha.Ma[N,]) - alpha),
quantile(alpha.Ma[N,],0.975) - quantile(alpha.Ma[N,],0.025))

}

#Procentagens
round(mean(porcent1),3)
porcent2 <- sum(SigM.resume[,4] <= sigma^2 & sigma^2 <= SigM.resume[,5])/Re
porcent3 <- sum(alpha.resume[,4] <= alpha & alpha <= alpha.resume[,5])/Re
#Tabelas
xtable(rbind(colMeans(alpha.resume), colMeans(SigM.resume)), digits = 3)
xtable(rbind(c(colMeans(alpha.resume2),mean(porcent3)),
c(colMeans(SigM.resume2),mean(porcent2))), digits = 3)

```

D.1.2 Seção 5.4 - Estimação de q , α e σ^2 .

```

require(LaplacesDemon, ggplot2, xtable, gridExtra)
Re <- 1000; N <- 100; L <- 200; X <- numeric(N); X[1] <- 20; q <- 1; alpha <- 1
kappa <- 0.25
sigma <- 2
tau <- sqrt(kappa * sigma^2)
for (i in 2:N) X[i] <- alpha * X[i - 1] + rnorm(1, 0, tau) #Simulando os estados

X.tilde.resume <- X.hat.resume <- alpha_resume <- q_resume <- Sig_resume <- list()
alpha.resume <- q.resume <- SigM.resume <- matrix(0, ncol = 5, nrow = Re)
alpha.resume2 <- q.resume2 <- SigM.resume2 <- matrix(0, ncol = 3, nrow = Re)
porcent1 <- porcent2 <- porcent3 <- porcent4 <- matrix(0, ncol = 1, nrow = Re)
theta <- seq(-5, 5, len = L)
theta1 <- seq(0.01, 20, len = L)
d.theta <- expand.grid(x = theta, y = theta1)
qtheta <- d.theta[, 1]
sigtheta <- d.theta[, 2]

for (j in 1:Re){
Y <- q * X + rnorm(N, 0, sigma) # Simulando as observações
X.tilde <- X.hat <- alpha.Matr <- q.Matr <- MatSig <- matrix(0,
ncol = L, nrow = N + 1)
# Distribuições a priori
alpha.Matr[1, ] <- rnorm(L, 1, 2)

```

```

# Passo inicial
Yi <- c(0, Y)
tau0 <- sqrt(kappa * rhalfcauchy(L, 25)^2)
X.hat[1, ] <- sapply(tau0, FUN = function(o) rnorm(1, 0, o))
log.theta <- dnorm(d.theta[, 1], 1, 5, log = T) + dhalfcauchy(sigtheta, 25, log = T)

#Algoritmos EnKF
for (i in 2:(N + 1)){
  X.tilde[i, ] <- alpha.Matr[i - 1, ] * X.hat[i - 1, ]

  # Amostragem por importância para estimar sigma/q
  log.theta <- dnorm(Yi[i], qtheta * mean(X.tilde[i, ]), sqrt(qtheta^2 *
    (var(X.tilde[i, ]) + kappa * sigtheta^2) + sigtheta^2), log = T) + log.theta
  wi <- exp(log.theta - max(log.theta))
  wi <- wi/sum(wi)
  indice <- sample(1:40000, size = L, replace = T, prob = wi)
  q.Matr[i, ] <- d.theta[indice, 1]
  MatSig[i, ] <- d.theta[indice, 2]^2
  Tau <- sqrt(kappa * MatSig[i, ])

  # EnKF conjunto
  X.tilde[i, ] <- X.tilde[i, ] + sapply(Tau, function(o) rnorm(1, 0, o))
  gain.alpha <- q.Matr[i, ] * cov(X.tilde[i, ], alpha.Matr[i - 1, ])/
    (q.Matr[i, ]^2 * var(X.tilde[i, ]) + MatSig[i, ])
  gain.Kalman <- q.Matr[i, ] * var(X.tilde[i, ])/
    (q.Matr[i, ]^2 * var(X.tilde[i, ]) + MatSig[i, ])
  yobs <- q.Matr[i, ] * X.tilde[i, ] - sapply(MatSig[i, ],
    function(o) rnorm(1, 0, sqrt(o)))
  alpha.Matr[i, ] <- alpha.Matr[i - 1, ] + gain.alpha * (Yi[i] - yobs)
  X.hat[i, ] <- X.tilde[i, ] + gain.Kalman * (Yi[i] - yobs)
}

