

Laboratório de Modelagem, Análise e Controle de Sistemas Não-Lineares
Departamento de Engenharia Eletrônica
Universidade Federal de Minas Gerais
Av. Antônio Carlos, 6627, 31270-901, Belo Horizonte, MG, Brasil
Fone: +55 31 3499-4866 - Fax: +55 31 3499-4850
aguirre@cpdee.ufmg.br



Métodos de Estimação Recursiva Baseados no Filtro de Kalman Aplicados a Sistemas Não-Lineares

Hudson Bitencourt Junior

Dissertação submetida à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, como parte dos requisitos necessários à obtenção do grau de Mestre em Engenharia Elétrica.

Orientadores: Dr. Leonardo A. B. Tôrres
Dr. Luis A. Aguirre

Belo Horizonte, 18 de Novembro de 2003.

*Aos meus queridos pais, pelo amor e dedicação
Às minhas queridas irmãs pelo carinho*

Agradecimentos

Em primeiro lugar agradeço a Deus por estar presente em todos os momentos de minha vida, e portanto em cada etapa da realização deste trabalho.

Agradeço com muito carinho aos meus pais, Udison e Jadismar, por todo amor, dedicação, incentivo, e apoio em toda minha vida acadêmica desde o pré-primário até o mestrado. Tenho certeza de que se estou tendo a graça de chegar até aqui, foi pela constante presença de meus pais sempre dispostos a me amparar no que fosse necessário. Agradeço pelas orações que fizeram por mim. Aproveito para pedir desculpas pelo sofrimento que lhes causei, ao me afastar de casa para ingressar nessa difícil jornada.

Agradeço à minha irmã Gabriele pela paciência e compreensão em me ver partir para outra cidade em um momento tão delicado de sua vida, que é a infância. Agradeço à minha irmã Michele e meu cunhado Jean por terem torcido por mim.

Agradeço à minha avó Maria de Lourdes pelas muitas e valiosas orações ao Senhor nosso Deus, sempre suplicando pela minha felicidade e segurança nas viagens e no dia-a-dia de quem mora em uma cidade grande.

Agradeço aos meus familiares que residem em Brasília, que me receberam com tanto carinho e entusiasmo na capital desse país, onde dei meus primeiros passos no mestrado. Agradeço de maneira especial à tia Nenir e tio Antônio que me acolheram em sua casa como se fosse um filho.

Gostaria de deixar registrado, meu eterno agradecimento ao vovô Nilo. Agradeço por ter me incentivado a estudar Matemática desde cedo, me trazendo com grande entusiasmo, problemas escritos em tirinhas de papel, e pedindo que eu os tentasse resolver. Quando ele nos deixou eu ainda não havia começado a graduação, mas parece que ele sabia que um dia eu precisaria daquele conhecimento para estudar problemas mais complexos, como o algoritmo que tornou possível ao homem chegar até a Lua!

Agradeço a todos os mestres e funcionários das instituições que contribuíram para minha formação e principalmente ao pessoal do Centro Universitário do Leste de MG - UnilesteMG. Agradeço em especial ao professor Marcelo Vieira pelo convívio durante quase a metade do curso de graduação sempre mostrando disponibilidade, pessoalmente ou por email, para discussões e questionamentos. Agradeço pelo grande incentivo e motivação para meu ingresso no mestrado. Agradeço ao professor Manuel Camela por ter me orientado quando eu dava meus primeiros passos na pesquisa como estagiário no UnilesteMG.

Agradeço aos professores que foram meus mestres na UFMG: Benjamin Menezes, Hani Yehia, Eduardo Mendes e Luis Aguirre. Agradeço também aos professores da UnB, onde cursei algumas disciplinas.

Agradeço aos colegas e funcionários do CPDEE e CPH pelo companheirismo, proporcionando um excelente ambiente de trabalho. Agradeço àqueles que estiveram mais próximos e que além de colegas tornaram-se amigos, dentre os quais destaco os amigos do grupo MACSIN: Dair, Edgar, Erivelton, Gladstone, Gleison, Mara, Márcio e Viviane. Aos companheiros de república agradeço pelo convívio.

Agradeço à CAPES pelo apoio financeiro.

Agradeço de maneira muito especial aos professores Dr. Leonardo Tôrres e Dr. Luis Aguirre por terem me incentivado a pesquisar sobre um assunto tão importante e pouco estudado em nosso país. Agradeço-lhes pela excelente orientação, procurando me mostrar qual o melhor caminho a seguir em cada etapa do trabalho, sempre pacientes em responder a meus longos e chatos emails relatando algum resultado ou pedindo ajuda. Agradeço em especial ao prof. Luis Aguirre pela oportunidade de fazer parte desse brilhante grupo de pesquisa dirigido por ele. Ao prof. Leonardo Tôrres agradeço também pelo convívio, disponibilidade e atenção, me motivando a seguir em frente ao demonstrar tanto entusiasmo pelo nosso trabalho.

Agradeço também àqueles que me esqueci de citar e registro meu pedido de perdão.

“Tudo posso naquele que me fortalece.”

Filipenses 4,13

Sumário

Resumo	xiii
Abstract	xv
Lista de Tabelas	xvii
Lista de Figuras	xx
Lista de Símbolos	xxi
Lista de Abreviações	xxv
1 Introdução	1
1.1 Motivação	3
1.2 Objetivos	5
1.3 Justificativa	5
1.4 Organização do Texto	7
2 Estimação Recursiva	9
2.1 Introdução	9
2.2 Modelos Estocásticos e Abordagem Bayesiana	10
2.2.1 Representação de Sistemas por Modelos Estocásticos .	12
2.2.2 Conceitos Básicos em Processos Estocásticos	12
2.2.3 A Abordagem Bayesiana	15

2.3	Estimadores Recursivos para Identificação de Sistemas	19
2.3.1	Atualização Recursiva	19
2.3.2	O Estimador Recursivo de Mínimos Quadrados	21
2.4	O Filtro de Kalman	22
2.4.1	O Algoritmo	24
2.4.2	O Algoritmo em uma Abordagem Bayesiana	27
2.5	Conclusões	30
3	O Filtro de Kalman para Sistemas Não-Lineares	31
3.1	Introdução	31
3.2	O Filtro de Kalman Estendido	33
3.3	O Filtro de Kalman Unscented	36
3.3.1	A Transformação Unscented	36
3.3.2	O Algoritmo do Filtro de Kalman Unscented	40
3.3.3	Estimação de Parâmetros	42
3.3.4	Aspectos da Implementação do Algoritmo	43
3.3.5	Exemplos de Aplicações	45
3.4	Conclusões	53
4	Estimação Recursiva: Aplicação ao Circuito de Chua	55
4.1	Introdução	55
4.2	Sistemas Caóticos	56
4.3	O Circuito de Chua	57
4.3.1	Descrição do Sistema	57
4.3.2	Montagem do Circuito Real	59
4.4	Estimando Estados	60
4.4.1	Metodologia	60
4.4.2	Sistema Simulado	62
4.4.3	Plataforma PCChua	69
4.5	Estimação Conjunta de Estados e Parâmetros	71

4.5.1	Metodologia	72
4.5.2	Resultados e Discussões	73
4.5.3	O Problema da Inicialização	78
4.6	Conclusões	80
5	Transmissão de Informação	83
5.1	Introdução	83
5.2	Transmissão e Recuperação de Informação	84
5.2.1	Codificação de Informação	84
5.2.2	Decodificação de Informação	85
5.2.3	Modificações no Algoritmo	85
5.3	Sistema de Lorenz	86
5.3.1	Metodologia	87
5.3.2	Resultados e Comentários	88
5.4	Circuito de Chua	93
5.4.1	Metodologia	94
5.4.2	Resultados e Comentários	95
5.5	Conclusões	100
6	Conclusão e Pesquisas Futuras	103
6.1	O Filtro de Kalman	103
6.2	Estimação em Sistemas Não-Lineares	104
6.3	Detecção de Variações nos Parâmetros	105
6.4	Conclusões	106
6.5	Pesquisas Futuras	106
	Bibliografia	109
A	Biografia: R. E. Kalman	117

Resumo

Estimação recursiva é uma técnica que processa os dados medidos de um determinado sistema seqüencialmente, fornecendo novas estimativas para as variáveis desejadas a cada iteração do algoritmo. Essa técnica pode ser usada para estimar os parâmetros do modelo em *Identificação de Sistemas*, ou para estimar as variáveis de estado de sistemas dinâmicos, procedimento necessário em técnicas de controle moderno quando não é possível medir diretamente todas as variáveis.

A estimação de estados pode ser realizada por diversos algoritmos recursivos. Neste trabalho apresenta-se o *Filtro de Kalman*, um estimador recursivo muito eficiente capaz de fornecer estimativas ótimas para os estados de sistemas dinâmicos, a partir de medições corrompidas por ruído. Além da descrição original do Filtro de Kalman para sistemas lineares, são apresentadas duas abordagens diferentes que permitem que o algoritmo seja aplicado a sistemas não-lineares: o Filtro de Kalman Estendido (EKF) e o Filtro de Kalman *Unscented* (UKF).

Os desempenhos dos dois algoritmos estudados neste trabalho (EKF e UKF), são comparados durante a aplicação em um sistema real, o *circuito de Chua*. Os algoritmos são utilizados para estimar as três variáveis de estado desse sistema. O UKF é capaz de propagar a matriz de covariância dos estados sem linearizar as funções do sistema, de forma que as estimativas para os parâmetros e estados são calculadas com mais precisão do que pelo EKF. Além dos resultados, discute-se as dificuldades encontradas durante a implementação. Outra aplicação interessante é a transmissão de informação. Mostra-se que é possível inserir códigos binários em séries temporais caóticas por meio de sucessivas variações de um ou mais parâmetros do sistema. A informação pode ser recuperada por estimação conjunta de estados e parâmetros via UKF.

Abstract

Recursive estimation is a technique that processes measured data from a system in a sequential way, providing new estimates to the desired variables at each algorithm step. This technique can be used to estimate the model parameters in *System Identification*. It can also be used to estimate the state variables of dynamic systems, which is required on modern control techniques when all the variables cannot be measured directly.

State estimation can be performed by many recursive algorithms. In this work a very efficient recursive estimator, capable of providing optimal estimates of system states from noisy measurements will be presented. It is known as *Kalman Filter*. In addition to the linear Kalman Filter original description, two approaches that make the algorithm applicable to nonlinear systems will be presented: the Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF).

The algorithms performances (EKF and UKF) are compared in a real application, the *Chua's circuit*. The algorithms are used to estimate the three state variables of this system. The UKF can propagate the covariance matrix of the states without linearizing the system functions, in such a way that parameters and unobserved states can be estimated with higher accuracy than EKF. In addition to the results, the difficulties found during implementation are discussed. Another interesting application is information transmission. It is shown that it is possible to insert binary codes in chaotic time series, by successive variation of one or more parameters of the system. Information can be recovered by jointly estimating states and parameters using the UKF.

Lista de Tabelas

4.1	Parâmetros da plataforma PCChua	59
4.2	RMSE – estimação via EKF – Chua simulado	64
4.3	RMSE – estimação via UKF – Chua simulado	65
4.4	RMSE – estimação via EKF/UKF – Chua real	69
4.5	Parâmetros estimados por UKF – Circ. de Chua	73
4.6	Parâmetros verdadeiros estimados por UKF – Circ. de Chua .	73
4.7	PRMSE – estimação via UKF – Chua real	74
4.8	PRMSE – estimação de estados via EKF/UKF com parâmetros estimados por UKF	77
5.1	Valores da resistência variável – primeiro experimento	96
5.2	Valores da resistência variável – segundo experimento	99

Lista de Figuras

2.1	Função densidade de probabilidade gaussiana	14
3.1	Transformação <i>Unscented</i>	40
3.2	Evolução dos parâmetros na estimação – Lotka-Volterra.	48
3.3	Estimação de estados – Lotka-Volterra	48
3.4	Plano de fase – Lotka-Volterra	49
3.5	Evolução dos parâmetros na estimação – Lorenz.	51
3.6	Filtragem de sinal ruidoso – Lorenz	51
3.7	Estimação de estados – Lorenz	52
3.8	Atrator de Lorenz	52
4.1	Circuito de Chua	58
4.2	Curva característica do diodo de Chua	58
4.3	Diagrama de montagem do PC-CHUA	60
4.4	Curva característica do diodo de Chua – Interpolação	63
4.5	Filtragem de sinal ruidoso via EKF – Chua simulado	64
4.6	Estimação de estados via EKF – Chua simulado	65
4.7	Filtragem de sinal ruidoso via UKF – Chua simulado	66
4.8	Estimação de estados via UKF – Chua simulado	66
4.9	Atrator dupla volta – Chua simulado	67
4.10	Relação sinal/ruído \times PRMSE para x_1	68
4.11	Relação sinal/ruído \times PRMSE – x_2 e x_3	68

4.12	Série temporal com ruído gaussiano adicionado	69
4.13	Filtragem de sinal ruidoso via EKF/UKF – Chua real	70
4.14	Estimação de estados via EKF/UKF – Chua real	70
4.15	Evolução dos parâmetros na estimação – Chua real	74
4.16	Estimação conjunta: filtragem do sinal medido via UKF	75
4.17	Estimação conjunta via UKF – Chua real	75
4.18	Projeção do atrator dupla volta – Chua real	76
4.19	Projeção do atrator reconstruído com parâmetros estimados	76
4.20	Filtragem de sinal ruidoso via EKF/UKF – Param. estim.	77
4.21	Estimação de estados via EKF/UKF – Param. estim.	78
4.22	Estimação de estado via UKF – resultado divergente	79
5.1	Variação de parâmetros – Lorenz – primeiro caso	88
5.2	Variação de parâmetros – Lorenz – terceiro caso	88
5.3	Estimação do estado simulado – Lorenz – primeiro caso	89
5.4	Erro de estimação – Lorenz – primeiro caso	90
5.5	Estimação do parâmetro variável – Lorenz – primeiro caso	90
5.6	Estimação do parâmetro variável – Lorenz – segundo caso	91
5.7	Estimação do estado simulado – Lorenz – terceiro caso	92
5.8	Erro de estimação – Lorenz – terceiro caso	92
5.9	Estimação de parâmetros variáveis – Lorenz – terceiro caso	93
5.10	Estimação do estado medido – Chua – primeiro experimento	95
5.11	Estimação do parâmetro variável – Chua – primeiro experimento	96
5.12	Estimação do estado medido – Chua – série temporal após convergência	97
5.13	Erro de estimação – Chua – primeiro experimento	97
5.14	Estimação do estado medido – Chua – segundo experimento	98
5.15	Erro de estimação – Chua – segundo experimento	98
5.16	Estimação do parâmetro variável – Chua – segundo exp.	99

Lista de Símbolos

Neste trabalho, as matrizes são indicadas por letras maiúsculas em itálico. Vetores são representados por letras minúsculas e em negrito. Escalares são representados por letras minúsculas em itálico. As letras utilizadas são dos alfabetos grego e latino. A seguir são listados os principais símbolos usados no texto.

$\text{cov}[\cdot]$	Covariância;
\mathbb{R}	Conjunto dos números reais;
σ	Conjunto de pontos sigma;
\mathcal{X}	Conjunto de vetores gerado na transformação <i>unscented</i> ;
κ	Constante do UKF;
σ_x	Desvio padrão da variável aleatória x ;
ξ	Erro de estimação;
$E[\cdot]$	Esperança matemática ou valor esperado;
f	Função não-linear do vetor de estados;
h	Função não-linear do vetor de estados;
K	Ganho de Kalman;
$A \cap B$	Intersecção entre os eventos A e B ;
\log	Logaritmo;
$ x_{\max} $	Maior valor absoluto do sinal x ;

P	Matriz de covariância dos estados;
\tilde{P}	Matriz de covariância ou variância predita;
P^+	Matriz de covariância dos estados estendida;
P^{xy}	Matriz de correlação cruzada entre \mathbf{x} e \mathbf{y} ;
P^{yy}	Matriz de covariância do vetor de saída;
Q	Matriz de covariância do ruído de processo;
R	Matriz de covariância do ruído de medição;
Ψ	Matriz de regressores;
\bar{x}	Média de x ;
$\bar{\mathbf{x}}$	Média do vetor \mathbf{x} ;
n_{SP}	Número de pontos sigma;
n	Número de variáveis ou dimensão do vetor de estados;
$w^{(i)}$	Pesos do UKF;
T_s	Período de amostragem;
$P(A)$	Probabilidade de ocorrência do evento A ;
$P(A B)$	Probabilidade condicional de A dado B ;
$\hat{\mathbf{x}}_{k k-1}$	Predição de \mathbf{x}_k usando informações tomadas até o instante $k - 1$;
y_k	Sinal medido ou observado no instante k ;
x_k	Sinal real;
\hat{x}_k	Sinal estimado;
\tilde{x}_k	Sinal predito;
\mathbf{x}^T	Transposto do vetor \mathbf{x} ;

\mathbf{x}	Vetor de estados;
\mathbf{x}^+	Vetor de estados estendido;
$\hat{\mathbf{x}}$	Vetor de estados estimado;
$\hat{\mathbf{x}}_{k k}$	Vetor de estados estimado após etapa de correção;
\mathbf{x}^*	Vetor de estados de referência, usado no EKF;
$\boldsymbol{\delta}_k$	Vetor de estados após linearização feita no EKF;
λ	Vetor de parâmetros em problemas de estimação de estados;
$\boldsymbol{\theta}$	Vetor de parâmetros em problemas de identificação;
$\hat{\boldsymbol{\theta}}$	Vetor de parâmetros estimado;
$\hat{\boldsymbol{\theta}}_{\text{MQ}}$	Vetor de parâmetros estimado por mínimos quadrados;
$\boldsymbol{\psi}$	Vetor de regressores;
$\boldsymbol{\xi}$	Vetor de resíduos resultante da estimação;
\mathbf{w}	Vetor de ruído de processo;
\mathbf{e}	Vetor de ruído de medição;
\mathbf{y}	Vetor de saídas do sistema;
$\boldsymbol{\nu}$	Vetor de saídas após linearização feita no EKF.

Lista de Abreviações

KF	Filtro de Kalman (<i>Kalman filter</i>);
EKF	Filtro de Kalman Estendido (<i>Extended Kalman Filter</i>);
UKF	Filtro de Kalman Unscented (<i>Unscented Kalman Filter</i>);
PDF	Função densidade de probabilidade (<i>Probability density function</i>);
SP	Ponto Sigma (<i>Sigma Point</i>);
RMSE	Raiz quadrada do erro quadrático médio (<i>Root Mean Square Error</i>);
PRMSE	Percentual da Raiz quadrada do erro quadrático médio (<i>Percentual Root Mean Square Error</i>);
SNR	Relação sinal/ruído (<i>Signal/noise ratio</i>).

Introdução

“No princípio, Deus criou os céus e a terra. A terra era informe e vazia; as trevas cobriam o abismo e o Espírito de Deus pairava sobre as águas. Deus disse: Faça-se a luz! E a luz foi feita.”

Gênesis, 1,1-3

Desde a antiguidade o homem procura compreender os fenômenos naturais na tentativa de prever e controlar suas consequências. No mundo moderno, a necessidade de se controlar algum *sistema* tornou-se ainda mais evidente devido ao grande avanço tecnológico.

O termo *sistema* é utilizado para delimitar a porção do universo que se deseja estudar. Um sistema é definido como sendo uma entidade que manipula um ou mais sinais (entradas) para realizar uma função, de forma a produzir novos sinais (saídas) (Haykin and Van Veen, 2001). Um sistema pode ser *linear* ou *não-linear*. A característica que os diferencia é que *sistemas lineares* satisfazem o *princípio da superposição*,¹ o que não ocorre quando os sistemas são *não-lineares*.

Para que se torne mais simples o estudo e controle de um sistema, é possível representá-lo por meio de um *modelo*. Um *modelo* matemático representa algumas das características observadas no sistema usando relações matemáticas (Aguirre, 2000). O desenvolvimento de técnicas para obter equações matemáticas capazes de reproduzir de forma aproximada o comportamento

¹Uma função $y = f(x)$ satisfaz o princípio da superposição, e portanto é linear, se e somente se $y_1 = f(x_1)$, $y_2 = f(x_2)$ e $ay_1 + by_2 = f(ax_1 + bx_2)$, para quaisquer constantes a e b .

dinâmico de sistemas reais é um dos temas mais fascinantes da cultura humana (Ljung, 1987), e de vital importância para o desenvolvimento da ciência e tecnologia (Corrêa, 2001). Em muitos casos, controlar um determinado sistema seria uma tarefa impossível sem o uso de um modelo matemático.

O controle de velocidade de uma máquina a vapor, desenvolvido por James Watt no século XVIII, é citado como sendo o primeiro trabalho significativo no qual se fez uso da teoria de controle (Ogata, 1993). As técnicas de controle, nos dias de hoje, são aplicadas nos mais diversos ramos da engenharia, como por exemplo em veículos espaciais, aviões, robôs, nos modernos processos industriais e em sistemas de comunicação.

As técnicas da *teoria de controle* se dividem basicamente em dois grupos:

1. controle clássico;
2. controle moderno.

O controle clássico se aplica a sistemas lineares, invariantes no tempo, e com apenas uma entrada e uma saída. Duas técnicas importantes que constituem o controle clássico são os métodos de *resposta em frequência* (Bode, 1945) e *lugar das raízes* (Evans, 1948, 1950).

O controle moderno pode ser aplicado a sistemas de entradas e saídas múltiplas, lineares ou não-lineares, invariantes ou variantes no tempo. A teoria de controle moderno baseia-se nas representações em *espaço de estados*, nas quais surge o conceito de *variáveis de estado* (Chen, 1984). As *variáveis de estado* de um sistema dinâmico constituem o menor conjunto de variáveis que, quando conhecidas em um determinado instante, determinam completamente o comportamento do sistema nos instantes futuros, desde que a entrada também seja conhecida em todos esses instantes.

É importante deixar claro que não é necessário que as variáveis de estado sejam fisicamente mensuráveis. Em muitos casos é difícil medir diretamente todas essas variáveis. Todavia, diversas técnicas de controle moderno requerem o conhecimento completo de todo o espaço de estados. Nesse contexto surge a necessidade de se estimar as variáveis desconhecidas.

Em 1960, Rudolph Emil Kalman, um dos maiores pesquisadores na área de controle moderno, propôs um algoritmo recursivo muito eficiente capaz de estimar variáveis de estado a partir de medições ruidosas de um determinado sistema (Kalman, 1960). No ano seguinte, Kalman e Bucy publicaram outro trabalho sobre a técnica (Kalman and Bucy, 1961). Uma das diferenças

entre esses dois trabalhos é que no primeiro é apresentada uma abordagem para sistemas de tempo discreto, enquanto que no segundo trabalha-se com sistemas contínuos. Atribui-se ao próprio autor do algoritmo a formulação e o estudo de conceitos como *controlabilidade* e *observabilidade* (Kalman et al., 1962), hoje já consolidados e considerados essenciais no estudo e aplicação das técnicas de controle moderno.

Esse algoritmo, amplamente estudado e empregado em diversos tipos de problemas, ficou conhecido como *Filtro de Kalman* (KF), do inglês *Kalman Filter*, ou *Filtro de Kalman-Bucy*, segundo alguns pesquisadores. A motivação para se estudar e dar tanto destaque ao Filtro de Kalman será discutida a seguir.

1.1 Motivação

Durante os anos 50, o grande interesse pelo programa espacial, motivou pesquisadores a buscarem alternativas para as limitações da teoria de Wiener, que foi desenvolvida para lidar com séries temporais estacionárias,² invariantes no tempo (Wiener, 1949). Outra desvantagem desta teoria era o elevado custo computacional. O KF surgiu para solucionar esses problemas, pois além da possibilidade de aplicação em problemas não-estacionários, trata-se de um algoritmo recursivo, o que reduz consideravelmente esforços computacionais.

O KF teve sua primeira grande aplicação prática no projeto Apollo da NASA³ com destaque para o projeto Apollo 11, no qual uma nave espacial tripulada foi levada à lua e trazida de volta à Terra em 1969. Pesquisadores da NASA implementaram o algoritmo do filtro de Kalman no computador de bordo responsável por guiar a nave espacial em sua trajetória. Diz-se que a checagem final do programa foi feita um dia antes dos astronautas Armstrong, Aldrin, e Collins partirem em sua missão (Cipra, 1993).

Depois dessa primeira aplicação, o KF passou a ser amplamente utilizado e se tornou peça fundamental para o sucesso do programa espacial. Como se não bastasse essa grande aplicação, foram encontradas aplicações para o KF nas mais diversas áreas, como por exemplo navegação marítima, instrumentação de usinas nucleares, modelagem demográfica, economia, robótica

²Um processo é dito estacionário quando as leis de probabilidade que o regem não variam com o tempo (Peebles, 1993).

³National Aeronautics and Space Administration.

e processos industriais em geral. O uso do KF passou a ser tão intenso que certa vez Bucy, pioneiro no desenvolvimento do algoritmo ao lado de Kalman, declarou: “*I think sometimes it’s overused!*”⁴ (Cipra, 1993).

A função básica do KF em qualquer aplicação é sempre fornecer estimativas ótimas dos estados do sistema, prever valores futuros de determinadas variáveis ou melhorar a estimativa de variáveis em instantes anteriores. A seguir são listados alguns trabalhos que descrevem aplicações do KF.

No trabalho de Campbell e cooperadores (1983) o KF foi usado na estimação da órbita da nave espacial Voyager até Júpiter. Foi utilizada uma combinação entre as observações obtidas da nave com aquelas obtidas a partir da Terra de forma a se obter uma estimação precisa da órbita.

Mealy e Tang (1983) apresentam os resultados de uma investigação da capacidade de desempenho de um sistema baseado no KF proposto para controle do movimento de um helicóptero em baixas altitudes. Sidar e Doolin (1983) estudam o problema da previsão de movimento usando o KF, aplicado ao problema da decolagem de aviões a partir de navios porta-aviões no mar.

Fung e Grimble (1983) desenvolveram um estimador baseado no KF aplicado no sistema de controle de posicionamento de navios.

Tylee (1983) descreve e avalia uma abordagem funcional *on-line* do KF para detecção de falhas na instrumentação de uma usina nuclear.

No trabalho de Wallace e Clarke (1983) é mostrada uma contribuição sobre a aplicação do KF no controle de sistemas de potência.

Leibundgut e colegas (1983) fazem uso do KF no estudo de modelos demográficos.

Uma técnica para identificação do posicionamento de um veículo na pista, usando uma combinação de informações do ângulo da direção do veículo e imagens externas, usando o KF, foi proposta por Mouri e colaboradores (2002). O objetivo era definir automaticamente qual seria a trajetória do veículo.

A manutenção preventiva é uma técnica cujo objetivo é detectar ou prever uma falha antes que esta venha a ocorrer. No trabalho de Yang (2002) são apresentados resultados experimentais de um método de previsão de falha para manutenção preventiva baseada na estimação de estados de um motor DC, usando o KF. Assimetrias no rotor de máquinas de indução perturbam muitos componentes, produzindo, por exemplo, perturbações no conjugado eletromagnético. O artigo escrito por Eltabach e outros (2002) trata do es-

⁴“Acho que às vezes ele [o algoritmo] é usado de forma exagerada!”

tudo da detecção de defeitos no rotor a partir da análise espectral do torque eletromagnético calculado por dois estimadores de fluxo do rotor.

As aplicações citadas acima contêm apenas alguns dos problemas que podem ser resolvidos utilizando-se o KF. A seguir, discutem-se os objetivos principais deste trabalho.

1.2 Objetivos

Os objetivos desse trabalho são:

1. investigar algoritmos para estimação recursiva, dando ênfase às técnicas baseadas no Filtro de Kalman (Capítulos 2 e 3);
2. aplicar as técnicas investigadas no problema de estimação de estados e parâmetros de sistemas não-lineares (Capítulo 4);
3. mostrar que o Filtro de Kalman pode ser usado em sistemas de comunicação para transmissão de informação utilizando osciladores caóticos (Capítulo 5).

1.3 Justificativa

Algoritmos recursivos são amplamente utilizados na solução de problemas comuns de engenharia, bem como de outras áreas. Na estimação recursiva os dados medidos podem ser utilizados seqüencialmente, ou seja, à medida em que são disponibilizados, o que torna esses algoritmos muito úteis quando os parâmetros do processo sofrem variações. Dessa forma, a aplicação de algoritmos recursivos em controle adaptativo e identificação *on-line* é bastante conveniente (Ljung, 1977).

Muitas técnicas de controle moderno, como por exemplo, controle por realimentação de estados (Ogata, 1993), requerem que todas as variáveis de estado estejam disponíveis durante a realização dos cálculos que determinam as ações de controle. A necessidade de obtenção dos valores dos estados a cada instante de tempo é uma das justificativas para o interesse em se estudar estimadores recursivos eficientes.

O Filtro de Kalman original (Kalman, 1960; Kalman and Bucy, 1961) foi concebido para ser aplicado em sistemas de dinâmica linear. A fim de

tornar possível a aplicação do KF em sistemas não-lineares, pesquisadores propuseram o Filtro de Kalman Estendido (EKF), o qual se mostrou eficiente para muitos sistemas cuja não-linearidade é pouco significativa.

Com relação ao EKF o próprio Bucy fez a seguinte declaração:

*“Researchers have developed a method, called the extended Kalman filter, that adjusts a linear approximation at each filter update. That works on some problems where the nonlinear effects are small. However, there are problems where this is not the case, and people use the filter and they have troubles with it. They try all sorts of fixes, but basically the problem is such that the linear theory does not apply.”*⁵ (Richard Bucy) (Cipra, 1993).

Os problemas do mundo real são quase sempre de natureza não-linear, justificando a busca constante por técnicas que se aplicam a essa classe de sistemas. Pelos motivos expostos acima, busca-se neste trabalho, investigar além do conhecido EKF, uma abordagem mais eficiente que torna possível a aplicação do KF em sistemas cuja dinâmica é bastante influenciada pela presença de elementos não-lineares.

A segurança em transmissão de informação tem sido uma questão crítica no mundo moderno. Os novos avanços tecnológicos trouxeram uma infinidade de situações nas quais informações críticas precisam ser protegidas do acesso comum. Por este motivo justifica-se a crescente necessidade de se transmitir informação de forma segura através de um meio inerentemente inseguro. Visando atingir esse objetivo, neste trabalho será mostrado que é possível transmitir informação usando sinais caóticos perturbados, o que supostamente torna mais difícil o acesso não autorizado à informação.

⁵“Pesquisadores desenvolveram um método, chamado Filtro de Kalman Estendido, que faz uma aproximação linear do sistema a cada etapa de atualização. Isso funciona em alguns problemas nos quais os efeitos não-lineares são pequenos. Entretanto, existem problemas nos quais esse não é o caso, e por isso não se tem sucesso quando se tenta utilizar o filtro nessas situações. Tenta-se todo tipo de ajustes, mas o problema é que a teoria linear não se aplica nesses casos.”

1.4 Organização do Texto

Esta dissertação está organizada da seguinte forma:

No capítulo 2 é feita uma revisão bibliográfica sobre alguns estimadores recursivos, com destaque para o Filtro de Kalman. Apresentam-se alguns conceitos importantes para o bom entendimento da representação de sistemas por meio de modelos estocásticos.

No capítulo 3 são investigados dois algoritmos, baseados no KF, para estimação em sistemas não-lineares. Procura-se destacar as vantagens de um método em relação ao outro, de acordo com o sistema em que são aplicados.

O capítulo 4 traz resultados de filtragem e estimação de estados de uma plataforma de laboratório. Realiza-se uma comparação entre os resultados obtidos utilizando-se os dois algoritmos abordados no capítulo 3.

No capítulo 5 mostra-se que é possível transmitir informação de forma segura através de um meio considerado inseguro, usando um dos algoritmos investigados no capítulo 3.

No capítulo 6 são apresentadas as conclusões do trabalho e propostas para pesquisas futuras.

O Apêndice traz uma breve biografia sobre R. E. Kalman.

Estimação Recursiva

“Ainda acredito na possibilidade de construirmos um modelo da realidade.”

Albert Einstein

2.1 Introdução

A *estimação de parâmetros* constitui uma das etapas dos procedimentos de *Identificação de Sistemas*. A estrutura do modelo é o conjunto de variáveis regressoras a ser usado. Após selecionada a estrutura é possível calcular os parâmetros, por meio de um algoritmo apropriado, a partir de um conjunto de dados coletados do sistema. Existem basicamente duas classes de técnicas que podem ser usadas para se estimar os parâmetros de um modelo. São elas:

- *estimação em batelada*;
- *estimação recursiva*.

