

DISSERTAÇÃO DE MESTRADO

Estimadores Robustos para Sistemas Lineares e Não-Lineares

Rodrigo Fontes Souto

Brasília, março de 2009

UNIVERSIDADE DE BRASÍLIA

FACULDADE DE TECNOLOGIA

UNIVERSIDADE DE BRASÍLIA
Faculdade de Tecnologia

DISSERTAÇÃO DE MESTRADO

Estimadores Robustos para Sistemas Lineares e Não-Lineares

Rodrigo Fontes Souto

*Relatório submetido ao Departamento de Engenharia
Elétrica como requisito parcial para obtenção
do grau de Mestre em Engenharia Elétrica*

Banca Examinadora

Prof. João Yoshiyuki Ishihara, ENE/UnB
Orientador

Prof. Vítor H. Nascimento, USP
Examinador externo

Prof. Geovany Araújo Borges, ENE/UnB
Examinador interno

Dedicatória

Aos meus pais, Maria de Fátima e Gilson, por simplesmente tudo.

Aos meus irmãos Rafael e Bruno.

Aos meus avós Terezinha e Antônio.

À memória dos avós Dolores e João.

A toda minha família.

Aos meus amigos.

À minha namorada Nolyanne, por sua companhia e carinho.

Rodrigo Fontes Souto

Agradecimentos

O trabalho nesta dissertação não teria sido possível sem o apoio, a influência e o encorajamento de vários indivíduos - família, professores, amigos e colegas - com os quais tive a grande honra de conviver neste período.

Em primeiro lugar, gostaria de expressar minha gratidão ao meu orientador, João Ishihara, por seu constante apoio e orientação, não apenas nos assuntos técnicos, mas também nos momentos em que uma palavra amiga é capaz de valer mais que seus ensinamentos técnicos. Agradeço-o por sua enorme paciência em fazer reuniões diárias por vários meses seguidos. E, por fim, agradeço-o por sua confiança, orientando-me mesmo cômico de todas as minhas dificuldades.

Aos professores Geovany Araújo Borges e Adolfo Bauchspiess por suas orientações durante a graduação e, principalmente, por suas amizades.

Ao Edson Hiroshi que, por sua insistência, obrigou-me a fazer uma cópia desta dissertação um dia antes de eu ter o meu computador pessoal furtado, literalmente salvando este trabalho.

Aos amigos Marçal, Gustavo, Luciano e Alexandre que, em busca de seus sonhos, encontram-se em terras situadas além-mares.

Aos amigos que deixei em Brasília, Mauro e Camilla, Léo e Dê, Carol, Gustavo Amaral, Romero, Paula Barbosa, Maria Helena, Dezan, Gustavo Sá, Gustavo Félix, Saulo, Henrique, André, Gabriel, Alfredo, Hermógenes, Gabi, Silas, Marquinho, Ana Luísa, Frederico, Magno, Vasco, Érika, Geovana, Adovaldo, Maridélia, Renata Silveira e Larissa.

Aos amigos do Instituto Nokia de Tecnologia: Felipe Peixe, Edgar, Gustavão, Léo, Robson, Daniel, André, Agostini, Stefania, Dax, Alexandre e David.

Aos novos amigos que encontrei Yuji, Henrique, Wellington, João Marcelo, Foltran, Paulo Ferrua e Paulo Mestre, Ernesto, Gilliard, Márcio, Latini, Filipe, Milena, Guiliano, Lazarini, Dayvis e tantos outros por seu apoio e incentivo a este trabalho.

E, sobretudo, agradecer a Deus pela oportunidade de experimentar esta chance única e maravilhosa que se chama viver e de colocar pessoas tão maravilhosas em minha vida.

Rodrigo Fontes Souto

RESUMO

O foco desta dissertação é analisar e desenvolver o projeto de estimadores robustos para sistemas lineares e não-lineares de tempo discreto e sujeitos a incertezas. Para o projeto dos estimadores, foi utilizada a abordagem por equação de Riccati, garantindo o desempenho do estimador para todo o conjunto de incertezas permitidas. Generalizou-se a formulação de modelos incertos para o caso em que não se conhecem as propriedades estatísticas dos ruídos, tais como média, covariância e covariância cruzada, uma situação muito comum na prática. No caso não-linear, as funções não-lineares também são supostas incertas, embora pertençam a uma região cônica conhecida.

A principal contribuição deste trabalho foi no sentido de generalizar alguns resultados da literatura que tratam do problema da estimação robusta linear e não-linear, bem como melhorar seu desempenho no sentido de torná-los menos conservadores.

ABSTRACT

The main focus of this dissertation is to develop and analyze robust estimators for linear and nonlinear discrete time systems subject to uncertainties. The design of the estimators, using the Riccati equation approach, leads to a guaranteed cost for all allowed uncertainties within a known set. Assuming norm bounded uncertainties, it is proposed a general robust estimator design to overcome uncertain noise statistics (such as mean, covariance and cross-covariance) and uncertain system linear dynamics, which are a common characteristic in real life applications. In the case of nonlinear systems, the nonlinear functions are also uncertain, but they are assumed to be within a known conic set.

The main contribution of this dissertation is generalize some results found on the literature about linear and nonlinear robust state estimation, as well as obtain less conservatives filters.

SUMÁRIO

1	INTRODUÇÃO	1
1.1	CONTEXTUALIZAÇÃO	1
1.2	MOTIVAÇÃO	2
1.3	VISÃO HISTÓRICA E TRABALHOS RECENTES	4
1.4	INCEZTEZAS E ROBUSTIFICAÇÃO	7
1.5	OBJETIVOS	11
1.6	APRESENTAÇÃO DA DISSERTAÇÃO	11
1.7	NOTAÇÃO E NOMENCLATURA	12
2	FUNDAMENTOS - ESTIMAÇÃO CLÁSSICA	13
2.1	INTRODUÇÃO	13
2.2	O PROBLEMA GERAL DE ESTIMAÇÃO DE ESTADOS	13
2.3	ESTIMAÇÃO LINEAR CLÁSSICA	14
2.3.1	BUSCANDO SIMPLIFICAÇÕES	22
2.4	ESTIMAÇÃO NÃO-LINEAR CLÁSSICA	24
2.4.1	O FILTRO DE KALMAN ESTENDIDO	25
3	ESTIMADORES ROBUSTOS LINEARES	28
3.1	INTRODUÇÃO À ESTIMAÇÃO LINEAR ROBUSTA	28
3.2	OBJETIVOS	29
3.3	MODELAGEM DE SISTEMAS SUJEITOS A INCEZTEZAS	30
3.3.1	O ESPAÇO DAS MÉDIAS E DAS COVARIÂNCIAS DOS RUÍDOS	31
3.3.2	SISTEMAS COM INCEZTEZAS NOS PARÂMETROS DA PLANTA E DOS RUÍDOS	34
3.4	O SISTEMA AUMENTADO INCERTO	37
3.5	COVARIÂNCIA DO SISTEMA AUMENTADO INCERTO	40
3.6	MAJORANTE PARA A EVOLUÇÃO DA COVARIÂNCIA DO SISTEMA AUMENTADO	41
3.6.1	LIMITANTE SUPERIOR PARA A COVARIÂNCIA INCERTA	44
3.7	PARTIÇÃO DO MAJORANTE	47
3.8	MAJORANTE DA COVARIÂNCIA DO ERRO DE ESTIMAÇÃO	48
3.9	MINIMIZAÇÃO DA COVARIÂNCIA DO ERRO DE ESTIMAÇÃO	50
3.10	BUSCANDO SIMPLIFICAÇÕES	50
3.11	REESCREVENDO AS EXPRESSÕES	51
4	ESTIMADORES ROBUSTOS NÃO-LINEARES	53
4.1	ESTIMAÇÃO NÃO-LINEAR ROBUSTA	53
4.2	OBJETIVOS	54
4.3	MODELAGEM DO SISTEMA NÃO-LINEAR INCERTO	54
4.4	MATRIZES DE COVARIÂNCIA NÃO-LINEARES	56
4.5	MAJORANTES PARA AS COVARIÂNCIAS DAS FUNÇÕES NÃO-LINEARES	57

4.6	MAJORANTES PARA OS RUÍDOS COM PROPRIEDADES INCERTAS	59
4.7	MINIMIZAÇÃO DOS MAJORANTES E OBTENÇÃO DO FILTRO	60
5	EXPERIMENTOS NUMÉRICOS	64
5.1	INDICADORES DE DESEMPENHO	64
5.2	INFLUÊNCIA DO PARÂMETRO ESCALAR.....	64
5.3	ESTIMAÇÃO LINEAR	65
5.4	ESTIMAÇÃO NÃO-LINEAR	66
5.4.1	DEMULADOR SEM DISTÚRBO DE FASE	68
5.4.2	DEMULADOR COM DISTÚRBO DE FASE.....	69
5.4.3	DEMULADOR COM VARIAÇÃO DE FREQUÊNCIA E COM DISTÚRBO DE FASE ...	71
6	CONCLUSÕES.....	74
	REFERÊNCIAS BIBLIOGRÁFICAS	75
	ANEXOS	84
I	RESULTADOS COMPLEMENTARES	85
I.1	LEMA DE INVERSÃO DE MATRIZES	85
I.2	MATRIZ INVERSA GENERALIZADA	86
I.3	FUNÇÕES DE LIPSCHITZ	92
I.4	ELIPSES DE CONFIANÇA	94
I.5	CÁLCULO MATRICIAL	95
I.6	RESULTADOS AUXILIARES	96
II	DESCRIÇÃO DO CONTEÚDO DO CD	98
III	ARTIGO PUBLICADO NO MED 2008	99
IV	ARTIGO PUBLICADO NO CBA 2008	100
V	ARTIGO PUBLICADO NO CDC 2008	101
VI	ARTIGO A SER PUBLICADO NA IEEE SIGNAL PROCESSING LETTERS	102
VII	ARTIGO A SER PUBLICADO NO ACC 2009	103

LISTA DE FIGURAS

1.1	Diagrama de blocos para o problema de estimação.	1
1.2	Diagrama de blocos para uma planta suposta conhecida.	8
1.3	Diagrama de blocos para uma planta incerta e com ruídos com propriedades incertas.	8
3.1	Passos para a obtenção do preditor robusto por meio da abordagem da equação de Riccati. .	30
5.1	Influência do parâmetro escalar sobre a variância do erro de estimação do estado 1.	65
5.2	Influência do parâmetro escalar sobre a variância do erro de estimação do estado 2.	65
5.3	Média das elipses de covariância após 500 experimentos.	67
5.4	Variância aproximada do erro de estimação para simulação com ruídos não-correlacionados.	67
5.5	Estado real do sistema e suas estimativas, segundo os preditores analisados.	67
5.6	Exemplo de sinal de saída para o demodulador.	69
5.7	Média dos estados após 500 experimentos.	70
5.8	Variância aproximada do erro de estimação após 500 experimentos.	70
5.9	Média dos estados após 1500 experimentos.	71
5.10	Variância aproximada do erro de estimação após 1500 experimentos.	72
5.11	Exemplo de sinal de saída para o demodulador com frequência variante no tempo de distúrbio de fase de até 25 graus.	73
5.12	Média da amplitude após 500 experimentos.	73
5.13	Variância aproximada do erro de estimação da amplitude após 500 experimentos.	73
I.1	Exemplo de uma função não-linear limitada em cone.	92

LISTA DE TABELAS

2.1	Conjuntos e significados dos vetores utilizados.....	13
2.2	Dimensões das matrizes utilizadas no modelo linear.....	15
2.3	Vetores e matrizes para a definição das covariâncias.....	15
2.4	Algoritmo do FK na sua forma preditora de 1 estágio.....	24
2.5	Algoritmo do FKE na sua forma preditora de 1 estágio.....	27
3.1	Conjuntos das matrizes utilizadas no modelo (3.40)-(3.41).....	34
3.2	Dimensões das matrizes e vetores utilizados no modelo (3.42)-(3.43).....	35
3.3	Preditor Linear Robusto.....	52
4.1	Algoritmo do Filtro Não-Linear Robusto.....	63

LISTA DE SÍMBOLOS

Notação

$\mathbb{R}, \mathbb{R}^+, \mathbb{C}$	Conjunto dos números reais, não-negativos e complexos, respectivamente
\mathbb{R}^n	Espaço euclidiano n -dimensional
$\mathbb{R}^{n \times m}$	Conjunto de matrizes reais com dimensões $n \times m$
\in	Indica que um elemento pertence a determinado conjunto
\forall	Para todo
$\{\bullet\}$	Sequência
$E\{\bullet\}$	Operador de esperança matemática
$cov\{\bullet\}$	Operador de covariância
$var\{\bullet\}$	Operador de variância
$tr(\bullet)$	Operador de traço
$diag(\bullet)$	Matriz bloco diagonal
$posto(\bullet)$	Posto de uma matriz
$\sigma_{max}\{\bullet\}$	Indica o valor singular máximo de uma matriz
$min J_k$	Problema de minimização em que se busca o mínimo de J_k
$min(x_1, x_2)$	O menor valor do conjunto compreendendo x_1 e x_2
$Z > 0$	Matriz definida positiva
$Z \geq 0$	Matriz semi-definida positiva
$\ z\ ^2$	Norma Euclidiana, <i>i.e.</i> , $z^T z$
$\ z\ _W^2$	Para uma matriz definida positiva W , indica a norma Euclidiana ponderada, <i>i.e.</i> , $z^T W z$
$ \bullet $	Módulo ou valor absoluto
δ_{kj}	Função delta de Kronecker
n	Dimensão do vetor de estados
m	Dimensão do vetor de saídas
p	Dimensão do vetor de ruídos de processo
q	Dimensão do vetor de ruídos de medição

Sobrescritos

Z^T	Matriz transposta da matriz Z
Z^{-1}	Matriz inversa da matriz Z
Z^{-T}	Matriz inversa da transposta de Z
Z^\dagger	Matriz pseudo-inversa, ou de Moore-Penrose, da matriz Z
Z^-	Matriz inversa generalizada de Z (ver Anexo I.2)
Z^*	Indica que a matriz Z minimiza um funcional

Matrizes e Vetores

$A_{\Delta,k}$	Matriz incerta de transição de estados
A_k	Matriz nominal de transição de estados
ΔA_k	Matriz que representa as incertezas sobre A_k
C_k	Matriz de saída
I	Matriz identidade com dimensão apropriada
I_n	Matriz identidade com dimensão $n \times n$
0	Matriz de zeros com dimensão apropriada
J_k	Função custo
K_k	Matriz de ganhos de Kalman
x_k	Vetor de estados
x_0	Condição inicial do vetor de estados
\bar{x}_0	Valor esperado do estado inicial x_0
$\hat{x}_{k+1 k}$	Estimação do vetor x_{k+1} no instante k , dadas as medições $\{y_0, y_1, \dots, y_k\}$
$\hat{x}_{k k}$	Estimação do vetor x_k no instante k , dadas as medições $\{y_0, y_1, \dots, y_k\}$
y_k	Vetor de saídas
w_k	Vetor de ruído de processo
\bar{w}_k	Valor esperado de w_k
v_k	Vetor de ruído de medição
\bar{v}_k	Valor esperado de v_k
u_k	Vetor de entradas exógenas e determinísticas
X_0	Matriz de covariância do estado inicial x_0
W_k	Matriz de covariância dos ruídos de processo
V_k	Matriz de covariância dos ruídos de medição

Siglas

ARE	Algebraic Riccati Equation
DARE	Discrete Algebraic Riccati Equation
FKE	Filtro de Kalman Estendido
FKL	Filtro de Kalman Linearizado
FKU	Filtro de Kalman Unscented
IQC	Integral Quadratic Constraint
LD	Linearmente Dependente
LI	Linearmente Independente
LFT	Linear Fractional Transformation
LMI	Linear Matrix Inequalities
LTI	Linear Time-Invariant
MMQ	Método dos Mínimos Quadrados
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
NFT	Nonlinear Fractional Transformation
QFT	Quantum Filtering Theory
REFK	Robust Extended Kalman Filter
SVD	Singular Value Decomposition
SPD	Semidefinite Program

1 INTRODUÇÃO

“As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality”. Albert Einstein (1879-1955) citado por J.R. Newman em The World of Mathematics. Este capítulo apresenta a motivação e os objetivos desta dissertação, bem como sua organização e apresentação.

1.1 CONTEXTUALIZAÇÃO

O foco desta dissertação é a Teoria de Estimação. De maneira geral, a Teoria de Estimação foi desenvolvida com o seguinte intuito: dados os valores de um sinal ¹ observado, também chamado de sinal medido, deseja-se estimar ² (suavizar, filtrar ou prever) os valores de um outro sinal (sinal desejado) que não pode ser diretamente acessado ou está corrompido por ruídos e/ou perturbações.

Vale ressaltar que, na solução do problema de estimação, deve haver uma relação (ou modelo) entre o sinal medido e o desejado. Também deve haver algum critério segundo o qual o sinal desejado é estimado. Neste sentido, é preciso escolher um critério que seja compatível com o modelo ou vice-versa. O problema de estimação é ilustrado resumidamente no diagrama da Figura 1.1.

Na Figura 1.1, deseja-se estimar o sinal x . O sinal y é o sinal medido a partir da planta. O sinal w indica um ruído de entrada, desconhecido, apresentando tipicamente um comportamento estocástico com propriedades estatísticas conhecidas. O problema de estimação consiste em projetar um estimador que seja capaz de fornecer estimativas \hat{x} , a partir das medidas y , que se aproximam de x para diversas realizações de y . Em se tratando de critério, o problema também pode ser vislumbrado classicamente como uma minimização da variância do erro de estimação e , definido como a diferença entre \hat{x} e x . Entretanto, no caso robusto a ser considerado neste estudo, utilizar-se-á a minimização de um limite superior para a variância do erro de estimação por motivos a serem apresentados mais adiante nesta dissertação.

¹A palavra sinal é empregada para se designar um vetor, ou conjunto, de dados.

²Ver seção 1.7 sobre a nomenclatura utilizada nesta dissertação.

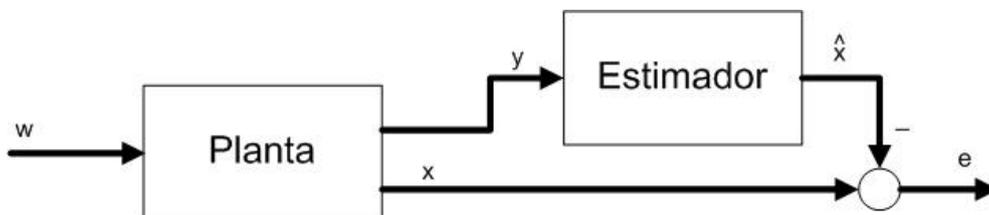


Figura 1.1: Diagrama de blocos para o problema de estimação.

1.2 MOTIVAÇÃO

Nesta dissertação, almeja-se fazer a estimação dos estados de um sistema dinâmico, cujo comportamento pode ser descrito por um conjunto de equações diferenciais. Há pelo menos dois motivos para se conhecer os estados de um sistema, a saber:

- Primeiramente, o conhecimento dos estados pode ser necessário a fim de se controlar o sistema. Como exemplo, um engenheiro eletricista precisa saber as correntes elétricas de um motor com o objetivo de se controlar sua rotação. Um engenheiro aeroespacial precisa estimar a velocidade de um satélite a fim de se controlar sua órbita. Um engenheiro biomédico necessita conhecer o nível de açúcar no sangue para controlar a taxa de injeção de insulina.
- Em outros casos, a necessidade de se conhecer os estados do sistema é interessante por si só. Um engenheiro eletricista que conhece os parâmetros de um sistema de potência pode prever probabilidades de falhas. Um engenheiro aeroespacial, conhecendo a posição de um satélite, pode agendar melhor suas atividades futuras. Um engenheiro biomédico com conhecimento dos níveis de proteínas no sangue pode avaliar melhor a saúde de seu paciente.

Classicamente, os problemas de estimação assumem que o modelo da planta é conhecido com exatidão. Entretanto, na maioria dos casos, modelos utilizados em engenharia fornecem aproximações do sistema real. Esta diferença entre a planta real e o seu modelo aparece por diversos fatores, *e.g.*,

- Mudanças imprevistas dos parâmetros do sistema.
- Dinâmicas não modeladas: não-linearidades e comportamentos dinâmicos que são simplificados, não considerados ou meramente desconhecidos, bem como uma incerteza intrínseca, devido aos procedimentos adotados para a identificação do sistema.
- Retardos, ou atrasos, não incluídos no modelo.
- Mudanças no ponto de equilíbrio (ponto de operação).
- Ruídos de sensores.
- Perturbações não previstas.

Por estes motivos, os estimadores devem ser projetados de maneira a não degradar seu desempenho na presença destas imperfeições e alterações do modelo. Tais estimadores são ditos robustos no sentido em que eles tentam limitar, de certa forma, os efeitos das incertezas no desempenho médio do estimador. Este tema também é conhecido como *Estimação Robusta*.

É cada vez mais notória a utilização do filtro de Kalman e suas extensões e formas em diversos campos do conhecimento. Além de áreas tipicamente da engenharia, *e.g.*, processamento de sinais, controle e telecomunicações, sua utilização tem alcançado novas áreas, tais como, física, economia e biomedicina. A seguir, somente para citar alguns casos, serão apresentadas algumas situações em que o filtro tem sido empregado com sucesso.

- **Controle.** Uma aplicação típica em controle consiste na estimação do vetor de estados. O controle de motores de indução também é estudo da aplicação do filtro de Kalman [1].
- **Telecomunicações.** Aplicações de filtragem robusta em telecomunicações aparecem em problemas de equalizadores de canal [2, 3], estimação de canal [4] e estratégias para controle de taxa e de potência [5].
- **Aeronáutica.** Aplicações em aeronáutica podem ser encontradas visando ao monitoramento de parâmetros que indicam a saúde de um motor turbofan, e.g. [6, 7, 8], e para a determinação de um algoritmo para o aviso de tesouras de vento perigosas próximas à aeronave [9]. Uma aplicação de rastreamento de alvos em solo é encontrada em [10]. Um filtro de Kalman não-linear é utilizado em [11] para estimar a velocidade angular, o ângulo de ataque ³, deflexão da superfície de controle e velocidade linear de um míssil. Em [12], compara-se o desempenho do Filtro de Kalman *Unscented* (FKU) com o Filtro de Kalman Estendido (FKE) para a estimação de estados e de parâmetros de um F-15 em uma manobra complexa e em outra manobra com falha, de forma que o FKU apresentou o resultado mais satisfatório. Outra aplicação do FKE consiste na determinação de órbitas de aeronaves espaciais, tais como apresentadas em [13] e [14]. Outros problemas em que o FKE é utilizado consistem na estimação de altitude [15]-[16], estimação de propriedades da massa [17], calibração de instrumentos com linha de visada [18] e pilotos automáticos de pouso [19].
- **Navegação.** Em [20], descreve-se o problema de navegação baseada em sonar de um robô móvel em que a utilização do FKE resultou em um tratamento menos conservador e melhorou a análise de convergência. O FKE também é utilizado com sucesso para a navegação e rastreamento de navios [21, 22], veículos submersíveis [23], veículos terrestres [24, 25, 26], manipuladores robóticos [27, 28, 29] e aparelhos celulares [30].
- **Redes Neurais e Lógica Difusa.** O treinamento de uma rede neural pode ser realizado por meio de uma técnica baseada no FKE, conforme mostrado em [31].
- **Processamento de Imagem.** A restauração de imagens por meio do filtro de Kalman é apresentada em [32]. O FKE também pode ser utilizado para o rastreamento em imagens digitais, conforme mostrado em [33, 34]. O rastreamento de estradas em imagens aéreas é mostrado em [35] e, em [36], aplica-se o filtro ao rastreamento de pedestres.
- **Processos Industriais.** Em [37], emprega-se o FKE para a estimação de vazamento em dutos.
- **Biomedicina.** Sinais fisiológicos, tais como eletrocardiogramas e pressão arterial, em unidades de tratamento intensivo, são frequentemente corrompidos de maneira severa por ruídos e medições perdidas, o que ocasiona grandes erros na estimação do batimento cardíaco. Em [38], é proposto um filtro de Kalman robusto capaz de lidar com estes problemas até em casos de extrema braquicardia e taquicardia.
- **Aquisição e Processamento de Dados Meteorológicos.** A fim de se estimar os estados de um modelo oceânico, neste caso, um conjunto de variáveis de prognóstico, propõe-se em [39] um filtro de Kalman não-linear de ordem reduzida. A motivação dos autores está no fato de que neste tipo de

³Ângulo entre a linha de referência da aeronave, ou míssil, e o vetor da velocidade do vento.

simulação, lida-se com um conjunto muito grande de estados, inviabilizando a utilização do filtro de Kalman de ordem completa.

- **Economia.** Recentemente, em [40], utilizou-se o filtro de Kalman tradicional para se estimar a demanda por reservas internacionais. Há ainda modelos em que um determinado fato causa impacto tanto nas variáveis de saída quanto nos estados do modelo, resultando em um sistema com ruídos correlacionados. Um exemplo é o mercado de taxas de câmbio em que uma nova informação altera o valor da taxa bem como o prêmio de risco. Esta nova informação pode ser atribuída a um novo regime monetário, declaração de moratória ou o anúncio da conversão da moeda de um país. Em [41], podem-se encontrar os modelos acima, bem como a utilização do filtro de Kalman com ruídos correlacionados para tais sistemas. Um *framework* para a estimação de parâmetros de volatilidade pode ser encontrado em [42]. Para outras aplicações de filtragem e controle robusto para modelos econômicos, consultar [43].
- **Fusão de dados.** Em [44], utilizam-se filtros robustos para fusão de dados com multi-sensores. Uma aplicação de fusão de dados para navegação é estudada em [26].
- **Física.** Até o final da década de 70, acreditava-se que os sistemas quânticos possuíam um comportamento incontrolável no processo de medida. O problema de controlabilidade quântica foi então relacionado ao problema de observabilidade quântica, tanto para tempo contínuo [45] quanto discreto [46], que pode ser resolvido por meio da filtragem quântica. Neste sentido, os filtros quânticos dinâmicos foram propostos para os problemas de controle ótimo quântico. A solução explícita [47] para a filtragem ótima linear quântica e o controle de processos de Markov-Gauss quânticos anteciparam a solução geral [48, 49] dos filtros de Markov quânticos. Os filtros quânticos não-lineares, para uma esperança condicional *a posteriori*, também foram representados na forma de equações de ondas estocásticas [48, 49]. A chamada *Quantum Filtering Theory* (QFT) e as equações estocásticas quânticas possuem hoje várias aplicações. Em [50], utiliza-se o QFT para a estimação e o controle de estados quânticos, tais como o spin e um escalar referente a um campo magnético. Em [51], o filtro de Kalman quântico é adotado para se melhorar um processo de magnetometria atômica. Um tutorial recente à filtragem quântica é desenvolvido em [52].

1.3 VISÃO HISTÓRICA E TRABALHOS RECENTES

O início da teoria de estimação, no sentido de tentar minimizar uma função relacionada ao erro cometido, é atribuído a Galileu Galilei em 1632 [53]. Mais de um século depois, em 1795, Gauss [54] provavelmente foi o primeiro a utilizar o Método dos Mínimos Quadrados (MMQ) com o objetivo de resolver sistemas de equações indeterminados⁴ com base em estimativas que melhor se aproximavam das observações no sentido de mínimos quadrados. Todavia, as primeiras publicações acerca do assunto são atribuídas a Legendre em 1805 [55] e Adrain em 1808 [56]. Desde então, uma extensa bibliografia sobre o MMQ determinístico e estimação MMQ com variáveis aleatórias desenvolveu-se e pode ser encontrada em diversos livros de Álgebra Linear, tais como, por exemplo, em [57].

⁴Sistemas lineares com mais equações que incógnitas

O problema de estimação de mínimos quadrados para processos estocásticos foi investigado pela primeira vez por Kolmogorov [58], Krein [59] e Wiener [60]. Segundo [61], embora a abordagem de Kolmogorov fosse mais fundamental, o trabalho de Wiener, notadamente o filtro de Wiener para a predição de processos estocásticos estacionários, tornou-se mais influente. A principal contribuição dos trabalhos de Kolmogorov e Wiener foi a introdução de noções de estatística nos problemas de estimação e controle. Neste contexto, assume-se que os sinais medidos são processos estocásticos com propriedades estatísticas de primeira e segunda ordens conhecidas e invariantes no tempo. O critério de desempenho adotado foi também o do mínimos quadrados. Esta teoria permitiu extrair um sinal a partir das observações do sinal e das perturbações. Por outro lado, a teoria de Kolmogorov-Wiener não fora aplicada de maneira extensa [62], pois a teoria necessitava da solução de uma equação integral, neste caso, a equação de Wiener-Hopf. Em problemas reais, a equação de Wiener-Hopf raramente apresenta uma solução analítica, e sua solução numérica não é fácil de se obter.