#Armazenando as quantidades
q.M <- q.Matr[-1, ]
SigM <- MatSig[-1, ]
alpha.Ma <- alpha.Matr[-1, ]
X..hat <- X.hat[-1, ]
X..tilde <- X.tilde[-1, ]
X.hat.resume[[j]] <- t(apply(X..hat, 1, FUN = function(z) c(mean(z),
  sd(z), min(z), quantile(z, c(0.25, 0.5, 0.75)), max(z)))))
X.tilde.resume[[j]] <- t(apply(X..tilde, 1, FUN = function(z) c(mean(z),
  sd(z), min(z), quantile(z, c(0.25, 0.5, 0.75)),
  max(z)))))
alpha.resume[[j]] <- t(apply(alpha.Ma, 1, FUN = function(z) c(mean(z),
  quantile(z, c(0.025, 0.975)))))
Sig.resume[[j]] <- t(apply(SigM, 1, FUN = function(z) c(mean(z),
  quantile(z, c(0.025, 0.975)))))
q.resume[[j]] <- t(apply(q.M, 1, FUN = function(z) c(mean(z), quantile(z,
  c(0.025, 0.975)))))

```

```

porcent1[j] <- sum(X.hat.resume[[j]][, 4] <= X & X <= X.hat.resume[[j]][, 6])/N
q.resume[j, ] <- c(mean(q.M[N, ]), sd(q.M[N, ]), quantile(q.M[N, ],
  c(0.5, 0.025, 0.975)))
SigM.resume[j, ] <- c(mean(SigM[N, ]), sd(SigM[N, ]), quantile(SigM[N, ],
  c(0.5, 0.025, 0.975)))
alpha.resume[j, ] <- c(mean(alpha.Ma[N, ]), sd(alpha.Ma[N, ]), quantile(alpha.Ma[N, ],
  c(0.5, 0.025, 0.975)))
# Quantidades
q.resume2[j, ] <- c(mean(q.M[N, ]) - q, abs(mean(q.M[N, ]) - q),
  quantile(q.M[N, ], 0.975) - quantile(q.M[N, ], 0.025))
SigM.resume2[j, ] <- c(mean(SigM[N, ]) - sigma^2, abs(mean(SigM[N, ]) - sigma^2),
  quantile(SigM[N, ], 0.975) - quantile(SigM[N, ], 0.025))
alpha.resume2[j, ] <- c(mean(alpha.Ma[N, ]) - alpha, abs(mean(alpha.Ma[N, ]) - alpha),
  quantile(alpha.Ma[N, ], 0.975) - quantile(alpha.Ma[N, ],
  ], 0.025))
}
# Procentagens
round(mean(porcent1), 3)
porcentq <- sum(q.resume[, 4] <= q & q <= q.resume[, 5])/Re
porcent2 <- sum(SigM.resume[, 4] <= sigma^2 & sigma^2 <= SigM.resume[, 5])/Re
porcent3 <- sum(alpha.resume[, 4] <= alpha & alpha <= alpha.resume[, 5])/Re
xtable(rbind(colMeans(q.resume), colMeans(alpha.resume), colMeans(SigM.resume)),
  digits = 3)
xtable(rbind(c(colMeans(q.resume2), mean(porcentq)), c(colMeans(alpha.resume2),
  mean(porcent3)), c(colMeans(SigM.resume2), mean(porcent2))), digits = 3)

```

D.2 Códigos R do capítulo 6

D.2.1 Seção 6.2 - Estimação de r , K e σ^2

```

require(invgamma, ggplot2, xtable, gridExtra)

Re <- 1000; n <- 100; L <- 200; X <- numeric(n)

X[1] <- 10
q <- 9
r <- 0.42
K <- 46
s <- 0.005
sigma <- 37
tau <- sqrt(s*sigma^2)

for(i in 2:n){ #simulando os estados
X[i] <- X[i-1] + r*X[i-1]*(1-X[i-1]/K ) + rnorm(1,0,tau)
}

theta <- seq(0.01,10000,len = 1000*L)

```

```

X.tilde.resume <- X.hat.resume <- r_resume <- K_resume <- Sig_resume <- list()
K.resume <- r.resume <- SigM.resume <- matrix(0,ncol = 5, nrow = Re)
K.resume2 <- r.resume2 <- SigM.resume2 <- matrix(0,ncol = 3, nrow = Re)
porcent1 <- porcent2 <- porcent3 <- matrix(0,ncol = 1, nrow = Re)

for(j in 1:Re){
Y <- q * X + rnorm(n,0,sigma) #simulando as observações
X.tilde <- X.hat <- K.Matr <- r.Matr <- MatSig <- matrix(0,ncol = L,nrow = n+1)