Nas técnicas de *estimação em batelada* todo o conjunto de dados deve estar disponível antes de se iniciar o processo de estimação, pois o vetor de parâmetros é calculado em um único conjunto de operações.

Na *estimação recursiva* os dados medidos podem ser utilizados seqüencialmente, ou seja, à medida que são disponibilizados, atualizando constantemente o vetor de parâmetros do modelo. A *estimação recursiva* é muito útil quando os parâmetros do processo variam lentamente, como por exemplo,

em função de não-linearidades decorrentes de desgastes e falhas. Essa classe de técnicas também pode ser usada na solução de problemas numéricos que são de difícil solução quando resolvidos *em batelada*. Este trabalho trata especificamente de algoritmos para estimação *recursiva*.

O algoritmo que realiza a estimação é conhecido na literatura como *estimador*. Existem diversos algoritmos capazes de realizar estimação recursiva, como por exemplo, *atualização recursiva*, o *estimador recursivo de mínimos quadrados*, o *estimador recursivo estendido de mínimos quadrados*, o *estimador recursivo de variáveis instrumentais*, dentre outros (Ljung, 1987; Aguirre, 2000). Neste capítulo são apresentados dois exemplos de *estimadores recursivos* que podem ser aplicados em problemas de identificação: atualização recursiva e o estimador recursivo de mínimos quadrados.

Em determinadas aplicações, como em controle de processos via realimentação de estados por exemplo, surge a necessidade de que todas as variáveis de estado do sistema estejam acessíveis ou disponíveis para medição, o que na prática é impossível em muitos casos. Em situações como essas torna-se necessário estimar os estados desconhecidos. Isso pode ser feito por meio de *estimadores recursivos* da mesma forma que parâmetros são estimados em problemas de identificação. O uso de estimadores recursivos na estimação de variáveis de estado é bastante conveniente, pois estes podem, em princípio, ser implementados em sistemas de tempo real. Um estimador recursivo bastante eficiente, capaz de estimar as variáveis de estado de sistemas dinâmicos lineares representados por modelos em espaço de estados, é o *Filtro de Kalman* (KF) (Kalman, 1960).

Neste trabalho será dada maior ênfase à estimação de variáveis de estado. Neste capítulo apresenta-se o *Filtro de Kalman* como uma técnica utilizada na estimação recursiva de estados para sistemas lineares, enquanto que no próximo capítulo serão mostradas duas de suas variações empregadas quando o sistema em questão apresenta dinâmica não-linear.

2.2 Modelos Estocásticos e Abordagem Bayesiana

Os *modelos determinísticos* são aqueles nos quais não são modeladas incertezas de nenhuma natureza. Nos *modelos estocásticos* as incertezas são modeladas na forma de variáveis aleatórias (Aguirre, 2000).

Nesta seção procura-se justificar a suposição de que é preferível trabalhar com modelos estocásticos ao invés de determinísticos. Com o objetivo de facilitar o entendimento das técnicas a serem apresentadas ao longo do trabalho, nesta seção também são introduzidos alguns conceitos importantes em *processos estocásticos*, além de ser apresentada uma abordagem muito utilizada nessa área: a *abordagem Bayesiana*.

Quando se deseja investigar o comportamento de um sistema real, observando como ele responde a determinados tipos de entradas, é freqüentemente útil desenvolver um modelo matemático que represente adequadamente alguns aspectos desse comportamento. Tal modelagem pode ser feita tomando-se como base as teorias determinísticas de forma tal que o modelo obtido é *determinístico*, ou então, a partir das técnicas de análise de processos estocásticos, o que resulta em um modelo *estocástico*.

Ao se fazer análise de sistemas ou projeto de controladores, seria mais simples utilizar modelos determinísticos, visto que nesses não são incluídos termos referentes às fontes de incerteza. Surge a partir de então uma pergunta: Por que aumentar o nível de complexidade do problema propondo modelos estocásticos? Para responder a esta pergunta deve-se fazer um exame das teorias determinísticas, o que torna possível perceber onde estas podem apresentar falhas.

Existem pelo menos três razões importantes indicando que modelos determinísticos não são representações totalmente adequadas e suficientes para se fazer análise de sistemas. São elas (Maybeck, 1979):

- *nenhum modelo matemático é perfeito*. Qualquer modelo representa somente as características que forem de interesse, ou seja, de forma a cumprir apenas o objetivo para o qual foi desenvolvido, o que é positivo em se tratando da necessidade de se manter uma baixa complexidade. As leis físicas proporcionam muitas vezes estruturas adequadas para os modelos, embora vários parâmetros não possam ser facilmente determinados. Desta forma, verifica-se que existem muitas fontes de incerteza no processo de modelagem;
- *a dinâmica dos sistemas não é determinada somente pelas entradas de controle*. É também influenciada por distúrbios que não podem ser controlados nem modelados de maneira determinística. O somatório desses distúrbios formam o chamado *ruído de processo*;

- *a precisão de qualquer sistema de medição é limitada.* A falta de precisão de alguns instrumentos de medição podem proporcionar medidas inexatas das variáveis desejadas, contaminando o processo com sinais não determinísticos. O conjunto desses sinais constitui o *ruído de medição*.

2.2.1 Representação de Sistemas por Modelos Estocásticos

Com base nos pontos discutidos acima fica claro que é bastante conveniente que se proponha o desenvolvimento de modelos estocásticos, pois estes apresentam em suas equações os termos referentes ao ruído de processo e ruído de medição, mencionados no parágrafo anterior:

$$\text{Tempo discreto} \quad \begin{cases} \mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{w}_k \\ \mathbf{y}_k = h(\mathbf{x}_k) + \mathbf{e}_k, \end{cases} \quad (2.1)$$

$$\text{Tempo contínuo} \quad \begin{cases} \dot{\mathbf{x}} = f(\mathbf{x}) + \mathbf{w} \\ \mathbf{y} = h(\mathbf{x}) + \mathbf{e}. \end{cases} \quad (2.2)$$

Os sistemas (2.1) e (2.2) representam processos estocásticos multidimensionais, sendo \mathbf{w}_k e \mathbf{w} os vetores de ruído de processo e \mathbf{e}_k e \mathbf{e} os vetores de ruído de medição.

É importante salientar que a análise das soluções destes sistemas ainda constitui objeto de pesquisa em matemática na atualidade, principalmente em se tratando de sistemas contínuos. Apesar deste fato, já é possível obter resultados promissores quanto a estimação de parâmetros e/ou estados mediante a adoção de hipóteses simplificadoras, tais como comportamento linear aproximado em torno do ponto de operação e funções de distribuição gaussiana para os processos estocásticos representados pelo ruído de processo e pelo ruído de medição.

2.2.2 Conceitos Básicos em Processos Estocásticos

O *filtro de Kalman*, algoritmo recursivo ao qual se dará maior enfoque neste trabalho, aborda o problema geral de se tentar estimar os estados de um sistema representado por um modelo *estocástico*. Na seção anterior justificou-se a preferência por se trabalhar com modelos *estocásticos*. Neste

ponto torna-se importante apresentar alguns conceitos básicos em processos estocásticos a fim de proporcionar uma boa compreensão dos algoritmos recursivos.

Média e Covariância

Uma determinada variável aleatória vetorial $\mathbf{x} \in \mathbb{R}^n$ tem o seu *valor médio*, ou *valor esperado*, dado por (Papoulis, 1984):

$$E[\mathbf{x}] = \bar{\mathbf{x}} = \int_{\mathbb{R}^n} \mathbf{x} f(\mathbf{x}) d\mathbf{x}, \quad (2.3)$$

sendo f a função densidade de probabilidade de \mathbf{x} .

O valor médio de uma variável aleatória é usado no cálculo de *momentos*. Os momentos em torno do *valor médio* de \mathbf{x} são conhecidos como *momentos centrais* e podem ser representados por:

$$E[(\mathbf{x} - \bar{\mathbf{x}})^n], \quad (2.4)$$

sendo n a ordem do momento. O momento central de segunda ordem é muito importante na área de *processos estocásticos* e recebe o nome de *variância*, no caso unidimensional, ou *covariância*, no caso multidimensional.

Desta forma, a matriz de covariância de um vetor $\mathbf{x} \in \mathbb{R}^n$, tem dimensão $n \times n$ e é definida por (Papoulis, 1984):

$$\text{cov}[\mathbf{x}] = E[(\mathbf{x} - E[\mathbf{x}])(\mathbf{x} - E[\mathbf{x}])^T], \quad (2.5)$$

sendo \mathbf{x}^T um vetor linha, resultado da transposição do vetor coluna \mathbf{x} .

Variáveis Gaussianas

Uma variável aleatória x é chamada de *gaussiana*¹ ou *normal* quando sua *função densidade de probabilidade*² (PDF), do inglês *probability density function*, é expressa por (Peebles, 1993):

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-(x-\bar{x})^2/2\sigma_x^2}, \quad (2.6)$$

¹Em homenagem ao matemático alemão Johann Friedrich Carl Gauss (1777-1855).

²A *função densidade de probabilidade* é definida como a derivada da *função distribuição de probabilidade*.

sendo que $\sigma_x = \sqrt{\text{cov}[x]}$ e \bar{x} são constantes reais, representando respectivamente o desvio padrão e a média da variável x .

A PDF de uma variável gaussiana apresenta o formato de um sino conforme pode ser visto na Figura 2.1.

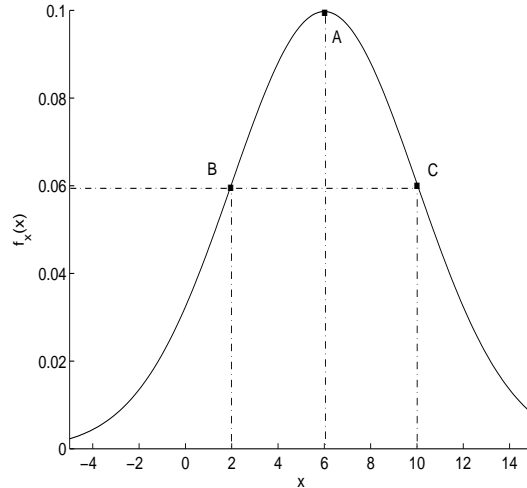


Figura 2.1: Função densidade de probabilidade gaussiana. Esta curva foi gerada com $\bar{x} = 6$, $\sigma_x = 4$, e x variando no intervalo de -5 a 15 . O ponto A corresponde ao valor de \bar{x} no eixo x e à constante $\frac{1}{\sqrt{2\pi\sigma_x^2}}$ no eixo f_x . Os pontos B e C equivalem à constante $\frac{0,607}{\sqrt{2\pi\sigma_x^2}}$ no eixo f_x , e no eixo x , correspondem respectivamente a $(\bar{x} - \sigma_x)$ e $(\bar{x} + \sigma_x)$.

A razão pela qual se opta por representar o ruído usando-se uma variável gaussiana é de origem física: o ruído geralmente é causado por um somatório de pequenas fontes de perturbação. Pode ser demonstrado matematicamente³ que quando uma certa quantidade de variáveis aleatórias são somadas, o efeito global pode ser representado por uma PDF gaussiana (Kennedy and Neville, 1986). Nenhum ruído real pode ser verdadeiramente gaussiano pelo fato de todo ruído ter amplitude limitada, ao contrário do que exige a distribuição gaussiana. Entretanto, a distribuição gaussiana ainda é uma boa aproximação pelas razões indicadas.

A representação da maior parte das funções densidade de probabilidade exige um número infinito de momentos, a fim de que o formato da curva seja

³Teorema do limite central (Peebles, 1993).

completamente determinado. Entretanto, para que se possa representar uma PDF gaussiana são necessários apenas os momentos de primeira e de segunda ordem, ou seja, média e covariância. Esse fato é de grande importância no procedimento do KF, conforme poderá ser constatado no decorrer deste trabalho. É possível atribuir valores aproximados para \bar{x} e σ_x em situações reais. Por essa razão, a representação de ruído por meio de uma PDF gaussiana é comum em problemas de engenharia.

O Ruído Branco

O termo *branco* implica que os valores do ruído não são correlacionados no tempo, ou seja, a partir do registro histórico do ruído, não é possível prever qual será o próximo valor. Um ruído branco apresenta potência espectral teoricamente constante em todas as frequências, o que implica um sinal com potência infinita. Como a potência de qualquer sinal é limitada, ruído branco não existe na realidade.

A razão por se trabalhar com ruído branco é que qualquer sistema tem uma certa faixa de passagem de frequências de interesse, ou seja, uma faixa na qual ele pode responder às entradas aplicadas. Fora dessa faixa, a entrada não produz efeito significativo na saída. Dessa forma, dentro da faixa de passagem de interesse, o ruído pode ser considerado branco, pois apresenta potência constante em todas as frequências. Esta hipótese traduz-se, na prática, em maior facilidade na manipulação das equações que descrevem o sistema estocástico.

2.2.3 A Abordagem Bayesiana

A aplicação de métodos estatísticos em problemas de engenharia pode ser feita por meio de dois tipos de abordagem distintos:

1. estatística clássica (ou frequentista);
2. estatística Bayesiana.

A estatística clássica ignora quaisquer informações existentes antes da realização de um determinado experimento. Já a estatística *Bayesiana* leva em consideração o conhecimento *a priori*,⁴ na forma de funções de densidade de probabilidade (PDF).

⁴Informação disponível antes da realização do experimento.

Embora bastante diferentes, ambas as abordagens têm seus campos específicos de aplicação, o que quer dizer que não se deve esperar que uma seja melhor do que a outra. Todavia, algumas observações podem ser feitas com respeito a vantagens e desvantagens de cada abordagem (Bekman and Costa, 1980):

- alguns dos procedimentos da estatística clássica só apresentam bons resultados com grandes amostras, ou seja, muitas observações; esta exigência não é restrição para os procedimentos Bayesianos;
- por outro lado, em vários casos, a abordagem Bayesiana esbarra na dificuldade matemática oriunda de seu procedimento, como por exemplo, na especificação da PDF inicial (Jaynes, 1996).

O Teorema de Bayes

A base da teoria Bayesiana pode ser representada pelo *Teorema de Bayes*.⁵ O teorema se baseia em alguns conceitos estatísticos, como *probabilidade condicional* e *probabilidade total*.

A probabilidade condicional de um evento B_n , dado A , sendo que o índice n se refere ao n -ésimo evento mutuamente exclusivo, é expressa por:

$$P(B_n|A) = \frac{P(B_n \cap A)}{P(A)}, \quad (2.7)$$

sendo que $B_n \cap A$ representa a intersecção entre B_n e A , e $P(A) \neq 0$. De forma análoga, pode-se escrever:

$$P(A|B_n) = \frac{P(A \cap B_n)}{P(B_n)}, \quad (2.8)$$

sendo $P(B_n) \neq 0$. Os termos $P(B_n|A)$ e $P(A|B_n)$ indicam que a probabilidade do primeiro evento dentro dos parênteses, depende do segundo evento.

A *probabilidade total* de um evento A é dada por:

$$P(A) = \sum_{n=1}^N P(A|B_n)P(B_n). \quad (2.9)$$

⁵Teorema proposto pelo teólogo e matemático inglês Thomas Bayes (1702-1761).

A partir das equações (2.7) e (2.8) obtém-se uma das formas do *Teorema de Bayes* (Peebles, 1993):

$$P(B_n|A) = \frac{P(A|B_n)P(B_n)}{P(A)}. \quad (2.10)$$

Substituindo a equação (2.9) em (2.10) pode-se obter uma outra representação deste teorema (Peebles, 1993; Bekman and Costa, 1980):

$$P(B_n|A) = \frac{P(A|B_n)P(B_n)}{\sum P(A|B_n)P(B_n)}. \quad (2.11)$$

O Teorema de Bayes calcula novas probabilidades de ocorrência dos eventos B_n em função do conhecimento adquirido, representado pela realização do evento A . O termo $P(B_n)$ representa a probabilidade *a priori*, ou seja, a probabilidade de ocorrência de B_n sem considerar o evento A . A expressão $P(B_n|A)$ representa a probabilidade *a posteriori*, que pode ser entendida como sendo a probabilidade de ocorrência dos eventos B_n levando-se em conta a ocorrência do evento A .

A partir do Teorema de Bayes, pode-se chegar a duas conclusões importantes (Bekman and Costa, 1980):

- a probabilidade não representa um estado das coisas (ou da natureza), mas sim um estado de conhecimento sobre o sistema;
- a partir do Teorema de Bayes é possível desenvolver toda a teoria da *inferência estatística*,⁶ uma vez que se pode atualizar as atribuições de probabilidade sempre que forem obtidas novas informações relevantes.

Abordagem Bayesiana na Estimação Recursiva

Os problemas de estimação recursiva podem ser tratados em uma abordagem Bayesiana, como será visto no decorrer deste trabalho. Em problemas desse tipo ocorrem sucessivas atualizações na PDF das variáveis que estão sendo estimadas, à medida em que novos dados medidos são disponibilizados. Esse procedimento será visto com maior clareza no algoritmo do KF, na seção 2.4.2.

⁶Inferência estatística pode ser entendida como sendo o processo de raciocínio pelo qual se procura tirar conclusões sobre o universo de interesse a partir do conhecimento de amostras.

As atualizações da PDF trazem novas informações sobre o sistema, as quais podem ser incorporadas ao modelo a partir do Teorema de Bayes. Um estimador recursivo genérico, na abordagem Bayesiana, é implementado em duas etapas (Arulampalam et al., 2002):

1. Predição ou propagação temporal da PDF.

Inicialmente conhece-se

$$P(\mathbf{x}_{k-1} | \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}). \quad (2.12)$$

Considerando que $(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{k-1})$ é uma seqüência de Markov de primeira ordem (isto é, $P(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{x}_{k-2}, \dots, \mathbf{x}_0) = P(\mathbf{x}_k | \mathbf{x}_{k-1})$), é possível calcular a nova PDF, resultado da evolução temporal do sistema, a partir da equação de Chapman-Kolmogorov (Papoulis, 1984):

$$P(\mathbf{x}_k | \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) = \int P(\mathbf{x}_k | \mathbf{x}_{k-1}) P(\mathbf{x}_{k-1} | \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) d\mathbf{x}_{k-1}. \quad (2.13)$$

Neste caso, usou-se o fato de que

$$P(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) = P(\mathbf{x}_k | \mathbf{x}_{k-1}).$$

2. Correção ou atualização.

Após a obtenção da medição \mathbf{y}_k , usando-se o Teorema de Bayes, é possível escrever que:

$$\begin{aligned} P(\mathbf{x}_k | \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}, \mathbf{y}_k) &= \frac{P(\mathbf{y}_k | \mathbf{x}_k) P(\mathbf{x}_k | \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1})}{P(\mathbf{y}_k | \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1})}, \\ &= \frac{P(\mathbf{y}_k | \mathbf{x}_k) P(\mathbf{x}_k | \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1})}{\int P(\mathbf{y}_k | \mathbf{x}_k) P(\mathbf{x}_k | \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) d\mathbf{x}_k}. \end{aligned} \quad (2.14)$$

Salienta-se que $P(\mathbf{x}_k | \mathbf{x}_{k-1})$ na equação (2.13) será determinada pela equação dinâmica do sistema (2.1a) e pelas características estatísticas do ruído de processo. Já $P(\mathbf{y}_k | \mathbf{x}_k)$ na equação (2.14) será determinada pela equação de medição (2.1b) e pelas características estatísticas do ruído de medição.

2.3 Estimadores Recursivos para Identificação de Sistemas

A Identificação de Sistemas é o estudo dos procedimentos que permitem desenvolver modelos matemáticos a partir de dados e sinais observados (Aguirre, 2000). As etapas do processo de identificação são as seguintes:

1. testes dinâmicos e coleta de dados;
2. escolha da representação matemática;
3. determinação da estrutura do modelo;
4. estimação de parâmetros;
5. validação do modelo.

Após escolhida a representação matemática e determinada a estrutura do modelo, torna-se necessário estimar os parâmetros do modelo a partir dos dados coletados. Essa estimação pode ser feita usando-se estimadores recursivos. Um dos estimadores recursivos mais importantes e mais utilizados em aplicações práticas é o estimador *recursivo de mínimos quadrados*.

Antes de se apresentar este estimador, será mostrado um algoritmo geral para estimação recursiva conhecido como *atualização recursiva* (Norton, 1986). Os estimadores recursivos apresentados nesta seção, serão abordados na visão dos princípios da estatística *clássica* ou *freqüentista*.

2.3.1 Atualização Recursiva

Atualização recursiva é um algoritmo capaz de estimar recursivamente os parâmetros de um modelo pré-definido. Para que se possa obter as equações desse algoritmo deve-se considerar um determinado modelo para o sistema a ser identificado: $y(k) = \boldsymbol{\psi}^T \boldsymbol{\theta} + e(k)$, sendo $y(k)$ o sinal medido, $\boldsymbol{\psi}$ o vetor de regressores, $\boldsymbol{\theta}$ o vetor de parâmetros e $e(k)$ é o ruído de medição. As propriedades estatísticas do ruído são: $E[e(k)] = 0$ e $\text{cov}[e(k)] = R$, sendo $E[\cdot]$ o operador esperança matemática e $\text{cov}[\cdot]$ a covariância.

A partir de uma estimativa $\hat{\boldsymbol{\theta}}_k$ para os parâmetros, este sistema pode ser representado pela seguinte equação:

$$y(k) = \boldsymbol{\psi}_k^T (k-1) \hat{\boldsymbol{\theta}}_k + \xi(k), \quad (2.15)$$

sendo que $y(k)$ é a saída medida, $\boldsymbol{\psi}_k(k-1)$ é o vetor de regressores formado na iteração k com informação disponível até a iteração $k-1$, $\hat{\boldsymbol{\theta}}_k$ é o vetor de parâmetros estimados e $\xi(k)$ é o valor do resíduo resultante da estimação.

O princípio básico dos estimadores recursivos é o cálculo de grandezas em um determinado instante k em função de valores obtidos no passado. Desta maneira, torna-se necessário expressar $\hat{\boldsymbol{\theta}}_k$ em função do último valor estimado $\hat{\boldsymbol{\theta}}_{k-1}$ de forma tal que a nova estimativa incorpore informação atualizada obtida da medição $y(k)$. A expressão obtida é a seguinte:

$$\hat{\boldsymbol{\theta}}_k = J_k \hat{\boldsymbol{\theta}}_{k-1} + K_k y(k). \quad (2.16)$$

As matrizes J_k e K_k devem ser determinadas garantindo que $\hat{\boldsymbol{\theta}}_k$ seja uma boa estimativa sob dois pontos de vista:

1. $\hat{\boldsymbol{\theta}}_k$ deve ser não polarizado, isto é, $\lim_{k \rightarrow \infty} \hat{\boldsymbol{\theta}}_k = \boldsymbol{\theta}$.
2. $\text{cov}[\hat{\boldsymbol{\theta}}_k]$ deve ser tão pequena quanto possível.

Levando-se em consideração as duas restrições citadas acima é possível, após algumas manipulações algébricas, obter-se um algoritmo recursivo para a estimação do vetor $\boldsymbol{\theta}$ (Aguirre, 2000):

$$\left. \begin{aligned} K_k &= P_{k-1} \boldsymbol{\psi}_k (\boldsymbol{\psi}_k^T P_{k-1} \boldsymbol{\psi}_k + R)^{-1}; \\ \hat{\boldsymbol{\theta}}_k &= \hat{\boldsymbol{\theta}}_{k-1} + K_k (y(k) - \boldsymbol{\psi}_k^T \hat{\boldsymbol{\theta}}_{k-1}); \\ P_k &= (I - K_k \boldsymbol{\psi}_k^T) P_{k-1} (I - K_k \boldsymbol{\psi}_k^T)^T + K_k R K_k^T. \end{aligned} \right\} \quad (2.17)$$

Neste algoritmo P_k é a matriz de covariância do vetor de parâmetros estimados: $P_k = \text{cov}[\hat{\boldsymbol{\theta}}_k]$ e R é a matriz de covariância do ruído: $R = \text{cov}[e(k)]$. Os elementos da diagonal principal da matriz de covariância P_k são as variâncias dos respectivos elementos no vetor de parâmetros e, portanto, indicam o grau de confiança que se tem nos valores estimados dos elementos de $\hat{\boldsymbol{\theta}}_k$.

O algoritmo recursivo (2.17) precisa ser iniciado antes de se realizar a primeira iteração. Para que a escolha da condição inicial $\hat{\boldsymbol{\theta}}_0$ não afete significativamente a qualidade da estimação, os elementos de P_0 devem ser iniciados com valores suficientemente elevados (Aguirre, 2000). Desta forma, uma boa maneira de se inicializar o algoritmo recursivo é usar um vetor de parâmetros nulo e uma matriz de covariância diagonal com elementos de valor elevado.

2.3.2 O Estimador Recursivo de Mínimos Quadrados

O *estimador recursivo de mínimos quadrados*, como o próprio nome indica, se baseia no *estimador de mínimos quadrados*.⁷

O estimador de mínimos quadrados pode ser expresso pela seguinte equação:

$$\hat{\boldsymbol{\theta}} = \left[\sum_{i=1}^k \boldsymbol{\psi}(i-1)\boldsymbol{\psi}^T(i-1) \right]^{-1} \left[\sum_{i=1}^k \boldsymbol{\psi}(i-1)y(i) \right], \quad (2.18)$$

sendo que,

$$P_k = \left[\sum_{i=1}^k \boldsymbol{\psi}(i-1)\boldsymbol{\psi}^T(i-1) \right]^{-1}, \quad (2.19)$$

$$\begin{aligned} P_k^{-1} &= \left[\sum_{i=1}^{k-1} \boldsymbol{\psi}(i-1)\boldsymbol{\psi}^T(i-1) \right] + \boldsymbol{\psi}(k-1)\boldsymbol{\psi}^T(k-1) \\ &= P_{k-1}^{-1} + \boldsymbol{\psi}(k-1)\boldsymbol{\psi}^T(k-1) \end{aligned} \quad (2.20)$$

Fazendo-se as substituições expressas nas equações (2.19) e (2.20), pode-se reescrever a equação (2.18) da seguinte maneira:

$$\hat{\boldsymbol{\theta}}_k = P_k \left[\sum_{i=1}^{k-1} \boldsymbol{\psi}(i-1)y(i) + \boldsymbol{\psi}(k-1)y(k) \right]. \quad (2.21)$$

Após algumas operações algébricas é possível obter uma expressão para cálculo do vetor de parâmetros estimados $\hat{\boldsymbol{\theta}}_k$, de forma recursiva:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_k &= P_k [P_{k-1}^{-1} \hat{\boldsymbol{\theta}}_{k-1} + \boldsymbol{\psi}(k-1)y(k)] \\ &= \hat{\boldsymbol{\theta}}_{k-1} + P_k \boldsymbol{\psi}(k-1) [y(k) - \boldsymbol{\psi}^T(k-1) \hat{\boldsymbol{\theta}}_{k-1}] \\ &= \hat{\boldsymbol{\theta}}_{k-1} + K_k \eta(k), \end{aligned} \quad (2.22)$$

sendo $K_k = P_k \boldsymbol{\psi}(k-1)$ uma matriz de ganho e $\eta(k) = y(k) - \boldsymbol{\psi}^T(k-1) \hat{\boldsymbol{\theta}}_{k-1}$ o erro de predição cometido pelo modelo ao predizer o valor medido correspondente ao instante k , usando o vetor de parâmetros conhecido no instante anterior. Este termo é conhecido como *inovação*.

Com o intuito de evitar a inversão da matriz na equação (2.19), a cada

⁷Este algoritmo é utilizado para realizar estimação de parâmetros em batelada.

iteração do algoritmo, pode-se expressar P_k de outra forma:⁸

$$P_k = P_{k-1} - P_{k-1} \boldsymbol{\psi}(k-1) [\boldsymbol{\psi}^T(k-1) P_{k-1} \boldsymbol{\psi}(k-1) + 1]^{-1} \boldsymbol{\psi}^T(k-1) P_{k-1}. \quad (2.23)$$

A partir da equação (2.20) obtém-se K_k :

$$K_k = \frac{P_{k-1} \boldsymbol{\psi}(k-1)}{\boldsymbol{\psi}^T(k-1) P_{k-1} \boldsymbol{\psi}(k-1) + 1}. \quad (2.24)$$

Organizando as equações obtidas e observando que $\boldsymbol{\psi}_k = \boldsymbol{\psi}(k-1)$ obtém-se o algoritmo do estimador recursivo de mínimos quadrados:

$$\left. \begin{aligned} K_k &= \frac{P_{k-1} \boldsymbol{\psi}_k}{\boldsymbol{\psi}_k^T P_{k-1} \boldsymbol{\psi}_k + 1} \\ \hat{\boldsymbol{\theta}}_k &= \hat{\boldsymbol{\theta}}_{k-1} + K_k [y(k) - \boldsymbol{\psi}_k^T \hat{\boldsymbol{\theta}}_{k-1}] \\ P_k &= P_{k-1} - K_k \boldsymbol{\psi}_k^T P_{k-1} \end{aligned} \right\}. \quad (2.25)$$

2.4 O Filtro de Kalman

Na seção 2.3, estimadores recursivos foram apresentados como ferramentas capazes de estimar parâmetros de modelos em problemas de identificação. Nesta seção, introduz-se uma outra classe de problemas, na qual o algoritmo estimador passa a ter uma função um pouco diferente. O objetivo agora é estimar as *variáveis de estado*⁹ de sistemas representados por modelos em *espaço de estados*. Tal procedimento é muito útil em situações nas quais é desejável que todas as variáveis de estado do modelo estejam disponíveis. A medição direta de todos os estados é uma tarefa impossível em muitas aplicações reais, fazendo com que se torne necessário estimá-los. Essa estimação pode ser realizada por *estimadores recursivos* apropriados, de maneira análoga à estimação de parâmetros realizada em procedimentos de Identificação de Sistemas.

Os problemas de estimação de estados geralmente são abordados de três maneiras distintas, e para cada uma dessas abordagens é usado um termo diferente na literatura. Para que se possa fazer uma distinção entre esses termos é interessante considerar a seguinte situação. Um sinal $y_k = y_k^i + e_k$, sendo y_k^i uma parcela ideal sem ruído e e_k uma parcela de ruído, é medido de

⁸Esta representação é obtida aplicando-se o lema da inversão (Aguirre, 2000).

⁹Neste trabalho as *variáveis de estado* de um modelo também serão chamadas simplesmente de *estados* ou *variáveis*.

forma que todos os valores y_0, \dots, y_k são conhecidos. A partir dessa medição, o que se pode inferir com relação ao valor do sinal em um determinado instante k_a , que pode ser menor, igual ou maior que k ? A determinação do instante até o qual se deseja estimar dá origem a três termos diferentes, normalmente encontrados na literatura quando o assunto é *estimação de estados*:

1. $k_a < k$. Deseja-se estimar os valores apenas do *passado* dos estados. Este é um problema de *interpolação*.
2. $k_a = k$. Estimação de valores do *presente*. O estimador, nesses casos, é chamado de *filtro*.
3. $k_a > k$. Quando necessita-se dos valores dos estados em instantes futuros, usa-se o termo *predição*.

O algoritmo recursivo que será introduzido nesta seção é conhecido na literatura como sendo um *filtro*. Entretanto, sabe-se que ele é capaz de realizar estimação de valores dos estados nos instantes presente, passado e futuro. Tal procedimento foi proposto pela primeira vez em um importante trabalho de um dos maiores pesquisadores da área de controle: Rudolph Emil Kalman.¹⁰

Em seu trabalho, Kalman descreve uma nova abordagem para problemas de filtragem linear e predição (Kalman, 1960). Por este motivo o algoritmo recebe o nome de *Filtro de Kalman* (KF). Ele é definido como um algoritmo recursivo *ótimo* capaz de estimar os estados de um sistema dinâmico linear¹¹ perturbado por ruído (de processo e de medição) branco e gaussiano.¹² O termo *ótimo*, neste caso, significa que o KF calcula as estimativas para as variáveis minimizando o valor esperado da soma quadrática dos resíduos de estimação. Será visto que esta minimização equivale a minimizar a matriz de covariância dos estados.

Na seção 2.4.1 será feita uma breve descrição do algoritmo, mostrando como Kalman obteve uma nova solução para a classe de problemas em questão.

¹⁰No apêndice A, encontra-se uma breve biografia de Kalman.

¹¹No capítulo 3 serão apresentados algoritmos baseados no KF próprios para aplicação em sistemas dinâmicos não-lineares.

¹²Uma rápida explicação sobre esses modelos de ruído pode ser encontrada nas seções 2.2.1 e 2.2.2.

2.4.1 O Algoritmo

A função do Filtro de Kalman é estimar as variáveis de estado de sistemas lineares a partir de medições contaminadas por ruído. Para facilitar a compreensão do algoritmo, ele será apresentado considerando-se um sistema unidimensional (Joseph, 1999), ou seja, tem-se apenas uma variável a estimar. Dessa forma, as grandezas envolvidas deixam de ser vetores, passando a ser escalares.