O início da utilização de computadores digitais influenciou de maneira evidente o desenvolvimento da teoria no final dos anos 50 e início dos 60 [62]. De fato, a hipótese central da teoria de Kolmogorov-Wiener, que assume processos estacionários, mostrou-se inadequada em aplicações em que os sinais e os ruídos são não-estacionários. Neste contexto, uma importante contribuição para a teoria de filtragem foi dada por Kalman [63]. Seus resultados permitiram uma solução recursiva para o problema de filtragem e de predição utilizando a capacidade de processamento e de armazenamento dos computadores digitais. Por sua vez, a execução dos filtros de Wiener utilizava processamento analógico por meio de elementos ativos e passivos [64]. Sua contribuição também era uma extensão dos trabalhos de Kolmogorov-Wiener no sentido de lidar com processos não-estacionário. Esta nova teoria passou a ser conhecida como *filtro de Kalman*.

Para a execução do filtro de Kalman, necessita-se resolver um problema de valor inicial para a equação de Riccati. Esta equação de Riccati é similar à encontrada na resolução de problemas de controle ótimo. De fato, os problemas de predição e de controle linear quadrático são matematicamente duais. Esta abordagem mostrou-se interessante em aplicações para trajetórias e controle de veículos aeroespaciais [61].

Até então, todos os modelos assumiam, por hipótese, um conhecimento perfeito de seus parâmetros e de suas propriedades estatísticas. Entretanto, em muitas aplicações, há erros de modelagem, *e.g.*, proveniente de simplificações e desgaste dos sistemas físicos e desconhecimento das propriedades estatísticas do modelo. Destarte, a aplicação direta dos filtros disponíveis resultava em uma queda do desempenho dos filtros. Este fato motivou os pesquisadores da época a pensar em como projetar estimadores eficientes, mesmo com as deficiências do modelo. Esta abordagem para o problema de filtragem passou a ser conhecida como *teoria de estimação robusta*.

Neste contexto, uma outra técnica desenvolvida com a finalidade de se obter um filtro robusto é a filtragem \mathcal{H}_∞ , em que as fontes de ruído são sinais de energia ou de potência média limitadas. Embora mais robusto que o filtro de Kalman, o filtro \mathcal{H}_∞ também é afetado pela mesma dependência sobre a certeza dos parâmetros do modelo e, assim, surgiu o estudo sobre os filtros \mathcal{H}_∞ robustos. Na filtragem \mathcal{H}_∞ robusta são aplicadas metodologias similares às empregadas para os filtros de Kalman robustos (ver, *e.g.*, [65] e [66]).

A partir da Segunda Guerra Mundial, o engenheiro possuía métodos clássicos de controle e de estimação, *i.e.*, transformadas- z e equações de Euler. A partir das décadas de 60 e 70, passou-se a utilizar métodos

“modernos”, com ênfase em algoritmos recursivos que pudessem ser executados por computadores. Por outro lado, até esta época, assumia-se que o modelo que descrevia determinado sistema real fosse verdadeiro, ou seja, não se fazia nenhuma alusão de que o modelo contivesse imperfeições. Neste contexto, surgiu um novo paradigma em que o engenheiro compartilha os mesmos objetivos de seus predecessores; porém, considerando que seu modelo é uma aproximação de um sistema real que fornece suas medições e que, de fato, é-lhe desconhecido e não pode ser determinado perfeitamente. Neste sentido, buscaram-se estimadores que funcionem bem para um conjunto de modelos que sejam próximos ao seu modelo aproximado.

Apresentam-se, a seguir, algumas recentes contribuições para o problema de estimação robusta.

Em [67] é estudado o projeto do filtro de Kalman robusto para horizontes finito e infinito. O modelo considerado é sujeito a incertezas limitadas. O projeto proposto garante um limite superior para a variância do erro. Também são investigadas as condições necessárias e suficientes para a existência do filtro.

Em [68], propõe-se a escolha de variáveis na equação de Riccati de forma a se obter otimização e robustez locais. O filtro realiza a minimização, em cada iteração, do pior caso para o quadrado da norma do resíduo. Os algoritmos foram apresentados de três formas distintas: medida-atualização, previsão e informação. Uma extensão deste estudo fora realizada em [69], em que as matrizes de regulação são determinadas de maneira a minimizar globalmente a matriz de covariância do erro de estimação. Em [70] é apresentada uma extensão destes resultados para sistemas descritores.

Em [71] estuda-se o problema do filtro de Kalman robusto para horizonte finito utilizando parâmetros escalares que podem ser otimizados por meio de programação semidefinida, melhorando a técnica vigente naquela época.

Em [72] são mostrados procedimentos para melhorar a margem de robustez de filtros robustos por meio de parâmetros escalares. Este procedimento foi aplicado em três filtros robustos: o filtro \mathcal{H}_∞ , o filtro com custo garantido e o filtro utilizado em [68]. O parâmetro escalar é escolhido como a raiz quadrada do inverso da solução definida positiva de determinada desigualdade matricial. Esta escolha foi motivada a fim de se obter um estimador com uma matriz estável em malha fechada, cujo maior valor singular é limitado superiormente pela unidade.

Em [73] é proposta uma abordagem para o problema da filtragem robusta de horizonte finito para sistemas discretos no tempo afetados por ruídos limitados e incertezas limitadas em todas as matrizes do sistema. Tal método também aplica-se para uma dependência não-linear das incertezas no modelos de espaço de estados. Tanto as estimativas quanto seus elipsóides de confiança são calculados em tempo polinomial por meio de otimização convexa, nos casos de minimização do volume e do traço.

Em [74], trata-se da filtragem robusta para sistemas estocásticos bilineares no tempo discreto sujeitos a incertezas. As incertezas são limitadas e aparecem nas matrizes de estado e de medida. O projeto do filtro é resolvido por meio da resoluções de desigualdades matriciais. Também são dadas as condições de existência do filtro.

Em [3] é apresentado um filtro de Kalman robusto recursivo para sistemas lineares variantes no tempo com incertezas estocásticas. O algoritmo possui um passo de previsão e minimiza um limite superior do erro quadrático médio a cada iteração, sendo que o problema de minimização é reduzido em um problema de otimização convexa baseado em LMIs (*Linear Matrix Inequalities*). Também é apresentada a conver-

gência quando o sistema é quadraticamente estável e as matrizes de estado são invariantes no tempo.

Em [2] são resolvidos dois problemas de filtragem ótima para sistemas contendo incertezas determinísticas e estocásticas simultaneamente. O primeiro problema consiste em projetar um filtro LTI (*Linear Time-Invariant*) que minimiza um limite superior do ganho de energia entre o ruído que afeta o sistema e o erro de estimação. O segundo problema trata do projeto de filtro, também linear e invariante no tempo, que minimiza um limite superior do erro assintótico quadrático médio, considerando a atuação de ruído branco na planta. Em ambos os problemas, os parâmetros do filtro são determinados por meio de otimização convexa baseada em LMIs.

Em [75, 76] é proposto um filtro robusto para sistemas discretos variantes no tempo sujeito a incertezas com base em uma formulação de mínimos quadrados com regulação. O algoritmo proposto garante variância mínima para o erro de estimação e é compatível para aplicações em tempo real. O problema é resolvido para duas classes de incertezas: *i*) limitadas pela norma na matriz de saída e *ii*) incertezas estocásticas e politópicas.

Em [77] é resolvido o problema de filtragem robusta multi-objetiva para sistemas discretos sujeitos a incertezas estocásticas e determinísticas conjuntamente, bem como sistemas com atraso e sistemas não-lineares. É desenvolvido um procedimento para projetar filtros exponencial e linearmente estáveis e com garantia de um limite mínimo para o erro da variância.

Em [78], propõe-se o projeto de filtros \mathcal{H}_2 e \mathcal{H}_∞ com incertezas estruturadas e não-estruturadas, limitadas e representadas a partir de uma LFT. Seu principal resultado é mostrar que o problema de minimizar os limites superiores em processos com incertezas LFT é convertido em um problema de otimização convexa envolvendo LMIs, estendendo resultados obtidos em [79] e [80]. Mostrou-se que, no caso de incertezas não-estruturadas e limitadas, os limites superiores podem ser minimizados diretamente, implicando em resultados menos conservadores do que os disponíveis na época.

Em [81] é considerado o problema de filtragem para uma classe de sistemas discretos e variantes no tempo com incertezas limitadas e com medições perdidas.

Em [82] é desenvolvido um filtro de Kalman robusto para sistemas lineares discretos e variantes no tempo com incertezas na matriz de covariâncias dos ruídos, dos estados e das saídas. As matrizes de estado e de ganho são otimizadas com o objetivos de resultar em um limite inferior na covariância do erro de estimação para todas as incertezas admissíveis.

Em [83] é projetado um filtro \mathcal{H}_2 robusto para sistemas lineares representado em espaço de estados sujeitos a parâmetros incertos e variantes no tempo com variação limitada.

1.4 INCERTEZAS E ROBUSTIFICAÇÃO

Vale ressaltar que a incerteza no modelo não é a única preocupação quando se trata de estimação robusta. Outras considerações incluem, por exemplo, falhas de sensores e atuadores e restrições físicas do sistema. Ainda mais: quando o projeto do estimador é baseado em uma otimização, a função custo pode não descrever satisfatoriamente o problema. Além disso, os algoritmos podem não ser robustos numericamente. Neste sentido, quando a palavra robustez for utilizada nesta dissertação, será com respeito

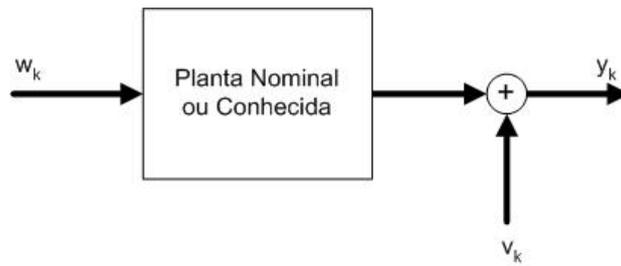


Figura 1.2: Diagrama de blocos para uma planta suposta conhecida.

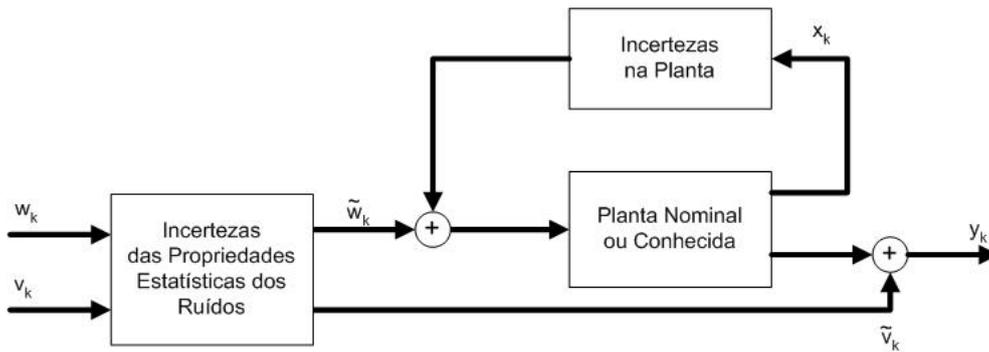


Figura 1.3: Diagrama de blocos para uma planta incerta e com ruídos com propriedades incertas.

às incertezas dos modelos e dos ruídos.

A Figura 1.2 mostra o diagrama de blocos de um sistema de tempo discreto sujeito a ruídos de processo $\{w_k\}$ e de medida $\{v_k\}$ e em que se supõe o conhecimento preciso da planta.

Por outro lado, a Figura 1.3 ilustra como fica o diagrama de blocos para o caso em que há incertezas no modelo da planta e das propriedades estatísticas dos ruídos de processo e de medida. Com o intuito de lidar com estes tipos de sistemas, assume-se que o comportamento da planta é descrito não somente por um único modelo, mas por um conjunto conhecido de modelos possíveis. Assim, apesar de a planta e as propriedades dos ruídos não serem conhecidas, supõe-se que eles pertençam a um conjunto conhecido.

As incertezas possuem diversas causas. De forma geral, podem ser classificadas em duas classes:

- Incertezas paramétricas ou estruturadas - a estrutura do modelo é conhecida, inclusive a sua ordem [84]. Entretanto, alguns parâmetros são incertos. A incerteza pode advir de diferentes fatores, *e.g.*, imprecisão na medida, erro de identificação, linearização em torno de um ponto de operação, variação temporal dos parâmetros.
- Incertezas não-paramétricas ou dinâmicas não modeladas - resultam da falta de modelagem da dinâmica da planta para altas frequências ou simplesmente do desconhecimento da sua dinâmica de ordem mais alta.

As técnicas de estimação robusta visam lidar com as incertezas do modelo. Melhorar a robustez do estimador significa melhorar o seu desempenho médio quando o sistema está sujeito a incertezas. Portanto, a descrição dos modelos das incertezas é um fator importante ao se considerar suas técnicas de resolução. Os seguintes modelos de incertezas são os mais encontrados na literatura:

- Incertezas com norma limitada - assume-se que os parâmetros incertos encontram-se em um conjunto com norma limitada, podendo ser estruturados ou não-estruturados.
- Incertezas politópicas - os parâmetros incertos do modelo estão contidos em um politopo, formando um conjunto estruturado.
- LFT (*Linear Fractional Transformation*) - trata-se de uma descrição mais geral e pode ser utilizada tanto para parâmetros incertos quanto para dinâmicas não-modeladas. Neste modelo, as partes conhecidas e desconhecidas do modelo da dinâmica da planta apresentam uma conexão parecida com uma realimentação.
- IQC (*Integral Quadratic Constraint*) - permite descrever ambos os modelos de incertezas descritos anteriormente. Consiste em utilizar conjuntos convexos de formas que satisfazem determinada condição envolvendo um integral quadrática [85] para caracterizar as incertezas, resultando em um problema de otimização convexa envolvendo o modelo nominal e o conjunto convexo de IQCs.

Em se tratando das técnicas de projeto de estimadores robustos, destacam-se as seguintes abordagens:

- Equações de Riccati.
- Construção de funções de Lyapunov, transformando-o em um problema de SPD (*Semidefinite Programs*).
- Otimização direta explorando características do filtro.

As duas primeiras técnicas baseiam-se em encontrar um limite superior para o critério de desempenho e, em seguida, tentar minimizá-lo. Por este motivo, estas técnicas também são conhecidas como métodos indiretos. A terceira abordagem, por outro lado, é considerada como um método direto.

Uma característica conhecida sobre o filtro de Kalman é que ele pode não apresentar uma robustez satisfatória quando, no modelo, estão associados distúrbios e incertezas [85, 86, 87, 73, 88, 71]. Este fato tem sido a motivação de diversas pesquisas no sentido de projetar filtros que sejam robustos às incertezas e distúrbios e que, apesar de não apresentarem um resultado ótimo para o modelo nominal, garantam um desempenho de estimação para todas as possíveis incertezas consideradas. Este problema é conhecido na literatura como *filtragem de Kalman robusta* ou *estimação \mathcal{H}_2 robusta* [86].

De acordo com [86], a vantagem da abordagem por **equação de Riccati** é que o efeito das incertezas na estrutura e no ganho do filtro é claramente perceptível, permitindo uma melhor compreensão do problema. Uma das técnicas utilizadas envolve a solução de duas equações de Riccati discretas, *e.g.*, [67, 74, 86, 89]. Outra possibilidade é baseada na solução de somente uma equação de Riccati, ver [90, 82]. Entretanto, no horizonte finito, ao contrário do filtro de Kalman clássico, o filtro robusto ótimo em k pode não conduzir a uma estimação ótima no instante $k + 1$, conforme constatado em [71].

A utilização de **LMI's** vem-se consolidando nos últimos anos como uma ferramenta útil na solução de diversos problemas em controle, *e.g.*, o problema de estimação. Para mais detalhes, vide [91, 92, 93]. De acordo com [85], a maior vantagem no uso de LMI's deve-se à sua flexibilidade para lidar com sistemas dinâmicos sujeitos a incertezas de maneira computacionalmente eficiente.

De fato, já é conhecido que o projeto de filtros robustos para sistemas com incertezas com norma limitada ou contidas em um politopo pode ser caracterizado como uma otimização convexa envolvendo LMI's. Entretanto, uma das desvantagens desta metodologia é que ela resulta, geralmente, em filtros mais conservadores [86]. Assim, grande esforço tem sido feito no sentido de se obterem projetos menos conservadores.

Na abordagem **MinMax**, a fim de se lidar com as incertezas do modelo, o projeto do estimador é caracterizado por uma solução de equilíbrio de um problema de otimização min-max [94]. A formulação matemática consiste na resolução do seguinte problema

$$\min_{\hat{x}_{k|k}} \max_{\Delta_k} J, \quad (1.1)$$

em que J é a medida de desempenho do estimador, $\hat{x}_{k|k}$ é o vetor estimado e Δ_k representa as incertezas do modelo. Para o caso tratado aqui, J é a norma \mathcal{H}_2 do erro de estimação. Outro índice é proposto em [95], em que se utiliza a taxa do erro quadrático médio. A estimação robusta é então encarada como um jogo homem-natureza, em que a natureza seleciona o pior valor possível das incertezas no sentido de maximizar o erro de estimação, *i.e.*, maximizando a sua norma \mathcal{H}_2 . O homem, por sua vez, tenta minimizar o erro de estimação escolhendo os parâmetros do estimador que irão minimizar o pior efeito possível das incertezas no modelo. A solução de equilíbrio, se existir, fornecerá então o melhor filtro robusto para o pior caso de incerteza. Entretanto, para o caso geral, o cálculo deste ponto de equilíbrio tem sido um desafio na literatura. Esta abordagem é estudada com mais detalhes em [96] e [97].

Outra possibilidade de estimação \mathcal{H}_2 robusta é obtida por meio de **soluções polinomiais**. Pode-se citar, por exemplo, o trabalho desenvolvido em [97]. Neste trabalho, apresenta-se o projeto de um filtro \mathcal{H}_2 robusto baseado em uma fatorização do espectro e em uma equação de Diophantine unilateral. O sistema adotado é linear e invariante no tempo, de forma que este estimador robusto também é conhecido como filtro de Wiener robusto.

Uma abordagem para o problema de estimação robusta é conhecida como *Estimação \mathcal{H}_∞* . Originalmente, a formulação desenvolvida por [98] fora aplicada para o problema de controle robusto. A teoria de controle robusto surgiu da necessidade de se projetar controladores que fossem insensíveis a erros de modelamento da planta e à falta de informações estatísticas sobre sinais externos que atuavam sobre o sistema. Os estimadores \mathcal{H}_∞ tiveram seu desenvolvimento no final de década de 80 e utilizam como critério de desempenho a minimização da norma \mathcal{H}_∞ .

A principal idéia da estimação \mathcal{H}_∞ é projetar estimadores para minimizar, ou limitar, no caso sub-ótimo, a energia máxima transferida das perturbações para o erro de estimação. Procedendo desta forma, consegue-se garantir que, se a energia das perturbações for pequena, então os erros de estimação serão os menores possíveis, em energia, independentemente de que tipo são as incertezas. A robustez dos estimadores \mathcal{H}_∞ advém do fato de eles serem projetos para o pior caso das incertezas possíveis, não fazendo qualquer hipótese sobre a distribuição e propriedades estatísticas dos sinais de distúrbio. Neste sentido, o projeto utilizando a abordagem \mathcal{H}_∞ pode resultar em estimadores mais conservadores.

Entretanto, estes estimadores também não são robustos em relação a incertezas no modelo [85]. Da mesma forma que no caso da estimação \mathcal{H}_2 , também foram desenvolvidos filtros \mathcal{H}_∞ robustos sujeitos a incertezas de norma limitada. Um dos primeiros projetos para o desenvolvimento de filtros \mathcal{H}_∞ robustos pode ser encontrado em [65, 99].

1.5 OBJETIVOS

O objetivo desta dissertação é analisar e desenvolver o projeto de estimadores robustos para sistemas de tempo discretos lineares e não-lineares sujeitos a incertezas. O escopo desta dissertação contempla estimadores de horizonte finito, incertezas de norma limitada e a abordagem por meio de Equação de Riccati, garantindo o desempenho do estimador para todo o conjunto de incertezas permitidas. Admitindo incertezas limitadas por norma, generalizou-se a formulação do sistema para o caso em que não se conhecem as propriedades estatísticas dos ruídos, tais como média, covariância e covariância cruzada, uma situação muito comum na prática. No caso não-linear, as funções não-lineares também são incertas, mas pertencendo a uma região cônica conhecida. Esta condição é conhecida como uma condição de Lipschitz nos resíduos dos estados e do sinal de controle do sistema. Neste sentido, deseja-se generalizar alguns resultados da literatura que trata do problema da estimação robusta linear e não-linear, aumentando o conjunto de sistemas possíveis que podem ser tratados por meio da estimação robusta.

Esta dissertação visa melhorar também o desempenho dos estimadores, por meio do acréscimo de parâmetros escalares, tornando-os menos conservadores, sem aumentar demasiadamente seu custo computacional.

Resultados desta dissertação foram publicados, ou foram aceitos para publicação, em quatro congressos e um periódico, a saber:

- *IEEE Signal Processing Letters.*
- *47th IEEE Conference on Decision and Control.*
- *2009 American Control Conference.*
- *16th Mediterranean Conference on Control and Automation.*
- XVII Congresso Brasileiro de Automática.

1.6 APRESENTAÇÃO DA DISSERTAÇÃO

A estrutura da dissertação é mostrada na sequência:

- O capítulo 2 apresenta os tradicionais estimadores de Kalman para sistemas lineares e não-lineares.
- No capítulo 3 são desenvolvidas e analisadas técnicas para aumentar a robustez do estimador de Kalman discreto para modelos lineares sujeitos a incertezas em seus parâmetros.
- No capítulo 4 são desenvolvidas e analisadas técnicas para aumentar a robustez do filtro de Kalman discreto para modelos não-lineares sujeitos a incertezas em seus parâmetros e em suas funções não-lineares.
- Os resultados numéricos que confirmam os resultados teóricos obtidos no capítulo 3 são apresentados no capítulo 5, tanto para sistemas lineares quanto para os não-lineares.

- As conclusões pertinentes são apresentadas no capítulo 6.
- O anexo I apresenta alguns resultados e conceitos complementares e auxiliares à leitura desta dissertação.
- O anexo II contém a descrição de materiais complementares contidos no CD entregue à Secretaria do Departamento de Engenharia Elétrica da Universidade de Brasília.

1.7 NOTAÇÃO E NOMENCLATURA

A nomenclatura aqui utilizada é padrão e encontrada em diversas referências acerca do tema. Para uma descrição da notação e dos símbolos adotados, olhar a lista de símbolos no início desta dissertação.

Na maior parte da dissertação, faz-se a distinção entre as palavras *estimador*, *filtro* e *preditor*. *Estimador* é um algoritmo matemático capaz de avaliar determinada massa de dados obtidos no presente, passado ou futuro. *Preditor* é um estimador que, somente com dados do presente, procura antever qual será o próximo dado a ser obtido. Um *filtro*, por sua vez, tenta melhorar uma predição a partir de dados que se tornaram disponíveis no presente. Esta questão de nomenclatura costuma gerar confusão na literatura e, não raro, é possível encontrar autores que a utilizam erroneamente, conforme constatado em [100]. Por fim, um *suavizador* é capaz de fornecer uma estimativa com base em dados passados, presentes e futuros.

Entretanto, o assunto é um pouco mais complicado. No caso clássico, a função de transferência do filtro e a do preditor de Kalman é a mesma, resultando em um sistema chamado "filtro de Kalman". O filtro de Kalman pode estar na forma preditora, filtrada, suavizada ou informação. Por outro lado, no caso robusto estudado nesta dissertação, os algoritmos do filtro e do preditor não resultam mais no mesmo "sistema filtro", e a distinção de nomenclatura faz-se necessária novamente.

2 FUNDAMENTOS - ESTIMAÇÃO CLÁSSICA

“Finally, we make some remarks on why linear systems are so important. The answer is simple: because we can solve them!” R. Feynman [101]. Neste capítulo, apresentar-se-á um dos estimadores mais conhecidos: o filtro de Kalman e suas variações para sistemas lineares e não-lineares.

2.1 INTRODUÇÃO

Neste capítulo será apresentado brevemente o problema geral de estimação de estados para sistemas sem incertezas. Em seguida, será apresentada sua solução para o caso linear com ruídos gaussianos, ou seja, o algoritmo do filtro de Kalman na sua forma preditora de 1 estágio. Há várias formas conhecidas na literatura de se demonstrar tal resultado, seja por argumentos estocásticos, determinísticos, geométricos ou probabilísticos. Em especial, a demonstração do preditor apresentada neste capítulo é de especial importância por facilitar o entendimento e a obtenção de uma forma robusta para o preditor de Kalman. Ao final do capítulo, mostrar-se o Filtro de Kalman Estendido (FKE) que é uma solução aproximada para sistemas não-lineares.

2.2 O PROBLEMA GERAL DE ESTIMAÇÃO DE ESTADOS

De maneira genérica, um sistema dinâmico discreto no tempo pode ser escrito como

$$x_{k+1} = f_k(x_k, u_k, w_k), \quad (2.1)$$

$$y_k = h_k(x_k, u_k, v_k), \quad (2.2)$$

em que os conjuntos e os significados de cada vetor são mostrados na Tabela 2.1.

Tabela 2.1: Conjuntos e significados dos vetores utilizados.

Vetor	Conjunto	Vetor
x_k	\mathbb{R}^{n_x}	estados do sistema
y_k	\mathbb{R}^{n_y}	saídas do sistema
u_k	\mathbb{R}^{n_u}	sinal de controle (entrada exógena)
w_k	\mathbb{R}^{n_w}	ruídos de processo
v_k	\mathbb{R}^{n_v}	ruídos de medida

Admite-se que as funções $f_k(\bullet)$ e $h_k(\bullet)$ podem ser não-lineares e o subscrito em $f_k(\bullet)$ é incluído para representar uma possível dependência do tempo. Desta forma, dados

- os sinais de entrada $\{u_k\}$ e de saída $\{y_k\}$ do sistema,

- as características estocásticas das perturbações $\{w_k\}$ e $\{v_k\}$,
- as características do estado inicial x_0 , e sua relação com as perturbações,

o problema de estimação de estados consiste em se determinar a seguinte densidade de probabilidade, ou densidade *a posteriori*,

$$p(x_k|U_r, Y_r), \quad (2.3)$$

em que $U_r = \{u_1, u_2, \dots, u_r\}$ representa uma sequência de entradas conhecidas e Y_r corresponde a uma sequência de observações do sistema da forma $Y_r = \{y_1, y_2, \dots, y_r\}$. Em outras palavras, dada uma realização da sequência de observações $\{y_1, y_2, \dots, y_r\}$ e de entradas $\{u_1, u_2, \dots, u_r\}$, o problema de estimação consiste em estimar o valor de x_k baseado nos conjuntos Y_r e U_r . Quando há mais medidas do que pontos a serem estimados, *i.e.*, $k < r$, este problema é conhecido como *suavização discreta* ou *interpolação*. No caso em que se almeja estimar x_k a partir de medidas y_r obtidas até o instante k , ou seja, $k = r$, tem-se um problema conhecido *filtragem discreta*. Entretanto, quando se estimar x_k a partir de medidas y_r obtidas até, no máximo, no instante $k - 1$, ou seja, $k > r$, o problema torna-se um problema de *predição discreta*.

A função de densidade condicional de x_k dados U_r, Y_r , ou seja, $p(x_k|U_r, Y_r)$ é a solução completa do problema de estimação. Isto é verdade, uma vez que $p(x_k|U_r, Y_r)$ contém todas as informações estatísticas a respeito de x_k a partir das medições disponíveis e da condição inicial.

No caso linear gaussiano, pode-se mostrar que a função de densidade (2.3) é gaussiana [13]. Consequentemente, ela pode ser caracterizada simplesmente por seu vetor de média e sua matriz de covariância. No caso não-linear, o problema é mais complexo: de maneira geral, não há um conjunto finito de parâmetros capazes de representar tal função densidade. Neste sentido, diz-se que no caso linear o estado do filtro é finito, enquanto no caso não-linear, o estado do filtro é dito infinito [13].