#distribuições a priori
K.Matr[1,] <- rlnorm(L, meanlog = 3.466195, sdlog = 0.2462207)
r.Matr[1,] <- rlnorm(L, meanlog = -1.138975, sdlog = 0.2462207)

#Passo inicial
Yi <- c(0,Y)
tau0 <- sqrt(s*rinvgamma(L, shape = 0.1, rate = 0.1))
X.hat[1,] <- sapply( tau0 , FUN = function(o) rnorm(1,0,o))

log.theta <- dinvgamma(theta, shape = 0.1, rate = 0.1, log = T)

for( i in 2:(n+1)){
X.tilde[i,] <- X.hat[i-1,] + r.Matr[i-1,]*X.hat[i-1,]*(1- X.hat[i-1,]/K.Matr[i-1,])

#amostragem por importância
log.theta <- dnorm(Yi[i],q*mean(X.tilde[i,]), sqrt(q^2*(var(X.tilde[i,]) +
s*theta) + theta) , log = T) + log.theta
wi <- exp(log.theta - max(log.theta))
wi <- wi/sum(wi)

MatSig[i,] <- sample(theta, size = L, replace = T, prob = wi)

Tau <- sqrt(s*MatSig[i,])

X.tilde[i,] <- pmax(X.tilde[i,] + sapply(Tau, function(o) rnorm(1,0,o)),0.001)

yobs <- q*X.tilde[i,] + sapply(MatSig[i,], function(o) rnorm(1,0,sqrt(o)))

#Ganho de Kalman
gain.K <- q*cov(X.tilde[i,],K.Matr[i-1,])/(q^2*var(X.tilde[i,]) + MatSig[i,])
gain.r <- q*cov(X.tilde[i,],r.Matr[i-1,])/(q^2*var(X.tilde[i,]) + MatSig[i,])
gain.Kalman <- q*var(X.tilde[i,])/(q^2*var(X.tilde[i,]) + MatSig[i,])

K.Matr[i,] <- pmax(K.Matr[i-1,] + gain.K*(Yi[i] - yobs ),0.001)
r.Matr[i,] <- pmax(r.Matr[i-1,] + gain.r*(Yi[i] - yobs ),0.001)
X.hat[i,] <- pmax( X.tilde[i,] + gain.Kalman*(Yi[i] - yobs ),0.001)
}
SigM <- sqrt(MatSig[-1,])
K.Ma <- K.Matr[-1,]

```

```

r.Ma <- r.Matr[-1,]
X..hat <- X.hat[-1,]
X..tilde <- X.tilde[-1,]

X.hat.resume[[j]] <- t(apply(X..hat, 1, FUN = function(z) c(mean(z),sd(z),min(z),
quantile(z,c(0.25,0.5,0.75)),max(z))))
X.tilde.resume[[j]] <- t(apply(X..tilde, 1, FUN = function(z) c(mean(z),sd(z),min(z),
quantile(z,c(0.25,0.5,0.75)),max(z))))
r_resume[[j]] <- t(apply( r.Ma, 1, FUN = function(z) c(mean(z),
quantile(z,c(0.025,0.975)))))
K_resume[[j]] <- t(apply( K.Ma, 1, FUN = function(z) c(mean(z),
quantile(z,c(0.025,0.975)))))
Sig_resume[[j]] <- t(apply(SigM, 1, FUN = function(z) c(mean(z),
quantile(z,c(0.025,0.975)))))

porcent1[j] <- sum(X.hat.resume[[j]][,4] <= X & X <= X.hat.resume[[j]][,6])/n

SigM.resume[j,] <- c(mean(SigM[n,]),sd(SigM[n,]),quantile(SigM[n,],c(0.5,0.025,0.975)))
r.resume[j,] <- c(mean(r.Ma[n,]),sd(r.Ma[n,]),quantile(r.Ma[n,],c(0.5,0.025,0.975)))
K.resume[j,] <- c(mean(K.Ma[n,]),sd(K.Ma[n,]),quantile(K.Ma[n,],c(0.5,0.025,0.975)))