Deseja-se estimar os valores de uma certa variável aleatória escalar x_k , a qual satisfaz a seguinte equação linear:

$$x_k = a x_{k-1} + w_{k-1}, \quad (2.26)$$

sendo que a tem valor conhecido e w_{k-1} é ruído branco, gaussiano, de média zero e variância q .

O primeiro passo é inicializar o algoritmo de forma adequada. Considera-se que a melhor estimativa inicial para x é conhecida, e será representada por \hat{x}_0 . Entretanto, supõe-se que haja um erro entre o valor real de x_0 e o valor estimado \hat{x}_0 , cuja variância pode ser expressa por:

$$p_0 = E[x_0 - \hat{x}_0]^2. \quad (2.27)$$

De posse das condições iniciais, é possível calcular uma estimativa para a amostra seguinte, que será designada por \tilde{x}_1 . De acordo com a equação (2.26), verifica-se que o valor real dessa amostra é:

$$x_1 = a x_0 + w_0. \quad (2.28)$$

Segundo Kalman, a melhor estimativa para x_1 é:

$$\tilde{x}_1 = a \hat{x}_0. \quad (2.29)$$

É possível constatar que essa é realmente a melhor estimativa, levando-se em conta o conjunto de informações que se tem disponíveis até então, pois não há como atribuir outro valor a w_0 que não seja a média, definida como sendo nula.

A variância do erro de estimação de \tilde{x}_1 é dada por:

$$\begin{aligned}\tilde{p}_1 &= E[x_1 - \tilde{x}_1]^2, \\ &= E[a x_0 + w_0 - a \hat{x}_0]^2, \\ &= a^2 E[x_0 - \hat{x}_0]^2 + E[w_0]^2 + 2a E[(x_0 - \hat{x}_0)w_0].\end{aligned}$$

Como não há correlação entre w e x , o último termo na equação anterior é nulo e a expressão simplificada é:

$$\tilde{p}_1 = a^2 p_0 + q. \quad (2.30)$$

Na etapa seguinte, obtém-se uma medida da variável x , contaminada por ruído, conforme a seguinte equação:

$$y_1 = c x_1 + e_1, \quad (2.31)$$

sendo que y_1 é a amostra do sinal medido, c é uma constante conhecida e e_1 ruído branco, cuja variância é r . Caso o valor de y_1 fosse estimado antes de se obter a medida, esse seria:

$$\tilde{y}_1 = c \tilde{x}_1. \quad (2.32)$$

De acordo com as equações (2.29) e (2.32), \tilde{x}_1 e \tilde{y}_1 foram obtidos a partir de valores obtidos no passado, ou seja, \hat{x}_0 nesse caso. Por esse motivo essas equações são chamadas de *equações de predição*.

Segundo Kalman, a partir da medição y_1 , é possível obter uma nova estimativa para x_1 , de acordo com a seguinte expressão:

$$\begin{aligned}\hat{x}_1 &= \tilde{x}_1 + k_1(y_1 - \tilde{y}_1), \\ &= \tilde{x}_1 + k_1(y_1 - c \tilde{x}_1),\end{aligned} \quad (2.33)$$

sendo que k_1 é o *ganho de Kalman* para o instante $k = 1$.

O termo $y_1 - \tilde{y}_1$, conhecido como *inovação*, representa o erro ao estimar y_1 , sendo portanto proveniente de duas fontes:

1. ruído de medição (equação 2.31);
2. erro ao calcular \tilde{x}_1 (equação 2.29).

Pode-se verificar, observando a equação (2.33), que o erro é multiplicado por

um ganho k_1 , para então ser adicionado à estimativa anterior \tilde{x}_1 , de forma tal que essa estimativa é corrigida. Por esse motivo, a equação (2.33) é chamada de *equação de correção*.

Para calcular o ganho k_1 deve-se levar em consideração a variância do erro de estimação:

$$\begin{aligned} E[x_1 - \hat{x}_1]^2 &= E[x_1 - \tilde{x}_1 - k_1(y_1 - c \tilde{x}_1)]^2, \\ &= E[x_1 - \tilde{x}_1 - k_1(c x_1 + e_1 - c \tilde{x}_1)]^2, \\ &= E[(1 - c K_1)^2(x_1 - \tilde{x}_1)^2 + (k_1 e_1)^2], \\ p_1 &= \tilde{p}_1(1 - c k_1)^2 + r k_1^2. \end{aligned} \quad (2.34)$$

Para que o erro de estimação seja mínimo e \hat{x} seja uma estimativa ótima, é preciso minimizar o valor de p_1 na equação (2.34). Isso é feito derivando-se p_1 em relação a k_1 e igualando-se a expressão resultante a zero. Seguindo-se esse desenvolvimento, verifica-se que o ganho de Kalman é dado por:

$$k_1 = c \tilde{p}_1 (c^2 \tilde{p}_1 + r)^{-1}. \quad (2.35)$$

As equações (2.29), (2.30), (2.32), (2.33), (2.34) e (2.35), reescritas de forma generalizada, podem ser reunidas compondo o algoritmo recursivo que recebe o nome de Filtro de Kalman:

$$\left. \begin{aligned} \tilde{x}_k &= a \hat{x}_{k-1} \\ \tilde{y}_k &= c \tilde{x}_k \\ \tilde{p}_k &= a^2 p_{k-1} + q \\ p_k &= c \tilde{p}_k (c^2 \tilde{p}_k + r)^{-1} \\ p_k &= \tilde{p}_k (1 - c k_k)^2 + r k_k^2 \\ \hat{x}_k &= \tilde{x}_k + k_k (y_k - \tilde{y}_k) \end{aligned} \right\}. \quad (2.36)$$

A fim de que possa ser aplicado em sistemas multidimensionais, torna-se necessário fazer pequenas alterações no algoritmo (2.36) permitindo que

sejam realizadas algumas operações matriciais:

$$\left. \begin{aligned} \tilde{\mathbf{x}}_k &= A \hat{\mathbf{x}}_{k-1} \\ \tilde{\mathbf{y}}_k &= C \tilde{\mathbf{x}}_k \\ \tilde{P}_k &= A P_{k-1} A^T + Q \\ K_k &= \tilde{P}_k C^T (C \tilde{P}_k C^T + R)^{-1} \\ P_k &= (I - K_k C) \tilde{P}_k (I - K_k C)^T + K_k R K_k^T \\ \hat{\mathbf{x}}_k &= \tilde{\mathbf{x}}_k + K_k (\mathbf{y}_k - \tilde{\mathbf{y}}_k) \end{aligned} \right\}. \quad (2.37)$$

Na seção seguinte o algoritmo será desenvolvido para aplicação em sistemas multidimensionais, porém seguindo uma outra linha de raciocínio, que é a abordagem Bayesiana.

2.4.2 O Algoritmo em uma Abordagem Bayesiana

Na seção 2.4.1 apresentou-se o algoritmo do Filtro de Kalman (KF) em uma abordagem clássica. Esse algoritmo é abordado na literatura de diversas formas diferentes, porém com o mesmo objetivo: obter a melhor estimativa das variáveis desejadas a partir de medições ruidosas.

Seguindo uma abordagem Bayesiana, o que o filtro faz é determinar a função densidade de probabilidade *a posteriori* $P(\mathbf{x}_k | \mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k)$ e sua evolução no tempo (Peterka, 1981). Esta abordagem é baseada no Teorema de Bayes, descrito na seção 2.2.3. Nesta seção, o KF será descrito no contexto da estimação Bayesiana.

Seja o sistema linear de ordem n representado pelo seguinte modelo:

$$\begin{cases} \mathbf{x}_k = A \mathbf{x}_{k-1} + \mathbf{w}_{k-1}, \\ \mathbf{y}_k = C \mathbf{x}_k + \mathbf{e}_k. \end{cases} \quad (2.38)$$

sendo \mathbf{x}_k o vetor de estados de dimensão $n \times 1$ e \mathbf{y}_k o vetor de saídas. O ruído de processo e o de observação, representados respectivamente por \mathbf{w}_{k-1} e \mathbf{e}_k , são processos aleatórios não-correlacionados e gaussianos. A e C são matrizes constantes de dimensão $n \times n$.

O ruído afeta o sistema de forma tal que os próprios estados se tornam variáveis aleatórias. Pelo fato de o sistema ser linear, a PDF que está sendo propagada é gaussiana em todas as observações até o instante k (Bar-Shalom and Fortmann, 1988). Uma PDF gaussiana pode ser completamente descrita por sua média e variância (ver seção 2.2.2), o que justifica o fato de ser

suficiente se estudar apenas a evolução dessas duas grandezas ao invés de todos os momentos, que por serem possivelmente infinitos, impossibilitariam a solução do problema.

Conforme pôde ser constatado nas seções 2.2.3 e 2.4.1, o KF é composto por duas etapas:

1. predição;
2. correção.

A partir das equações de estado do modelo, durante a etapa de predição calcula-se a média da PDF, ou seja, a predição para a estimativa do estado $\hat{\mathbf{x}}_{k|k-1}$ e a observação associada a ela $\hat{\mathbf{y}}_{k|k-1}$, usando informação de todas as observações *a priori*, ou seja, observações tomadas até o instante $k - 1$, e designada por:

$$Y_{k-1} = \mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{k-1}. \quad (2.39)$$

As predições ótimas são dadas pelas operações conhecidas como *esperança matemática condicional* expressas pelas equações abaixo (Bar-Shalom and Fortmann, 1988):

$$\hat{\mathbf{x}}_{k|k-1} = E[\mathbf{x}_k | Y_{k-1}] = E[A \mathbf{x}_{k-1} | Y_{k-1}], \quad (2.40)$$

$$\hat{\mathbf{y}}_{k|k-1} = E[\mathbf{y}_k | Y_{k-1}] = E[C \mathbf{x}_k | Y_{k-1}]. \quad (2.41)$$

Como o sistema é linear, as equações (2.40) e (2.41) podem ser resolvidas analiticamente, ou seja:

$$\hat{\mathbf{x}}_{k|k-1} = E[A \mathbf{x}_{k-1} | Y_{k-1}] = A E[\mathbf{x}_{k-1} | Y_{k-1}] = A \hat{\mathbf{x}}_{k-1|k-1}. \quad (2.42)$$

Observando-se a equação anterior, verifica-se que o *estado predito* no instante k é função da predição no instante anterior $k - 1$.

As matrizes de covariância ou erros de predição são calculadas pelas equações:

$$P_{k|k-1} = E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T | Y_{k-1}], \quad (2.43)$$

$$P_{k|k-1}^{\mathbf{y}\mathbf{y}} = E[(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})^T | Y_{k-1}], \quad (2.44)$$

$$P_{k|k-1}^{\mathbf{x}\mathbf{y}} = E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})^T | Y_{k-1}]. \quad (2.45)$$

Um modelo nunca é capaz de representar perfeitamente todas as características do comportamento de um determinado sistema. As imperfeições

inerentes ao modelo podem provocar um significativo acúmulo de erros, o que justifica a necessidade de uma etapa de correção no algoritmo do KF (Brunke, 2001). A etapa de correção atualiza as predições para o estado e erro de estimação usando a nova observação \mathbf{y}_k :

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + K_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}), \quad (2.46)$$

$$P_{k|k} = P_{k|k-1} - K_k P_{k|k-1}^{\mathbf{y}\mathbf{y}} K_k^T. \quad (2.47)$$

A equação (2.46) equivale à equação (2.37f), definida na seção anterior, pois $\hat{\mathbf{x}}_{k|k-1} = \tilde{\mathbf{x}}_k$ e $\hat{\mathbf{y}}_{k|k-1} = \tilde{\mathbf{y}}_k$. Existe também uma relação entre a equação (2.47) e (2.37e). A prova da correspondência entre essas duas equações pode ser obtida a partir da equação (2.44).

As incertezas provenientes dos erros gerados durante a etapa de predição são levadas em consideração e corrigidas usando-se a matriz do ganho de Kalman, dada pela equação abaixo:

$$K_k = (P_{k|k-1}^{\mathbf{x}\mathbf{y}})(P_{k|k-1}^{\mathbf{y}\mathbf{y}})^{-1}, \quad (2.48)$$

que atua como sendo um peso para a *inovação* $\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}$ em (2.46). O ganho de Kalman escrito desta forma, é equivalente à equação (2.37d). Isso pode ser demonstrado a partir de um extenso desenvolvimento matemático, após substituição das equações (2.44) e (2.45) em (2.48).

As equações (2.46), (2.47) e (2.48) podem ser agrupadas e colocadas em uma ordem apropriada, constituindo a etapa de correção do algoritmo do KF em uma abordagem Bayesiana:

$$\left. \begin{aligned} K_k &= (P_{k|k-1}^{\mathbf{x}\mathbf{y}})(P_{k|k-1}^{\mathbf{y}\mathbf{y}})^{-1} \\ P_{k|k} &= P_{k|k-1} - K_k P_{k|k-1}^{\mathbf{y}\mathbf{y}} K_k^T \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + K_k (\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \end{aligned} \right\}. \quad (2.49)$$

Partindo de condições iniciais para o vetor de estados e sua matriz de covariância, o algoritmo estima recursivamente o vetor de estados $\hat{\mathbf{x}}_{k|k}$ e a matriz $P_{k|k}$, usando as informações disponibilizadas pelas novas medidas \mathbf{y}_k .

2.5 Conclusões

Neste capítulo foram apresentados alguns conceitos básicos sobre modelagem estocástica e abordagem Bayesiana para problemas de estimação recursiva. Estes conceitos foram apresentados com o intuito de tornar mais clara a estrutura dos algoritmos estudados.

Foram apresentados dois estimadores recursivos para aplicação em problemas de Identificação de Sistemas, dentre os quais se destaca o estimador recursivo de mínimos quadrados, um dos mais utilizados em aplicações práticas.

O tema central do capítulo foi a descrição de um estimador recursivo capaz de estimar as variáveis de estado de sistemas lineares: o Filtro de Kalman (Kalman, 1960). Essa descrição foi apresentada levando-se em consideração um sistema unidimensional, o que deixa o princípio de funcionamento do algoritmo claro e intuitivo. O KF também foi analisado em uma abordagem Bayesiana.

O KF, da forma em que foi proposto originalmente, não pode ser aplicado em sistemas representados por modelos não-lineares. Para ampliar a classe de problemas nos quais o método pode ser usado, foram propostas algumas variações do algoritmo original, tema este que será assunto do próximo capítulo.

O Filtro de Kalman para Sistemas Não-Lineares

“Aprendemos a voar como pássaros, e a nadar como peixes, mas não aprendemos a conviver como irmãos.”

Martin Luther King

3.1 Introdução

Conforme visto no capítulo 2, o Filtro de Kalman (KF) é um algoritmo recursivo muito eficiente, capaz de estimar de forma ótima os estados de sistemas lineares. No KF, toda a informação disponibilizada é incorporada a partir de medições ruidosas para estimar as variáveis desejadas. Pressupõe-se que o sistema seja perturbado por ruído branco e gaussiano, de forma que os estados possam ser tratados como variáveis aleatórias gaussianas. O caráter ótimo da estimação depende criticamente destes pressupostos.

O que acontece se o processo a ser estimado e/ou a relação entre o processo e as medições apresentarem uma dinâmica não-linear? Esta pergunta pode ser respondida considerando-se que transformações não-lineares de variáveis aleatórias gaussianas não são necessariamente gaussianas. Além disso, o algoritmo clássico do KF foi concebido para sistemas lineares, o que impede o uso direto do KF original em sistemas não-lineares.

A fim de se obter uma solução ótima para o problema de estimação em sistemas não-lineares seria necessária uma descrição completa da função densidade de probabilidade condicional dos valores estimados, a cada iteração

(Kushner, 1967b). Infelizmente, esta descrição completa exigiria o estudo da evolução de um número infinito de momentos da distribuição de probabilidade, o que faz com que seja mais viável algum tipo de linearização. Por este motivo, os problemas de estimação para essa classe de sistemas não podem apresentar soluções ótimas.

Várias aproximações foram propostas no passado, e também recentemente (Kushner, 1967a; Sorenson and Stubberud, 1968; Jazwinski, 1970; Maybeck, 1982; Daum, 1988; Gordon et al., 1993; Julier and Uhlmann, 1997).

É importante destacar que algumas das aplicações mais interessantes e que obtiveram resultados bastante satisfatórios com a aplicação do KF são exatamente aquelas nas quais algum tipo de não-linearidade se faz presente.

Neste capítulo serão abordadas duas extensões do KF para sistemas não-lineares. Essas técnicas são conhecidas como:

- Filtro de Kalman Estendido (EKF);¹
- Filtro de Kalman *Unscented* (UKF).²

Além de realizarem estimação de estados de sistemas não-lineares, os dois algoritmos mencionados também são dotados da capacidade de estimar os parâmetros do modelo.

O EKF é um algoritmo que se caracteriza por ser uma simples extensão do KF para aplicação em sistemas não-lineares. Este algoritmo é, possivelmente, o mais conhecido e mais utilizado na maior parte das aplicações.

Por outro lado, o UKF é um algoritmo proposto recentemente por Julier e Uhlmann (Julier and Uhlmann, 1997), como alternativa para solucionar os problemas apresentados pelo EKF. Estes problemas vão desde dificuldades de implementação até divergência de resultados, passando também por falta de precisão, como consequência dos procedimentos de linearização.

Nas seções seguintes serão apresentados os algoritmos EKF e UKF. Em seguida, serão mostrados dois exemplos de aplicação do UKF, destacando os bons resultados obtidos na estimação de estados e parâmetros de sistemas não-lineares simulados.

¹*Extended Kalman Filter.*

²*Unscented Kalman Filter.*

3.2 O Filtro de Kalman Estendido

O *Filtro de Kalman Estendido* (EKF) é uma extensão do algoritmo do KF, proposta para aplicação em sistemas não-lineares. O algoritmo se baseia na linearização das funções, gerando um novo sistema linear aproximado, a partir do qual se torna possível a aplicação do KF clássico (Welch and Bishop, 2002; Morrell, 1997). Este procedimento será explicado a seguir.

O sistema não-linear discreto pode ser representado pelo seguinte modelo:

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1}, \quad (3.1)$$

$$\mathbf{y}_k = h(\mathbf{x}_k) + \mathbf{e}_k, \quad (3.2)$$

sendo que \mathbf{x}_k é o vetor de estados e \mathbf{y}_k o vetor de saídas. Os vetores \mathbf{w}_{k-1} e \mathbf{e}_k representam, respectivamente, ruído de processo e ruído de observação, e são processos aleatórios não-correlacionados e gaussianos.

A não-linearidade pode entrar no problema na dinâmica do processo, por meio da função não-linear f em (3.1), ou na equação de observação (3.2) por meio da função não-linear h (Brown and Hwang, 1997).

Um novo sistema, porém linear, pode ser gerado a partir das equações (3.1) e (3.2) da seguinte maneira. Seja \mathbf{x}_k^* um vetor de referência, a princípio escolhido arbitrariamente. Esse vetor pode ser expresso pela equação (3.1) sem o termo referente ao ruído \mathbf{w}_{k-1} :

$$\mathbf{x}_k^* = f(\mathbf{x}_{k-1}^*). \quad (3.3)$$

A diferença entre \mathbf{x}_k e \mathbf{x}_k^* é definida como:

$$\boldsymbol{\delta}_k = \mathbf{x}_k - \mathbf{x}_k^*. \quad (3.4)$$

Substituindo as equações (3.1) e (3.3) em (3.4), obtém-se:

$$\begin{aligned} \boldsymbol{\delta}_k &= \mathbf{x}_k - \mathbf{x}_k^* \\ &= f(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} - f(\mathbf{x}_{k-1}^*). \end{aligned} \quad (3.5)$$

É possível obter uma aproximação linear para a equação (3.5) a partir de

uma expansão em série de Taylor³ de $f(\mathbf{x}_{k-1})$ em torno de \mathbf{x}_{k-1}^* :

$$f(\mathbf{x}_{k-1}) \approx f(\mathbf{x}_{k-1}^*) + F_{k-1}(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^*), \quad (3.6)$$

sendo F_{k-1} a matriz jacobiana:

$$F_{k-1} = \left. \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_{k-1}^*}. \quad (3.7)$$

Substituindo a equação (3.6) em (3.5) obtém-se uma equação de diferenças linear:

$$\begin{aligned} \boldsymbol{\delta}_k &\approx f(\mathbf{x}_{k-1}^*) + F_{k-1}(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^*) + \mathbf{w}_{k-1} - f(\mathbf{x}_{k-1}^*) \\ &= F_{k-1}\boldsymbol{\delta}_{k-1} + \mathbf{w}_{k-1}. \end{aligned} \quad (3.8)$$

De maneira muito semelhante é possível obter uma aproximação linear para a equação (3.2). O termo não-linear $h(\mathbf{x}_k)$ pode ser aproximado por:

$$h(\mathbf{x}_k) \approx h(\mathbf{x}_k^*) + H_k(\mathbf{x}_k - \mathbf{x}_k^*), \quad (3.9)$$

sendo H_k a matriz jacobiana:

$$H_k = \left. \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_k^*}. \quad (3.10)$$

A equação de observação (3.2) pode então ser expressa como:

$$\mathbf{y}_k \approx h(\mathbf{x}_k^*) + H_k(\mathbf{x}_k - \mathbf{x}_k^*) + \mathbf{e}_k, \quad (3.11)$$

³A expansão de uma função $f(x)$, em série de Taylor, em torno de um ponto x_0 , é definida pela seguinte expressão:

$$f(x) = f(x_0) + \left. \frac{\partial f(x)}{\partial x} \right|_{x=x_0} (x - x_0) + \text{termos de ordem superior em } (x - x_0).$$

a partir da qual obtém-se uma equação linear para a nova variável $\boldsymbol{\nu}$:

$$\begin{aligned}\boldsymbol{\nu}_k &= \mathbf{y}_k - h(\mathbf{x}_k^*) \\ &= h(\mathbf{x}_k) + \mathbf{e}_k - h(\mathbf{x}_k^*) \\ &\approx H_k(\mathbf{x}_k - \mathbf{x}_k^*) + \mathbf{e}_k \\ \boldsymbol{\nu}_k &= H_k \boldsymbol{\delta}_k + \mathbf{e}_k.\end{aligned}\tag{3.12}$$

As equações geradas por meio dos procedimentos de linearização acima podem ser agrupadas compondo o seguinte modelo linear:

$$\begin{cases} \boldsymbol{\delta}_k = F_{k-1} \boldsymbol{\delta}_{k-1} + \mathbf{w}_{k-1}, \\ \boldsymbol{\nu}_k = H_k \boldsymbol{\delta}_k + \mathbf{e}_k. \end{cases}\tag{3.13}$$

Após obtido este sistema linearizado, é possível aplicar normalmente as equações do KF clássico para calcular uma estimativa para o novo vetor de estados $\hat{\boldsymbol{\delta}}_k$. Como $\hat{\boldsymbol{\delta}}_k$ é uma estimativa da diferença entre \mathbf{x}_k e \mathbf{x}_k^* , a estimativa do vetor original $\hat{\mathbf{x}}_k$, é calculada de acordo com a expressão:

$$\hat{\mathbf{x}}_k = \mathbf{x}_k^* + \hat{\boldsymbol{\delta}}_k.\tag{3.14}$$

O vetor de referência \mathbf{x}_k^* varia a cada iteração. No primeiro passo ele é escolhido como sendo o valor esperado do vetor \mathbf{x}_0 . Desta forma,

$$E[\boldsymbol{\delta}_0] = 0,$$

e a primeira estimativa para o vetor \mathbf{x} é:

$$\hat{\mathbf{x}}_0 = \mathbf{x}_0^*.\tag{3.15}$$

Nas próximas iterações, \mathbf{x}_k^* deve ser atualizado com a estimativa obtida no passo anterior: $\hat{\mathbf{x}}_{k-1}$.

Devido às aproximações realizadas, o EKF não é capaz de fornecer uma estimativa ótima para o vetor de estados desejados, sendo estimativa ótima aquela que minimiza o valor esperado da soma quadrática dos resíduos de estimação. Entretanto, a técnica se mostrou eficaz em diversas aplicações diferentes envolvendo sistemas não-lineares. Por outro lado, é sabido que os procedimentos de linearização das equações podem conduzir a resultados

insatisfatórios ou até mesmo à divergência do filtro. Além disso, o cálculo de matrizes jacobianas não é trivial na maior parte das aplicações, o que frequentemente gera uma certa dificuldade na implementação.

Essas falhas identificadas no EKF motivaram a busca por uma nova abordagem para problemas de estimação em sistemas não-lineares baseada no KF. Esse assunto será tratado na próxima seção.

3.3 O Filtro de Kalman Unscented

O Filtro de Kalman Unscented (UKF) foi proposto por Simon J. Julier e Jeffrey K. Uhlmann com o intuito de superar os problemas verificados quando se faz uso do EKF em determinadas aplicações (Julier and Uhlmann, 1996, 1997; Julier et al., 1995, 2000). Algumas modificações para o UKF foram propostas posteriormente de forma a adaptá-lo de acordo com o tipo de aplicação. As adaptações mais interessantes encontram-se nos trabalhos de Eric A. Wan e Rudolph van der Merwe (Merwe and Wan, 2000, 2001a,b; Merwe et al., 2000; Wan et al., 2000).

O UKF é um algoritmo capaz de realizar estimação de estados em sistemas dotados de alguma não-linearidade, sem a necessidade de se linearizar as funções presentes no modelo. Essa estimação se torna possível graças a uma técnica de transformação que gera um conjunto finito de vetores que, ao passarem por uma transformação não-linear, apresentam média amostral e covariância amostral iguais à média e covariância que seriam calculadas para o novo conjunto de vetores transformados, supondo uma transformação de 2ª ordem. Esta técnica é conhecida como *Transformação Unscented* e será detalhada na próxima seção.

3.3.1 A Transformação Unscented

A *Transformação Unscented* (UT), do inglês *Unscented Transformation*, é uma técnica que foi proposta por Julier e Uhlmann (Julier and Uhlmann, 1996) para calcular as propriedades estatísticas de variáveis aleatórias que passam por alguma transformação não-linear.

O problema pode ser apresentado de forma mais específica da seguinte maneira: Uma variável aleatória \mathbf{x} , de dimensão n , possui média e covariância representadas respectivamente por $\bar{\mathbf{x}}$ e P . Deseja-se calcular a média $\bar{\mathbf{y}}$ e a covariância $P^{\mathbf{y}\mathbf{y}}$ de uma variável aleatória \mathbf{y} , de dimensão m , sendo que \mathbf{y}

está relacionada a \mathbf{x} por meio de uma função não-linear g :

$$\mathbf{y} = g(\mathbf{x}). \quad (3.16)$$

Segundo o autor, é muito mais simples aproximar uma distribuição gaussiana do que linearizar uma função não-linear arbitrária. Conforme já foi dito anteriormente (seção 2.2.2), para que uma PDF gaussiana possa ser completamente descrita, são necessários apenas os momentos de primeira e de segunda ordem, ou seja, média e covariância. Desta forma, o objetivo passa a ser o de encontrar alguma parametrização que ao mesmo tempo, capture média e covariância, permitindo a propagação direta da informação por meio das equações não-lineares.

O objetivo almejado pode ser alcançado, gerando-se um conjunto de vetores a partir das linhas ou colunas⁴ das matrizes $\pm\sqrt{nP}$, ou seja, das matrizes compostas pela raiz quadrada positiva e negativa da matriz de covariância P multiplicada pela dimensão do vetor de variáveis aleatórias n .

Esse novo conjunto de vetores é normalmente representado pela matriz $n \times 2n$:

$$\boldsymbol{\sigma} = \begin{bmatrix} \sqrt{nP} & -\sqrt{nP} \end{bmatrix}, \quad (3.17)$$

sendo que cada coluna, isto é, cada vetor é conhecido na literatura como *Ponto Sigma* (SP), do inglês *Sigma Point*. Por este motivo, alguns pesquisadores tem o hábito de se referir ao UKF como *Sigma Points Kalman Filter* (SPKF) (Merwe and Wan, 2001a).

Devido à maneira como é construído, esse conjunto de vetores possui covariância amostral igual a P e média amostral zero. Desta forma, basta adicionar a média $\bar{\mathbf{x}}$ a cada coluna da matriz $\boldsymbol{\sigma}$ para que o novo conjunto de vetores formado possua a mesma média dos vetores \mathbf{x} originais. Como o novo conjunto é simétrico, seus momentos centrais ímpares são nulos, fazendo com que os três primeiros momentos sejam iguais aos da distribuição original.

É possível ainda adicionar-se um fator de escala κ no cálculo da raiz quadrada da matriz de covariância, sem que o resultado da nova covariância se altere. A razão pela qual essa alteração no algoritmo pode ser útil, bem como os valores que são tipicamente atribuídos a κ serão apresentados na seção 3.3.4.

⁴A matriz de covariância P é uma matriz simétrica, bem como a matriz \sqrt{P} .

O novo conjunto de vetores é gerado da seguinte maneira:

$$\begin{aligned}\boldsymbol{\sigma} &= \left[\sqrt{(n + \kappa)P} \quad - \sqrt{(n + \kappa)P} \right], \\ \mathcal{X}^{(0)} &= \bar{\mathbf{x}}, \\ \mathcal{X}^{(i)} &= \bar{\mathbf{x}} + \boldsymbol{\sigma}^{(i)},\end{aligned}\tag{3.18}$$

sendo $\boldsymbol{\sigma}^{(i)}$ a i -ésima coluna da matriz $\boldsymbol{\sigma}$. Desta forma i varia de 1 até $2n$.

A nova matriz de covariância é dada por:

$$P = \frac{1}{2(n + \kappa)} \sum_{i=1}^{2n} (\mathcal{X}^{(i)} - \bar{\mathbf{x}})(\mathcal{X}^{(i)} - \bar{\mathbf{x}})^T.\tag{3.19}$$

Desenvolvendo-se a equação (3.19) é possível verificar que a covariância de \mathcal{X} é igual à covariância original P .

A UT pode ser efetuada seguindo-se o procedimento abaixo (Julier and Uhlmann, 1996, 1997):

1. Cálculo dos SP e novo conjunto de vetores \mathcal{X} :

$$\begin{aligned}\mathcal{X}^{(0)} &= \bar{\mathbf{x}}, \\ \mathcal{X}^{(i)} &= \bar{\mathbf{x}} + \boldsymbol{\sigma}^{(i)}.\end{aligned}\tag{3.20}$$

2. Cálculo dos pesos w :

$$\begin{aligned}w^{(0)} &= \frac{\kappa}{n + \kappa}, \\ w^{(i)} &= \frac{1}{2(n + \kappa)}.\end{aligned}\tag{3.21}$$

3. Cada novo vetor gerado $\mathcal{X}^{(i)}$ é propagado usando-se a função não-linear g , formando um conjunto de SP transformado:

$$\mathcal{Y}^{(i)} = g(\mathcal{X}^{(i)}).\tag{3.22}$$

4. A nova média é dada por:

$$\bar{\mathbf{y}} = \sum_{i=0}^{2n} w^{(i)} \mathcal{Y}^{(i)}.\tag{3.23}$$

5. A covariância do conjunto de vetores transformado $\mathcal{Y}^{(i)}$ é:

$$P^{yy} = \sum_{i=0}^{2n} w^{(i)} (\mathcal{Y}^{(i)} - \bar{\mathbf{y}}) (\mathcal{Y}^{(i)} - \bar{\mathbf{y}})^T. \quad (3.24)$$

Observando-se a equação (3.20), verifica-se que o número de pontos sigma (n_{SP}) é dado por:

$$n_{SP} = 2n + 1. \quad (3.25)$$

É importante notar que as estimativas (3.23) e (3.24) são exatas somente para transformações de até 2ª ordem (Julier and Uhlmann, 1996).

Representação Gráfica da UT

Seja \mathbf{x} uma variável aleatória bidimensional cujas média e covariância são dadas, respectivamente, por $\bar{\mathbf{x}}$ e P . Deseja-se determinar a média $\bar{\mathbf{y}}$ e a covariância P^{yy} da variável aleatória \mathbf{y} , a qual está relacionada com \mathbf{x} pela função:

$$\mathbf{y} = \arctan(\mathbf{x}). \quad (3.26)$$

Como as variáveis aleatórias envolvidas são bidimensionais, calcula-se, a partir da equação (3.25), que o número de SP é $n_{SP} = 5$.