2.3 ESTIMAÇÃO LINEAR CLÁSSICA

Conforme mencionado na seção 2.2, o problema de estimação está resolvido, se a função de densidade de probabilidade for determinada. Esta é a essência da abordagem desenvolvida em [63] por meio de argumentos geométricos. Entretanto, esta abordagem possui seus pontos fracos e/ou restritivos, tais como

- hipótese de um modelo linear,
- hipótese de processos gaussianos,
- uma interpretação com pouco ou nenhum significado em alguns contextos, tais como em análise de séries temporais.

Neste sentido, surgiram maneiras alternativas de se obter as expressões do filtro de Kalman. Uma delas é supor que o preditor possui uma estrutura fixa. Ressalta-se que o objetivo do projeto é escolher os parâmetros do preditor que minimizem a variância do erro de estimação. Esta técnica é especialmente útil e será utilizada mais adiante para o desenvolvimento de estimadores robustos. Inspirado pelo formato da

solução fornecida pelo filtro de Kalman, assume-se que o preditor possua o seguinte formato

$$\hat{x}_{0|-1} = E \{x_0\} = \bar{x}_0, \quad (2.4)$$

$$\hat{x}_{k+1|k} = \Phi_k \hat{x}_{k|k-1} + K_k (y_k - C_k \hat{x}_{k|k-1}), \forall k \in (0, N], \quad (2.5)$$

em que Φ_k e K_k são os parâmetros do estimador.

Assume-se que a planta possua uma dinâmica linear e discreta no tempo descrita por

$$x_{k+1} = A_k x_k + B_{w,k} w_k, \quad (2.6)$$

$$y_k = C_k x_k + D_{v,k} v_k, \quad (2.7)$$

em que as dimensões e os significados dos vetores são mostrados na Tabela 2.1. Admitem-se matrizes A_k , $B_{w,k}$, C_k e $D_{v,k}$ conhecidas e variantes no tempo, cujas dimensões são apresentadas na Tabela 2.2.

Tabela 2.2: Dimensões das matrizes utilizadas no modelo linear.

Matriz	Conjunto
A_k	$\mathbb{R}^{n_x \times n_x}$
$B_{w,k}$	$\mathbb{R}^{n_x \times n_w}$
C_k	$\mathbb{R}^{n_y \times n_x}$
$D_{v,k}$	$\mathbb{R}^{n_y \times n_v}$

Além disso, por hipótese, os vetores $\{w_k\}$, $\{v_k\}$ e x_0 são não-correlacionados, *i.e.*,

$$E \left\{ \begin{bmatrix} w_k \\ v_k \\ x_0 \end{bmatrix} \right\} = \begin{bmatrix} \bar{w}_k \\ \bar{v}_k \\ \bar{x}_0 \end{bmatrix}, \quad (2.8)$$

$$E \left\{ \begin{bmatrix} w_k - \bar{w}_k \\ v_k - \bar{v}_k \\ x_0 - \bar{x}_0 \end{bmatrix} \begin{bmatrix} w_j - \bar{w}_j \\ v_j - \bar{v}_j \\ x_0 - \bar{x}_0 \end{bmatrix}^T \right\} = \begin{bmatrix} W_k \delta_{kj} & 0 & 0 \\ 0 & V_k \delta_{kj} & 0 \\ 0 & 0 & X_0 \end{bmatrix}, \quad (2.9)$$

em que as dimensões e os significados das matrizes envolvidas são apresentados na Tabela 2.3 e δ_{kj} é a função delta de Kronecker, *i.e.*, $\delta_{kj} = 1$ se $k = j$ e $\delta_{kj} = 0$, caso contrário.

Tabela 2.3: Vetores e matrizes para a definição das covariâncias.

Matriz	Conjunto	Significado
\bar{w}_k	\mathbb{R}^{n_w}	Valor esperado do ruído de processo
\bar{v}_k	\mathbb{R}^{n_v}	Valor esperado do ruído de medida
\bar{x}_0	\mathbb{R}^{n_x}	Valor esperado do estado inicial
W_k	$\mathbb{R}^{n_w \times n_w}$	Covariância do ruído de processo
V_k	$\mathbb{R}^{n_v \times n_v}$	Covariância do ruído de medida
X_0	$\mathbb{R}^{n_x \times n_x}$	Covariância do estado inicial

O próximo passo é obter um modelo aumentado a partir das estruturas do sistema e do preditor. Uma das formas possíveis de se obter um modelo aumentado para o sistema é mostrada no Lema 2.3.1.

Lema 2.3.1 Considere um sistema dado por (2.6)-(2.7) e um estimador de estados com a estrutura (2.5), se for definido um vetor de estados aumentado da forma $\tilde{x}_k = \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix}$, então

$$\tilde{x}_{k+1} = \tilde{A}_k \tilde{x}_k + \tilde{B}_k w_k + \tilde{D}_k v_k, \quad (2.10)$$

em que

$$\tilde{A}_k = \begin{bmatrix} A_k & 0 \\ K_k C_k & \Phi_k - K_k C_k \end{bmatrix}, \tilde{B}_k = \begin{bmatrix} B_{w,k} \\ 0 \end{bmatrix}, \tilde{D}_k = \begin{bmatrix} 0 \\ K_k D_{v,k} \end{bmatrix}. \quad (2.11)$$

Prova Seja \tilde{x}_k o vetor de estado definido. Portanto, no instante $k + 1$,

$$\tilde{x}_{k+1} = \begin{bmatrix} x_{k+1} \\ \hat{x}_{k+1|k} \end{bmatrix}. \quad (2.12)$$

Substituindo (2.6) e (2.5) em (2.12), obtém-se (2.10). ■

Neste ponto, deseja-se obter a covariância do sistema aumentado (2.6). Conhecendo a variância do estado inicial, dada em (2.9), e considerando que \tilde{x}_k , w_k e v_k são não-correlacionados, ou seja,

$$E \left\{ \begin{bmatrix} \tilde{x}_k - E \{ \tilde{x}_k \} \\ w_k - \bar{w}_k \\ v_k - \bar{v}_k \end{bmatrix} \begin{bmatrix} \tilde{x}_j - E \{ \tilde{x}_j \} \\ w_j - \bar{w}_j \\ v_j - \bar{v}_j \end{bmatrix}^T \right\} = \begin{bmatrix} \tilde{P}_k & 0 & 0 \\ 0 & W_k & 0 \\ 0 & 0 & V_k \end{bmatrix} \delta_{kj}, \quad (2.13)$$

obtém-se, com o auxílio do Lema 2.3.2, a recursão para a matriz de covariância do vetor de estados aumentado.

Lema 2.3.2 Dado o sistema em (2.6), com propriedades estatísticas (2.13) e condições iniciais (2.4) e (2.8), então sua covariância é dada por

$$\tilde{P}_{0|-1} = \begin{bmatrix} X_0 & 0 \\ 0 & 0 \end{bmatrix} \quad (2.14)$$

$$\tilde{P}_{k+1|k} = \tilde{A}_k \tilde{P}_{k|k-1} \tilde{A}_k^T + \tilde{B}_k W_k \tilde{B}_k^T + \tilde{D}_k V_k \tilde{D}_k^T. \quad (2.15)$$

Prova A partir de (2.6) e utilizando (2.8), chega-se a

$$E \{ \tilde{x}_{k+1} \} = \tilde{A}_k E \{ \tilde{x}_k \} + \tilde{B}_k \bar{w}_k + \tilde{D}_k \bar{v}_k, \quad (2.16)$$

de forma que

$$\tilde{x}_{k+1} - E \{ \tilde{x}_{k+1} \} = \tilde{A}_k (\tilde{x}_k - E \{ \tilde{x}_k \}) + \tilde{B}_k (w_k - \bar{w}_k) + \tilde{D}_k (v_k - \bar{v}_k). \quad (2.17)$$

Portanto, utilizando (2.13), obtém-se (2.15). A condição inicial é obtida por meio da Nota 2.3.3, a seguir. Esta mesma nota também mostra uma incoerência no resultado apresentado por [82]. ■

Nota 2.3.3 De maneira genérica, $\tilde{x}_0 = \begin{bmatrix} x_0 \\ \hat{x}_{0|-1} \end{bmatrix}$, de forma que

$$E \{ \tilde{x}_0 \} = \begin{bmatrix} E \{ x_0 \} \\ E \{ \hat{x}_{0|-1} \} \end{bmatrix} = \begin{bmatrix} \bar{x}_0 \\ E \{ \hat{x}_{0|-1} \} \end{bmatrix}. \quad (2.18)$$

Portanto,

$$\begin{aligned} \tilde{P}_{0|-1} &:= E \{ \{ \tilde{x}_0 - E \{ \tilde{x}_0 \} \} \{ \tilde{x}_0 - E \{ \tilde{x}_0 \} \}^T \} \\ &= \begin{bmatrix} E \{ (x_0 - \bar{x}_0) (x_0 - \bar{x}_0)^T \} & E \{ (x_0 - \bar{x}_0) (\hat{x}_{0|-1} - E \{ \hat{x}_{0|-1} \})^T \} \\ E \{ (\hat{x}_{0|-1} - E \{ \hat{x}_{0|-1} \}) (x_0 - \bar{x}_0)^T \} & E \{ (\hat{x}_{0|-1} - E \{ \hat{x}_{0|-1} \}) (\hat{x}_{0|-1} - E \{ \hat{x}_{0|-1} \})^T \} \end{bmatrix}. \end{aligned} \quad (2.19)$$

A condição inicial do preditor é um valor determinístico dado por $\hat{x}_{0|-1} = \bar{x}_0$, logo $E \{ \hat{x}_{0|-1} \} = \bar{x}_0$. Desta maneira, utilizando a covariância do estado inicial dada em (2.8), o valor inicial da matriz de covariância em (2.19) torna-se

$$\tilde{P}_{0|-1} = \begin{bmatrix} E \{ (x_0 - \bar{x}_0) (x_0 - \bar{x}_0)^T \} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} X_0 & 0 \\ 0 & 0 \end{bmatrix}, \forall \bar{x}_0. \quad (2.20)$$

Em [82], apesar de o problema ser explicitamente estabelecido com \bar{x}_0 genérico, de fato, ao longo do artigo, assume-se $\bar{x}_0 = 0$, uma vez que o autor deixa explícito que

$$\tilde{P}_{0|-1} = \begin{bmatrix} E \{ x_0 x_0^T \} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} X_0 & 0 \\ 0 & 0 \end{bmatrix}, \forall \bar{x}_0. \quad (2.21)$$

■

O Lema 2.3.4 mostra que a matriz $\tilde{P}_{k+1|k}$, determinada no lema 2.3.2, pode ser particionada em uma matrix-bloco. Esta partição da matriz $\tilde{P}_{k+1|k}$ é interessante no sentido de se obter a covariância do erro de estimação a partir da covariância do sistema aumentado.

Lema 2.3.4 Considere a matriz de covariância dada em (2.15) com as identificações mostradas em (2.11). Então $\tilde{P}_{k+1|k}$ pode ser particionada como em (2.22) com as identificações (2.23).

$$\tilde{P}_{k+1|k} = \begin{bmatrix} P_{11,k+1|k} & P_{12,k+1|k} \\ P_{12,k+1|k}^T & P_{22,k+1|k} \end{bmatrix} \geq 0, \quad (2.22)$$

em que

$$P_{11,k+1|k} = A_k P_{11,k|k-1} A_k^T + \Psi_{1,k}, \quad (2.23)$$

$$P_{12,k+1|k} = (A_k S_{1,k} C_k^T + \Psi_{2,k}) K_k^T + A_k P_{12,k|k-1} \Phi_k^T, \quad (2.24)$$

$$\begin{aligned} P_{22,k+1|k} &= K_k (C_k S_{3,k} C_k^T + \Psi_{3,k}) K_k^T + K_k C_k S_{2,k} \Phi_k^T \\ &+ \Phi_k S_{2,k}^T C_k^T K_k^T + \Phi_k P_{22,k|k-1} \Phi_k^T, \end{aligned} \quad (2.25)$$

com as definições

$$S_{1,k} := P_{11,k|k-1} - P_{12,k|k-1}, \quad S_{2,k} := P_{12,k|k-1} - P_{22,k|k-1}, \quad S_{3,k} := S_{1,k} - S_{2,k}^T \quad (2.26)$$

e as identificações ¹

$$\Psi_{1,k} := B_{w,k} W_k B_{w,k}^T, \quad \Psi_{2,k} := 0, \quad \Psi_{3,k} := D_{v,k} V_k D_{v,k}^T. \quad (2.27)$$

Prova Efetuando as multiplicações dos termos da matriz de covariância aumentada (2.15), chega-se a

$$\begin{aligned} \tilde{A}_k \tilde{P}_{k|k-1} \tilde{A}_k^T &= \begin{bmatrix} L_{11,k} & L_{12,k} \\ L_{21,k} & L_{22,k} \end{bmatrix}, \quad \tilde{B}_k W_k \tilde{B}_k^T = \begin{bmatrix} \Psi_{1,k} & 0 \\ 0 & 0 \end{bmatrix}, \\ \tilde{D}_k V_k \tilde{D}_k^T &= \begin{bmatrix} 0 & 0 \\ 0 & K_k \Psi_{3,k} K_k^T \end{bmatrix}, \end{aligned} \quad (2.28)$$

em que

$$L_{11,k} = A_k P_{11,k|k-1} A_k^T, \quad (2.29)$$

$$L_{12,k} = A_k P_{11,k|k-1} C_k^T K_k^T + A_k P_{12,k|k-1} (\Phi_k - K_k C_k)^T \quad (2.30)$$

$$L_{21,k} = K_k C_k P_{11,k|k-1} A_k^T + (\Phi_k - K_k C_k) P_{21,k|k-1} A_k^T \quad (2.31)$$

$$\begin{aligned} L_{22,k} &= K_k C_k P_{11,k|k-1} C_k^T K_k^T + K_k C_k P_{12,k|k-1} (\Phi_k - K_k C_k)^T \\ &+ (\Phi_k - K_k C_k) P_{21,k|k-1} C_k^T K_k^T + (\Phi_k - K_k C_k) P_{22,k|k-1} (\Phi_k - K_k C_k)^T. \end{aligned} \quad (2.32)$$

Realizando a soma das matrizes (2.28), verifica-se que

$$P_{11,k+1|k} = A_k P_{11,k+1|k} A_k^T + \Psi_{1,k}, \quad (2.33)$$

$$P_{12,k+1|k} = A_k P_{11,k|k-1} C_k^T K_k^T + A_k P_{12,k|k-1} (\Phi_k - K_k C_k)^T, \quad (2.34)$$

$$P_{21,k+1|k} = K_k C_k P_{11,k|k-1} A_k^T + (\Phi_k - K_k C_k) P_{21,k|k-1} A_k^T, \quad (2.35)$$

$$\begin{aligned} P_{22,k+1|k} &= K_k C_k P_{11,k|k-1} C_k^T K_k^T + K_k C_k P_{12,k|k-1} (\Phi_k - K_k C_k)^T \\ &+ (\Phi_k - K_k C_k) P_{21,k|k-1} C_k^T K_k^T \\ &+ (\Phi_k - K_k C_k) P_{22,k|k-1} (\Phi_k - K_k C_k)^T + K_k \Psi_{3,k} K_k^T. \end{aligned} \quad (2.36)$$

Nota-se, portanto, que $P_{21,k|k-1} = P_{12,k|k-1}^T$. Simplificando a expressão para $P_{12,k+1|k}$,

$$P_{12,k+1|k} = A_k P_{12,k|k-1} \Phi_k^T + A_k S_{1,k} C_k^T K_k^T. \quad (2.37)$$

Simplificando a expressão para $P_{22,k+1|k}$,

$$\begin{aligned} P_{22,k+1|k} &= K_k C_k (P_{11,k|k-1} - P_{12,k|k-1} - P_{21,k|k-1} + P_{22,k|k-1}) C_k^T K_k^T \\ &+ K_k C_k (P_{12,k|k-1} - P_{22,k|k-1}) \Phi_k^T + \Phi_k (P_{21,k|k-1} - P_{22,k|k-1}) C_k^T K_k^T \\ &+ \Phi_k P_{22,k|k-1} \Phi_k^T + K_k \Psi_{3,k} K_k^T. \end{aligned} \quad (2.38)$$

Lembrando que $P_{21,k|k-1} = P_{12,k|k-1}^T$, pode-se simplificar a expressão para

$$P_{22,k+1|k} = \Phi_k P_{22,k|k-1} \Phi_k^T + K_k C_k S_{2,k} \Phi_k^T + \Phi_k S_{2,k}^T C_k^T K_k^T + K_k (C_k S_{3,k} C_k^T + \Psi_{3,k}) K_k^T. \quad \blacksquare$$

¹Vale ressaltar que estas identificações estão sendo feitas para este modelo, e o resultado do lema é válido para matrizes $\Psi_{1,k}$, $\Psi_{2,k}$ e $\Psi_{3,k}$ genéricas.

Nota 2.3.5 A expressão (2.23) fornece a evolução da matriz de covariância dos estados. ■

A seguir, o Lema 2.3.6 mostra as relações existentes entre a matriz de covariância do erro de estimação e a matriz de covariância do sistema aumentado. A partir destas informações, o lema 2.3.7 fornece uma expressão para a matriz de covariância do erro de estimação, a saber

$$\begin{aligned}\bar{P}_{k+1|k} &= (A_k - K_k C_k) P_{11,k|k-1} (A_k - K_k C_k)^T - (A_k - K_k C_k) P_{12,k|k-1} (\Phi_k - K_k C_k)^T \\ &\quad - (\Phi_k - K_k C_k) P_{12,k|k-1}^T (A_k - K_k C_k)^T + (\Phi_k - K_k C_k) P_{22,k|k-1} (\Phi_k - K_k C_k)^T \\ &\quad + B_{w,k} W_k B_{w,k}^T + K_k D_{v,k} V_k D_{v,k}^T K_k^T.\end{aligned}\quad (2.39)$$

Lema 2.3.6 Partindo das definições do estado aumentado \tilde{x}_k e de sua matriz de covariância, $\tilde{P}_{k+1|k}$, então as seguintes relações são válidas para a covariância do erro de estimação:

$$\bar{P}_{k+1|k} = \begin{bmatrix} I & -I \end{bmatrix} \tilde{P}_{k+1|k} \begin{bmatrix} I & -I \end{bmatrix}^T \quad (2.40)$$

$$= P_{11,k+1|k} - P_{12,k+1|k} - P_{12,k+1|k}^T + P_{22,k+1|k}, \quad (2.41)$$

com a condição inicial $\bar{P}_{0|-1} = X_0$.

Prova Partindo da definição de covariância e da partição sugerida em (2.22), pode-se mostrar que

$$\tilde{P}_{k+1|k} := E \left\{ \begin{bmatrix} x_{k+1} - E\{x_{k+1}\} \\ \hat{x}_{k+1|k} - E\{\hat{x}_{k+1|k}\} \end{bmatrix} \begin{bmatrix} x_{k+1} - E\{x_{k+1}\} \\ \hat{x}_{k+1|k} - E\{\hat{x}_{k+1|k}\} \end{bmatrix}^T \right\} = \begin{bmatrix} P_{11,k+1|k} & P_{12,k+1|k} \\ P_{12,k+1|k}^T & P_{22,k+1|k} \end{bmatrix} \quad (2.42)$$

Multiplicando as expressões mostradas em (2.42) por matrizes constantes $\begin{bmatrix} I & -I \end{bmatrix}$ e $\begin{bmatrix} I & -I \end{bmatrix}^T$, encontra-se

$$\begin{aligned}\bar{P}_{k+1|k} &:= E \left\{ (x_{k+1} - \hat{x}_{k+1|k} - E\{x_k - \hat{x}_{k+1|k}\}) (x_{k+1} - \hat{x}_{k+1|k} - E\{x_k - \hat{x}_{k+1|k}\})^T \right\} \\ &= P_{11,k+1|k} - P_{12,k+1|k} - P_{12,k+1|k}^T + P_{22,k+1|k}.\end{aligned}\quad (2.43)$$

Lema 2.3.7 A expressão (2.41), com as identificações (2.23)-(2.25), também pode ser escrita como

$$\begin{aligned}\bar{P}_{k+1|k} &= (A_k - K_k C_k) P_{11,k|k-1} (A_k - K_k C_k)^T - (A_k - K_k C_k) P_{12,k|k-1} (\Phi_k - K_k C_k)^T \\ &\quad - (\Phi_k - K_k C_k) P_{12,k|k-1}^T (A_k - K_k C_k)^T + (\Phi_k - K_k C_k) P_{22,k|k-1} (\Phi_k - K_k C_k)^T \\ &\quad + \Psi_{1,k} - \Psi_{2,k} K_k^T - K_k \Psi_{2,k}^T + K_k \Psi_{3,k} K_k^T.\end{aligned}\quad (2.44)$$

Prova Por substituição direta de (2.23)-(2.25) em (2.41). ■

A matriz de covariância do erro de estimação apresentada em (2.39) é válida para quaisquer parâmetros Φ_k e K_k do preditor. Neste sentido, estas matrizes podem ser escolhidas de maneira a minimizar a covariância do erro de estimação $\bar{P}_{k+1|k}$. Esta minimização será realizada calculando-se as derivadas parciais de $\bar{P}_{k+1|k}$ com respeito a Φ_k e K_k e fazendo

$$\frac{\partial}{\partial \Phi_k} \bar{P}_{k+1|k} = 0 \quad (2.45)$$

$$\frac{\partial}{\partial K_k} \bar{P}_{k+1|k} = 0. \quad (2.46)$$

Os lemas 2.3.8 e 2.3.11 fornecem os resultados destas diferenciações. A prova de que se trata de um ponto de mínimo global pode ser demonstrado por meio da convexidade de (2.39), uma vez que $\tilde{P}_{k+1|k} > 0$, conforme (2.22), $W_k > 0$ e $V_k > 0, \forall k$. Assim, definindo $S_k := P_{11,k|k-1} - P_{12,k|k-1}P_{22,k|k-1}^\dagger P_{12,k|k-1}^T$, conclui-se que

$$\Phi_k^* = A_k P_{12,k|k-1} P_{22,k|k-1}^\dagger - K_k^* C_k S_{2,k} P_{22,k|k-1}^\dagger, \quad (2.47)$$

$$K_k^* = A_k S_k C_k^T (C_k S_k C_k^T + D_{v,k} V_k D_{v,k}^T)^{-1}. \quad (2.48)$$

Utilizando a matriz hessiana de $\bar{P}_{k+1|k}$, mostrada a seguir, os resultados dos lemas 2.3.8 e 2.3.11, conclui-se que, de fato, trata-se de um ponto de mínimo.

$$He(\bar{P}_{k+1|k}) := \begin{bmatrix} \frac{\partial^2}{\partial^2 \Phi_k} \bar{P}_{k+1|k} & \frac{\partial^2}{\partial^2 [\Phi_k, K_k]} \bar{P}_{k+1|k} \\ \frac{\partial^2}{\partial^2 [K_k, \Phi_k]} \bar{P}_{k+1|k} & \frac{\partial^2}{\partial^2 K_k} \bar{P}_{k+1|k} \end{bmatrix} = \begin{bmatrix} 2P_{22,k|k-1} & 2C_k S_{2,k} \\ 2S_{2,k}^T C_k^T & C_k S_k C_k^T + \Psi_{3,k} \end{bmatrix} > 0.$$

Lema 2.3.8 *Considere novamente a expressão (2.41) com as identificações apresentadas em (2.23)-(2.25). Portanto, a expressão de Φ_k que minimiza (2.41) é dada por*

$$\Phi_k^* = A_k P_{12,k|k-1} P_{22,k|k-1}^\dagger - K_k C_k S_{2,k} P_{22,k|k-1}^\dagger, \quad (2.49)$$

$$= A_k + (A_k - K_k C_k) \left(P_{12,k|k-1} P_{22,k|k-1}^\dagger - I \right). \quad (2.50)$$

Prova O valor de Φ_k que minimiza (2.41) pode ser obtido por meio da diferenciação termo a termo de (2.41), ou seja,

$$\frac{\partial}{\partial \Phi_k} \bar{P}_{k+1|k} = \frac{\partial}{\partial \Phi_k} P_{11,k+1|k} - \frac{\partial}{\partial \Phi_k} P_{12,k+1|k} - \frac{\partial}{\partial \Phi_k} P_{21,k+1|k} + \frac{\partial}{\partial \Phi_k} P_{22,k+1|k}. \quad (2.51)$$

Utilizando as regras do cálculo matricial², tem-se que

$$\frac{\partial}{\partial \Phi_k} P_{11,k+1|k} = 0. \quad (2.52)$$

$$\frac{\partial}{\partial \Phi_k} P_{12,k+1|k} = A_k P_{12,k|k-1}. \quad (2.53)$$

$$\frac{\partial}{\partial \Phi_k} P_{21,k+1|k} = \frac{\partial}{\partial \Phi_k} P_{12,k+1|k}^T = A_k P_{12,k|k-1}. \quad (2.54)$$

$$\frac{\partial}{\partial \Phi_k} P_{22,k+1|k} = 2K_k C_k S_{2,k} + 2\Phi_k P_{22,k|k-1}. \quad (2.55)$$

Substituindo (2.52) a (2.55) em (2.41),

$$\frac{\partial}{\partial \Phi_k} \bar{P}_{k+1|k} = -2A_k P_{12,k|k-1} + 2K_k C_k S_{2,k} + 2\Phi_k P_{22,k|k-1}. \quad (2.56)$$

Para encontrar o ponto de mínimo, faz-se (2.56) igual a zero, ou seja,

$$\frac{\partial}{\partial \Phi_k} \bar{P}_{k+1|k} = 0 \Rightarrow \Phi_k^* P_{22,k|k-1} = A_k P_{12,k|k-1} - K_k C_k S_{2,k}. \quad (2.57)$$

Destarte, o valor de Φ_k que minimiza (2.41) deve satisfazer à equação (2.57). Entretanto, a expressão para Φ_k não pode ser obtida diretamente, pois exige que $P_{22,k|k-1}^{-1}$ exista para todo instante k . Olhando as

²Consultar Anexo I.5

condições iniciais $P_{12,0|-1} = P_{12,0|-1}^T = P_{22,0|-1} = 0$, conclui-se que $P_{22,0} = 0$ e, portanto, sua matriz inversa não existe para todo instante k . De fato, pode-se mostrar que existe uma matriz pseudo-inversa³ para todo instante k . Neste caso, uma expressão para Φ_k é dada por:

$$\Phi_k^* = A_k P_{12,k|k-1} P_{22,k|k-1}^\dagger - K_k C_k S_{2,k} P_{22,k|k-1}^\dagger, \quad (2.58)$$

ou, olhando a Nota 2.3.9, obtém-se a expressão mostrada em (2.50).

Calculando a segunda derivada de (2.56) e utilizando o resultado da Nota 2.3.10, obtém-se que

$$\frac{\partial^2}{\partial^2 \Phi_k} \bar{P}_{k+1|k} = 2P_{22,k|k-1}^T = 2P_{22,k|k-1} \geq 0. \quad (2.59)$$

Portanto, Φ_k^* trata-se de um ponto de mínimo. ■

Nota 2.3.9 Sabendo que $S_{2,k} = P_{12,k|k-1} - P_{22,k|k-1}$, então

$$\Phi_k^* P_{22,k|k-1} = A_k P_{12,k|k-1} - K_k C_k S_{2,k} \Rightarrow \Phi_k^* = A_k + (A_k - K_k C_k) \left(P_{12,k|k-1} P_{22,k|k-1}^\dagger - I \right) \quad (2.60)$$
■

Nota 2.3.10 Como $P_{22,k|k-1}$ é um bloco da diagonal de uma matriz de covariância, ela satisfaz, por definição, $P_{22,k|k-1} = P_{22,k|k-1}^T \geq 0$. ■

Lema 2.3.11 Considere a expressão (2.41) e $P_{11,k+1|k}$ definida em (2.61). Ainda,

$$P_{12,k+1|k} = (A_k S_{1,k} C_k^T + \Psi_{2,k}) K_k^T + A_k P_{12,k|k-1} (\Phi_k^*)^T, \quad (2.61)$$

$$\begin{aligned} P_{22,k+1|k} &= K_k (C_k S_{3,k} C_k^T + \Psi_{3,k}) K_k^T + K_k C_k S_{2,k} (\Phi_k^*)^T + \Phi_k^* S_{2,k}^T C_k^T K_k^T \\ &+ \Phi_k^* P_{22,k|k-1} (\Phi_k^*)^T, \end{aligned} \quad (2.62)$$

em que Φ_k^* satisfaz (2.58). Portanto, a expressão de K_k que minimiza (2.41) é dada por

$$K_k^* = A_k S_k C_k^T (C_k S_k C_k^T + D_{v,k} V_k D_{v,k}^T)^{-1}, \quad (2.63)$$

em que $S_k := P_{11,k|k-1} - P_{12,k|k-1} P_{22,k|k-1}^\dagger P_{12,k|k-1}^T$.