SigM.resume2[j,] <- c(mean(SigM[n,]) - sigma, abs(mean(SigM[n,]) - sigma),
quantile(SigM[n,],0.975) - quantile(SigM[n,],0.025))
r.resume2[j,] <- c(mean(r.Ma[n,]) - r, abs(mean(r.Ma[n,]) - r),
quantile(r.Ma[n,],0.975) - quantile(r.Ma[n,],0.025))
K.resume2[j,] <- c(mean(K.Ma[n,]) - K, abs(mean(K.Ma[n,]) - K),
quantile(K.Ma[n,],0.975) - quantile(K.Ma[n,],0.025))
}

#Procentagens
round(mean(porcent1),3)
porcent2 <- sum(SigM.resume[,4] <= sigma & sigma <= SigM.resume[,5])/Re
porcent3 <- sum( r.resume[,4] <= r & r <= r.resume[,5])/Re
porcent4 <- sum( K.resume[,4] <= K & K <= K.resume[,5])/Re

#Tabelas
xtable(rbind( colMeans(r.resume),
colMeans(K.resume),
colMeans(SigM.resume)), digits = 3)

xtable(rbind( c(colMeans(r.resume2),mean(porcent3)),
c(colMeans(K.resume2),mean(porcent4)),
c(colMeans(SigM.resume2),mean(porcent2))), digits = 3)

```

D.2.2 Seção 6.3 - Estimação de r , K e σ^2

```

require(invgamma, ggplot2, xtable, gridExtra)
n <- 32
L <- 200
X <- numeric(n)
s <- 0.005
q <- 9

Y <- c(0.11, 0.144, 0.14, 0.089, 0.089, 0.085, 0.106, 0.107, 0.103, 0.106, 0.108,
      0.117, 0.11, 0.141, 0.125, 0.134, 0.107, 0.089, 0.061, 0.074, 0.205, 0.203, 0.234, 0.234, 0.326,
      0.284, 0.383, 0.367, 0.326, 0.311, 0.325, 0.351)*1000

Ht <- c(2.942, 3.394, 5.562, 2.321, 1.384, 1.627, 2.303, 2.616, 3.042, 4.157, 4.413, 4.46, 4.176,
      5.458, 6.262, 6.163, 5.05, 3.729, 2.459, 1.103, 1.195, 1.706, 1.681, 2.454, 3.135, 3.197, 3.055,
      3.033, 3.488, 3.173, 2.603, 3.088)

theta <- seq(0.01, 10000, len = 1000*L)
X.tilde <- X.hat <- K.Matr <- r.Matr <- MatSig <- Bmsy <- msy <-
Hmsy <- matrix(0, ncol = L, nrow = n+1)

#priors
K.Matr[1,] <- rlnorm(L, meanlog = 3.466195, sdlog = 0.2462207)
r.Matr[1,] <- rlnorm(L, meanlog = -1.138975, sdlog = 0.2462207)

#Passo inicial
Yi <- c(0, Y)
Hti <- c(0, Ht)
tau0 <- sqrt(s*rinvgamma(L, shape = 0.1, rate = 0.1))
X.hat[1,] <- sapply(tau0, FUN = function(o) rnorm(1, 0, o))
log.theta <- dinvgamma(theta, shape = 0.1, rate = 0.1, log = T)
for( i in 2:(n+1)){
  X.tilde[i,] <- X.hat[i-1,] + r.Matr[i-1,]*X.hat[i-1,]*(1- X.hat[i-1,]/
  K.Matr[i-1,]) - Hti[i]

  #Amostragem por importância
  log.theta <- dnorm(Yi[i], q*mean(X.tilde[i,]), sqrt(q^2*(var(X.tilde[i,]) +
  s*theta) + theta) , log = T) + log.theta
  wi <- exp(log.theta - max(log.theta))
  wi <- wi/sum(wi)

  MatSig[i,] <- sample(theta, size = L, replace = T, prob = wi)
  Tau <- sqrt(s*MatSig[i,])
  X.tilde[i,] <- pmax(X.tilde[i,] + sapply(Tau, function(o) rnorm(1, 0, o)), 0.001)
  yobs <- q*X.tilde[i,] + sapply(MatSig[i,], function(o) rnorm(1, 0, sqrt(o)))

  gain.K <- q*cov(X.tilde[i,], K.Matr[i-1,])/(q^2*var(X.tilde[i,]) + MatSig[i,])
  gain.r <- q*cov(X.tilde[i,], r.Matr[i-1,])/(q^2*var(X.tilde[i,]) + MatSig[i,])
  gain.Kalman <- q*var(X.tilde[i,])/(q^2*var(X.tilde[i,]) + MatSig[i,])

```

