

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA APLICADA

Estimação dos Estados de Biorreatores Anaeróbicos

por

Camilla da Silva Poletto

Trabalho submetido como requisito parcial
para a obtenção do grau de
Mestre em Matemática Aplicada

Prof. Dr. Diego Eckhard
Orientador

Porto Alegre, janeiro de 2018.

CIP - CATALOGAÇÃO NA PUBLICAÇÃO

Poleto, Camilla da Silva

Estimação dos Estados de Biorreatores Anaeróbicos / Camilla da Silva Poleto.—Porto Alegre: PPGMAP da UFRGS, 2018.

97 p.: il.

Dissertação (Mestrado) —Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Matemática Aplicada, Porto Alegre, 2018.

Orientador: Eckhard, Diego

Dissertação: Matemática Aplicada,
Mestrado, Dissertação

Estimação dos Estados de Biorreatores Anaeróbicos

por

Camilla da Silva Poletto

Trabalho submetido ao Programa de Pós-Graduação em Matemática Aplicada do Instituto de Matemática e Estatística da Universidade Federal do Rio Grande do Sul, como requisito parcial para a obtenção do grau de

Mestre em Matemática Aplicada

Linha de Pesquisa: Sistemas não-Lineares e Biomatemática

Orientador: Prof. Dr. Diego Eckhard

Banca examinadora:

Prof. Dr. Luiz Emilio Allem
PPGMAp-UFRGS

Prof. Dr. Alexandre Sanfelici Bazanella
PPGEE-UFRGS

Prof. Dr. Guilherme Araujo Pimentel
Faculdade de Engenharia-PUCRS

Dissertação apresentada e aprovada em
janeiro de 2018.

Prof. Dr. Carlos Hoppen
Coordenador

À minha mãe.

AGRADECIMENTO

Ao meu pai, Eduino Poletto, pelo apoio incondicional. Ao meu orientador Diego Eckhard, pelo conhecimento transmitido e pela paciência. Ao colega e amigo Walter Haselein, por toda ajuda dispensada. A todos que, de alguma forma, contribuíram com este trabalho.

Sumário

LISTA DE FIGURAS	x
LISTA DE TABELAS	xii
LISTA SIGLAS E SÍMBOLOS	xiii
RESUMO	xviii
ABSTRACT	xix
1 INTRODUÇÃO	1
2 BIORREATORES	3
2.1 Tipos de biorreatores	3
2.1.1 Aeróbicos	3
2.1.2 Anaeróbicos	4
2.1.3 Híbridos	4
2.2 Modos de operação	4
2.2.1 Batelada	5
2.2.2 Semibatelada	5
2.2.3 Contínuo	6
2.3 Processo de digestão anaeróbica	7
2.3.1 Hidrólise	8
2.3.2 Acidogênese	8
2.3.3 Acetogênese	8
2.3.4 Metanogênese	9
2.4 Modelo matemático	9
3 INTRODUÇÃO AOS PROCESSOS ALEATÓRIOS	16

3.1	Probabilidade	16
3.2	Variáveis aleatórias	18
3.2.1	As distribuições Uniforme e Normal de probabilidade	22
3.2.2	Variáveis aleatórias múltiplas	24
3.2.3	Variáveis aleatórias multivariadas	26
3.3	Processos estocásticos	29
3.4	Estimação Bayesiana	31
4	FILTRO DE KALMAN	36
4.1	Propriedades dos estados do sistema	37
4.2	Propagação dos valores esperados e das covariâncias	39
4.3	O problema de filtragem	42
4.3.1	Resumo e solução do problema de filtragem de Kalman	43
4.4	Dedução das equações do Filtro de Kalman	44
4.4.1	Exemplos	51
4.4.1.1	Condição inicial correta ($\bar{x}_0 = x_0$) e covariância nula ($P_0 = 0$) . . .	52
4.4.1.2	Condição inicial incorreta ($\bar{x}_0 = [2 \ 2]^T$) e covariância não-nula ($P_0 \neq 0$)	54
4.4.1.3	Condição inicial incorreta ($\bar{x}_0 = [2 \ 2]^T$) e covariância nula ($P_0 = 0$)	55
4.5	Filtro de Kalman Estendido	55
4.5.1	Resultados	57
4.5.1.1	Condição inicial correta ($\bar{x}_0 = x_0$) e covariância nula ($P_0 = 0$) . . .	60
4.5.1.2	Condição inicial incorreta ($\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]^T$) e covariância não-nula ($P_0 \neq 0$)	67
4.5.1.3	Condição inicial incorreta ($\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]^T$) e covariância nula ($P_0 = 0$)	70
5	FILTRO DE PARTÍCULAS	73

5.1	Importance Sampling (IS)	74
5.2	Amostragem de Importância Sequencial (SIS)	76
5.3	Empobrecimento da amostra e reamostragem	78
5.3.1	Roughening	79
5.3.2	Resultados	79
5.3.2.1	Condição inicial correta ($\bar{x}_0 = x_0$) e covariância nula ($P_0 = 0$) . . .	80
5.3.2.2	Condição inicial incorreta ($\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]^T$) e covariância não-nula ($P_0 \neq 0$)	82
5.3.2.3	Condição inicial incorreta ($\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]^T$) e covariância nula ($P_0 = 0$)	85
6	DISCUSSÃO DOS RESULTADOS	88
6.1	Condição inicial correta ($\bar{x}_0 = x_0$) e covariância nula ($P_0 = 0$) .	88
6.2	Condição inicial incorreta ($\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]^T$) e covariância não-nula ($P_0 \neq 0$)	89
6.3	Condição inicial incorreta ($\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]^T$) e covariância nula ($P_0 = 0$)	91
7	CONCLUSÕES	93
	REFERÊNCIAS BIBLIOGRÁFICAS	95

Lista de Figuras

Figura 3.1	Distribuição Uniforme	23
Figura 3.2	Distribuição Normal com $E[X] = 10$ e $\sigma^2 = 9$	24
Figura 4.1	Simulação com $\bar{x}_0 = x_0$	53
Figura 4.2	Simulação com $\bar{x}_0 \neq x_0$ e $P_0 \neq 0$	54
Figura 4.3	Simulação com $\bar{x}_0 \neq x_0$ e $P_0 = 0$	55
Figura 4.4	Saída de gás com ruído adicionado	59
Figura 4.5	Simulação do S_2 com $\bar{x}_0 = x_0$	61
Figura 4.6	Simulação com $\bar{x}_0 = x_0$	62
Figura 4.7	Simulação do S_2 com $\bar{x}_0 = x_0$ e $Q = I_5/100$	63
Figura 4.8	Simulação com $\bar{x}_0 = x_0$ e $Q = I_5/100$	64
Figura 4.9	Simulação do S_2 com $\bar{x}_0 = x_0$ e $R/10000$	65
Figura 4.10	Simulação com $\bar{x}_0 = x_0$ e $R/10000$	66
Figura 4.11	Simulação com $\bar{x}_0 \neq x_0$ e $P_0 = I_5/100$	68
Figura 4.12	Simulação com $\bar{x}_0 \neq x_0$ e $P_0 = I_5/100$	69
Figura 4.13	Simulação do S_2 com $\bar{x}_0 \neq x_0$ e $P_0 = I_5/100$	70
Figura 4.14	Simulação com $\bar{x}_0 \neq x_0$ e $P_0 = 0$	71
Figura 4.15	Simulação do S_2 com $\bar{x}_0 \neq x_0$ e $P_0 = 0$	72
Figura 5.1	Aproximação de uma distribuição não-gaussiana pelo método de Monte Carlo [17]	74
Figura 5.2	Simulação do S_2 com $\bar{x}_0 = x_0$, $P_0 = 0$ e $L = 10000$	80
Figura 5.3	Simulação com $\bar{x}_0 = x_0$, $P_0 = 0$ e $L = 10000$	81
Figura 5.4	Simulação do S_2 com $\bar{x}_0 \neq x_0$, $P_0 = I_5$ e $L = 10000$	82
Figura 5.5	Simulação com $\bar{x}_0 \neq x_0$, $P_0 = I_5$ e $L = 10000$	83
Figura 5.6	Exemplo de simulação em que o filtro diverge	84

Figura 5.7	Simulação do S_2 com $\bar{x}_0 \neq x_0$, $P_0 = 0$ e $L = 10000$	86
Figura 5.8	Simulação com $\bar{x}_0 \neq x_0$, $P_0 = 0$ e $L = 10000$	87

Lista de Tabelas

Tabela 2.1	Relação dos estados e parâmetros do modelo	13
Tabela 4.1	Erros médios com $\bar{x}_0 = x_0$ e matriz Q manipulada	53
Tabela 4.2	Erros médios com $\bar{x}_0 = x_0$ e valor de R manipulado	54
Tabela 4.3	Erros médios com $\bar{x}_0 = x_0$	61
Tabela 4.4	Erros médios com $\bar{x}_0 = x_0$ e matriz Q manipulada	63
Tabela 4.5	Erros médios com $\bar{x}_0 = x_0$ e valor de R modificado	65
Tabela 4.6	Erros médios com $\bar{x}_0 \neq x_0$ e $P_0 \neq 0$	67
Tabela 4.7	Erros médios com $\bar{x}_0 \neq x_0$ e $P_0 = 0$	70
Tabela 5.1	Resultados para $\bar{x}_0 = x_0$ e $P_0 = 0$	80
Tabela 5.2	Resultados para $\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]$, $P_0 = I_5$	82
Tabela 5.3	Resultados para $\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]$, $P_0 = I_5/100$	85
Tabela 5.4	Resultados para $\bar{x}_0 \neq x_0$, $P_0 = 0$	86
Tabela 6.1	Comparação dos resultados com $\bar{x}_0 = x_0$ e $P_0 = 0$	88
Tabela 6.2	Comparação dos resultados com $\bar{x}_0 = x_0$, $P_0 = 0$ e matrizes Q e R manipuladas	89
Tabela 6.3	Comparação dos resultados com $\bar{x}_0 \neq x_0$ e $P_0 = I_5$	90
Tabela 6.4	Comparação dos resultados com $\bar{x}_0 \neq x_0$ e $P_0 = I_5/100$	90
Tabela 6.5	Comparação dos resultados com $\bar{x}_0 \neq x_0$ e $P_0 = 0$	91

LISTA SIGLAS E SÍMBOLOS

Lista de Siglas

COD	Demanda química de oxigênio
VFA	Ácidos graxos voláteis
CO_2	Gás carbônico
CH_4	Gás metano
pH	Potencial hidrogeniônico (acidez)
PDF	Função distribuição de probabilidade
pdf	Função densidade de probabilidade
EKF	Filtro de Kalman Estendido
IS	Amostragem de importância
SIS	Amostragem de importância sequencial
SIR	Reamostragem de importância sequencial

Lista de Símbolos

x_1	Concentração de bactérias acidogênicas
x_2	Concentração de bactérias metanogênicas
S_1	Concentração de substrato orgânico
S_2	Concentração de ácidos graxos
C	Concentração de gás carbônico

Z	Alcalinidade
D	Taxa de diluição
k_1	Coefficiente de degradação do substrato orgânico
k_2	Coefficiente de produção de ácidos graxos
k_2	Coefficiente de produção de ácidos graxos
k_3	Coefficiente de degradação de ácidos graxos
k_4	Coefficiente de produção de CO_2 a partir de x_1
k_5	Coefficiente de produção de CO_2 a partir de x_2
k_6	Coefficiente de produção de CH_4
q_M	Taxa de fluxo de CH_4
q_c	Taxa de fluxo de CO_2
k_c	Coefficiente de consumo de CO_2
v_1	Velocidade de reação na acidogênese
v_2	Velocidade de reação na metanogênese
ν_1	Taxa de crescimento das bactérias acidogênicas
ν_2	Taxa de crescimento das bactérias metanogênicas
K_{S1}	Parâmetro de saturação associado a S_1
K_{S2}	Parâmetro de saturação associado a S_2
K_{I1}	Parâmetro de inibição associado a S_1
K_{I2}	Parâmetro de inibição associado a S_2

c_1	Taxa de mortalidade das bactérias acidogênicas
c_2	Taxa de mortalidade das bactérias metanogênicas
$P(E)$	Probabilidade de ocorrência do evento E
$P(A, B)$	Probabilidade de ocorrência do evento A e do evento B
$P(A B)$	Probabilidade condicional de A e dado que B ocorreu
$F_X(x)$	Função distribuição de probabilidade
$p_X(x)$	Função densidade de probabilidade
$p_{X Y}$	Densidade de probabilidade condicional de X dado Y
σ^2	Variância
C_{XY}	Covariância entre as variáveis aleatórias X e Y
R_{XY}	Correlação entre as variáveis aleatórias X e Y
R_X	Autocorrelação entre os elementos do vetor X
R_k	Covariância do ruído v_k
Q_k	Covariância do ruído w_k
δ_{kl}	Matriz delta de Kronecker
\bar{x}_k	Valor esperado de x_k
P_k	Covariância de x_k
$P_{k,l}$	Covariância entre x_k e x_l
$\Phi_{k,l}$	Matriz de transição dos estados
$\hat{x}_{k k-1}$	Estimativa a priori de x_k

$\hat{x}_{k k}$	Estimativa a posteriori de x_k
$P_{k k-1}$	Covariância do erro da estimativa $x_{k k-1}$
$P_{k k}$	Covariância do erro da estimativa $x_{k k}$
K_k	Matriz ganho de Kalman

RESUMO

Biorreatores anaeróbicos são equipamentos que degradam matéria orgânica, produzindo gás metano e fertilizante, cujo processo é modelado por sistemas não-lineares. Este trabalho tem como objetivo estimar, através do estudo de modelos dinâmicos e da observação de dados coletados, a quantidade de substratos, bactérias e gás carbônico presentes em um biorreator, nos diversos pontos de sua operação. Para isso, foram utilizados e comparados dois filtros: o de Kalman Estendido e o de Partículas. Esses filtros visam amenizar o efeito do ruído nos dados coletados do experimento e geram estimativas dos estados do sistema, as quais podem ser utilizadas para operar o biorreator em sua máxima capacidade de produção.

ABSTRACT

Anaerobic bioreactors are devices that degrade organic matter, producing methane and fertilizer, whose process is modelled by non-linear systems. This work goals to estimate, through the study of dynamic models and the observation of collected data, the current amount of substrates, bacteria and carbon dioxide in a bioreactor, in the various points of its operation. For this, two filters were used and compared: Extended Kalman and Particles. These filters intent to attenuate the effect of noise on the collected data from the experiment and generate estimates of the system states, which can be used to operate the bioreactor at its maximum production capacity.

1 INTRODUÇÃO

O crescimento mundial vem ocasionando uma série de danos ao meio ambiente. A produção industrial e agrícola tem preocupado órgãos ambientais por conta do grande volume de resíduos gerado, fazendo com que leis mais severas sejam aplicadas para controlar o que é descartado na natureza. Além disso, o aumento na demanda por energia elétrica tem incentivado a busca por novas alternativas de fonte de energia, que tenham menor impacto ambiental e que não sejam passíveis de esgotamento, como o carvão mineral e o petróleo. Aqui entra em questão o conceito de desenvolvimento sustentável, que, no caso da produção de energia, combina o aumento da oferta com a preservação do meio ambiente [3].

Biorreatores são equipamentos que aliam a degradação de matéria orgânica (como resíduos da indústria de alimentos ou da agricultura) e a geração de um produto. No caso dos biorreatores anaeróbicos, nos quais a degradação ocorre na ausência de oxigênio, esse produto consiste, basicamente, em gás metano, gás carbônico e fertilizante. Três famílias de bactérias atuam em um biorreator anaeróbico: acidogênicas, acetogênicas e metanogênicas. As bactérias acidogênicas degradam a matéria orgânica, produzindo ácidos graxos voláteis (VFA); as bactérias acetogênicas transformam VFA em acetato; por fim, as bactérias metanogênicas transformam acetato em gás metano. A taxa de produção de um biorreator está diretamente ligada à concentração de VFA. O processo não é linear e altas concentrações de VFA acarretam na diminuição e até mesmo na interrupção da produção. Por isso, em geral, os biorreatores acabam sendo operados bem abaixo da sua capacidade máxima. Nesse sentido, é pertinente entender e buscar modelos matemáticos que descrevam a dinâmica desses sistemas e, assim, otimizar o seu funcionamento.

Na prática, o que é viável de ser medido em um biorreator é apenas a saída de gás. Para precisar a concentração de substrato e bactérias presentes no

equipamento, são necessários processos de análise custosos e demorados, como a cromatografia gasosa. Estimar os estados do sistema, através de um modelo matemático do biorreator e das observações das medidas de gás, acaba sendo uma alternativa para esse problema.

Os problemas de filtragem ou de estimação consistem exatamente nisso: estimar os estados de um sistema a partir de um sinal medido, normalmente perturbado por ruído. Quando o sistema é descrito por um modelo linear com ruídos gaussianos, o filtro de Kalman dá a solução ótima para o problema de estimação. No entanto, assim como a maioria dos sistemas reais, o modelo de um biorreator é não-linear e os ruídos associados à saída de gás não são, necessariamente, gaussianos.

Neste trabalho, serão apresentados e implementados dois tipos de filtragem: o Filtro de Kalman Estendido, em que o sistema é linearizado através de uma aproximação local e os ruídos são considerados gaussianos; e o Filtro de Partículas, que admite sistemas não-lineares e quaisquer tipos de ruídos. O Filtro de Kalman Estendido é um dos mais utilizados na literatura, devido a sua simplicidade de implementação. Já o Filtro de Partículas tende à solução ótima quando o número de partículas é muito grande, mas sua implementação é mais complexa, além de demandar um alto custo computacional. Os filtros serão aplicados a um modelo que descreve um biorreator anaeróbico do tipo semibatelada, cuja validação foi feita em um trabalho anterior [11].

2 BIORREATORES

Biorreatores são equipamentos nos quais ocorrem uma série de reações químicas, realizadas por bactérias, que convertem matéria orgânica em algum produto [11]. Essa matéria orgânica, também chamada de biomassa, pode ser de origem florestal (como madeira), agrícola (arroz, soja, milho, cana-de-açúcar, etc.) ou de rejeitos urbanos e industriais (resíduos da indústria alimentícia ou o próprio lixo, por exemplo) [3, 8]. Os primeiros biorreatores derivaram dos equipamentos denominados fermentadores, os quais foram, há muitas décadas, desenvolvidos para cultivo de fungos e bactérias para fins industriais [21].

2.1 Tipos de biorreatores

Os biorreatores podem ser classificados de acordo com o seu processo de digestão: aeróbicos, anaeróbicos ou híbridos. Geralmente, ao se projetar um equipamento, além da escolha de um dos três tipos de digestão, que está ligada à finalidade do biorreator, são levados em conta outros fatores: espaço físico disponível, viabilidade financeira, regulamentação ambiental, entre outros [11].

2.1.1 Aeróbicos

A digestão aeróbica se beneficia da degradação de resíduos por bactérias que necessitam de oxigênio para viver. Esse processo é, essencialmente, uma operação de compostagem em larga escala. Para acelerar a estabilização do sistema, ar e líquidos são adicionados ao substrato a fim de promover condições favoráveis de temperatura e umidade para a decomposição aeróbica [12]. Esse tipo de digestão é ideal para aterros que não geram gás metano em quantidades suficientes para a conversão em energia [11].

2.1.2 Anaeróbicos

Em um biorreator anaeróbico, a decomposição da matéria orgânica ocorre através de bactérias que não utilizam oxigênio na sua digestão. Uma das principais vantagens desse processo é a produção de gás metano como produto da degradação dos insumos. Esse biogás pode ser utilizado como combustível veicular, gás de cozinha ou na geração de energia elétrica [20]. Esse tipo de processo também é vantajoso por apresentar maior capacidade de degradar substratos concentrados e complexos, e gerar pouco resíduo (que ainda pode ser utilizado como fertilizante) [8, 16].

2.1.3 Híbridos

A digestão híbrida combina os dois processos de anteriores: a parte superior do substrato é adicionada de oxigênio para acelerar a decomposição e a parte inferior produz metano através da digestão anaeróbica [11]. A finalidade dessa implementação híbrida é a rápida degradação de resíduos orgânicos no estágio aeróbico, o que reduz a produção de ácidos orgânicos e resulta na geração acelerada de metano no estágio anaeróbico [12].

2.2 Modos de operação

Os três principais modos de operação de um biorreator são: batelada (ou de fermentação descontínua), semibatelada (ou de fermentação semicontínua) e contínuo (ou fermentação contínua). A escolha entre um desses processos fermentativos deve visar a produtividade do biorreator, o que depende de muitos fatores, tais como: microorganismo utilizado, método de preparo da biomassa, temperatura, pH, entre outros [6].

2.2.1 Batelada

Nesse modo de operação, todo o substrato é introduzido no equipamento no início. Nada é adicionado ou retirado ao longo do processo, o que caracteriza uma reação a volume constante. As únicas ações que podem ocorrer estão ligadas ao controle de variáveis ambientais (pH, velocidade da agitação, temperatura, etc.). A fermentação será cessada quando o substrato inicial tiver sido suficientemente consumido [7].

Como constitui um processo simples e exige poucos recursos na sua implementação, a fermentação descontínua tem se mostrado bastante atrativa, principalmente no meio industrial, como na produção de iogurte, chucrute, cerveja e vinho [6, 7]. Outras vantagens desse processo são os menores riscos de contaminação (quando comparado à fermentação contínua), flexibilidade de operação (por viabilizar a fermentação de diferentes substratos) e a possibilidade de utilizar o mesmo recipiente para realizar fases sucessivas [6].

No entanto, há um limite da quantidade aceitável de substrato a ser utilizada no processo. A adição de uma grande quantidade de substrato pode causar a inibição do crescimento das bactérias que o consomem e levar a baixos rendimentos [7]. Outra desvantagem desse processo são os chamados *tempos mortos*, como o tempo de carga e descarga ou lavagem e esterização do equipamento [6].

2.2.2 Semibatelada

Esse modo de operação difere do anterior pela alimentação do biorreator no decorrer do processo e foi desenvolvido com o intuito de reduzir ou eliminar os problemas de inibição que ocorrem no modo batelada, aumentando a sua produtividade [11].

Dentre as vantagens da fermentação semicontínua, pode-se destacar: a possibilidade de operar o biorreator por longos períodos sem a necessidade de adicionar mais bactérias; a mobilidade no cronograma de operação (retirada e adição de substrato), viabilizando um aumento na produtividade; e, uma vez conhecidas as condições ótimas de operação, a possibilidade de uma produção significamente maior, comparada a obtida em fermentação descontínua [6].

No entanto, a necessidade da alimentação por um fluxo controlado acaba limitando o uso desse tipo de fermentação, principalmente em escala industrial [7]. Há desvantagem também em relação ao risco de contaminação, que acaba aumentando em função dos acréscimos de substrato. Além disso, a frequente alteração do meio de cultivo dificulta a busca de um ponto ótimo de operação [11].

2.2.3 Contínuo

Esse processo caracteriza-se pela adição contínua de substrato, a uma determinada vazão constante, sendo mantido o volume da reação através da retirada contínua de produto fermentado. É de extrema importância que o volume de líquido no biorreator seja constante, com o intuito de que o sistema atinja a condição de estado estacionário, no qual as variáveis de estado permanecem constantes ao longo do processo [6]. No entanto, a manutenção do volume constante no biorreator é praticamente impossível de se obter na prática, pois necessita de vazões idênticas de inserção e retirada de substrato e produto, respectivamente. Por isso, geralmente são utilizados sistemas de retirada de líquido por transbordamento, a fim de manter o nível de líquido constante [7].

As principais vantagens da fermentação contínua - em relação à fermentação descontínua, principalmente - decorrem da possibilidade de operação em estado estacionário, dentre as quais pode-se citar: aumento da produtividade do processo, em decorrência da diminuição de tempos mortos ou não-produtivos; manu-

tenção das células em um mesmo estado fisiológico - o que torna o processo contínuo uma excelente ferramenta para estudos de mecanismos de regulação metabólica ou otimização da composição do substrato; maior facilidade no emprego de controles avançados; e menor necessidade de mão de obra [16].

Todavia, o processo de fermentação contínua exige um maior investimento inicial na planta, é mais suscetível a contaminações, pode apresentar dificuldade na manutenção da homogeneidade do substrato e até mesmo de operação em estado estacionário em algumas situações (formação de espuma, crescimento de microorganismos nas paredes do equipamento ou nos sistemas de entrada e saída de substrato) [6].

Apesar dos problemas citados, a utilização da fermentação contínua possui grande aplicação prática, como o tratamento de efluentes industriais de fábricas de cervejas e refrigerantes, laticínios e de indústrias alimentícias em geral.

O processo de fermentação contínua normalmente começa como um processo descontínuo, ou seja: carrega-se o biorreator com substrato e, após algum período de operação descontínua, inicia-se a alimentação e retirada de caldo, iniciando assim, efetivamente, o processo contínuo. Dependendo do instante em que se inicie, de fato, o processo contínuo, o sistema poderá atingir o estado estacionário com maior ou menor rapidez [7].

2.3 Processo de digestão anaeróbica

A digestão anaeróbica ocorre, basicamente, em quatro fases: hidrólise, acidogênese, acetogênese e metanogênese. Essas fases são sequenciais e dependentes entre si, pois micro-organismos atuantes em cada uma delas geram componentes que servem de alimento para os micro-organismos da fase sucessora. Em consequência disso, se houver qualquer tipo de inibição nas taxas de crescimento de uma população

de micro-organismos atuantes em qualquer uma das fases, o sistema todo acaba sendo comprometido [11, 14].

2.3.1 Hidrólise

A hidrólise é a primeira fase da digestão anaeróbica e pode ser dividida em duas etapas: na primeira, componentes do substrato (lipídios, carboidratos e proteínas) são decompostos em substâncias menos complexas, através da reação com as moléculas de água; na segunda, essas substâncias mais simples são convertidas em monossacarídeos, aminoácidos e ácidos graxos pelas enzimas que as bactérias excretam [11].

2.3.2 Acidogênese

Nessa fase, os componentes gerados na hidrólise são convertidos em ácidos graxos voláteis (como ácido acético, propanoico e butanoico), álcoois, gás carbônico, hidrogênio, amônia, entre outros. Além disso, há a produção de novas células bacterianas [13]. A quantidade de ácidos produzidos nessa etapa pode acarretar na inibição da atividade das bactérias acetogênicas, devido à redução do pH do meio, e na consequente interrupção do processo de digestão [11].