Prova O valor de K_k que minimiza (2.41) pode ser obtido por meio da diferenciação de cada termo:

$$\frac{\partial}{\partial K_k} \bar{P}_{k+1|k} = \frac{\partial}{\partial K_k} P_{11,k+1|k} - \frac{\partial}{\partial K_k} P_{12,k+1|k} - \frac{\partial}{\partial K_k} P_{21,k+1|k} + \frac{\partial}{\partial K_k} P_{22,k+1|k}. \quad (2.64)$$

Utilizando as regras do cálculo matricial⁴, tem-se que

$$\frac{\partial}{\partial K_k} P_{11,k+1|k} = 0, \quad (2.65)$$

$$\frac{\partial}{\partial K_k} P_{12,k+1|k} = A_k S_{1,k} C_k^T + \Psi_{2,k}, \quad (2.66)$$

$$\frac{\partial}{\partial K_k} P_{21,k+1|k} = A_k S_{1,k} C_k^T + \Psi_{2,k}, \quad (2.67)$$

$$\frac{\partial}{\partial K_k} P_{22,k+1|k} = 2K_k (C_k S_{3,k} C_k^T + \Psi_{3,k}) + 2\Phi_k^* S_{2,k}^T C_k^T. \quad (2.68)$$

³ Consultar o Anexo I.2 para maiores detalhes sobre matrizes pseudo-inversa e inversas generalizadas

⁴ Consultar Anexo I.5

Substituindo (2.65) a (2.68) em (2.41),

$$\frac{\partial}{\partial K_k} \bar{P}_{k+1|k} = -2 (A_k S_{1,k} C_k^T + \Psi_{2,k}) + 2K_k (C_k S_{3,k} C_k^T + \Psi_{3,k}) + 2\Phi_k^* S_{2,k}^T C_k^T. \quad (2.69)$$

Para encontrar o ponto de mínimo, faz-se (2.69) igual a zero, ou seja,

$$\frac{\partial}{\partial K_k} \bar{P}_{k+1|k} = 0 \Rightarrow K_k^* (C_k S_{3,k} C_k^T + \Psi_{3,k}) = (A_k S_{1,k} C_k^T + \Psi_{2,k}) - \Phi_k^* S_{2,k}^T C_k^T. \quad (2.70)$$

Substituindo o valor de Φ_k^* em (2.70), obtém-se

$$K_k^* \left(C_k \left(S_{3,k} - S_{2,k} P_{22,k|k-1}^\dagger S_{2,k}^T \right) C_k^T + \Psi_{3,k} \right) = A_k \left(S_{1,k} - P_{12,k|k-1} P_{22,k|k-1}^\dagger S_{2,k}^T \right) C_k^T + \Psi_{2,k} \quad (2.71)$$

Logo, observando o resultado da Nota 2.3.12, tem-se o resultado do lema. Calculando a segunda derivada da expressão (2.71), nota-se que K_k^* é um ponto de mínimo:

$$\frac{\partial^2}{\partial^2 K_k} \bar{P}_{k+1|k} = C_k S_k C_k^T + \Psi_{3,k} > 0. \quad (2.72)$$

■

Nota 2.3.12 *Sejam as expressões*

$$f_k := S_{3,k} - S_{2,k} P_{22,k|k-1}^\dagger S_{2,k}^T, \quad (2.73)$$

$$g_k := S_{1,k} - P_{12,k|k-1} P_{22,k|k-1}^\dagger P_{12,k|k-1}^T, \quad (2.74)$$

em que $S_{1,k}$, $S_{2,k}$, $S_{3,k}$ foram definidos em (2.26). Portanto,

$$f_k = S_{3,k} - S_{2,k} P_{22,k|k-1}^\dagger S_{2,k}^T = g_k = P_{11,k|k-1} - P_{12,k|k-1} P_{22,k|k-1}^\dagger P_{12,k|k-1}^T. \quad (2.75)$$

■

2.3.1 Buscando Simplificações

Neste ponto, sabe-se que os parâmetros ótimos do preditor devem ser escolhidos conforme (2.47) e (2.48). A partir destas escolhas, é possível simplificar as expressões dos parâmetros do preditor e a recursão da covariância. Por meio dos resultados dos lemas 2.3.13 e 2.3.14, mostra-se que

$$P_{12,k+1|k} = P_{12,k+1|k}^T = P_{22,k+1|k} \quad (2.76)$$

$$\begin{aligned} &= (A_k S_k C_k^T + \Psi_{2,k}) (C_k S_k C_k^T + \Psi_{3,k})^{-1} (A_k S_k C_k^T + \Psi_{2,k})^T \\ &+ A_k P_{12,k|k-1} P_{22,k|k-1}^\dagger P_{12,k|k-1}^T A_k^T. \end{aligned} \quad (2.77)$$

Como $P_{12,k+1|k} = P_{12,k+1|k}^T = P_{22,k+1|k}$, para qualquer $P_{k|k-1}$ simétrica, começando-se com uma matriz $P_{n|n-1}$ satisfazendo $P_{12,n|n-1} = P_{12,n|n-1}^T = P_{22,n|n-1}$ para algum $n \geq 0$, pode-se concluir que $P_{12,k+1|k} = P_{12,k+1|k}^T = P_{22,k+1|k}$ é válido para qualquer $k \geq n$. Utilizando este fato, as expressões para Φ_k^* , K_k^* e $\bar{P}_{k+1|k}$ podem ser simplificadas. Como consequência de (2.76), nota-se que

$$\bar{P}_{k+1|k} := P_{11,k|k-1} - P_{12,k|k-1} - P_{12,k|k-1}^T + P_{22,k|k-1} = P_{11,k|k-1} - P_{12,k|k-1}^T. \quad (2.78)$$

Logo,

$$S_k := P_{11,k|k-1} - P_{12,k|k-1} P_{22,k|k-1}^\dagger P_{12,k|k-1}^T = P_{11,k|k-1} - P_{12,k|k-1}^T = \bar{P}_{k+1|k}. \quad (2.79)$$

Lema 2.3.13 Considere a expressão de $P_{12,k+1|k}$ dada em (2.24). Se $P_{21,k+1|k} = P_{12,k+1|k}^T$, $\Psi_{3,k} = \Psi_{3,k}^T$ e Φ_k^* e K_k^* satisfazem, respectivamente, (2.47)-(2.48), então

$$P_{12,k+1|k} = P_{12,k+1|k}^T = P_{21,k+1|k}. \quad (2.80)$$

Prova Substituindo a expressão (2.47) em $P_{12,k+1|k}$, dada em (2.24),

$$P_{12,k+1|k} = \left(A_k \left(S_{1,k} - P_{12,k|k-1} P_{22,k|k-1}^\dagger S_{2,k}^T \right) C_k^T + \Psi_{2,k} \right) K_k^T + A_k P_{12,k|k-1} P_{22,k|k-1}^\dagger P_{12,k|k-1}^T A_k^T. \quad (2.81)$$

Utilizando o resultado da Nota 2.3.12, conclui-se que

$$P_{12,k+1|k} = \left(A_k S_k C_k^T + \Psi_{2,k} \right) K_k^T + A_k P_{12,k|k-1} P_{22,k|k-1}^\dagger P_{12,k|k-1}^T A_k^T. \quad (2.82)$$

Diretamente, observa-se que

$$K_k^T = \left(C_k S_k C_k^T + \Psi_{3,k} \right)^{-1} \left(A_k S_k C_k^T + \Psi_{2,k} \right)^T. \quad (2.83)$$

Substituindo (2.83) em (2.82), chega-se à expressão para $P_{12,k+1|k}$ mostrada em (2.77). Dado que $P_{21,k+1|k} = P_{12,k+1|k}^T$, é possível concluir que (2.80) é válida. ■

Lema 2.3.14 Considere as expressões de $P_{12,k+1|k}$ e $P_{22,k+1|k}$ apresentadas respectivamente em (2.77) e (2.25). Se Φ_k^* e K_k^* satisfazem, respectivamente, (2.47)-(2.48), então $P_{22,k+1|k} = P_{12,k+1|k}$.

Prova Substituindo a expressão (2.47) em (2.25) e utilizando as propriedades da pseudo-inversa⁵, tem-se que

$$P_{22,k+1|k} = K_k \left(C_k \left(S_{3,k} - S_{2,k} P_{22,k|k-1}^\dagger S_{2,k}^T \right) C_k^T + \Psi_{3,k} \right) K_k^T + A_k P_{12,k|k-1} P_{22,k|k-1}^\dagger P_{12,k|k-1}^T A_k^T. \quad (2.84)$$

Lembrando que $S_k := S_{3,k} - S_{2,k} P_{22,k|k-1}^\dagger S_{2,k}^T$, então

$$P_{22,k+1|k} = K_k \left(C_k S_k C_k^T + \Psi_{3,k} \right) K_k^T + A_k P_{12,k|k-1} P_{22,k|k-1}^\dagger P_{12,k|k-1}^T A_k^T. \quad (2.85)$$

Substituindo (2.48) e sua matriz transposta em (2.85), obtém-se

$$P_{22,k+1|k} = \left(A_k S_k C_k^T + \Psi_{2,k} \right) \left(C_k S_k C_k^T + \Psi_{3,k} \right)^{-1} \left(A_k S_k C_k^T + \Psi_{2,k} \right)^T + A_k P_{12,k|k-1} P_{22,k|k-1}^{-1} P_{12,k|k-1}^T A_k^T. \quad (2.86)$$

Dada a expressão (2.77), então $P_{22,k+1|k} = P_{12,k+1|k}$. ■

Por fim, a partir do fato de que $P_{12,k+1|k} = P_{12,k+1|k}^T = P_{22,k+1|k}$, pode-se simplificar as expressões para Φ_k^* , K_k^* e $\bar{P}_{k+1|k}$. Observando a expressão de Φ_k^* mostrada em (2.50), segue diretamente que $\Phi_k^* = A_k$. Para a expressão simplificada do ganho, parte-se de (2.48) e se utiliza (2.79) para se obter

$$K_k^* = A_k \bar{P}_{k|k-1} C_k^T \left(C_k \bar{P}_{k|k-1} C_k^T + D_{v,k} V_k D_{v,k}^T \right)^{-1}. \quad (2.87)$$

Para a equação de Riccati, substitui-se $\Phi_k^* = A_k$ em (2.39) para se obter

$$\bar{P}_{k+1|k} = \left(A_k - K_k^* C_k \right) \bar{P}_{k|k-1} \left(A_k - K_k^* C_k \right)^T + B_{w,k} W_k B_{w,k}^T + K_k^* D_{v,k} V_k D_{v,k}^T \left(K_k^* \right)^T. \quad (2.88)$$

⁵Consultar anexo I.1.

Por fim, a expressão mínima para a covariância do erro de estimação é obtida diretamente após a substituição de (2.87) em (2.88), resultando em

$$\begin{aligned}\bar{P}_{k+1|k} &= A_k \bar{P}_{k|k-1} A_k^T + B_{w,k} W_k B_{w,k}^T \\ &- A_k \bar{P}_{k|k-1} C_k^T (C_k \bar{P}_{k|k-1} C_k^T + D_{v,k} V_k D_{v,k}^T)^{-1} (A_k \bar{P}_{k|k-1} C_k^T)^T.\end{aligned}\quad (2.89)$$

Neste ponto, já se tem o algoritmo do FK na forma preditora de 1 estágio. Este algoritmo é apresentado na Tabela 2.4. O FK (formas preditora e preditora-corretora) para sistemas de tempo discreto pode também ser encontrado nas referências [62], [102], [103], [64] e [104].

Tabela 2.4: Algoritmo do FK na sua forma preditora de 1 estágio.

<p><i>Passo 0 (Condições Iniciais):</i></p> $\hat{x}_{0 -1} = \bar{x}_0,$ $\bar{P}_{0 -1} = X_0.$ <p><i>Passo 1: Calcular os parâmetros do preditor</i></p> $\Phi_k = A_k,$ $K_k = A_k \bar{P}_{k k-1} C_k^T (C_k \bar{P}_{k k-1} C_k^T + D_{v,k} V_k D_{v,k}^T)^{-1}.$ <p><i>Passo 2: Atualizar $\{\hat{x}_{k+1 k}, \bar{P}_{k+1 k}\}$ como</i></p> $\hat{x}_{k+1 k} = \Phi_k \hat{x}_{k k-1} + K_k (y_k - C_k \hat{x}_{k k-1}),$ $\bar{P}_{k+1 k} = A_k \bar{P}_{k k-1} A_k^T + B_{w,k} W_k B_{w,k}^T \\ - A_k \bar{P}_{k k-1} C_k^T (C_k \bar{P}_{k k-1} C_k^T + D_{v,k} V_k D_{v,k}^T)^{-1} (A_k \bar{P}_{k k-1} C_k^T)^T.$
--

2.4 ESTIMAÇÃO NÃO-LINEAR CLÁSSICA

Os estimadores de Kalman são bastante utilizados em aplicações reais, incluindo robótica, sistemas de comunicação, GPS, navegação inercial, controle de plantas químicas, predição do tempo, fusão de dados multi-sensor, rastreamento de aeronaves, satélites, navios, mísseis, carros, pessoas, gado, salmão, bem como aplicações mais recentes, como a tentativa de predição do mercado de ações. Além do mais, são relativamente fáceis de se projetar e codificar e, geralmente, fornecem uma boa estimativa. Por outro lado, por diversas razões, seu desempenho pode ser bastante ruim em algumas aplicações. Segundo [105], alguns fatores que conhecidamente prejudicam o desempenho do estimador são

- Não-linearidades nas equações que descrevem o modelo físico.
- Mal condicionamento numérico da matriz de covariância.
- Modelos imprecisos e/ou incompletos dos sistemas físicos.

O preditor de Kalman foi proposto há quase cinco décadas, consolidando-se como uma ferramenta eficiente para estimação, mas limitada para os sistemas lineares. Infelizmente, sistemas lineares não existem

na prática - de fato, todos os sistemas reais são não-lineares. Entretanto, muitos sistemas possuem comportamento aproximadamente linear, de maneira que as abordagens lineares fornecem bons resultados. Por outro lado, muitas vezes é difícil identificar quão linear é o comportamento de um sistema. Outra situação ocorre para sistemas que são "aproximadamente" lineares em uma pequena faixa de operação e não-lineares em uma faixa maior. A partir desta motivação, originou-se o estudo dos filtros não-lineares.

Na prática, a principal questão dos estimadores não-lineares é a sua complexidade computacional em tempo real para se atingir uma estimativa com uma precisão satisfatória. A capacidade de processamento e de armazenamento dos computadores cresceu enormemente quando comparada com os computadores disponíveis na década de 60, época em que Kalman publicou seu famoso artigo [63]. A partir deste desenvolvimento da informática, atualmente já é possível executar estimadores não-lineares em algumas aplicações em tempo real. A capacidade de processamento em tempo real depende de vários fatores, como, por exemplo,

- Dimensão n do vetor de estados a ser estimado.
- Presença de uma matriz de covariância esparsa ou o acoplamento distinto entre componentes do vetor de estados.
- Taxa de aquisição das observações das medidas.
- Precisão desejada para a estimativa do vetor de estados.
- Grau de não-linearidade do modelo.
- Forma das densidades de probabilidades, *e.g.*, unimodal ou multimodal.

Em alguns problemas de baixa dimensão, já é factível executar uma aproximação numérica do filtro não-linear ótimo em tempo real em um único PC, dependendo das condições previamente mencionadas [105]. Todavia, para muitas outras aplicações, este cálculo ainda não é possível ou o seu custo é muito alto. Neste contexto, é que se buscam soluções mais interessantes para o problema da estimação não-linear.

2.4.1 O Filtro de Kalman Estendido

Seja um sistema discreto não-linear descrito por

$$x_{k+1} = f_k(x_k, u_k) + B_{w,k}w_k, \quad (2.90)$$

$$y_k = h_k(x_k) + D_{v,k}v_k, \quad (2.91)$$

em que os conjuntos e os significados de cada vetor são mostrados na Tabela 2.1. Admite-se que as funções $f_k(\bullet)$ e $h_k(\bullet)$ sejam não-lineares e suficientemente suaves, podendo ser expandidas em séries de Taylor em torno das estimativas $\hat{x}_{k|k}$ e $\hat{x}_{k|k-1}$, respectivamente. Desta forma,

$$f_k(x_k, u_k) = f_k(\hat{x}_{k|k}, u_k) + A_k(x_k - \hat{x}_{k|k}) + \dots, \quad (2.92)$$

$$h_k(x_k) = h_k(\hat{x}_{k|k-1}) + C_k(x_k - \hat{x}_{k|k-1}) + \dots, \quad (2.93)$$

em que

$$A_k = \left. \frac{\partial}{\partial x_k} f_k(x_k, u_k) \right|_{x_k = \hat{x}_{k|k}}, \quad (2.94)$$

$$C_k = \left. \frac{\partial}{\partial x_k} h_k(x_k) \right|_{x_k = \hat{x}_{k|k-1}}. \quad (2.95)$$

Desprezando os termos de ordem mais alta, obtém-se as seguintes aproximações

$$f_k(x_k, u_k) \approx f_k(\hat{x}_{k|k}, u_k) + A_k(x_k - \hat{x}_{k|k}), \quad (2.96)$$

$$h_k(x_k) \approx h_k(\hat{x}_{k|k-1}) + C_k(x_k - \hat{x}_{k|k-1}), \quad (2.97)$$

em que se assume o conhecimento de $\hat{x}_{k|k}$ e $\hat{x}_{k|k-1}$. Portanto, utilizando (2.96), pode-se reescrever (2.90) da seguinte forma

$$x_{k+1} = A_k x_k + u_{c,k} + B_{w,k} w_k, \quad (2.98)$$

em que $u_{c,k} := f_k(\hat{x}_{k|k}, u_k) - A_k \hat{x}_{k|k}$. A substituição de (2.97) em (2.91) fornece

$$y_{c,k} = C_k x_k + D_{v,k} v_k, \quad (2.99)$$

em que $y_{c,k} := y_k - h_k(\hat{x}_{k|k-1}) + C_k \hat{x}_{k|k-1}$. Seguindo este procedimento, o modelo (2.90)-(2.91) pode ser reescrito por

$$x_{k+1} = A_k x_k + u_{c,k} + B_{w,k} w_k, \quad (2.100)$$

$$y_{c,k} = C_k x_k + D_{v,k} v_k. \quad (2.101)$$

Percebe-se que o modelo agora é linear em x_k . Desta forma, pode-se aplicar as equações do filtro de Kalman. Nota-se, também, que $u_{c,k}$ e $y_{c,k}$ são conhecidos no instante k . Outra observação é que nenhuma consideração sobre as propriedades estocásticas dos ruídos foi necessária até este ponto. O FKE é obtido a partir da aplicação do filtro linear ao sistema (2.100)-(2.101). O algoritmo do FKE na sua forma preditora é apresentado na Tabela 2.5 e é o mesmo encontrado na literatura. *e.g.*, [103, 64]. A forma predito-corretora do FKE para tempo discreto pode ser encontrada em [103], [64], [104].

Tabela 2.5: Algoritmo do FKE na sua forma preditora de 1 estágio.

Passo 0 (Condições Iniciais):

$$\begin{aligned}\hat{x}_{0|-1} &= \bar{x}_0, \\ \bar{P}_{0|-1} &= X_0.\end{aligned}$$

Passo 1: Obter matrizes linearizadas

$$\begin{aligned}A_k &= \left. \frac{\partial}{\partial x_k} f_k(x_k, u_k) \right|_{x_k = \hat{x}_{k|k-1}}, \\ C_k &= \left. \frac{\partial}{\partial x_k} h_k(x_k) \right|_{x_k = \hat{x}_{k|k-1}}.\end{aligned}$$

Passo 2: Calcular o ganho do preditor

$$K_k = A_k \bar{P}_{k|k-1} C_k^T \left(C_k \bar{P}_{k|k-1} C_k^T + D_{v,k} V_k D_{v,k}^T \right)^{-1}.$$

Passo 3: Atualizar $\{\hat{x}_{k+1|k}, \bar{P}_{k+1|k}\}$ como

$$\begin{aligned}\hat{x}_{k+1|k} &= f_k(x_{k|k-1}, u_k) + K_k (y_k - h_k(\hat{x}_{k|k-1})), \\ \bar{P}_{k+1|k} &= A_k \bar{P}_{k|k-1} A_k^T + B_{w,k} W_k B_{w,k}^T \\ &\quad - A_k \bar{P}_{k|k-1} C_k^T \left(C_k \bar{P}_{k|k-1} C_k^T + D_{v,k} V_k D_{v,k}^T \right)^{-1} (A_k \bar{P}_{k|k-1} C_k^T)^T.\end{aligned}$$

3 ESTIMADORES ROBUSTOS LINEARES

“The beginner...should not be discouraged if he finds that he does not have the prerequisites for reading the prerequisites”. P. Halmos (1916-2006) citado em [106]. Neste capítulo, será apresentado o estimador robusto de Kalman para modelos lineares sujeitos a incertezas.

3.1 INTRODUÇÃO À ESTIMAÇÃO LINEAR ROBUSTA

Nos últimos anos, o projeto e a análise de estimadores de Kalman robustos para sistemas lineares discretos no tempo têm atraído bastante atenção e diversos resultados têm sido obtidos. A motivação para o projeto de estimadores robustos deve-se à queda de desempenho do filtro de Kalman clássico ao tratar com modelos sujeitos a incertezas [73, 88]. Neste contexto, uma outra técnica desenvolvida para a estimação robusta é a estimação \mathcal{H}_∞ , em que as fontes de ruído são sinais de energia ou de potência média limitadas. Embora mais robusto que o estimador de Kalman, o estimador \mathcal{H}_∞ também é afetado pela mesma dependência sobre a exatidão dos parâmetros do modelo e, assim, surgiu o estudo sobre os estimadores \mathcal{H}_∞ robustos. Na estimação \mathcal{H}_∞ robusta são aplicadas metodologias similares às empregadas para os estimadores de Kalman robustos (ver, *e.g.*, [65, 66]).

As próximas seções tratam do problema de estimação robusta para sistemas lineares e discretos no tempo sujeitos a incertezas, desconhecidas e com norma suposta limitada, que afetam todas as matrizes do modelo. Se um estimador consegue lidar com as incertezas do modelo, ou seja, se ele pode ser utilizado com um desempenho garantido para todo o conjunto de incertezas permitidas, este estimador é conhecido na literatura como *estimador robusto* [86].

Essencialmente, as pesquisas atuais têm utilizado duas abordagens para a estimação robusta: uma por desigualdades matriciais lineares (LMIs) e outra por equação de Riccati. O projeto de estimadores robustos pela abordagem de LMIs é capaz de lidar com incertezas politópicas ou com norma limitada. Esforços atuais nesta área são no sentido de se projetar estimadores menos conservadores, *e.g.*, utilizando funções de Lyapunov paramétricas [107], e de se generalizar a teoria para sistemas descritores com incertezas politópicas [108].

Alternativamente, de acordo com [86], a vantagem da abordagem por equação de Riccati é permitir uma melhor compreensão do problema, uma vez que obtidas as expressões do preditor, o efeito das incertezas na estrutura e no ganho do estimador é diretamente perceptível. Mais adiante, será observado que a presença de incertezas no modelo tem o efeito de alterar as matrizes de covariância associadas aos ruídos e ao erro de estimação, bem como termos aditivos na equação de Riccati. Ainda, para o caso clássico sem incertezas, mostra-se que as expressões obtidas por meio da abordagem por equação de Riccati recaem no Filtro de Kalman tradicional. Abordagens numéricas, tais como a utilização de LMIs, não permitem a visualização dos efeitos das incertezas na estrutura e nos parâmetros do preditor. Uma das técnicas utilizadas envolve a solução de duas equações de Riccati discretas ([67], [74], [86] e [89]). Outra possibilidade é baseada na solução de somente uma equação de Riccati, ver [90, 82].

O problema da estimação com custo garantido consiste em projetar um estimador linear de maneira a assegurar um limite superior para a variância do erro de estimação para todo o conjunto de incertezas permitidas. Um dos primeiros estimadores de custo garantido foi proposto por [109] para modelos contínuos e invariantes no tempo com incertezas na equação de estados. Em [88], desenvolveu-se um preditor de Kalman robusto para sistemas a tempo discreto sujeitos a incertezas variantes e limitadas nas equações de estado e de saída. A estimação de estados para sistemas contínuos e discretos no tempo e com incertezas com norma limitada é tratada em [90]. Condições necessárias e suficientes para o projeto de estimadores robustos nos horizontes finito e infinito são dadas em [67]. Recentemente, em [82] foi proposto um preditor robusto com custo garantido para um modelo sujeito a incertezas nas matrizes de estado e de saída e na covariância dos ruídos. Entretanto, uma condição do projeto é que os ruídos de estado e de saída tenham a mesma dimensão.

3.2 OBJETIVOS

Esta seção propõe uma melhoria no projeto de preditores robustos de horizonte finito para a estimação de estados de um sistema sujeito a incertezas. O objetivo é propor um estimador de custo garantido que seja:

- menos conservador que o desenvolvido em [82],
- capaz de lidar com ruídos com matrizes de covariância e covariância cruzada incertas e
- permitir que os sinais dos ruídos possuam médias incertas e dimensões diferentes.

Uma das técnicas geralmente utilizadas para o problema de estimação robusta chama-se *Método-S*, ou *S-Procedure*, que substitui blocos de matrizes de incertezas por valores escalares [71]. Esta abordagem resulta em um limite superior para a matriz de covariância do erro de estimação. Este projeto de estimadores também é conhecido como projeto com custo garantido. Na literatura, encontram-se dois tipos de parâmetros escalares: constantes ou variantes no tempo. Parâmetros constantes são utilizados para problemas de horizonte finito ou em regime permanente [71]. Entretanto, um ponto fraco desta abordagem é que seu conservadorismo pode levar a estimadores imprecisos com o tempo. Parâmetros variantes no tempo são capazes de reduzir o conservadorismo, desde que sejam corretamente escolhidos. O foco deste trabalho será o projeto de preditores robustos com parâmetros escalares variantes no tempo.

Tomando por base a estrutura do preditor de Kalman clássico, supõe-se que o preditor possua a seguinte estrutura

$$\hat{x}_{0|-1} = \bar{x}_0, \quad (3.1)$$

$$\hat{x}_{k+1|k} = \Phi_k \hat{x}_{k|k-1} + K_k (y_k - C_k \hat{x}_{k|k-1}). \quad (3.2)$$

Os parâmetros do preditor devem ser escolhidos de forma a assegurar um limite superior para a variância do erro de estimação. Em outras palavras, deve-se garantir a existência de uma sequência de matrizes definidas-positivas $\bar{P}_{k|k-1}$ que, para todas as incertezas permitidas, satisfaçam

$$\text{cov} \{x_{k+1} - \hat{x}_{k+1|k}\} \leq \bar{P}_{k+1|k}. \quad (3.3)$$

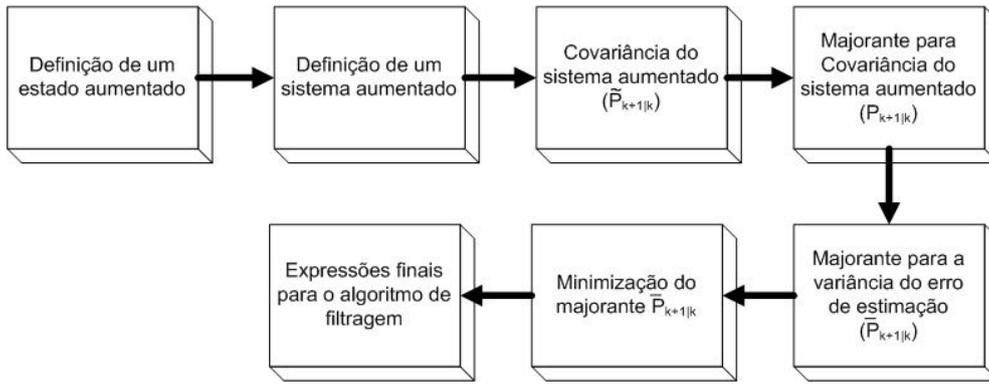


Figura 3.1: Passos para a obtenção do preditor robusto por meio da abordagem da equação de Riccati.