```

K.Matr[i,] <- pmax(K.Matr[i-1,] + gain.K*(Yi[i] - yobs ),0.001)
r.Matr[i,] <- pmax(r.Matr[i-1,] + gain.r*(Yi[i] - yobs ),0.001)
X.hat[i,] <- pmax( X.tilde[i,] + gain.Kalman*(Yi[i] - yobs ),0.001)
}
#Armazenando as quantidades
SigM <- MatSig[-1,]
K.Ma <- K.Matr[-1,]
r.Ma <- r.Matr[-1,]
X..hat <- X.hat[-1,]
X..tilde <- X.tilde[-1,]
X.hat.resume <- t(apply(X..hat, 1, FUN = function(z) c(mean(z),sd(z),min(z),
quantile(z,c(0.25,0.5,0.75))),max(z))))
X.tilde.resume <- t(apply(X..tilde, 1, FUN = function(z) c(mean(z),sd(z),min(z),
quantile(z,c(0.25,0.5,0.75))),max(z))))

SigM.resume <- c(mean(sqrt(SigM[n,])),sd(sqrt(SigM[n,])),quantile(sqrt(SigM[n,]),
c(0.5,0.025,0.975)))
r.resume <- c(mean(r.Ma[n,]),sd(r.Ma[n,]),quantile(r.Ma[n,],c(0.5,0.025,0.975)))
K.resume <- c(mean(K.Ma[n,]),sd(K.Ma[n,]),quantile(K.Ma[n,],c(0.5,0.025,0.975)))
Bmsy <- c(mean(K.Ma[n,]/2),sd(K.Ma[n,]/2),quantile(K.Ma[n,]/2,c(0.5,0.025,0.975)))
msy <-c(mean(K.Ma[n,]*r.Ma[n,]/4),sd(K.Ma[n,]*r.Ma[n,]/4),
quantile(K.Ma[n,]*r.Ma[n,]/4,c(0.5,0.025,0.975)))
Hmsy <- c(mean(r.Ma[n,]/2),sd(r.Ma[n,]/2),quantile(r.Ma[n,]/2,c(0.5,0.025,0.975)))
#Tabela

xtable(rbind( r.resume, K.resume, SigM.resume ), digits = 3)
xtable(rbind( Bmsy, Hmsy,msy ), digits = 3)

```

D.3 Códigos do Apêndice B.2

```

require(invgamma,ggplot2,xtable,gridExtra,dlm)

par(mfrow = c(4,2))
Tempo <-matrix(NA,ncol = 5,nrow = 2)
r.resume <- SigM.resume <- matrix(0,ncol=5,nrow = 10)
n <- c(30,50,100,250,500)
L <- 200

Ar1Gibbs <- function(y, alpha.mu0 , alpha.tau0 , shapeY, rateY, burn , MC){

#####MCMC - Inicialização#####
M <- length(y)
gibbsX <- list()
gibbsalpha <- gibbsV <- c()

alpha.init <- rnorm(1,mean=alpha.mu0 ,sd=sqrt(alpha.tau0))
V.init <- 1/rgamma(1,shapeY,rateY)

```



```

mod          <- dlm(m0=0, C0=0.25*V.init, FF=1, V = V.init , GG = alpha.init,
W=0.25*V.init )

#####Gibbs Sampling#####

for (i in 1:MC){

if(i%%10==0){ cat('# Iteration', i,'- phi:',round(median(gibbsalpha[1:i],
na.rm=TRUE),3),'\n'); flush.console()}
# Gerando os estados por FFBS

modFilt          <- dlmFilter(y, mod, simplify=TRUE) #Estimativa do modelo filtragem
X.estado         <- dlmBSample(modFilt) #amostra a posteriori dos estados
gibbsX[[i]]      <- X.estado  # armazenando a amostra

# generate the V_i
y.center         <- rep(NA,M)
for(t in 1:M) y.center[t] <- y[t] - X.estado[t+1,]

SSy <- sum(y.center^2)
gibbsV[i] <- 1/rgamma(1, shape=shapeY+M/2, rate=rateY+SSy/2)
mod$V <- gibbsV[i]
mod$W <- 0.25*gibbsV[i]

alpha.tau <- 1/(1/alpha.tau0 + 1/mod$W * crossprod(X.estado[-1,1]))

alpha.mu <- alpha.tau * (1/mod$W * sum((X.estado[-1,1])*X.estado[-M,1]) +
alpha.mu0/alpha.tau0)

mod$GG <- gibbsalpha[i] <- rnorm(1, alpha.mu, sd=sqrt(alpha.tau))
}

return( list(x.estado = gibbsX[-(1:burn)], Alpha = gibbsalpha[-(1:burn)],
V = gibbsV[-(1:burn)]))
}