2.3.3 Acetogênese

Durante essa fase, os compostos da fase anterior são oxidados pelas bactérias acetogênicas em substratos apropriados para as bactérias metanogênicas. Os produtos gerados nessa etapa são acetatos, hidrogênio, gás carbônico e água. De todos os produtos metabolizados pelas bactérias acidogênicas, apenas o hidrogênio e o acetato podem ser utilizados diretamente pelas metanogênicas [13]. A relação entre a produção de hidrogênio nas fases de acidogênese e acetogênese e o seu consumo pela fase metanogênica é de suma importância para a estabilidade do sistema [11].

2.3.4 Metanogênese

A fase final do processo de digestão anaeróbica é efetuada pelas bactérias metanogênicas, as quais consomem os compostos gerados na fase da acetogênese e produzem metano, gás carbônico e água [11]. As bactérias metanogênicas podem ser divididas em dois grupos: um que forma metano a partir de ácido acético ou metanol, e outro que produz metano a partir de hidrogênio e gás carbônico. O biogás resultante consiste em, aproximadamente, de 50 a 75% de metano; e de 25 a 45% de gás carbônico - além de vapor d'água e outros gases em volumes menores [13].

2.4 Modelo matemático

A modelagem matemática de um biorreator tem como principal objetivo prever o comportamento dinâmico e estacionário do processo fermentativo - até mesmo em condições que ainda não foram testadas. Essa predição é possível através da simulação computacional, e permite determinar as condições operacionais ótimas do sistema, contribuindo no projeto e no ajuste dos algoritmos de controle [6]. Por isso, a obtenção de modelos que representem a dinâmica de um biorreator é tão importante para a elaboração, operação e otimização da produção [8].

O modelo de um sistema é um conjunto de relações, normalmente expressas sob a forma de equações, entre as variáveis que o compõem. Em alguns processos, essas relações podem ser especificadas como relação de causa e efeito. As variáveis envolvidas podem abranger qualquer parâmetro de importância para o processo, como pH, temperatura, substrato, concentração, taxa de alimentação, entre outras.

A estrutura do modelo, bem como a sua complexidade, deve corresponder ao objetivo para o qual foi elaborado - como ajudar na compreensão dos mecanismos envolvidos, estimar variáveis que não são medidas, identificar os parâmetros

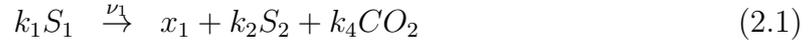
do processo, entre outros. Além disso, o modelo também deve ser adaptado aos dados disponíveis. Um modelo muito complexo, que necessite de uma grande quantidade de parâmetros, exigirá uma grande quantidade de dados a serem identificados e validados [7].

Uma série de características (como as baixas concentrações e velocidades de reação quando o meio é diluído, o conhecimento insuficiente de vários dos fenômenos limitantes das velocidades de produção e problemas de esterilidade) diferenciam os processos fermentativos dos processos químicos tradicionais, o que pode justificar as dificuldades encontradas para formular modelos matemáticos que os representem adequadamente [6]. No entanto, verifica-se, na literatura, grande aceitação por modelos mais simples, os quais exigem menor custo computacional, facilitando a simulação e análises numéricas, e conseguem descrever com boa precisão a produção de gás no biorreator [8].

Outro fator determinante na complexidade do modelo a ser desenvolvido é o número de famílias de bactérias envolvidas no processo de digestão. Como o processo de acetogênese acontece muito rápido, as bactérias acetogênicas são omitidas no processo e assume-se que a população de bactérias é composta por dois grupos com características homogêneas [11]. A digestão pode ser descrita em dois estágios:

1. Acidogênese: as bactérias acidogênicas (x_1) consomem o substrato orgânico (S_1) e produzem gás carbônico (CO_2) e ácidos graxos (S_2);
2. Metanogênese: as bactérias metanogênicas (x_2) consomem os ácidos graxos produzidos na etapa anterior e produzem CO_2 e metano (CH_4).

Esses dois estgios podem ser representados, respectivamente, pelas seguintes reaoes [5]:



onde ν_1 e ν_2 descrevem as velocidades de reaao nos estgios de acidognese e metanognese, respectivamente [11].

A sada de gs  composta, principalmente, por gs carbnico (CO_2) e metano (CH_4). Por conta da baixa solubilidade do gs metano, assume-se que no h concentraao de metano dissolvido e que sua taxa de sada q_M  proporcional  taxa de reaao da metanognese, sendo dada por [5]:

$$q_M = k_6 \nu_2 x_2 \quad (2.3)$$

J para obter a taxa de fluxo do gs carbnico,  necessrio utilizar a lei de Henry, que expressa o fluxo molar q_G de um composto gasoso G , da fase lquida para a fase gasosa [7]:

$$q_G = k_L \alpha (G - G^*),$$

onde $k_L \alpha$ (1/h)  o coeficiente de evaporaao, o qual depende das condioes de operaao - agitaao, presso e transferncia entre as fases lquida e gasosa; G^*  a concentraao de saturaao de G dissolvido e pode ser escrita como:

$$G^* = K_H P_G,$$

sendo K_H a constante de Henry, que depende da temperatura e do meio de cultura, e P_G a presso parcial do gs.

Desta forma, utilizando a lei de Henry e alguns resultados sobre a alcalinidade e a presso total do sistema [5], a sada do gs carbnico q_C  dada por [7]:

$$q_C = k_L \alpha (C + S_2 - Z - K_H P_C) \quad (2.4)$$

onde $k_L\alpha$ é o coeficiente de evaporação, Z é a alcalinidade do sistema, K_H é a constante de Henry e P_C é a pressão parcial de gás carbônico.

Como não existem leis que descrevem o desenvolvimento de microorganismos com precisão, a tarefa de modelar sistemas biológicos se torna um tanto complexa. Porém, esses sistemas obedecem a determinadas leis físicas, como a conservação de massa, eletroneutralidade (o estado de ter exatamente o mesmo número de positivos e negativos cargas elétricas), etc [7]. Por conta disso, os modelos que descrevem um biorreator são, geralmente, elaborados através do balanço de massa e energia do sistema, nos quais parâmetros definem as taxas de crescimentos das bactérias, a velocidade das reações, entre outros fatores [8].

No modelo balanço de massa, utilizando as equações (2.1) e (2.2), os estados são descritos por (os estados e parâmetros do sistema estão especificados na tabela 2.1 [7]:

$$\begin{cases} \dot{x}_1(t) = [\nu_1(S_1(t)) - \alpha D]x_1(t) \\ \dot{x}_2(t) = [\nu_2(S_2(t)) - \alpha D]x_2(t) \\ \dot{S}_1(t) = D(S_1^{in}(t) - S_1(t)) - k_1\nu_1(S_1(t))x_1(t) \\ \dot{S}_2(t) = D(S_2^{in}(t) - S_2(t)) + k_2\nu_1(S_1(t))x_1(t) - k_3\nu_2(S_2(t))x_2(t) \\ \dot{Z}(t) = D(Z^{in}(t) - Z(t)) \\ \dot{C}(t) = D(C^{in}(t) - C(t)) - q_C(t) + k_4\nu_1(S_1(t))x_1(t) + k_5\nu_2(S_2(t))x_2(t) \end{cases} \quad (2.5)$$

A saída do modelo é dada pela saída total de gás, que será a soma das equações de saída de gás carbônico (2.4) e de gás metano (2.3):

$$q_T = k_6\nu_2x_2 + k_L\alpha(C + S_2 - Z - K_HP_C)$$

Cada equação em (2.5) modela a taxa de variação de um dos estados do sistema. A primeira delas indica que as bactérias acidogênicas (x_1) têm seu crescimento atrelado ao consumo do substrato orgânico (S_1) e a diminuição da sua

$x_1(t)$	concentração de bactérias acidogênicas (mg/L)
$x_2(t)$	concentração de bactérias metanogênicas (mg/L)
$S_1(t)$	concentração do substrato orgânico (COD) (mg/L)
$S_2(t)$	concentração de ácidos graxos voláteis (VFA) ($mmol/L$)
$Z(t), Z^{in}$	alcalinidade total do sistema ($mmol/L$)
$C(t)$	concentração de carbono inorgânico ($mmol/L$)
D	taxa de diluição (dia^{-1})
α	parâmetro proporcional determinado experimentalmente ($0 < \alpha \leq 1$)
S_1^{in}, S_2^{in}	concentração dos influentes S_1 e S_2 , respectivamente
k_1	coeficiente de degradação do substrato orgânico ($mgCOD/mgx_1$)
k_2	coeficiente de produção de ácidos graxos ($mgVFA/mgx_1$)
k_3	coeficiente de consumo de ácidos graxos ($mgVFA/mgx_2$)
k_4	coeficiente de produção de CO_2 a partir de S_1 ($mmolCO_2/mgS_1$)
k_5	coeficiente de produção de CO_2 a partir de S_2 ($mmolCO_2/mgS_2$)
k_6	coeficiente de produção de CH_4 ($mmol/mg$)
q_C	taxa de fluxo de CO_2 ($mmol/Ldia^{-1}$)
ν_1, ν_2	taxa de crescimento das bactérias acidogênicas e metanogênicas (dia^{-1})

Tabela 2.1: Relação dos estados e parâmetros do modelo

concentração é ocasionada pela taxa de diluição D . De forma análoga, a segunda equação relaciona o aumento das bactérias metanogênicas (x_2) com o consumo de ácidos graxos (S_2), e a sua diminuição à mesma taxa D . A terceira equação indica que o substrato S_1 apenas diminui, à medida que é consumido pelas bactérias acidogênicas, e é também afetado pela taxa de diluição D . No entanto, se for adotado o modo de operação semi-batelada ou contínuo, haverá a inserção de mais substrato durante o processo, o que é indicado por S_1^{in} . Da mesma forma acontece com o VFA S_2 , como expresso na quarta equação, com a diferença do termo positivo, que indica o aumento de VFA pela sua produção na fase da acidogênese. A alcalinidade (Z) total do sistema diminui com a diluição, e, nos casos em que o biorreator é alimentado, aumenta com a adição de substratos. A última equação relaciona o aumento da concentração de gás carbônico (C) com a degradação dos substratos S_1 e S_2 e com a inserção de mais substrato, se for o caso; a diminuição de C fica por conta do próprio fluxo do gás (q_C) e também da taxa de diluição D [11]. Esse modelo pode ser usado em qualquer um dos modos de operação de biorreatores. No caso do modo

batelada, no qual todo o substrato e a biomassa são inseridos apenas no início do processo, $D = S_1^{in} = S_2^{in} = Z^{in} = C^{in} = 0$ [8].

O comportamento não-linear do processo se dá pelas taxas de crescimento específicas para as bactérias x_1 e x_2 , dadas por $\nu_1(S_1(t))$ e $\nu_2(S_2(t))$, respectivamente. Essas taxas podem ser expressas pela lei de Monod, a qual relaciona o crescimento bacteriano apenas à concentração de substratos no biorreator [20]:

$$\nu_1(S_1(t)) = \mu_{m1} \frac{S_1(t)}{K_{S1} + S_1(t)} \quad (2.6)$$

$$\nu_2(S_2(t)) = \mu_{m2} \frac{S_2(t)}{K_{S2} + S_2(t)} \quad (2.7)$$

onde μ_{m1} (dia^{-1}) é a taxa de crescimento máxima da biomassa acidogênica, μ_{m2} (dia^{-1}) é a taxa de crescimento máxima da biomassa metanogênica, K_{S1} (mg/L) K_{S2} (mmol/L) são os parâmetros de saturação associados a $S_1(t)$ e a $S_2(t)$, respectivamente.

De acordo com o modelo de Monod, a reação atingirá sua velocidade máxima quando não houver saturação ($K_{S1} = K_{S2} = 0$). Na prática, isso geralmente ocorre em biorreatores com baixa concentração de substrato. Por outro lado, a ausência de substratos acarreta na interrupção do processo. A inconsistência da lei de Monod com dados experimentais em alguns trabalhos tem gerado a busca de outros modelos para representação das velocidades da reações. Dentre estes, o modelo a seguir, que considera o fenômeno de inibição devido ao excesso de substrato, proposto por Haldane [7, 11]:

$$\nu_1(S_1(t)) = \mu_{m1} \frac{S_1(t)}{K_{S1} + S_1(t) + K_{I1}S_1(t)^2} \quad (2.8)$$

$$\nu_2(S_2(t)) = \mu_{m2} \frac{S_2(t)}{K_{S2} + S_2(t) + K_{I2}S_2(t)^2} \quad (2.9)$$

onde são acrescentados no modelo de Monod os parâmetros de inibição K_{I1} (mg/L) e K_{I2} (mmol/L), relacionados a S_1 e a S_2 , respectivamente. Desta forma, a reação só atingirá a velocidade máxima quando o sistema não sofrer efeitos de inibição e nem de saturação.

Nos capítulos de métodos numéricos 4 e 5, serão apresentados algoritmos que visam estimar os estados do modelo (2.5), através de medições ruidosas de saída de gás.

3 INTRODUÇÃO AOS PROCESSOS ALEATÓRIOS

Este capítulo apresenta uma breve introdução à probabilidade e processos aleatórios - conteúdos que serão utilizados no decorrer deste trabalho.

3.1 Probabilidade

O conjunto de todos os resultados possíveis de um experimento aleatório é chamado de *espaço amostral* (Ω). Um *evento* E é uma possível saída desse experimento [2]. Considere, por exemplo, o lançamento de um dado. O espaço amostral, nesse caso, são todas as possibilidades de resultado, ou seja, $\Omega = \{1, 2, 3, 4, 5, 6\}$. Um evento poderia ser $E = \{5\}$: ocorrer o número 5 ou $E = \{2, 4, 6\}$: ocorrer número par.

A probabilidade de um evento E ocorrer é denotada por $P(E)$ e é calculada através da fórmula $P(E) = \frac{n(E)}{n(\Omega)}$, onde $n(E)$ é o número de elementos em E e $n(\Omega)$ é o número de elementos de Ω . Num limite, quando o número de tentativas é grande, pode-se pensar $P(E)$ como a relação entre o número de vezes que E ocorreu e o número de vezes que o experimento foi realizado. Se todas as possíveis saídas de um experimento forem denotadas por E_i , com $i = 1, 2, \dots, n$, então [9]:

$$0 \leq P(E_i) \leq 1$$

e

$$\sum_{i=1}^n P(E_i) = 1.$$

Denota-se por $P(A, B)$ a probabilidade de que ambos os eventos, A e B , ocorram. Se os eventos A e B forem *mutuamente independentes* entre si (a

ocorrência de um não depende da ocorrência do outro), a probabilidade $P(A, B)$ é dada pelo produto das probabilidades, conhecida como regra da multiplicação. Nesse caso,

$$P(A, B) = P(A)P(B)$$

A probabilidade de ocorrência do evento A ou do evento B é denotada por $P(A+B)$. No caso em que esses eventos são *mutuamente exclusivos* (a ocorrência de um deles automaticamente impede a ocorrência dos demais), essa probabilidade é dada pela soma de $P(A)$ e $P(B)$, ou seja [9]:

$$P(A + B) = P(A) + P(B)$$

Se dois eventos A e B não são mutuamente exclusivos, deve-se subtrair a probabilidade de ocorrência de ambos, ou seja [9]:

$$P(A + B) = P(A) + P(B) - P(A, B)$$

Considere agora o exemplo de uma sala de aula com 20 alunos, sendo 12 meninos e 8 meninas. Entre os meninos, 5 têm olhos claros e 7 têm olhos escuros. Quanto às meninas, 2 têm olhos claros e 6 têm olhos escuros. Qual a probabilidade de que um aluno de olhos claros, escolhido ao acaso, seja menino? Note que, aqui, a probabilidade será interferida pelo fato de ser conhecido que o aluno tem olhos claros.

Para eventos que não são independentes, é necessário o conceito de *probabilidade condicional*. A probabilidade de um evento A ocorrer, dado que o evento B ocorreu, é denotada por $P(A|B)$ e definida como [9]:

$$P(A|B) = \frac{P(A, B)}{P(B)}, \quad (3.1)$$

Se A e B forem intercambiáveis, segue da equação (3.1) que

$$P(B|A) = \frac{P(B, A)}{P(A)} \Rightarrow P(B, A) = P(B|A)P(A)$$

Como $P(A, B) = P(B, A)$ - probabilidade de ocorrência de ambos os eventos A e B , pode-se reescrever a equação (3.1) como [9]:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \quad (3.2)$$

resultado conhecido como *Teorema de Bayes*.

No exemplo anterior, deseja-se saber a probabilidade de um aluno da sala ser menino (A), sendo que o aluno em questão tem olhos claros (B). Nesse caso, tem-se:

- probabilidade de ser menino: $P(A) = \frac{12}{20}$;
- probabilidade de ter olhos claros: $P(B) = \frac{7}{20}$;
- probabilidade de ter olhos claros, sendo que é menino: $P(B|A) = \frac{5}{12}$

Logo,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{\frac{5}{12} \frac{12}{20}}{\frac{7}{20}} = \frac{5}{7}$$

A probabilidade de um único evento ocorrer ($P(A)$, por exemplo) é chamada de uma probabilidade *a priori* porque se aplica à probabilidade de um evento além de qualquer informação conhecida anteriormente. A probabilidade condicional é chamada de uma probabilidade *a posteriori*, pois se aplica a uma probabilidade dado o fato de que algumas informações sobre um evento possivelmente relacionado são conhecidas [19]. Esses conceitos serão importantes nos capítulos de métodos numéricos 4 e 5.

3.2 Variáveis aleatórias

Uma *variável aleatória* (v.a.) X pode ser definida como um mapa funcional de um conjunto de resultados experimentais - o próprio espaço amostral Ω -

para um conjunto de números reais. O lançamento de um dado, por exemplo, pode ser interpretado como uma v.a. se relacionarmos o aparecimento da face com um ponto ao número 1, da face com dois pontos ao número 2, e assim por diante [19]. Desta forma, $X = \{1, 2, 3, 4, 5, 6\}$. Após o lançamento do dado, no entanto, o valor obtido é certo - e não mais uma variável aleatória. Se o valor obtido for 4, por exemplo, esse resultado é chamado de *realização* de X . A variável aleatória existe independente de suas realizações [22].

Quando o conjunto de valores possíveis de uma v.a. X for finito ou enumerável, ela é denominada uma variável aleatória discreta. Caso contrário, ou seja, quando o conjunto for infinito ou não-enumerável, ela é denominada variável aleatória contínua. O lançamento de um dado é uma v.a. discreta, pois suas realizações pertencem a um conjunto discreto de valores $\{1, 2, 3, 4, 5, 6\}$. Já a temperatura máxima de uma cidade é uma variável aleatória contínua, porque suas realizações pertencem a um conjunto contínuo de valores. Ou seja, a temperatura varia continuamente - e não aumenta subitamente de 30° para 31° , por exemplo [15, 19].

O comportamento de uma variável aleatória é determinado pela sua *função distribuição de probabilidade* (PDF), que é definida por [19]:

$$F_X(x) = P(X \leq x),$$

onde x é uma variável não-aleatória ou uma constante e $F_X(x) \in [0, 1]$.

Nos casos em que a variável aleatória X é contínua, utiliza-se a *função densidade de probabilidade* (pdf) - a qual será de extrema importância para o estudo dos problemas de filtragem desse trabalho, definida pela derivada da função distribuição de probabilidade, ou seja [19]:

$$p_X(x) = \frac{dF_X(x)}{dx},$$

sendo $p_X(x) \geq 0$ e $\int_{-\infty}^{+\infty} p_X(x)dx = 1$.

O conceito de probabilidade condicional pode ser estendido para variáveis aleatórias. A distribuição e a densidade de probabilidade condicional de uma v.a. X , dado que o evento A ocorreu são dadas por [19]:

$$F_X(x|A) = P(X \leq x|A) = \frac{P(X \leq x, A)}{P(A)}$$

$$p_X(x|A) = \frac{dF_X(x|A)}{dx}$$

O Teorema de Bayes (3.2) pode ser generalizado para o caso de densidades condicionais. Sejam X e Y duas variáveis aleatórias e, a priori (ou, seja, antes de saber-se alguma informação acerca de Y), X tem densidade de probabilidade $p_X(x)$. Se $Y = y$ modifica a densidade de probabilidade de X , essa densidade modificada é denominada *densidade de probabilidade condicional* e é dada por [2, 19]:

$$p_{X|Y}(x|y) = P[(X \leq x)|(Y = y)] = \frac{p_{X,Y}(x, y)}{p_Y(y)}, \quad (3.3)$$

assumindo-se que $p_Y(y) \neq 0$.

O *valor esperado* ou *esperança* de X ($E[X]$) é definido como o seu valor médio em um grande número de experimentos. Seja X uma v.a. discreta tal que $X = \{x_1, x_2, \dots\}$ e $P(x_i)$ a probabilidade de $X = x_i$. O valor esperado de X é dado por [1]:

$$E[X] = \sum_{i=1}^{\infty} x_i P(x_i)$$

Se X assumir um número finito de valores, o valor esperado pode ser considerado como uma média ponderada dos possíveis valores de x_i , cujos pesos são as probabilidades correspondentes $P(x_i)$:

$$E[X] = \sum_{i=1}^n x_i P(x_i)$$

É possível calcular, por exemplo, o valor esperado para a saída de um dado. Anteriormente, vimos que a variável aleatória, nesse caso, é $X = \{1, 2, 3, 4, 5, 6\}$.

Espera-se ver cada resultado possível em $\frac{1}{6}$ das vezes (ou seja, $P(1) = \dots = P(6) = \frac{1}{6}$). Assim, o valor esperado será dado por:

$$E[X] = (1)\frac{1}{6} + (2)\frac{1}{6} + \dots + (6)\frac{1}{6} = \frac{21}{6} = 3,5$$

Note que, mesmo que o valor esperado seja 3,5, esse nunca será o resultado de um lançamento de dado. Portanto, o valor esperado não é, necessariamente, o que se espera obter como resultado de um experimento.

Para os casos em que X é uma v.a. contínua, o valor esperado é dado pela integral [19]:

$$E[X] = \int_{-\infty}^{+\infty} xp_X(x)dx,$$

onde $p_X(x)$ é a pdf de X .

Como exemplo, considere X uma v.a. cuja pdf é dada por [1]:

$$p_X(x) = \begin{cases} \frac{1}{2}x, & 0 \leq x \leq 2 \\ 0, & x < 0 \text{ ou } x > 2 \end{cases}$$

O valor esperado, nesse caso, será:

$$E[X] = \int_{-\infty}^{+\infty} xp_X(x)dx = \int_0^2 x\frac{1}{2}x dx = \int_0^2 \frac{1}{2}x^2 dx = \frac{4}{3}$$

O valor esperado é um operador linear, ou seja [9]:

$$\begin{aligned} E[cX] &= cE[X], c \text{ constante} \\ E[X + Y] &= E[X] + E[Y] \end{aligned}$$

Também faz sentido definir o *valor esperado condicional* de X , dado que $Y = y$. Para isso, utiliza-se a pdf condicional (3.3), e segue que [2]:

$$E[X|Y = y] = \int_{-\infty}^{+\infty} xp_{X|Y}(x|y)dx$$

Considere, por exemplo, X como o tempo de vida útil de uma pilha fabricada por uma empresa, e $E[X] = 500$ horas. Isso pode significar que a maioria das pilhas tem vida útil entre 450 e 550 horas, mas também que metade das pilhas tem vida útil de 200 horas e outra metade tem vida útil de 800 horas. Para que se possa discernir as situações, é necessário introduzir o conceito de variância.

A *variância* (σ^2) de uma v.a. X é uma medida do quanto esperamos que X varie de sua média, ao quadrado. Considere, por exemplo, o lançamento de um dado viciado, cuja saída é sempre 1. Nesse caso, $\sigma^2(X) = 0$. A variância é definida como o desvio médio quadrado entre uma v.a. e o seu valor esperado, e é dada por [9, 19]:

$$\sigma^2 = E[(X - E[X])^2]$$

No caso de variáveis aleatórias discretas, tem-se [22]:

$$\sigma^2 = E[(X - E[X])^2] = \sum_{i=1}^{\infty} (x_i - E[X])^2$$

Para variáveis aleatórias contínuas, a variância é dada por [2]:

$$\sigma^2 = E[(X - E[X])^2] = \int_{-\infty}^{+\infty} (x - E[X])^2 p_X(x) dx$$

Note que a variância é dada pelo quadrado da unidade de X , o que acaba dificultando a interpretação do resultado. Por isso, é comum utilizar-se o *desvio-padrão* (σ), que nada mais é do que a raiz quadrada da variância e, por isso, tem a mesma unidade de medida que a v.a. X [22].

3.2.1 As distribuições Uniforme e Normal de probabilidade

Duas importantes formas de distribuição de probabilidade para variáveis aleatórias contínuas são a *uniforme* e a *normal* (ou gaussiana). Na distribuição uniforme, a pdf é um valor constante num determinado intervalo finito e é definida por

[19]:

$$p_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & x < a \text{ ou } x > b \end{cases}$$

Isso indica que a variável aleatória X tem a mesma probabilidade de obter qualquer valor entre a e b , e probabilidade zero de obter um valor fora desse intervalo. A figura (3.1) ilustra um exemplo de distribuição uniforme de probabilidade no intervalo $[1, 3]$, ou seja [19]:

$$p_X(x) = \begin{cases} \frac{1}{2}, & 1 \leq x \leq 3 \\ 0, & x < 1 \text{ ou } x > 3 \end{cases}$$

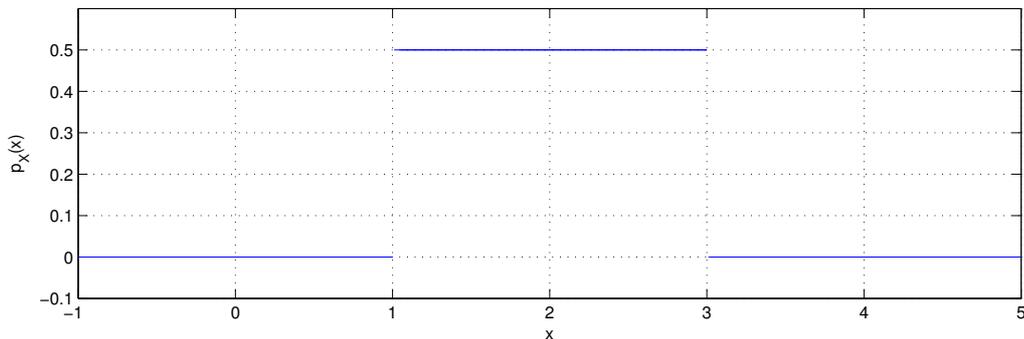


Figura 3.1: Distribuição Uniforme

O valor esperado dessa distribuição é dado por:

$$E[X] = \int_{-\infty}^{+\infty} xp_X(x)dx = \int_1^3 x \frac{1}{2} dx = 2$$

A variância é dada por:

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - E[X])^2 p_X(x) dx = \int_1^3 (x - 2)^2 \frac{1}{2} dx = \frac{1}{3}$$

A distribuição normal ou gaussiana de probabilidade depende de dois parâmetros: o valor esperado $E[X]$ e a variância σ^2 , e tem a seguinte forma analítica

[9]:

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-E[X])^2}{2\sigma^2}} \quad (3.4)$$

A figura (3.2) mostra um exemplo de distribuição normal, com valor esperado 10 e variância 9 (ou, equivalentemente, desvio-padrão 3).