Os procedimentos adotados para o desenvolvimento do preditor robusto são mostrados na Figura 3.1. O primeiro passo consiste em se obter um estado aumentado apropriado, envolvendo, de alguma forma, o estado real do sistema e a estimativa. O erro de estimação pode estar de maneira explícita no estado aumentado ou deve ser obtido por meio de uma transformação de similaridade. O segundo passo é a obtenção da matriz de covariância do sistema aumentado, a qual será denotada por $\tilde{P}_{k+1|k}$. A seguir, é necessário encontrar um limite superior para a matriz de covariância do sistema aumentado ($P_{k+1|k}$), visando garantir que $\tilde{P}_{k+1|k} \leq P_{k+1|k}$ para todo o instante k . A partir deste majorante, o próximo passo é obter um limite superior para o erro de estimação $\bar{P}_{k+1|k}$, de tal forma a satisfazer (3.3). Utilizando a estrutura do preditor, as matrizes Φ_k e K_k podem ser determinadas com o objetivo de se minimizar $\bar{P}_{k+1|k}$. Por fim, são determinadas as expressões finais para o preditor e para a equação de Riccati. Após este breve comentário sobre a resolução do problema da predição robusta, as seções a seguir mostrarão cada passo com maiores detalhes.

3.3 MODELAGEM DE SISTEMAS SUJEITOS A INCERTEZAS

Nesta seção, serão apresentados aspectos sobre a modelagem de sistemas sujeitos a incertezas aditivas em seus parâmetros. Inicialmente, considere a seguinte classe de sistemas incertos discretos no tempo

$$x_{k+1} = A_{\Delta,k}x_k + \tilde{w}_k, \quad (3.4)$$

$$y_k = C_{\Delta,k}x_k + \tilde{v}_k, \quad (3.5)$$

em que $x_k \in \mathbb{R}^{n_x}$ representa o vetor de estados, $y_k \in \mathbb{R}^{n_y}$ é o vetor de saída e $\tilde{w}_k \in \mathbb{R}^{n_x}$ e $\tilde{v}_k \in \mathbb{R}^{n_y}$ são os sinais de ruído de processo e de medida respectivamente. Admite-se que o sistema possua as seguintes incertezas:

1. Incertezas aditivas na dinâmica do modelo representadas por $A_{\Delta,k} = A_k + \Delta A_k$, em que se admite A_k como uma matriz conhecida, ou nominal, e ΔA_k como a incerteza sobre a matriz nominal.
2. Incertezas aditivas na matriz de saída do modelo, de forma que $C_{\Delta,k} = C_k + \Delta C_k$, em que C_k seja a matriz nominal de saída e ΔC_k caracterize a incerteza sobre C_k .
3. Incertezas nas médias, covariâncias e covariâncias cruzadas dos ruídos \tilde{w}_k e \tilde{v}_k . Considerando que

os ruídos sejam não-correlacionados com o estado inicial, tem-se que

$$E \left\{ \begin{bmatrix} \tilde{w}_k \\ \tilde{v}_k \\ x_0 \end{bmatrix} \right\} = \begin{bmatrix} E \{ \tilde{w}_k \} \\ E \{ \tilde{v}_k \} \\ \bar{x}_0 \end{bmatrix}, \quad (3.6)$$

$$E \left\{ \begin{bmatrix} \tilde{w}_k - E \{ \tilde{w}_k \} \\ \tilde{v}_k - E \{ \tilde{v}_k \} \\ x_0 - \bar{x}_0 \end{bmatrix} \begin{bmatrix} \tilde{w}_j - E \{ \tilde{w}_j \} \\ \tilde{v}_j - E \{ \tilde{v}_j \} \\ x_0 - \bar{x}_0 \end{bmatrix}^T \right\} = \begin{bmatrix} \tilde{W}_k \delta_{kj} & \tilde{S}_k \delta_{kj} & 0 \\ \tilde{S}_k^T \delta_{kj} & \tilde{V}_k \delta_{kj} & 0 \\ 0 & 0 & X_0 \end{bmatrix}, \quad (3.7)$$

em que \tilde{W}_k , \tilde{V}_k e X_0 denotam as matrizes de covariância dos ruídos e do estado inicial, \tilde{S}_k é a covariância cruzada e δ_{kj} é a função delta de Kronecker.

Embora os valores das médias e das covariâncias sejam desconhecidos, assume-se que pertençam a um conjunto conhecido. Os conjuntos das covariâncias serão denotados conforme (3.8).

$$\tilde{W}_k \in \mathcal{W}_k, \quad \tilde{V}_k \in \mathcal{V}_k, \quad \tilde{S}_k \in \mathcal{S}_k. \quad (3.8)$$

Ou, de forma mais compacta,

$$\tilde{N}_k := \begin{bmatrix} \tilde{W}_k & \tilde{S}_k \\ \tilde{S}_k^T & \tilde{V}_k \end{bmatrix} \in \begin{bmatrix} \mathcal{W}_k & \mathcal{S}_k \\ \mathcal{S}_k^T & \mathcal{V}_k \end{bmatrix}. \quad (3.9)$$

Nota 3.3.1 Em [82], considera-se que os ruídos de processo e de medida são não-correlacionados e com média nula. Em outras palavras,

$$E \left\{ \begin{bmatrix} \tilde{w}_k \\ \tilde{v}_k \\ x_0 \end{bmatrix} \right\} = \begin{bmatrix} 0 \\ 0 \\ \bar{x}_0 \end{bmatrix}, \quad (3.10)$$

$$E \left\{ \begin{bmatrix} \tilde{w}_k \\ \tilde{v}_k \\ x_0 - \bar{x}_0 \end{bmatrix} \begin{bmatrix} \tilde{w}_j \\ \tilde{v}_j \\ x_0 - \bar{x}_0 \end{bmatrix}^T \right\} = \begin{bmatrix} \tilde{W}_k \delta_{kj} & 0 & 0 \\ 0 & \tilde{V}_k \delta_{kj} & 0 \\ 0 & 0 & X_0 \end{bmatrix}. \quad (3.11)$$

Na próxima subseção, mostrar-se-á uma maneira de caracterizar um sistema de covariância incerta como um sistema de covariância fixa, mas com parâmetros incertos.

3.3.1 O espaço das médias e das covariâncias dos ruídos

Nesta subseção, serão analisadas algumas características das incertezas nos ruídos. A subseção começa com o tratamento de incertezas sem estrutura alguma. Em seguida, assume-se que as incertezas sejam aditivas, ou seja, seu efeito é adicionado à matriz conhecida, mas ainda pertencentes a um conjunto desconhecido.

Neste trabalho, permite-se que os ruídos \tilde{w}_k e \tilde{v}_k sejam correlacionados com média, covariância e covariância cruzada desconhecidas, mas pertencendo a um conjunto conhecido. Conforme será mostrado, estas propriedades podem ser alcançadas a partir da seguinte estrutura para os ruídos

$$\tilde{w}_k = B_{\Delta w, k} w_k + B_{\Delta v, k} v_k, \quad (3.12)$$

$$\tilde{v}_k = D_{\Delta w, k} w_k + D_{\Delta v, k} v_k. \quad (3.13)$$

Os ruídos podem ser escritos de forma mais compacta, tal que

$$\tilde{\eta}_k := \begin{bmatrix} \tilde{w}_k \\ \tilde{v}_k \end{bmatrix} = \begin{bmatrix} B_{\Delta w,k} & B_{\Delta v,k} \\ D_{\Delta w,k} & D_{\Delta v,k} \end{bmatrix} \begin{bmatrix} w_k \\ v_k \end{bmatrix} = N_{\Delta,k} \eta_k. \quad (3.14)$$

Assume-se também que as condições iniciais x_0 não sejam correlacionadas com os ruídos $\{w_k\}$ e $\{v_k\}$, possuindo propriedades estatísticas, tais que

$$E \left\{ \begin{bmatrix} w_k \\ v_k \\ x_0 \end{bmatrix} \right\} = \begin{bmatrix} \bar{w}_k \\ \bar{v}_k \\ \bar{x}_0 \end{bmatrix}, \quad (3.15)$$

$$E \left\{ \begin{bmatrix} w_k - \bar{w}_k \\ v_k - \bar{v}_k \\ x_0 - \bar{x}_0 \end{bmatrix} \begin{bmatrix} w_j - \bar{w}_j \\ v_j - \bar{v}_j \\ x_0 - \bar{x}_0 \end{bmatrix}^T \right\} = \begin{bmatrix} W_k \delta_{kj} & S_k \delta_{kj} & 0 \\ S_k^T \delta_{kj} & V_k \delta_{kj} & 0 \\ 0 & 0 & X_0 \end{bmatrix}, \quad (3.16)$$

em que W_k , V_k e X_0 denotam as matrizes de covariância dos ruídos e do estado inicial e S_k representa a matriz de covariância cruzada dos ruídos.

Portanto, utilizando as propriedades (3.15) e (3.16), e as definições dos ruídos fornecidas por (3.12) e (3.13), conclui-se que os ruídos \tilde{w}_k e \tilde{v}_k possuem média incerta da forma

$$E \{ \tilde{w}_k \} = B_{\Delta w,k} \bar{w}_k + B_{\Delta v,k} \bar{v}_k, \quad (3.17)$$

$$E \{ \tilde{v}_k \} = D_{\Delta w,k} \bar{w}_k + D_{\Delta v,k} \bar{v}_k. \quad (3.18)$$

Suas covariâncias também são incertas e dadas por

$$E \left\{ \begin{bmatrix} \tilde{w}_k - E \{ \tilde{w}_k \} \\ \tilde{v}_k - E \{ \tilde{v}_k \} \end{bmatrix} \begin{bmatrix} \tilde{w}_j - E \{ \tilde{w}_j \} \\ \tilde{v}_j - E \{ \tilde{v}_j \} \end{bmatrix}^T \right\} = \begin{bmatrix} \tilde{W}_k \delta_{kj} & \tilde{S}_k \delta_{kj} \\ \tilde{S}_k^T \delta_{kj} & \tilde{V}_k \delta_{kj} \end{bmatrix} = \tilde{N}_k \delta_{kj}. \quad (3.19)$$

Para ruídos caracterizados por (3.12) e (3.13), tem-se que

$$\begin{bmatrix} \tilde{W}_k \delta_{kj} & \tilde{S}_k \delta_{kj} \\ \tilde{S}_k^T \delta_{kj} & \tilde{V}_k \delta_{kj} \end{bmatrix} = \begin{bmatrix} B_{\Delta w,k} & B_{\Delta v,k} \\ D_{\Delta w,k} & D_{\Delta v,k} \end{bmatrix} \begin{bmatrix} W_k \delta_{kj} & S_k \delta_{kj} \\ S_k^T \delta_{kj} & V_k \delta_{kj} \end{bmatrix} \begin{bmatrix} B_{\Delta w,k} & B_{\Delta v,k} \\ D_{\Delta w,k} & D_{\Delta v,k} \end{bmatrix}^T. \quad (3.20)$$

Observa-se que a notação acima é redundante. De fato, sem perda de generalidade, a covariância cruzada pode ser representada de outras duas formas:

1. Adotando

$$\begin{bmatrix} B_{\Delta w,k} & B_{\Delta v,k} \\ D_{\Delta w,k} & D_{\Delta v,k} \end{bmatrix} = \begin{bmatrix} B_{\Delta w,k} & 0 \\ 0 & D_{\Delta v,k} \end{bmatrix}. \quad (3.21)$$

Neste caso, a matriz de covariâncias pode ser representada como

$$\begin{bmatrix} \tilde{W}_k \delta_{kj} & \tilde{S}_k \delta_{kj} \\ \tilde{S}_k^T \delta_{kj} & \tilde{V}_k \delta_{kj} \end{bmatrix} = \begin{bmatrix} B_{\Delta w,k} W_k B_{\Delta w,k}^T & B_{\Delta w,k} S_k D_{\Delta v,k}^T \\ D_{\Delta v,k} S_k^T B_{\Delta w,k}^T & D_{\Delta v,k} V_k D_{\Delta v,k}^T \end{bmatrix} \delta_{kj}. \quad (3.22)$$

2. A covariância cruzada também pode ser representada, se

$$\begin{bmatrix} W_k \delta_{kj} & S_k \delta_{kj} \\ S_k^T \delta_{kj} & V_k \delta_{kj} \end{bmatrix} = \begin{bmatrix} W_k \delta_{kj} & 0 \\ 0 & V_k \delta_{kj} \end{bmatrix} = N_k \delta_{kj}. \quad (3.23)$$

Nesta situação, a matriz de covariâncias é dada por

$$\begin{bmatrix} \widetilde{W}_k \delta_{kj} & \widetilde{S}_k \delta_{kj} \\ \widetilde{S}_k^T \delta_{kj} & \widetilde{V}_k \delta_{kj} \end{bmatrix} = \begin{bmatrix} B_{\Delta w,k} W_k B_{\Delta w,k}^T + B_{\Delta v,k} V_k B_{\Delta v,k}^T & B_{\Delta w,k} W_k D_{\Delta w,k}^T + B_{\Delta v,k} V_k D_{\Delta v,k}^T \\ D_{\Delta w,k} W_k B_{\Delta w,k}^T + D_{\Delta v,k} V_k B_{\Delta v,k}^T & D_{\Delta w,k} W_k D_{\Delta w,k}^T + D_{\Delta v,k} V_k D_{\Delta v,k}^T \end{bmatrix} \times \delta_{kj}. \quad (3.24)$$

Até o momento, não foi feita suposição alguma acerca das incertezas nos modelos dos ruídos (3.12) e (3.13). A exemplo das incertezas nas matrizes de dinâmica e da saída, serão consideradas incertezas aditivas nas estruturas dos ruídos, de forma que

$$B_{\Delta w,k} = B_{w,k} + \Delta B_{w,k}, \quad B_{\Delta v,k} = B_{v,k} + \Delta B_{v,k}, \quad (3.25)$$

$$D_{\Delta w,k} = D_{w,k} + \Delta D_{w,k}, \quad D_{\Delta v,k} = D_{v,k} + \Delta D_{v,k}, \quad (3.26)$$

em que $B_{w,k}$, $B_{v,k}$, $D_{w,k}$ e $D_{v,k}$ denotam as matrizes nominais, e suas incertezas são caracterizadas por $\Delta B_{w,k}$, $\Delta B_{v,k}$, $\Delta D_{w,k}$ e $\Delta D_{v,k}$, respectivamente. Supondo que as incertezas possuam as estruturas (3.25)-(3.26), então os ruídos podem ser escritos da forma

$$\widetilde{w}_k = (B_{w,k} + \Delta B_{w,k}) w_k + (B_{v,k} + \Delta B_{v,k}) v_k, \quad (3.27)$$

$$\widetilde{v}_k = (D_{w,k} + \Delta D_{w,k}) w_k + (D_{v,k} + \Delta D_{v,k}) v_k. \quad (3.28)$$

De modo alternativo,

$$\widetilde{\eta}_k := \left(\begin{bmatrix} B_{w,k} & B_{v,k} \\ D_{w,k} & D_{v,k} \end{bmatrix} + \begin{bmatrix} \Delta B_{w,k} & \Delta B_{v,k} \\ \Delta D_{w,k} & \Delta D_{v,k} \end{bmatrix} \right) \begin{bmatrix} w_k \\ v_k \end{bmatrix} = (B_{\eta,k} + \Delta B_{\eta,k}) \eta_k. \quad (3.29)$$

Observa-se, também neste caso, que as médias dos ruídos dependem de parâmetros incertos do modelo. O mesmo é válido para a matriz de covariâncias.

Nota 3.3.2 *As incertezas podem ser representadas por um efeito multiplicativo, ou seja,*

$$\widetilde{w}_k = B_{w,k} (I + \Delta B_{w,k}^0) w_k + B_{v,k} (I + \Delta B_{v,k}^0) v_k, \quad (3.30)$$

$$\widetilde{v}_k = D_{w,k} (I + \Delta D_{w,k}^0) w_k + D_{v,k} (I + \Delta D_{v,k}^0) v_k. \quad (3.31)$$

Para esta situação, consideram-se especificamente que existam matrizes de transformação, tais que

$$\Delta B_{w,k} = B_{w,k} \Delta B_{w,k}^0, \quad \Delta B_{v,k} = B_{v,k} \Delta B_{v,k}^0, \quad (3.32)$$

$$\Delta D_{w,k} = D_{w,k} \Delta D_{w,k}^0, \quad \Delta D_{v,k} = D_{v,k} \Delta D_{v,k}^0. \quad (3.33)$$

Nota 3.3.3 *Em [82] são considerados ruídos não-correlacionados, ou seja, assume-se*

$$B_{v,k} = 0, \quad \Delta B_{v,k} = 0, \quad (3.34)$$

$$D_{w,k} = 0, \quad \Delta D_{w,k} = 0. \quad (3.35)$$

3.3.2 Sistemas com Incertezas nos Parâmetros da Planta e dos Ruídos

Nesta subseção serão realizadas considerações a cerca de uma classe de sistema, de modo que tanto a planta como os ruídos de processo e de medida estejam sujeitos a incertezas em seus parâmetros. Adotando os modelos de ruídos definidos em (3.12) e (3.13), o sistema (3.4)-(3.5) pode ser escrito como

$$x_{k+1} = A_{\Delta,k}x_k + B_{\Delta w,k}w_k + B_{\Delta v,k}v_k, \quad (3.36)$$

$$y_k = C_{\Delta,k}x_k + D_{\Delta w,k}w_k + D_{\Delta v,k}v_k. \quad (3.37)$$

De forma alternativa,

$$x_{k+1} = A_{\Delta,k}x_k + B_{\Delta,k}\eta_k, \quad (3.38)$$

$$y_k = C_{\Delta,k}x_k + D_{\Delta,k}\eta_k, \quad (3.39)$$

em que $B_{\Delta,k} := \begin{bmatrix} B_{\Delta w,k} & B_{\Delta v,k} \end{bmatrix}$ e $D_{\Delta,k} := \begin{bmatrix} D_{\Delta w,k} & D_{\Delta v,k} \end{bmatrix}$. Entretanto, estes sistemas descrevem as incertezas de maneira muito genérica. Considerando que tanto o modelo quanto os ruídos estejam sujeitos a incertezas do tipo aditivas, então, utilizando (3.4)-(3.5) e (3.27)-(3.28), o sistema (3.36)-(3.37) pode ser reescrito como

$$x_{k+1} = (A_k + \Delta A_k)x_k + (B_{w,k} + \Delta B_{w,k})w_k + (B_{v,k} + \Delta B_{v,k})v_k, \quad (3.40)$$

$$y_k = (C_k + \Delta C_k)x_k + (D_{w,k} + \Delta D_{w,k})w_k + (D_{v,k} + \Delta D_{v,k})v_k. \quad (3.41)$$

Embora w_k e v_k sejam independentes, o modelo (3.40)-(3.41) é equivalente a um modelo com apenas um vetor de ruído nas equações da dinâmica e de saída com correlação explícita entre eles, conforme mostrado na seção 3.3.1. Para maiores detalhes, consultar [64]. O preditor proposto em [82] é restrito para o caso em que $n_w = n_v$. Neste estudo, esta restrição é eliminada, permitindo-se que os ruídos w_k e v_k possuam dimensões distintas. As matrizes A_k , $B_{w,k}$, $B_{v,k}$, C_k , $D_{w,k}$ e $D_{v,k}$ são conhecidas e variantes no tempo. As matrizes ΔA_k , $\Delta B_{w,k}$, $\Delta B_{v,k}$, ΔC_k , $\Delta D_{w,k}$ e $\Delta D_{v,k}$ representam as incertezas associadas. As dimensões dos conjuntos associados às incertezas são mostradas na Tabela 3.1. As dimensões dos demais vetores e matrizes foram apresentadas na Tabela 2.1.

Tabela 3.1: Conjuntos das matrizes utilizadas no modelo (3.40)-(3.41).

Matriz	Conjunto
ΔA_k	$\mathbb{R}^{n_x \times n_x}$
$\Delta B_{w,k}$	$\mathbb{R}^{n_x \times n_w}$
$\Delta B_{v,k}$	$\mathbb{R}^{n_x \times n_v}$
ΔC_k	$\mathbb{R}^{n_y \times n_x}$
$\Delta D_{w,k}$	$\mathbb{R}^{n_y \times n_w}$
$\Delta D_{v,k}$	$\mathbb{R}^{n_y \times n_v}$

O modelo (3.40)-(3.41) pode ser representado de maneira mais compacta. Uma alternativa é considerar

$$x_{k+1} = (A_k + \Delta A_k)x_k + (B_k + \Delta B_k)\eta_k, \quad (3.42)$$

$$y_k = (C_k + \Delta C_k)x_k + (D_k + \Delta D_k)\eta_k, \quad (3.43)$$

em que as dimensões são apresentadas na Tabela 3.2. É possível observar as seguintes identificações

$$B_k := \begin{bmatrix} B_{w,k} & B_{v,k} \end{bmatrix}, \quad \Delta B_k := \begin{bmatrix} \Delta B_{w,k} & \Delta B_{v,k} \end{bmatrix}, \quad (3.44)$$

$$D_k := \begin{bmatrix} D_{w,k} & D_{v,k} \end{bmatrix}, \quad \Delta D_k := \begin{bmatrix} \Delta D_{w,k} & \Delta D_{v,k} \end{bmatrix}. \quad (3.45)$$

Tabela 3.2: Dimensões das matrizes e vetores utilizados no modelo (3.42)-(3.43).

Matriz ou Vetor	Conjunto
η_k	$\mathbb{R}^{n_w+n_v}$
$B_k + \Delta B_k$	$\mathbb{R}^{n_x \times (n_w+n_v)}$
$D_k + \Delta D_k$	$\mathbb{R}^{n_y \times (n_w+n_v)}$

Até o momento, considerou-se apenas que as incertezas aditivas pertencessem a um conjunto conhecido. A partir de agora, são necessárias mais informações sobre as incertezas. De maneira geral, as incertezas são consideradas de norma limitada ou contidas em um polítopo. O segundo modelo apresenta maior complexidade de tratamento, apesar de o conjunto de incertezas de norma limitada estar contido no conjunto representado por um polítopo.

Neste trabalho, serão consideradas incertezas com norma limitada. Para o caso mais geral, cada incerteza do sistema planta + ruídos pode ser representada como

$$\Delta A_k = H_{A,k} F_{A,k} G_{A,k}, \quad (3.46)$$

$$\Delta B_{w,k} = H_{Bw,k} F_{Bw,k} G_{Bw,k}, \quad (3.47)$$

$$\Delta B_{v,k} = H_{Bv,k} F_{Bv,k} G_{Bv,k}, \quad (3.48)$$

$$\Delta C_k = H_{C,k} F_{C,k} G_{C,k}, \quad (3.49)$$

$$\Delta D_{w,k} = H_{Dw,k} F_{Dw,k} G_{Dw,k}, \quad (3.50)$$

$$\Delta D_{v,k} = H_{Dv,k} F_{Dv,k} G_{Dv,k}. \quad (3.51)$$

com $H_{A,k}$, $H_{Bw,k}$, $H_{Bv,k}$, $H_{C,k}$, $H_{Dw,k}$, $H_{Dv,k}$, $G_{x,k}$, $G_{w,k}$ e $G_{v,k}$ conhecidas e com dimensões apropriadas. A estrutura adotada desta maneira permite que se admita $n_w \neq n_v$. As matrizes $F_{A,k}$, $F_{Bw,k}$, $F_{Bv,k}$, $F_{C,k}$, $F_{Dw,k}$ e $F_{Dv,k}$ são desconhecidas, variantes com o tempo e com norma limitada, *i.e.*,

$$F_{A,k}^T F_{A,k} \leq I, F_{Bw,k}^T F_{Bw,k} \leq I, F_{Bv,k}^T F_{Bv,k} \leq I, F_{C,k}^T F_{C,k} \leq I, F_{Dw,k}^T F_{Dw,k} \leq I, F_{Dv,k}^T F_{Dv,k} \leq I. \quad (3.52)$$

As incertezas também podem ser representadas no formato matricial, a saber

$$\begin{aligned}
& \begin{bmatrix} \Delta A_k & \Delta B_{w,k} & \Delta B_{v,k} \\ \Delta C_k & \Delta D_{w,k} & \Delta D_{v,k} \end{bmatrix} = \begin{bmatrix} H_{A,k}F_{A,k}G_{A,k} & H_{Bw,k}F_{Bw,k}G_{Bw,k} & H_{Bv,k}F_{Bv,k}G_{Bv,k} \\ H_{C,k}F_{C,k}G_{C,k} & H_{Dw,k}F_{Dw,k}G_{Dw,k} & H_{Dv,k}F_{Dv,k}G_{Dv,k} \end{bmatrix} \\
& = \begin{bmatrix} H_{A,k} & H_{Bw,k} & H_{Bv,k} & 0 & 0 & 0 \\ 0 & 0 & 0 & H_{C,k} & H_{Dw,k} & H_{Dv,k} \end{bmatrix} \\
& \times \text{diag}\{F_{A,k}, F_{Bw,k}, F_{Bv,k}, F_{C,k}, F_{Dw,k}, F_{Dv,k}\} \begin{bmatrix} G_{A,k} & 0 & 0 \\ 0 & G_{Bw,k} & 0 \\ 0 & 0 & G_{Bv,k} \\ G_{C,k} & 0 & 0 \\ 0 & G_{Dw,k} & 0 \\ 0 & 0 & G_{Dv,k} \end{bmatrix}. \quad (3.53)
\end{aligned}$$

Uma forma mais genérica para incertezas com norma limitada foi apresentada em (3.53). Entretanto, uma outra forma de representar tais incertezas, por meio da escolha apropriada das matrizes H , pode ser obtida do seguinte modo

$$\begin{bmatrix} \Delta A_k \\ \Delta C_k \end{bmatrix} = \begin{bmatrix} H_{A,k} \\ H_{C,k} \end{bmatrix} F_{x,k} G_{x,k} \quad (3.54)$$

$$\begin{bmatrix} \Delta B_{w,k} \\ \Delta D_{w,k} \end{bmatrix} = \begin{bmatrix} H_{Bw,k} \\ H_{Dw,k} \end{bmatrix} F_{w,k} G_{w,k} \quad (3.55)$$

$$\begin{bmatrix} \Delta B_{v,k} \\ \Delta D_{v,k} \end{bmatrix} = \begin{bmatrix} H_{Bv,k} \\ H_{Dv,k} \end{bmatrix} F_{v,k} G_{v,k}, \quad (3.56)$$

em que $F_{x,k}^T F_{x,k} \leq I$, $F_{w,k}^T F_{w,k} \leq I$ e $F_{v,k}^T F_{v,k} \leq I$. Reescrevendo as incertezas em uma única estrutura,

$$\begin{aligned}
& \begin{bmatrix} \Delta A_k & \Delta B_{w,k} & \Delta B_{v,k} \\ \Delta C_k & \Delta D_{w,k} & \Delta D_{v,k} \end{bmatrix} = \begin{bmatrix} H_{A,k}F_{x,k}G_{x,k} & H_{Bw,k}F_{w,k}G_{w,k} & H_{Bv,k}F_{v,k}G_{v,k} \\ H_{C,k}F_{x,k}G_{x,k} & H_{Dw,k}F_{w,k}G_{w,k} & H_{Dv,k}F_{v,k}G_{v,k} \end{bmatrix} \\
& = \begin{bmatrix} H_{A,k} & H_{Bw,k} & H_{Bv,k} \\ H_{C,k} & H_{Dw,k} & H_{Dv,k} \end{bmatrix} \begin{bmatrix} F_{x,k} & 0 & 0 \\ 0 & F_{w,k} & 0 \\ 0 & 0 & F_{v,k} \end{bmatrix} \begin{bmatrix} G_{x,k} & 0 & 0 \\ 0 & G_{w,k} & 0 \\ 0 & 0 & G_{v,k} \end{bmatrix}. \quad (3.57)
\end{aligned}$$

Como anteriormente, serão analisadas incertezas de norma limitada. Definindo as perturbações

$$\begin{bmatrix} \Delta B_k \\ \Delta D_k \end{bmatrix} := \begin{bmatrix} H_{B,k} \\ H_{D,k} \end{bmatrix} F_{\eta,k} G_{\eta,k}, \quad (3.58)$$

em que $F_{\eta,k}^T F_{\eta,k} \leq I$, conclui-se que

$$\begin{aligned}
& \begin{bmatrix} \Delta A_k & \Delta B_k \\ \Delta C_k & \Delta D_k \end{bmatrix} = \begin{bmatrix} H_{A,k}F_{x,k}G_{x,k} & H_{B,k}F_{\eta,k}G_{\eta,k} \\ H_{C,k}F_{x,k}G_{x,k} & H_{D,k}F_{\eta,k}G_{\eta,k} \end{bmatrix} \\
& = \begin{bmatrix} H_{A,k} & H_{B,k} \\ H_{C,k} & H_{D,k} \end{bmatrix} \begin{bmatrix} F_{x,k} & 0 \\ 0 & F_{\eta,k} \end{bmatrix} \begin{bmatrix} G_{x,k} & 0 \\ 0 & G_{\eta,k} \end{bmatrix}. \quad (3.59)
\end{aligned}$$

Assim, exceto no caso em que $F_{x,k} = F_{\eta,k}$, em geral, $\begin{bmatrix} F_{x,k} & 0 \\ 0 & F_{\eta,k} \end{bmatrix}$ possui estrutura bloco-diagonal.