Na Figura 3.1 mostra-se como ocorre a UT. Na parte (a) representa-se as propriedades estatísticas de \mathbf{x} , enquanto que na parte (b) estão representadas as propriedades estatísticas da variável transformada \mathbf{y} . Os SP são representados pelos 4 \times mais a média, que é representada por um pequeno círculo, totalizando os 5 SP, em ambas as partes da figura. O triângulo na parte (b) da figura se refere à média real após transformação. As covariâncias são representadas pelas elipsóides, sendo que aquela traçada em (—) é a covariância dos SP, enquanto que a que foi traçada em (---) é a covariância real.

Para que se possa comparar a UT com métodos que linearizam a função não-linear, como o EKF, mostram-se a média (quadrado) e a covariância (— · —) da variável \mathbf{y} obtidas por linearização.

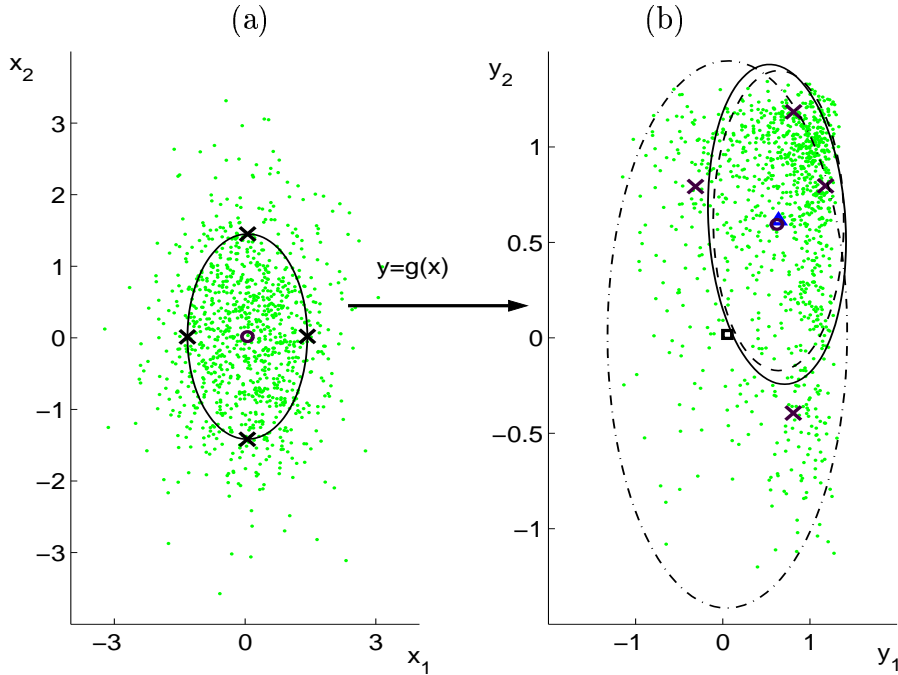


Figura 3.1: Transformação Unscented – Propagação da média e covariância – Parte (a) SP (\times), média (círculo), covariância (elipsóide). Parte (b) média (círculo) e covariância dos SP (elipsóide (—)), média real (triângulo), covariância real (elipsóide (---)), média linearizada (quadrado), covariância linearizada (elipsóide (- · -)).

3.3.2 O Algoritmo do Filtro de Kalman Unscented

O algoritmo UKF se baseia na UT (Julier and Uhlmann, 1996), apresentada na seção anterior. Da mesma maneira que na seção 3.2, o sistema não-linear cujos estados se deseja estimar pode ser representado pelas equações (3.1) e (3.2).

Conforme estabelecido na seção 2.4.2, a primeira etapa do KF corresponde a um procedimento de *predição*, no qual calcula-se a média de uma PDF, ou seja, a predição para a estimativa do estado $\hat{\mathbf{x}}_{k|k-1}$ e a observação associada a ela $\hat{\mathbf{y}}_{k|k-1}$, usando informação de todas as observações *a priori* até o instante $k-1$. A predição pode ser realizada usando-se a UT de acordo com o seguinte algoritmo:

1. Cálculo dos pontos sigma e novo conjunto de vetores:

$$\boldsymbol{\sigma} = \left[\sqrt{(n + \kappa)P} \quad - \sqrt{(n + \kappa)P} \right],$$

$$\begin{aligned} \mathcal{X}_{k-1|k-1}^{(0)} &= \hat{\mathbf{x}}_{k-1|k-1}, \\ \mathcal{X}_{k-1|k-1}^{(i)} &= \hat{\mathbf{x}}_{k-1|k-1} + \left[\sqrt{(n + \kappa)P_{k-1|k-1}} \right]^{(i)}. \end{aligned} \quad (3.27)$$

2. Cálculo dos pesos w ; ver equação (3.21).

3. Propagação de cada vetor $\mathcal{X}^{(i)}$ usando-se as funções não-lineares f e h :

$$\mathcal{X}_{k|k-1}^{(i)} = f(\mathcal{X}_{k-1|k-1}^{(i)}), \quad (3.28)$$

$$\mathcal{Y}_{k|k-1}^{(i)} = h(\mathcal{X}_{k|k-1}^{(i)}). \quad (3.29)$$

4. Cálculo das médias:

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{i=0}^{2n} w^{(i)} \mathcal{X}_{k|k-1}^{(i)}, \quad (3.30)$$

$$\hat{\mathbf{y}}_{k|k-1} = \sum_{i=0}^{2n} w^{(i)} \mathcal{Y}_{k|k-1}^{(i)}. \quad (3.31)$$

5. Cálculo das matrizes de covariância e correlação cruzada:

$$P_{k|k-1} = \sum_{i=0}^{2n} w^{(i)} (\mathcal{X}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1})(\mathcal{X}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1})^T, \quad (3.32)$$

$$P_{k|k-1}^{\mathbf{x}\mathbf{y}} = \sum_{i=0}^{2n} w^{(i)} (\mathcal{X}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1})(\mathcal{Y}_{k|k-1}^{(i)} - \hat{\mathbf{y}}_{k|k-1})^T, \quad (3.33)$$

$$P_{k|k-1}^{\mathbf{y}\mathbf{y}} = \sum_{i=0}^{2n} w^{(i)} (\mathcal{Y}_{k|k-1}^{(i)} - \hat{\mathbf{y}}_{k|k-1})(\mathcal{Y}_{k|k-1}^{(i)} - \hat{\mathbf{y}}_{k|k-1})^T. \quad (3.34)$$

A média $\hat{\mathbf{x}}_{k|k-1}$ e covariância $P_{k|k-1}$ preditas são atualizadas pelas equações de correção do KF tradicional (seção 2.4.2), para gerar uma estimativa

para o estado $\hat{\mathbf{x}}_{k|k}$ e seu erro de estimação $P_{k|k}$:

$$K_k = (P_{k|k-1}^{\mathbf{x}\mathbf{y}})(P_{k|k-1}^{\mathbf{y}\mathbf{y}})^{-1}, \quad (3.35)$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + K_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}), \quad (3.36)$$

$$P_{k|k} = P_{k|k-1} - K_k P_{k|k-1}^{\mathbf{y}\mathbf{y}} K_k^T. \quad (3.37)$$

O processo de estimação começa com a determinação de condições iniciais para o vetor de estados \mathbf{x}_0 e matriz de covariância dos estados P_0 , conforme será visto na seção 3.3.4 e nos exemplos apresentados neste trabalho. Após a primeira iteração, o vetor \mathbf{x}_k e a matriz P_k passam a ser atualizados pelas equações (3.36) e (3.37), respectivamente.

3.3.3 Estimação de Parâmetros

Em problemas de estimação de estados existem situações em que os parâmetros do modelo não são precisamente conhecidos, ou até mesmo totalmente desconhecidos. Nessas situações, os resultados de evolução das variáveis de estado obtidos por estimação podem não ser satisfatórios. Uma alternativa para contornar esse problema, ou seja, reduzir o erro de estimação para os estados, é estimar também os parâmetros do modelo. O UKF é capaz de realizar essa tarefa ao mesmo tempo em que estima os estados.

Essa operação pode ser realizada considerando-se os parâmetros como variáveis de estado artificiais. Neste caso, o comportamento dinâmico do vetor de parâmetros é dado pela equação:

$$\boldsymbol{\lambda}_k = \boldsymbol{\lambda}_{k-1}. \quad (3.38)$$

O ruído de processo \mathbf{w}_k e o de medição \mathbf{e}_k também podem ser tratados como variáveis de estado artificiais, fazendo-se:

$$\mathbf{w}_k = \mathbf{w}_{k-1}, \quad (3.39)$$

$$\mathbf{e}_k = \mathbf{e}_{k-1}, \quad (3.40)$$

caso se deseje estimar também estes sinais. O interesse na estimação desses sinais pode ser justificado em algumas aplicações, como por exemplo em *separação de sinais*. Neste caso a variável aleatória \mathbf{w}_k adicionada na equação de estado do sistema deixaria de ser considerada ruído de processo, passando a ser um sinal desejado que deve ser separado do estado \mathbf{x}_k .

Desta maneira, é formado um novo vetor de estados \mathbf{x}_k^+ , freqüentemente chamado de *vetor de estados estendido*, que é composto pelos estados reais \mathbf{x}_k , parâmetros $\boldsymbol{\lambda}_k$ e variáveis aleatórias \mathbf{w}_k e \mathbf{e}_k :

$$\mathbf{x}_k^+ = \begin{bmatrix} \mathbf{x}_k \\ \boldsymbol{\lambda}_k \\ \mathbf{w}_k \\ \mathbf{e}_k \end{bmatrix}. \quad (3.41)$$

Após formado o vetor de estados estendido, deve-se criar também uma matriz de covariância dos estados estendida P_k^+ , a qual deve incluir elementos correspondentes aos parâmetros e ruído. A partir desse ponto o algoritmo UKF, apresentado na seção anterior, é utilizado normalmente para realizar a estimação do vetor \mathbf{x}_k^+ . Os parâmetros e sinais de ruído estimados variam a cada iteração do algoritmo até que a convergência seja atingida de forma a garantir que as equações (3.38), (3.39) e (3.40) sejam satisfeitas.

3.3.4 Aspectos da Implementação do Algoritmo

O UKF apresenta uma série de vantagens com relação ao EKF, conforme foi demonstrado por Julier e Uhlmann (Julier and Uhlmann, 1996). Por outro lado, alguns fatores podem dificultar a implementação ou conduzir a resultados indesejados. Nesta seção será feita uma análise do algoritmo destacando os aspectos mais importantes da sua implementação prática.

Precisão

Os procedimentos de linearização, utilizados no EKF, provocam erros nos termos de segunda ordem no cálculo da média e nos termos de quarta ordem ao se calcular a covariância das variáveis em questão (Julier and Uhlmann, 1996). Os autores da técnica demonstraram que a UT produz erro nos termos de quarta ordem tanto nos cálculos da média como da covariância. Isto sugere que o UKF reduz bastante os erros de estimação, conduzindo a resultados bem mais precisos que os obtidos usando-se o EKF.

Um outro fato interessante que pode conduzir a resultados mais precisos, reduzindo o erro de estimação, é a influência da constante κ . Pode ser demonstrado (Julier and Uhlmann, 1996) que quando o sistema for escalar, escolhendo-se $\kappa = 2$, os erros de média e covariância só aparecem nos ter-

mos de sexta ordem. Quando se tem um sistema multi-dimensional, os erros surgem a partir dos termos de quarta ordem, e nestes casos é interessante selecionar $\kappa = 3 - n$.

Este parâmetro pode ser usado para incorporar informação proveniente de momentos de ordem superior. Por esse motivo, o método também pode ser aplicado quando as variáveis aleatórias de ruído possuem outro tipo de distribuição além da gaussiana (Julier and Uhlmann, 1996).

O Cálculo da Raiz Quadrada

O cálculo da raiz quadrada da matriz de covariância é uma operação crítica na determinação dos SP. Em determinados casos podem ocorrer problemas de instabilidade numérica e mal condicionamento da matriz resultante. Por outro lado, existe uma certa liberdade na escolha do método a ser usado no cálculo dessa raiz quadrada,⁵ conduzindo a diversas alternativas de cálculo, como por exemplo via *decomposição de Cholesky* (Gill et al., 1991; Press et al., 1992):

$$P = A A^T, \quad (3.42)$$

sendo que A é uma matriz triangular inferior, e P deve ser simétrica e *definida positiva* (Chen, 1984).

Há também uma alternativa mais sofisticada, que faz uso tanto da decomposição de Cholesky como também de outras ferramentas da álgebra linear, pela qual se propaga diretamente a matriz raiz quadrada, sem a necessidade de se recalculá-la a cada iteração. Esse algoritmo é conhecido como *Square-root Unscented Kalman filter* (SRUKF) (Merwe and Wan, 2001b). Esses métodos, considerados mais estáveis no cálculo da matriz raiz quadrada são muito importantes para aplicações em tempo real (Julier and Uhlmann, 1996).

Inicialização do Algoritmo

Para dar início ao processo de estimação todo estimador recursivo precisa ser inicializado. Esta inicialização inclui a determinação de condições iniciais para formar o vetor de estados estendido \mathbf{x}_0^+ e uma matriz de covariância inicial adequada P_0^+ . Normalmente, quando não se conhece os parâmetros

⁵Esta liberdade é possível porque uma determinada matriz raiz quadrada pode ser gerada a partir de qualquer outra aplicando-se uma transformação ortonormal (Chen, 1984). Portanto, a matriz raiz quadrada não é única.

do modelo, o que se faz é inicializar esse vetor com valores nulos e a matriz de covariância com valores elevados, indicando que se tem pouca ou nenhuma certeza sobre o valor real do parâmetro.

No caso do UKF, a situação parece ser um pouco mais complexa. A escolha da estimativa inicial para a matriz de covariância influencia a taxa de convergência do algoritmo. De acordo com a matriz inicial, a convergência pode ser mais lenta, mais rápida ou pode não ocorrer.

Na realidade, há pouca informação sobre essa questão disponível na literatura. Trata-se de um problema ainda em estudo por alguns pesquisadores e omitido por outros, por não ter se apresentado como um empecilho relevante na sua aplicação do algoritmo. Um fato conhecido é que, ao contrário do EKF, não se pode inicializar a matriz de covariância com valores elevados porque isto pode causar grande impacto na etapa de predição (Julier, 2003). Desta forma, não é qualquer matriz de covariância inicial suficientemente grande que vai levar o algoritmo a convergir para os valores desejados.

É importante deixar claro que quando se refere à inicialização da matriz de covariância, os elementos significativos são aqueles que se encontram na diagonal desta matriz, pois estes realmente correspondem à covariância de cada variável. Aos demais elementos geralmente são atribuídos valores nulos.

A sensibilidade à inicialização se apresenta como uma desvantagem do UKF e será melhor caracterizada no capítulo 4, onde se descreve sobre a aplicação do UKF em um sistema de laboratório.

3.3.5 Exemplos de Aplicações

Nesta seção reproduz-se os resultados apresentados por Andre Sitz e colegas (Sitz et al., 2002), com o objetivo investigar os aspectos práticos da implementação do UKF.

O UKF será utilizado para estimar, ao mesmo tempo, os estados e os parâmetros de dois sistemas não-lineares, a partir de séries temporais ruidosas. Esses sistemas são:

- o sistema de Lotka-Volterra;
- o sistema de Lorenz.

Geração de dados

Os dados, para cada sistema estudado, foram gerados de acordo com o seguinte procedimento:

1. simulação das equações diferenciais do sistema durante 100 segundos, com intervalo de integração variável e amostragem $T = 0,01s$.
2. adição de ruído gaussiano ao estado x_1 de cada sistema, com desvio padrão igual a 10% do desvio padrão da série temporal dos estados sem ruído.

O sinal resultante, gerado desta forma, pode ser considerado como uma boa aproximação para um sinal medido em um processo real, com ruído de medição.

No que se segue, serão apresentadas as particularidades de cada sistema e os respectivos resultados de estimação proporcionados pelo UKF.

O Sistema de Lotka-Volterra

O *sistema de Lotka-Volterra* é descrito pelas seguintes equações diferenciais não-lineares (Hofbauer and Sigmund, 1988):

$$\begin{cases} \dot{x}_1 = \lambda_1 x_1 - \lambda_2 x_1 x_2, \\ \dot{x}_2 = \lambda_2 x_1 x_2 - \lambda_3 x_2. \end{cases} \quad (3.43)$$

Ele pode ser representado em espaço de estados por:

$$\begin{bmatrix} \mathbf{x}_k \\ \boldsymbol{\lambda}_k \end{bmatrix} = \begin{bmatrix} f(\mathbf{x}_{k-1}, \boldsymbol{\lambda}_{k-1}) \\ \boldsymbol{\lambda}_{k-1} \end{bmatrix} \quad (3.44)$$

$$y_k = x_k^{(1)} + e_k \quad (3.45)$$

sendo que f representa as equações (3.43) do instante $k - 1$ até o instante k ; \mathbf{x}_k é o vetor de estados no instante k e $\boldsymbol{\lambda}_k$ é o vetor de parâmetros do modelo. Na equação de saída (3.45), $x_k^{(1)}$ representa o valor da variável x_1 na iteração k , e e_k ruído de medição.

Os valores reais dos parâmetros (usados para gerar os dados) foram:

$$\begin{aligned}\lambda_1 &= 1; \\ \lambda_2 &= 1,5; \\ \lambda_3 &= 2.\end{aligned}$$

O UKF foi utilizado para estimar os estados x_1 e x_2 , e parâmetros λ_1 , λ_2 e λ_3 ao mesmo tempo, a partir da série temporal simulada x_1 com ruído de medição.

O vetor inicial de estados estendido foi escolhido como:

$$\mathbf{x}_0^+ = \begin{bmatrix} y_1 \\ y_1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad (3.46)$$

sendo que os dois estados são iniciados com a primeira amostra do sinal simulado y_1 e os três últimos valores se referem às condições iniciais escolhidas para os parâmetros.

A matriz inicial de covariância do vetor de estados estendido usada para inicializar o algoritmo foi:

$$P_0 = \begin{bmatrix} 0,025 & 0 & 0 & 0 & 0 \\ 0 & 0,025 & 0 & 0 & 0 \\ 0 & 0 & 0,3 & 0 & 0 \\ 0 & 0 & 0 & 0,3 & 0 \\ 0 & 0 & 0 & 0 & 0,3 \end{bmatrix}. \quad (3.47)$$

Os parâmetros obtidos por estimação apresentaram os seguintes valores:

$$\begin{aligned}\lambda_1 &= 0,9966; \\ \lambda_2 &= 1,5025; \\ \lambda_3 &= 2,0029.\end{aligned}$$

Na Figura 3.2 mostra-se a evolução desses parâmetros até atingir a convergência. Pode ser notado que os parâmetros foram estimados com exatidão.

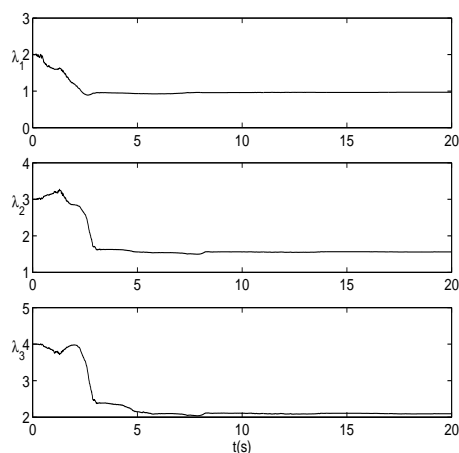


Figura 3.2: Evolução dos parâmetros na estimação – Lotka-Volterra.

Na Figura 3.3 são comparados os estados estimados com os estados obtidos por simulação.

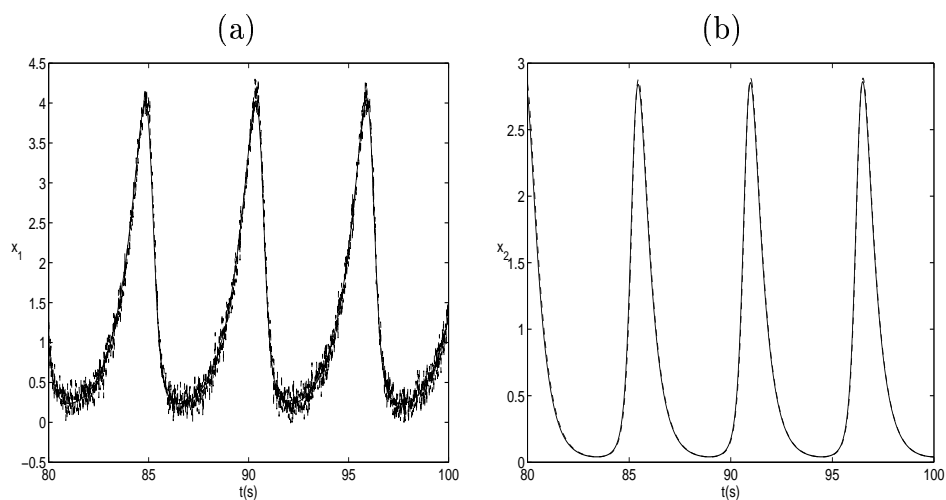


Figura 3.3: Estimação de estados – Lotka-Volterra. (a) Evolução do estado x_1 , simulado (---) e estimado (—) e (b) Evolução do estado x_2 , simulado (---) e estimado (—).

A Figura 3.3a refere-se ao estado x_1 , ou seja, sinal simulado com ruído. É interessante notar que o ruído é eliminado durante a estimação, ou seja, este sinal é “filtrado”. Na Figura 3.3b as curvas se referem a x_2 estimado e simulado. Verifica-se que estas curvas se encontram praticamente sobrepostas,

como consequência do bom resultado proporcionado pelo algoritmo.

Os erros percentuais máximos de estimação, em relação aos valores máximos das séries temporais x_1 e x_2 , foram de 0,92 % e 1,66 %, respectivamente.

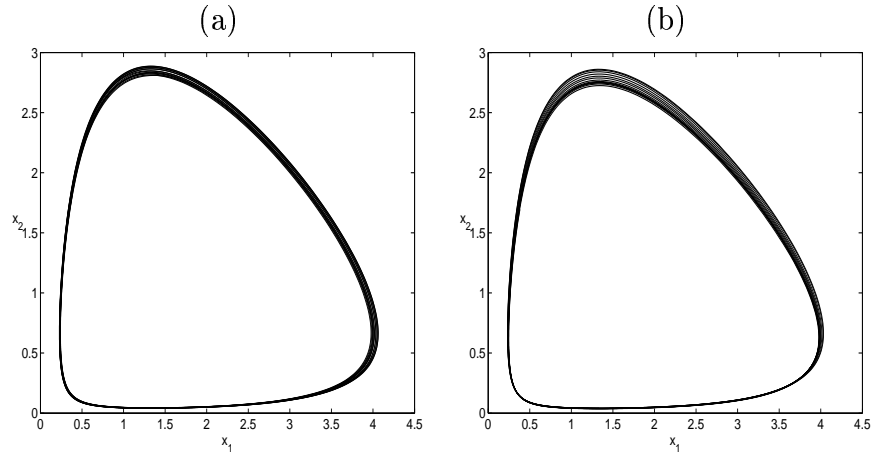


Figura 3.4: Plano de fase do sistema de Lotka-Volterra. (a) original e (b) estimado.

Na Figura 3.4 representa-se o plano de fase $x_1 \times x_2$ do sistema de Lotka-Volterra. O plano projetado a partir dos estados estimados é muito semelhante ao original, pois os estados foram estimados com bastante precisão.

A inicialização da matriz de covariância dos estados teve alguma influência na velocidade de convergência do algoritmo. Alguns testes foram feitos nos quais variou-se tal matriz inicial, a partir dos quais foi possível perceber que melhores resultados podem ser obtidos com valores mais baixos. O melhor resultado foi obtido com P_0^+ dada pela equação (3.47).

O Sistema de Lorenz

O *sistema de Lorenz* é um sistema que pode apresentar dinâmica caótica e é representado pelas seguintes equações diferenciais (Lorenz, 1963):

$$\begin{cases} \dot{x}_1 = -\lambda_1 x_1 + \lambda_1 x_2, \\ \dot{x}_2 = \lambda_2 x_1 - x_2 - x_1 x_3, \\ \dot{x}_3 = -\lambda_3 x_3 + x_1 x_2. \end{cases} \quad (3.48)$$

O sistema pode ser representado pelo seguinte modelo em espaço de estados:

$$\begin{bmatrix} \mathbf{x}_k \\ \boldsymbol{\lambda}_k \end{bmatrix} = \begin{bmatrix} f(\mathbf{x}_{k-1}, \boldsymbol{\lambda}_{k-1}) \\ \boldsymbol{\lambda}_{k-1} \end{bmatrix}, \quad (3.49)$$

$$y_k = x_k^{(1)} + e_k. \quad (3.50)$$

Os valores reais dos parâmetros (usados na geração de dados) foram:

$$\lambda_1 = 12;$$

$$\lambda_2 = 40;$$

$$\lambda_3 = 4.$$

O objetivo neste exemplo, da mesma maneira que no sistema de Lotka-Volterra, foi utilizar o UKF para estimar os estados x_1 , x_2 e x_3 , e parâmetros λ_1 , λ_2 e λ_3 , a partir da simulação da variável de estado x_1 contaminada por ruído.

O vetor de estados estendido \mathbf{x}_0^+ , escolhido para inicializar o algoritmo foi:

$$\mathbf{x}_0^+ = \begin{bmatrix} y_1 \\ y_1 \\ y_1 \\ 20 \\ 15 \\ 10 \end{bmatrix}. \quad (3.51)$$

O algoritmo se mostrou mais sensível à inicialização da matriz de covariância neste caso do que no anterior. Percebe-se que o algoritmo converge para uma faixa menor de valores. Esta situação não foi tão crítica no sistema de Lotka-Volterra. Esse problema será mostrado em maiores detalhes a partir da análise de outro sistema, no capítulo 4.

Os parâmetros estimados, usando como condições iniciais o vetor de estados estendido dado pela equação (3.51), foram:

$$\lambda_1 = 12,01;$$

$$\lambda_2 = 39,99;$$

$$\lambda_3 = 3,99.$$

A Figura 3.5 mostra a evolução desses parâmetros.

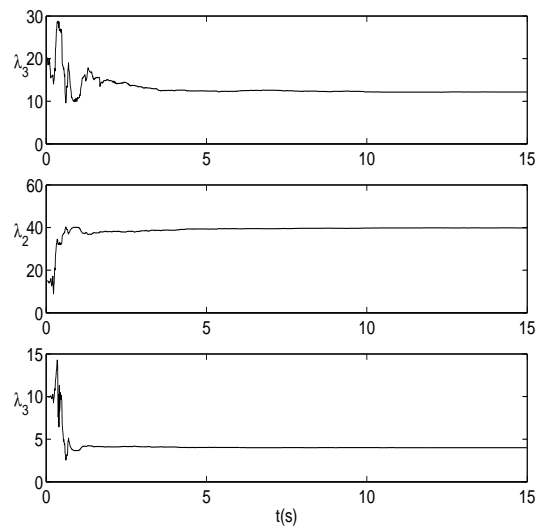


Figura 3.5: Evolução dos parâmetros na estimação – Lorenz.

Assim como no caso anterior, também houve convergência para os valores usados na geração de dados. Desta forma, os estados também foram estimados com boa precisão, conforme pode ser notado nas Figuras 3.6 e 3.7.

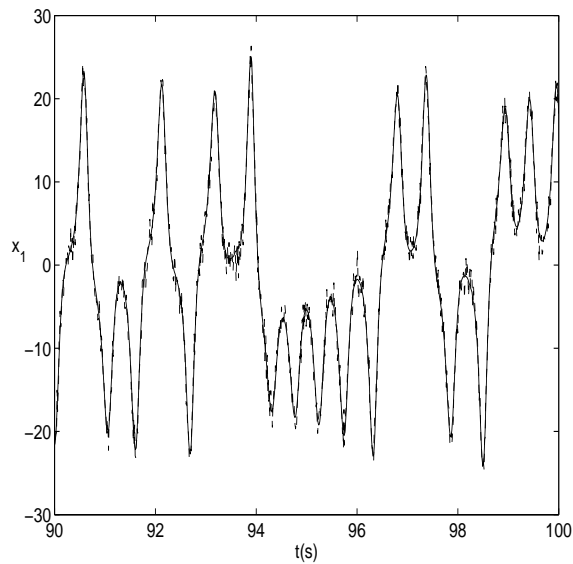


Figura 3.6: Filtragem de sinal ruidoso – Lorenz. Sinal simulado (---) e estimado (—).

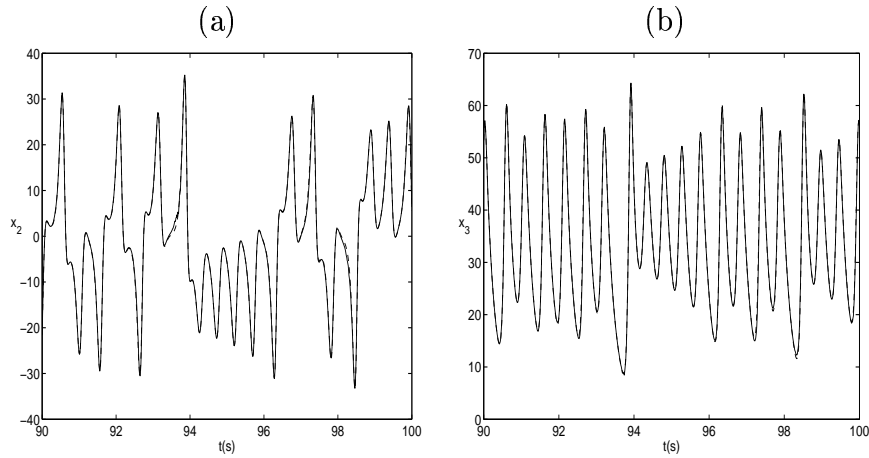


Figura 3.7: Evolução do estados simulados (---) e estimados (—), (a) x_2 e (b) x_3 .

Os gráficos da Figura 3.7, se referem aos estados x_2 e x_3 . As linhas que descrevem as variáveis simuladas e estimadas encontram-se sobrepostas como resultado da boa precisão obtida na estimação. Os erros percentuais máximos de estimação, em relação aos valores máximos das séries temporais x_1 , x_2 e x_3 , foram de 0,88 %, 1,15 % e 0,62 %, respectivamente.

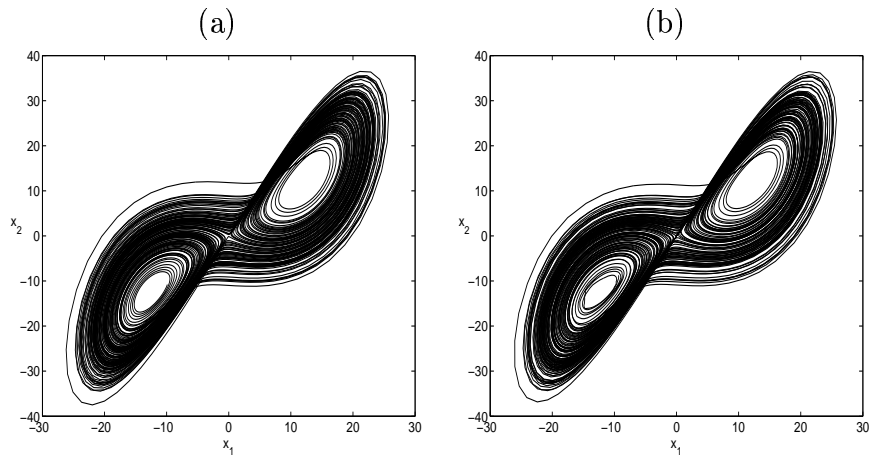


Figura 3.8: (a) Atrator do sistema de Lorenz original e (b) estimado.

A Figura 3.8 mostra o atrator do sistema de Lorenz. A partir da comparação entre os atratores original e estimado, também é possível perceber o quão precisa foi a estimação.

3.4 Conclusões

Neste capítulo foram apresentados dois algoritmos recursivos para estimação de estados em sistemas não-lineares baseados no KF. São eles: O Filtro de Kalman Estendido (EKF) e o Filtro de Kalman *Unscented* (UKF).

O EKF (seção 3.2) é um algoritmo bastante conhecido e, provavelmente, é a variante do KF mais utilizada. Por este motivo, conclui-se que a técnica é eficaz para muitas aplicações. Por outro lado sabe-se que em determinadas aplicações podem ocorrer problemas de imprecisão e divergência, como consequência dos procedimentos de linearização de funções não-lineares, presentes no algoritmo.

O UKF, revisto na seção 3.3, foi proposto com a promessa de proporcionar resultados mais precisos, com menos custo computacional e de mais fácil implementação do que o EKF (Julier and Uhlmann, 1997). Na seção 3.3.5 reproduziram-se os resultados encontrados em (Sitz et al., 2002), nos quais o UKF foi utilizado para estimação de estados e parâmetros proporcionando bons resultados.