#Comparação

for(k in 1:length(n)){
X <- numeric(n[k])

X[1] <- 20
q <- 1
r <- 1
s <- 0.25
sigma <- 2
tau <- sqrt(s*sigma^2)

```

```

for(i in 2:n[k]){
  set.seed(1000*i)
  X[i] <- r*X[i-1] + rnorm(1,0,tau)
}

Y <- q * X + rnorm(n[k],0,sigma)

theta <- seq(0.01,20,len = 2*L)

X.tilde <- X.hat <- r.Matr <- MatSig <- matrix(0,ncol = L,nrow = n[k]+1)

#priors
r.Matr[1,] <- rnorm(L, 0,10)

#Passo inicia
Yi <- c(0,Y)
tau0 <- sqrt(s*rinvgamma(L, shape = 0.1, rate = 0.1))
X.hat[1,] <- sapply( tau0, FUN = function(o) rnorm(1,0,o))

log.theta <- dinvgamma(theta,shape = 0.1, rate = 0.1, log = T)

Tempo1 <- system.time(
for( i in 2:(n[k]+1)){
  X.tilde[i,] <- r.Matr[i-1,]*X.hat[i-1,]

  log.theta <- dnorm(Yi[i],q*mean(X.tilde[i,]), sqrt(q^2*(var(X.tilde[i,]) + s*theta) +
  wi <- exp(log.theta - max(log.theta))
  wi <- wi/sum(wi)

  MatSig[i,] <- sample(theta, size = L, replace = T, prob = wi)

  Tau <- sqrt(s*MatSig[i,])

  X.tilde[i,] <- X.tilde[i,] + sapply(Tau, function(o) rnorm(1,0,o))

  gain.r <- q*cov(X.tilde[i,],r.Matr[i-1,])/(q^2*var(X.tilde[i,]) + MatSig[i,])
  gain.Kalman <- q*var(X.tilde[i,])/(q^2*var(X.tilde[i,]) + MatSig[i,])

  yobs <- q*X.tilde[i,] - sapply(MatSig[i,], function(o) rnorm(1,0,sqrt(o)))
  r.Matr[i,] <- r.Matr[i-1,] + gain.r*(Yi[i] - yobs )
  X.hat[i,] <- X.tilde[i,] + gain.Kalman*(Yi[i] - yobs )

}
)

```

```

SigM <- MatSig[-1,]
r.Ma <- r.Matr[-1,]
X..hat <- X.hat[-1,]
X..tilde <- X.tilde[-1,]

X.hat.resume <- t(apply(X..hat, 1, FUN = function(z) c(mean(z),sd(z),min(z),
quantile(z,c(0.25,0.5,0.75)),max(z)))))
X.tilde.resume <- t(apply(X..tilde, 1, FUN = function(z) c(mean(z),sd(z),min(z),
quantile(z,c(0.25,0.5,0.75)),max(z)))))

r.resume[k,] <- c(min(r.Ma[n[k],]),quantile(r.Ma[n[k],],c(0.25,0.5,0.75)),
max(r.Ma[n[k],]))
SigM.resume[k,] <- c(min(SigM[n[k],]),quantile(SigM[n[k],],c(0.25,0.5,0.75)),
max(SigM[n[k],]))

Tempo[1,k] <- Tempo1[3]

Tempo2 <- system.time(
Poster <- Ar1Gibbs(Y, alpha.mu0 = 0, alpha.tau0 = 10,shapeY = 0.1,rateY = 0.1,
burn = 50, MC =150)
)

plot(Poster$Alpha,type = "l")
plot(Poster$V,type = "l")
Tempo[2,k] <- Tempo2[3]
r.resume[5+k,] <- c(min(Poster$Alpha), quantile(Poster$Alpha,c(0.25,0.5,0.75)),
max(Poster$Alpha) )
SigM.resume[5+k,] <- c(min(Poster$V) ,quantile(Poster$V,c(0.25,0.5,0.75)),
max(Poster$V))
}

box.r <- data.frame(x = rep(c(30,50,100,250,500), 2),r.resume)
box.sigma2 <- data.frame(x =rep(c(30,50,100,250,500),2),SigM.resume)

xtable(Tempo, digits = 3)