3.4 O SISTEMA AUMENTADO INCERTO

Nesta seção, serão apresentados alguns sistemas aumentados do sistema compreendendo a planta e o preditor. Inicialmente, será tratada a escolha do vetor de estados aumentado. Logo depois, serão apresentadas alternativas para a representação em espaço de estado do sistema aumentado.

Considere as seguintes escolhas possíveis, e normalmente encontradas na literatura, para o vetor de estados aumentado

$$\tilde{x}_k = \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix}, \tilde{x}_k = \begin{bmatrix} x_k - \hat{x}_{k|k-1} \\ \hat{x}_{k|k-1} \end{bmatrix}. \quad (3.60)$$

Nota-se que existe uma transformação de similaridade que faz a correspondência entre os dois vetores aumentados. Esta matriz de transformação, e a sua inversa, são dadas por

$$T = \begin{bmatrix} I & I \\ 0 & I \end{bmatrix}, T^{-1} = \begin{bmatrix} I & -I \\ 0 & I \end{bmatrix}. \quad (3.61)$$

A seguir, serão apresentados os sistemas aumentados a partir de modelos incertos. Primeiramente, mostrar-se-ão os sistemas para as incertezas modeladas sem considerações sobre sua estrutura, conforme (3.12)-(3.13). Em seguida, serão mostrados os sistemas aumentados em que é adotada a hipótese de incertezas aditivas, como em (3.27)-(3.28). Por fim, as incertezas são supostas com norma limitada, e os respectivos sistemas aumentados são obtidos.

Lema 3.4.1 *Para o sistema formado pela planta e pelos ruídos dado por*¹

$$x_{k+1} = A_{\Delta,k}x_k + B_{\Delta,k}\eta_k = A_{\Delta,k}x_k + B_{\Delta w,k}w_k + B_{\Delta v,k}v_k, \quad (3.62)$$

$$y_k = C_{\Delta,k}x_k + D_{\Delta,k}\eta_k = C_{\Delta,k}x_k + D_{\Delta w,k}w_k + D_{\Delta v,k}v_k, \quad (3.63)$$

e com o estimador, com a estrutura definida em (3.2), pode haver as seguintes representações aumentadas

$$\begin{aligned} \begin{bmatrix} x_{k+1} \\ \hat{x}_{k+1|k} \end{bmatrix} &= \begin{bmatrix} A_{\Delta,k} & 0 \\ K_k C_{\Delta,k} & \Phi_k - K_k C_k \end{bmatrix} \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix} + \begin{bmatrix} B_{\Delta w,k} \\ K_k D_{\Delta w,k} \end{bmatrix} w_k + \begin{bmatrix} B_{\Delta v,k} \\ K_k D_{\Delta v,k} \end{bmatrix} v_k \\ &= \begin{bmatrix} A_{\Delta,k} & 0 \\ K_k C_{\Delta,k} & \Phi_k - K_k C_k \end{bmatrix} \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix} + \begin{bmatrix} B_{\Delta,k} \\ K_k D_{\Delta,k} \end{bmatrix} \eta_k. \end{aligned} \quad (3.64)$$

ou, ainda,

$$\begin{aligned} \begin{bmatrix} x_{k+1} - \hat{x}_{k+1|k} \\ \hat{x}_{k+1|k} \end{bmatrix} &= \begin{bmatrix} A_{\Delta,k} - K_k C_{\Delta,k} & A_{\Delta,k} - \Phi_k + K_k C_k - K_k C_{\Delta,k} \\ K_k C_{\Delta,k} & \Phi_k - K_k C_k + K_k C_{\Delta,k} \end{bmatrix} \begin{bmatrix} x_k - \hat{x}_{k|k-1} \\ \hat{x}_{k|k-1} \end{bmatrix} \\ &+ \begin{bmatrix} B_{\Delta w,k} - K_k D_{\Delta w,k} \\ K_k D_{\Delta w,k} \end{bmatrix} w_k + \begin{bmatrix} B_{\Delta v,k} - K_k D_{\Delta v,k} \\ K_k D_{\Delta v,k} \end{bmatrix} v_k \\ &= \begin{bmatrix} A_{\Delta,k} - K_k C_{\Delta,k} & A_{\Delta,k} - \Phi_k + K_k C_k - K_k C_{\Delta,k} \\ K_k C_{\Delta,k} & \Phi_k - K_k C_k + K_k C_{\Delta,k} \end{bmatrix} \begin{bmatrix} x_k - \hat{x}_{k|k-1} \\ \hat{x}_{k|k-1} \end{bmatrix} \\ &+ \begin{bmatrix} B_{\Delta,k} - K_k D_{\Delta,k} \\ K_k D_{\Delta,k} \end{bmatrix} \eta_k. \end{aligned} \quad (3.65)$$

¹Modelos definidos em (3.36)-(3.37) e (3.38)-(3.39), mas repetidos aqui para facilidade de leitura.

Prova Substituição direta de (3.62), (3.63) e (3.2) no vetor de estados aumentado. ■

Nota 3.4.2 As influências das incertezas e dos parâmetros podem ser mais bem visualizadas, reescrevendo (3.64) da forma:

$$\begin{bmatrix} x_{k+1} \\ \hat{x}_{k+1|k} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \Psi_k \begin{bmatrix} 0 \\ I \end{bmatrix} \end{bmatrix} \begin{bmatrix} A_{\Delta,k} \\ C_{\Delta,k} \end{bmatrix} \begin{bmatrix} 0 \\ \Psi_k \begin{bmatrix} I \\ -C_k \end{bmatrix} \end{bmatrix} \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix} \\ + \begin{bmatrix} I & 0 \\ 0 & \Psi_k \begin{bmatrix} 0 \\ I \end{bmatrix} \end{bmatrix} \begin{bmatrix} B_{\Delta w,k} \\ D_{\Delta w,k} \end{bmatrix} w_k + \begin{bmatrix} I & 0 \\ 0 & \Psi_k \begin{bmatrix} 0 \\ I \end{bmatrix} \end{bmatrix} \begin{bmatrix} B_{\Delta v,k} \\ D_{\Delta v,k} \end{bmatrix} v_k, \quad (3.66)$$

em que os parâmetros do preditor são representados por Ψ_k , tal que $\Psi_k := \begin{bmatrix} \Phi_k & K_k \end{bmatrix}$. Observando (3.66), a parte da covariância correspondente à

1. Incerteza na matriz dinâmica só envolve $A_{\Delta,k}$ e $C_{\Delta,k}$.
2. Incerteza na matriz do ruído na dinâmica só envolve $B_{\Delta w,k}$ e $D_{\Delta w,k}$.
3. Incerteza na matriz do ruído na saída só envolve $B_{\Delta v,k}$ e $D_{\Delta v,k}$.

Neste sentido, parece razoável definir estruturas separadas para $\begin{bmatrix} A_{\Delta,k} \\ C_{\Delta,k} \end{bmatrix}$, $\begin{bmatrix} B_{\Delta w,k} \\ D_{\Delta w,k} \end{bmatrix}$ e $\begin{bmatrix} B_{\Delta v,k} \\ D_{\Delta v,k} \end{bmatrix}$, de tal forma que, se $\begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix}$, w_k e v_k forem não-correlacionados dois a dois, obtém-se uma expressão menos conservadora a ser utilizada para o cálculo da covariância. ■

Considerando que as incertezas do modelo são do tipo aditivas, conforme (3.25)-(3.26), então o modelo aumentado (3.64) com a planta e o estimador pode ser escrito em função das incertezas como

$$\begin{bmatrix} x_{k+1} \\ \hat{x}_{k+1|k} \end{bmatrix} = \left(\begin{bmatrix} A_k & 0 \\ K_k C_k & \Phi_k - K_k C_k \end{bmatrix} + \begin{bmatrix} \Delta A_k & 0 \\ K_k \Delta C_k & 0 \end{bmatrix} \right) \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix} \\ + \left(\begin{bmatrix} B_{w,k} \\ K_k D_{w,k} \end{bmatrix} + \begin{bmatrix} \Delta B_{w,k} \\ K_k \Delta D_{w,k} \end{bmatrix} \right) w_k + \left(\begin{bmatrix} B_{v,k} \\ K_k D_{v,k} \end{bmatrix} + \begin{bmatrix} \Delta B_{v,k} \\ K_k \Delta D_{v,k} \end{bmatrix} \right) v_k \\ = \left(\begin{bmatrix} A_k & 0 \\ K_k C_k & \Phi_k - K_k C_k \end{bmatrix} + \begin{bmatrix} \Delta A_k & 0 \\ K_k \Delta C_k & 0 \end{bmatrix} \right) \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix} \\ + \left(\begin{bmatrix} B_k \\ K_k D_k \end{bmatrix} + \begin{bmatrix} \Delta B_k \\ K_k \Delta D_k \end{bmatrix} \right) \eta_k. \quad (3.67)$$

Nota 3.4.3 A partir de (3.67), percebe-se que a parte da covariância correspondente à

1. Incerteza na matriz dinâmica só envolve ΔA_k e ΔC_k .
2. Incerteza na matriz do ruído na dinâmica só envolve $\Delta B_{w,k}$ e $\Delta D_{w,k}$.
3. Incerteza na matriz do ruído na saída só envolve $\Delta B_{v,k}$ e $\Delta D_{v,k}$.

Com base na nota anterior, faz sentido escolher estruturas separadas para as incertezas nas matrizes de dinâmica, do ruído de dinâmica e do ruído de saída. Portanto, se o sistema aumentado incerto (3.67) possuir incertezas aditivas com norma limitada satisfazendo (3.54)-(3.56), então o modelo planta + estimador pode ser escrito como

$$\begin{aligned} \begin{bmatrix} x_{k+1} \\ \hat{x}_{k+1|k} \end{bmatrix} &= \left(\begin{bmatrix} A_k & 0 \\ K_k C_k & \Phi_k - K_k C_k \end{bmatrix} + \begin{bmatrix} H_{A,k} \\ K_k H_{C,k} \end{bmatrix} F_{x,k} \begin{bmatrix} G_{x,k} & 0 \end{bmatrix} \right) \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix} \\ &+ \left(\begin{bmatrix} B_{w,k} \\ K_k D_{w,k} \end{bmatrix} + \begin{bmatrix} H_{Bw,k} \\ K_k H_{Dw,k} \end{bmatrix} F_{w,k} G_{w,k} \right) w_k \\ &+ \left(\begin{bmatrix} B_{v,k} \\ K_k D_{v,k} \end{bmatrix} + \begin{bmatrix} H_{Bv,k} \\ K_k H_{Dv,k} \end{bmatrix} F_{v,k} G_{v,k} \right) v_k. \end{aligned} \quad (3.68)$$

De forma mais compacta,

$$\tilde{x}_{k+1} = \left(\tilde{A}_k + \tilde{H}_{x,k} \tilde{F}_{x,k} \tilde{G}_{x,k} \right) \tilde{x}_k + \left(\tilde{B}_{w,k} + \tilde{H}_{w,k} \tilde{F}_{w,k} \tilde{G}_{w,k} \right) w_k + \left(\tilde{B}_{v,k} + \tilde{H}_{v,k} \tilde{F}_{v,k} \tilde{G}_{v,k} \right) v_k, \quad (3.69)$$

em que $\tilde{F}_{x,k} = F_{x,k}$, $\tilde{F}_{w,k} = F_{w,k}$, $\tilde{F}_{v,k} = F_{v,k}$, $\tilde{G}_{w,k} = G_{w,k}$, $\tilde{G}_{v,k} = G_{v,k}$,

$$\tilde{x}_k = \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix}, \quad (3.70)$$

$$\tilde{A}_k = \begin{bmatrix} A_k & 0 \\ K_k C_k & \Phi_k - K_k C_k \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \Psi_k \end{bmatrix} \begin{bmatrix} A_k \begin{bmatrix} I & 0 \\ 0 & I \\ C_k & 0 \end{bmatrix} \end{bmatrix}, \quad (3.71)$$

$$\tilde{B}_{w,k} = \begin{bmatrix} B_{w,k} \\ K_k D_{w,k} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \Psi_k \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \begin{bmatrix} 0 \\ I \end{bmatrix} \end{bmatrix} \begin{bmatrix} B_{w,k} \\ D_{w,k} \end{bmatrix}, \quad (3.72)$$

$$\tilde{B}_{v,k} = \begin{bmatrix} B_{v,k} \\ K_k D_{v,k} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \Psi_k \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \begin{bmatrix} 0 \\ I \end{bmatrix} \end{bmatrix} \begin{bmatrix} B_{v,k} \\ D_{v,k} \end{bmatrix}, \quad (3.73)$$

$$\tilde{H}_{x,k} = \begin{bmatrix} H_{A,k} \\ K_k H_{C,k} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \Psi_k \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \begin{bmatrix} 0 \\ I \end{bmatrix} \end{bmatrix} \begin{bmatrix} H_{A,k} \\ H_{C,k} \end{bmatrix}, \quad (3.74)$$

$$\tilde{H}_{w,k} = \begin{bmatrix} H_{Bw,k} \\ K_k H_{Dw,k} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \Psi_k \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \begin{bmatrix} 0 \\ I \end{bmatrix} \end{bmatrix} \begin{bmatrix} H_{Bw,k} \\ H_{Dw,k} \end{bmatrix}, \quad (3.75)$$

$$\tilde{H}_{v,k} = \begin{bmatrix} H_{Bv,k} \\ K_k H_{Dv,k} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \Psi_k \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \begin{bmatrix} 0 \\ I \end{bmatrix} \end{bmatrix} \begin{bmatrix} H_{Bv,k} \\ H_{Dv,k} \end{bmatrix}, \quad (3.76)$$

$$\tilde{G}_{x,k} = \begin{bmatrix} G_{x,k} & 0 \end{bmatrix}. \quad (3.77)$$

De maneira similar, considere novamente a equação (3.67), só que, desta vez, na sua formulação com o vetor auxiliar η_k . Nesta situação,

1. A incerteza na matriz dinâmica só envolve ΔA_k e ΔC_k .
2. A incerteza na matriz do ruído na dinâmica envolve ΔB_k e ΔD_k que, por sua vez, dependem das matrizes $\Delta B_{w,k}$, $\Delta D_{w,k}$, $\Delta B_{v,k}$ e $\Delta D_{v,k}$.

Assim, se o sistema aumentado incerto (3.67) possuir incertezas aditivas com norma limitada satisfazendo (3.54) e (3.58), então o modelo planta + estimador pode ser escrito como

$$\begin{bmatrix} x_{k+1} \\ \hat{x}_{k+1|k} \end{bmatrix} = \left(\begin{bmatrix} A_k & 0 \\ K_k C_k & \Phi_k - K_k C_k \end{bmatrix} + \begin{bmatrix} H_{A,k} \\ K_k H_{C,k} \end{bmatrix} F_{x,k} \begin{bmatrix} G_{x,k} & 0 \end{bmatrix} \right) \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix} + \left(\begin{bmatrix} B_k \\ K_k D_k \end{bmatrix} + \begin{bmatrix} H_{B,k} \\ K_k H_{D,k} \end{bmatrix} F_{\eta,k} G_{\eta,k} \right) \eta_k. \quad (3.78)$$

Uma solução mais especializada, sugerida em [110], consiste em considerar a seguinte estrutura para as incertezas

$$\begin{bmatrix} \Delta A_k & \Delta B_k \\ \Delta C_k & \Delta D_k \end{bmatrix} = \begin{bmatrix} H_{1,k} \\ H_{2,k} \end{bmatrix} F_k \begin{bmatrix} G_{x,k} & G_{\eta,k} \end{bmatrix}. \quad (3.79)$$

A escolha das incertezas da forma mostrada em (3.79) resultado no seguinte sistema aumentado

$$\tilde{x}_{k+1} = \left(\tilde{A}_k + \tilde{H}_k \tilde{F}_k \tilde{G}_{x,k} \right) \tilde{x}_k + \left(\tilde{B}_k + \tilde{H}_k \tilde{F}_k \tilde{G}_{\eta,k} \right) \eta_k, \quad (3.80)$$

em que $\tilde{F}_k = F_k$, $\tilde{G}_{\eta,k} = G_{\eta,k}$ e

$$\begin{aligned} \tilde{x}_k &= \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix}, & \tilde{A}_k &= \begin{bmatrix} A_k & 0 \\ K_k C_k & \Phi_k - K_k C_k \end{bmatrix}, \\ \tilde{B}_k &= \begin{bmatrix} B_k \\ K_k D_k \end{bmatrix}, & \tilde{H}_k &= \begin{bmatrix} H_{1,k} \\ K_k H_{2,k} \end{bmatrix}. \end{aligned} \quad (3.81)$$

3.5 COVARIÂNCIA DO SISTEMA AUMENTADO INCERTO

Conhecendo a variância do estado inicial, dada em (3.15), e considerando que \tilde{x}_k , w_k e v_k são não-correlacionados, conforme (3.16) e (3.23), utiliza-se o mesmo procedimento adotado no lema 2.3.2 e conclui-se que a matriz de covariância do sistema aumentado (3.69) é dada por

$$\tilde{P}_{0|-1} = \begin{bmatrix} X_0 & 0 \\ 0 & 0 \end{bmatrix}, \quad (3.82)$$

$$\begin{aligned} \tilde{P}_{k+1|k} &= \left(\tilde{A}_k + \tilde{H}_{x,k} \tilde{F}_{x,k} \tilde{G}_{x,k} \right) \tilde{P}_{k|k-1} \left(\tilde{A}_k + \tilde{H}_{x,k} \tilde{F}_{x,k} \tilde{G}_{x,k} \right)^T \\ &+ \left(\tilde{B}_{w,k} + \tilde{H}_{w,k} \tilde{F}_{w,k} \tilde{G}_{w,k} \right) W_k \left(\tilde{B}_{w,k} + \tilde{H}_{w,k} \tilde{F}_{w,k} \tilde{G}_{w,k} \right)^T \\ &+ \left(\tilde{B}_{v,k} + \tilde{H}_{v,k} \tilde{F}_{v,k} \tilde{G}_{v,k} \right) V_k \left(\tilde{B}_{v,k} + \tilde{H}_{v,k} \tilde{F}_{v,k} \tilde{G}_{v,k} \right)^T. \end{aligned} \quad (3.83)$$

Para o sistema aumentado em (3.80), tem-se que a condição inicial (3.83) e a seguinte recursão para a matriz de covariância

$$\begin{aligned}\tilde{P}_{k+1|k} &= \left(\tilde{A}_k + \tilde{H}_k \tilde{F}_k \tilde{G}_{x,k} \right) \tilde{P}_{k|k-1} \left(\tilde{A}_k + \tilde{H}_k \tilde{F}_k \tilde{G}_k \right)^T \\ &+ \left(\tilde{B}_k + \tilde{H}_k \tilde{F}_k \tilde{G}_{\eta,k} \right) N_k \left(\tilde{B}_k + \tilde{H}_k \tilde{F}_k \tilde{G}_{\eta,k} \right)^T,\end{aligned}\quad (3.84)$$

em que, como já fora definido em (3.23), $N_k := \begin{bmatrix} W_k & 0 \\ 0 & V_k \end{bmatrix} \geq 0$.

Nota 3.5.1 Para o caso em que se tem N_k cheia, nota-se que (3.83) é uma especialização de (3.84) para o caso em que W_k e V_k não são correlacionados.

3.6 MAJORANTE PARA A EVOLUÇÃO DA COVARIÂNCIA DO SISTEMA AUMENTADO

Esta é uma seção fundamental para a obtenção do estimador robusto. Será desenvolvido e apresentado um limite superior para a recursão da covariância do sistema aumentado. Por se tratar de um majorante para uma recursão, tem-se que demonstrar que o majorante é válido para todo tempo discreto k . Neste contexto, os Teoremas 3.6.1 e 3.6.2 apresentam um limite superior para a evolução recursiva da covariância do sistema aumentado na sua forma geral (3.84) e especializada (3.83). A demonstração do teorema é análoga para os dois casos. Ambos teoremas também utilizam os resultados da subseção 3.6.1, em que são analisados e desenvolvidos majorantes para a covariância de um sistema linear perturbado. Vale ressaltar, com base no desenvolvimento da seção 3.4, que o Teorema 3.6.2 fornece um majorante menos conservador para W_k e V_k não-correlacionados.

É interessante notar que a partir deste ponto começam a aparecer as “expressões corrigidas”² para o caso do estimador robusto. Em outras palavras, são expressões modificadas em relação ao estimador clássico com o objetivo de lidar com o sistema incerto. Este tipo de notação é bastante útil quando se podem utilizar os mesmos lemas do caso clássico.

Teorema 3.6.1 (Referente ao modelo (3.42)-(3.43) e a covariância aumentada (3.84).) Suponha $N_k \geq 0$ e considere uma sequência de matrizes simétricas $\{\tilde{P}_{k+1|k}\}$ gerada pela recursão

$$\tilde{P}_{0|-1} = \begin{bmatrix} X_0 & 0 \\ 0 & 0 \end{bmatrix} \geq 0, \quad (3.85)$$

$$\begin{aligned}\tilde{P}_{k+1|k} &= \left(\tilde{A}_k + \tilde{H}_{x,k} \tilde{F}_{x,k} \tilde{G}_{x,k} \right) \tilde{P}_{k|k-1} \left(\tilde{A}_k + \tilde{H}_{x,k} \tilde{F}_{x,k} \tilde{G}_{x,k} \right)^T \\ &+ \left(\tilde{B}_k + \tilde{H}_k \tilde{F}_k \tilde{G}_{\eta,k} \right) N_k \left(\tilde{B}_k + \tilde{H}_k \tilde{F}_k \tilde{G}_{\eta,k} \right)^T, \forall k > 0.\end{aligned}\quad (3.86)$$

²Será utilizada a notação com um c subscrito.

Para cada seqüência de matrizes simétricas $\{P_{k+1|k}\}$ da recursão

$$P_{0|-1} = \begin{bmatrix} X_0 & 0 \\ 0 & 0 \end{bmatrix}, \quad (3.87)$$

$$P_{k+1|k} = \tilde{A}_k P_{k|k-1} \tilde{A}_k^T + \tilde{A}_k P_{k|k-1} \tilde{G}_{x,k}^T \left(\alpha_{x,k}^{-1} I - \tilde{G}_{x,k} P_{k|k-1} \tilde{G}_{x,k}^T \right)^{-1} \tilde{G}_{x,k} P_{k|k-1} \tilde{A}_k^T \\ + \tilde{B}_k N_{c,k} \tilde{B}_k^T + \alpha_{x,k}^{-1} \tilde{H}_{x,k} \tilde{H}_{x,k}^T + \alpha_{\eta,k}^{-1} \tilde{H}_k \tilde{H}_k^T, \quad (3.88)$$

em que $N_{c,k} := N_k + N_k \tilde{G}_{\eta,k}^T \left(\alpha_{\eta,k}^{-1} I - \tilde{G}_{\eta,k} N_k \tilde{G}_{\eta,k}^T \right)^{-1} \tilde{G}_{\eta,k} N_k$ e em que $\alpha_{x,k}^{-1}$ e $\alpha_{\eta,k}^{-1} \in \mathbb{R}^+$ satisfazem

$$\alpha_{x,k}^{-1} I - \tilde{G}_{x,k} P_{k|k-1} \tilde{G}_{x,k}^T > 0, \quad (3.89)$$

$$\alpha_{\eta,k}^{-1} I - \tilde{G}_{\eta,k} N_k \tilde{G}_{\eta,k}^T > 0, \quad (3.90)$$

tem-se que $P_{k+1|k}$ é um limite superior de $\tilde{P}_{k+1|k}$, de forma que $0 \leq \tilde{P}_{k+1|k} \leq P_{k+1|k}$, $\forall k > 0$ e $\forall \tilde{F}_{x,k}, \tilde{F}_k$ tal que $\tilde{F}_{x,k}^T \tilde{F}_{x,k} \leq I$, $\tilde{F}_k^T \tilde{F}_k \leq I$.

Teorema 3.6.2 (Referente ao modelo (3.40)-(3.41) e a covariância aumentada (3.82).) Suponha $W_k \geq 0$, $V_k \geq 0$ e considere uma seqüência de matrizes simétricas $\{\tilde{P}_{k+1|k}\}$ gerada pela recursão

$$\tilde{P}_{0|-1} = \begin{bmatrix} X_0 & 0 \\ 0 & 0 \end{bmatrix} \geq 0, \quad (3.91)$$

$$\tilde{P}_{k+1|k} = \left(\tilde{A}_k + \tilde{H}_{x,k} \tilde{F}_{x,k} \tilde{G}_{x,k} \right) \tilde{P}_{k|k-1} \left(\tilde{A}_k + \tilde{H}_{x,k} \tilde{F}_{x,k} \tilde{G}_{x,k} \right)^T \\ + \left(\tilde{B}_{w,k} + \tilde{H}_{w,k} \tilde{F}_{w,k} \tilde{G}_{w,k} \right) W_k \left(\tilde{B}_{w,k} + \tilde{H}_{w,k} \tilde{F}_{w,k} \tilde{G}_{w,k} \right)^T \\ + \left(\tilde{B}_{v,k} + \tilde{H}_{v,k} \tilde{F}_{v,k} \tilde{G}_{v,k} \right) V_k \left(\tilde{B}_{v,k} + \tilde{H}_{v,k} \tilde{F}_{v,k} \tilde{G}_{v,k} \right)^T, \quad \forall k > 0. \quad (3.92)$$

Para cada seqüência de matrizes simétricas $\{P_{k+1|k}\}$ da recursão

$$P_{0|-1} = \begin{bmatrix} X_0 & 0 \\ 0 & 0 \end{bmatrix}, \quad (3.93)$$

$$P_{k+1|k} = \tilde{A}_k P_{k|k-1} \tilde{A}_k^T + \tilde{A}_k P_{k|k-1} \tilde{G}_{x,k}^T \left(\alpha_{x,k}^{-1} I - \tilde{G}_{x,k} P_{k|k-1} \tilde{G}_{x,k}^T \right)^{-1} \tilde{G}_{x,k} P_{k|k-1} \tilde{A}_k^T \\ + \tilde{B}_{w,k} W_{c,k} \tilde{B}_{w,k}^T + \tilde{B}_{v,k} V_{c,k} \tilde{B}_{v,k}^T + \alpha_{x,k}^{-1} \tilde{H}_{x,k} \tilde{H}_{x,k}^T + \alpha_{w,k}^{-1} \tilde{H}_{w,k} \tilde{H}_{w,k}^T + \alpha_{v,k}^{-1} \tilde{H}_{v,k} \tilde{H}_{v,k}^T, \quad (3.94)$$

em que

$$W_{c,k} := W_k + W_k \tilde{G}_{w,k}^T \left(\alpha_{w,k}^{-1} I - \tilde{G}_{w,k} W_k \tilde{G}_{w,k}^T \right)^{-1} \tilde{G}_{w,k} W_k, \quad (3.95)$$

$$V_{c,k} := V_k + V_k \tilde{G}_{v,k}^T \left(\alpha_{v,k}^{-1} I - \tilde{G}_{v,k} V_k \tilde{G}_{v,k}^T \right)^{-1} \tilde{G}_{v,k} V_k \quad (3.96)$$

e em que $\alpha_{x,k}^{-1}$, $\alpha_{w,k}^{-1}$ e $\alpha_{v,k}^{-1} \in \mathbb{R}^+$ satisfazem

$$\alpha_{x,k}^{-1} I - \tilde{G}_{x,k} P_{k|k-1} \tilde{G}_{x,k}^T > 0, \quad (3.97)$$

$$\alpha_{w,k}^{-1} I - \tilde{G}_{w,k} W_k \tilde{G}_{w,k}^T > 0, \quad (3.98)$$

$$\alpha_{v,k}^{-1} I - \tilde{G}_{v,k} V_k \tilde{G}_{v,k}^T > 0, \quad (3.99)$$

tem-se que $P_{k+1|k}$ é um limite superior de $\tilde{P}_{k+1|k}$, de forma que $0 \leq \tilde{P}_{k+1|k} \leq P_{k+1|k}$, $\forall k > 0$ e $\forall \tilde{F}_{x,k}, \tilde{F}_{w,k}, \tilde{F}_{v,k}$ tal que $\tilde{F}_{x,k}^T \tilde{F}_{x,k} \leq I$, $\tilde{F}_{w,k}^T \tilde{F}_{w,k} \leq I$, $\tilde{F}_{v,k}^T \tilde{F}_{v,k} \leq I$.