Embora tenha sido possível obter resultados promissores com o UKF, alguns problemas foram verificados durante o procedimento. O algoritmo é muito sensível à inicialização da matriz de covariância. Esta característica se tornou mais evidente no sistema de Lorenz. Esta questão será tratada de forma mais detalhada no capítulo 4, no qual procura-se caracterizar melhor o problema e discutir possíveis soluções.

Estimação Recursiva: Aplicação ao Circuito de Chua

“A falsa ciência gera ateus; a verdadeira ciência leva os homens a se curvarem diante da divindade.”

Voltaire

4.1 Introdução

Um problema de engenharia de grande relevância, é o controle de sistemas via realimentação de estados (Phillips and Nagle, 1995). No caso de sistemas lineares por exemplo, pode-se aplicar a técnica da *imposição de pólos*. A aplicação dessa técnica começa com a determinação dos pólos de malha fechada desejados baseada nos requisitos de resposta transitória ou resposta em frequência, tais como velocidade de resposta, sobre-sinal máximo, tempo de acomodação, dentre outros, bem como nos requisitos de regime estacionário. Escolhendo-se uma matriz de ganho apropriada para realimentação de estados, é possível forçar o sistema a ter pólos de malha fechada nas posições desejadas, contanto que o sistema original seja completamente controlável (Kailath, 1980; Chen, 1984).

Uma condição para que se possa aplicar essa técnica de controle é que todas as variáveis de estado sejam mensuráveis e estejam disponíveis para realimentação. Em muitos casos, por razões práticas, torna-se impossível medir todos os estados. Neste caso, surge a necessidade de estimá-los. Essa

necessidade motivou o desenvolvimento de técnicas determinísticas para a estimação de estados (Luenberger, 1966).

Em aplicações práticas, é comum que mesmo os estados que podem ser diretamente acessados, sejam medidos com algum ruído. Essas distorções nos sinais medidos podem ser provenientes da falta de precisão dos próprios dispositivos de medição. Nesses casos, procedimentos determinísticos como o observador de Luenberger devem ser substituídos por técnicas estocásticas. Tais métodos são capazes de lidar com o ruído presente nos dados de maneira a conseguir estimar os estados a despeito do ruído. Por essa razão, tais algoritmos são chamados de filtros, sendo que o Filtro de Kalman é possivelmente o exemplo mais conhecido.

Neste capítulo serão tratados os problemas de filtragem de sinais ruidosos e estimação de estados desconhecidos do circuito de Chua. Este circuito eletrônico é um sistema não-linear capaz de apresentar diferentes comportamentos dinâmicos, inclusive regime caótico. A fim de comparar desempenho, os dois algoritmos descritos no capítulo anterior, o Filtro de Kalman Estendido (EKF) e o Filtro de Kalman Unscented (UKF), são utilizados na solução desses problemas de estimação. A seguir será dada uma breve explicação sobre *caos*, que pode ser útil na análise do circuito de Chua.

4.2 Sistemas Caóticos

Sistemas *caóticos* apresentam comportamentos de extrema complexidade. Esses sistemas são muito sensíveis às condições iniciais, o que faz com que sejam totalmente imprevisíveis mesmo conhecendo-se as equações que descrevem seu comportamento.

O complexo comportamento caótico pode ser observado em sistemas representados por equações simples. Isso tem motivado a busca por modelos não-lineares capazes de prever o comportamento de fenômenos que antes eram considerados como puramente aleatórios (Gleick, 1987).

O estudo de tais sistemas teve início com as pesquisas de Henri Poincaré, no século XIX, que investigava as possíveis trajetórias de três corpos no espaço sujeitos a forças gravitacionais entre eles (Poincaré, 1957). A partir do século seguinte, com a descoberta de novas ferramentas matemáticas e com o advento de sistemas de computação mais poderosos, a teoria do caos passou a ser estudada com mais afinco (Lorenz, 1963).

Um texto claro e objetivo que descreve brevemente conceitos introdutórios

sobre sistemas caóticos, além de alguns exemplos, pode ser encontrado em Tôres (2001). É possível também encontrar abordagens mais detalhadas sobre o assunto, como por exemplo em (Fiedler-Ferrara and Prado, 1994; Monteiro, 2002).

4.3 O Circuito de Chua

O Circuito de Chua (Matsumoto, 1984; Chua, 1994) é um sistema caótico bastante conhecido e muito útil na análise de dinâmica não-linear e caos devido à sua robustez, relativa facilidade de implementação (Tôres and Aguirre, 1996; Tôres, 2001) e grande variedade de regimes dinâmicos (Tôres, 2001). O circuito de Chua é um circuito eletrônico capaz de exibir oscilações periódicas ou caóticas (Kennedy, 1992). Desde sua criação, o circuito de Chua tem se tornado uma plataforma de testes para os mais diversos algoritmos e sistemas de controle. Não é de se admirar que existam diversas realizações distintas (Kiliç, 2003). Maiores detalhes do circuito são apresentados a seguir.

4.3.1 Descrição do Sistema

O circuito de Chua, representado na Figura 4.1, é composto por dois capacitores C_1 e C_2 , um resistor variável R , um indutor L e um elemento não-linear conhecido como *diodo de Chua*. O diodo de Chua possui resistência negativa e é o elemento que confere energia ao sistema, permitindo assim, manter o circuito oscilando autonomamente. O diodo de Chua, que pode ser implementado na prática utilizando-se amplificadores operacionais e outros componentes simples (Tôres and Aguirre, 2000), apresenta a característica linear por partes mostrada na Figura 4.2.

As equações que descrevem o circuito são (Chua, 1994):

$$\begin{cases} C_1 \frac{dv_{C_1}}{dt} = \frac{(v_{C_2} - v_{C_1})}{R} - i_d(v_1), \\ C_2 \frac{dv_{C_2}}{dt} = \frac{(v_{C_1} - v_{C_2})}{R} + i_L, \\ L \frac{di_L}{dt} = -v_{C_2}, \end{cases} \quad (4.1)$$

sendo que v_{C_1} é a tensão sobre o capacitor C_1 , v_{C_2} a tensão sobre o capacitor

C_2 e i_L a corrente que passa através do indutor.

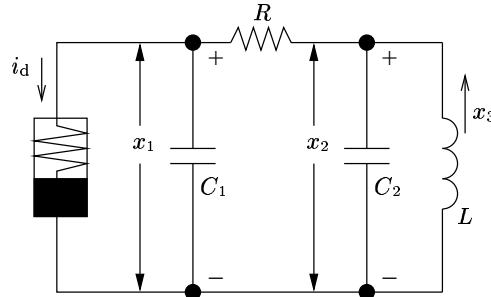


Figura 4.1: Circuito de Chua.

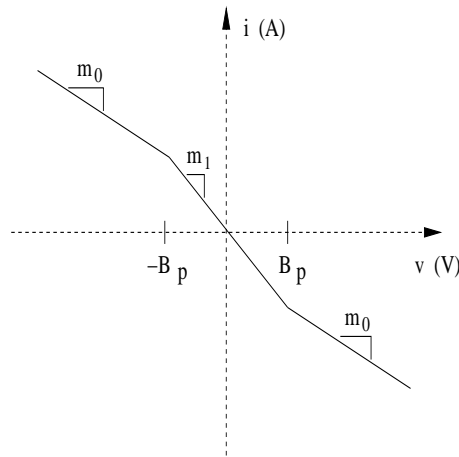


Figura 4.2: Curva característica do diodo de Chua. $i \equiv i_d$ e $v \equiv x_1$.

A corrente que passa pelo diodo de Chua pode ser expressa por:

$$i_d(v_{C_1}) = \begin{cases} m_0 v_{C_1} + B_p(m_0 - m_1) & \text{para } v_{C_1} \leq -B_p, \\ m_1 v_{C_1} & \text{para } -B_p < v_{C_1} < B_p, \\ m_0 v_{C_1} + B_p(m_1 - m_0) & \text{para } v_{C_1} \geq B_p, \end{cases} \quad (4.2)$$

ou, equivalentemente, pela seguinte função (Cruz et al., 2001):

$$i_d(v_{C_1}) = m_0 v_{C_1} + \frac{1}{2}(m_1 - m_0)(|v_{C_1} + B_p| - |v_{C_1} - B_p|). \quad (4.3)$$

Representando o circuito de Chua em espaço de estados, obtém-se as

seguintes equações:

$$\begin{cases} \dot{x}_1 = -\frac{1}{RC_1}x_1 + \frac{1}{RC_1}x_2 - \frac{1}{C_1}i_d(x_1), \\ \dot{x}_2 = \frac{1}{RC_2}x_1 - \frac{1}{RC_2}x_2 + \frac{1}{C_2}x_3, \\ \dot{x}_3 = -\frac{1}{L}x_2, \end{cases} \quad (4.4)$$

sendo $x_1 = v_{C_1}$, $x_2 = v_{C_2}$, e $x_3 = i_L$.

4.3.2 Montagem do Circuito Real

Neste trabalho foram coletados dados a partir de uma implementação real do circuito de Chua. Trata-se do Protótipo para Controle do Circuito de Chua - PCChua (Tôres, 2001). Os valores dos parâmetros selecionados para a montagem estão listados na Tabela 4.1.

Tabela 4.1: Parâmetros da plataforma PCChua.

Parâmetro	Valor
R	1,5 k Ω a 2,0 k Ω
C_1	(25 \pm 1) μ F
C_2	(236 \pm 1) μ F
L	(42,3 \pm 0,5) H
m_0	(-0,406 \pm 0,001) mS
m_1	(-0,755 \pm 0,001) mS
B_p	1,87 \pm 0,01 V

A resistência variável R foi escolhida em uma faixa que permite a exibição de diversos regimes dinâmicos. Desta forma, variando-se o valor de R , é possível fazer com que o sistema oscile caoticamente.

A descrição completa da montagem e a razão pela qual os valores dos parâmetros foram selecionados conforme listado na Tabela 4.1, se encontram na tese de doutorado de Leonardo Tôres (Tôres, 2001). Salienta-se que o alto valor da indutância L foi obtido utilizando-se um circuito eletrônico ativo que se comporta como um indutor na faixa de frequências de interesse para o circuito de Chua (Tôres and Aguirre, 2000). O diagrama esquemático da montagem do circuito pode ser visto na Figura 4.3, onde é possível verificar como foram implementados o diodo de Chua e o indutor simulado.

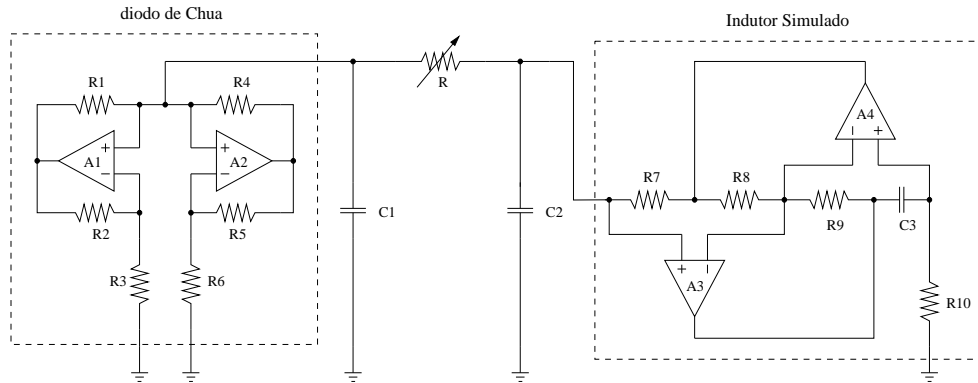


Figura 4.3: Diagrama de montagem do PC-CHUA (Tôrres, 2001).

4.4 Estimando Estados

Foi visto na seção anterior que o circuito de Chua possui três variáveis de estado. O objetivo, neste exemplo, é estimar os valores desses três estados à medida que o sistema evolui. Será suposto que o estado x_1 pode ser medido diretamente, enquanto que os demais, x_2 e x_3 , devem ser estimados.

4.4.1 Metodologia

Os dados utilizados na realização deste experimento foram gerados de duas maneiras distintas. No primeiro caso as equações diferenciais do sistema foram simuladas de forma a se obter os dados desejados. Já no segundo caso, dados foram coletados diretamente da plataforma PCChua, descrita na seção 4.3.2.

Os valores dos parâmetros do protótipo real foram utilizados também na simulação (primeiro caso). Esses parâmetros, que foram determinados por Tôrres (2001), encontram-se listados na Tabela 4.1. O valor da resistência variável R foi escolhido garantindo que o sistema exibisse um comportamento caótico. O valor selecionado foi $R = 1659 \Omega$.

Em ambos os casos, os estados foram estimados pelos dois algoritmos estudados no capítulo anterior, o Filtro de Kalman Estendido (EKF) (seção 3.2) e o Filtro de Kalman Unscented (UKF) (seção 3.3).

Critério de Avaliação do Erro

A fim de comparar o desempenho dos algoritmos EKF e UKF, ao estimar os estados do sistema nos dois casos estudados, real e simulação, foi usado um critério para avaliação do erro de estimação. Esse critério se baseia na *raiz quadrada do erro quadrático médio* (RMSE).¹

O valor RMSE, para cada variável de estado, é dado por:

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N (x_k - \hat{x}_k)^2}, \quad (4.5)$$

sendo x_k o valor real da variável na iteração k , \hat{x}_k o valor estimado e N o número total de amostras.

Como os algoritmos em estudo são recursivos, o erro de estimação pode atingir valores extremamente elevados no início do processo de estimação. A matriz de covariância dos estados estimados varia a cada iteração do algoritmo e o erro de estimação tende a diminuir até a convergência dessa matriz. Desta forma, para que se possa ter uma idéia melhor do desempenho dos dois estimadores em cada caso, o valor RMSE foi calculado após a convergência da matriz de covariância dos estados.

Para verificar a convergência da matriz de covariância, determinou-se a norma *euclídeana*² do vetor constituído pelos elementos da diagonal da matriz de covariância. Calcula-se, a cada iteração, a diferença entre a norma desse vetor no instante atual e a norma no instante anterior. Quando essa diferença for menor que um limite pré-definido, considera-se que a convergência da matriz foi atingida. Este limite deve ser um número bem próximo de zero, o qual pode ser determinado por meio de experimentos.

Algumas variáveis, por apresentarem baixa ordem de grandeza, podem deixar a falsa impressão de que o erro de estimação é pequeno. Para contornar este problema, o RMSE também foi calculado em relação ao máximo valor absoluto da série temporal correspondente à variável estimada. Este índice foi chamado de RMSE percentual (PRMSE) e é dado por:

¹RMSE – Root Mean Square Error.

²A norma *euclídeana* de um vetor é definida pela raiz quadrada da soma dos elementos do vetor ao quadrado (Chen, 1984).

$$PRMSE = RMSE \times \frac{100}{|x_{\max}|}, \quad (4.6)$$

sendo $|x_{\max}|$ o maior valor absoluto do sinal real.

4.4.2 Sistema Simulado

Neste primeiro caso os dados foram gerados a partir da simulação das equações diferenciais do sistema. A integração das equações foi feita usando o método de *Runge-Kutta* durante um período de 100 segundos, com intervalo de integração variável e intervalo de amostragem $T_s = 0,01s$. Para aproximar uma medição contaminada por ruído, adicionou-se ruído gaussiano ao estado x_1 . O desvio padrão do ruído adicionado equivale a 5 % do desvio padrão da série temporal correspondente ao estado x_1 .

Estimação de estados usando EKF

A curva característica do diodo de Chua possui dois pontos de derivada descontínua (Figura 4.2), os quais podem provocar erro no resultado final da estimação. Para evitar esse problema foi feita uma aproximação polinomial para tal curva conforme a seguinte equação:

$$i_d = \alpha_1 x_1^3 + \alpha_2 x_1^2 + \alpha_3 x_1 + \alpha_4, \quad (4.7)$$

sendo α_1 , α_2 , α_3 e α_4 os coeficientes do polinômio de terceiro grau ajustado aos dados sobre a não-linearidade descontínua. A substituição da não-linearidade original por uma suave é procedimento comum na literatura (Koslov et al., 1996; Zhong, 1994).

Na Figura 4.4 são mostradas curvas características do diodo de Chua obtidas a partir da equação verdadeira (4.3) e a partir do polinômio interpolado dado pela equação (4.7).

Para dar início ao processo de estimação é necessário definir condições iniciais para o vetor de estados e para a matriz de covariância. O vetor de estados escolhido foi:

$$\hat{\mathbf{x}}_0 = \begin{bmatrix} x_1 \\ x_1 \\ x_1 \end{bmatrix}, \quad (4.8)$$

sendo que y_1 é a primeira amostra do sinal coletado do sistema simulado (primeira variável de estado contaminada artificialmente com ruído).

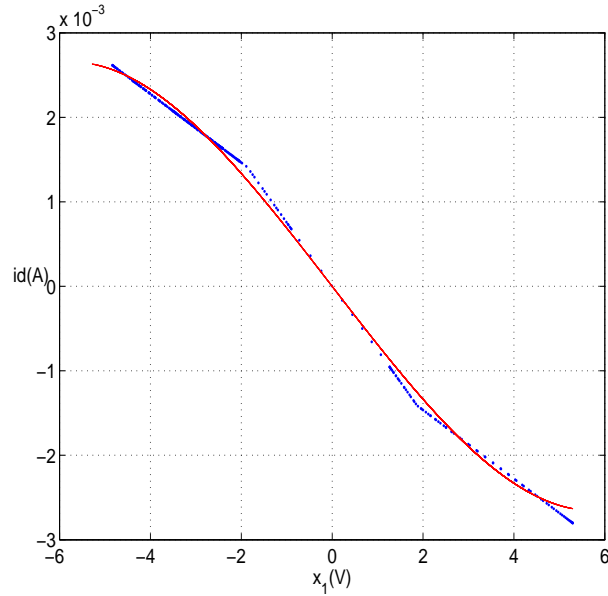


Figura 4.4: Curva característica do diodo de Chua obtida pela equação verdadeira ($\cdot\cdot\cdot$) e por interpolação polinomial ($—$).

A matriz de covariância inicial usada foi:

$$P_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.9)$$

A escolha da matriz de covariância inicial não teve grande influência na convergência do algoritmo. Algumas tentativas foram feitas, reduzindo ou aumentando os elementos da diagonal desta matriz, mas essas variações não proporcionaram nenhum resultado melhor do que aquele obtido quando a escolha foi uma matriz identidade.

Os erros entre os estados estimados pelo EKF e os estados simulados foram avaliados de acordo com o critério RMSE, descrito na seção 4.4.1. Os valores RMSE (absoluto e percentual) encontram-se listados na Tabela 4.2. Nas Figuras 4.5 e 4.6, mostra-se a evolução dos estados estimados.

É possível constatar que existe um erro sistemático entre estado simulado e estimado. Este resultado pode ser constatado tanto no estado x_1 utilizado como sinal de entrada para o algoritmo EKF (Figura 4.5), quanto nos estados

Tabela 4.2: Valor RMSE na estimação de estados via EKF – dados simulados.

Estado	RMSE absoluto	PRMSE
x_1	0,82 V	15,50 %
x_2	0,17 V	15,33 %
x_3	$6,09 \times 10^{-4}$ A	12,42 %

x_2 e x_3 que foram somente estimados³ (Figura 4.6).

Supõe-se que a origem do problema reside no fato de que as verdadeiras média e covariância dos estados não tenham sido propagadas corretamente devido ao procedimento de linearização presente no EKF, provocando erro no resultado final da estimação. É importante ressaltar que o erro ocorreu mesmo com a aproximação polinomial para a curva característica do diodo de Chua, dada pela equação (4.2). Não houve variação significativa entre os resultados obtidos com ou sem aproximação polinomial.

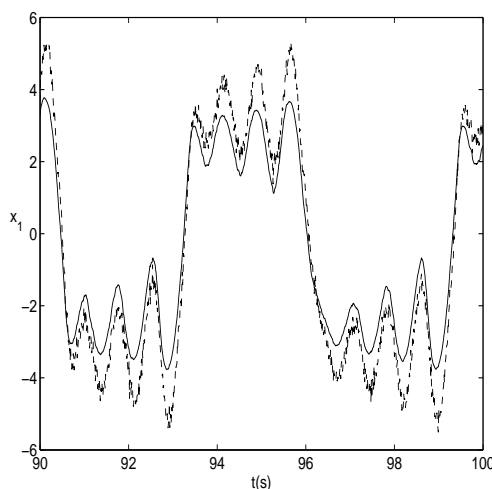


Figura 4.5: Filtragem de sinal ruidoso via EKF – dados simulados. Estado x_1 simulado com ruído (---) e estimado (—).

O pobre desempenho do EKF neste exemplo motivou a busca por uma nova técnica na literatura que conduzisse a resultados melhores, conforme será visto a seguir.

³Inicialmente foi feita a suposição de que os estados x_2 e x_3 não podem ser medidos diretamente. Entretanto, seus valores foram usados para fins de comparação com os valores estimados.

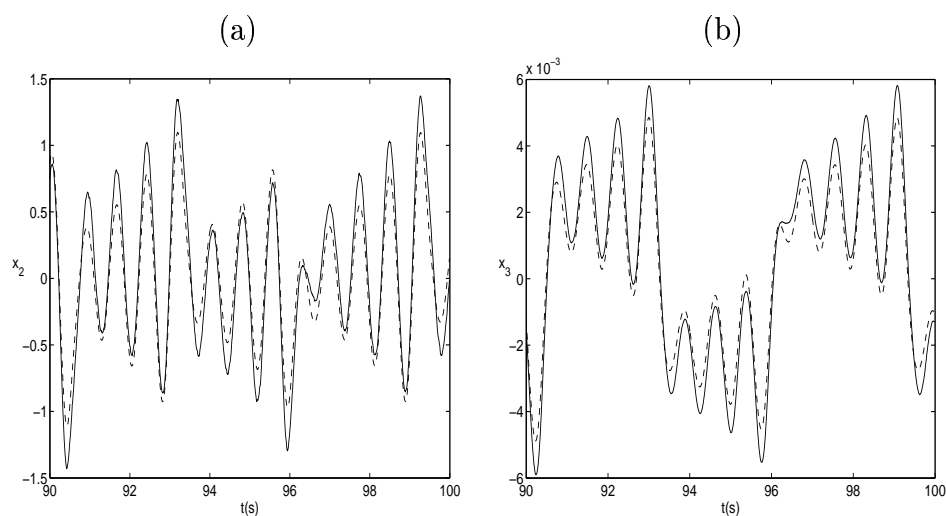


Figura 4.6: Estimação de estados via EKF. Estado simulado (---) e estimado (—) (a) x_2 e (b) x_3 .

Estimação de estados usando UKF

A inicialização do algoritmo foi feita da mesma maneira que no EKF. A determinação de valores iniciais para o vetor de estados estimado e para matriz de covariância praticamente não produziu efeito significativo na convergência do algoritmo. Escolhas diferentes para esses valores iniciais apenas levaram o algoritmo a convergir ligeiramente mais ou menos rápido, mas sempre garantindo a convergência desejada.

O desempenho do algoritmo UKF foi superior ao do EKF, como pode ser visto nas Figuras 4.7 e 4.8, e na Tabela 4.3, na qual são mostrados os erros de acordo com o critério RMSE.

Tabela 4.3: Valor RMSE na estimação de estados via UKF – dados simulados.

Estado	RMSE absoluto	PRMSE
x_1	0,07 V	1,38 %
x_2	0,02 V	2,22 %
x_3	$6,96 \times 10^{-5}$ A	1,42 %

O gráfico da Figura 4.7 refere-se ao estado x_1 , ou seja, sinal simulado com adição de ruído. As curvas referentes aos estados simulados e estimados encontram-se praticamente sobrepostas, destacando a exatidão atingida pelo

algoritmo (Figura 4.8). Comparando-se os atratores simulado e estimado, também é possível confirmar esse resultado (Figura 4.9).

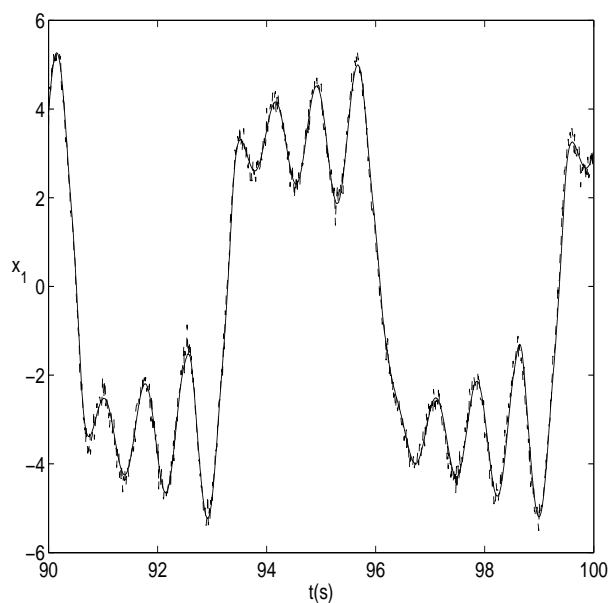


Figura 4.7: Filtragem de sinal ruidoso via UKF. Estado x_1 simulado com ruído (---) e estimado (—).

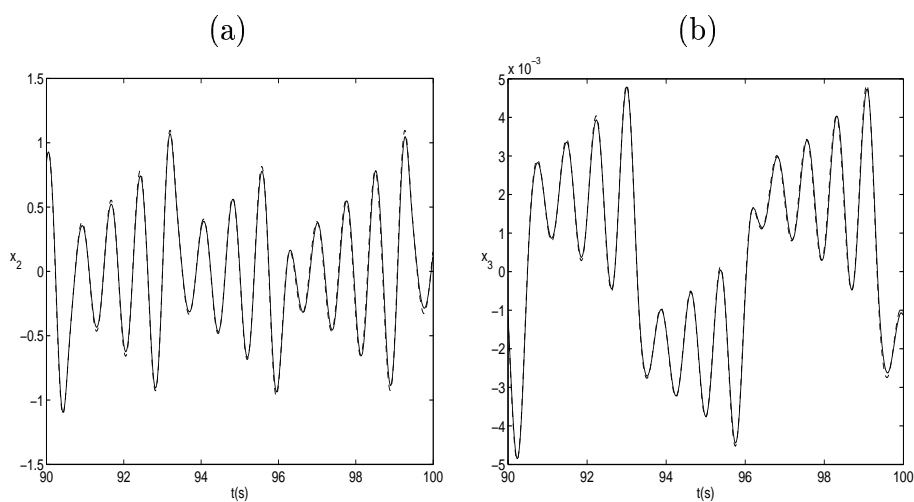


Figura 4.8: Estados estimados via UKF. Estado simulado (---) e estimado (—)
(a) x_2 e (b) x_3 .

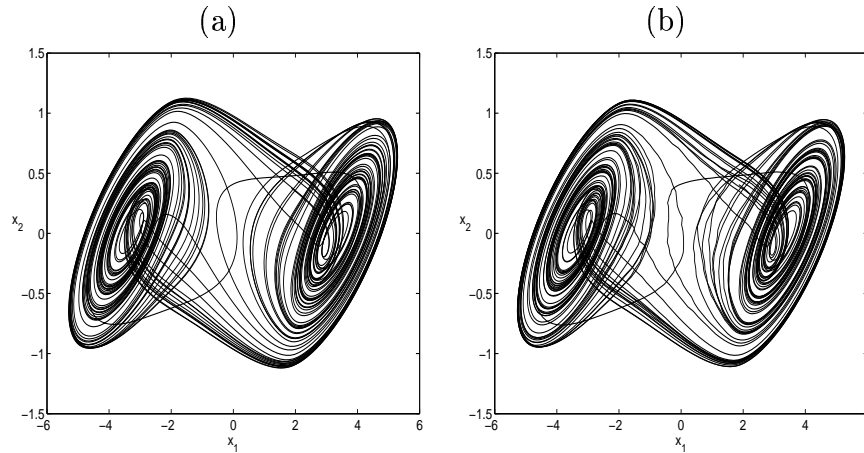


Figura 4.9: Projeção do atrator dupla volta do circuito de Chua simulado (a) e estimado (b).

Variando o Nível de Ruído

A relação sinal/ruído (SNR)⁴ pode ser definida como a razão entre o desvio padrão da série temporal do sinal e o desvio padrão do ruído, conforme a seguinte equação:

$$\text{SNR} = 20 \log \left(\frac{\sigma_x}{\sigma_e} \right), \quad (4.10)$$

sendo σ_x o desvio padrão do sinal x , σ_e o desvio padrão do ruído e e a unidade de SNR é decibéis (dB).

Nesse caso simulado, em que apenas os estados e não os parâmetros do circuito são estimados, foram feitas algumas simulações com diferentes níveis de ruído, a fim de avaliar a robustez do desempenho dos filtros com respeito ao ruído. Os resultados são apresentados nas Figuras 4.10 e 4.11.

As Figuras 4.10 e 4.11 mostram os valores do erro percentual PRMSE em função da relação sinal ruído (SNR) das séries temporais x_1 , x_2 e x_3 . Para fins de comparação, os erros foram calculados a partir das séries temporais estimadas pelos algoritmos EKF e UKF. Pode ser notado que mesmo aumentando consideravelmente o ruído, os algoritmos ainda são capazes de convergir.

Conclui-se que o UKF produz estimativas melhores do que o EKF desde que a relação sinal/ruído seja maior do que 5 dB. Salienta-se que, para

⁴SNR – *Signal/noise ratio*.

valores de SNR menores do que 5 dB, podem ser encontrados valores de PRMSE diferentes daqueles mostrados nas Figuras 4.10 e 4.11 a cada nova realização do ruído. Desta forma, não se pode mais confiar nos resultados produzidos pelo UKF quando SNR for menor do que 5 dB, para este caso particular do circuito de Chua.

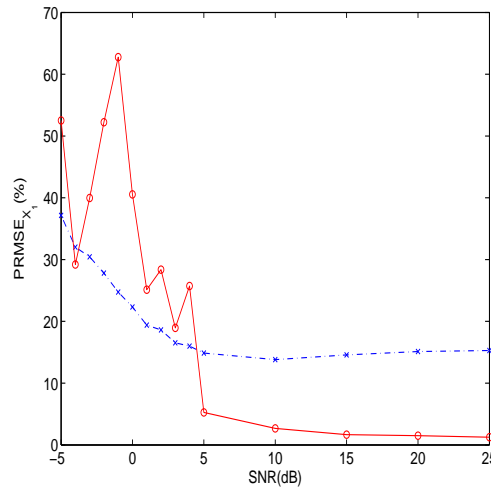


Figura 4.10: Relação sinal/ruído \times PRMSE para x_1 . Estimação via EKF ($- \times -$) e estimação via UKF ($- o -$).

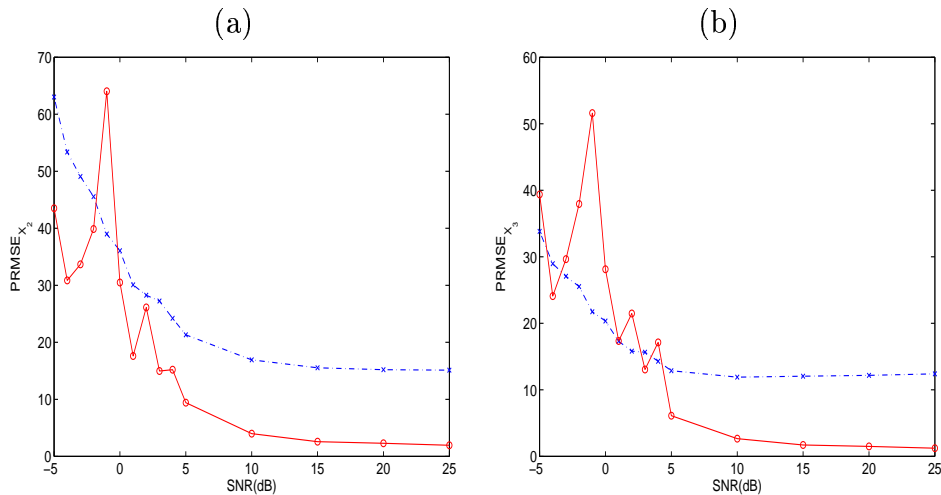


Figura 4.11: Relação sinal/ruído \times PRMSE – (a) x_2 e (b) x_3 . Estimação via EKF ($- \times -$) e estimação via UKF ($- o -$).

Na figura 4.12 mostra-se como fica a série temporal do estado x_1 com adição de ruído gaussiano com $\text{SNR} = 5$ dB.