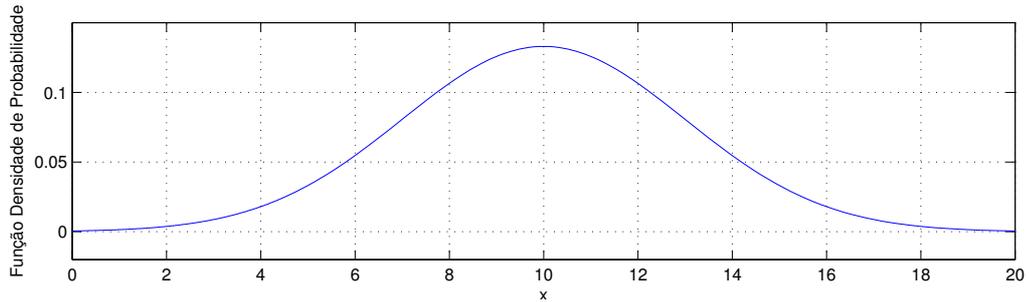


Figura 3.2: Distribuição Normal com $E[X] = 10$ e $\sigma^2 = 9$

3.2.2 Variáveis aleatórias múltiplas

Considere duas variáveis aleatórias X e Y , cujas distribuições de probabilidade são dadas por [19]:

$$F_X(x) = P(X \leq x)$$

$$F_Y(y) = P(Y \leq y)$$

A probabilidade de que $X \leq x$ e $Y \leq y$ é definida como a função distribuição de probabilidade conjunta de X e Y e é dada por:

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

A função densidade de probabilidade conjunta é dada pela derivada de F_{XY} :

$$p_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial y \partial x}$$

Na seção 3.1, foi inserido o conceito de eventos independentes (quando a ocorrência de um evento não interfere na probabilidade de ocorrência do outro). Seguindo esse raciocínio, pode-se dizer que duas variáveis aleatórias X e Y são independentes se, para qualquer x e y , a seguinte relação for satisfeita [19]:

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

A partir dessa definição e da definição de funções densidade, temos que:

$$F_{XY}(x, y) = F_X(x)F_Y(y)$$

$$p_{XY}(x, y) = p_X(x)p_Y(y)$$

Além disso, quando X e Y são variáveis independentes, vale a seguinte propriedade para o valor esperado [9]:

$$E[XY] = E[X]E[Y] \tag{3.5}$$

O Teorema do Limite Central afirma que a soma de variáveis aleatórias independentes tende a uma v.a. gaussiana, independentemente da pdf de cada uma das variáveis aleatórias. Na natureza, muitas variáveis aleatórias são a soma de outras, individuais e independentes. A temperatura máxima de um determinado local, por exemplo, tende a seguir uma distribuição Gaussiana. Isso ocorre porque a temperatura máxima é afetada por nuvens, precipitação, vento, pressão do ar, umidade e outros fatores. Cada um desses fatores é, por sua vez, determinado por outros fatores aleatórios. A combinação de muitas variáveis aleatórias independentes determina a temperatura máxima, que possui uma pdf gaussiana [19].

A *covariância* indica o grau de relação entre duas variáveis aleatórias, e é dada pelo valor esperado do produto dos seus desvios e suas esperanças:

$$\begin{aligned} C_{XY} &= E[(X - E[X])(Y - E[Y])] = E[XY - XE[Y] - YE[X] + E[X]E[Y]] = \\ &= E[XY] - E[XE[Y]] - E[YE[X]] + E[X]E[Y] = \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

Já a *correlação* entre duas variáveis aleatórias é definida por:

$$R_{XY} = E[XY]$$

Se $R_{XY} = E[X]E[Y]$, as variáveis aleatórias X e Y são ditas *descorrelacionadas*.

De acordo com a definição de variáveis aleatórias independentes (3.5), segue que a independência implica a não-correlação. Porém, a falta de correlação não implica, necessariamente, a independência - a exceção do caso especial em que as variáveis aleatórias são também gaussianas [2, 19]

3.2.3 Variáveis aleatórias multivariadas

Todos os conceitos que vimos até agora podem ser estendidos para quando as variáveis aleatórias são vetores. A seguir, serão apresentados os principais deles.

Considere as variáveis aleatórias X e Y , dadas por vetores coluna de n e m elementos, respectivamente. Ou seja, $X = [x_1 \dots x_n]^T$ e $Y = [y_1 \dots y_m]^T$.

O valor esperado, nesse caso, é tomado em cada componente. No caso da variável aleatória X , por exemplo, tem-se:

$$E[X] = E \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} E[x_1] \\ \vdots \\ E[x_n] \end{bmatrix}$$

A correlação entre X e Y é dada por [19]:

$$R_{XY} = E[XY^T] = E \begin{bmatrix} x_1y_1 & \cdots & x_1y_m \\ \vdots & \ddots & \vdots \\ x_ny_1 & \cdots & x_ny_m \end{bmatrix} = \begin{bmatrix} E[x_1y_1] & \cdots & E[x_1y_m] \\ \vdots & \ddots & \vdots \\ E[x_ny_1] & \cdots & E[x_ny_m] \end{bmatrix} =$$

E a covariância é definida como:

$$C_{XY} = E[(X - E[X])(Y - E[Y])^T] = E[XY^T] - E[X]E[Y]^T$$

Como, agora, cada variável aleatória é um vetor, faz sentido verificar a correlação entre os próprios elementos desse vetor, a qual é denominada *autocorrelação*. Por exemplo, no caso da v.a. X , tem-se:

$$R_X = E[XX^T] = E \begin{bmatrix} x_1x_1 & \cdots & x_1x_n \\ \vdots & \ddots & \vdots \\ x_nx_1 & \cdots & x_nx_n \end{bmatrix} = \begin{bmatrix} E[x_1^2] & \cdots & E[x_1x_n] \\ \vdots & \ddots & \vdots \\ E[x_nx_1] & \cdots & E[x_n^2] \end{bmatrix}$$

Note que $R_X = R_X^T$, pois $E[x_ix_j] = E[x_jx_i]$, $i = 1, \dots, n$. Qualquer matriz de autorrelação satisfaz essa propriedade e, portanto, será sempre simétrica.

É possível definir também a *autocovariância* entre os elementos de uma variável aleatória. No caso da v.a. X , tem-se [19]:

$$\begin{aligned} C_X &= E[(X - E[X])(X - E[X])^T] = \\ &= \begin{bmatrix} E[(x_1 - E[x_1])^2] & \cdots & E[(x_1 - E[x_1])(x_n - E[x_n])] \\ \vdots & \ddots & \vdots \\ E[(x_n - E[x_n])(x_1 - E[x_1])] & \cdots & E[(x_n - E[x_n])^2] \end{bmatrix} \end{aligned}$$

A diagonal da matriz C_X é composta pela variância de cada elemento x_i de X . Os demais elementos da matriz representam a covariância entre x_i e x_j , $i, j = 1, \dots, n$. Como a covariância entre duas variáveis é a mesma, independente da ordem, ou seja, $E[(x_i - E[x_i])(x_j - E[x_j])] = E[(x_j - E[x_j])(x_i - E[x_i])]$, tem-se que a matriz de autovariância também é simétrica [19, 22].

Uma variável aleatória X é dita gaussiana se a sua pdf for definida como:

$$p(X) = \frac{1}{\sqrt{(2\pi)^n C_X}} e^{-\frac{1}{2}(X - E[X])^T C_X^{-1} (X - E[X])} \quad (3.6)$$

Note que a versão multivariada é análoga à versão apresentada na seção 3.2.1, sendo que o valor esperado e a variância, antes escalares, foram substituídos por um vetor e por uma matriz (da autocovariância entre os elementos de X), respectivamente.

Ainda no contexto de variáveis aleatórias multidimensionais, pode-se falar de valor esperado e covariância condicionais. O seguinte teorema [2] trata disso, quando as variáveis são assumidas gaussianas, e será usado no capítulo 4, na dedução das equações do Filtro de Kalman.

Teorema 3.1. *Se X e Y são gaussianas, com $Z = \begin{bmatrix} X \\ Y \end{bmatrix}$, com valor esperado m e covariância Σ dados por*

$$m = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} \text{ e } \Sigma = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$

então a variável aleatória X , quando condicionada à informação que $Y = y$, é gaussiana, com valor esperado e covariância condicionais dadas por

$$E[X|Y] = \bar{x} + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \bar{y})$$

$$\Sigma_{X|Y} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}$$

Dem.: 1. *Assumindo a não-singularidade da matriz Σ , tem-se*

$$p_{X|Y}(x|y) = \frac{p_{XY}(x, y)}{p_Y(y)} =$$

$$= \frac{1}{\sqrt{(2\pi)^N}} \frac{\sqrt{|\Sigma_{yy}|} e^{-\frac{1}{2}[x^T - \bar{x}^T; y^T - \bar{y}^T]\Sigma^{-1}[x^T - \bar{x}^T; y^T - \bar{y}^T]^T}}{\sqrt{|\Sigma|} e^{-\frac{1}{2}(y - \bar{y})^T \Sigma_{yy}^{-1} (y - \bar{y})}},$$

onde N é a dimensão de X . Essa pdf pode ser escrita usando-se a seguinte fórmula:

$$\begin{bmatrix} I & -\Sigma_{xy}\Sigma_{yy}^{-1} \\ 0 & I \end{bmatrix} \Sigma \begin{bmatrix} I & 0 \\ -\Sigma_{yy}^{-1}\Sigma_{yx}^T & I \end{bmatrix} = \begin{bmatrix} \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} & 0 \\ 0 & \Sigma_{yy} \end{bmatrix}$$

Note que, tomando o determinante dos dois lados da equação, tem-se:

$$|\Sigma| = |\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}| |\Sigma_{yy}|$$

Além disso, segue que

$$\begin{aligned}
& [x^T - \bar{x}^T : y^T - \bar{y}^T] \Sigma^{-1} [x^T - \bar{x}^T : y^T - \bar{y}^T]^T = \\
= [x^T - \bar{x}^T : y^T - \bar{y}^T] & \begin{bmatrix} I & 0 \\ -\Sigma_{yy}^{-1} \Sigma_{xy}^T & I \end{bmatrix} \begin{bmatrix} (\Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx})^{-1} & 0 \\ 0 & \Sigma_{yy}^{-1} \end{bmatrix} \\
& \times \begin{bmatrix} I & -\Sigma_{xy} \Sigma_{yy}^{-1} \\ 0 & I \end{bmatrix} [x^T - \bar{x}^T : y^T - \bar{y}^T]^T = \\
& = (x^T - \bar{x}^T) (\Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx})^{-1} (x - \bar{x}),
\end{aligned}$$

onde $\bar{\bar{x}} = \bar{x} + \Sigma_{xy} \Sigma_{yy}^{-1} (y - \bar{y})$.

Desta forma, pode-se reescrever a pdf como

$$p_{X|Y}(x|y) = \frac{1}{\sqrt{(2\pi)^N} \sqrt{|\Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx}|}} e^{-\frac{1}{2} (x^T - \bar{\bar{x}}^T) (\Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx})^{-1} (x - \bar{\bar{x}})}$$

Note que a equação acima, se comparada à equação 3.6, mostra que X é uma variável condicionalmente gaussiana, com variância e média dadas, respectivamente, por

$$\begin{aligned}
& \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx} \\
& \bar{\bar{x}} + \Sigma_{xy} \Sigma_{yy}^{-1} (y - \bar{y})
\end{aligned}$$

3.3 Processos estocásticos

Um *processo estocástico*, também chamado de processo aleatório, é uma generalização do conceito de variável aleatória. Um processo estocástico $X(t)$ é uma v.a. X que muda com o tempo. Um processo estocástico pode ser de um destes quatro tipos [19]:

- Se a v.a. em cada tempo for contínua e o tempo for contínuo, então $X(t)$ é um processo aleatório contínuo. Por exemplo, a temperatura

em cada momento do dia é um processo aleatório contínuo porque a temperatura e o tempo são contínuos;

- Se a v.a. em cada momento for discreta e o tempo for contínuo, então $X(t)$ é um processo aleatório discreto. Por exemplo, o número de pessoas em um determinado edifício em cada momento do dia é um processo aleatório discreto, porque o número de pessoas é uma variável discreta e o tempo é contínuo;
- Se a v.a. em cada tempo for contínua e o tempo for discreto, então $X(t)$ é uma sequência aleatória contínua. Por exemplo, a temperatura máxima a cada dia é uma sequência aleatória contínua porque a temperatura é contínua, mas o tempo é discreto (dia um, dia dois, etc.);
- Se a v.a. em cada momento for discreta e o tempo for discreto, então $X(t)$ é uma sequência aleatória discreta. Por exemplo, o maior número de pessoas em um determinado edifício a cada dia é uma sequência aleatória discreta porque o número de pessoas é uma variável discreta e o tempo também é discreto.

Se um processo estocástico é uma v.a. que muda com o tempo, ele possui uma função de distribuição e densidade que são funções do tempo. A função distribuição de probabilidade de $X(t)$ será dada por

$$F_X(x, t) = P(X(t) \leq x)$$

Se $X(t)$ é um vetor aleatório, então a desigualdade acima será aplicada elemento a elemento. Por exemplo, se $X(t)$ tem n elementos, tem-se

$$F_X(x, t) = P[X_1(t) \leq x_1, \dots, X_n(t) \leq x_n]$$

A função densidade de probabilidade (pdf) de $X(t)$ é dada por

$$p_X(x, t) = \frac{dF_X(x, t)}{dx}$$

Novamente, se $X(t)$ for um vetor aleatório, então a derivada acima é tomada uma vez em relação a cada elemento de x . Por exemplo, se $X(t)$ tem n elementos [19],

$$f_X(x, t) = \frac{d^{(n)}F_X(x, t)}{dx_1 \cdots dx_n}$$

Uma classe de processos estocásticos muito estudada na literatura são os *processos Markovianos* [15]. De forma sucinta, um processo é dito Markoviano se, dado que o presente é conhecido, o passado não afeta o futuro, ou seja [2]:

$$p_{x_{k+1}|x_1, x_2, \dots, x_k}(x_{k+1}|x_1, x_2, \dots, x_k) = p_{x_{k+1}|x_k}(x_{k+1}|x_k)$$

3.4 Estimação Bayesiana

Os problemas de estimação, também chamados de problemas de filtragem, aplicam-se a sistemas dinâmicos, a fim de prever estados desconhecidos através de medições da saída, geralmente perturbadas por ruído. As entradas do sistema também podem ser ruidosas [2].

Um exemplo de filtragem é a forma como o cérebro humano aborda o problema da leitura de algo manuscrito por outra pessoa. Cada palavra é abordada sequencialmente, e quando se chega a uma palavra difícil de interpretar, as palavras anteriores (e posteriores) à palavra em questão podem ser usadas para tentar deduzí-la. Um outro exemplo de aplicação de filtragem no processamento de informação pelo cérebro humano é a situação em que deseja-se pegar uma bola. Para isso, é necessário prever a trajetória do objeto para posicionar as mãos corretamente. Essa tarefa se torna mais difícil à medida que a bola é sujeita a distúrbios aleatórios, como rajadas de vento. Em geral, qualquer previsão torna-se mais difícil à medida que o ambiente se torna mais ruidoso [2].

Em situações reais, frequentemente as observações, medições e controle de sistemas são feitas em tempo discreto. Por isso, neste trabalho, todos os sistemas serão admitidos em tempo discreto.

Considere, então, o sistema não-linear descrito pelas seguintes equações [19]:

$$x_{k+1} = f_k(x_k, w_k) \quad (3.7)$$

$$y_k = h_k(x_k, v_k), \quad (3.8)$$

onde x_k é o estado, w_k é o ruído do processo, y_k corresponde às medições e v_k ao ruído das medições. As funções $f_k(\cdot)$ e $h_k(\cdot)$ são não-lineares. As sequências dos ruídos $\{w_k\}$ e $\{v_k\}$ são assumidas como independentes e brancas, com pdf's conhecidas. A estimação Bayesiana visa aproximar a pdf do caminho $X_k = x_0, \dots, x_k$, condicionada às medidas y_1, y_2, \dots, y_k . Essa pdf condicional é denotada por $p(X_k|Y_k)$, sendo $Y_k = \{y_1, y_2, \dots, y_n\}$.

Para isso, se faz necessário uma teoria mais geral, que se aplique a distribuições genéricas (e não, necessariamente, gaussianas) [18].

Dado o sistema descrito pelas equações (3.7) e (3.8), deseja-se determinar:

$$p(x_{0:k}|Y_k),$$

que é, na verdade, a probabilidade de todos os estados, dadas todas as observações, até o tempo k .

O objetivo é ter uma relação recursiva, a qual utilize apenas a estimativa do estado anterior e a última medida para estimar o próximo estado, ou seja [18]:

$$p(x_{0:k}|Y_k) = f[p(x_{0:k-1}|Y_{k-1}), y_k] \quad (3.9)$$

Para isso, será necessário definir algumas notações, identidades e efetuar algumas suposições que serão usadas na prova dessa recursão. Primeiramente, tem-

se:

$$X_k = x_{0:k} = x_0 \text{ E } x_1 \text{ E } \dots \text{ E } x_k = x_0, x_1, \dots, x_k = x_0 \cap x_1 \cap \dots \cap x_k$$

É necessário também determinar duas identidades. A **identidade 1**, a qual já foi discutida na seção 3.1, é definida por:

$$p(x, y) = p(x|y)p(y)$$

Desta forma, é possível escrever $p(a, b, c)$ das seguintes formas [18]:

$$p(a, b, c) = p(\{a, b\}, c) = p(\{a, b\}|c)p(c) = p(a, b|c)p(c) \quad (3.10)$$

$$p(a, b, c) = p(a, \{b, c\}) = p(a|\{b, c\})p(\{b, c\}) = p(a|b, c)p(b|c)p(c) \quad (3.11)$$

Igualando as equações (3.10) e (3.11), e simplificando os termos $p(c)$, temos a **identidade 2**:

$$p(a, b|c) = p(a|b, c)p(b|c)$$

Também será assumido que o processo em questão é um processo de Markov. Como foi brevemente comentado na seção 3.3, isso significa que o próximo estado depende somente do estado atual, e não de todos os estados anteriores, ou seja:

$$p(x_k|X_{k-1}) = p(x_k|x_{k-1})$$

A segunda suposição é que as observações são independentes. Isso significa que a observação no estado atual depende apenas do estado atual [18]:

$$p(y_k|x_k, \text{ outras informações}) = p(y_k|x_k)$$

Agora, de acordo com o Teorema de Bayes, discutido na seção 3.1:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)},$$

é possível escrever a probabilidade que desejamos escrever recursivamente, na equação (3.9) da seguinte forma:

$$p(X_k|Y_k) = \frac{p(Y_k|X_k)p(X_k)}{p(Y_k)}$$

As observações de 0 até k podem ser agrupadas como o conjunto de todas as observações de 0 até $k - 1$ e a observação em k [18]:

$$p(X_k|Y_k) = \frac{p(y_k, Y_{k-1}|X_k)p(X_k)}{p(y_k, Y_{k-1})}$$

Agora, é possível aplicar a identidade 2 no primeiro termo no numerador:

$$p(X_k|Y_k) = \frac{p(y_k|Y_{k-1}, X_k)p(Y_{k-1}|X_k)p(X_k)}{p(y_k, Y_{k-1})}$$

E a identidade 1 no termo do denominador:

$$p(X_k|Y_k) = \frac{p(y_k|Y_{k-1}, X_k)p(Y_{k-1}|X_k)p(X_k)}{p(y_k|Y_{k-1})p(Y_{k-1})} \quad (3.12)$$

O termo central do numerador pode ser reescrito, usando a regra de Bayes, da seguinte forma:

$$p(Y_{k-1}|X_k) = \frac{p(X_k|Y_{k-1})p(Y_{k-1})}{p(X_k)}$$

Substituindo na equação (3.12), tem-se:

$$\begin{aligned} p(X_k|Y_k) &= \frac{p(y_k|Y_{k-1}, X_k)p(X_k|Y_{k-1})p(Y_{k-1})p(X_k)}{p(y_k|Y_{k-1})p(Y_{k-1})p(X_k)} = \\ &= \frac{p(y_k|Y_{k-1}, X_k)p(X_k|Y_{k-1})}{p(y_k|Y_{k-1})} \end{aligned}$$

De acordo com hipótese de observações são independentes, o primeiro termo do numerador pode ser reescrito como $p(y_k|Y_{k-1}, X_k) = p(y_k|x_k)$ e tem-se:

$$p(X_k|Y_k) = \frac{p(y_k|x_k)p(X_k|Y_{k-1})}{p(y_k|Y_{k-1})}$$

Os estados de 0 até k , no segundo termo do numerador, podem ser agrupados como o conjunto dos estados de 0 até $k - 1$ e o estado em k , o que permite usar a identidade 2 [18]:

$$\begin{aligned} p(X_k|Y_k) &= \frac{p(y_k|x_k)p(x_k, X_{k-1}|Y_{k-1})}{p(y_k|Y_{k-1})} = \\ &= \frac{p(y_k|x_k)p(x_k|X_{k-1}, Y_{k-1})p(X_{k-1}|Y_{k-1})}{p(y_k|Y_{k-1})p(Y_{k-1})} \end{aligned}$$

O segundo termo do numerador pode ser reescrito, usando a hipótese de que o processo é markoviano, como $p(x_k|x_{k-1})$. Assim, tem-se:

$$p(X_k|Y_k) = \frac{p(y_k|x_k)p(x_k|x_{k-1})p(X_{k-1}|Y_{k-1})}{p(y_k|Y_{k-1})p(Y_{k-1})}$$

Agora, apenas rearranjando os termos, segue que:

$$p(X_k|Y_k) = \frac{p(x_k|x_{k-1})p(y_k|x_k)}{p(y_k|Y_{k-1})}p(X_{k-1}|Y_{k-1}) \quad (3.13)$$

Essa é uma solução recursiva para a equação (3.9), a qual é denominada *estimativa bayesiana recursiva*, e é aplicável para qualquer distribuição, não apenas gaussianas [18]. No capítulo 5, esse resultado será utilizado para mostrar o funcionamento do Filtro de Partículas.

4 FILTRO DE KALMAN

Os problemas de filtragem ou de estimação, conforme discutido na seção 3.4, têm como objetivo prever estados desconhecidos de um sistema, através de medições de saída, geralmente perturbadas por ruído. Os próprios estados, e até mesmo as entradas do sistema, podem ser admitidos como ruidosos.

Normalmente, na prática, as observações e medições do sistema são feitas em intervalos definidos de tempo (a cada hora ou diariamente, por exemplo). Por isso, o problema de filtragem será aplicado ao seguinte sistema discreto, para $k > 0$ [2, 19]:

$$x_{k+1} = F_k x_k + G_k u_k + w_k \quad (4.1)$$

$$y_k = H_k^T x_k + v_k, \quad (4.2)$$

onde x_k é o estado do sistema no tempo k e F_k é a matriz de transição dos estados; $G_k u_k$ é a entrada do sistema, considerada determinística; H_k^T é a matriz que converte o estado x_k na medida e y_k é a saída do sistema; w_k e v_k são os ruídos relacionados aos estados e às medidas, respectivamente.

Será assumido aqui que as sequências dos ruídos $\{w_k\}$ e $\{v_k\}$:

- têm valor esperado zero ($E[w_k] = 0$ e $E[v_k] = 0$), são gaussianos e têm covariância conhecida;
- são brancos. Isso significa que, para $l \neq k$, v_l e v_k são variáveis aleatórias independentes, assim como w_l e w_k . Ou seja, $E[v_k v_l^T] = E[v_k] E[v_l^T] = 0$ e, analogamente, $E[w_k w_l^T] = 0$;
- são independentes, ou seja, $E[v_k] E[w_k] = E[v_k w_k^T] = 0$.

Será denotado aqui que a covariância do ruído w_k é dada por

$$Q_k = E[(w_k - E[w_k])(w_k - E[w_k])^T] = E[w_k w_k^T]$$

Desta forma, a covariância do processo $\{w_k\}$ é dada por:

$$E[(w_k - E[w_k])(w_l - E[w_l])^T] = E[w_k w_l^T] = Q_k \delta_{kl}, \quad (4.3)$$

para todo k e l , onde δ_{kl} é a matriz delta de Kronecker, na qual tem-se 1 para $k = l$ e 0 para $k \neq l$.

Analogamente, tem-se que a covariância de v no instante k é dada por

$$R_k = E[(v_k - E[v_k])(v_k - E[v_k])^T] = E[v_k v_k^T]$$

e, para todo k e l ,

$$E[(v_k - E[v_k])(v_l - E[v_l])^T] = E[v_k v_l^T] = R_k \delta_{kl} \quad (4.4)$$

O estado inicial x_0 será assumido como uma variável aleatória gaussiana de valor esperado $E[x_0]$ e covariância P_0 conhecidas, dados por:

$$\begin{aligned} \bar{x}_0 &= E[x_0] \\ P_0 &= E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] \end{aligned}$$

Também será assumido que x_0 é independente dos ruídos v_k e w_k .

4.1 Propriedades dos estados do sistema

Três importantes propriedades do processo aleatório $\{x_k\}$ serão consideradas para a dedução do Filtro de Kalman [2].