Prova A seguir, será mostrada a prova do Teorema 3.6.2. A prova do Teorema 3.6.1 é análoga, aplicando-se os mesmos procedimentos e argumentos. Inicialmente, defina

$$p_k \left(\tilde{P}_{k|k-1} \right) := \left(\tilde{A}_k + \tilde{H}_{x,k} \tilde{F}_{x,k} \tilde{G}_{x,k} \right) \tilde{P}_{k|k-1} \left(\tilde{A}_k + \tilde{H}_{x,k} \tilde{F}_{x,k} \tilde{G}_{x,k} \right)^T, \quad (3.100)$$

$$q_k := \left(\tilde{B}_{w,k} + \tilde{H}_{w,k} \tilde{F}_{w,k} \tilde{G}_{w,k} \right) W_k \left(\tilde{B}_{w,k} + \tilde{H}_{w,k} \tilde{F}_{w,k} \tilde{G}_{w,k} \right)^T, \quad (3.101)$$

$$r_k := \left(\tilde{B}_{v,k} + \tilde{H}_{v,k} \tilde{F}_{v,k} \tilde{G}_{v,k} \right) V_k \left(\tilde{B}_{v,k} + \tilde{H}_{v,k} \tilde{F}_{v,k} \tilde{G}_{v,k} \right)^T, \quad (3.102)$$

$$f_k \left(\tilde{P}_{k|k-1} \right) := p_k \left(\tilde{P}_{k|k-1} \right) + q_k + r_k. \quad (3.103)$$

Como $\tilde{P}_{k|k-1}$ é simétrica, aplicando o Lema 3.6.5, apresentado na seção seguinte, para $p_k \left(\tilde{P}_{k|k-1} \right)$, tem-se que, para todo $\alpha_{x,k} \in \mathbb{R}^+$ satisfazendo $\alpha_{x,k}^{-1} I - \tilde{G}_{x,k} \tilde{P}_{k|k-1} \tilde{G}_{x,k}^T > 0$, é válida a desigualdade

$$\begin{aligned} p_k \left(\tilde{P}_{k|k-1} \right) &\leq \tilde{A}_k \tilde{P}_{k|k-1} \tilde{A}_k^T + \tilde{A}_k \tilde{P}_{k|k-1} \tilde{G}_{x,k}^T \left(\alpha_{x,k}^{-1} I - \tilde{G}_{x,k} \tilde{P}_{k|k-1} \tilde{G}_{x,k}^T \right)^{-1} \tilde{G}_{x,k} \tilde{P}_{k|k-1} \tilde{A}_k^T \\ &+ \alpha_{x,k}^{-1} \tilde{H}_{x,k} \tilde{H}_{x,k}^T. \end{aligned} \quad (3.104)$$

A matriz W_k também é simétrica e, aplicando o Lema 3.6.5 em q_k , para todo $\alpha_{w,k} \in \mathbb{R}^+$ satisfazendo $\alpha_{w,k}^{-1} I - \tilde{G}_{v,k} W_k \tilde{G}_{v,k}^T > 0$, pode-se afirmar que

$$q_k \leq \tilde{B}_{w,k} W_k \tilde{B}_{w,k}^T + \tilde{B}_{w,k} W_k \tilde{G}_{v,k}^T \left(\alpha_{w,k}^{-1} I - \tilde{G}_{v,k} W_k \tilde{G}_{v,k}^T \right)^{-1} \tilde{G}_{v,k} W_k \tilde{B}_{w,k}^T + \alpha_{w,k}^{-1} \tilde{H}_{w,k} \tilde{H}_{w,k}^T. \quad (3.105)$$

Utilizando a definição (3.95), a desigualdade (3.105) pode ser reescrita como $q_k \leq \tilde{B}_{w,k} W_{c,k} \tilde{B}_{w,k}^T + \alpha_{w,k}^{-1} \tilde{H}_{w,k} \tilde{H}_{w,k}^T$. A matriz V_k também é simétrica e, aplicando o Lema 3.6.5 em r_k , para todo $\alpha_{v,k} \in \mathbb{R}^+$ satisfazendo $\alpha_{v,k}^{-1} I - \tilde{G}_{v,k} V_k \tilde{G}_{v,k}^T > 0$, demonstra-se que $r_k \leq \tilde{B}_{v,k} V_{c,k} \tilde{B}_{v,k}^T + \alpha_{v,k}^{-1} \tilde{H}_{v,k} \tilde{H}_{v,k}^T$, em que $V_{c,k}$ é definida em (3.96). Definindo

$$\begin{aligned} g_k \left(\tilde{P}_{k|k-1} \right) &:= \tilde{A}_k \tilde{P}_{k|k-1} \tilde{A}_k^T + \tilde{A}_k \tilde{P}_{k|k-1} \tilde{G}_{x,k}^T \left(\alpha_{x,k}^{-1} I - \tilde{G}_{x,k} \tilde{P}_{k|k-1} \tilde{G}_{x,k}^T \right)^{-1} \tilde{G}_{x,k} \tilde{P}_{k|k-1} \tilde{A}_k^T \\ &+ \alpha_{x,k}^{-1} \tilde{H}_{x,k} \tilde{H}_{x,k}^T + \tilde{B}_{v,k} V_{c,k} \tilde{B}_{v,k}^T + \alpha_{w,k}^{-1} \tilde{H}_{w,k} \tilde{H}_{w,k}^T + \tilde{B}_{v,k} V_{c,k} \tilde{B}_{v,k}^T \\ &+ \alpha_{v,k}^{-1} \tilde{H}_{v,k} \tilde{H}_{v,k}^T, \end{aligned} \quad (3.106)$$

então $f_k \left(\tilde{P}_{k|k-1} \right) = p_k \left(\tilde{P}_{k|k-1} \right) + q_k + r_k \leq g_k \left(\tilde{P}_{k|k-1} \right)$, ou seja, $f_k \left(\tilde{P}_{k|k-1} \right) \leq g_k \left(\tilde{P}_{k|k-1} \right)$ para todo $\tilde{P}_{k|k-1}$ e todo $\alpha_{x,k}, \alpha_{w,k}, \alpha_{v,k} \in \mathbb{R}^+$ satisfazendo (3.97)-(3.99). Por (3.103), a recursão (3.92) pode ser escrita como $\tilde{P}_{k+1|k} = f_k \left(\tilde{P}_{k|k-1} \right)$.

Nota-se que f_k é não-decrescente, *i.e.*, para X e $Y \in \mathbb{R}^{n \times n}$, se $X \leq Y$, então $f_k(X) \leq f_k(Y)$. Considere a definição $P_{k+1|k} := g_k \left(P_{k|k-1} \right)$. Suponha que, para algum k , valha a relação $\tilde{P}_{k|k-1} \leq P_{k|k-1}$. Portanto, observando o Lema 3.6.3, é correto afirmar que

$$\tilde{P}_{k+1|k} = f_k \left(\tilde{P}_{k|k-1} \right) \leq f_k \left(P_{k|k-1} \right) \leq g_k \left(P_{k|k-1} \right) = P_{k+1|k}, \quad (3.107)$$

para todo $P_{k|k-1}$ e todo $\alpha_{x,k}, \alpha_{w,k}, \alpha_{v,k} \in \mathbb{R}^+$ satisfazendo (3.97)-(3.99). Assim, definindo $P_{0|-1} := \tilde{P}_{0|-1}$, pode-se concluir que se $\left\{ \tilde{P}_{k+1|k} \right\}$ é gerado por $\tilde{P}_{k+1|k} = f_k \left(\tilde{P}_{k|k-1} \right)$ e $\left\{ P_{k+1|k} \right\}$ é gerada por $P_{k+1|k} = g_k \left(P_{k|k-1} \right)$, logo $\tilde{P}_{k|k-1} \leq P_{k|k-1}, \forall k \geq 0$. Desta forma, $P_{k+1|k}$ é um limite superior de $\tilde{P}_{k+1|k}$ para todo instante k . ■

Lema 3.6.3 (Estendido do Lema 3.2 de [90]). Seja $\mathbb{S}^{n \times n}$ o conjunto das matrizes reais simétricas de dimensão $n \times n$. Considere duas seqüências $\{f_k\}$ e $\{g_k\}$ de funções $f_k, g_k : \mathbb{S}^{n \times n} \rightarrow \mathbb{S}^{n \times n}$. Admitindo que, para todo k , f_k seja não-decrescente, para matrizes A e $B \in \mathbb{S}^{n \times n}$, se $A \leq B$, então $f_k(A) \leq f_k(B)$. Supondo que, para cada k , g_k forneça um majorante para f_k , é correto afirmar que $f_k(B) \leq g_k(B), \forall B \in \mathbb{S}^{n \times n}$. Portanto, para $A_0 = B_0 \in \mathbb{S}^{n \times n}$, as soluções $\{A_k\}$ e $\{B_k\}$ das seguintes equações de diferenças $A_{k+1} = f_k(A_k)$ e $B_{k+1} = g_k(B_k)$ satisfazem $A_k \leq B_k$ para todo $k \geq 0$.

Prova O resultado do lema segue por indução. É imediato notar que $A_0 \leq B_0$. Suponha que, para algum k , A_k e $B_k \in \mathbb{S}^{n \times n}$ satisfaçam $A_k \leq B_k$. Assim, $A_{k+1} = f_k(A_k) \leq f_k(B_k) \leq g_k(B_k) = B_{k+1}$. ■

Nota 3.6.4 Nesta versão do lema foi estendido o resultado de [90] que considera as hipóteses $A_0 = B_0 > 0$, $A < B$. Portanto, $f_k(A) \leq f_k(B)$ e $f_k(B) \leq g_k(B), \forall B > 0$. ■

3.6.1 Limitante Superior para a Covariância Incerta

Defina uma variável aleatória z , segundo o seguinte modelo linear perturbado $z = A_\Delta x$, em que x é uma variável aleatória de covariância P e A_Δ possui incertezas aditivas de norma limitada, pertencendo ao conjunto $\{A_\Delta = A + \Delta A = A + HFG : F^T F \leq I\}$, em que F é uma matriz de valor desconhecido. Portanto, a covariância de z é desconhecida, mas pertencente ao conjunto

$$\left\{ \text{cov}(z) = (A + HFG) P (A + HFG)^T : F^T F \leq I \right\}. \quad (3.108)$$

Neste contexto, almeja-se encontrar um limite superior para a covariância de z . A literatura sobre estimadores robustos fornece, dentre outras, duas sugestões de limites superiores que poderão ser escolhidos. Estes majorantes são apresentados nos lemas 3.6.5 e 3.6.8. Pelas provas, é de se esperar (não foi provado) que o majorante de 3.6.5 seja menos conservador do que o majorante de 3.6.8 e será o majorante escolhido para esta dissertação.

Lema 3.6.5 (Lema 2 de [111])³ Considere matrizes reais A, H, G, F e P com dimensões compatíveis. Suponha que $F^T F \leq I$ e que P seja uma matriz simétrica. Então, é correto afirmar que

$$(A + HFG) P (A + HFG)^T \leq APA^T + APG^T (\epsilon^{-1} I - GPG^T)^{-1} GPA^T + \epsilon^{-1} HH^T \quad (3.109)$$

para todo $\epsilon \in \mathbb{R}^+$, tal que $\epsilon^{-1} I - GPG^T > 0$.

Prova Seja, por definição, $f(F) := (A + HFG) P (A + HFG)^T$. Portanto,

$$f(F) = APA^T + APG^T F^T H^T + HFGPA^T + HFGPG^T F^T H^T. \quad (3.110)$$

Somando e subtraindo $\epsilon^{-1} HFF^T H^T$, obtém-se que

$$f(F) = APA^T + APG^T F^T H^T + HFGPA^T - HF (\epsilon^{-1} I - GPG^T) F^T H^T + \epsilon^{-1} HFF^T H^T. \quad (3.111)$$

Defina também

$$g(F) := APG^T F^T H^T + HFGPA^T - HF (\epsilon^{-1} I - GPG^T) F^T H^T, \quad (3.112)$$

$$h(F) := \epsilon^{-1} HFF^T H^T. \quad (3.113)$$

³Em [111], é suposto $P > 0$. Pelo desenvolvimento realizado aqui, o resultado vale para qualquer P simétrica.

Deve-se determinar majorantes g_{mF} e h_{mF} para $g(F)$ e $h(F)$, respectivamente, de tal forma que

$$f(F) = APA^T + g(F) + h(F) \leq APA^T + g_{mF} + h_{mF}. \quad (3.114)$$

Como $F^T F \leq I$, é imediato que $h(F) := \epsilon^{-1} H F F^T H^T \leq h_{mF} := \epsilon^{-1} H H^T$. Para todo ϵ , tal que $\epsilon^{-1} I - G P G^T > 0$, a Nota 3.6.6 fornece um majorante para $g(F)$. Com os majorantes g_{mF} e h_{mF} , segue

$$f(F) = APA^T + g(F) + h(F) \leq APA^T + g_{mF} + h_{mF}. \quad (3.115)$$

Logo, $f(F) \leq APA^T + APG^T (\epsilon^{-1} I - G P G^T)^{-1} G P A^T + \epsilon^{-1} H H^T$. ■

Nota 3.6.6 Para todo ϵ , tal que $\epsilon^{-1} I - G P G^T > 0$, um majorante para

$$g(F) := APG^T F^T H^T + H F G P A^T - H F (\epsilon^{-1} I - G P G^T) F^T H^T \quad (3.116)$$

é dado por $g_{mF} := APG^T (\epsilon^{-1} I - G P G^T)^{-1} G P A^T$.

Prova Somando e subtraindo $APG^T (\epsilon^{-1} I - G P G^T)^{-1} G P A^T$ em (3.116) e escolhendo ϵ tal que $\epsilon^{-1} I - G P G^T$ seja invertível, tem-se que

$$\begin{aligned} g(F) &= APG^T (\epsilon^{-1} I - G P G^T)^{-1} G P A^T \\ &\quad - \left(APG^T (\epsilon^{-1} I - G P G^T)^{-1} G P A^T - APG^T F^T H^T - H F G P A^T + H F (\epsilon^{-1} I - G P G^T) F^T H^T \right). \end{aligned} \quad (3.117)$$

Suponha X simétrica invertível. A Nota 3.6.7 fornece o seguinte resultado

$$B X^{-1} B^T - B A^T - A B^T + A X A^T = (A - B X^{-1}) X (A - B X^{-1})^T. \quad (3.118)$$

Sendo P simétrica, e aplicando o resultado da Nota 3.6.7, tem-se que

$$\begin{aligned} g(F) &= APG^T (\epsilon^{-1} I - G P G^T)^{-1} G P A^T \\ &\quad - \left(H F - APG^T (\epsilon^{-1} I - G P G^T)^{-1} \right) (\epsilon^{-1} I - G P G^T) \left(H F - APG^T (\epsilon^{-1} I - G P G^T)^{-1} \right)^T. \end{aligned} \quad (3.119)$$

Portanto, supondo $\epsilon^{-1} I - G P G^T > 0$, tem-se $g(F) \leq APG^T (\epsilon^{-1} I - G P G^T)^{-1} G P A^T$. ■

Nota 3.6.7 Suponha X simétrica invertível. Então, pode-se afirmar que

$$A B^T + B A^T - A X A^T = B X^{-1} B^T - (A - B X^{-1}) X (A - B X^{-1})^T. \quad (3.120)$$

Prova Como X é simétrica, segue que

$$(A - B X^{-1}) X (A - B X^{-1})^T = A X A^T - A B^T - B A^T + B X^{-1} B^T. \quad (3.121)$$

Assim, $A B^T + B A^T - A X A^T = B X^{-1} B^T - (A - B X^{-1}) X (A - B X^{-1})^T$. ■

Lema 3.6.8 (Lema apresentado em [90]). Dadas matrizes A, H, G e F com dimensões compatíveis e $F^T F \leq I$. Seja $\epsilon > 0$ um escalar positivo, então

$$(A + HFG)P(A + HFG)^T \leq APA^T + APG^T GPA^T + (1 + \epsilon)HH^T \quad (3.122)$$

para todo ϵ , tal que $\epsilon I - GPG^T \geq 0$.

Em particular, escolhendo-se $\epsilon = \sigma_{\max}(GPG^T)$, tem-se que o resultado de [90], ou seja,

$$(A + HFG)P(A + HFG)^T \leq APA^T + APG^T GPA^T + (1 + \sigma_{\max}(GPG^T))HH^T \quad (3.123)$$

Prova Seja, por definição, $f(F) := (A + HFG)P(A + HFG)^T$. Portanto,

$$f(F) = APA^T + APG^T F^T H^T + HFGPA^T + HFGPG^T F^T H^T. \quad (3.124)$$

Defina também

$$g(F) := APG^T F^T H^T + HFGPA^T, \quad (3.125)$$

$$h(F) := HFGPG^T F^T H^T. \quad (3.126)$$

Deve-se determinar majorantes g_{mF} e h_{mF} para $g(F)$ e $h(F)$, respectivamente, de tal forma que

$$f(F) = APA^T + g(F) + h(F) \leq APA^T + g_{mF} + h_{mF}. \quad (3.127)$$

Na Nota 3.6.9, demonstra-se que $h(F) := HFGPG^T F^T H^T \leq h_{mF} := \epsilon HH^T$. A Nota 3.6.10 fornece o seguinte resultado $g_{mF} := APG^T GPA^T + HH^T$. Com os majorantes g_{mF} e h_{mF} , segue que

$$\begin{aligned} f(F) &= APA^T + g(F) + h(F) \leq APA^T + g_{mF} + h_{mF} \\ \Rightarrow f(F) &\leq APA^T + APG^T GPA^T + HH^T + \epsilon HH^T \\ \Rightarrow f(F) &\leq APA^T + APG^T GPA^T + (1 + \epsilon)HH^T. \end{aligned} \quad (3.128)$$

■

Nota 3.6.9 Para todo ϵ , tal que $\epsilon I - FGPG^T F^T \geq 0$, tem-se que

$$h(F) := HFGPG^T F^T H^T \leq h_{mF} := \epsilon HH^T. \quad (3.129)$$

Nota-se que $\sigma_{\max}(FGPG^T F^T) \leq \sigma_{\max}(GPG^T)$. Assim, escolhendo $\epsilon \geq \sigma_{\max}(GPG^T)$, garante-se que $\epsilon I - FGPG^T F^T \geq 0$.

Prova Somando e subtraindo ϵHH^T , tem-se que

$$h(F) = \epsilon HH^T - H(\epsilon I - FGPG^T F^T)H^T. \quad (3.130)$$

Assim, para $\epsilon I - FGPG^T F^T \geq 0$, chega-se a

$$h(F) = \epsilon HH^T - H(\epsilon I - FGPG^T F^T)H^T \leq \epsilon HH^T. \quad (3.131)$$

■

Nota 3.6.10 Um majorante para $g(F) := APG^T F^T H^T + HFGPA^T$ é dado por

$$g_{mF} := APG^T GPA^T + HH^T. \quad (3.132)$$

Prova De fato, somando e subtraindo HH^T e $APG^T F^T FGPA^T$, tem-se que

$$\begin{aligned} g(F) &= APG^T F^T FGPA^T + HH^T - (H - APG^T F^T) (H - APG^T F^T)^T \\ &\leq APG^T F^T FGPA^T + HH^T. \end{aligned} \quad (3.133)$$

Nota-se que, como $F^T F \leq I$, então

$$APG^T F^T FGPA^T \leq APG^T GPA^T. \quad (3.134)$$

Assim, $g(F) \leq APG^T GPA^T + HH^T$. ■

3.7 PARTIÇÃO DO MAJORANTE

O objetivo desta seção é reescrever a matriz de covariância do sistema aumentado com a finalidade de se obter um majorante para a variância do erro de estimação.

Nota-se que o resultado do lema é bem parecido com o caso clássico, exceto por:

- Surgem expressões modificadas para as matrizes $P_{11,k+1|k}$, $P_{12,k+1|k}$ e $P_{22,k+1|k}$ no tratamento do caso incerto.
- As matrizes $S_{1,k}$, $S_{2,k}$ e $S_{3,k}$ agora dependem das expressões modificadas de $P_{11,k+1|k}$, $P_{12,k+1|k}$ e $P_{22,k+1|k}$.
- As matrizes de covariância também aparecem modificadas para o caso robusto.
- O aparecimento de termos aditivos $\Delta_{1,k}$, $\Delta_{2,k}$ e $\Delta_{3,k}$ decorrentes do processo de se garantir o limite superior da covariância.
- Os termos cruzados aparecem neste sistema pelo fato de os ruídos serem correlacionados, o que não fora considerado na demonstração do caso clássico. No entanto, não é difícil perceber que estes termos cruzados também apareceriam na demonstração do caso clássico, se tal hipótese fosse assumida.

Quando o modelo é conhecido precisamente, as expressões modificadas recaem no caso clássico, e o Lema 3.7.1 torna-se igual ao lema clássico apresentado no Lema 2.3.4.

Lema 3.7.1 Considere a matriz de covariância dada em (3.86) com as identificações mostradas em (3.70)-(3.77). Então, $P_{k+1|k}$ pode ser particionada como em (3.135) com as identificações (3.136)-(3.149).

$$P_{k+1|k} = \begin{bmatrix} P_{11,k+1|k} & P_{12,k+1|k} \\ P_{12,k+1|k}^T & P_{22,k+1|k} \end{bmatrix} \geq 0, \quad (3.135)$$

em que

$$P_{11,k+1|k} = A_k P_{11,c} A_k^T + B_k N_{c,k} B_k^T + \Delta_{3,k}, \quad (3.136)$$

$$P_{12,k+1|k} = A_k P_{12,c} \Phi_k^T + (A_k S_{1,k} C_k^T + B_k N_{c,k} D_k + \Delta_{1,k}) K_k^T, \quad (3.137)$$

$$\begin{aligned} P_{22,k+1|k} &= \Phi_k P_{22,c} \Phi_k^T + K_k C_k S_{2,k} \Phi_k^T + \Phi_k S_{2,k}^T C_k^T K_k^T \\ &+ K_k (C_k S_{3,k} C_k^T + D_{v,k} N_{c,k} D_{v,k}^T + \Delta_{2,k}) K_k^T, \end{aligned} \quad (3.138)$$

em que se definem os termos

$$N_{c,k} := \begin{bmatrix} W_{c,k} & 0 \\ 0 & V_{c,k} \end{bmatrix}, \quad (3.139)$$

$$\Delta_{1,k} := \alpha_{x,k}^{-1} H_{A,k} H_{C,k}^T + \alpha_{w,k}^{-1} H_{Bw,k} H_{Dw,k}^T + \alpha_{v,k}^{-1} H_{Bv,k} H_{Dv,k}^T, \quad (3.140)$$

$$\Delta_{2,k} := \alpha_{x,k}^{-1} H_{C,k} H_{C,k}^T + \alpha_{w,k}^{-1} H_{Dw,k} H_{Dw,k}^T + \alpha_{v,k}^{-1} H_{Dv,k} H_{Dv,k}^T, \quad (3.141)$$

$$\Delta_{3,k} := \alpha_{x,k}^{-1} H_{A,k} H_{A,k}^T + \alpha_{w,k}^{-1} H_{Bw,k} H_{Bw,k}^T + \alpha_{v,k}^{-1} H_{Bv,k} H_{Bv,k}^T, \quad (3.142)$$

$$M_k := G_{x,k}^T \left(\alpha_{x,k}^{-1} I - G_{x,k} P_{11,k|k-1} G_{x,k}^T \right)^{-1} G_{x,k}, \quad (3.143)$$

$$P_{11c,k} := P_{11,k|k-1} + P_{11,k|k-1} M_k P_{11,k|k-1}, \quad (3.144)$$

$$P_{12c,k} := P_{12,k|k-1} + P_{11,k|k-1} M_k P_{12,k|k-1}, \quad (3.145)$$

$$P_{22c,k} := P_{22,k|k-1} + P_{12,k|k-1}^T M_k P_{12,k|k-1}, \quad (3.146)$$

$$S_{1,k} := P_{11c,k} - P_{12c,k}, \quad (3.147)$$

$$S_{2,k} := P_{12c,k} - P_{22c,k}, \quad (3.148)$$

$$S_{3,k} := S_{1,k} - S_{2,k}^T. \quad (3.149)$$

Prova Seguindo o mesmo procedimento do Lema 2.3.4 para o preditor clássico, e utilizando os resultados dos lemas anteriores, substituem-se as identificações (3.70)-(3.77) na matriz de covariância (3.86), adotando-se a partição (3.135). ■

3.8 MAJORANTE DA COVARIÂNCIA DO ERRO DE ESTIMAÇÃO

O objetivo desta seção é encontrar um limite superior para a covariância do erro de estimação a partir do limite superior da covariância do sistema aumentado. A diferença para o caso clássico é que neste busca-se a covariância do erro propriamente dita a partir da covariância do sistema aumentado. No caso robusto, trabalha-se com seu limite superior.

Neste contexto, o Lema 3.8.1 fornece um majorante para a covariância do erro de estimação a partir de um majorante da covariância do sistema aumentado. Utilizando a partição do sistema aumentado no resultado do Lema 3.8.1, mostra-se diretamente que o limite superior da covariância do erro de estimação é dado, tal qual no caso clássico, por

$$\bar{P}_{k+1|k} = P_{11,k+1|k} - P_{12,k+1|k} - P_{12,k+1|k}^T + P_{22,k+1|k}. \quad (3.150)$$

Nota-se que, tendo o mesmo formato do caso clássico, pode-se aplicar o mesmo Lema 2.3.7 utilizado

para o preditor usual, resultando em

$$\begin{aligned}
\bar{P}_{k+1|k} &= (A_k - K_k C_k) P_{11,c} (A_k - K_k C_k)^T - (A_k - K_k C_k) P_{12,c} (\Phi_k - K_k C_k)^T \\
&- (\Phi_k - K_k C_k) P_{12,c}^T (A_k - K_k C_k)^T + (\Phi_k - K_k C_k) P_{22,c1} (\Phi_k - K_k C_k)^T \\
&+ (B_{w,k} - K_k D_{w,k}) W_{c,k} (B_{w,k} - K_k D_{w,k})^T \\
&+ (B_{v,k} - K_k D_{v,k}) V_{c,k} (B_{v,k} - K_k D_{v,k})^T \\
&+ \alpha_{x,k}^{-1} (H_{A,k} - K_k H_{C,k}) (H_{A,k} - K_k H_{C,k})^T \\
&+ \alpha_{w,k}^{-1} (H_{Bw,k} - K_k H_{Dw,k}) (H_{Bw,k} - K_k H_{Dw,k})^T \\
&+ \alpha_{v,k}^{-1} (H_{Bv,k} - K_k H_{Dv,k}) (H_{Bv,k} - K_k H_{Dv,k})^T
\end{aligned} \tag{3.151}$$

com $\bar{P}_{0|-1} = X_0$. Novamente, é possível visualizar que o majorante (3.151) recai no caso clássico, quando não há incertezas no modelo.