Simu_r <- data.frame(box.r ,Metod= rep(c("EnKF","MCMC"), each = 5))
Simu_sig2 <- data.frame(box.sigma2,Metod= rep(c("EnKF","MCMC"), each = 5))

my_theme <- theme(axis.title = element_text(size =15, face="bold"),
# legend.position="top",
panel.border = element_rect( size = 2,fill = NA),
panel.background = element_rect("white"),
panel.grid.minor = element_line(size=0.5, colour="darkgrey",linetype = "dotted"),
panel.grid.major = element_line(size=0.5, colour="darkgrey",linetype = "dotted"),
plot.title = element_text(hjust = 0.5))

```

```

G1 <- ggplot(Simu_r, aes(x = factor(x))) +
my_theme +
theme(legend.position="top") +
theme(legend.text = element_text( size = 12, face = "bold"))+
theme(text = element_text(size=13)) +
geom_boxplot(
aes(ymin = X1, lower = X2, middle = X3, upper = X4, ymax = X5,   linetype=Metod),
stat = "identity", alpha=0.4) +
scale_linetype_manual(" ",values=c("EnKF" = "solid","MCMC"="twodash"),
breaks= c("EnKF","MCMC") ) +
coord_cartesian(ylim = c(0.97, 1.03)) +
xlab("N") +
ylab(expression( alpha)) +
scale_y_continuous(breaks = c(0.97,0.99,1, 1.02))

G2 <- ggplot(Simu_sig2, aes(x = factor(x) )) +
my_theme +
theme(legend.position="top") +
theme(legend.text = element_text( size = 12, face = "bold"))+
theme(text = element_text(size=13)) +
geom_boxplot(
aes(ymin = X1, lower = X2, middle = X3, upper = X4, ymax = X5,   linetype=Metod),
  stat = "identity", alpha=0.4) +
scale_linetype_manual(" ",values=c("EnKF" = "solid","MCMC"="twodash"),
breaks= c("EnKF","MCMC") ) +
coord_cartesian(ylim = c(1, 8)) +
xlab("N") +
scale_y_continuous(breaks = c(2,4,6, 8))+
ylab(expression(sigma^2))