A primeira é que x_k é uma variável aleatória gaussiana. Isso significa que o ruído do sistema $\{w_k\}$ e o estado inicial x_0 colaboram para que x_k seja gaussiano [2]. Note que, de acordo com a equação (4.1), é possível escrever x_k como

segue:

$$\begin{aligned}
x_k &= F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} = \\
&= F_{k-1}(F_{k-2}x_{k-2} + G_{k-2}u_{k-2} + w_{k-2}) + G_{k-1}u_{k-1} + w_{k-1} \\
&= F_{k-1}F_{k-2}x_{k-2} + F_{k-1}(G_{k-2}u_{k-2} + w_{k-2}) + G_{k-1}u_{k-1} + w_{k-1} \\
&= F_{k-1}F_{k-2}(F_{k-3}x_{k-3} + G_{k-3}u_{k-3} + w_{k-3}) + \\
&\quad + F_{k-1}(G_{k-2}u_{k-2} + w_{k-2}) + G_{k-1}u_{k-1} + w_{k-1} = \\
&\quad \vdots \\
&= F_{k-1} \dots F_0 x_0 + F_{k-1} \dots F_1 (G_0 u_0 + w_0) + F_{k-1} \dots F_2 (G_1 u_1 + w_1) + \\
&\quad + \dots + G_{k-1} u_{k-1} + w_{k-1}
\end{aligned}$$

Desta forma, x_k pode ser reescrito como [2, 19]:

$$x_k = \Phi_{k,0}x_0 + \sum_{l=0}^{k-1} (\Phi_{k,l+1}G_l u_l + \Phi_{k,l+1}w_l), \quad (4.5)$$

onde $\Phi_{k,l}$ é a matriz de transição da equação homogênea $x_{k+1} = F_k x_k$, dada por:

$$\Phi_{k,l} = F_{k-1}F_{k-2} \dots F_l, k > l$$

Note que:

$$\Phi_{k,l} = F_{k-1} \dots F_m F_{m-1} \dots F_l = \Phi_{k,m} \Phi_{m,l}, k > m > l$$

$$\Phi_{k,k} = I$$

$$\Phi_{k,l} = \Phi_{l,k}^T, k < l$$

A equação (4.5) expressa x_k como uma combinação linear de x_0 , $\{w_k\}$ e $\{u_k\}$. Como a entrada do sistema é determinística, então a sequência $\{G_k u_k\}$ pode ser considerada como uma sequência de variáveis aleatórias gaussianas de covariância zero. Foi assumido também que x_0 e a sequência $\{w_k\}$ são variáveis aleatórias gaussianas. Como a combinação linear de variáveis aleatórias gaussianas preserva essa característica, segue que x_k também é variável aleatória gaussiana [2, 19].

A segunda propriedade é que $\{x_k\}$ é um processo aleatório gaussiano, o que decorre diretamente da propriedade anterior.

A terceira propriedade é a hipótese de $\{x_k\}$ ser um processo de Markov.

Ou seja:

$$p(x_k | x_{k-1}, x_{k-2}, \dots, x_{k-m}) = p(x_k | x_{k-1}), k > m$$

4.2 Propagação dos valores esperados e das covariâncias

Da equação (4.5) e do fato de assumir-se que $E[w_k] = 0$, tem-se [2]:

$$\bar{x}_k = E[x_k] = \Phi_{k,0}\bar{x}_0 + \sum_{l=0}^{k-1} \Phi_{k,l+1}G_l u_l \quad (4.6)$$

E, da equação (4.1), segue que:

$$\bar{x}_{k+1} = E[x_{k+1}] = E[F_k x_k + G_k u_k + w_k] = F_k \bar{x}_k + G_k u_k \quad (4.7)$$

Já da equação (4.2) e da hipótese de que $E[v_k] = 0$, tem-se:

$$\bar{y}_k = E[y_k] = E[H_k^T x_k + v_k] = H_k^T \bar{x}_k \quad (4.8)$$

A covariância $P_{k,l}$, para $k \geq l$, é dada por:

$$P_{k,l} = E[(x_k - \bar{x}_k)(x_l - \bar{x}_l)^T] \quad (4.9)$$

Usando as equações (4.5) e (4.6), é possível reescrever $P_{k,l}$ como:

$$\begin{aligned}
P_{k,l} = E \left[\left(\Phi_{k,0}x_0 + \sum_{i=0}^{k-1} \Phi_{k,i+1}G_iu_i + \right. \right. \\
\left. \left. \sum_{i=0}^{k-1} \Phi_{k,i+1}w_i - \Phi_{k,0}\bar{x}_0 - \sum_{i=0}^{k-1} \Phi_{k,i+1}G_iu_i \right) \right. \\
\left. \left(\Phi_{l,0}x_0 + \sum_{i=0}^{l-1} \Phi_{l,i+1}G_iu_i + \sum_{i=0}^{l-1} \Phi_{l,i+1}w_i - \right. \right. \\
\left. \left. - \Phi_{l,0}\bar{x}_0 - \sum_{i=0}^{l-1} \Phi_{l,i+1}G_iu_i \right)^T \right]
\end{aligned}$$

Os somatórios referentes às entradas se cancelam e segue que:

$$\begin{aligned}
P_{k,l} &= E \left[\left(\Phi_{k,0}(x_0 - \bar{x}_0) + \sum_{i=0}^{k-1} \Phi_{k,i+1}w_i \right) \right. \\
&\quad \left. \left(\Phi_{l,0}(x_0 - \bar{x}_0) + \sum_{i=0}^{l-1} \Phi_{l,i+1}w_i \right)^T \right] = \\
&= \Phi_{k,0}E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T]\Phi_{l,0}^T + \\
&\quad + \Phi_{k,0}E \left[(x_0 - \bar{x}_0) \left(\sum_{i=0}^{l-1} \Phi_{l,i+1}w_i \right)^T \right] + \\
&\quad + E \left[\left(\sum_{i=0}^{k-1} \Phi_{k,i+1}w_i \right) (x_0 - \bar{x}_0)^T \right] \Phi_{l,0}^T + \\
&\quad + E \left[\sum_{i=0}^{k-1} \Phi_{k,i+1}G_iw_i \left(\sum_{i=0}^{l-1} \Phi_{l,i+1}w_i \right)^T \right]
\end{aligned}$$

Como $(x_0 - \bar{x}_0)$ e $\{w_k\}$ são variáveis aleatórias independentes, dois termos da equação acima serão nulos e segue que:

$$\begin{aligned}
P_{k,l} &= \Phi_{k,0}E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T]\Phi_{l,0}^T + E \left[\sum_{i=0}^{k-1} \Phi_{k,i+1}w_i \left(\sum_{i=0}^{l-1} \Phi_{l,i+1}w_i \right)^T \right] = \\
&= \Phi_{k,0}P_0\Phi_{l,0}^T + E \left[(\Phi_{k,1}w_0 + \dots + \Phi_{k,k}w_{k-1}) \right. \\
&\quad \left. (\Phi_{l,1}w_0 + \dots + \Phi_{l,l}w_{l-1})^T \right]
\end{aligned}$$

Como $k \geq l$ e o ruído w_k é branco, ou seja, para $w \neq l$, $E[w_k w_l^T] = 0$, a equação acima poderá ser reescrita como:

$$\begin{aligned}
P_{k,l} &= \Phi_{k,0} P_0 \Phi_{l,0}^T + \Phi_{k,1} w_0 w_0^T \Phi_{l,1}^T + \dots + \Phi_{l,l} w_{l-1} w_{l-1}^T \Phi_{l,l}^T = \\
&= \Phi_{k,0} P_0 \Phi_{l,0}^T + \sum_{i=0}^{l-1} \Phi_{k,i+1} Q_i \Phi_{l,i+1}^T = \\
&= \Phi_{k,l} \Phi_{l,0} P_0 \Phi_{l,0}^T + \sum_{i=0}^{l-1} \Phi_{k,l} \Phi_{l,i+1} Q_i \Phi_{l,i+1}^T
\end{aligned}$$

E, finalmente,

$$P_{k,l} = \Phi_{k,l} \left(\Phi_{l,0} P_0 \Phi_{l,0}^T + \sum_{i=0}^{l-1} \Phi_{l,i+1} Q_i \Phi_{l,i+1}^T \right) \quad (4.10)$$

Para encontrar uma equação mais fácil para $P_{k,l}$, primeiro note que $P_{k,k}$ pode ser escrita usando $k = l$ na equação acima, e o fato de que $\Phi_{k,k} = I$:

$$P_{k,k} = \Phi_{k,0} P_0 \Phi_{k,0}^T + \sum_{i=0}^{k-1} \Phi_{k,i+1} Q_i \Phi_{k,i+1}^T \quad (4.11)$$

Mas $P_{k,k}$ é a covariância de x_k , ou seja:

$$P_k = P_{k,k} = E[(x_k - \bar{x}_k)(x_k - \bar{x}_k)^T] \quad (4.12)$$

Usando as equações (4.5) e (4.6), segue que:

$$\begin{aligned}
P_k &= E[(F_{k-1} x_{k-1} + G_{k-1} u_{k-1} + w_{k-1} - F_{k-1} \bar{x}_{k-1} - G_{k-1} u_{k-1}) \\
&\quad (F_{k-1} x_{k-1} + G_{k-1} u_{k-1} + w_{k-1} - F_{k-1} \bar{x}_{k-1} - G_{k-1} u_{k-1})^T] = \\
&= E[(F_{k-1}(x_{k-1} - \bar{x}_{k-1}) + w_{k-1})(F_{k-1}(x_{k-1} - \bar{x}_{k-1}) + w_{k-1})^T] = \\
&= E[F_{k-1}(x_{k-1} - \bar{x}_{k-1})(F_{k-1}(x_{k-1} - \bar{x}_{k-1}))^T] + \\
&\quad + E[F_{k-1}(x_{k-1} - \bar{x}_{k-1})(w_{k-1})^T] + \\
&\quad + E[w_{k-1}(F_{k-1}(x_{k-1} - \bar{x}_{k-1}))^T] + E[w_{k-1}(w_{k-1})^T] = \\
&= E[F_{k-1}(x_{k-1} - \bar{x}_{k-1})(x_{k-1} - \bar{x}_{k-1})^T F_{k-1}^T] + E[w_{k-1} w_{k-1}^T] = \\
&= F_{k-1} E[(x_{k-1} - \bar{x}_{k-1})(x_{k-1} - \bar{x}_{k-1})^T] F_{k-1}^T + E[w_{k-1} w_{k-1}^T]
\end{aligned}$$

E, finalmente,

$$P_k = F_{k-1}P_{k-1}F_{k-1}^T + Q_{k-1} \quad (4.13)$$

É possível, então, escrever P_k em termos de P_{k-1} , e assim sucessivamente. Essa equação recursiva para a covariância permite determinar P_k a partir de P_0 .

Como $k \geq l$, conhecida P_l , pode-se reescrever $P_{k,l}$, usando a equação (4.10) e a equação (4.11):

$$P_{k,l} = \Phi_{k,l}P_l, k \geq l \quad (4.14)$$

Como $P_{k,l} = P_{l,k}^T$, tem-se:

$$P_{k,l} = P_k\Phi_{l,k}^T, k < l \quad (4.15)$$

4.3 O problema de filtragem

O problema de filtragem consiste em produzir uma estimativa do estado do sistema x_k , usando medidas até o tempo $k - 1$. Ou seja: utilizar a sequência de medições y_0, y_1, \dots, y_{k-1} , a qual será denotada por Y_{k-1} , para prever o valor de x_k . Depois, utiliza-se a medida y_k para atualizar a estimativa [2, 19].

A estimativa a priori de x_k , que será denotada por $\hat{x}_{k|k-1}$, utiliza as medidas até o instante $k - 1$ e é dada pelo cálculo do valor esperado condicional:

$$\hat{x}_{k|k-1} = E[x_k|Y_{k-1}]$$

Após a utilização com a medida no tempo k , obtém-se uma estimativa a posteriori, denotada por $\hat{x}_{k|k}$ e dada por [19]:

$$\hat{x}_{k|k} = E[x_k|Y_k]$$

Ao mesmo passo que deseja-se conhecer $\hat{x}_{k|k-1}$ e $\hat{x}_{k|k}$, também é importante saber o quão boas são essas estimativas. É possível mensurá-las pelas matrizes da covariância do erro, que serão denotadas por $P_{k|k-1}$ e $P_{k|k}$, dadas por [2, 19]:

$$P_{k|k-1} = E[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T]$$

$$P_{k|k} = E[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T]$$

A estimativa do estado inicial x_0 , antes de qualquer medida estar disponível, e a covariância do erro desta estimativa são dadas [2]:

$$\bar{x}_0 = E[x_0]$$

$$P_0 = E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T]$$

Em geral, P_0 representa a incerteza da estimativa inicial. Se o estado inicial fosse conhecido com precisão, então $P_0 = 0$ [19].

4.3.1 Resumo e solução do problema de filtragem de Kalman

Para o sistema linear em tempo discreto dado pelas equações:

$$x_{k+1} = F_k x_k + G_k u_k + w_k$$

$$y_k = H_k^T x_k + v_k,$$

definido para $k \geq 0$, supõe-se que w_k e v_k são independentes, têm valor esperado zero e são ruídos brancos gaussianos, com

$$E[w_k w_l^T] = Q_k \delta_{kl}$$

$$E[v_k v_l^T] = R_k \delta_{kl}$$

Supõe-se também que o estado inicial x_0 é uma variável aleatória gaussiana, com valor esperado \bar{x}_0 e covariância P_0 , e independente de w_k e v_k .

A partir dessas suposições, deseja-se determinar as estimativas de x_k , a priori e a posteriori, dadas, respectivamente, por:

$$\hat{x}_{k|k-1} = E[x_k|Y_{k-1}] \quad (4.16)$$

$$\hat{x}_{k|k} = E[x_k|Y_k] \quad (4.17)$$

e as suas matrizes de covariância do erro associadas $P_{k|k-1}$ e $P_{k|k}$, respectivamente.

O Filtro de Kalman é dividido em duas etapas: predição e correção. Na predição, a estimativa do estado no instante anterior $k-1$ é usada para prever a estimativa no instante atual k , através da equação do processo. Na correção, utiliza-se a equação de saída e a medida da observação para atualizar a estimativa no tempo k .

A solução para o Filtro de Kalman é dada pelas equações, para $k > 0$ [2, 19]:

$$\begin{aligned} \hat{x}_{k|k-1} &= F_{k-1}\hat{x}_{k-1|k-1} + G_{k-1}u_{k-1} \\ P_{k|k-1} &= F_{k-1}P_{k-1|k-1}F_{k-1}^T + Q_{k-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(y_k - H_k\hat{x}_{k|k-1}) \\ P_{k|k} &= (I - K_kH_k^T)P_{k|k-1} \end{aligned}$$

Sendo K_k , denominada matriz ganho de Kalman, dada por [2]:

$$K_k = P_{k|k-1}H_k(H_k^T P_{k|k-1}H_k + R_k)^{-1}, \quad (4.18)$$

assumindo que a inversa existe [2].

Essas equações são inicializadas por $\bar{x}_0 = \hat{x}_{0|-1}$ e $P_0 = P_{0|-1}$.

4.4 Dedução das equações do Filtro de Kalman

Para esta seção, será usado o resultado do Teorema 3.1 do capítulo 3, o qual segue:

Teorema: Se X e Y são gaussianas, com $Z = \begin{bmatrix} X \\ Y \end{bmatrix}$, com valor esperado m e covariância Σ

$$m = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} \text{ e } \Sigma = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$

então a variável aleatória X , quando condicionada à informação que $Y = y$, é gaussiana, com valor esperado e covariância condicionais dadas por

$$E[X|Y] = \bar{x} + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \bar{y})$$

$$\Sigma_{X|Y} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}$$

Portanto, para usar o teorema, será necessário verificar o valor esperado e a covariância da variável aleatória a ser condicionada, e da condição a ela submetida. Serão analisados alguns casos específicos e depois será feita a generalização.

Serão usadas as equações do sistema (4.1) e (4.2); as afirmações acerca da covariância dos ruídos w_k e v_k em (4.3) e (4.4); a independência entre $x_0 - \bar{x}_0$ e os ruídos w_k e v_k ; as equações a respeito das propagações dos valores esperados de x_k e y_k , dadas em (4.6) e (4.8); as equações sobre a covariância $P_{k,l}$, dadas por (4.14) e (4.15); e as definições da covariância P_k , dada em (4.12) e (4.13). Procede-se, então, da seguinte forma [2]:

1. Inicialmente, deseja-se verificar o valor esperado e a covariância de x_0 condicionado à medida y_0 . Para isso, considere a variável aleatória $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$. O valor esperado e a covariância dessa v.a. são dados, respectivamente, por:

$$E \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} \bar{x}_0 \\ H_0^T \bar{x}_0 \end{bmatrix}$$

$$\begin{aligned}
& E \left[\left(\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - E \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \right) \left(\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - E \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \right)^T \right] = \\
& = E \left[\begin{bmatrix} x_0 - \bar{x}_0 \\ H_0^T x_0 + v_0 - H_0^T \bar{x}_0 \end{bmatrix} \begin{bmatrix} x_0 - \bar{x}_0 \\ H_0^T x_0 + v_0 - H_0^T \bar{x}_0 \end{bmatrix}^T \right] = \\
& = E \left[\begin{bmatrix} x_0 - \bar{x}_0 \\ H_0^T (x_0 - \bar{x}_0) + v_0 \end{bmatrix} \begin{bmatrix} x_0 - \bar{x}_0 \\ H_0^T (x_0 - \bar{x}_0) + v_0 \end{bmatrix}^T \right] = \\
& = \begin{bmatrix} E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] & E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] H_0 \\ H_0^T E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] & H_0 E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] H_0^T + E[v_0 v_0^T] \end{bmatrix} \\
& = \begin{bmatrix} P_0 & P_0 H_0 \\ H_0^T P_0 & H_0^T P_0 H_0 + R_0 \end{bmatrix}
\end{aligned}$$

Então, de acordo com o Teorema 3.1, x_0 condicionado a y_0 tem valor esperado $\hat{x}_{0|0}$ e covariância $P_{0|0}$ dados, respectivamente, por:

$$\begin{aligned}
\hat{x}_{0|0} &= E[x_0|Y_0] = \bar{x}_0 + P_0 H_0 (H_0^T P_0 H_0 + R_0)^{-1} (y_0 - H_0^T \bar{x}_0) \\
P_{0|0} &= E[(x_0 - \hat{x}_{0|0})(x_0 - \hat{x}_{0|0})^T] = P_0 - P_0 H_0 (H_0^T P_0 H_0 + R_0)^{-1} H_0^T P_0
\end{aligned}$$

2. O próximo passo é estimar x_1 , dada a medida y_0 . Neste caso, deve-se considerar a variável aleatória $\begin{bmatrix} x_1 \\ y_0 \end{bmatrix}$. Usando as equações (4.6) e (4.8), segue que o valor esperado será:

$$E \begin{bmatrix} x_1 \\ y_0 \end{bmatrix} = \begin{bmatrix} \bar{x}_1 \\ H_0^T \bar{x}_0 \end{bmatrix} = \begin{bmatrix} F_0 \bar{x}_0 + G_0 u_0 \\ H_0^T \bar{x}_0 \end{bmatrix}$$

E, a partir das afirmações acerca da propagação de $P_{k,l}$ em (4.14) e (4.15), tem-se que a covariância de $[x_1 \ y_0]^T$ é dada por:

$$\begin{aligned}
& E \left[\left(\begin{bmatrix} x_1 \\ y_0 \end{bmatrix} - E \begin{bmatrix} x_1 \\ y_0 \end{bmatrix} \right) \left(\begin{bmatrix} x_1 \\ y_0 \end{bmatrix} - E \begin{bmatrix} x_1 \\ y_0 \end{bmatrix} \right)^T \right] = \\
& = E \left[\begin{bmatrix} x_1 - \bar{x}_1 \\ H_0^T x_0 + v_0 - H_0^T \bar{x}_0 \end{bmatrix} \begin{bmatrix} x_1 - \bar{x}_1 \\ H_0^T x_0 + v_0 - H_0^T \bar{x}_0 \end{bmatrix}^T \right] = \\
& = E \left[\begin{bmatrix} x_1 - \bar{x}_1 \\ H_0^T (x_0 - \bar{x}_0) + v_0 \end{bmatrix} \begin{bmatrix} x_1 - \bar{x}_1 \\ H_0^T (x_0 - \bar{x}_0) + v_0 \end{bmatrix}^T \right] = \\
& = \begin{bmatrix} P_1 & P_{1,0} H_0 \\ H_0^T P_{0,1} & H_0^T P_0 H_0 + R_0 \end{bmatrix} = \\
& = \begin{bmatrix} P_1 & \Phi_{1,0} P_0 H_0 \\ H_0^T \Phi_{0,1} P_0 & H_0^T P_0 H_0 + R_0 \end{bmatrix} = \\
& = \begin{bmatrix} F_0 P_0 F_0^T + Q_0 & F_0 P_0 H_0 \\ H_0^T P_0 F_0^T & H_0^T P_0 H_0 + R_0 \end{bmatrix}
\end{aligned}$$

Desta forma, usando o teorema 3.1, x_1 condicionado a y_0 tem valor esperado $\hat{x}_{1|0}$ e covariância $P_{1|0}$ dados por:

$$\begin{aligned}
\hat{x}_{1|0} & = E[x_1|Y_0] = F_0 \bar{x}_0 + G_0 u_0 + F_0 P_0 H_0 (H_0^T P_0 H_0 + R_0)^{-1} (y_0 - H_0^T \bar{x}_0) = \\
& = F_0 (\bar{x}_0 + P_0 H_0 (H_0^T P_0 H_0 + R_0)^{-1} (y_0 - H_0^T \bar{x}_0)) + B_0 u_0 = \\
& = F_0 \hat{x}_{0|0} + G_0 u_0 \\
P_{1|0} & = E[(x_1 - \hat{x}_{1|0})(x_1 - \hat{x}_{1|0})^T] = \\
& = F_0 P_0 F_0^T + Q_0 - F_0 P_0 H_0 (H_0^T P_0 H_0 + R_0)^{-1} H_0^T P_0 F_0^T = \\
& = F_0 (P_0 - P_0 H_0 (H_0^T P_0 H_0 + R_0)^{-1} H_0^T P_0) F_0^T + Q_0 = \\
& = F_0 P_{0|0} F_0^T + Q_0
\end{aligned}$$

3. Agora deve-se utilizar a medida y_1 para atualizar a estimativa de x_1 . Para isso, serão consideradas duas etapas: condiciona-se y_1 a y_0 e, depois,

condiciona-se x_1 às duas medidas. Note que a variável aleatória $\begin{bmatrix} y_1 \\ y_0 \end{bmatrix}$ tem valor esperado

$$E \begin{bmatrix} y_1 \\ y_0 \end{bmatrix} = \begin{bmatrix} H_1^T \bar{x}_1 \\ H_0^T \bar{x}_0 \end{bmatrix} = \begin{bmatrix} H_1^T F_0 \bar{x}_0 \\ H_0^T \bar{x}_0 \end{bmatrix}$$

e covariância

$$\begin{aligned} & E \begin{bmatrix} \begin{bmatrix} H_1^T x_1 + v_1 - H_1^T \bar{x}_1 \\ H_0^T x_0 + v_0 - H_0^T \bar{x}_0 \end{bmatrix} \begin{bmatrix} H_1^T x_1 + v_1 - H_1^T \bar{x}_1 \\ H_0^T x_0 + v_0 - H_0^T \bar{x}_0 \end{bmatrix}^T \\ &= E \begin{bmatrix} \begin{bmatrix} H_1^T (x_1 - \bar{x}_1) + v_1 \\ H_0^T (x_0 - \bar{x}_0) + v_0 \end{bmatrix} \begin{bmatrix} H_1^T (x_1 - \bar{x}_1) + v_1 \\ H_0^T (x_0 - \bar{x}_0) + v_0 \end{bmatrix}^T \\ &= \begin{bmatrix} H_1^T P_1 H_1 + R_1 & H_1^T P_{1,0} H_0 \\ H_0^T P_{0,1} H_1 & H_0^T P_0 H_0 + R_0 \end{bmatrix} \\ &= \begin{bmatrix} H_1^T (F_0 P_0 F_0^T + Q_0) H_1 + R_1 & H_1^T \Phi_{1,0} P_0^- H_0 \\ H_0^T P_0^- \Phi_{1,0}^T & H_0^T P_0^- H_0 + R_0 \end{bmatrix} \\ &= \begin{bmatrix} H_1^T (F_0 P_0 F_0^T + Q_0) H_1 + R_1 & H_1^T F_0 P_0 H_0 \\ H_0^T P_0 F_0^T & H_0^T P_0 H_0 + R_0 \end{bmatrix} \end{aligned}$$

Assim, usando o teorema 3.1, segue que y_1 condicionado a y_0 tem valor esperado $\hat{y}_{1|0}$ e covariância $cov(y_1|y_0)$ dados por:

$$\begin{aligned} \hat{y}_{1|0} &= E[y_1|Y_0] = H_1^T F_0 \bar{x}_0 + H_1^T F_0 P_0 H_0 (H_0^T P_0 H_0 + R_0)^{-1} (y_0 - H_0^T \bar{x}_0) = \\ &= H_1^T F_0 (\bar{x}_0 + P_0 H_0 (H_0^T P_0 H_0 + R_0)^{-1} (y_0 - H_0^T \bar{x}_0)) = \\ &= H_1^T F_0 \hat{x}_{0|0} = H_1^T \hat{x}_{1|0} \end{aligned}$$

$$\begin{aligned}
cov(y_1|y_0) &= H_1^T(F_0P_0F_0^T + G_0Q_0G_0^T)H_1 + R_1 - \\
&\quad - H_1^T F_0 P_0^- H_0 (H_0^T P_0^- H_0 + R_0)^{-1} H_0^T P_0^- F_0^T H_1 = \\
&= H_1^T(F_0P_0^-F_0^T + G_0Q_0G_0^T - \\
&\quad - F_0P_0^-H_0(H_0^T P_0^- H_0 + R_0)^{-1} H_0^T P_0^- F_0^T)H_1 + R_1 = \\
&= H_1^T P_1^- H_1 + R_1
\end{aligned}$$

Desta forma, covariância entre x_1 e y_1 , ambos condicionados a y_0 , será dada por:

$$\begin{aligned}
&E[(x_1 - \hat{x}_{1|0})(y_1 - \hat{y}_{1|0})^T] = \\
&= E[(x_1 - \hat{x}_{1|0})(H_1^T x_1 + v_1 - H_1^T \hat{x}_{1|0})^T] = \\
&= E[(x_1 - \hat{x}_{1|0})(H_1^T(x_1 - \hat{x}_{1|0}) + v_1)^T] = \\
&= E[(x_1 - \hat{x}_{1|0})(x_1 - \hat{x}_{1|0})^T H_1] = \\
&= E[(x_1 - \hat{x}_{1|0})(x_1 - \hat{x}_{1|0})^T] H_1 = P_{1|0} H_1
\end{aligned}$$