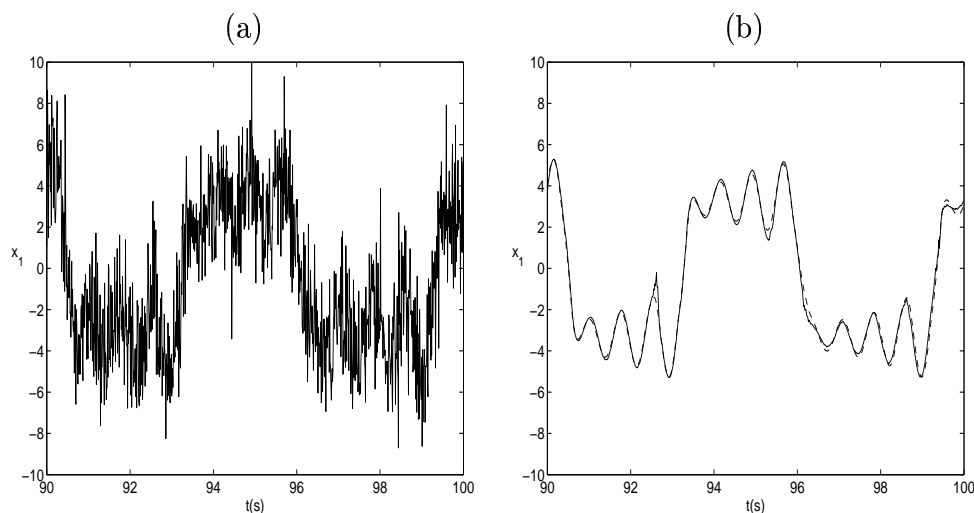


Figura 4.12: (a) Série temporal do estado x_1 com ruído gaussiano adicionado com $\text{SNR} = 5$ Db. (b) Série temporal do estado x_1 sem ruído (---) e estimada por UKF (—).

4.4.3 Plataforma PCChua

No segundo caso estudado, os algoritmos EKF e UKF foram executados com dados obtidos a partir do circuito de Chua real, descrito na seção 4.3.2.

Os dados foram coletados com o sistema operando em regime caótico. Da mesma maneira que no caso anterior, a variável medida foi a tensão no capacitor C_1 , ou seja, $y = x_1$, enquanto que as outras duas foram apenas estimadas. Os resultados obtidos são apresentados na Tabela 4.4 e nas Figuras 4.13 e 4.14.

Tabela 4.4: Valor RMSE – estimação via EKF/UKF – Dados medidos.

Estado	PRMSE (EKF)	PRMSE (UKF)
x_1	16,17 %	34,27 %
x_2	18,27 %	27,61 %
x_3	15,26 %	25,96 %

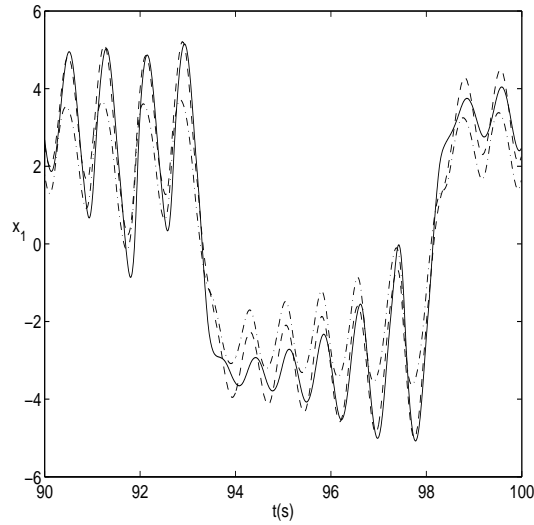


Figura 4.13: Filtragem de sinal ruidoso via EKF/UKF – dados medidos. Estado x_1 real (---), estimado com EKF (- · -) e estimado com UKF (—).

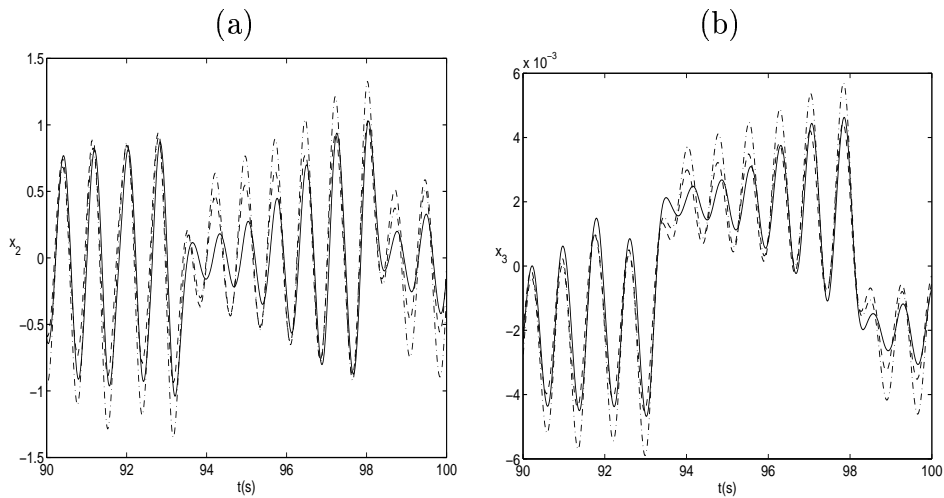


Figura 4.14: Estimação de estados via EKF/UKF – dados medidos. Estado real (---), estimado com EKF (- · -) e estimado com UKF (—) (a) x_2 e (b) x_3 .

Para fins de melhor visualização e comparação de desempenho entre os algoritmos, as curvas referentes ao estado x_1 real, estimado com EKF e estimado com UKF são apresentadas no mesmo gráfico (Figura 4.13). O resul-

tado proporcionado pelo EKF apresentou um erro ainda maior que no caso anterior como já era esperado, pois como se tratam de dados reais existem mais fontes de incerteza do que no caso do sistema simulado. Ao contrário do que se esperava, o UKF produziu estimativas piores que aquelas fornecidas pelo EKF, diferentemente do que ocorreu quando os dados foram gerados por simulação. Um resultado análogo pode ser observado na Figura 4.14, na qual mostra-se as séries temporais das variáveis x_2 e x_3 .

A explicação para este resultado pode estar relacionada com os valores dos parâmetros. Nos dois casos (simulado e real), foi suposto que os parâmetros são conhecidos e o UKF foi usado para estimar apenas os estados. No primeiro caso, os mesmos valores de parâmetros foram utilizados tanto na simulação das equações diferenciais para gerar os dados, quanto no algoritmo do UKF para realizar a estimação dos estados. No segundo caso, como os dados utilizados foram reais, não se pode garantir que todos os valores dos parâmetros tenham sido medidos no circuito real com exatidão, ou que estes não tenham sofrido alterações desde o seu levantamento original.

A possibilidade de haver divergência entre os parâmetros reais e os que foram utilizados no algoritmo pode ter impossibilitado o UKF de fornecer boas estimativas para os estados do sistema. Uma possível solução para este problema será proposta na próxima seção.

4.5 Estimação Conjunta de Estados e Parâmetros

Na seção 4.4.2 mostrou-se que o UKF foi superior ao EKF ao fornecer melhores estimativas para os estados do circuito de Chua quando os dados foram gerados por simulação das equações diferenciais do sistema. No caso de dados reais esse resultado não se repetiu.

A falha foi atribuída à possibilidade de se estar utilizando nos algoritmos, valores de parâmetros diferentes dos reais. Uma alternativa para contornar esse obstáculo é estimar os parâmetros juntamente com os estados. Os resultados desse procedimento serão discutidos nesta seção. O algoritmo escolhido foi o UKF, que foi mais eficaz do que o EKF no caso de dados simulados, além de ser capaz de realizar estimação conjunta de maneira relativamente simples, conforme descrito na seção 3.3.3.

4.5.1 Metodologia

Algumas modificações devem ser feitas antes de executar o algoritmo UKF. Os parâmetros, que antes eram conhecidos, tornam-se incógnitas e o circuito passa a ser representado pelas seguintes equações:

$$\left\{ \begin{array}{l} \dot{x}_1 = \lambda_1 (-x_1 + x_2 - \lambda_4 i_d), \\ \dot{x}_2 = \lambda_2 (x_1 - x_2 + \lambda_4 x_3), \\ \dot{x}_3 = -(1/\lambda_3) x_2, \\ \dot{\lambda}_1 = 0, \\ \dot{\lambda}_2 = 0, \\ \dot{\lambda}_3 = 0, \\ \dot{\lambda}_4 = 0, \\ \dot{\lambda}_5 = 0, \\ \dot{\lambda}_6 = 0, \\ \dot{\lambda}_7 = 0, \end{array} \right. \quad (4.11)$$

$$i_d = \lambda_5 x_1 + 0,5(\lambda_6 - \lambda_5)(|x_1 + \lambda_7| - |x_1 - \lambda_7|), \quad (4.12)$$

sendo que λ_i são os parâmetros a serem estimados:

$$\begin{aligned} \lambda_1 &= \frac{1}{RC_1}, & \lambda_5 &= m_0, \\ \lambda_2 &= \frac{1}{RC_2}, & \lambda_6 &= m_1, \\ \lambda_3 &= L, & \lambda_7 &= B_p. \\ \lambda_4 &= R, \end{aligned}$$

Da mesma forma que anteriormente, é preciso definir condições iniciais para o vetor de estados estendido \mathbf{x}_0^+ . É muito importante que se tenha uma boa idéia da ordem de grandeza desses parâmetros “desconhecidos”. Quanto mais próximos os valores iniciais estiverem dos valores reais, melhor será a convergência do algoritmo. Antes de começar o processo de estimação, é preciso ainda inicializar a matriz de covariância do vetor de estados estendido P_0^+ . A escolha dessa matriz inicial também pode ser crucial na obtenção de convergência.

4.5.2 Resultados e Discussões

As condições iniciais escolhidas e os respectivos parâmetros estimados encontram-se listados na Tabela 4.5. Na Tabela 4.6, listam-se os parâmetros verdadeiros, ou seja, aqueles que representam grandezas físicas reais do circuito, estimados por UKF.

Tabela 4.5: Parâmetros estimados por UKF – Condições iniciais e valores estimados – Circuito de Chua.

Parâmetro	Condições iniciais	Valores estimados	Valores esperados
λ_1	30	19,827	24,11
λ_2	3	3,218	2,55
λ_3	50	52,282	42,3
λ_4	2000	1673,6	1659
λ_5	$-0,5 \times 10^{-3}$	$-0,365 \times 10^{-3}$	$-0,406 \times 10^{-3}$
λ_6	$-0,9 \times 10^{-3}$	$-0,801 \times 10^{-3}$	$-0,755 \times 10^{-3}$
λ_7	2	1,739	1,87

Tabela 4.6: Parâmetros do circuito de Chua – Valores estimados por UKF.

Parâmetro	Valor
R	1673,6 Ω
C_1	30,136 μF
C_2	185,66 μF
L	52,282 H
m_0	-0,365 mS
m_1	-0,801 mS
B_p	1,739 V

Como pode ser notado, os parâmetros do sistema convergiram para valores diferentes daqueles que foram determinados por medição na placa do circuito de Chua real. Na Figura 4.15 mostra-se a evolução dos parâmetros estimados ao longo do tempo até atingirem convergência.

As séries temporais referentes aos estados x_1 , x_2 e x_3 foram estimadas com baixos valores de PRMSE, conforme listado na Tabela 4.7. Como pode ser notado o desempenho do UKF ao estimar x_2 foi pior do que com relação às demais variáveis. Este resultado pode ser justificado pela maior dificuldade

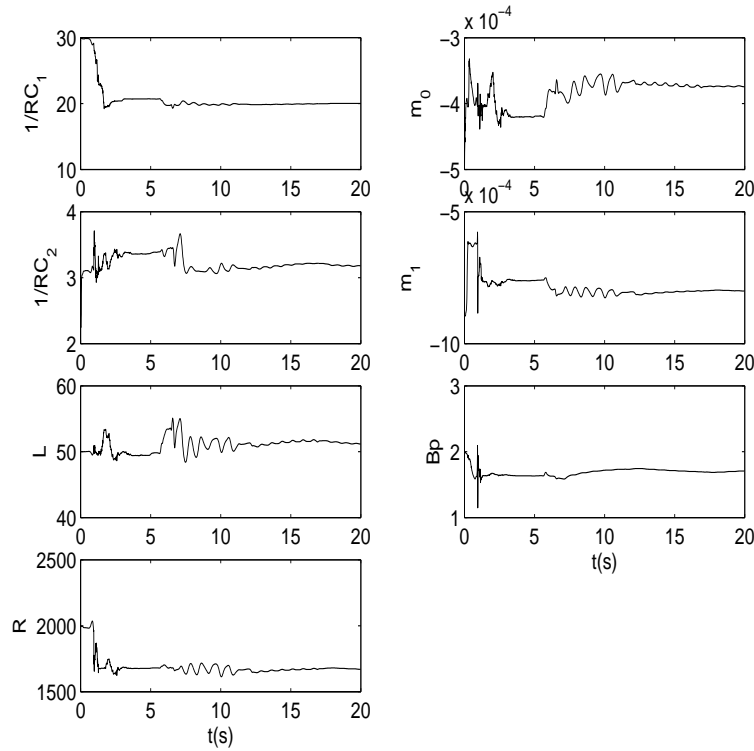


Figura 4.15: Evolução dos parâmetros do circuito de Chua na estimação.

em extrair informações do atrator utilizando-se esta variável (Aguirre et al., 1997).

Tabela 4.7: Valor PRMSE – estimação conjunta de estados e parâmetros via UKF – dados medidos.

Estado	PRMSE
x_1	1,95 %
x_2	10,85 %
x_3	1,75 %

Observando-se a Tabela 4.7 e as Figuras 4.16 e 4.17 é possível notar que o UKF foi melhor quando se realiza estimação conjunta de estados e parâmetros do circuito de Chua real em comparação com o caso em que os parâmetros são considerados conhecidos. Este fato também pode ser constatado observando-se a Figura 4.18.

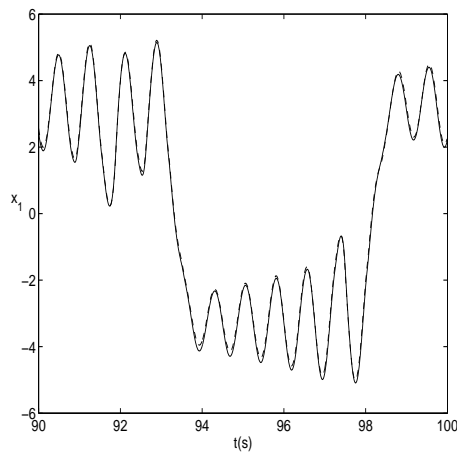


Figura 4.16: Estimação conjunta: filtragem do sinal medido via UKF – dados medidos. Estado x_1 real (---) e estimado com UKF (—).

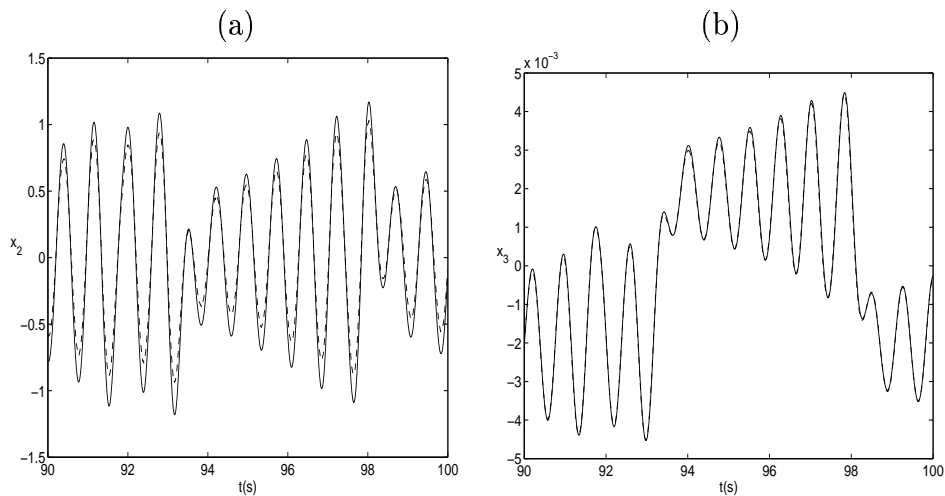


Figura 4.17: Estimação conjunta de estados e parâmetros via UKF – dados medidos. Estado real (---) e estimado com UKF (—) (a) x_2 e (b) x_3 .

A análise dos resultados sugere que os parâmetros reais do circuito de Chua, quando este foi usado para coletar os dados, diferem dos valores apresentados na Tabela 4.1. Essa suposição pode ser justificada simulando-se as equações diferenciais do circuito de Chua com os valores dos parâmetros estimados pelo UKF (Tabela 4.6).

É possível perceber que o atrator reconstruído com parâmetros estimados pelo UKF (Figura 4.19a) é muito mais semelhante ao atrator real (Figura 4.18a) do que aquele que foi simulado com parâmetros cujos valores se acreditava que estivessem corretos (Figura 4.19b).

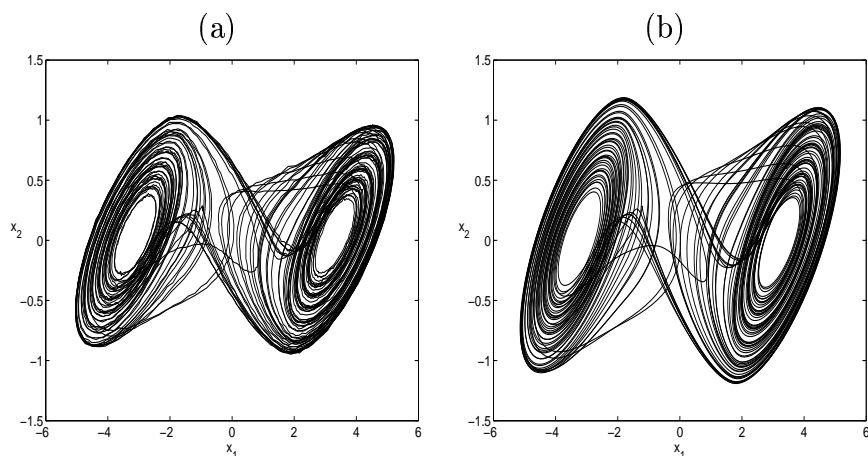


Figura 4.18: (a) Projeção do atrator dupla volta do circuito de Chua real e (b) Projeção do atrator dupla volta estimado por UKF.

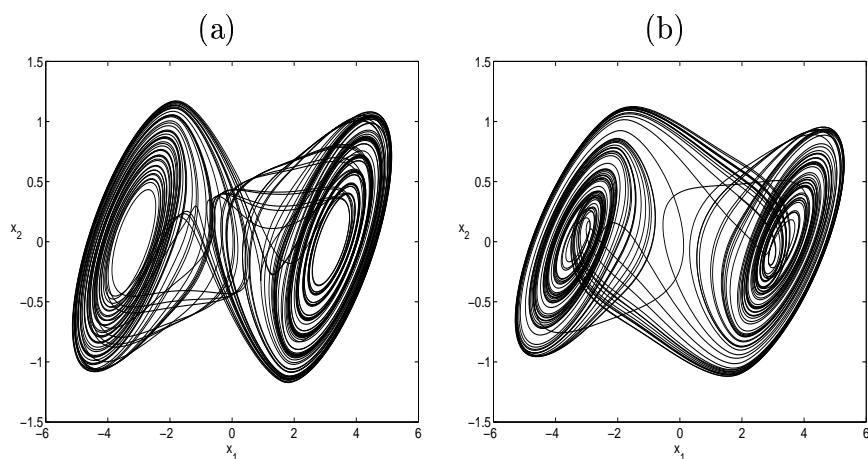


Figura 4.19: (a) Projeção do atrator dupla volta simulado com parâmetros estimados pelo UKF e (b) Projeção do atrator dupla volta simulado com parâmetros medidos no circuito de Chua real. Notar como o atrator em (a) está muito mais próximo do atrator em 4.18a do que o atrator em (b).

Estimando Novamente os Estados

Os parâmetros estimados pelo UKF (Tabela 4.6), foram inseridos no algoritmo do EKF na expectativa de se obter melhores resultados com este algoritmo do que aqueles apresentados na Tabela 4.4. Para fins de comparação o UKF também foi usado para estimar apenas os estados. Na Tabela 4.8 e nas Figuras 4.20 e 4.21 mostra-se que, mesmo com os novos valores de parâmetros, não foi possível obter melhores resultados usando o EKF.

Tabela 4.8: Valor PRMSE – estimação de estados via EKF/UKF com parâmetros previamente estimados por UKF.

Estado	PRMSE (EKF)	PRMSE (UKF)
x_1	15,91 %	2,23 %
x_2	22,47 %	9,94 %
x_3	12,06 %	1,84 %

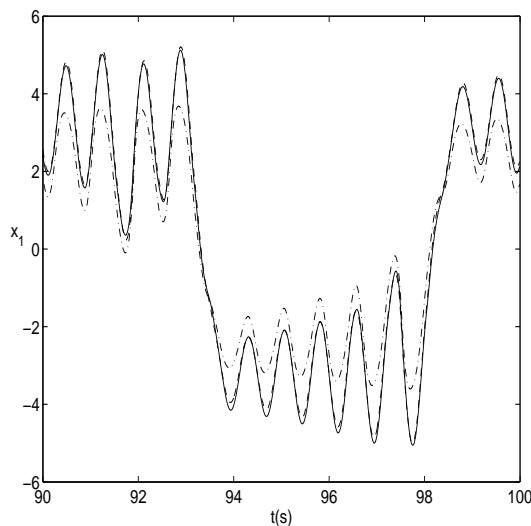


Figura 4.20: Filtragem de sinal ruidoso via EKF/UKF com parâmetros previamente estimados por UKF – dados medidos. Estado x_1 real (---), estimado com EKF (- · -) e estimado com UKF (—).

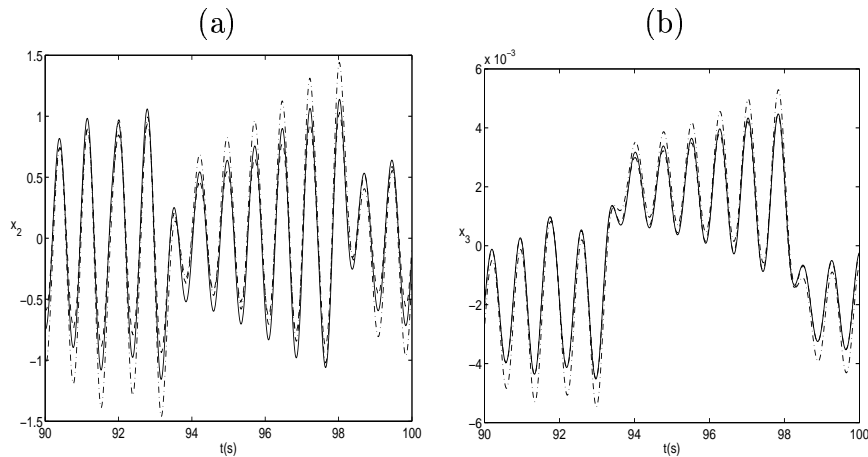


Figura 4.21: Estimação de estados via EKF/UKF com parâmetros previamente estimados por UKF – dados medidos. Estado real (---), estimado com EKF (-·-) e estimado com UKF (—) (a) x_2 e (b) x_3 .

4.5.3 O Problema da Inicialização

A inicialização do UKF teve fundamental importância na obtenção dos resultados apresentados na seção anterior. Este fato foi verificado sob dois aspectos:

1. condições iniciais para parâmetros,
2. matriz de covariância inicial.

No caso dos parâmetros, verificou-se que seus valores iniciais devem ser escolhidos os mais próximos possíveis de seus valores reais.

A matriz de covariância estendida inicial P_0^+ não interfere apenas na velocidade de convergência, como no caso do sistema simulado, em que se estimavam apenas os estados (seção 4.4.2), mas também na qualidade final dos resultados. Pôde ser notado que, ao estimar conjuntamente estados e parâmetros do circuito de Chua, a situação parece ser bem mais complexa. Uma pequena variação na matriz P_0^+ pode fazer com que o algoritmo não convirja, conduzindo a resultados totalmente incorretos, como o apresentado na Figura 4.22.

Acompanhando a evolução da variável x_3 , verifica-se que o algoritmo diverge muito rapidamente. Nos primeiros 10 segundos de estimação já se observa a presença de valores muito distantes do valor real da variável x_3 .

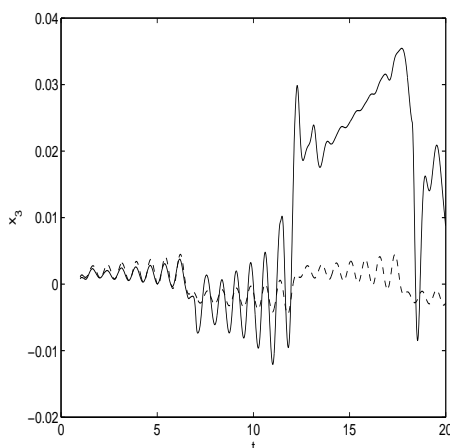


Figura 4.22: Estimação de estado via UKF – resultado divergente. Estado x_3 real (---) e estimado com UKF (—).

Possíveis Causas

As observações feitas com relação à inicialização foram tomadas principalmente com base nos testes feitos no circuito de Chua, um sistema que apresenta comportamento bastante complexo. Além disso, o UKF foi usado para estimar um grande número de parâmetros, inclusive aqueles que fazem parte da função não-linear do diodo de Chua.

Uma possível causa para o problema pode estar no nível de ruído. Nos casos simulados, onde foi possível modificar o nível de ruído, foi notado que para relações sinal/ruído (SNR) maiores, ou seja, quando se tem menos ruído nos dados, torna-se mais fácil encontrar uma matriz de covariância inicial que garanta convergência. Por outro lado, o problema foi verificado com os dados reais do circuito de Chua, cujo nível de ruído é considerado baixo.

O problema de sensibilidade à inicialização também foi verificado nos exemplos apresentados ao final do capítulo anterior. Apesar de não ter sido tão problemático quanto no circuito de Chua, a alta sensibilidade também foi verificada no sistema caótico de Lorenz. O mesmo não ocorreu no sistema de Lotka-Volterra, onde só a velocidade de convergência foi afetada. Desta forma, isso sugere que cada caso deve ser avaliado isoladamente.

É Possível Garantir Convergência?

Infelizmente ainda não existe uma regra para se escolher a matriz de covariância inicial, conforme já declarado na seção 3.3.4. Este problema também foi recentemente relatado em (Sitz et al., 2002) e ainda é objeto de estudo.

A partir da experiência adquirida durante a realização dos diversos testes ao longo deste trabalho, foi possível selecionar uma matriz de covariância para a qual o algoritmo convergiu, produzindo os resultados apresentados na seção 4.5.2. A escolha dessa matriz se deu por tentativa e erro. Foi possível notar que os valores dos elementos da diagonal da matriz de covariância estendida P_0^+ não podem ser muito elevados, ao contrário do que pode ocorrer em outros estimadores recursivos, como o próprio EKF. A sugestão é que seja observada a ordem de grandeza dos elementos do vetor de estados estendido \mathbf{x}^+ , de forma que estes sejam maiores (cerca de 10 a 100 vezes) do que os elementos da diagonal de P_0^+ .

Uma maneira de se garantir convergência, mas que pode ser difícil na prática, está relacionada ao nível de ruído. Conforme já foi exposto, para SNR mais elevadas torna-se mais fácil arbitrar uma matriz de covariância inicial que leve o algoritmo a convergir.

4.6 Conclusões

Neste capítulo mostrou-se uma aplicação prática do filtro de Kalman em um sistema não-linear. O problema proposto foi a filtragem e estimação de estados desconhecidos do circuito de Chua. Os dois algoritmos discutidos no capítulo 3 foram comparados em duas situações distintas. Na primeira, foram utilizados dados gerados por simulação, enquanto que na segunda as medições foram obtidas a partir da implementação real do circuito de Chua, descrita na seção 4.3.2.

Usando-se os dados gerados por simulação, foi possível verificar que o UKF apresentou melhor desempenho do que o EKF, quando o objetivo era estimar estados. Já com os dados reais, esse resultado não se repetiu. Ao contrário do que se esperava, o desempenho do UKF foi pior.

Como foi suposto que os problemas do UKF, quando se trabalhava com dados reais, eram devidos aos valores incorretos dos parâmetros, partiu-se para a estimação conjunta. Estimando ao mesmo tempo estados e parâme-

tros, o UKF voltou a apresentar desempenho superior ao EKF. Alguns parâmetros convergiram para valores bastante diferentes daqueles que estavam sendo utilizados quando se estimava apenas os estados.

Como forma de validar os resultados obtidos, o sistema foi simulado com os valores de parâmetros estimados pelo UKF. O atrator dupla volta apresentou uma aparência bastante semelhante ao atrator obtido a partir dos dados reais.

Apesar dos autores afirmarem que o esforço computacional do UKF é baixo (Julier and Uhlmann, 1997), verificou-se que o procedimento de cálculo dos pontos sigma (ver seção 3.3.1), pode tornar o algoritmo lento na medida em que se aumenta a complexidade do problema (quantidade de parâmetros e estados).

De maneira geral o UKF aparentou ser um algoritmo bastante eficaz principalmente quando se realiza estimação conjunta de estados e parâmetros em sistemas reais.

A despeito desse fato, foi identificada uma desvantagem do algoritmo: grande sensibilidade à inicialização. Valores ligeiramente diferentes, definidos como condições iniciais, podem ter uma grande influência no resultado final proporcionado pelo algoritmo. Após algumas pesquisas, buscando encontrar algum material na literatura que investigasse esse problema, foi possível concluir que a sensibilidade à inicialização ainda se encontra em estudo por pesquisadores da área, e além disso, não se faz presente em muitos sistemas não-lineares já estudados. Talvez por isso o problema ainda não tenha merecido maior atenção.

Transmissão de Informação

“Nunca ande pelo caminho traçado, pois ele conduz somente até onde os outros foram.”

Alexander Graham Bell

5.1 Introdução

O homem está vivendo na era da informação. Essa afirmação, tão frequentemente ouvida nos dias de hoje, pode ser justificada principalmente pela forte presença de computadores, telefones celulares, e diversos outros dispositivos de comunicação na vida das pessoas. Com esse grande avanço tecnológico que vem sendo rapidamente introduzido, principalmente no campo da informática aplicada à engenharia, existe uma grande necessidade de se transmitir informação de forma segura. Por esse motivo, a busca por soluções que atendam a essas expectativas tem sido constante.

Na atualidade, verifica-se grande interesse pelo uso de sistemas caóticos para transmissão de informação (Bollt, 2003). A vantagem em se usar esses sistemas é que o sinal portador da informação apresenta características semelhantes às de um sinal aleatório, o que de certa forma ajuda a garantir privacidade para o usuário do sistema, trazendo proteção contra acesso não autorizado (Tôres, 2001).

Neste capítulo será mostrado que o Filtro de Kalman Unscented (UKF), descrito na seção 3.3, pode ser usado para recuperar informação inserida em sinais caóticos. Dois sistemas são usados para gerar esses sinais: o sistema de Lorenz e o oscilador caótico de Chua. No primeiro exemplo, geram-se os

dados do sistema de Lorenz por simulação. Já no segundo exemplo, utiliza-se dados reais obtidos a partir de medições de variáveis de estado do circuito de Chua implementado, descrito na seção 4.3.2. Na próxima seção descreve-se um procedimento que pode ser usado quando se deseja transmitir algum sinal de informação.

5.2 Transmissão e Recuperação de Informação

No capítulo anterior foi visto que o UKF é capaz de estimar estados e parâmetros de sistemas caóticos. A partir do instante em que a estimação converge, ou seja, o estado estimado se aproxima o suficiente do sinal medido, pode-se dizer que o UKF entrou em *sincronismo* com o sistema que gerou os dados. O conjunto *sistema real-filtro de Kalman* pode ser visto como sendo um conjunto *transmissor-receptor* de informação. Mas como inserir informação no sinal caótico transmitido? Essa pergunta será respondida na próxima seção.

5.2.1 Codificação de Informação

As equações que compõem o modelo de um sistema caótico são dotadas de parâmetros que podem ser alterados durante sua evolução. Entretanto, essa variação deve ocorrer dentro de certos limites, garantindo que o sistema continue exibindo comportamento caótico.

A variação de parâmetros pode codificar informação da seguinte maneira: durante a geração de dados, varia-se um determinado parâmetro θ do modelo entre dois valores fixos θ_1 e θ_2 . Caso o modelo possua outros parâmetros, estes são, em princípio, mantidos constantes. Sempre que o valor do parâmetro variável for θ_1 atribui-se bit 0 ou 1, e quando este mudar para θ_2 , atribui-se o bit oposto ao que foi atribuído anteriormente. Repetindo-se esse procedimento, é possível estabelecer uma correspondência entre a série temporal caótica perturbada e códigos binários que representem alguma informação.