grid.arrange(G1,G2,ncol = 2)

```

Referências Bibliográficas

- Anderson, J. L. (2001). An ensemble adjustment kalman filter for data assimilation. *Monthly weather review*, **129**(12), 2884–2903. Citado na pág. 3, 20, 58
- Brockwell, P. J. e Davis, R. A. (2016). *Introduction to time series and forecasting*. springer. Citado na pág. 4, 6
- Butala, M. D., Yun, J., Chen, Y., Frazin, R. A. e Kamalabadi, F. Asymptotic convergence of the ensemble kalman filter. Em *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, páginas 825–828. IEEE. Citado na pág. 20
- Campagnoli, P., Petrone, S. e Petris, G. (2009). *Dynamic Linear Models with R*. Citado na pág. 6, 29, 30
- Carlin, B. P., Polson, N. G. e Stoffer, D. S. (1992). A monte carlo approach to nonnormal and nonlinear state-space modeling. *Journal of the American Statistical Association*, **87**(418), 493–500. Citado na pág. 9
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, **99**(C5), 10143–10162. Citado na pág. 2, 3, 15, 18, 20
- Evensen, G. (1997). Advanced data assimilation for strongly nonlinear dynamics. *Monthly weather review*, **125**(6), 1342–1354. Citado na pág. 3, 18
- Evensen, G. (2003). The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, **53**(4), 343–367. Citado na pág. 20
- Evensen, G. (2009). *Data assimilation: the ensemble Kalman filter*. Springer Science & Business Media. Citado na pág. 20, 63
- Evensen, G. (2009). The ensemble kalman filter for combined state and parameter estimation. *IEEE Control Systems*, **29**(3). Citado na pág. 3, 18
- Evensen, G. e Van Leeuwen, P. J. (1996). Assimilation of geosat altimeter data for the agulhas current using the ensemble kalman filter with a quasigeostrophic model. *Monthly Weather Review*, **124**(1), 85–96. Citado na pág. 3, 18
- Fox, W. W. (1970). An exponential surplus-yield model for optimizing exploited fish populations. *Transactions of the American Fisheries Society*, **99**(1), 80–88. Citado na pág. 10, 13
- Gelman, A. *et al.* (2006). Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper). *Bayesian analysis*, **1**(3), 515–534. Citado na pág. 31

- Gordon, N. J., Salmond, D. J. e Smith, A. F. Novel approach to nonlinear/non-gaussian bayesian state estimation. Em *IEE Proceedings F (Radar and Signal Processing)*, volume 140, páginas 107–113. IET. Citado na pág. 9
- Hall, B. (2008). Laplacesdemon: An r package for bayesian inference. Citado na pág. 31
- Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*. Cambridge university press. Citado na pág. 9
- Ho, Y. e Lee, R. (1964). A bayesian approach to problems in stochastic estimation and control. *IEEE Transactions on Automatic Control*, **9**(4), 333–339. Citado na pág. 16
- Houtekamer, P. L. e Mitchell, H. L. (1998). Data assimilation using an ensemble kalman filter technique. *Monthly Weather Review*, **126**(3), 796–811. Citado na pág. 3, 18
- Houtekamer, P. e Zhang, F. (2016). Review of the ensemble kalman filter for atmospheric data assimilation. *Monthly Weather Review*, **144**(12), 4489–4532. Citado na pág. 3
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, **82**(1), 35–45. Citado na pág. 2, 15, 16
- Katzfuss, M., Stroud, J. R. e Wikle, C. K. (2016). Understanding the ensemble kalman filter. *The American Statistician*, **70**(4), 350–357. Citado na pág. 2, 3, 14
- Kitagawa, G. (1996). Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of computational and graphical statistics*, **5**(1), 1–25. Citado na pág. 9
-). Multivariate analysis. 1979. *Probability and mathematical statistics*. Academic Press Inc. Citado na pág. 17
- Maybeck, P. S. Stochastic models, estimation, and control, volume 141 of mathematics in science and engineering., 1979. Citado na pág. 2, 15
- Mitchell, H. L. e Houtekamer, P. L. (2000). An adaptive ensemble kalman filter. *Monthly Weather Review*, **128**(2), 416–433. Citado na pág. 3, 18
- Montenegro, C. e Branco, M. (2016). Bayesian state-space approach to biomass dynamic models with skewed and heavy-tailed error distributions. *Fisheries Research*, **181**, 48–62. Citado na pág. 14, 46, 50, 55, 56
- Montenegro, C., Maunder, M. N. e Zilleruelo, M. (2009). Improving management advice through spatially explicit models and sharing information. *Fisheries Research*, **100**(3), 191–199. Citado na pág. 46
- Pella, J. J. e Tomlinson, P. K. (1969). A generalized stock production model. *Inter-American Tropical Tuna Commission Bulletin*, **13**(3), 416–497. Citado na pág. 10, 13
- Petris, G. (2010). An R package for dynamic linear models. *Journal of Statistical Software*, **36**(12), 1–16. URL <http://www.jstatsoft.org/v36/i12/>. Citado na pág. 63
- Pitcher, T. e Hart, P. J. (1982). *Fisheries ecology*. Number 597: 504.4 PIT. Citado na pág. 10

- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019. URL <https://www.R-project.org/>. Citado na pág. 25
- Schaefer, M. B. (1954). Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. *Inter-American Tropical Tuna Commission Bulletin*, **1**(2), 23–56. Citado na pág. 10, 11
- Stroud, J. R. e Bengtsson, T. (2007). Sequential state and variance estimation within the ensemble kalman filter. *Monthly Weather Review*, **135**(9), 3194–3208. Citado na pág. 3, 24
- Stroud, J. R., Katzfuss, M. e Wikle, C. K. (2016). A bayesian adaptive ensemble kalman filter for sequential state and parameter estimation. *arXiv preprint arXiv:1611.03835*. Citado na pág. 3, 20, 22, 58
- Tanizaki, H. (2001). Nonlinear and non-gaussian state space modeling using sampling techniques. *Annals of the Institute of Statistical Mathematics*, **53**(1), 63–81. Citado na pág. 9
- West, M. e Harrison, J. (2006). *Bayesian forecasting and dynamic models*. Springer Science & Business Media. Citado na pág. 6, 29, 30
- Wikle, C. K. e Berliner, L. M. (2007). A bayesian tutorial for data assimilation. *Physica D: Nonlinear Phenomena*, **230**(1-2), 1–16. Citado na pág. 3