E a variável aleatória $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$, condicionada a y_0 , tem valor esperado:

$$\begin{bmatrix} E[x_1|y_0] \\ E[y_1|y_0] \end{bmatrix} = \begin{bmatrix} \hat{x}_{1|0} \\ \hat{y}_{1|0} \end{bmatrix} = \begin{bmatrix} \hat{x}_{1|0} \\ H_1^T \hat{x}_{1|0} \end{bmatrix}$$

e covariância

$$\begin{aligned}
&E \left[\begin{bmatrix} x_1 - \hat{x}_{1|0} \\ H_1^T x_1 + v_1 - H_1^T \hat{x}_{1|0} \end{bmatrix} \begin{bmatrix} x_1 - \hat{x}_{1|0} \\ H_1^T x_1 + v_1 - H_1^T \hat{x}_{1|0} \end{bmatrix}^T \right] = \\
&E \left[\begin{bmatrix} x_1 - \hat{x}_{1|0} \\ H_1^T(x_1 - \hat{x}_{1|0}) + v_1 \end{bmatrix} \begin{bmatrix} x_1 - \hat{x}_{1|0} \\ H_1^T(x_1 - \hat{x}_{1|0}) + v_1 \end{bmatrix}^T \right] = \\
&= \begin{bmatrix} P_{1|0} & P_{1|0} H_1 \\ H_1^T P_{1|0} & H_1^T P_{1|0} H_1 + R_1 \end{bmatrix}
\end{aligned}$$

Assim, aplicando novamente o teorema, conclui-se que x_1 condicionado a $Y_1 = y_0, y_1$ tem valor esperado $\hat{x}_{1|1}$ e covariância $P_{1|1}$ dados por:

$$\begin{aligned}\hat{x}_{1|1} &= E[x_1|Y_1] = \hat{x}_{1|0} + P_{1|0}H_1(H_1^T P_{1|0}H_1 + R_1)^{-1}(y_1 - H_1^T x_{1|0}) \\ P_{1|1} &= P_{1|0} - P_{1|0}H_1(H_1^T P_{1|0}H_1 + R_1)^{-1}H_1^T P_{1|0}\end{aligned}$$

4. Para realizar a predição de x_2 , pode-se aplicar o novamente o passo 2:

$$\begin{aligned}\hat{x}_{2|1} &= E[x_2|Y_1] = F_1\hat{x}_{1|1} + G_1u_1 \\ P_{2|1} &= F_1P_{1|1}F_1^T + Q_1\end{aligned}$$

5. E o passo 3 pode ser usado para atualizar a estimativa de x_2 , usando a medida y_2 :

$$\begin{aligned}\hat{x}_{2|2} &= \hat{x}_{2|1} + P_{2|1}H_2(H_2^T P_{2|1}H_2 + R_2)^{-1}(y_2 - H_2^T \hat{x}_{2|1}) \\ P_{2|2} &= P_{2|1} - P_{2|1}H_2(H_2^T P_{2|1}H_2 + R_2)^{-1}H_2^T P_{2|1}\end{aligned}$$

6. Mais genericamente, a repetição do passo 2 produzirá a parte da predição, que consiste nas estimativas a priori - nas quais se estima o estado e a covariância no tempo k a partir das estimativas em $k - 1$ e da própria equação do estado (4.1):

$$\begin{aligned}\hat{x}_{k|k-1} &= F_{k-1}\hat{x}_{k-1|k-1} + G_{k-1}u_{k-1} \\ P_{k|k-1} &= F_{k-1}P_{k-1|k-1}F_{k-1}^T + Q_{k-1}\end{aligned}$$

E a repetição do passo 3 será responsável pelas estimativas a posteriori, nas quais atualiza-se as estimativas no tempo k , usando a medida no tempo y_k :

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + P_{k|k-1}H_k(H_k^T P_{k|k-1}H_k + R_k)^{-1}(y_k - H_k^T \hat{x}_{k|k-1}) \\ P_{k|k} &= P_{k|k-1} - P_{k|k-1}H_k(H_k^T P_{k|k-1}H_k + R_k)^{-1}H_k^T P_{k|k-1}\end{aligned}$$

Note que, usando a equação (4.18), do ganho de Kalman, pode-se reescrever $\hat{x}_{k|k}$ e $P_{k|k}$ da seguinte forma:

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + P_{k|k-1}H_k(H_k^T P_{k|k-1}H_k + R_k)^{-1}(y_k - H_k^T \hat{x}_{k|k-1}) = \\ &= \hat{x}_{k|k-1} + K_k(y_k - H_k^T \hat{x}_{k|k-1}) \\ P_{k|k} &= P_{k|k-1} - P_{k|k-1}H_k(H_k^T P_{k|k-1}H_k + R_k)^{-1}H_k^T P_{k|k-1} = \\ &= P_{k|k-1} - K_k H_k^T P_{k|k-1} = (I - K_k H_k^T)P_{k|k-1}\end{aligned}$$

que condizem com as equações apresentadas na seção 4.3.1.

4.4.1 Exemplos

Para ilustrar o comportamento do Filtro em diferentes situações, serão apresentados cinco exemplos: três nos quais a condição inicial é a correta e sabe-se disso (ou seja, assume-se que a covariância P_0 é nula), mas se varia a matriz e o valor da covariância Q e R , associados aos ruídos do sistema e das medições, respectivamente; um no qual a condição inicial é incorreta e também sabe-se disso (ou seja, a covariância $P_0 \neq 0$); e um terceiro no qual a condição inicial é incorreta, mas acredita-se que esteja correta ($P_0 = 0$).

Nesses exemplos, com o intuito de simplificar a visualização, o modelo utilizado para simulação terá apenas dois estados (S_1 e S_2) e será descrito pelo seguinte sistema, em tempo discreto:

$$\begin{cases} S_{1_{k+1}} &= S_{1_k} + h(-0,9S_{1_k} + u_k) \\ S_{2_{k+1}} &= S_{2_k} + h(0,9S_{1_k} - 0,8S_{2_k}), \end{cases}$$

onde u_k representa a entrada do sistema e h representa o tamanho do passo na discretização. Em todos os exemplos, foi utilizado $h = 1/10$. Além disso, a condição inicial usada no modelo foi $x_0 = \begin{bmatrix} S_1 & S_2 \end{bmatrix}^T = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$.

As observações serão dadas por

$$q_k = 0,8S_{2_k}$$

Na implementação do filtro, os valores usados para o cálculo do erro e também para o gráfico foram os de $\hat{x}_{k|k}$ - ou seja, a estimativa de x após a atualização. Dessa forma, não ficará explícito o \bar{x}_0 (estado inicial fornecido ao filtro) nas figuras de cada exemplo, pois este equivale às estimativas do tipo $\hat{x}_{k|k-1}$ (note que a inicialização do filtro é dada por $\bar{x}_0 = \hat{x}_{0|-1}$).

Para o erro da estimação E_i de cada um dos estados, em uma simulação, tomou-se a raiz do erro médio quadrático, ou seja:

$$E_i = \sqrt{\frac{\sum_{k=1}^N (x_k - \hat{x}_{k|k})^2}{N}},$$

sendo $N = 100$ amostras, x_k o valor real de x e $\hat{x}_{k|k}$ o valor estimado de x , no tempo k . Como foram feitas 100 simulações para cada exemplo, o erro médio de cada estado foi obtido através da média aritmética dos erros E_i .

4.4.1.1 Condição inicial correta ($\bar{x}_0 = x_0$) e covariância nula ($P_0 = 0$)

Dentro dessa situação, serão analisados três casos:

- *Matrizes Q e R coerentes com os ruídos adicionados*

É possível perceber na figura (4.1) que, quando a condição inicial dada ao filtro é a mesma dada ao modelo e a matriz de covariância P_0 é nula - o que representa confiança na estimativa \bar{x}_0 do estado inicial, o filtro consegue uma boa estimativa dos estados do sistema desde o início da simulação. Obteve-se um erro médio de 0,0429 para S_1 e de 0,0178 para S_2 .

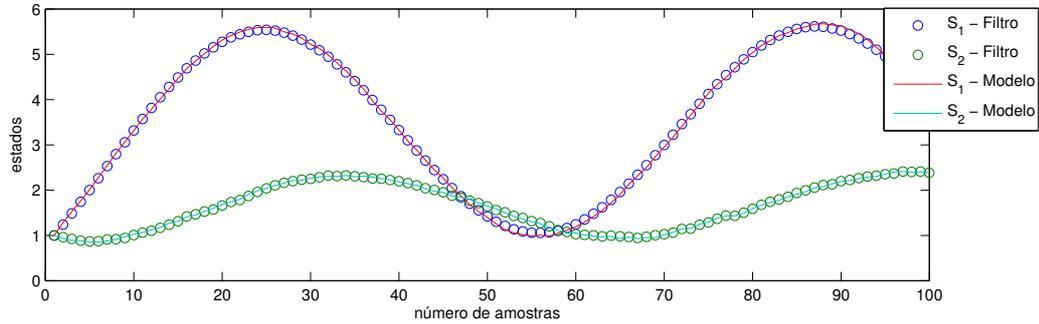


Figura 4.1: Simulação com $\bar{x}_0 = x_0$

- *Matriz Q manipulada*

Nesses exemplos, a matriz Q foi multiplicada por $1/100$, $1/10000$ e por 100 , o que influencia na covariância do ruído adicionado ao sistema. Quando multiplica-se Q por um número menor do que 1, admite-se que o ruído adicionado é menor; quando Q é multiplicada por um valor maior do que 1, assume-se que o ruído é maior.

Os gráficos desses exemplos não apresentaram muita diferença em relação ao da figura (4.1), e por isso foram omitidos. Os erros médios obtidos constam na tabela (4.1). É possível notar que, comparados ao caso anterior, não há grande variação em relação a S_1 , com mudança na terceira casa decimal. Em relação a S_2 , a mudança aparece na segunda casa decimal e o menor erro para esse estado apareceu quando Q foi multiplicado por 100.

Estado	Erro médio		
	$Q/100$	$Q/10000$	$Q \times 100$
S_1	0,0470	0,0466	0,0463
S_2	0,0412	0,0485	0,0246

Tabela 4.1: Erros médios com $\bar{x}_0 = x_0$ e matriz Q manipulada

- *Matriz R manipulada*

Quando o valor de R é manipulado (λR), admite-se que o ruído adi-

cionado à saída do modelo é maior ou menor, quando $\lambda > 1$ e $\lambda < 1$, respectivamente. Nesses exemplos, assim como no caso anterior, foi usado $\lambda = 1/100$, $\lambda = 1/10000$ e $\lambda = 100$. Novamente não há muita variação no erro médio de S_1 , quando comparado ao primeiro caso. Em relação a S_2 , há mudança a partir da segunda casa decimal. Os erros médios constam na tabela 4.2.

Estado	Erro médio		
	$R/100$	$R/10000$	$R \times 100$
S_1	0,0470	0,0432	0,0453
S_2	0,0238	0,0251	0,0413

Tabela 4.2: Erros médios com $\bar{x}_0 = x_0$ e valor de R manipulado

4.4.1.2 Condição inicial incorreta ($\bar{x}_0 = [2 \ 2]^T$) e covariância não-nula ($P_0 \neq 0$)

Mesmo quando a condição inicial dada ao filtro é diferente da condição inicial usada no modelo, se a matriz P_0 não for nula (nesse exemplo, utilizou-se $P_0 = I_2$, ou seja, a matriz identidade 2×2), em poucas amostras o filtro converge para a solução do modelo, como ilustra a figura 4.2. Os erros médios foram de 0,1410 para S_1 e 0,0183 para S_2 .

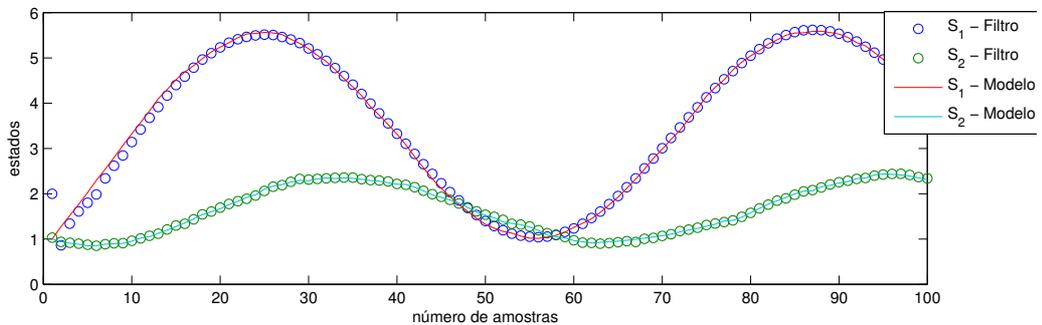


Figura 4.2: Simulação com $\bar{x}_0 \neq x_0$ e $P_0 \neq 0$

4.4.1.3 Condição inicial incorreta ($\bar{x}_0 = [2 \ 2]^T$) e covariância nula ($P_0 = 0$)

Quando $\bar{x}_0 \neq x_0$ e a matriz P_0 for nula (o que significa que acredita-se que a condição inicial está correta), a figura 4.3 mostra que o filtro acaba convergindo para a solução do modelo, porém demora mais. Como esperado, os erros médios apresentados foram maiores do que os do caso anterior (em que $P_0 \neq 0$). Obteve-se, para S_1 , um erro médio de 0.2373, e para S_2 , de 0.1221.

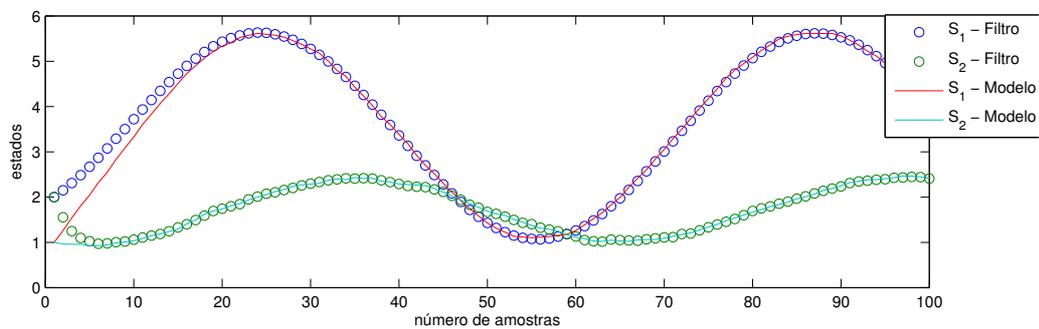


Figura 4.3: Simulação com $\bar{x}_0 \neq x_0$ e $P_0 = 0$

4.5 Filtro de Kalman Estendido

O filtro de Kalman aplica-se diretamente a sistemas lineares. Quando aplicado a um sistema não-linear, a filtragem tende a ser mais complexa e menos precisa. A ideia-chave do Filtro de Kalman Estendido (EKF) é linearizar o sistema, através de séries de Taylor, e utilizar as técnicas de estimação linear do Filtro de Kalman convencional [2, 19].

Será utilizado, então, o seguinte modelo [2]:

$$x_{k+1} = f_k(x_k, u_k, w_k) \quad (4.19)$$

$$y_k = h_k(x_k, v_k) \quad (4.20)$$

sendo, em geral, f_k e h_k não-lineares; x_k é o estado, u_k é a entrada (determinística) e w_k é o ruído do sistema, no tempo k ; y_k é a saída e v_k é o ruído relacionado

a ela. As seguintes propriedades permanecem: os ruídos w_k e v_k têm média zero, são gaussianos e brancos. Além disso, v_k , w_k e x_0 são mutualmente independentes, $E[w_k w_k^T] = Q_k$ e $E[v_k v_k^T] = R_k$.

As funções não-lineares e diferenciáveis f_k e h_k podem ser expandidas por série de Taylor em torno das estimativas $\hat{x}_{k|k}$ e $\hat{x}_{k|k-1}$ e de $w_k = 0$ e $v_k = 0$ como segue [19]:

$$\begin{aligned} f_k(x_k, u_k, w_k) &= f_k(\hat{x}_{k|k}, u_k, 0) + \left. \frac{\partial f_k}{\partial x} \right|_{x=\hat{x}_{k|k}} (x_k - \hat{x}_{k|k}) + \left. \frac{\partial f_k}{\partial w} \right|_{w=\hat{x}_{k|k}} w_k + \dots = \\ &= f_k(\hat{x}_{k|k}, u_k, 0) + F_k(x_k - \hat{x}_{k|k}) + L_k w_k \dots \\ h_k(x_k, v_k) &= h_k(\hat{x}_{k|k-1}, 0) + \left. \frac{\partial h_k}{\partial x} \right|_{x=\hat{x}_{k|k-1}} (x_k - \hat{x}_{k|k-1}) + \left. \frac{\partial h_k}{\partial v} \right|_{v=\hat{x}_{k|k-1}} v_k + \dots = \\ &= h_k(\hat{x}_{k|k-1}, 0) + H_k^T(x_k - \hat{x}_{k|k-1}) + M_k v_k + \dots \end{aligned}$$

Desconsiderando os termos de ordem mais alta e assumindo que $\hat{x}_{k|k}$ e $\hat{x}_{k|k-1}$ são conhecidos, pode-se aproximar o modelo dado pelas equações (4.19) e (4.20) por

$$x_{k+1} = f_k(\hat{x}_{k|k}, u_k, 0) + F_k x_k - F_k \hat{x}_{k|k} + L_k w_k = F_k x_k + \tilde{u}_k + \tilde{w}_k \quad (4.21)$$

$$y_k = h_k(\hat{x}_{k|k-1}, 0) + H_k^T x_k - H_k^T \hat{x}_{k|k-1} + M_k v_k = H_k^T x_k + z_k + \tilde{v}_k \quad (4.22)$$

com

$$\begin{aligned} \tilde{u}_k &= f_k(\hat{x}_{k|k}, u_k, 0) - F_k \hat{x}_{k|k} \\ z_k &= h_k(\hat{x}_{k|k-1}, 0) - H_k^T \hat{x}_{k|k-1} \end{aligned}$$

Os ruídos $\tilde{w}_k = L_k w_k$ e $\tilde{v}_k = M_k v_k$ têm valor esperado zero e covariância dadas, respectivamente, por:

$$\begin{aligned} E[(\tilde{w}_k - E[\tilde{w}_k])(\tilde{w}_k - E[\tilde{w}_k])^T] &= E[\tilde{w}_k \tilde{w}_k^T] = E[L_k w_k (L_k w_k)^T] = L_k Q_k L_k^T \\ E[(\tilde{v}_k - E[\tilde{v}_k])(\tilde{v}_k - E[\tilde{v}_k])^T] &= E[\tilde{v}_k \tilde{v}_k^T] = E[M_k v_k (M_k v_k)^T] = M_k R_k M_k^T \end{aligned}$$

Desta forma, com base nas equações do Filtro de Kalman convencional, tem-se as seguintes equações para o Filtro de Kalman Estendido [2, 19]:

$$\begin{aligned}\hat{x}_{k|k-1} &= f_k(\hat{x}_{k-1|k-1}, u_k, 0) \\ P_{k|k-1} &= F_{k-1}P_{k-1|k-1}F_{k-1}^T + L_{k-1}Q_{k-1}L_{k-1}^T \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(y_k - (H_k^T\hat{x}_{k|k-1} + z_k)) = \hat{x}_{k|k-1} + K_k(y_k - h_k(\hat{x}_{k|k-1}, 0)) \\ P_{k|k} &= (I - K_kH_k)P_{k|k-1}\end{aligned}$$

Sendo K_k a matriz ganho de Kalman, dada por:

$$K_k = P_{k|k-1}H_k(H_k^T P_{k|k-1}H_k + M_kR_kM_k^T)^{-1}$$

A inicialização é dada por $P_{0|-1} = P_0$ e $\hat{x}_{0|-1} = \bar{x}_0$.

4.5.1 Resultados

Nessa seção, o Filtro de Kalman Estendido será aplicado em um modelo que descreve um biorreator anaeróbico, operado no modo semibatelada.

A sua configuração é baseada no modelo geral de balanço de massa (2.5), aplicável a qualquer tipo de biorreator anaeróbico, apresentado no capítulo 2. No entanto, os exemplos serão aplicados a um do tipo semibatelada, alimentado com sacarose a cada 7 dias, numa proporção de 1% do seu volume total - cuja validação e identificação dos parâmetros foram feitas em [11]. Para representar esse caso específico, é necessário fazer alguns ajustes no modelo geral.

O parâmetro D , que representa a taxa de diluição do sistema e faz sentido apenas para biorreatores de alimentação contínua, é considerado nulo. Contudo, isso leva ao crescimento ininterrupto da população de bactérias. Para contornar essa situação, foram acrescentados termos de mortalidade - que talvez estivessem implícitos na taxa de diluição - c_1 e c_2 , associados às bactérias acidogênicas (x_1) e

metanogênicas (x_2), respectivamente. Além disso, a taxa de fluxo do gás carbônico (2.4), a qual aparece na sexta linha do modelo geral em (2.5), será considerada como $q_C = k_c C$, onde k_c denota uma constante de consumo de gás carbônico pelo meio. Desta forma, assume-se que o sistema não depende da alcalinidade, o que acarreta na desconsideração do estado Z no modelo geral.

Desta forma, o modelo usado nos exemplos será descrito pelo seguinte sistema [11], em tempo discreto:

$$\begin{cases} x_{1_{k+1}} &= x_{1_k} + [\nu_1(S_{1_k}) - c_1]x_{1_k} h \\ x_{2_{k+1}} &= x_{2_k} + [\nu_2(S_{2_k}) - c_2]x_{2_k} h \\ S_{1_{k+1}} &= S_{1_k} - k_1 \nu_1(S_{1_k})x_{1_k} h + S_1^{in} \\ S_{2_{k+1}} &= S_{2_k} + (k_2 \nu_1(S_{1_k})x_{1_k} - k_3 \nu_2(S_{2_k})x_{2_k}) h \\ C_{k+1} &= C_k + (-k_c C_k + k_4 \nu_1(S_{1_k})x_{1_k} + k_5 \nu_2(S_{2_k})x_{2_k}) h \end{cases}$$

onde h representa o tamanho do passo na discretização. Em todos os exemplos, foi utilizado $h = 1/40$, o que representa, na prática, que foram coletadas 40 amostras por dia. As simulações foram feitas para um total de 1400 amostras, ou seja, 35 dias. A entrada de 1% de sacarose é representada por S_1^{in} e é dada por um vetor que acresce 1 a S_1 a cada sete dias simulados, sendo que a primeira inserção é no início do processo.

Para representar o crescimento bacteriano de x_1 e x_2 , dado no sistema por $\nu_1(S_{1_k})$ e $\nu_2(S_{2_k})$, respectivamente, utilizou-se o modelo proposto por Haldane (2.9) - o qual admite parâmetros de inibição K_{I1} e K_{I2} associados a S_1 e S_2 .

Inicialmente, para o modelo, assumiu-se que as concentrações das bactérias x_1 e x_2 têm valor adimensional 1, o que denota 100%. Os demais estados foram inicializados como zero. Desta forma, a condição inicial usada foi

$$x_0 = \begin{bmatrix} x_1 & x_2 & S_1 & S_2 & C \end{bmatrix}^T = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T.$$

As observações serão dadas por:

$$q_k = k_6 \nu_2(S_{2_k}) x_{2_k} + k_c C_k$$

Foi adicionado um ruído na saída de gás, como ilustra a figura 4.4, visto que as medidas experimentais são suscetíveis a erros. Esse ruído foi gerado de uma distribuição normal, de valor esperado 0 e desvio padrão 1/30.

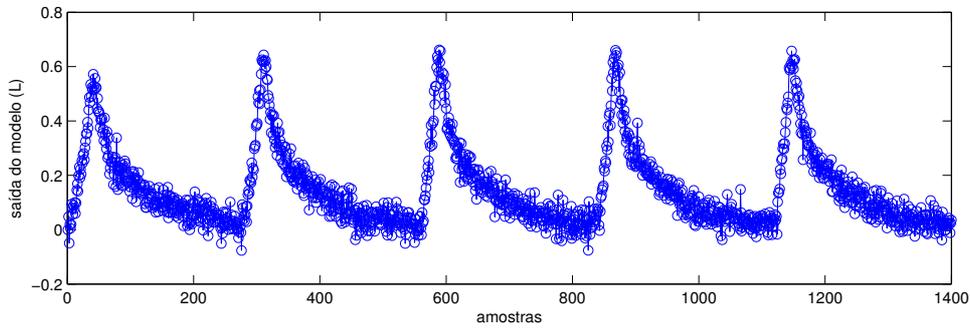


Figura 4.4: Saída de gás com ruído adicionado

O vetor de parâmetros p usado será [11]:

$$\begin{aligned} p &= \left[\mu_{m1} \quad K_{S1} \quad K_{I1} \quad \mu_{m2} \quad K_{S2} \quad K_{I2} \quad c_1 \quad c_2 \quad k_1 \quad k_2 \quad k_3 \quad k_6 \quad k_4 \quad k_c \quad k_5 \right] = \\ &= \left[172,036 \quad 21,3658 \quad 124,679 \quad 39,4958 \quad 735,921 \quad 0,0836 \quad 0,1328 \right. \\ &\quad \left. 0,009481 \quad 0,34878 \quad 0,0132 \quad 100 \quad 497,8528 \quad 0,29771 \quad 0,4879 \quad 19,7825 \right] \end{aligned}$$

O estado inicial dado ao filtro \bar{x}_0 não ficará explícito nas figuras de cada exemplo, pois os valores usados para o gráfico (e também para o cálculo do erro médio) foram os de $\hat{x}_{k|k}$, ou seja, a estimativa de x após a atualização.

Para calcular o erro da estimação $E_j, j = 1, \dots, 5$ de cada um dos estados, em uma simulação, tomou-se a raiz do erro médio quadrático, ou seja:

$$E_j = \sqrt{\frac{\sum_{k=1}^N (x_k - \hat{x}_{k|k})^2}{N}},$$

sendo $N = 1400$ amostras, x_k o valor real de x no tempo k e $\hat{x}_{k|k}$ o valor estimado de x no tempo k . Como foram feitas 100 simulações para cada exemplo, o erro médio de cada estado foi obtido através da média aritmética dos erros E_j .

Algumas figuras serão omitidas, pois, de um modo geral, acabam sendo muito parecidas. Como visto no capítulo 2, o estado S_2 (VFA) tem papel importante na bioquímica da digestão e pode acarretar a interrupção do processo pela redução do pH do meio e conseqüente inibição da atividade das bactérias metanogênicas. Como o estado assume valores muito pequenos, o que dificulta a visualização na figura do sistema completo (com os cinco estados), em alguns casos será conveniente mostrá-lo em uma figura a parte, junto com a sua estimativa.

Seguindo o mesmo padrão do Filtro de Kalman, cinco exemplos serão apresentados aqui: três nos quais a condição inicial é a correta e sabe-se disso, mas se varia as matrizes de covariância dos ruídos do sistema e das medições de gás, Q e R , respectivamente; um no qual a condição inicial é incorreta e também sabe-se disso; e um terceiro no qual a condição inicial é incorreta, mas acredita-se que esteja correta.