Esse sinal caótico perturbado pode ser transmitido ao receptor usando-se algum meio de transmissão para então ser decodificado, de forma que a informação transmitida seja recuperada.

5.2.2 Decodificação de Informação

Foi mostrado na seção 3.3.3, que o UKF pode estimar os parâmetros de um modelo, simplesmente estendendo o vetor de estados de forma que este passe a conter também os parâmetros. Dessa maneira, os parâmetros passam a ser estimados ao mesmo tempo que os estados.

Após algumas modificações, o UKF é capaz de detectar a variação de parâmetros induzida, pela informação binária, no sinal medido, durante a estimação. Desta forma, o código inserido no sinal é recuperado, ou seja, a informação é *decodificada*.

A transmissão de informação por meio de sinais caóticos, seguindo um procedimento semelhante ao descrito nessa seção, foi realizada por Sobiski e Thorp (Sobiski and Thorp, 1998), porém usando o Filtro de Kalman Estendido (EKF). O sistema caótico utilizado naquele trabalho foi o sistema de *Duffing*.

Neste trabalho, procura-se atingir o mesmo objetivo de Sobiski e Thorp, porém usando outros dois sistemas não-lineares: o sistema de Lorenz e o circuito de Chua. No caso do circuito de Chua pretende-se ir um pouco além, trabalhando-se com dados da plataforma PCChua. A fim de se atingir tal objetivo, o algoritmo escolhido é o UKF, que se mostrou mais eficaz quando o sistema em questão é o circuito de Chua (capítulo 4). A fim de permitir que se tenha um primeiro contato com essa técnica de codificação/decodificação de informação, utiliza-se primeiramente o sistema de Lorenz, cujos dados foram gerados por simulação.

5.2.3 Modificações no Algoritmo

Nesta seção apresentam-se duas propostas de modificação no UKF que podem possibilitar a estimação de parâmetros variantes no tempo.

Redefinindo a Matriz de Covariância

O erro de estimação aumenta consideravelmente nos instantes em que ocorre variação de parâmetro. Idealmente, em tais instantes a matriz estendida de covariância dos estados já deveria ter convergido e, portanto, seus elementos devem encontrar-se em seus menores valores.

A matriz de covariância precisa ser redefinida, elevando novamente os valores de seus elementos e trazendo para o algoritmo a informação de que a

partir daquele instante têm-se menos certeza sobre os valores reais de estados e parâmetros. Esse reajuste da matriz pode ser feito automaticamente com base no erro de estimação:

$$\xi_k = x_k^{(1)} - \hat{x}_k^{(1)}, \quad (5.1)$$

sendo $x_k^{(1)}$ o valor da variável x_1 no instante k e $\hat{x}_k^{(1)}$ o valor estimado.

Quando o valor de ξ_k aumenta a ponto de ultrapassar um limite pré-definido, assume-se que a razão é o uso de valores de parâmetros que não mais correspondem à realidade. Por outro lado, o ajuste dos parâmetros não é eficaz pois a matriz de covariância tem valores pequenos. Uma forma de contornar este problema é aumentar os valores dos elementos da matriz. Ou seja, P_k^+ precisa ser redefinida ($P_k^+ = P_0^+$), reativando assim o mecanismo de estimação de parâmetros que, agora, poderá se aproximar mais rapidamente do novo parâmetro. Desta forma, é possível detectar variações de parâmetros, conforme será mostrado nos exemplos apresentados nesse capítulo.

Limitando a Matriz de Covariância

Outra maneira de permitir que o UKF esteja sempre sensível a variações de parâmetros é impedir que a matriz de covariância P_k^+ atinja valores muito baixos. Fazendo-se isso, o ganho de Kalman K_k também fica limitado. Para limitar a matriz P_k^+ pode-se adicionar a ela uma certa matriz D , a cada iteração do algoritmo. Essa matriz D , que pode ser escolhida por meio de experimentos, garante um valor mínimo para os elementos de P_k . A idéia de se somar uma matriz mínima à matriz de covariância com o intuito de limitá-la, foi proposta por Young (1970) para o estimador de variáveis instrumentais.

Antes de se aplicar este método, é preciso verificar se o UKF é sensível à inicialização da matriz P_0^+ , para o sistema no qual estiver sendo aplicado. Se a sensibilidade for alta, pode ser necessário determinar uma matriz de covariância inicial que funcione, retornando a ela sempre que o erro aumentar, conforme descrito na 1ª opção de modificação do UKF.

5.3 Sistema de Lorenz

Na seção 3.3.5 mostrou-se que o UKF apresentou bom desempenho na estimação de estados e parâmetros do sistema de Lorenz (Lorenz, 1963).

Nesta seção pretende-se usar o procedimento descrito acima para demonstrar que é possível transmitir informação por meio de variações paramétricas.

A fim de se atingir este objetivo, o UKF será utilizado para detectar variações em um ou mais parâmetros do seguinte modelo:

$$\begin{cases} \dot{x}_1 &= -\lambda_1 x_1 + \lambda_1 x_2, \\ \dot{x}_2 &= \lambda_2 x_1 - x_2 - x_1 x_3, \\ \dot{x}_3 &= -\lambda_3 x_3 + x_1 x_2. \end{cases} \quad (5.2)$$

5.3.1 Metodologia

O sistema pode ser representado em espaço de estados de acordo com:

$$\begin{bmatrix} \mathbf{x}_k \\ \boldsymbol{\lambda}_k \end{bmatrix} = \begin{bmatrix} f(\mathbf{x}_{k-1}, \boldsymbol{\lambda}_{k-1}) \\ \boldsymbol{\lambda}_{k-1} \end{bmatrix}, \quad (5.3)$$

$$y_k = x_k^{(1)} + e_k, \quad (5.4)$$

sendo que f representa as equações do sistema de Lorenz, do instante $k - 1$ até k , \mathbf{x}_k é o vetor de estados no instante k , $\boldsymbol{\lambda}_k$ o vetor de parâmetros, e e_k ruído de medição. Na equação de saída (5.4), $x_k^{(1)}$ representa o valor da variável x_1 na iteração k e e_k , o ruído de medição.

Varição de Parâmetros

Inicialmente (caso 1) foram mantidos constantes: $\lambda_1 = 12$ e $\lambda_2 = 40$. Variou-se o parâmetro λ_3 , alternando seu valor entre $\lambda_3 = 3$ e $\lambda_3 = 4$, sendo que a variação foi realizada conforme mostrado na Figura 5.1.

Em um segundo momento (caso 2), variou-se λ_3 da mesma forma que no caso anterior, embora com um tempo maior (100 segundos) e adicionando-se ruído na série temporal simulada. A seqüência de variação foi: 3-3-4-3-4-3-3-4-4-3.

No terceiro caso, manteve-se constante apenas $\lambda_1 = 12$, com variação de λ_2 entre $\lambda_2 = 40$ e $\lambda_2 = 45$. Também foi adicionado ruído aos dados. Variou-se λ_3 dentro da mesma faixa anterior, conforme Figura 5.2.

É importante salientar que todos esses valores, escolhidos para os três parâmetros do modelo, garantem que o sistema opere em regime caótico (Fiedler-Ferrara and Prado, 1994).

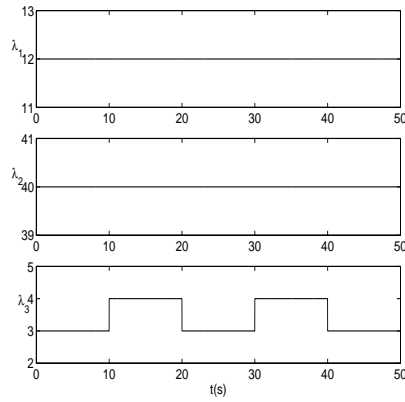


Figura 5.1: Variação de parâmetros do sistema de Lorenz – primeiro caso.

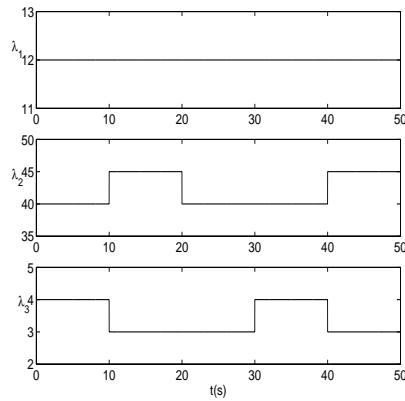


Figura 5.2: Variação de parâmetros do sistema de Lorenz – terceiro caso.

5.3.2 Resultados e Comentários

Nesta seção apresenta-se os resultados obtidos utilizando-se o UKF para os três casos estudados.

Primeiro caso – Variação de λ_3

Neste caso, não foi verificada muita sensibilidade à inicialização, ou seja, o algoritmo convergiu para uma ampla faixa de matrizes P_0^+ distintas (de $10^{-4}I$ a 10^2I , sendo I uma matriz identidade), mesmo aumentando-se o nível de ruído nos dados. Desta maneira, não houve necessidade de se redefinir P_k^+

sempre que o erro de estimação aumentava. Por este motivo, a modificação escolhida para tornar o algoritmo capaz de detectar variação de parâmetros foi a 2ª proposta na seção 5.2.3.

Os resultados obtidos são apresentados abaixo. Salienta-se que, neste primeiro caso, não foi adicionado ruído à série temporal simulada. Na Figura 5.3a mostra-se o intervalo completo no qual realizou-se a estimação. Para melhor visualização, destaca-se na Figura 5.3b uma pequena amostra desse resultado, indicando que o sistema de Lorenz e o UKF estão praticamente sincronizados durante a transmissão de informação.

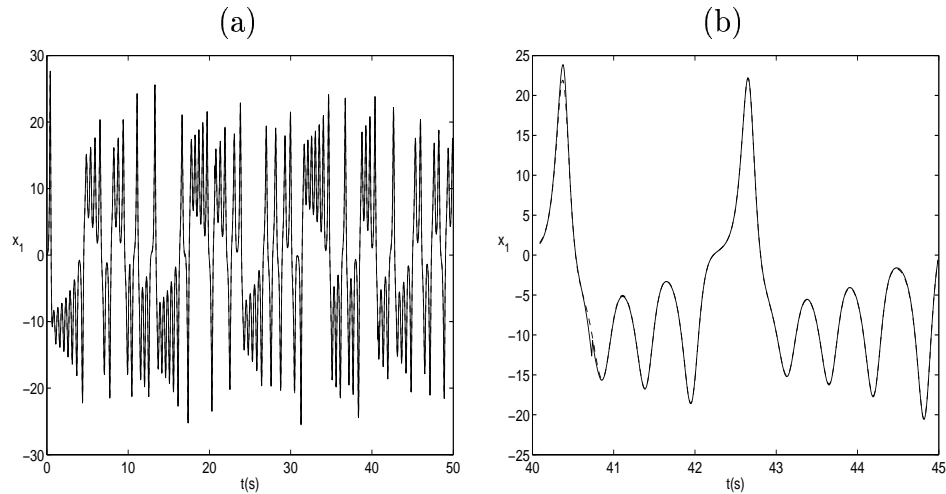


Figura 5.3: Estimação do estado simulado – Sistema de Lorenz – primeiro caso. Sinal simulado (---) e estimado (—). As mudanças de λ_3 se deram em $t = 10, 20, 30$ e 40 s.

A diferença entre o sinal simulado e o estado estimado, ou seja, o erro de estimação ξ aumenta nos instantes em que ocorre variação de λ_3 . A elevação de ξ permanece por alguns instantes até que se obtenha convergência novamente. Esse fato pode ser observado na Figura 5.4, na qual é possível notar a presença de picos, significando elevação no valor de ξ , exatamente nos instantes em que ocorre variação do parâmetro ($t = 10, 20, 30, 40$).

Na Figura 5.5 encontra-se o resultado desejado. Percebe-se claramente que o UKF é capaz de detectar variações de λ_3 , recuperando o código inserido antes da transmissão do sinal.

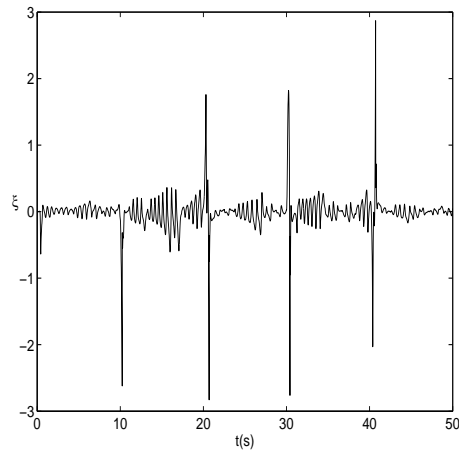


Figura 5.4: Erro de estimação do estado x_1 – Sistema de Lorenz – primeiro caso.

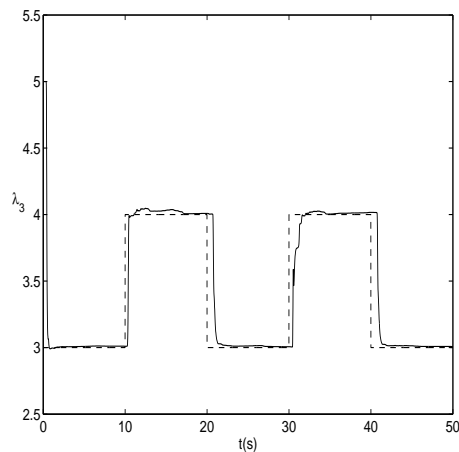


Figura 5.5: Estimação do parâmetro variável λ_3 do sistema de Lorenz – primeiro caso. Variação provocada em λ_3 (---) e $\hat{\lambda}_3$ (—).

Segundo caso – Variação do nível de ruído

A fim de investigar a robustez da técnica de transmissão de informação ao ruído, foram feitas algumas simulações com diferentes níveis de ruído (SNR). Constatou-se que para SNR = 25, 20, 15, 10, e 5 dB nenhum bit foi recuperado incorretamente. Para SNR=0 db dos 10 bits transmitidos 4 foram decodificados de forma errada.

Na Figura 5.6 apresenta-se o resultado obtido com a pior SNR que não

apresentou bits errados na decodificação (SNR=5 dB).

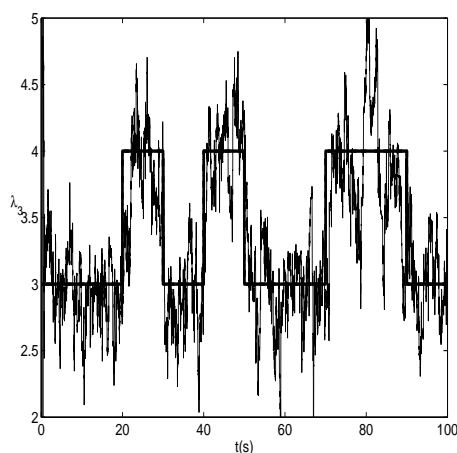


Figura 5.6: Estimação do parâmetro variável λ_3 do sistema de Lorenz – segundo caso.

Aplicando-se um algoritmo para detecção de variação na série temporal do parâmetro estimado verificou-se que o código binário recuperado foi: 0–0–1–0–1–0–0–1–1–0, o qual está totalmente de acordo com o código transmitido citado na seção 5.3.1. Este algoritmo de detecção calcula a média do sinal estimado a cada intervalo de tempo (neste caso 10s). De posse da média, verifica-se se esta está mais próxima do valor mínimo ou máximo, tornando possível decodificar a informação.

Terceiro caso – Variação de λ_2 e λ_3

Neste caso, os dados foram gerados enquanto eram variados dois parâmetros (λ_2 e λ_3), em instantes de tempo diferentes, conforme Figura 5.2. Além disso, foi adicionado ruído ao sinal simulado.

Conforme dito anteriormente, variou-se λ_2 entre dois valores: $\lambda_2 = 40$ e $\lambda_2 = 45$; e λ_3 entre $\lambda_3 = 3$ e $\lambda_3 = 4$. O ruído somado ao sinal simulado foi gerado com distribuição gaussiana e desvio padrão igual a 10% do desvio padrão da série temporal x_1 sem ruído, ou seja, SNR=20 dB.

Foi verificada grande sensibilidade à inicialização, o que impediu que se obtivesse sucesso com a 2ª modificação do algoritmo. Desta forma, a modificação adotada foi aquela que reajusta a matriz de covariância quando o erro de estimação aumenta.

Nas Figuras 5.7 e 5.8, mostra-se respectivamente, o sinal estimado e o erro de estimação. Como pode ser notado, o erro é bem maior que no caso 1. Isso se deve à presença de ruído nos dados. Mesmo assim, o algoritmo foi capaz de estimar os estados, detectando variações nos dois parâmetros em instantes bem próximos dos instantes em que as variações efetivamente ocorreram na etapa de transmissão.

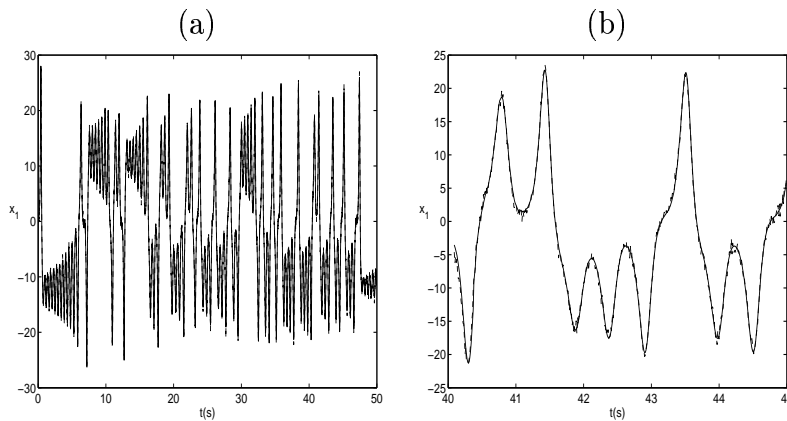


Figura 5.7: Estimação do estado simulado – Sistema de Lorenz – terceiro caso. Sinal simulado (---) e estimado (—).

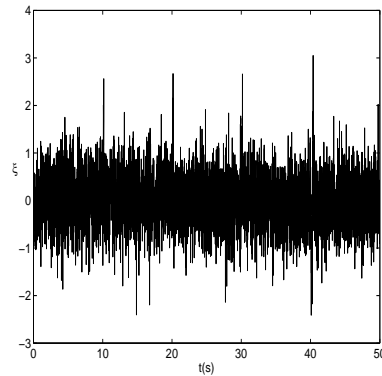


Figura 5.8: Erro de estimação do estado x_1 – Sistema de Lorenz – terceiro caso.

A evolução dos dois parâmetros durante a estimação é apresentada na Figura 5.9. Esse resultado é muito importante, pois a capacidade de transmissão de informação está sendo duplicada. O volume de informação codificada

em um único sinal pode aumentar ainda mais desde que seja possível estimar um número maior de parâmetros. Por outro lado, é importante salientar que, em aplicações reais, o número de parâmetros passíveis de serem estimados é limitado. Foi mostrado no capítulo 4 que existe uma certa dificuldade em se estimar parâmetros com o UKF à medida que se aumenta a quantidade destes. Naquele capítulo essa dificuldade pôde ser notada principalmente na escolha de valores para o vetor de estados e matriz de covariância iniciais, para inicializar o algoritmo.

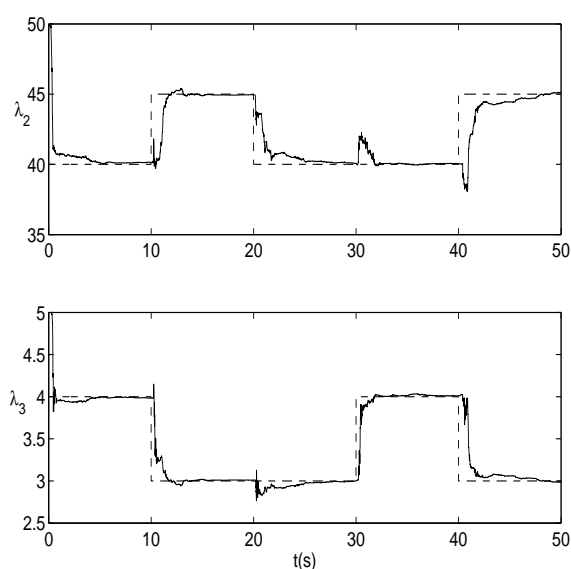


Figura 5.9: Estimação de parâmetros variáveis do sistema de Lorenz – terceiro caso. Valores dos parâmetros usados na simulação (---) e valores estimados (—).

5.4 Circuito de Chua

Na seção anterior foi mostrado, utilizando-se dados gerados a partir do sistema de Lorenz, que o UKF é capaz de detectar variação de parâmetros durante a geração da série temporal caótica, tornando possível a transmissão de informação por meio de códigos que podem ser recuperados estimando parâmetros. Nesta seção pretende-se ir um pouco mais além: a série caótica será gerada a partir de um circuito de Chua real, descrito na seção 4.3.2.

5.4.1 Metodologia

O circuito de Chua pode ser representado pelas seguintes equações de estado:

$$\begin{cases} \dot{x}_1 = -\frac{1}{RC_1}x_1 + \frac{1}{RC_1}x_2 - \frac{1}{C_1}i_d(x_1), \\ \dot{x}_2 = \frac{1}{RC_2}x_1 - \frac{1}{RC_2}x_2 + \frac{1}{C_2}x_3, \\ \dot{x}_3 = -\frac{1}{L}x_2, \end{cases} \quad (5.5)$$

sendo $i_d(x_1) = m_0x_1 + \frac{1}{2}(m_1 - m_0)(|x_1 + B_p| - |x_1 - B_p|)$.

Neste sistema, a exemplo do sistema de Lorenz, mediu-se a série temporal caótica x_1 , que será a portadora de informação.

Variação de Parâmetro

A placa do circuito de Chua real foi construída de forma a permitir variações no resistor R (Tôres, 2001). Por este motivo, esse parâmetro do modelo foi escolhido como parâmetro variável. Os demais parâmetros, cujos valores encontram-se listados na Tabela 4.6, são mantidos constantes durante todo o tempo de geração de dados.

Foram feitos dois experimentos com o circuito de Chua, nos quais variou-se a resistência manualmente.

1. Dados foram gerados durante 120 segundos com $R = 1.549 \Omega$, e o sistema operando em regime caótico dupla volta. Nesse instante, ainda com o circuito ligado, variou-se a resistência para um determinado valor, a princípio desconhecido, porém mantendo regime dupla volta. Após passados mais 120 segundos, desligou-se o circuito e mediu-se o valor atual da resistência. O valor obtido foi $R = 1.615 \Omega$.
2. Os dados foram gerados durante 120 segundos. Durante todo esse período, houve variações de R a cada intervalo de 30 segundos, totalizando portanto 3 alterações, porém sempre comutando, aproximadamente, entre dois valores. Os valores medidos foram: $R = 1.540 \Omega$ e $R = 1.647 \Omega$.

5.4.2 Resultados e Comentários

Nesta seção são apresentados os resultados obtidos nos dois experimentos descritos anteriormente. Como já era esperado, o problema de sensibilidade à inicialização voltou a aparecer neste caso, de forma que foi preciso adotar a 1ª modificação no UKF, para que a matriz de covariância fosse redefinida sempre que o erro de estimação aumentasse.

Na Figura 5.10, é apresentado o gráfico do sinal medido x_1 e estimado \hat{x}_1 pelo UKF. A partir da análise deste gráfico, verifica-se o sincronismo entre transmissor e receptor. A alteração no valor de R durante a geração de dados ocorreu aproximadamente em $t = 120$ s. Percebe-se que nas proximidades desse instante de tempo ocorreu um princípio de perda de sincronismo, o qual foi corrigido alguns segundos depois (Figura 5.10b).

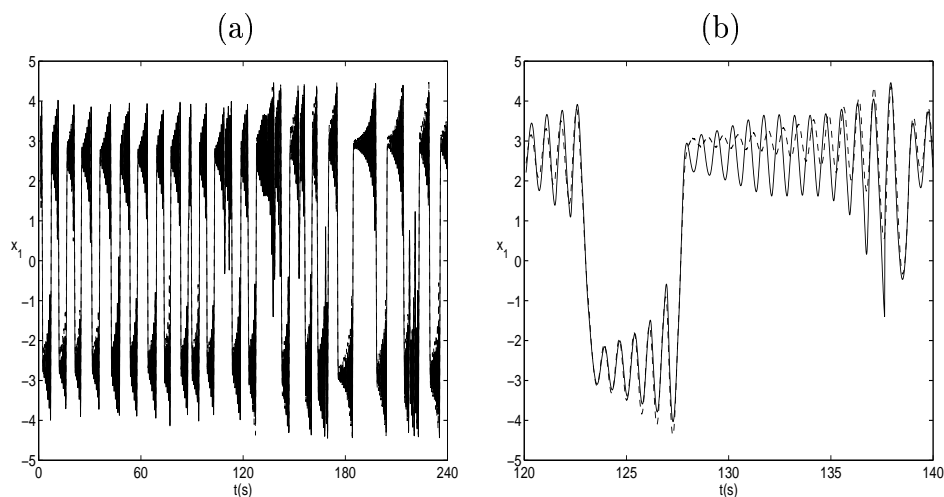


Figura 5.10: Estimação do estado medido – Circuito de Chua – primeiro experimento. Sinal medido (---) e estimado (—). A variação de R ocorreu em aproximadamente $t = 120$ s.

Na Figura 5.11 mostra-se que o resultado desejado foi obtido. Após a convergência do algoritmo e até o instante $t = 120$ s, o valor estimado para o resistor foi aproximadamente $\hat{R} = 1.540 \Omega$ e o erro percentual com relação ao valor medido de R foi de apenas 0,58 %. Após o instante $t = 120$ s, o algoritmo detecta uma elevação no erro de estimação, o que faz com que a matriz de covariância seja redefinida. O valor estimado passou a ser de aproximadamente $\hat{R} = 1.594 \Omega$ depois da convergência, enquanto que o erro percentual com relação ao valor medido foi de 1,30 % (ver Tabela 5.1).

Tabela 5.1: Valores da resistência R do circuito de Chua – medição e estimação.

Parâmetro	Valor medido	Valor estimado	Erro
R_{\min}	1.549 Ω	1.540 \pm 1,5 Ω	0,58 %
R_{\max}	1.615 Ω	1.594 \pm 3 Ω	1,30 %

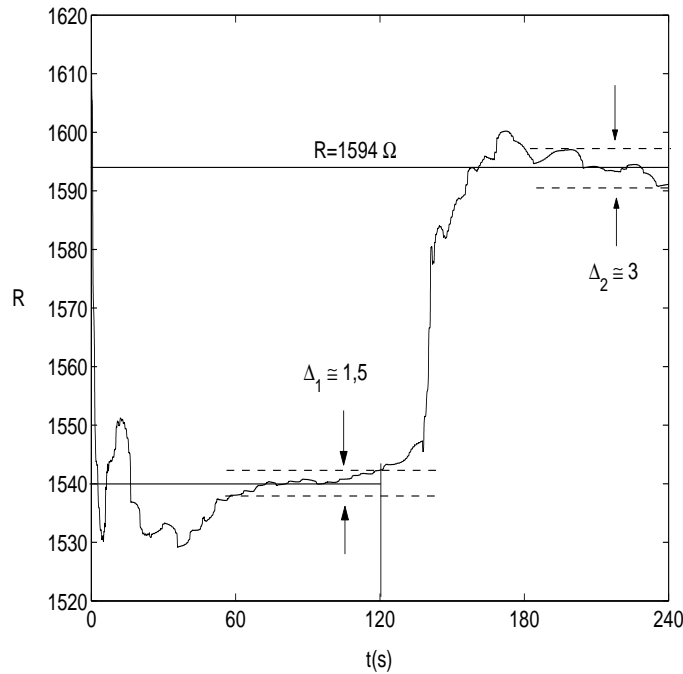


Figura 5.11: Estimação do parâmetro variável do circuito de Chua – primeiro experimento. A variação de R ocorreu em aproximadamente $t = 120$ s.

Na Figura 5.12 é mostrado que após a correção do parâmetro R e convergência do algoritmo, o sinal medido x_1 aproxima-se do sinal medido x_1 , diferentemente do resultado apresentado na Figura 5.10b, nas proximidades do instante em que houve variação paramétrica. Na Figura 5.13 mostra-se o erro de estimação.

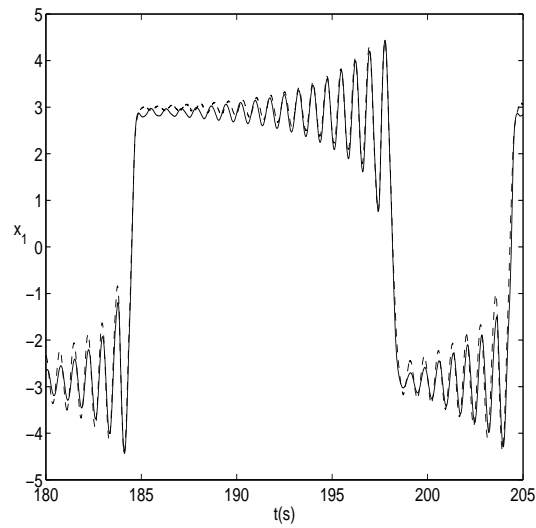


Figura 5.12: Estimação do estado medido – série temporal após convergência.

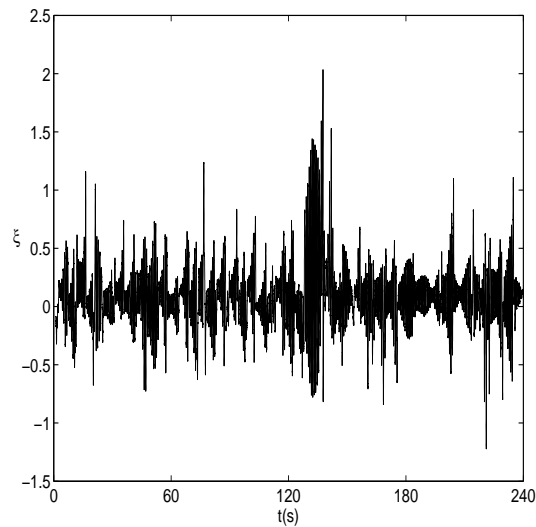


Figura 5.13: Erro de estimação do estado x_1 – Circuito de Chua – primeiro experimento.

No segundo experimento, conforme já descrito, variou-se o valor de R mais de uma vez durante a geração de dados, com o intuito de produzir uma seqüência binária.

Nas Figuras 5.14 e 5.15 são apresentados os sinais medido x_1 e estimado

\hat{x}_1 e erro de estimação ξ , respectivamente.

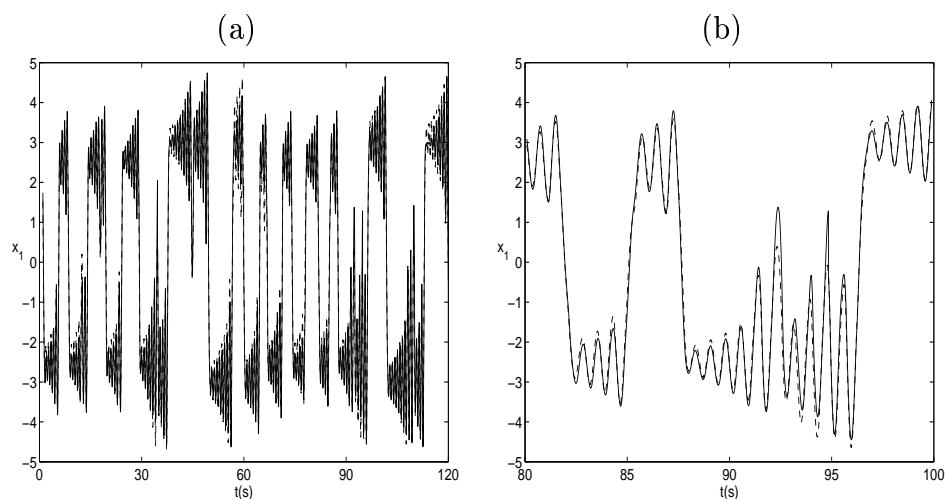


Figura 5.14: Estimação do estado medido – Circuito de Chua – segundo experimento. Sinal medido (---) e estimado (—). A variação de R ocorreu nos instantes 30, 60 e 90 segundos.