4.5.1.1 *Condição inicial correta ($\bar{x}_0 = x_0$) e covariância nula ($P_0 = 0$)*

Da mesma forma como foi feito nos exemplos de Filtro de Kalman, a fim de verificar o peso da escolha das matrizes Q e R na qualidade das estimativas, três casos serão analisados aqui:

- *Matrizes Q e R coerentes com os ruídos adicionados*

A figura 4.6 mostra que, ao ser dada ao filtro a mesma condição inicial dada ao modelo, a estimativa dos estados fica bem satisfatória durante toda a simulação. A figura 4.5 mostra com maior detalhe a estimativa de S_2 . Os erros médios obtidos nesse caso constam na tabela 4.3, e foram muito próximos de zero. Por ser o caso no qual espera-se os

melhores resultados do filtro, os erros aqui apresentados acabam sendo um parâmetro de comparação para os outros resultados.

Estado	Erro médio
x_1	$2,1729 \times 10^{-5}$
x_2	$2,9795 \times 10^{-7}$
S_1	$2,1138 \times 10^{-6}$
S_2	$6,3566 \times 10^{-8}$
C	$2,6083 \times 10^{-6}$

Tabela 4.3: Erros médios com $\bar{x}_0 = x_0$

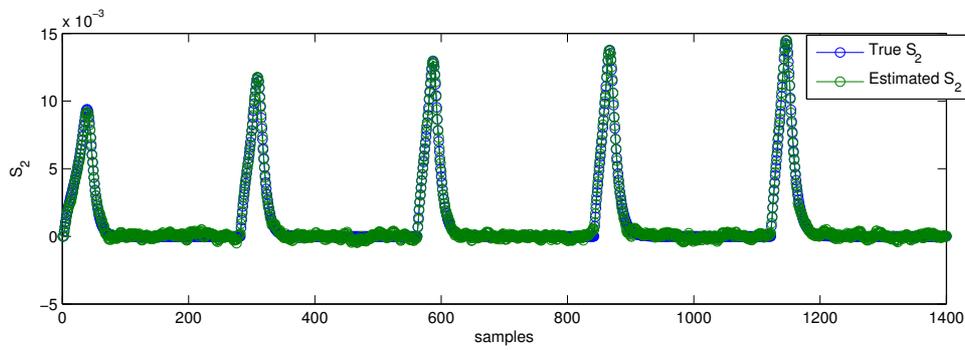


Figura 4.5: Simulação do S_2 com $\bar{x}_0 = x_0$

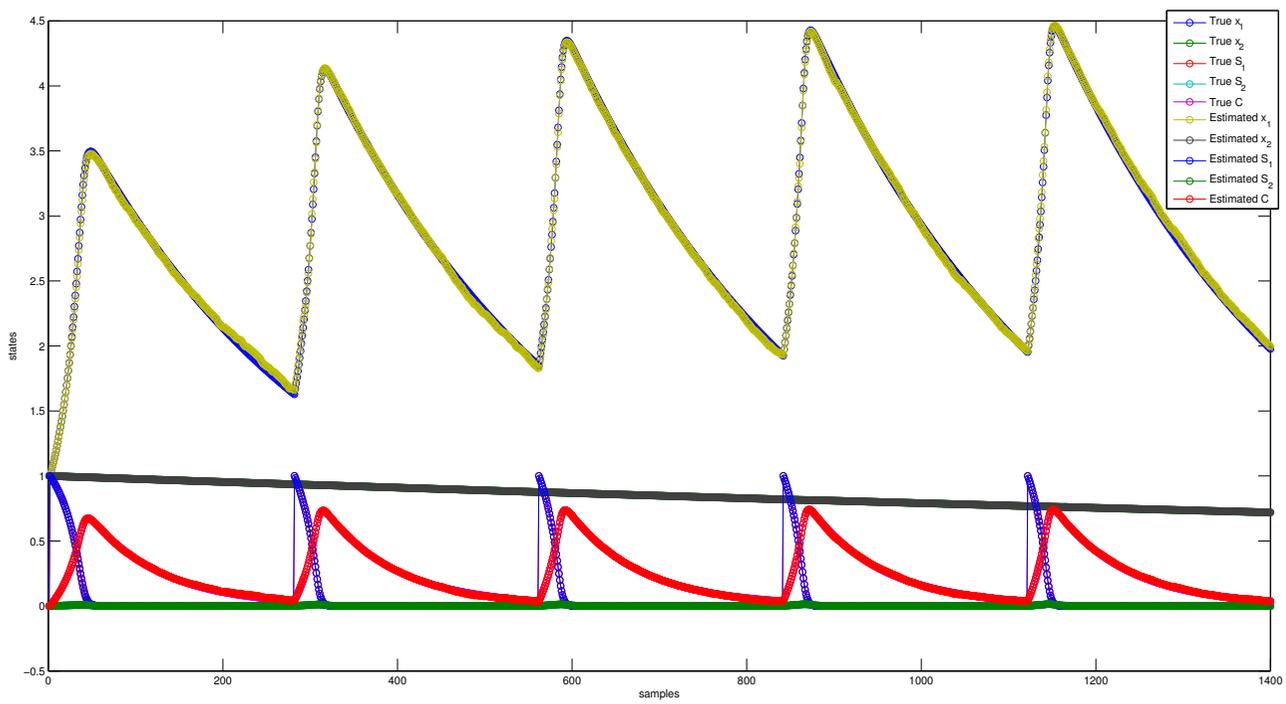


Figura 4.6: Simulação com $\bar{x}_0 = x_0$

- *Matriz Q manipulada*

A matriz Q refere-se à covariância do ruído adicionado aos estados do modelo. Ou seja, quanto maior o valor de Q , maior admite-se esse ruído e, conseqüentemente, menor a confiança no modelo utilizado. Nesses exemplos, utilizou-se $Q = I_5$, $Q = I_5/100$. Os erros médios obtidos estão apresentados na tabela 4.4. Se comparados com os erros da tabela 4.4, é possível perceber que, nesses dois exemplos, as estimativas ficaram piores. A figura 4.8 ilustra uma simulação com $Q = I_5/100$, na qual é possível observar que a estimativa do estado x_2 se afasta consideravelmente do x_2 do modelo, e a estimativa de x_1 também se afasta da solução do modelo em alguns intervalos. Nos demais estados, a estimativa também parece ficar pior com o passar do tempo. A figura 4.8 mostra com maior detalhe como o filtro começa a ficar insatisfatório depois de um tempo no estado S_2 .

Estado	Erro médio	
	$Q = I_5$	$Q = I_5/100$
x_1	2,3246	0,2512
x_2	0,7298	0,4477
S_1	2,7089	0,0304
S_2	0,0332	0,0049
C	0,1314	0,0354

Tabela 4.4: Erros médios com $\bar{x}_0 = x_0$ e matriz Q manipulada

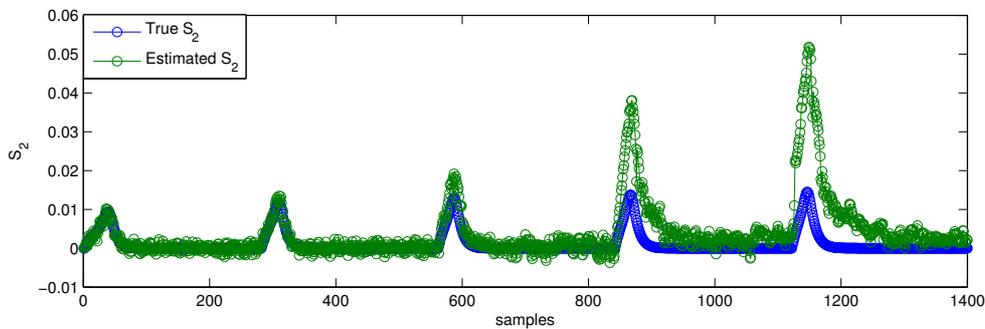


Figura 4.7: Simulação do S_2 com $\bar{x}_0 = x_0$ e $Q = I_5/100$

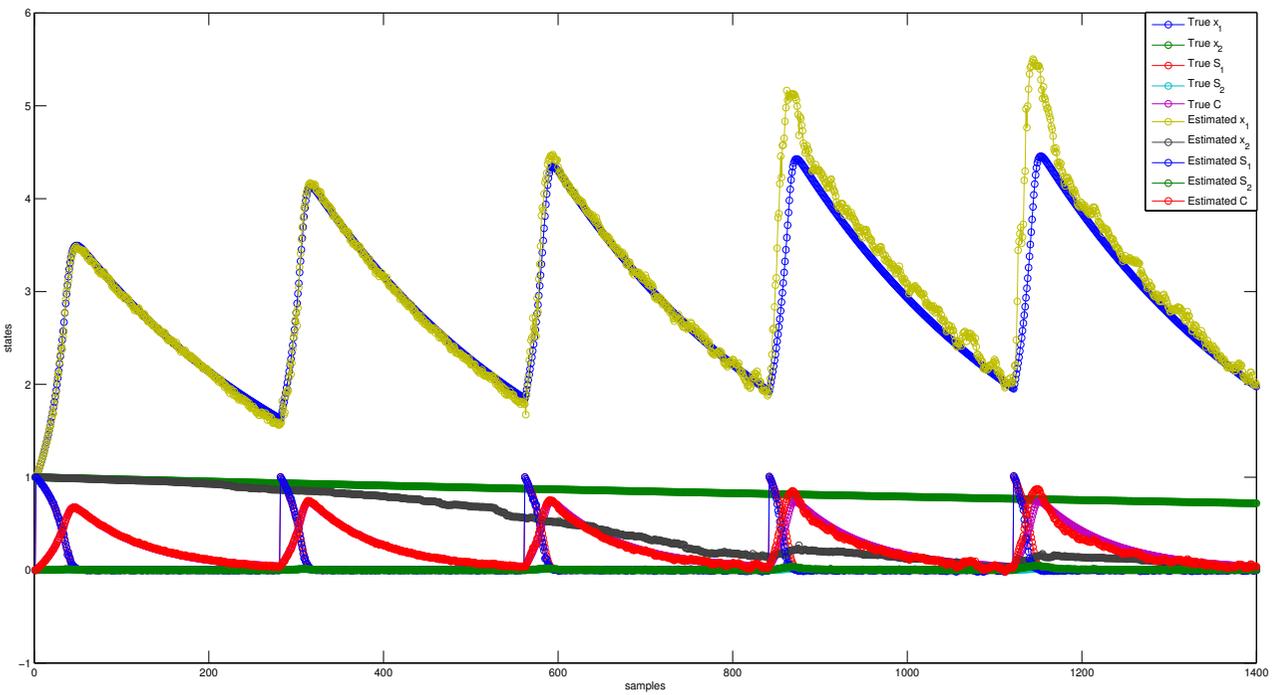


Figura 4.8: Simulação com $\bar{x}_0 = x_0$ e $Q = I_5/100$

- *Valor de R manipulado*

O valor de R representa a covariância do ruído na medida do gás. Assim como quando manipula-se a matriz Q , ao alterar-se o valor de R , admite-se que esse ruído é maior ou menor. A tabela 4.5 apresenta os erros médios obtidos ao multiplicar-se R por $1/100$, $1/10000$ e 100 . Os melhores resultados foram obtidos quando multiplicou-se R por 100 - mas não superaram o primeiro caso (tabela 4.3). Ao dividir-se o valor de R por 10000 , os erros médios ficam significativamente maiores. A figura 4.10 mostra que as estimativas ficam, de fato, insatisfatórias nesse caso. A figura 4.9 detalha a estimativa de S_2 e o quanto ela se afastou do S_2 do modelo.

Estado	Erro médio		
	$R/100$	$R/10000$	$R \times 100$
x_1	0,2505	2,3152	0,0020
x_2	0,4499	0,7299	$2,8291 \times 10^{-5}$
S_1	0,0301	2,6140	$1,9571 \times 10^{-4}$
S_2	0,0050	0,0328	$6,0615 \times 10^{-6}$
C	0,0354	0,1306	$2,4830 \times 10^{-4}$

Tabela 4.5: Erros médios com $\bar{x}_0 = x_0$ e valor de R modificado

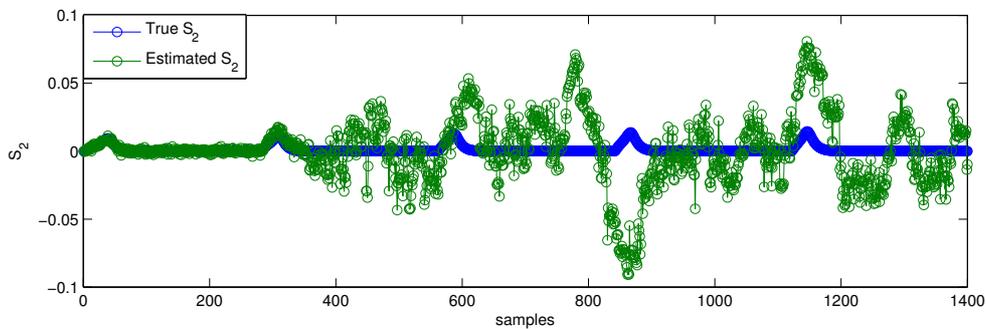


Figura 4.9: Simulação do S_2 com $\bar{x}_0 = x_0$ e $R/10000$

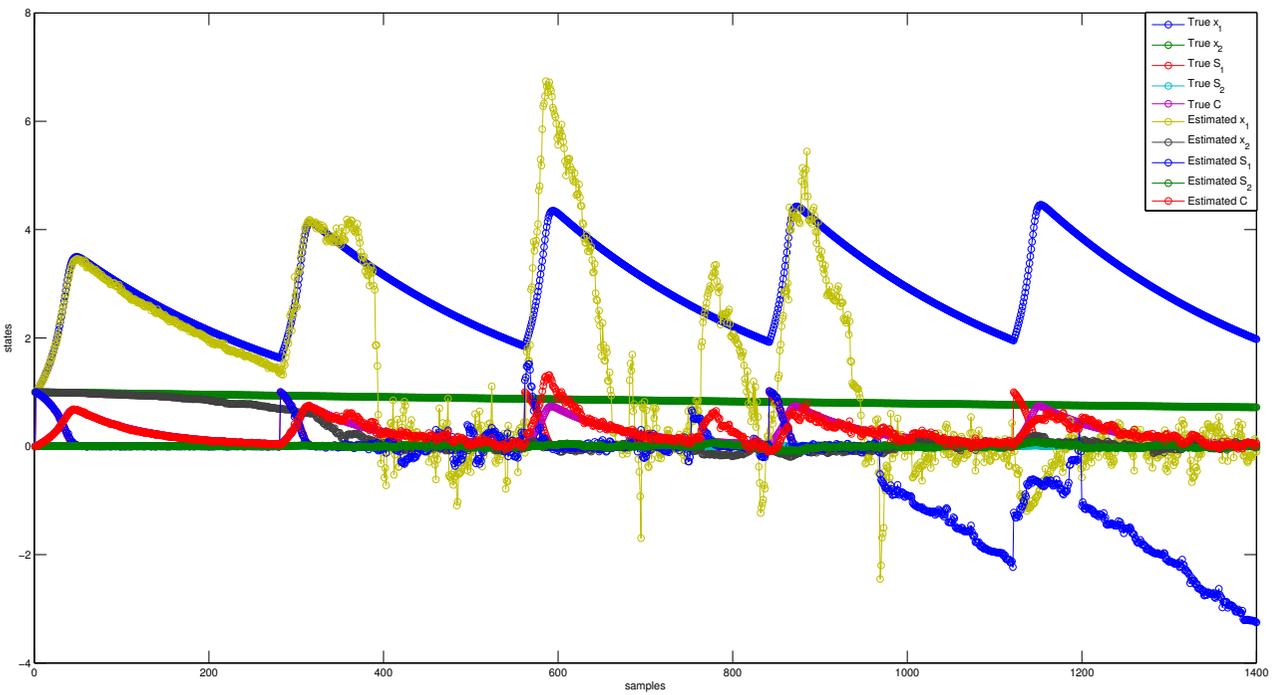


Figura 4.10: Simulação com $\bar{x}_0 = x_0$ e $R/10000$

4.5.1.2 *Condição inicial incorreta ($\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]^T$) e covariância não-nula ($P_0 \neq 0$)*

Se a $\bar{x}_0 \neq x_0$ e admite-se que a covariância da condição inicial $P_0 \neq 0$ (aqui, foi usado $P_0 = I_5$), o que indica que assume-se a existência de um erro na estimativa, a tendência é que o filtro convirja para a solução do modelo mais rapidamente. A figura 4.11 ilustra isso, no caso em que $P_0/100$ (exemplo que obteve os menores erros médios). O estado que mais demorou para se aproximar da solução do modelo foi x_2 - e, ainda assim, não ficou totalmente satisfatório ao fim dos 35 dias. A figura 4.13 mostra especificamente a simulação do estado S_2 e como a estimativa se aproxima do modelo. Já a figura 4.12 mostra um exemplo de simulação com $P_0 = I_5$. É possível perceber que, nesse caso, o filtro erra bastante nas primeiras estimativas, mas acaba convergindo para a solução do modelo.

De acordo com a tabela 4.6, o estado x_1 foi o que apresentou maior oscilação no erro médio: 0,1611 para $P_0/100$ e 2,0023 para $P_0 \times 100$. É possível observar também que não há uma generalização dos erros de caso para caso. Ou seja, quando manipula-se a matriz P_0 , os erros de alguns estados diminuem e outros aumentam.

Estado	Erro médio			
	$P_0/100$	$P_0/10000$	$P_0 \times 100$	P_0
x_1	0,1611	0,2542	2,4540	2,0023
x_2	0,1275	0,1734	0,3489	0,3929
S_1	0,0937	0,0213	1,4264	0,3318
S_2	$6,4658 \times 10^{-4}$	$7,5966 \times 10^{-4}$	0,2350	0,2265
C	0,0123	0,0220	0,0220	0,1829

Tabela 4.6: Erros médios com $\bar{x}_0 \neq x_0$ e $P_0 \neq 0$

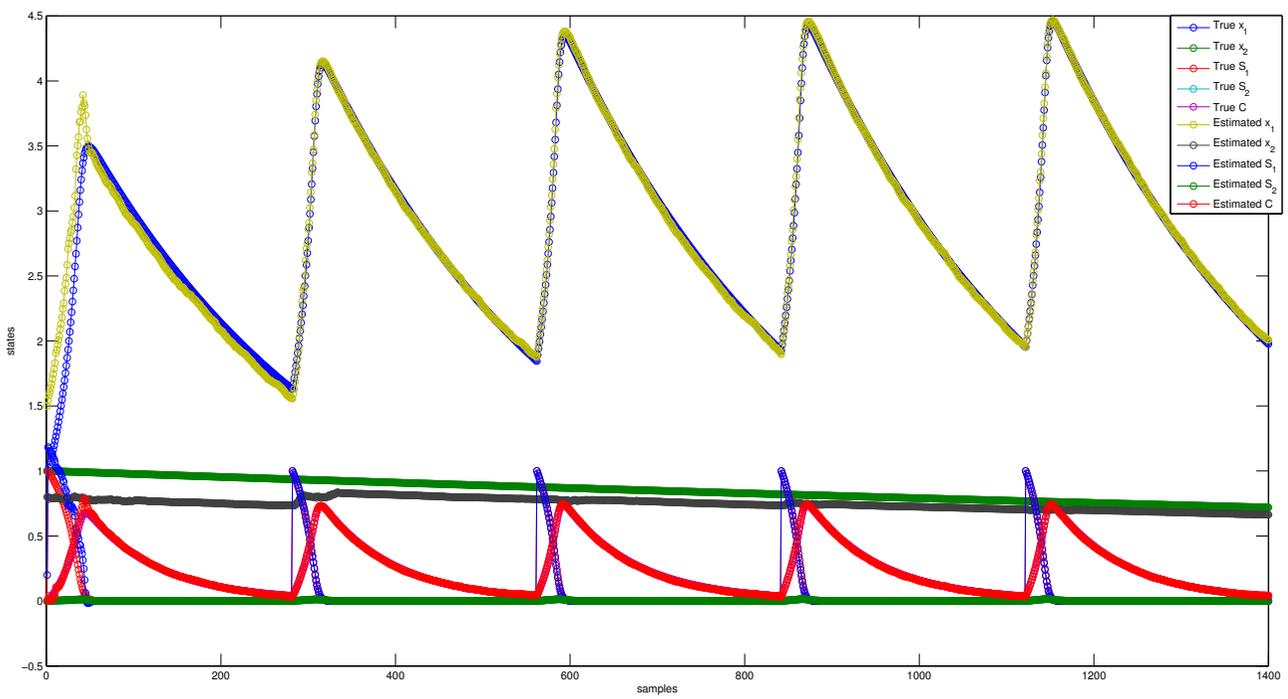


Figura 4.11: Simulação com $\bar{x}_0 \neq x_0$ e $P_0 = I_5/100$

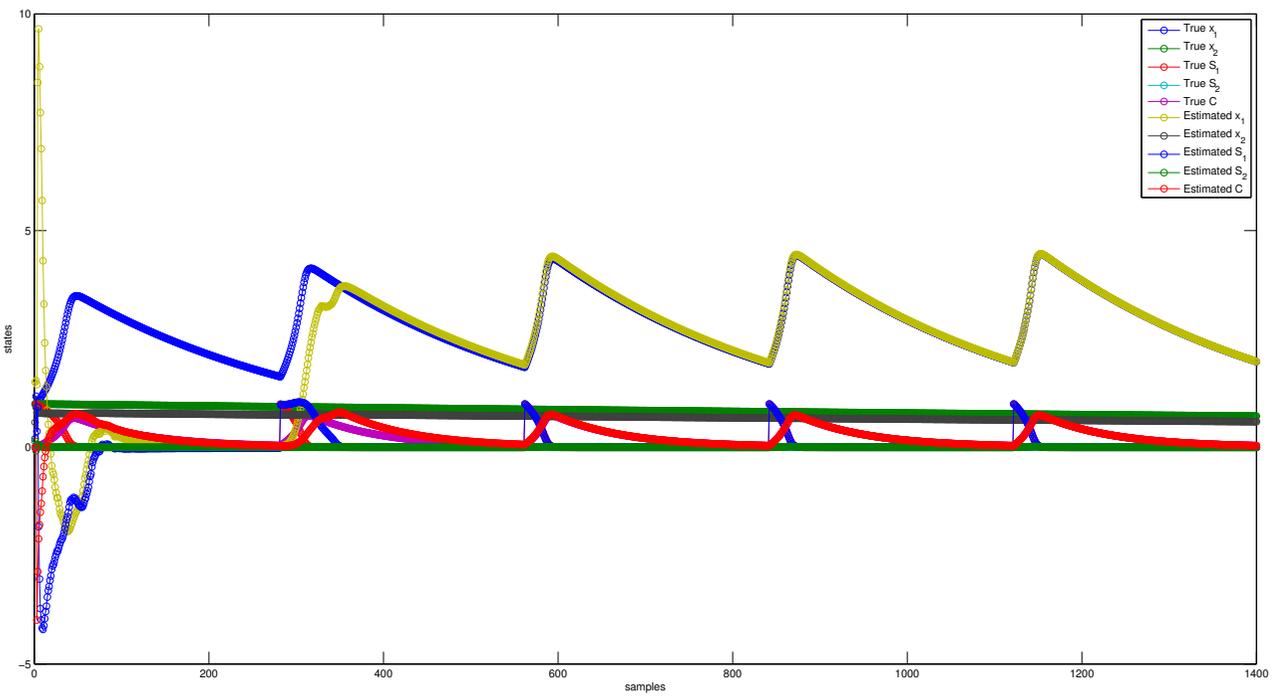


Figura 4.12: Simulação com $\bar{x}_0 \neq x_0$ e $P_0 = I_5/100$

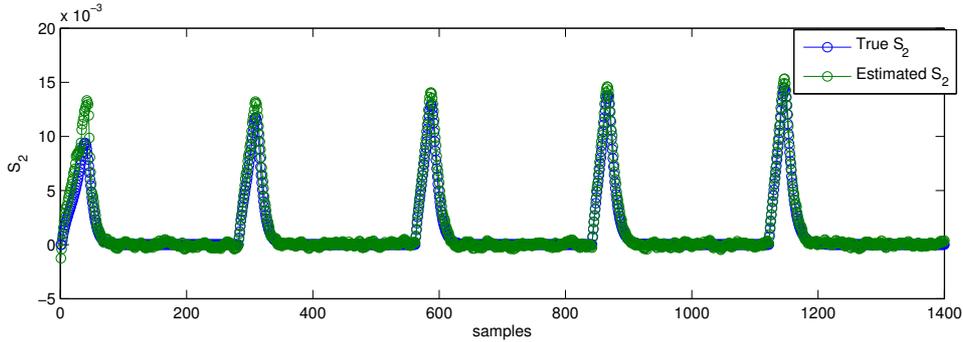


Figura 4.13: Simulação do S_2 com $\bar{x}_0 \neq x_0$ e $P_0 = I_5/100$

4.5.1.3 Condição inicial incorreta ($\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]^T$) e covariância nula ($P_0 = 0$)

Ao considerar a matriz da covariância da condição inicial como nula, na teoria, admite-se que a estimativa é exatamente igual ao estado inicial verdadeiro - o que não é o caso nesse exemplo. Os erros encontrados constam na tabela 4.7. Quando comparados à tabela 4.6, percebe-se que foram mais eficientes para $P_0 = I_5$ e $P_0 = I_5 \times 100$ (nesse, com exceção do estado C), e foram muito parecidos para $P_0 = I_5/100000$. Já o caso em que $P_0 = I_5/100$ apresentou erros médios menores para quatro estados, sendo menos eficaz apenas na estimação do estado S_1 .

Estado	Erro médio
x_1	0,2505
x_2	0,1698
S_1	0,0206
S_2	$9,4013 \times 10^{-4}$
C	0,0249

Tabela 4.7: Erros médios com $\bar{x}_0 \neq x_0$ e $P_0 = 0$

A figura 4.14 mostra que as estimativas demoram um pouco mais para convergir para a solução do modelo - principalmente a do estado x_1 , quando comparados à figura 4.11. A estimativa de x_2 parece ir se aproximando do modelo, embora muito lentamente. A estimativa de S_2 está ilustrada na figura 4.15.