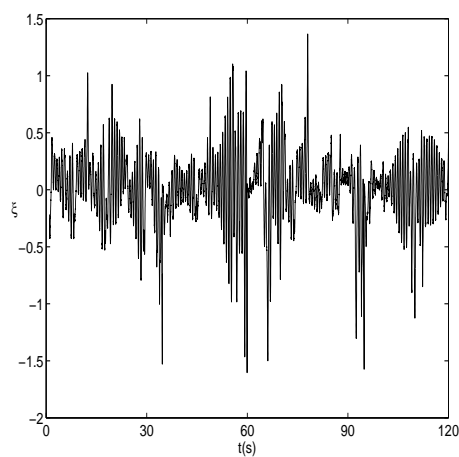


Figura 5.15: Erro de estimação do estado x_1 – Circuito de Chua – segundo experimento.

A evolução do parâmetro R durante a estimação é apresentada na Figura 5.16.

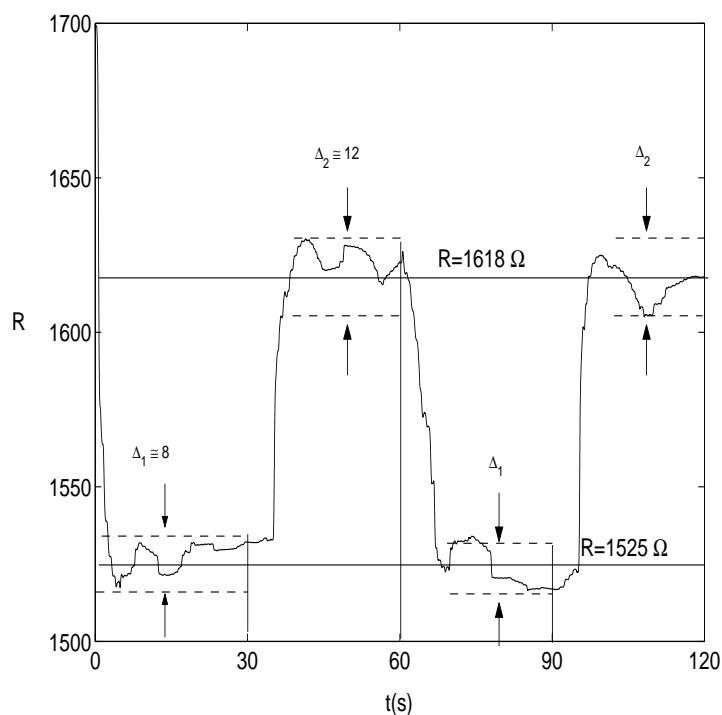


Figura 5.16: Estimação do parâmetro variável do circuito de Chua – segundo experimento. A variação de R ocorreu nos instantes 30, 60 e 90 segundos.

Assim como no experimento anterior, o UKF detectou as variações na resistência do circuito. Conforme já explicado anteriormente, cada variação pode ser vista como um símbolo binário. Ou seja, neste exemplo pode-se supor que quando o valor da resistência for $R \cong 1.540 \Omega$ tem-se bit 0 e quando $R \cong 1.647 \Omega$ tem-se bit 1. Desta forma, o código transmitido foi 0 – 1 – 0 – 1.

Tabela 5.2: Valores de R - medição e estimação.

Parâmetro	Valor medido	Valor estimado	Erro
R_{\min}	1.540Ω	$1.525 \pm 8 \Omega$	0,97 %
R_{\max}	1.647Ω	$1.618 \pm 12 \Omega$	1,76 %

Observando os gráficos das Figuras 5.11 e 5.16, nota-se a presença de pequenas oscilações na curva de evolução da resistência estimada \hat{R} . Essas

oscilações tendem a diminuir à medida que se permite um tempo maior de medição/estimação, o que por outro lado torna o processo de decodificação cada vez mais lento. Entretanto, verifica-se que nos dois experimentos da presente seção, essas oscilações se concentraram em uma faixa de valores cujos limites estão muito próximos do valor real de R .

Mesmo com as pequenas oscilações verificadas, considerou-se que a convergência já havia sido atingida após passados 120 s no primeiro experimento. No segundo experimento, verificou-se que as oscilações ocorreram dentro de uma faixa um pouco mais ampla do que no primeiro experimento. Isso ocorreu pelo fato das variações paramétricas, no segundo experimento, terem ocorrido a cada intervalo de 30 s. Considerou-se que após passados 30 s a convergência já havia sido atingida, o que trouxe grande economia no tempo de geração de dados e estimação, melhorando a taxa de transmissão de informação. Essa consideração foi possível uma vez que o erro na estimação de R foi baixo, conforme Tabela 5.2.

A maior dificuldade em se obter convergência de parâmetros para o circuito de Chua em relação ao exemplo do sistema de Lorenz (seção 5.3), pode estar relacionada a certos fatores que conduzem a um aumento de complexidade do problema quando se trabalha com dados reais. Tais fatores podem ser ruído de processo e de medição, por exemplo.

Uma das diferenças na etapa de geração de dados, que pode ter influenciado diretamente a convergência dos resultados, está relacionada com o procedimento de variação de parâmetros. No caso do sistema de Lorenz foi possível mudar o valor do parâmetro de forma instantânea, pois os dados foram gerados por simulação. No circuito de Chua a variação foi feita gradativamente, por meio de um procedimento manual, passando por uma faixa de valores intermediários até atingir o valor desejado. Entretanto, procurou-se realizar essa variação rapidamente, aproximando-a de uma variação instantânea.

5.5 Conclusões

Neste capítulo mostrou-se um procedimento que pode ser usado para transmissão de informação em sistemas de comunicação. Códigos binários são inseridos em séries temporais caóticas a partir de sucessivas variações de um ou mais parâmetros do sistema. Quando a série temporal chega ao receptor, o código transmitido é recuperado por estimação conjunta de estados

e parâmetros via UKF.

Dois exemplos de aplicação da técnica foram apresentados. Na seção 5.3 mostrou-se um exemplo no qual a série temporal caótica portadora da informação foi gerada por simulação das equações do sistema de Lorenz. No segundo exemplo, apresentado da seção 5.4, utilizou-se um circuito de Chua real, descrito na seção 4.3.2, para gerar os dados.

Em ambos os casos (sistema de Lorenz e circuito de Chua), o UKF foi capaz de detectar variações nos parâmetros dos modelos, o que torna possível a transmissão digital de informação utilizando-se estes sistemas em conjunto com o UKF. Por outro lado, salienta-se que as dificuldades encontradas para inicializar a matriz de covariância podem se apresentar como um forte empecilho para a implementação do algoritmo.

Resultados semelhantes aos apresentados neste capítulo foram obtidos por Sobiski e Thorp (Sobiski and Thorp, 1998). As diferenças entre aquele trabalho e o presente capítulo são o sistema não-linear utilizado e o algoritmo escolhido para realizar estimação. Sobiski e Thorp trabalharam com o sistema de *Duffing* e usaram o EKF para estimar os parâmetros do modelo. Neste capítulo foram usados, além do sistema de Lorenz, dados reais da plataforma PCChua, sendo que o algoritmo escolhido foi o UKF, método considerado mais eficaz do que o EKF quando se trata do Circuito de Chua.

É interessante destacar que o mesmo princípio de detecção de variação de parâmetros apresentado neste capítulo pode ser usado com o objetivo de detectar falhas (Camargo, 1997) em sistemas desenvolvidos para as mais diversas aplicações.

Conclusão e Pesquisas Futuras

“Mesmo que eu tivesse o dom da profecia, e conhecesse todos os mistérios e toda a ciência; mesmo que tivesse toda a fé, a ponto de transportar montanhas, se não tiver amor, não sou nada.”

I Coríntios, 13,2

Estimação recursiva é um assunto de grande interesse em ciência e tecnologia, sendo objeto de estudo em diversas áreas. Algoritmos para estimação recursiva podem ser aplicados em diferentes situações. Na Identificação de Sistemas é possível utilizar estimadores recursivos para estimar os parâmetros do modelo. Existem vários algoritmos que foram propostos para desempenhar essa função, sendo que o mais conhecido é o *estimador recursivo de mínimos quadrados*, descrito na seção 2.3.2. Estimadores recursivos também podem ser usados na filtragem de sinais e estimação dos estados desconhecidos de sistemas dinâmicos lineares ou não-lineares. Dentre os estimadores desenvolvidos para estimar estados, destaca-se o *Filtro de Kalman*, que foi o assunto principal deste trabalho.

6.1 O Filtro de Kalman

Por volta dos anos 50, aumentou-se o interesse do homem em conquistar o espaço. Pesquisadores daquela época tinham problemas na estimação da trajetória de veículos espaciais. Como solução para problemas desse tipo surgiu o Filtro de Kalman (Kalman, 1960). Desde a época em que foi proposto, tal algoritmo tem sido aplicado em diversas áreas da ciência. Na seção 1.1

foram listadas algumas dessas aplicações desenvolvidas no passado e também recentemente.

O KF é um algoritmo recursivo *ótimo*¹ capaz de estimar os estados de sistemas dinâmicos perturbados por ruído (de processo e de medição) branco e gaussiano. Como pode ser notado, para que o algoritmo de Kalman possa ser empregado na solução de algum problema, geralmente adotam-se algumas hipóteses simplificadoras com relação aos sinais aleatórios que afetam o processo.

Outra questão importante é o fato de que o KF da forma como foi proposto (Kalman, 1960; Kalman and Bucy, 1961), pressupõe que o sistema em questão seja modelado por equações lineares (seção 2.4.1). Essa característica pode trazer fortes limitações no uso do algoritmo, visto que grande parte dos sistemas do mundo real são de natureza não-linear. Podem ser encontradas na literatura diversas abordagens do KF adequadas para aplicação em sistemas não-lineares. Neste trabalho foram apresentados dois algoritmos baseados no mesmo procedimento proposto por Kalman, mas que podem ser aplicados a sistemas não-lineares: o Filtro de Kalman Estendido e o Filtro de Kalman *Unscented*.

6.2 Estimação em Sistemas Não-Lineares

O Filtro de Kalman Estendido (EKF) lineariza as funções do sistema não-linear, dando origem a um novo sistema de equações. É necessário calcular matrizes jacobianas a cada iteração do algoritmo. A partir desse novo modelo linear pode-se aplicar normalmente as equações do Filtro de Kalman original (seção 3.2).

O procedimento de linearização das funções pode provocar erro de estimação ou divergência, dependendo da influência que elementos não-lineares exercem sobre a dinâmica do sistema (Cipra, 1993; Julier and Uhlmann, 1996). Esse fato pôde ser verificado na prática por meio de um exemplo de aplicação apresentado nesse trabalho. O sistema escolhido foi o circuito de Chua, um sistema caótico de construção relativamente simples, mas cujo comportamento é bastante complexo.

O Filtro de Kalman *Unscented* (UKF) baseia-se no fato de que a apro-

¹O termo *ótimo*, neste caso, significa que o KF calcula as estimativas para as variáveis minimizando o valor esperado da soma quadrática dos resíduos de estimação.

ximação da distribuição gaussiana dos estados é capaz de proporcionar melhores resultados do que a linearização das funções não-lineares. Os autores desse algoritmo (Julier and Uhlmann, 1997), sugeriram uma técnica eficiente capaz de propagar a média e a matriz de covariância dos estados usando-se as próprias funções não-lineares do sistema. Desta forma, o que ocorre é uma espécie de “linearização estatística” e não uma linearização das funções do sistema. Essa técnica é conhecida como Transformação *Unscented* (seção 3.3.1).

Os resultados obtidos com os algoritmos EKF e UKF foram comparados no capítulo 4 deste trabalho. Foram utilizados dados gerados por simulação e dados gerados e medidos usando uma plataforma de laboratório. Trabalhando-se com dados obtidos por simulação, o UKF foi capaz de estimar os estados do circuito de Chua proporcionando resultados melhores do que aqueles fornecidos pelo EKF. Com dados reais não se obteve êxito de imediato. Pelo fato de não ser possível garantir que os valores de todos os parâmetros do sistema real sejam precisamente conhecidos, foi necessário estimá-los juntamente com os estados. Realizando-se esse procedimento melhores resultados puderam ser obtidos.

6.3 Detecção de Variações nos Parâmetros

A capacidade do UKF em estimar os parâmetros do sistema, detectando variações em seus valores, dá origem a diversas possibilidades de aplicações do algoritmo. A transmissão de informação é uma dessas possíveis aplicações. No capítulo anterior mostrou-se que é possível inserir códigos binários em séries temporais caóticas fazendo-se sucessivas variações de um ou mais parâmetros do sistema. Quando a série temporal chega ao receptor o código transmitido é recuperado por estimação conjunta de estados e parâmetros via UKF. A preferência por se utilizar sistemas caóticos é que o sinal portador da informação apresenta características semelhantes às de um sinal aleatório, dificultando o acesso não autorizado à informação, pelo menos em princípio.

Uma outra aplicação interessante é a detecção de falhas em processos. Estimando-se constantemente os parâmetros de um determinado sistema, é possível detectar possíveis variações em seus valores, indicando onde e quando pode ter ocorrido uma falha.

6.4 Conclusões

A revisão bibliográfica e a experiência adquirida com os exemplos práticos estudados neste trabalho mostraram que o uso do Filtro de Kalman para estimação de estados e parâmetros de sistemas não-lineares pode proporcionar ótimos resultados. De maneira geral, foi possível concluir que o UKF é mais eficaz do que o EKF, principalmente quando se trata de sistemas mais complexos e cuja dinâmica sofre forte influência de elementos não-lineares. Por outro lado, é preciso ter cautela ao se fazer uso do UKF, pois foram encontrados problemas na implementação. Para os sistemas não-lineares estudados, principalmente o circuito de Chua, o UKF apresentou alta sensibilidade à inicialização. Verificou-se que pequenas variações no vetor de estados inicial e matriz de covariância dos estados inicial, podem fazer com que o resultado não convirja. Este fato foi verificado durante a estimação simultânea de estados e parâmetros. Estimando somente estados, o que simplifica as operações, este problema não se mostrou relevante.

Espera-se que este trabalho sirva de motivação para novos estudos na área de estimação recursiva, e principalmente, seja um ponto de partida para pesquisadores interessados em investigar melhor o funcionamento do UKF, uma vez que a técnica foi proposta muito recentemente e por esse motivo ainda é pouco explorada.

6.5 Pesquisas Futuras

Este trabalho foi o primeiro trabalho realizado no grupo MACSIN sobre o tema, principalmente em se tratando do UKF. Como propostas de continuidade são apresentadas as seguintes sugestões:

- Aplicar o UKF em outros sistemas não-lineares, principalmente naqueles que não apresentam dinâmica caótica, para filtragem de sinais e estimação de estados desconhecidos. Verificar se surgem dificuldades na implementação, como por exemplo o problema da inicialização, encontrado nos sistemas estudados neste trabalho.
- Investigar o problema da sensibilidade do UKF às condições iniciais, para o vetor de estados e matriz de covariância, na tentativa de propor um método para cálculo dessas condições ou modificações no algoritmo que o tornem menos sensível.

-
- Propor e implementar um procedimento completo para transmissão de informação usando o circuito de Chua real, baseado na variação automática de um ou mais parâmetros.
 - Usar o procedimento de detecção de variações de parâmetros em sistemas não-lineares com o objetivo de detectar falhas no processo.
 - Buscar melhores formas de tornar o algoritmo capaz de detectar variações paramétricas, que não estejam atreladas à inicialização da matriz de covariância.

Bibliografia

- Aguirre, L. A. (2000). *Introdução à Identificação de Sistemas-Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*, chapter Estudo de Casos, pages 439–453. Editora UFMG, Belo Horizonte.
- Aguirre, L. A., Rodrigues, G. G., and Mendes, E. M. A. M. (1997). Non-linear identification and cluster analysis of chaotic attractors from a real implementation of Chua's Circuit. *International Journal of Bifurcation and Chaos*, 7(6):1411–1423.
- Arulampalam, S., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for *on-line* non-linear/non-gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188.
- Bar-Shalom, Y. and Fortmann, T. E. (1988). *Tracking and Data Association*. Academic Press, Orlando, FL.
- Bekman, O. R. and Costa, P. L. O. N. (1980). *Análise Estatística da Decisão*. Editora Edgard Blucher Ltda, São Paulo.
- Bode, H. W. (1945). *Network Analysis and Feedback Amplifier Design*. Van Nostrand Reinhold, New York.
- Boltt, E. M. (2003). Review of chaos communication by feedback control of symbolic dynamics. *International Journal of Bifurcation and Chaos*, 13(2):269–285.
- Brown, R. G. and Hwang, P. Y. C. (1997). *Introduction to Random Signals and Applied Kalman Filtering*. John Wiley e Sons, 3 edition.

- Brunke, S. S. (2001). *Nonlinear Filtering and System Identification Algorithms for Autonomous Systems*. PhD thesis, University of Washington, Washington.
- Camargo, H. A. (1997). *Detecção e Isolamento de Falhas Utilizando Estimaco de Parâmetros de Modelos Contínuos e Limites de Confiança Adaptativos - Aplicaco a uma Planta Piloto*. Dissertaco de Mestrado do PPGEE, Universidade Federal de Minas Gerais, Belo Horizonte.
- Campbell, J. K., Synnott, S. P., and Bierman, G. J. (1983). Voyager orbit determination at Jupiter. *IEEE Transactions on Automatic Control*, AC-28:256–268.
- Chen, C.-T. (1984). *Linear System Theory and Desing*. CBS College Publishing, New York.
- Chua, L. O. (1994). Chua circuit: An overview 10 years later. *Journal of Circ., Sys and Computers*, 4(2):117–159.
- Cipra, B. (1993). Engineers look to Kalman filtering for guidance. *SIAM News*, 26(5).
- Correa, M. V. (2001). *Identificaco Caixa-Cinza de Sistemas No-Lineares usando Representaces NARMAX Racionais e Polinomiais*. Tese de Doutorado do PPGEE, Universidade Federal de Minas Gerais, Belo Horizonte.
- Cruz, C., Nijmeijer, H., and Aguilar, A. (2001). Synchronization of a noisy Chua circuit via two switching Kalman filters. *Instrumentation and Development*, 5(3):162–169.
- Daum, F. E. (1988). New exact nonlinear filters. In J. C. Spall, editor, *Bayesian Analysis of Time Series and Dynamic Models*, chapter 8, pages 199–226. Marcel Drekker, Inc.
- Eltabach, M., Charara, A., and Zein, I. (2002). Kalman filtering and torque spectral analysis for broken bar detection in induction motors. In *15th Triennial World Congress*, Barcelona, Spain.
- Evans, W. R. (1948). Graphical analysis of control systems. *AIEE Transactions*, Part II(67):547–551.

- Evans, W. R. (1950). Control system synthesis by root locus method. *AIEE Transactions*, Part II(69):66–69.
- Fiedler-Ferrara, N. and Prado, C. P. C. (1994). *Caos - Uma Introdução*. Editora Edgard Blucher Ltda.
- Fung, P. T. K. and Grimble, M. J. (1983). Dynamic ship positioning using a self-tuning Kalman filter. *IEEE Transactions on Automatic Control*, AC-28:339–349.
- Gill, P. E., Murray, W., and Wright, M. H. (1991). *Numerical Linear Algebra and Optimization*, volume 1. Addison–Wesley, Redwood City, CA.
- Gleick, J. (1987). *Chaos: Making a New Science*. Penguin Books, New York.
- Gordon, N. J., Salmond, D. J., and Smith, A. F. M. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEE Proceedings-F*, volume 140, pages 107–113.
- Haykin, S. and Van Veen, B. (2001). *Sinais e Sistemas*. Tradução: Santos, J.C.B. Bookman, Porto Alegre.
- Hofbauer, J. and Sigmund, K. (1988). *Theory of Evolution and Dynamical Systems: Mathematical Aspects of Selection*. Cambridge University Press, Cambridge, MA.
- Jaynes, E. T. (1996). *Probability Theory: The Logic of Science*. Washington University (Fragmented Edition).
- Jazwinski, A. H. (1970). *Stochastic Processes and Filtering Theory*. Academic Press.
- Joseph, P. D. (1999). The one dimensional Kalman filter. Available in <<http://ourworld.compuserve.com/homepages/PDJoseph/kalman.htm>>. Accessed in: September, 04, 2003.
- Julier, S. J. (2003). Kalman filter initialization [comunicação pessoal]. Mensagem recebida por <HUDSONBITENCOURT@YAHOO.COM.BR> em 08 de Junho de 2003.

- Julier, S. J. and Uhlmann, J. K. (1996). A general method for approximating nonlinear transformations of probability distributions. Technical report, Robotics Research Group, Department of Engineering Science, University of Oxford.
- Julier, S. J. and Uhlmann, J. K. (1997). A new extension of the Kalman filter to nonlinear systems. In *Proceedings of Aerosense: The 11th International Symposium on Aerospace/Defence Sensing, Simulation and Controls*, Orlando, FL.
- Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F. (1995). A new approach for filtering nonlinear systems. In *The Proceedings of the American Control Conference*, pages 1628–1632, Seattle, Washington.
- Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F. (2000). A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3):447–482.
- Kailath, T. (1980). *Linear Systems*. Prentice Hall, New Jersey.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering*, 82:35–45.
- Kalman, R. E. and Bucy, R. S. (1961). New results in linear filtering and prediction theory. *Transactions of the ASME-Journal of Basic Engineering*, 83:95–108.
- Kalman, R. E., HO, Y. C., and Narendra, K. S. (1962). *Controllability of Linear Dynamic Systems*, volume 1. Wiley-Interscience Publishers, New York.
- Kennedy, J. B. and Neville, A. M. (1986). *Basic Statistical Methods for Engineers and Scientists*. Harper & Row, New York, 3 edition.
- Kennedy, M. P. (1992). Robust op amp realization of chua’s circuit. *Frequenz*, 46(3–4):66–80.
- Kiliç, R. (2003). A comparative study on realization of Chua’s Circuit combining the circuit topologies proposed for Chua’s diode and inductor elements. *International Journal of Bifurcation and Chaos*, 13(6):1475–1493.

- Koslov, A. K., Shalfeev, V. D., and Chua, L. O. (1996). Exact synchronization of mismatched chaotic systems. *International Journal of Bifurcation and Chaos*, 6(3):569–580.
- Kushner, H. J. (1967a). Approximations to optimal nonlinear filters. *IEEE Transactions on Automatic Control*, AC-12(5):546–556.
- Kushner, H. J. (1967b). Dynamical equations for optimum nonlinear filtering. *Journal of Differential Equations*, 3:179–190.
- Leibundgut, B. G., Rault, A., and Gendreau, F. (1983). Application of Kalman filtering to demographic models. *IEEE Transactions on Automatic Control*, AC-28:427–434.
- Ljung, L. (1977). Analysis of recursive stochastic algorithms. *IEEE Transactions on Automatic Control*, 22(4):551–575.
- Ljung, L. (1987). *System Identification, Theory for the User*. Prentice Hall, New Jersey.
- Lorenz, E. N. (1963). Deterministic nonperiodic flows. *Journal of Atmospheric Sciences*, 20:130–141.
- Luenberger, D. (1966). Observers for multivariable systems. *IEEE Transactions on Automatic Control*, 11:190–197.
- Matsumoto, T. (1984). A chaotic attractor from Chua's Circuit. *IEEE Transactions on Circuits and Systems*, CAS 31:1055–1058.
- Maybeck, P. S. (1979). *Stochastic Models, Estimation, and Control*, volume 1, chapter Introduction. Academic Press, New York.
- Maybeck, P. S. (1982). *Stochastic Models, Estimation and Control*, volume 2. Academic Press.
- Mealy, G. L. and Tang, W. (1983). Application of multiple model estimation to a recursive terrain height correlation system. *IEEE Transactions on Automatic Control*, AC-28:323–331.
- Merwe, R. V., Freitas, J. F. G., Doucet, A., and Wan, E. A. (2000). The unscented particle filter. Technical report, CUED/F-INFENG/TR 380, Cambridge University Engineering Department, Cambridge, England.

- Merwe, R. V. and Wan, E. A. (2000). The unscented Kalman filter for non-linear estimation. In *Proceedings of IEEE Symposium 2000 (AS-SPCC)*, Lake Louise, Alberta, Canada.
- Merwe, R. V. and Wan, E. A. (2001a). *Kalman Filtering and Neural Networks*, chapter 7. Wiley.
- Merwe, R. V. and Wan, E. A. (2001b). The square-root unscented Kalman filter for state and parameter estimation. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, Piscataway, NJ. IEEE Signal Processing Society.
- Monteiro, L. H. A. (2002). *Sistemas Dinâmicos*. Editora Livraria da Física, São Paulo.
- Morrell, D. (1997). Extended Kalman filter. Filtering of stochastic process: lecture notes. Available in <<http://www.eas.asu.edu/morrell>>. Accessed in September, 18, 2002.
- Mouri, H., Satoh, S., Furusho, H., and Nagai, M. (2002). Investigation of automatic path tracking using an extended Kalman filter. *JSAE Review*, 23:61–67.
- Norton, J. P. (1986). *An Introduction to Identification*. Academic Press, London.
- Ogata, K. (1993). *Engenharia de Controle Moderno*, chapter Projeto de Sistemas de Controle pelos Métodos de Espaço de Estados, pages 630–670. Prentice Hall do Brasil, Rio de Janeiro, 2 edition.
- Papoulis, A. (1984). *Probability, Random Variables and Stochastic Processes*, chapter Functions of One Random Variable, pages 89–120. McGraw-Hill, 2 edition.
- Peebles, P. Z. (1993). *Probability, Random Variables, and Random Signal Principles*, chapter The Random Variable. McGraw-Hill, 3 edition.
- Peterka, V. (1981). *Trends and Progress in System Identification*, chapter Bayesian approach to system identification, pages 239–304. Pergamon, Oxford.

- Phillips, C. L. and Nagle, H. T. (1995). *Digital Control System Analysis and Design*. Prentice Hall, New Jersey, 3^a edition.
- Poincaré, H. (1957). *Les Methodes Nouvelles de la Mechanique Celeste*. Dover, New York.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. (1992). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 2 edition.
- Sidar, M. M. and Doolin, B. F. (1983). On the feasibility of real-time prediction of aircraft carrier motion at sea. *IEEE Transactions on Automatic Control*, AC-28:350–355.
- Sitz, A., Schwartz, U., Kurths, J., and Voss, H. U. (2002). Estimation of parameters and unobserved components for nonlinear systems from noisy time series. *Physical Review*, E66.
- Sobiski, D. J. and Thorp, J. S. (1998). PDMA-1: Chaotic communication via the Extended Kalman Filter. *IEEE Transactions on Circuits and Systems*, 45(2):194–197.
- Sorenson, H. W. and Stubberud, A. R. (1968). Nonlinear filtering by approximation of the a posteriori density. *International Journal of Control*, 8(1):33–51.
- Tôrres, L. A. B. (2001). *Controle e Sincronismo de Osciladores Caóticos*. Tese de Doutorado do PPGEE, Universidade Federal de Minas Gerais, Belo Horizonte.
- Tôrres, L. A. B. and Aguirre, L. A. (1996). Um protótipo para o estudo de sistemas dinâmicos não-lineares: montagem e sintonia. In *XI Congresso Brasileiro de Automática*, volume 2, pages 881–886, São Paulo, SP.
- Tôrres, L. A. B. and Aguirre, L. A. (2000). Inductorless chua's circuit. *Electronics Letters*, 36(23):1915–1916.
- Tylee, J. L. (1983). On-line failure detection in nuclear power plant instrumentation. *IEEE Transactions on Automatic Control*, AC-28:406–415.

-
- Wallace, J. N. and Clarke, R. (1983). The application of Kalman filtering estimation techniques in power station control systems. *IEEE Transactions on Automatic Control*, AC-28:416–427.
- Wan, E. A., Merwe, R. V., and Nelson, A. T. (2000). *Dual Estimation and the Unscented Transformation*, volume Neural Information Processing Systems, pages 666–672. MIT Press.
- Welch, G. and Bishop, G. (2002). An introduction to the Kalman filter. Technical Report TR 95-041, University of North Carolina at Chapel Hill, Chapel Hill, NC.
- Wiener, N. (1949). *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Wiley, New York.
- Yang, S. K. (2002). An experiment of state estimation for predictive maintenance using Kalman filter on a DC motor. *Reliability Engineering and System Safety*, 75:103–111.
- Young, P. C. (1970). An instrumental variable method for real-time identification of a noisy process. *Automatica*, 6:271–287.
- Zhong, G. Q. (1994). Implementation of Chua’s Circuit with a cubic nonlinearity. *IEEE Transactions on Circuits I*, 41(12):934–941.

Biografia: R. E. Kalman

*“Mentes pequenas discutem pessoas;
Mentes medianas discutem eventos;
Grandes mentes discutem idéias.”*

R. E. Kalman

Rudolf Emil Kalman, filho de Otto e Ursula Kalman nasceu em Budapeste, Hungria, em 19 de Maio de 1930. Sua família mudou-se para os Estados Unidos durante a Segunda Guerra Mundial.

Kalman recebeu o grau de bacharel e mestre em Engenharia Elétrica pelo MIT (*Massachusetts Institute of Technology*) em 1953 e 1954 respectivamente. Dando continuidade a seus estudos, ele recebeu o título de PhD na Universidade de Columbia em 1957.

Durante os anos 60, Kalman foi um dos expoentes no desenvolvimento da teoria de controle moderno. Entre suas inúmeras contribuições pode-se citar a formulação e o estudo dos conceitos mais fundamentais sobre *espaço de estados*, como por exemplo os conceitos de *controlabilidade* e *observabilidade*. Os paradigmas formulados por Kalman e os resultados obtidos por ele se tornaram parte intrínseca dos fundamentos da teoria de controle, e hoje são ferramentas indispensáveis apresentadas em livros-texto de engenharia usados desde cursos de graduação até os níveis mais avançados.

Ele foi pesquisador no *Research Institute for Advanced Study* (RIAS) em Baltimore, entre 1958–1964, professor da Universidade de Stanford entre 1964–1971, professor pesquisador e diretor do *Center for Mathematical System Theory* da Universidade da Flórida entre 1971–1992. Além disso, Kalman lecionou Teoria Matemática de Sistemas no *Swiss Federal Institute of Technology* (ETH).

Kalman recebeu muitos prêmios, incluindo a Medalha de Honra do IEEE (1974), Medalha do Centenário do IEEE (1984), o Prêmio de Kyoto em

Alta Tecnologia da Fundação Inamori do Japão (1985), o Prêmio *Steele* da Sociedade Americana de Matemática (1987) e o Prêmio Bellman (1997). Ele é membro da Academia Nacional de Ciências (EUA), Academia Nacional de Engenharia (EUA) e Academia Americana de Artes e Ciências. Além disso, é membro estrangeiro das Academias de Ciências Húngara, Francesa e Russa e recebeu muitos doutorados honorários.

Índice

- Atualização recursiva, 19
- Bayes
 - abordagem Bayesiana, 15, 17
 - teorema de, 16
- Cholesky
 - decomposição de, 44
- Conhecimento *a priori*, 15
- Covariância, 13
- Equação de correção, 26
- Equação de predição, 25
- Erro
 - critério de avaliação, 61
- Espaço de estados, 2
- Estatística
 - Bayesiana, 15, 17
 - clássica, 15
- Estimação, 9
 - de estados
 - sist. lineares, 22
 - sist. não-lineares, 31
 - de parâmetros, 19, 42
 - em batelada, 9
 - recursiva, 9
- Estimador rec. mín. quadrados, 21
- Estimativa ótima, 35
- Filtro de Kalman, 10, 22, 23
 - algoritmo, 24, 27
 - Estendido, 33
 - algoritmo, 33
 - aplicações, 55
 - estimação de parâmetros, 42
 - sistemas lineares, 10
 - sistemas não-lineares, 31
 - Unscented, 36
 - algoritmo, 40, 85
 - aplicações, 45, 55, 83
 - inicialização, 44, 53, 65, 78
- Função densidade de prob., 13
- Ganho de Kalman, 25
- Identificação de Sistemas, 19
- Inferência estatística, 17
- Informação
 - codificação, 84
 - decodificação, 85
 - recuperação, 84
 - transmissão, 84
- Inovação, 21, 25, 29
- Interpolação, 23
- Lorenz, 49, 86
- Lotka-Volterra, 46
- Modelo
 - determinístico, 10

- estocástico, 10
- Momento, 13, 32
 - central, 13
 - central de segunda ordem, 13
- Ponto sigma, 37
- Probabilidade
 - condicional, 16
 - total, 16
- Ruído
 - branco, 15
 - de medição, 12
 - de processo, 11
 - gaussiano, 13
- Sistema, 1
- Sistema linear, 1
- Sistema não-linear, 1
- Transformação *unscented*, 36
- Transmissão de informação, 83
- Valor esperado, 13
- Valor médio, 13
- Variáveis de estado, 2
- Variável aleatória, 13
- Variância, 13