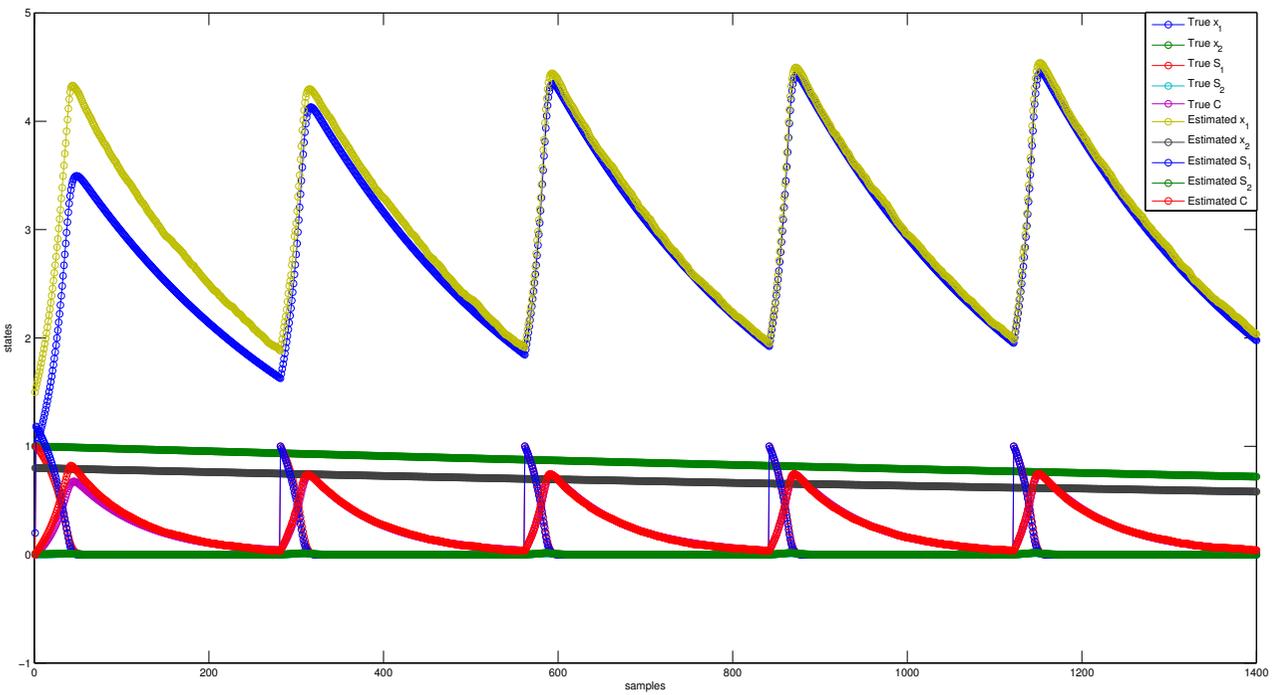


Figura 4.14: Simulação com $\bar{x}_0 \neq x_0$ e $P_0 = 0$

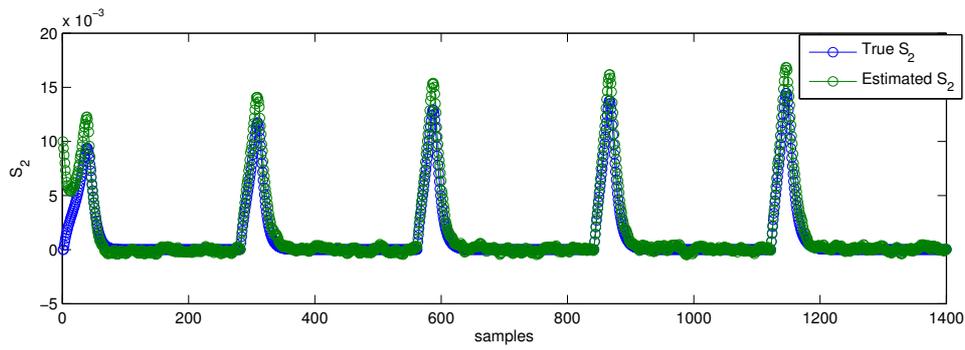


Figura 4.15: Simulação do S_2 com $\bar{x}_0 \neq x_0$ e $P_0 = 0$

Esses resultados serão discutidos no capítulo 6, juntamente com os resultados que serão apresentados no capítulo 5.

5 FILTRO DE PARTÍCULAS

O Filtro de Partículas foi introduzido como uma aproximação numérica para os problemas de estimação de estados não-lineares, principalmente nos casos em que o Filtro de Kalman Estendido não apresenta bons resultados. Os filtros de partículas utilizam o método chamado Sequential Importance Sampling (SIS), que é uma técnica de implementação recursiva por simulações de Monte Carlo, para aproximar uma determinada distribuição do estado, sem restrições sobre o modelo. O modelo pode ser não-linear e os estados iniciais e sinais de ruído podem ter qualquer distribuição, desde que conhecidas [19].

Na literatura, são encontradas diversas variações de filtros de partículas. Uma delas, a qual será usada nos exemplos na seção 5.3.2, é o Filtro de Partículas com Reamostragem ou Sequential Importance Resampling (SIR). A grande diferença entre os métodos SIR e SIS está exatamente na reamostragem das partículas a cada passo, a fim de evitar um fator chamado *empobrecimento da amostra*, que acaba fazendo com que a amostra inicial se degenere e restem poucas ou apenas uma partícula com peso significativo.

Retomando o problema apresentado na seção (3.4), considere o seguinte sistema não-linear:

$$x_{k+1} = f_k(x_k, w_k) \quad (5.1)$$

$$y_k = h_k(x_k, v_k) \quad (5.2)$$

onde x_k é o estado e w_k é o ruído do processo; y_k são as medições e v_k é o ruído associado a elas; $f_k(\cdot)$ e $h_k(\cdot)$ são funções não-lineares. $\{w_k\}$ e $\{v_k\}$ são ruídos brancos, independentes, e com pdf's conhecidas.

A pdf da condição $p(x_0)$ deve ser conhecida e a estimação é inicializada com $p(x_0|Y_0) = p(x_0)$ e o objetivo é estimar a pdf $p(x_k|Y_k)$. A solução desse pro-

blema, pela abordagem do método de Filtro de Partículas será desenvolvida neste capítulo. A ideia principal é representar a pdf a posteriori $p(x_k|Y_k)$ por um conjunto de amostras aleatórias com pesos associados e calcular estimativas com base nessas amostras e pesos. À medida que o número de amostras se torna muito grande, o filtro SIS se aproxima da estimativa bayesiana ideal [4].

5.1 Importance Sampling (IS)

De maneira simples, o princípio da amostragem de importância (IS) é calcular o valor esperado de uma determinada pdf alvo através da aproximação por uma média ponderada de sorteios aleatórios de outra distribuição [23]. A figura 5.1 ilustra um exemplo de distribuição difícil de ser descrita analiticamente. Porém, é possível aproximá-la usando o método de Monte Carlo, que consiste em gerar um conjunto $\{x^i, \omega^i\}_{i=1}^L$, onde x^i representa o estado e ω^i representa o peso associado a ele. Desta forma, a pdf pode ser aproximada por:

$$p(x|y) \approx \sum_{i=1}^L \omega^i \delta(x - x^i),$$

sendo $\delta(x - x^i)$ a função delta de Dirac.

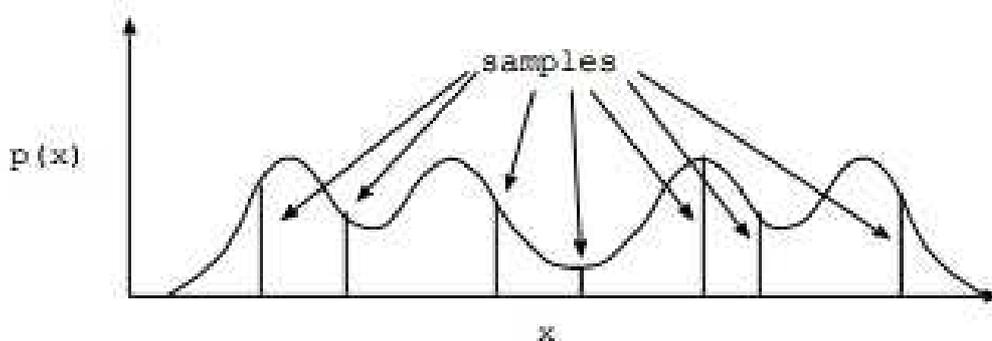


Figura 5.1: Aproximação de uma distribuição não-gaussiana pelo método de Monte Carlo [17]

Em um problema de filtragem, a principal distribuição de interesse é a probabilidade de estado dada uma medida, denotada como $p(x|y)$. Esta é a pdf que deseja-se aproximar usando o conjunto $\{x^i, \omega^i\}_{i=1}^L$. As amostras x^i são geradas aleatoriamente, como se fossem um palpite no estado atual. Porém, o peso w^i a ser atribuído a cada amostra é uma incógnita. Se a pdf $p(x|y)$ fosse conhecida, então o peso ω^i poderia ser calculado como [17]:

$$\omega^i = p(x^i|y),$$

ou seja: o peso é a probabilidade desse estado ser o real estado, dada a medida y . Porém, a pdf $p(x|y)$ não é conhecida.

Considere, então, o valor esperado condicional:

$$E_p[x] = \int xp(x|y)dx$$

A amostragem de importância usa uma identidade simples. Para qualquer densidade de probabilidade $q(x|y)$, pode-se reescrever [23]:

$$E_p[x] = \int x \frac{p(x|y)}{q(x|y)} q(x|y) dx = \int x \omega(x) q(x|y) dx = E_q[x\omega(x)] \quad (5.3)$$

e agora $E_q[\cdot]$ denota o valor esperado em relação a $q(x|y)$, que é chamada de proposta de distribuição (ou distribuição de amostragem), a qual deve ser conhecida - poderia ser uma distribuição gaussiana ou uniforme, por exemplo.

A equação (5.3) indica que amostrar x da pdf $p(x|y)$ é equivalente a amostrar $x\omega(x)$ da pdf $q(x|y)$.

Portanto, uma amostra de partículas independentes x^1, x^2, \dots, x^i , extraída de $q(x|y)$ pode ser usada pra estimar o valor esperado, como segue [23]:

$$E_p[X] \approx \frac{1}{L} \sum_{m=1}^L \omega(x^i) x^i,$$

onde os pesos ω^i são calculados como

$$\omega^i = \frac{p(x^i|y)}{q(x^i|y)}$$

Porém, embora a pdf $q(\cdot)$, por hipótese, seja conhecida, a pdf $p(\cdot)$ não é. A amostragem de importância sequencial (SIS) trata esse problema de forma iterativa, como será visto na próxima seção.

5.2 Amostragem de Importância Sequencial (SIS)

O Filtro de Partículas visa estimar todo o caminho $x_{0:k}$, a qual denotaremos por X_k , em vez de apenas o estado atual x_k . Isso significa que um conjunto $\{X_k^i\}_{i=1}^N$ de N trajetórias diferentes é gerado e avaliado [10].

Considere, então, o peso ω_k^i de uma amostra no instante k [17]:

$$\omega_k^i = \frac{p(X_k^i|Y_k)}{q(X_k^i|Y_k)} \quad (5.4)$$

A equação (3.13), da estimação recursiva bayesiana, dada por

$$p(X_k|Y_k) = \frac{p(x_k|x_{k-1})p(y_k|x_k)}{p(y_k|Y_{k-1})}p(X_{k-1}|Y_{k-1})$$

permite reescrever a equação do peso w_k^i como segue:

$$\omega_k^i = \frac{p(x_k^i|x_{k-1}^i)p(y_k|x_k^i)p(X_{k-1}^i|Y_{k-1})}{p(y_k|Y_{k-1})q(X_k^i|Y_k)} \quad (5.5)$$

Note que os termos $p(x^i|x_{k-1}^i)$ e $p(y_k|x_k^i)$ são conhecidos, respectivamente, pelas equações dos estados (5.1) e das observações (5.2) - então podem ser calculados. O termo $p(X_{k-1}^i|Y_{k-1})$ é a estimativa anterior de estado - então, como trata-se de um método recursivo, deve ser conhecido. O termo $p(y_k|Y_{k-1})$, no entanto, não é conhecido. Mas, ainda assim, é possível reescrever a equação (5.5) como:

$$\omega_k^i \propto \tilde{\omega}_k^i = \frac{p(x_k^i|x_{k-1}^i)p(y_k|x_k^i)p(X_{k-1}^i|Y_{k-1})}{q(X_k^i|Y_k)} \quad (5.6)$$

A pdf do denominador pode ser escolhida de tal forma que

$$q(X_k^i|Y_k) = q(x_k^i|X_{k-1}^i, Y_k)q(X_{k-1}^i|Y_{k-1}),$$

onde $q(X_{k-1}^i|Y_{k-1})$ representa a distribuição no tempo anterior $k-1$, e $q(x_k^i|X_{k-1}^i, Y_k)$ representa a probabilidade de transição para o estado x_k^i , dada a nova medida y_k .

A partir da equação anterior, pode-se reescrever a equação (5.6) como:

$$\tilde{\omega}_k^i = \frac{p(x_k^i|x_{k-1}^i)p(y_k|x_k^i)p(X_{k-1}^i|Y_{k-1})}{q(x_k^i|X_{k-1}^i, Y_k)q(X_{k-1}^i|Y_{k-1})}$$

Note que os últimos termos do numerador e denominador denotam ω_{k-1}^i , de acordo com a equação (5.4):

$$\omega_{k-1}^i = \frac{p(X_{k-1}^i|Y_{k-1})}{q(X_{k-1}^i|Y_{k-1})}$$

Desta forma, tem-se:

$$\tilde{\omega}_k^i = \frac{p(x_k^i|x_{k-1}^i)p(y_k|x_k^i)}{q(x_k^i|X_{k-1}^i, Y_k)}\omega_{k-1}^i \quad (5.7)$$

Agora que temos uma equação iterativa para os pesos ω_k^i , eles devem ser normalizados. Para isso,

$$\omega_k^i = \frac{\tilde{\omega}_k^i}{\sum_{m=1}^i \tilde{\omega}_k^m}$$

A equação (5.7) permite o cálculo dos pesos ω_k^i sem utilizar a pdf $P(X_k|Y_k)$, porém envolve uma outra distribuição $q(x_k^i|X_{k-1}^i, Y_k)$, a qual apenas admitiu-se ser conhecida e fácil de ser amostrada. Para concluir o algoritmo, é necessário escolher a distribuição q .

A proposta mais simples para a pdf $q(x_k^i|X_{k-1}^i, Y_k)$ é a própria pdf da transição de estado $p(x_k^i|x_{k-1}^i)$, pois a equação (5.7) torna-se [17, 10]:

$$\tilde{\omega}_k^i = \frac{p(x_k^i|x_{k-1}^i)p(y_k|x_k^i)}{q(x_k^i|X_{k-1}^i, Y_k)}\omega_{k-1}^i = p(y_k|x_k^i)\omega_{k-1}^i$$

Desta forma, a pdf $p(x_k|Y_k)$ pode ser aproximada por:

$$p(x_k|Y_k) \approx \sum_{i=1}^L \omega_k^i \delta(x_k - x_k^i)$$

e quanto maior o valor de L , melhor será essa aproximação [4].

Essa foi a escolha de $q(\cdot)$ usada nos exemplos que serão apresentados na seção 5.3.2, mas obviamente não é a única opção. Na teoria, a pdf $q(\cdot)$ deve ser escolhida de modo que tenha uma boa cobertura da distribuição original $p(x|y)$, ou seja, ter boa parte da sua distribuição sobreposta à $p(x|y)$. Mas na prática, a distribuição $q(\cdot)$ normalmente é escolhida a fim de simplificar a equação de atualização de peso e facilitar os cálculos [17].

5.3 Empobrecimento da amostra e reamostragem

A reamostragem é um passo importante no Filtro de Partículas, pois, sem ela, ocorre o empobrecimento da amostra - depois de um tempo, a maioria das partículas terão pesos insignificantes. Isso acontece porque apenas algumas das partículas estarão em um espaço da região de estado onde a pdf computada tem um valor significativo. Isso significa que o processo de reamostragem selecionará apenas algumas partículas a priori distintas para se tornarem partículas a posteriori [19]. A reamostragem resolve esse problema, mas cria outro, pois acaba destruindo informações e, conseqüentemente, aumentando a incerteza na amostragem aleatória. Por isso, é interessante começar o processo de reamostragem somente quando é realmente necessário. O método SIR, no entanto, utiliza a reamostragem a cada atualização[10]. A idéia básica por trás da reamostragem é eliminar as partículas que têm pesos pequenos e concentrar-se nas partículas com pesos significativos [4].

Uma forma bem direta de fazer a reamostragem é a seguinte: para $i = 1, \dots, N$,

1. Gera-se um número aleatório r uniformemente distribuído no intervalo $[0, 1]$;

2. Acumula-se a probabilidade q_i em um somatório, uma por vez, até que a soma acumulada seja maior que r . Ou seja, $\sum_{m=1}^{j-1} q_m < r$, mas $\sum_{m=1}^j q_m \geq r$. A nova partícula $x_{k,i}^+$ é definida da mesma forma que a partícula antiga $x_{k,j}^-$.

5.3.1 Roughening

Esse método também pode ser usado para evitar o empobrecimento da amostra e consiste em adicionar um ruído aleatório a cada partícula depois da reamostragem. Assim, as partículas a posteriori são modificadas da seguinte forma [19]:

$$\begin{aligned} x_{k,i}^+(m) &= x_{k,i}^+(m) + \Delta x(m), m = 1, \dots, n \\ \Delta x(m) &\sim (0, KM(m)N^{-1/n}), \end{aligned}$$

onde $\Delta x(m)$ é uma variável aleatória de média zero (normalmente gaussiana), e covariância $KM(m)N^{-1/n}$, sendo K um escalar, n a dimensão do espaço dos estados e M um vetor contendo a diferença máxima entre as partículas antes do processo.

5.3.2 Resultados

O modelo utilizado para os exemplos do Filtro de Partículas foi o mesmo usado nos exemplos do Filtro de Kalman Estendido. Maiores detalhes sobre a modelagem do sistema e considerações acerca das simulações constam na seção 4.5.1.

A condição inicial usada também foi a mesma, ou seja:

$$x_0 = \begin{bmatrix} x_1 & x_2 & S_1 & S_2 & C \end{bmatrix}^T = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T.$$

O ruído adicionado à saída de gás foi considerado o mesmo - gerado de uma distribuição normal, de valor esperado 0 e desvio padrão 1/30.

Serão apresentados três grandes casos: condição inicial correta e $P_0 = 0$, condição inicial incorreta e $P_0 \neq 0$, e condição inicial incorreta e $P_0 = 0$. Em cada um dos casos, serão utilizados diferentes número de partículas (L) para comparar a eficácia do filtro. Serão utilizados $L = 100$, $L = 1000$ e $L = 10000$.

5.3.2.1 Condição inicial correta ($\bar{x}_0 = x_0$) e covariância nula ($P_0 = 0$)

Como esperado, a figura 5.3 mostra que ao utilizar-se a mesma condição inicial para o filtro e para o modelo, e assumir isso como verdade ($P_0 = 0$), a estimativa do filtro fica próxima dos estados do modelo durante toda a simulação. Comparando-se os erros médios apresentados na tabela 5.1, a melhor estimativa foi obtida com $L = 10000$, mas com diferença apenas na terceira casa decimal (exceto para S_2 , que apresenta diferença na sexta casa decimal) para a estimativa com $L = 1000$.

	$L = 100$	$L = 1000$	$L = 10000$
x_1	0,0537	0,0290	0,0231
x_2	0,0099	0,0070	0,0030
S_1	0,0036	0,0023	0,0020
S_2	$2,0741 \times 10^{-4}$	$1,8462 \times 10^{-4}$	$1,8098 \times 10^{-4}$
C	0,0060	0,0043	0,0039
Tempo de simulação (s)	709	7419	96708

Tabela 5.1: Resultados para $\bar{x}_0 = x_0$ e $P_0 = 0$

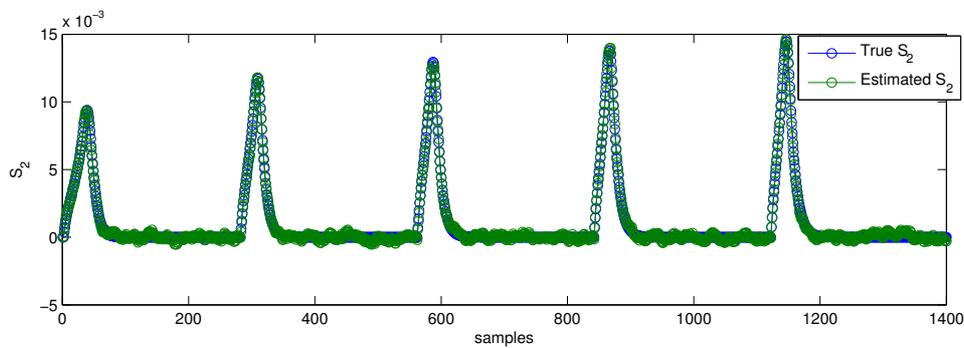


Figura 5.2: Simulação do S_2 com $\bar{x}_0 = x_0$, $P_0 = 0$ e $L = 10000$

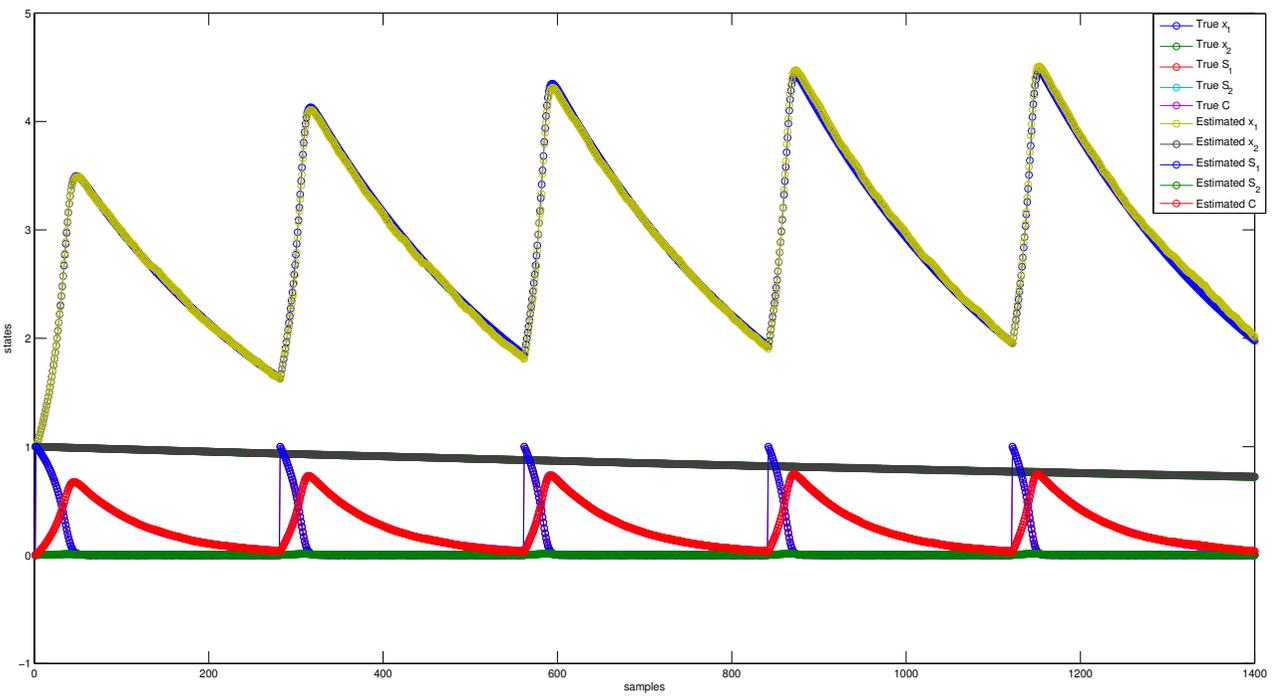


Figura 5.3: Simulação com $\bar{x}_0 = x_0$, $P_0 = 0$ e $L = 10000$

5.3.2.2 *Condição inicial incorreta ($\bar{x}_0 = [1, 5 \ 0,8 \ 0,2 \ 0,01 \ 0]^T$) e covariância não-nula ($P_0 \neq 0$)*

Ao utilizar-se $\bar{x}_0 \neq x_0$ e a matriz $P_0 = I_5$, algumas simulações começaram a apresentar erros. Por conta disso, consta na tabela 5.2, além dos erros médios das simulações que foram concluídas, o número de simulações que apresentaram erro (de um total de 100). É possível perceber que os erros médios apresentados foram os mais altos dentre todos vistos até aqui, chegando a quase 10 no estado C , quando $L = 10000$. Nesses exemplos, algumas simulações divergiram, como mostra a figura 5.6, contribuindo negativamente para o cálculo do erro médio das estimativas. Já a figura 5.5 apresenta um exemplo em que, com exceção de x_2 , as estimativas dos estados convergiram para os estados do modelo, mesmo o filtro errando bastante no início. A figura 5.4 apenas detalha a estimativa de S_2 nesse mesmo exemplo.

	$L = 100$	$L = 1000$	$L = 10000$
x_1	1,0838	0,7521	2,0522
x_2	0,6163	0,6945	0,7681
S_1	0,4662	0,1418	1,8033
S_2	1,0456	2,8177	6,5975
C	1,8162	4,4946	9,9928
Tempo de simulação (s)	611	6489	84410
Número de simulações com erro	47	55	79

Tabela 5.2: Resultados para $\bar{x}_0 = [1, 5 \ 0,8 \ 0,2 \ 0,01 \ 0]$, $P_0 = I_5$

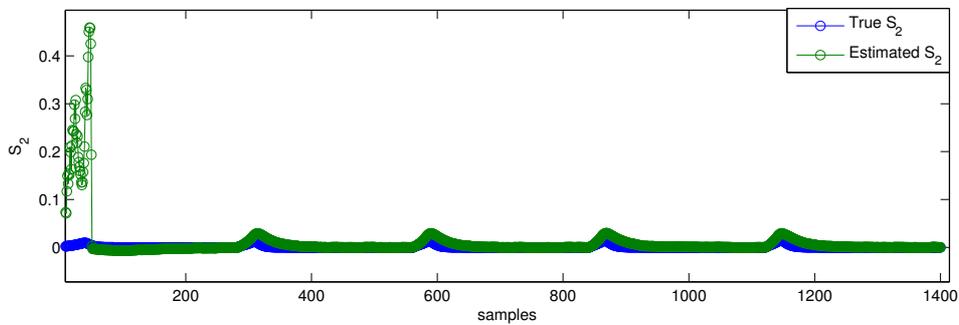


Figura 5.4: Simulação do S_2 com $\bar{x}_0 \neq x_0$, $P_0 = I_5$ e $L = 10000$

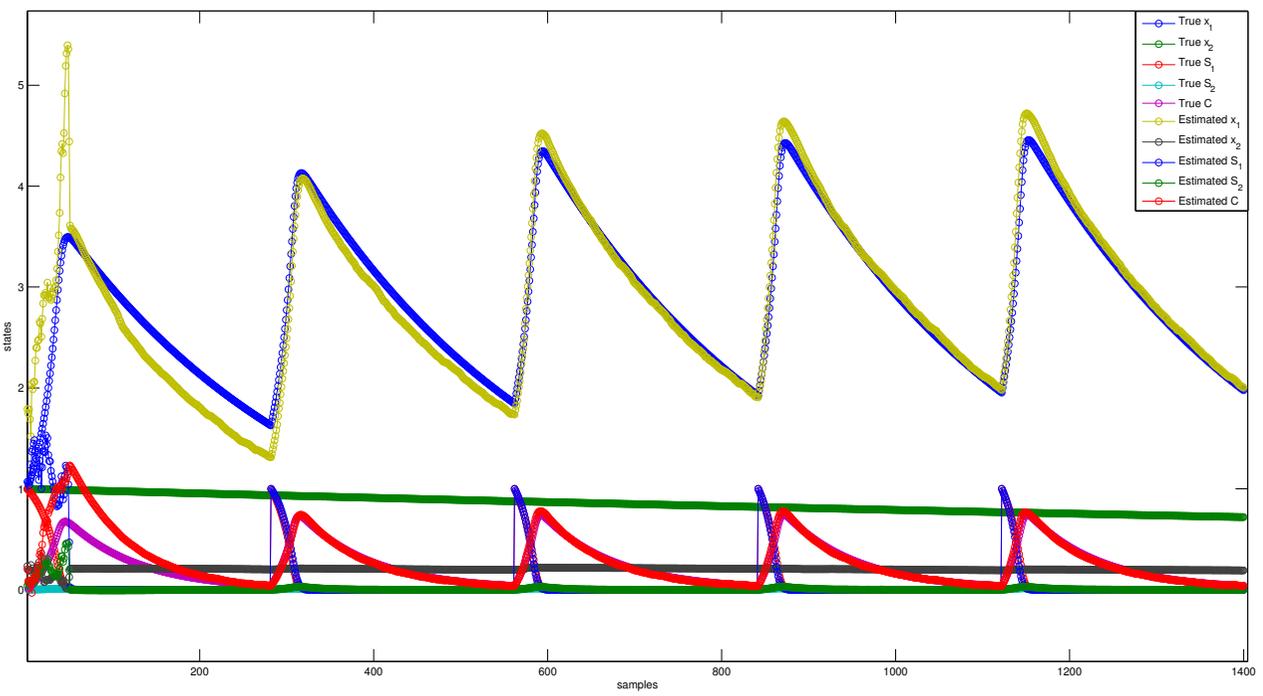


Figura 5.5: Simulação com $\tilde{x}_0 \neq x_0$, $P_0 = I_5$ e $L = 10000$

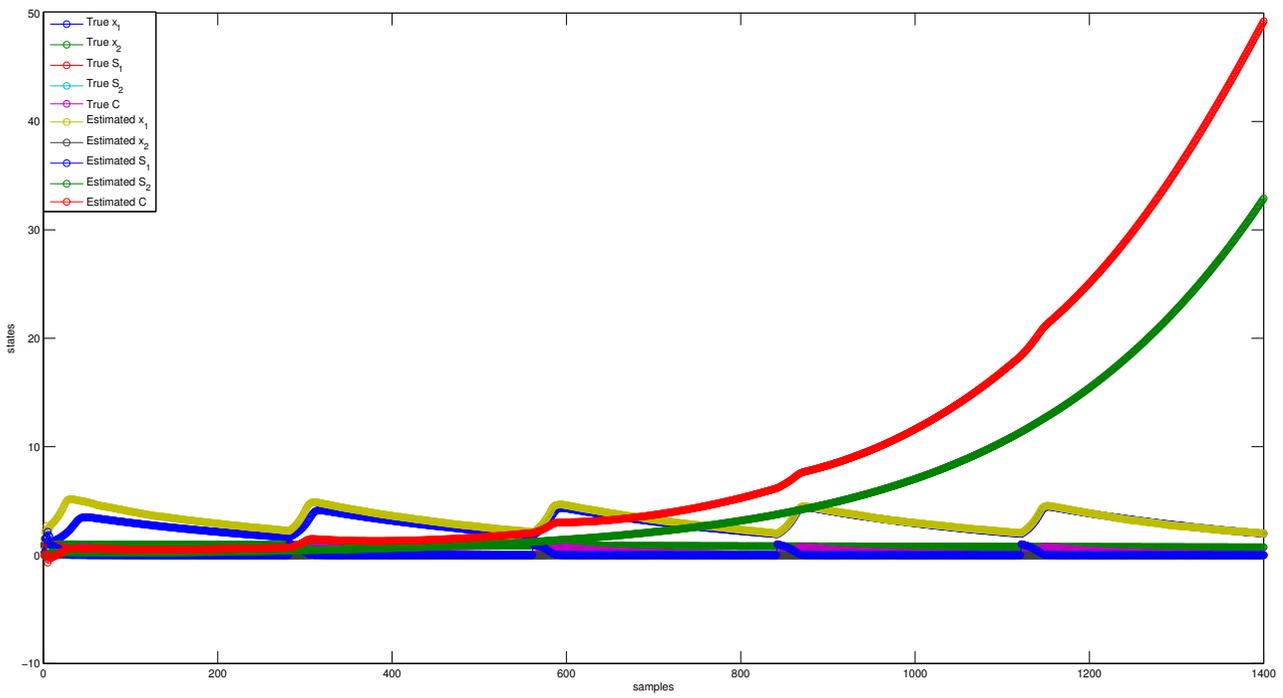


Figura 5.6: Exemplo de simulação em que o filtro diverge

Em busca de melhores resultados, e com base no erros médios apresentados no exemplo correspondente no capítulo 4, foram realizadas outras simulações, com $P_0 = I_5/100$ (a P_0 que apresentou, em geral, menores erros no Filtro de Kalman Estendido). Com essa P_0 , não foram apresentados erros nas simulações - ou seja, todas as 100 foram concluídas. Os erros médios apresentados constam na tabela 5.3, e foram mais satisfatórios do que os apresentados para $P_0 = I_5$. É possível perceber que não houve uma escolha de L que tenha se destacado por gerar os menores erros em todos os estados. O menor erro para x_1 apareceu em $L = 10000$, e para S_1 em $L = 100$, por exemplo.

	$L = 100$	$L = 1000$	$L = 10000$
x_1	0,3175	0,3281	0,2543
x_2	0,1986	0,1893	0,2223
S_1	0,0293	0,0363	0,0380
S_2	0,0011	0,0011	0,0012
C	0,0320	0,0310	0,0217
Tempo de simulação (s)	707	7242	99156
Número de simulações com erro	0	0	0

Tabela 5.3: Resultados para $\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]$, $P_0 = I_5/100$

5.3.2.3 Condição inicial incorreta ($\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]^T$) e covariância nula ($P_0 = 0$)

Ao dar ao filtro uma condição inicial diferente da usada no modelo e considerar $P_0 = 0$, é possível perceber na figura 5.8 que a estimativa do estado x_2 está se aproximando, embora lentamente, da solução do modelo; o estado x_1 demora cerca de 8 dias (o que equivale a 320 amostras) para convergir; os demais estados convergem já nos primeiros dias. De acordo com a tabela 5.4, os menores erros médios foram obtidos para $L = 10000$, mas com pouca diferença em relação a $L = 1000$ e $L = 100$.

No próximo capítulo, esses resultados serão comparados com os resultados apresentados pelo Filtro de Kalman Estendido, no capítulo 4.

	$L = 100$	$L = 1000$	$L = 10000$
x_1	0,3138	0,2912	0,2785
x_2	0,1715	0,1704	0,1703
S_1	0,0236	0,0218	0,0212
S_2	$9,7851 \times 10^{-4}$	$9,4856 \times 10^{-4}$	$9,4240 \times 10^{-4}$
C	0,0298	0,0282	0,0272
Tempo de simulação (s)	748	7223	101807

Tabela 5.4: Resultados para $\bar{x}_0 \neq x_0$, $P_0 = 0$

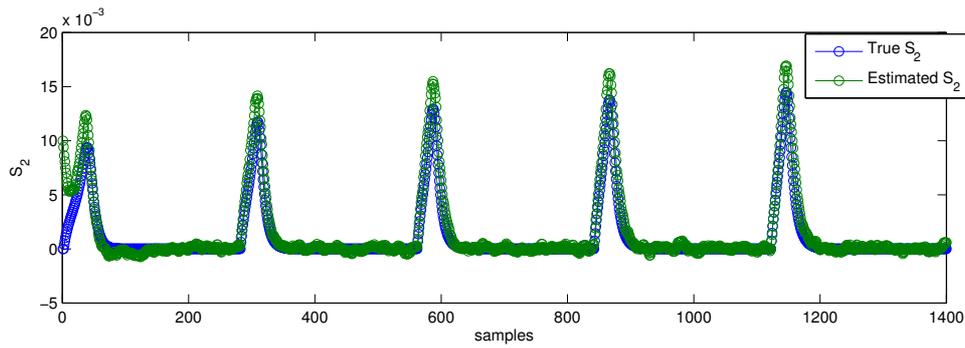


Figura 5.7: Simulação do S_2 com $\bar{x}_0 \neq x_0$, $P_0 = 0$ e $L = 10000$

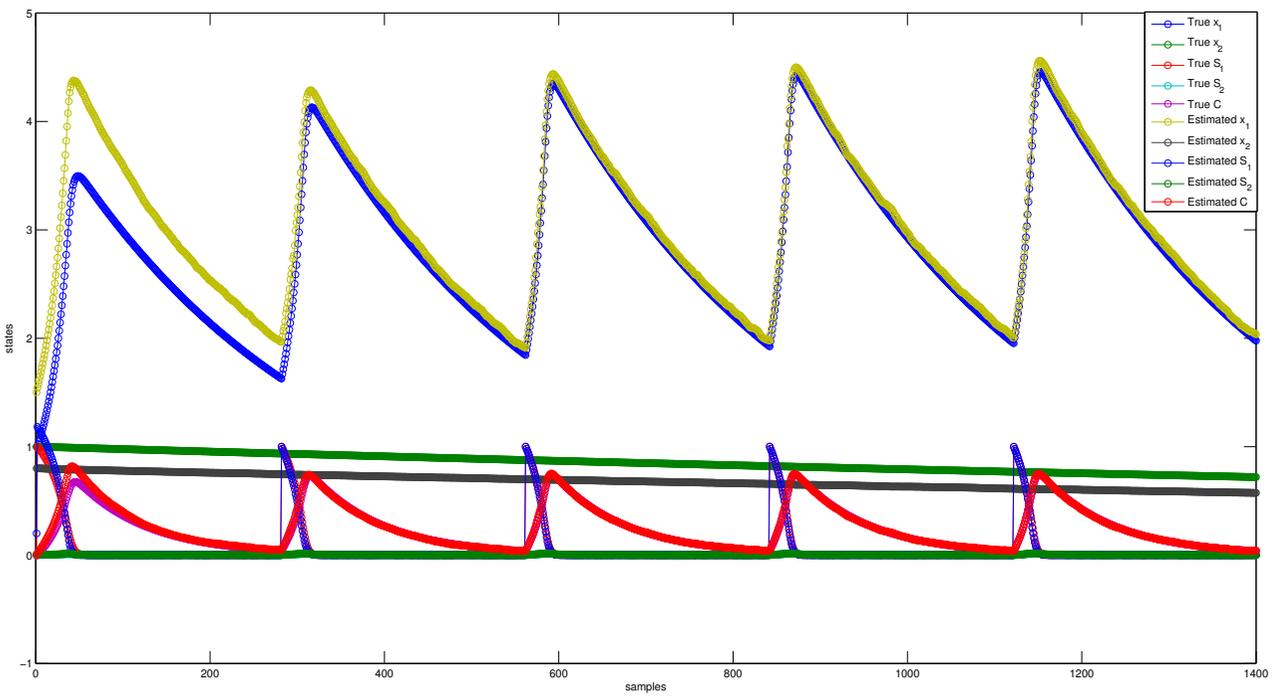


Figura 5.8: Simulação com $\bar{x}_0 \neq x_0$, $P_0 = 0$ e $L = 10000$

6 DISCUSSÃO DOS RESULTADOS

Neste capítulo, os resultados obtidos na estimação dos estados pelo Filtro de Kalman Estendido (EKF) e o Filtro de Partículas com Reamostragem (SIR), nos capítulos 4 e 5 serão comparados e discutidos.

6.1 Condição inicial correta ($\bar{x}_0 = x_0$) e covariância nula ($P_0 = 0$)

Esse é o caso em que se espera ter os melhores resultados para a estimação, visto que os filtros já iniciarão com a mesma condição inicial do modelo e a matriz $P_0 = 0$ indica a certeza de que $\bar{x}_0 = x_0$, pois afirma como nula a covariância da estimativa inicial. No entanto, esses são exemplos ilustrativos - visto que, na prática, é pouco provável que o estado inicial seja conhecido de forma precisa.

A tabela 6.1 mostra os erros médios apresentados pelos filtros. O SIR apresentou os menores erros médios para $L = 10000$. Quando foram usadas as matrizes de covariância Q e R apropriadas, o EKF apresentou erros ainda menores - além de um tempo de simulação mais atrativo.

	EKF	SIR		
		$L = 100$	$L = 1000$	$L = 10000$
x_1	$2,1729 \times 10^{-5}$	0,0537	0,0290	0,0231
x_2	$2,9795 \times 10^{-7}$	0,0099	0,0070	0,0030
S_1	$2,1138 \times 10^{-6}$	0,0036	0,0023	0,0020
S_2	$6,3566 \times 10^{-8}$	$2,0741 \times 10^{-4}$	$1,8462 \times 10^{-4}$	$1,8098 \times 10^{-4}$
C	$2,6083 \times 10^{-6}$	0,0060	0,0043	0,0039
Tempo de simulação (s)	64	709	7419	96708

Tabela 6.1: Comparação dos resultados com $\bar{x}_0 = x_0$ e $P_0 = 0$

Porém, quando as matrizes de covariância Q e R usadas no EKF não são consistentes com os ruídos considerados, os resultados ficam menos satisfatórios. Alguns casos constam na tabela 6.2, juntamente com os erros médios apresentados pelo SIR, com $L = 1000$ (que ficaram bem próximos dos erros com $L = 10000$, mas com um tempo de simulação mais de dez vezes menor). É possível perceber que, nesse caso, o SIR estimou melhor os estados.

	EKF		SIR
	$Q = I_5$	$R/10000$	$L = 1000$
x_1	2,3246	2,3152	0,0290
x_2	0,7298	0,7299	0,0070
S_1	2,7089	2,6140	0,0023
S_2	0,0332	0,0328	$1,8462 \times 10^{-4}$
C	0,1314	0,1306	0,0043
Tempo aproximado de simulação (s)	64	65	7419

Tabela 6.2: Comparação dos resultados com $\bar{x}_0 = x_0$, $P_0 = 0$ e matrizes Q e R manipuladas

6.2 Condição inicial incorreta ($\bar{x}_0 = [1,5 \ 0,8 \ 0,2 \ 0,01 \ 0]^T$) e covariância não-nula ($P_0 \neq 0$)

Esse é um caso mais condizente com a realidade: o estado inicial a ser usado nos filtros não é igual o estado inicial do modelo. Ao ser fornecida ao filtro uma matriz de covariância desse estado inicial $P_0 = I_5/100$, novamente o EKF apresentou os menores erros médios. As estimativas do SIR ficaram piores do que o esperado, principalmente quando $L = 10000$ - que demanda um maior tempo de simulação e não apresenta grande melhora nas estimativas, quando comparadas com as obtidas com $L = 100$ e $L = 1000$.

	EKF	SIR		
		$L = 100$	$L = 1000$	$L = 10000$
x_1	0,1611	0,3175	0,3281	0,2543
x_2	0,1275	0,1986	0,1893	0,2223
S_1	0,0937	0,0293	0,0363	0,0380
S_2	$6,4658 \times 10^{-4}$	0,0011	0,0011	0,0012
C	0,0123	0,0320	0,0310	0,0217
Tempo de simulação (s)	63	707	7242	99156

Tabela 6.3: Comparação dos resultados com $\bar{x}_0 \neq x_0$ e $P_0 = I_5$

	EKF	SIR		
		$L = 100$	$L = 1000$	$L = 10000$
x_1	2,0023	1,0838	0,7521	2,0522
x_2	0,3929	0,6163	0,6945	0,7681
S_1	0,3318	0,4662	0,1418	1,8033
S_2	0,2265	1,0456	2,8177	6,5975
C	0,1829	1,8162	4,4946	9,9928
Tempo de simulação (s)	67	611	6489	84410
Número de simulações com erro	0	47	55	79

Tabela 6.4: Comparação dos resultados com $\bar{x}_0 \neq x_0$ e $P_0 = I_5/100$

Ao utilizar $P_0 = I_5$, algumas simulações do SIR apresentaram erros, provavelmente devido a inconsistências matemáticas (como divisão por uma matriz nula, por exemplo). Desta forma, o erro médio foi calculado apenas para as simulações que foram concluídas, e algumas delas divergiram, como ilustra a figura 5.6, na seção 5.3.2. Para o caso em que $L = 10000$, por exemplo, apenas 21 simulações foram consideradas para determinar os erros médios. Isso, junto ao fato das divergências, colaborou para que as estimativas não ficassem boas. Já o EKF

concluiu as 100 simulações propostas e apresentou erros médios menores para maioria os estados.

6.3 Condição inicial incorreta ($\bar{x}_0 = [1, 5 \ 0, 8 \ 0, 2 \ 0, 01 \ 0]^T$) e covariância nula ($P_0 = 0$)

Esse foi o caso em que os resultados entre o EKF e o SIR, com $L = 10000$ ficaram mais parecidos. Além disso, os erros apresentados pelo SIR estão em concordância com a ideia do método: quanto maior o número de partículas L , melhor a estimativa. Porém, dado o tempo de simulação e a sutil diferença nos erros médios (principalmente entre $L = 1000$ e $L = 10000$), usar um número tão grande de partículas acaba não sendo uma boa opção. Na verdade, quando comparado ao EKF, que além de obter uma estimativa melhor, tem um tempo de simulação muito menor, o SIR não é atrativo nem mesmo com $L = 100$.

	EKF	SIR		
		$L = 100$	$L = 1000$	$L = 10000$
x_1	0,2505	0,3138	0,2912	0,2785
x_2	0,1698	0,1715	0,1704	0,1703
S_1	0,0206	0,0236	0,0218	0,0212
S_2	$9,4013 \times 10^{-4}$	$9,7851 \times 10^{-4}$	$9,4856 \times 10^{-4}$	$9,4240 \times 10^{-4}$
C	0,0249	0,0298	0,0282	0,0272
Tempo de simulação (s)	64	748	7223	101807

Tabela 6.5: Comparação dos resultados com $\bar{x}_0 \neq x_0$ e $P_0 = 0$

Mesmo usando uma aproximação linear do sistema, e com exceção do caso em que $\bar{x}_0 \neq x_0$ e $P_0 = I_5$, o Filtro de Kalman Estendido apresentou as melhores estimativas para os estados. Além disso, o tempo de simulação é cerca de 10 vezes menor quando comparado ao caso mais rápido do Filtro de Partículas

(com $L = 100$). No entanto, foi possível perceber que a escolha das matrizes Q e R afeta sensivelmente os resultados do EKF. Mesmo quando o estado inicial dado ao filtro é igual ao usado no modelo ($\bar{x}_0 = x_0$) e a matriz P_0 é nula, se as matrizes Q e R são incoerentes com os ruídos dos estados e da saída do sistema, os erros médios acabam sendo menos satisfatórios.

Com a escolha da proposta de distribuição q e a opção de reamostrar a cada passo, o Filtro de Partículas não se destacou pela qualidade das estimativas e ainda demandou muito mais tempo. Além disso, em alguns casos, apresentou erros nas simulações ou divergiu completamente.

7 CONCLUSÕES

Analisar os fenômenos inerentes ao processo de digestão em um biorreator é fundamental para compreender a sua dinâmica e descrevê-la através de modelos matemáticos. O processo é não-linear e a concentração de ácidos graxos voláteis (VFA) é determinante para o bom funcionamento do sistema. Como altas concentrações de VFA podem acarretar na redução da produção ou até mesmo cessá-la, os biorreatores acabam sendo operados bem abaixo da sua capacidade. Assim, se faz necessário identificar as concentrações de bactérias e a quantidade de substratos, principalmente VFA, no interior de um biorreator para otimizar o funcionamento.

A aplicação dos problemas de filtragem em um biorreator anaeróbico visa, através da modelagem do sistema e das observações da saída do gás metano, estimar as concentrações de bactérias, substrato, VFA e gás carbônico - o que é uma alternativa plausível, visto que, na prática, efetuar as medições dos estados exige, em geral, técnicas custosas e demoradas.

Nesse trabalho, duas propostas de filtragem foram utilizadas: o Filtro de Kalman Estendido (EKF) e o Filtro de Partículas com Reamostragem (SIR). Diferentes casos foram considerados, a fim de comparar a eficácia dos algoritmos. O Filtro de Kalman Estendido, embora utilize uma aproximação linear do sistema, estimou com mais eficiência os estados do sistema em praticamente todos os exemplos apresentados. Além do mais, o filtro de Kalman tem uma implementação mais simples e um tempo de simulação bem menor - cerca de 10 vezes menor, quando comparado ao exemplo de Filtro de Partículas que demandou menos tempo ($L = 100$).

O SIR tem variações que interferem diretamente na sua eficácia e no tempo de simulação: a escolha da proposta de distribuição q , quando efetuar a reamostragem e o próprio número de partículas. A proposta de q e a alternativa

de reamostrar a cada atualização não geraram resultados satisfatórios, a ponto de tornar o SIR mais atrativo do que EKF nos exemplos realizados.

A partir dos resultados obtidos, é possível utilizar o Filtro de Kalman Estendido para a estimação dos estados de um biorreator anaróbico semibatelada e, assim, otimizar seu funcionamento - pois possibilita decidir quando deve ser feita uma nova inserção de substrato. Isso também viabiliza o controle da operação. Pode-se, ainda, utilizar outras propostas de reamostragem e escolhas de q no Filtro de Partículas para melhorar suas estimativas. Por fim, também é possível utilizar os algoritmos de filtragem em um modelo de biorreator operado em modo contínuo.

Referências Bibliográficas

- [1] ACTION, P. Probabilidades. <http://www.portalaction.com.br/probabilidades> (2017).
- [2] ANDERSON, B. D. O., AND MOORE, J. B. *Optimal Filtering*, 1st ed. Prentice-Hall, INC., 1979.
- [3] ANEEL. *ATLAS da Energia Elétrica no Brasil*, 3rd ed. Agência Nacional de Energia Elétrica, 2008.
- [4] ARULAMPALAM, M. S., MASKELL, S., GORDON, N., AND CLAPP, T. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEE Transactions on Signal Processing* 50, 2 (2002), 174–188.
- [5] BERNARD, O., STEYER, J., AND DOCHAIN, D. Dynamical model development and parameter identification for anaerobic wastewater treatment process. *Biotechnology and Bioengineering* (2001).
- [6] DE CARVALHO, J. C. M., AND SATO, S. *Biotecnologia Industrial: engenharia bioquímica Vol. 2*. Editora Edgar Blücher Ltda., 2001.
- [7] DOCHAIN, D. *Automatic Control of Bioprocesses*, 1st ed. ISTE Ltd, John Wiley and Sons Inc, 2008.
- [8] ECKHARD, D., AND CAMPESTRINI, L. Análise do uso de modelos discretizados para identificação de modelos para biorreatores anaeróbicos. *DIN-CON - Conferência Brasileira de Dinâmica, Controle e Aplicações* (2015).
- [9] GELB, A. *Applied Optimal Estimation*, 1st ed. The Analytic Sciences Corporation, 1974.
- [10] GUSTAFSSON, F. Particle filter theory and practice with positioning applications. *IEEE A&E Systems* 25, 7 (2010), 53–70.

- [11] HASELEIN, W. M. *Identificação de Modelos Dinâmicos para Biorreatores Anaeróbicos*. UFRGS, 2017.
- [12] KIDMAN, T., BEDDINGFIELD, E., MINUCCI, C., COOK, M., AND ABEL, C. V. *Issue Paper: Aerobic and Anaerobic Bioreactor Project Protocol*. Science Application International Corporation, 2011.
- [13] KOCH, F. F. *Avaliação da viabilidade técnica quanto a obtenção de biometano através da purificação de biogás em meio aquoso: um estudo de caso do Projeto Consórcio Verde Brasil*. UFRGS, 2014.
- [14] MELVILLE, L., WEGER, A., WIESGICKL, S., AND FRANKE, M. *Anaerobic Digestion, in Transformation of Biomass: Theory to Practice*, 1st ed. John Wiley and Sons, Ltd., 2014.
- [15] RUI, R. *Estimação de Modelos Afins por Partes em Espaço de Estados*. UFRGS, 2016.
- [16] SALVATORI, T. *Modelagem Caixa-preta de Biorreatores em modo descontínuo utilizando modelos polinomiais do tipo NAR e NARMA*. UFRGS, 2016.
- [17] SHARMA, S. Importance sampling. <http://cecas.clemson.edu/~ahoover/ece854/lecture-notes/lecture-import-samp.pdf> (2017).
- [18] SHARMA, S. Recursive bayesian estimate. <http://cecas.clemson.edu/~ahoover/ece854/> (2017).
- [19] SIMON, D. *Optimal State Estimation: Kalman, H infinity, and Nonlinear Approaches*, 1st ed. John Wiley and Sons, Inc., 2006.
- [20] SMITH, J. E. *Biotechnology*, 5th ed. Cambridge University Press, 2009.

- [21] TEIXEIRA, J. B. Biorreatores. *Biotecnologia, Ciência e Desenvolvimento* pp.36-41 (2002).
- [22] TOCCHETTO, M. A. D. *Estimador de Estados para Robô Diferencial*. UFRGS, 2017.
- [23] TOKDAR, S. T., AND KASS, R. E. Importance sampling: A review. <http://www2.stat.duke.edu/st118/Publication/impsamp.pdf> (2017).