Lema 3.8.1 *Seja $P_{k+1|k}$ um limite superior de $\tilde{P}_{k+1|k}$ da forma*

$$P_{k+1|k} \geq \tilde{P}_{k+1|k} \geq 0, \forall k, \tag{3.152}$$

em que $\tilde{P}_{k+1|k}$ é a covariância do estado aumentado

$$\tilde{P}_{k+1|k} := E \left\{ \begin{bmatrix} x_{k+1} - E \{x_{k+1}\} \\ \hat{x}_{k+1|k} - E \{\hat{x}_{k+1|k}\} \end{bmatrix} \begin{bmatrix} x_{k+1} - E \{x_{k+1}\} \\ \hat{x}_{k+1|k} - E \{\hat{x}_{k+1|k}\} \end{bmatrix}^T \right\}. \tag{3.153}$$

Considerando a definição $e_{k+1|k} := x_{k+1|k} - \hat{x}_{k+1|k}$, então

$$\begin{aligned}
\text{cov}(e_{k+1|k}) &:= E \left\{ (e_{k+1|k} - E \{e_{k+1|k}\}) (e_{k+1|k} - E \{e_{k+1|k}\})^T \right\} \\
&\leq \bar{P}_{k+1|k} := \begin{bmatrix} I & -I \end{bmatrix} P_{k+1|k} \begin{bmatrix} I \\ -I \end{bmatrix}^T,
\end{aligned} \tag{3.154}$$

Prova Substituindo (3.153) em (3.152),

$$\tilde{P}_{k+1|k} \leq P_{k+1|k} \Rightarrow E \left\{ \begin{bmatrix} x_{k+1} - E \{x_{k+1}\} \\ \hat{x}_{k+1|k} - E \{\hat{x}_{k+1|k}\} \end{bmatrix} \begin{bmatrix} x_{k+1} - E \{x_{k+1}\} \\ \hat{x}_{k+1|k} - E \{\hat{x}_{k+1|k}\} \end{bmatrix}^T \right\} \leq P_{k+1|k}. \tag{3.155}$$

O Lema I.6.1, do anexo I.6, permite escrever (3.155) da forma

$$\begin{bmatrix} I & -I \end{bmatrix} E \left\{ \begin{bmatrix} x_{k+1} - E \{x_{k+1}\} \\ \hat{x}_{k+1|k} - E \{\hat{x}_{k+1|k}\} \end{bmatrix} \begin{bmatrix} x_{k+1} - E \{x_{k+1}\} \\ \hat{x}_{k+1|k} - E \{\hat{x}_{k+1|k}\} \end{bmatrix}^T \right\} \begin{bmatrix} I \\ -I \end{bmatrix} \leq \begin{bmatrix} I & -I \end{bmatrix} P_{k|k-1} \begin{bmatrix} I \\ -I \end{bmatrix}. \tag{3.156}$$

Como as matrizes $\begin{bmatrix} I & -I \end{bmatrix}$ e $\begin{bmatrix} I & -I \end{bmatrix}^T$ são constantes,

$$\begin{aligned}
&E \left\{ (x_{k+1} - \hat{x}_{k+1|k} - E \{x_{k+1} - \hat{x}_{k+1|k}\}) (x_{k+1} - \hat{x}_{k+1|k} - E \{x_{k+1} - \hat{x}_{k+1|k}\})^T \right\} \\
&\leq \begin{bmatrix} I & -I \end{bmatrix} P_{k+1|k} \begin{bmatrix} I \\ -I \end{bmatrix}^T.
\end{aligned} \tag{3.157}$$

Sendo a covariância do erro de estimação definida como

$$\text{cov}(e_{k+1|k}) := E \left\{ (e_{k+1|k} - E \{e_{k+1|k}\}) (e_{k+1|k} - E \{e_{k+1|k}\})^T \right\}, \quad (3.158)$$

em que $e_{k+1|k} := x_{k+1|k} - \hat{x}_{k+1|k}$, se for definido

$$\bar{P}_{k+1|k} := \begin{bmatrix} I & -I \\ & -I \end{bmatrix} P_{k+1|k} \begin{bmatrix} I \\ -I \end{bmatrix}, \quad (3.159)$$

então $\text{cov}(x_{k+1|k} - \hat{x}_{k+1|k}) \leq \bar{P}_{k+1|k}$. Alternativamente, utilizando o Lema I.6.2, do anexo I.6,

$$E \left\{ (x_k - \hat{x}_{k|k-1} - E \{x_k - \hat{x}_{k|k-1}\})^T (x_k - \hat{x}_{k|k-1} - E \{x_k - \hat{x}_{k|k-1}\}) \right\} \leq \text{tr}(\bar{P}_{k|k-1}). \quad (3.160)$$

■

3.9 MINIMIZAÇÃO DA COVARIÂNCIA DO ERRO DE ESTIMAÇÃO

Esta seção é inteiramente análoga ao caso clássico, bastando utilizar as expressões modificadas no lugar das expressões clássicas e incluir apropriadamente os termos aditivos provenientes do processo de majoração. Portanto, utilizando os lemas 2.3.8 e 2.3.11, obtêm-se os seguintes pontos de mínimo

$$K_k^* = (A_k S_k C_k^T + \Psi_{1,k}) (C_k S_k C_k^T + \Psi_{2,k})^{-1}, \quad (3.161)$$

$$\Phi_k^* = A_k + (A_k - K_k^* C_k) (P_{12c,k} P_{22c,k}^{-1} - I), \quad (3.162)$$

em que

$$S_k := P_{11c,k} - P_{12c,k} P_{22c,k}^{-1} P_{12c,k}^T, \quad (3.163)$$

$$\Psi_{1,k} := B_{w,k} W_{c,k} D_{w,k}^T + B_{v,k} V_{c,k} D_{v,k}^T + \Delta_{1,k}, \quad (3.164)$$

$$\Psi_{2,k} := D_{w,k} W_{c,k} D_{w,k}^T + D_{v,k} V_{c,k} D_{v,k}^T + \Delta_{2,k}. \quad (3.165)$$

Mais uma vez, a notação utilizada permite ao leitor perceber que os parâmetros do preditor robusto recaem no algoritmo clássico, caso não haja incertezas no modelo.

3.10 BUSCANDO SIMPLIFICAÇÕES

Até este ponto, já são conhecidas as expressões dos parâmetros do preditor que minimizam o limite superior escolhido para a covariância do erro de estimação. Substituindo os parâmetros do preditor (3.161)-(3.162) nos blocos de $P_{k+1|k}$ e utilizando os resultados dos lemas 2.3.13 e 2.3.14, mostra-se que

$$P_{12,k+1|k} = P_{12,k+1|k}^T = P_{22,k+1|k} = (A_k S_k C_k^T + \Psi_{1,k}) (C_k S_k C_k^T + \Psi_{2,k})^{-1} (A_k S_k C_k^T + \Psi_{1,k})^T + A_k P_{12c,k} P_{22c,k}^{-1} P_{12c,k}^T A_k^T. \quad (3.166)$$

Por meio dos mesmos argumentos adotados para o estimador clássico, conclui-se que 3.166 é válida para qualquer $k \geq 0$. Pode-se demonstrar, a partir de (3.166), a seguinte simplificação

$$S_k = \bar{P}_{c,k|k-1} := \bar{P}_{k|k-1} + \bar{P}_{k|k-1} G_{x,k}^T \left(\alpha_{x,k}^{-1} I - G_{x,k} \bar{P}_{k|k-1} G_{x,k}^T \right)^{-1} G_{x,k} \bar{P}_{k|k-1}. \quad (3.167)$$

De fato, a matriz de covariância do erro de estimação também possui sua expressão modificada para lidar com o sistema incerteza. Mais uma vez, é imediato notar que esta expressão recai no caso clássico para modelos sem incertezas.

A condição inicial sobre $\alpha_{x,k}$, apresentada em (3.89), também pode ser simplificada, de forma que

$$\alpha_{x,k}^{-1}I - G_{x,k}\bar{P}_{k|k-1}G_{x,k}^T > 0. \quad (3.168)$$

3.11 REESCREVENDO AS EXPRESSÕES

A partir das simplificações obtidas na seção 3.10, é possível obter expressões mais simples para os parâmetros do preditor, tais quais

$$\Phi_k^* = A_k + (A_k - K_k C_k) \bar{P}_{k|k-1} G_{x,k}^T \left(\alpha_{x,k}^{-1} I - G_{x,k} \bar{P}_{k|k-1} G_{x,k}^T \right)^{-1} G_{x,k}, \quad (3.169)$$

$$K_k^* = (A_k \bar{P}_{c,k|k-1} C_k^T + B_k N_{c,k} D_k^T + \Delta_{1,k}) (C_k \bar{P}_{c,k|k-1} C_k^T + D_k N_{c,k} D_k^T + \Delta_{2,k})^{-1} \quad (3.170)$$

A recursão da Riccati também pode ser simplificada. Substituindo Φ_k^* ,

$$\begin{aligned} \bar{P}_{k+1|k} &= (A_k - K_k^* C_k) \bar{P}_{c,k|k-1} (A_k - K_k^* C_k)^T + (B_{w,k} - K_k^* D_{w,k}) W_{c,k} (B_{w,k} - K_k^* D_{w,k})^T \\ &+ (B_{v,k} - K_k^* D_{v,k}) V_{c,k} (B_{v,k} - K_k^* D_{v,k})^T + \alpha_{x,k}^{-1} (H_{A,k} - K_k^* H_{C,k}) (H_{A,k} - K_k^* H_{C,k})^T \\ &+ \alpha_{w,k}^{-1} (H_{Bw,k} - K_k^* H_{Dw,k}) (H_{Bw,k} - K_k^* H_{Dw,k})^T \\ &+ \alpha_{v,k}^{-1} (H_{Bv,k} - K_k^* H_{Dv,k}) (H_{Bv,k} - K_k^* H_{Dv,k})^T. \end{aligned} \quad (3.171)$$

Substituindo a expressão de K_k^* em $\bar{P}_{k+1|k}$,

$$\begin{aligned} \bar{P}_{k+1|k} &= A_k \bar{P}_{c,k|k-1} A_k^T + B_k N_{c,k} B_k^T + \Delta_{3,k} \\ &- (A_k \bar{P}_{c,k|k-1} C_k^T + \Delta_{1,k}) (C_k \bar{P}_{c,k|k-1} C_k^T + D_k N_{c,k} D_k^T + \Delta_{2,k})^{-1} (A_k \bar{P}_{c,k|k-1} C_k^T + \Delta_{1,k})^T. \end{aligned} \quad (3.172)$$

Portanto, o preditor é dado por

$$\hat{x}_{k+1|k} = \Phi_k \hat{x}_{k|k-1} + K_k (y_k - C_k \hat{x}_{k|k-1}). \quad (3.173)$$

De outro modo, substituindo (3.169) em (3.173),

$$\hat{x}_{k+1|k} = A_{c,k} \hat{x}_{k|k-1} + K_k (y_k - C_{c,k} \hat{x}_{k|k-1}). \quad (3.174)$$

em que

$$A_{c,k} := A_k + A_k \bar{P}_{k|k-1} G_{x,k}^T \left(\alpha_{x,k}^{-1} I - G_{x,k} \bar{P}_{k|k-1} G_{x,k}^T \right)^{-1} G_{x,k}, \quad (3.175)$$

$$C_{c,k} := C_k + C_k \bar{P}_{k|k-1} G_{x,k}^T \left(\alpha_{x,k}^{-1} I - G_{x,k} \bar{P}_{k|k-1} G_{x,k}^T \right)^{-1} G_{x,k}. \quad (3.176)$$

Novamente, nota-se que o preditor clássico é um caso particular do preditor robusto para sistemas sem incertezas. O preditor robusto é apresentado na Tabela 3.3.

Tabela 3.3: Preditor Linear Robusto.

Passo 0 (Condições iniciais): $\hat{x}_{0|-1} = \bar{x}_0$ e $\bar{P}_{0|-1} = X_0$.

Passo 1: Obter escalares $\alpha_{x,k}$, $\alpha_{w,k}$ e $\alpha_{v,k}$ que satisfaçam

$$\alpha_{w,k}^{-1}I - G_{w,k}W_kG_{w,k}^T > 0,$$

$$\alpha_{v,k}^{-1}I - G_{v,k}V_kG_{v,k}^T > 0,$$

$$\alpha_{x,k}^{-1}I - G_{x,k}\bar{P}_{k|k-1}G_{x,k}^T > 0.$$

A seguir, definir

$$\Delta_{1,k} = \alpha_{x,k}^{-1}H_{A,k}H_{C,k}^T + \alpha_{w,k}^{-1}H_{Bw,k}H_{Dw,k}^T + \alpha_{v,k}^{-1}H_{Bv,k}H_{Dv,k}^T,$$

$$\Delta_{2,k} = \alpha_{x,k}^{-1}H_{C,k}H_{C,k}^T + \alpha_{w,k}^{-1}H_{Dw,k}H_{Dw,k}^T + \alpha_{v,k}^{-1}H_{Dv,k}H_{Dv,k}^T,$$

$$\Delta_{3,k} = \alpha_{x,k}^{-1}H_{A,k}H_{A,k}^T + \alpha_{w,k}^{-1}H_{Bw,k}H_{Bw,k}^T + \alpha_{v,k}^{-1}H_{Bv,k}H_{Bv,k}^T.$$

Passo 2: Calcular as correções, devido à presença das incertezas

$$\bar{P}_{c,k|k-1} = \bar{P}_{k|k-1} + \bar{P}_{k|k-1}G_{x,k}^T \left(\alpha_{x,k}^{-1}I - G_{x,k}\bar{P}_{k|k-1}G_{x,k}^T \right)^{-1} G_{x,k}\bar{P}_{k|k-1},$$

$$W_{c,k} = W_k + W_kG_{w,k}^T \left(\alpha_{w,k}^{-1}I - G_{w,k}W_kG_{w,k}^T \right)^{-1} G_{w,k}W_k,$$

$$V_{c,k} = V_k + V_kG_{v,k}^T \left(\alpha_{v,k}^{-1}I - G_{v,k}V_kG_{v,k}^T \right)^{-1} G_{v,k}V_k.$$

Passo 3: Definir matrizes aumentadas

$$B_k = \begin{bmatrix} B_{w,k} & B_{v,k} \end{bmatrix}, D_k = \begin{bmatrix} D_{w,k} & D_{v,k} \end{bmatrix}, U_{c,k} = \text{diag} \{W_{c,k}, V_{c,k}\}.$$

Passo 4: Calcular os parâmetros do preditor

$$K_k = (A_k\bar{P}_{c,k|k-1}C_k^T + B_kU_{c,k}D_k^T + \Delta_{1,k}) (C_k\bar{P}_{c,k|k-1}C_k^T + D_kU_{c,k}D_k^T + \Delta_{2,k})^{-1},$$

$$\Phi_k = A_k + (A_k - K_kC_k)\bar{P}_{k|k-1}G_{x,k}^T \left(\alpha_{x,k}^{-1}I - G_{x,k}\bar{P}_{k|k-1}G_{x,k}^T \right)^{-1} G_{x,k}.$$

Passo 5: Atualizar $\{\hat{x}_{k+1|k}\}$ e $\{\bar{P}_{k+1|k}\}$ como

$$\hat{x}_{k+1|k} = \Phi_k\hat{x}_{k|k-1} + K_k(y_k - C_k\hat{x}_{k|k-1}),$$

$$\begin{aligned} \bar{P}_{k+1|k} &= A_k\bar{P}_{c,k|k-1}A_k^T + B_kU_{c,k}B_k^T + \Delta_{3,k} \\ &\quad - (A_k\bar{P}_{c,k|k-1}C_k^T + \Delta_{1,k}) \\ &\quad \times (C_k\bar{P}_{c,k|k-1}C_k^T + D_kU_{c,k}D_k^T + \Delta_{2,k})^{-1} \\ &\quad \times (A_k\bar{P}_{c,k|k-1}C_k^T + \Delta_{1,k})^T. \end{aligned}$$

4 ESTIMADORES ROBUSTOS NÃO-LINEARES

“Depois da meia-noite, a gente vira abóbora. Fica só falando abobrinha...” João Y. Ishihara (13/03/2008), na véspera do deadline do Conference on Decision and Control 2008. Neste capítulo, será apresentado o estimador robusto de Kalman para modelos não-lineares sujeitos a incertezas.

4.1 ESTIMAÇÃO NÃO-LINEAR ROBUSTA

Assim como a estimação linear robusta, a estimação não-linear robusta tem sido pesquisada pela comunidade científica desde o início da década de 70. Bem como no caso linear, os estimadores não-lineares também possuem seu desempenho associado à precisão dos valores do modelo. Em aplicações reais, por exemplo, é comum encontrar sistemas em que as propriedades estatísticas dos ruídos não são conhecidas com certeza. Neste sentido, surgiu a chamada *Estimação Não-Linear Robusta*.

Analogamente ao caso linear, várias abordagens têm sido empregadas para a estimação não-linear robusta. Em [112], utiliza-se a abordagem de IQCs para o desenvolvimento de um estimador robusto com custo garantido. Estimadores robustos foram recentemente propostos em [113] para sistemas incertos representados no formato de *Nonlinear Fractional Transformation* (NFT), que é uma extensão da representação clássica LFT. A partir de uma formulação de LMIs, os autores desenvolveram o problema para filtros \mathcal{H}_2 , \mathcal{H}_∞ e $\mathcal{H}_2/\mathcal{H}_\infty$.

Diversas aplicações têm surgido para estimadores não-lineares robustos. Apenas para citar algumas aplicações, em [114] estuda-se o REKF para a estimação da posição de dispositivos móveis, notadamente por causa de não-linearidades elevadas advindas ao multipercurso¹ e sinais sem linha de visada. O estudo de alinhamento em voo utilizando REKF é realizado em [115]. Outro projeto robusto de um FKE para tempo discreto é desenvolvido por meio da aplicação de métodos de projetos lineares robustos baseados na minimização \mathcal{H}_∞ [116]. Também foram propostos REKFs para a estimação de parâmetros e estados de uma aeronave [117], sistema inercial de navegação do tipo *strapdown* (SDINS) [118] e estimação de altitude de um usuário móvel para predição de célula com o objetivo de melhorar a confiabilidade e a eficiência da banda de um sistema móvel de telecomunicações [119]. A estimação de estados de um sistema incerto e integrado de navegação INS/GPS baseada em estimação robusta é proposta em [120]. Uma formulação do REKF baseada em *Singular Value Decomposition* (SVD) da matriz de covariância/informação é apresentada em [117] com o objetivo de aumentar a estabilidade numérica e a eficiência computacional do FKE. Todavia, o projeto não considera incertezas no sistema.

¹Em telecomunicações móveis, multipercurso, ou *multipath*, é o fenômeno de propagação em que sinais de rádio chegam até a antena receptora por dois ou mais caminhos.

4.2 OBJETIVOS

O estimador estudado nesta seção é uma extensão do filtro robusto apresentado em [64]. De fato, o filtro em [64] é uma versão para tempo discreto do filtro em tempo contínuo desenvolvido em [121]. Considerando os sistemas de tempo contínuo, foram apresentados resultados tanto para um limite superior da covariância do erro de estimação [121] como para um limite inferior [122]. Motivados pelo estudo de estimadores lineares, esta seção propõe uma extensão do estimador não-linear robusto com custo garantido para sistemas discretos, incluindo ruídos com média, covariância e covariância cruzada incertas. Também é feita uma generalização do filtro para condições de Lipschitz com média não-nula. Tomando por base a estrutura do filtro de Kalman clássico para sistemas com entradas exógenas, supõe-se que o estimador robusto possua a seguinte estrutura

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (y_k - h_k(\hat{x}_{k|k-1}, u_k)), \quad (4.1)$$

$$\hat{x}_{k+1|k} = f_k(\hat{x}_{k|k}, u_k), \quad (4.2)$$

com condição inicial $\hat{x}_{0|-1} = \bar{x}_0$.

Deseja-se projetar um filtro de horizonte finito que garanta limites superiores para as variâncias dos erros *a priori* e *a posteriori*, ou seja, almeja-se encontrar uma sequência de matrizes positivas definidas $\{\bar{P}_{k+1|k}\}$ e $\{\bar{P}_{k|k}\}$ que, para todo o conjunto de incertezas permitidas, e para cada instante de tempo, satisfaçam

$$\text{cov} \{x_{k+1} - \hat{x}_{k+1|k}\} \leq \bar{P}_{k+1|k}, \quad (4.3)$$

$$\text{cov} \{x_k - \hat{x}_{k|k}\} \leq \bar{P}_{k|k}. \quad (4.4)$$

O ganho K_k é variante no tempo e deve ser determinado com a finalidade de minimizar os limites $\bar{P}_{k|k}$ e $\bar{P}_{k+1|k}$ no sentido da norma da matriz, obtendo um filtro de mínima variância no erro de estimação nos estágios de predição e correção.

É importante notar a relevância deste algoritmo de custo garantido para o sistema não-linear, dadas as dificuldades de se obterem informações satisfatórias a respeito do desempenho dos estimadores não-lineares clássicos. No caso robusto apresentado nesta dissertação, garante-se o desempenho do estimador.

4.3 MODELAGEM DO SISTEMA NÃO-LINEAR INCERTO

Seja um sistema não-linear em tempo discreto descrito por

$$x_{k+1} = f_k(x_k, u_k) + \tilde{w}_k, \quad (4.5)$$

$$y_k = h_k(x_k, u_k) + \tilde{v}_k, \quad (4.6)$$

lembrando que $\tilde{w}_k \in \mathbb{R}^{n_w}$ e $\tilde{v}_k \in \mathbb{R}^{n_v}$ são os sinais de ruído de processo e de medida respectivamente com incertezas nas médias, covariâncias e covariâncias cruzadas. Admitem-se funções não-lineares $f_k(\bullet) : \mathbb{R}_x^n \rightarrow \mathbb{R}_x^n$ e $h_k(\bullet) : \mathbb{R}_x^n \rightarrow \mathbb{R}_y^n$ diferenciáveis [64]. O subscrito em $f_k(\bullet)$ é incluído para representar uma possível dependência do tempo.

A classe de sistemas considerada aqui é uma subclasse de (4.5)-(4.6) em que as funções não-lineares $f_k(\bullet) : \mathfrak{R}^{n_x \times n_u} \rightarrow \mathfrak{R}_x^n$ e $h_k(\bullet) : \mathfrak{R}^{n_x \times n_u} \rightarrow \mathfrak{R}_y^n$ são uniformemente Lipschitz e variantes no tempo em x_k e u_k . A função $f_k(\bullet)$ descreve a evolução dos estados e $h_k(\bullet)$ modela as observações realizadas a partir dos estados. Neste ponto, impõem-se restrições nas funções não-lineares com o objetivo de limitá-las de certa forma. Tais restrições podem ser feitas assumindo limites superiores no erro de estimação. De maneira a quantificar o grau de não-linearidade no sistema, assume-se que o sistema (4.5)-(4.6) satisfaça os seguintes limites cônicos:

$$\|p_k - E\{p_k\}\| \leq \|\Delta_{a,k}(\mathbf{d}_{x,k} - E\{\mathbf{d}_{x,k}\})\| + \|\Delta_{b,k}(\mathbf{d}_{u,k} - E\{\mathbf{d}_{u,k}\})\|, \quad (4.7)$$

$$\|q_k - E\{q_k\}\| \leq \|\Delta_{c,k}(\mathbf{d}_{x,k} - E\{\mathbf{d}_{x,k}\})\| + \|\Delta_{d,k}(\mathbf{d}_{u,k} - E\{\mathbf{d}_{u,k}\})\|, \quad (4.8)$$

em que

$$p_k := f_k(\mathbf{x}_k + \mathbf{d}_{x,k}, u_k + \mathbf{d}_{u,k}) - f_k(\mathbf{x}_k, u_k) - A_k \mathbf{d}_{x,k} - B_k \mathbf{d}_{u,k}, \quad (4.9)$$

$$q_k := h_k(\mathbf{x}_k + \mathbf{d}_{x,k}, u_k + \mathbf{d}_{u,k}) - h_k(\mathbf{x}_k, u_k) - C_k \mathbf{d}_{x,k} - D_k \mathbf{d}_{u,k}, \quad (4.10)$$

para todo \mathbf{x}_k , $\mathbf{d}_{x,k}$ and $\mathbf{d}_{u,k}$ e para algumas matrizes $\Delta_{a,k}$, $\Delta_{b,k}$, $\Delta_{c,k}$ and $\Delta_{d,k}$. As matrizes $A_k \in \mathfrak{R}^{n_x \times n_x}$, $B_k \in \mathfrak{R}^{n_x \times n_u}$, $C_k \in \mathfrak{R}^{n_y \times n_x}$ e $D_k \in \mathfrak{R}^{n_y \times n_u}$ parametrizam o sistema nominal não-linear (4.5)-(4.6). Estes limites são um caso de média não-nula dos limites apresentados em [121] e maiores detalhes e limitações podem ser encontrados nesta referência.

Admite-se, também, que os sinais dos ruídos \tilde{w}_k e \tilde{v}_k são correlacionados com média, covariância e covariância cruzadas variantes no tempo e incertezas, mas pertencentes a conjuntos limitados e conhecidos². Por hipótese, estes conjuntos podem ser descritos como

$$\tilde{w}_k = (B_{w,k} + \Delta B_{w,k}) w_k + (B_{v,k} + \Delta B_{v,k}) v_k, \quad (4.11)$$

$$\tilde{v}_k = (D_{w,k} + \Delta D_{w,k}) w_k + (D_{v,k} + \Delta D_{v,k}) v_k. \quad (4.12)$$

Assume-se que as condições iniciais x_0 e os ruídos $w_k \in \mathfrak{R}^{n_w}$, $v_k \in \mathfrak{R}^{n_v}$ são não-correlacionados e com propriedades estatísticas

$$E \left\{ \begin{bmatrix} w_k^T & v_k^T & x_0^T \end{bmatrix}^T \right\} = \begin{bmatrix} \bar{w}_k^T & \bar{v}_k^T & \bar{x}_0^T \end{bmatrix}^T, \quad (4.13)$$

$$E \left\{ \begin{bmatrix} w_k - \bar{w}_k \\ v_k - \bar{v}_k \\ x_0 - \bar{x}_0 \end{bmatrix} \begin{bmatrix} w_j - \bar{w}_j \\ v_j - \bar{v}_j \\ x_0 - \bar{x}_0 \end{bmatrix}^T \right\} = \begin{bmatrix} W_k \delta_{kj} & 0 & 0 \\ 0 & V_k \delta_{kj} & 0 \\ 0 & 0 & X_0 \end{bmatrix}. \quad (4.14)$$

Em [64], utiliza-se $\bar{w}_k = 0$ e $\bar{v}_k = 0$ para todo k . Utilizando o modelamento de ruídos (4.11)-(4.12), o sistema (4.5)-(4.6) pode ser escrito como

$$x_{k+1} = f_k(x_k, u_k) + (B_{w,k} + \Delta B_{w,k}) w_k + (B_{v,k} + \Delta B_{v,k}) v_k, \quad (4.15)$$

$$y_k = h_k(x_k, u_k) + (D_{w,k} + \Delta D_{w,k}) w_k + (D_{v,k} + \Delta D_{v,k}) v_k. \quad (4.16)$$

Assume-se que as matrizes nominais $B_{w,k}$, $B_{v,k}$, $D_{w,k}$ e $D_{v,k}$ são conhecidas e variantes no tempo. As matrizes $\Delta B_{w,k}$, $\Delta B_{v,k}$, $\Delta D_{w,k}$ e $\Delta D_{v,k}$ representam as incertezas associadas e assume-se que possuam

²Consultar a seção 3.3.1 para maiores detalhes sobre o espaço das médias e das covariâncias incertas dos ruídos.

a seguinte estrutura

$$\begin{bmatrix} \Delta B_{w,k} & \Delta B_{v,k} \\ \Delta D_{w,k} & \Delta D_{v,k} \end{bmatrix} = \begin{bmatrix} H_{Bw,k} & H_{Bv,k} \\ H_{Dw,k} & H_{Dv,k} \end{bmatrix} F_k \begin{bmatrix} G_{w,k} & 0 \\ 0 & G_{v,k} \end{bmatrix}, \quad (4.17)$$

em que $F_k = \text{diag}\{F_{w,k}, F_{v,k}\}$. As matrizes $H_{Bw,k}$, $H_{Bv,k}$, $H_{Dw,k}$, $H_{Dv,k}$, $G_{w,k}$ e $G_{v,k}$ também são conhecidas e com dimensões apropriadas. Novamente, esta estrutura permite utilizar $n_w \neq n_v$. As matrizes $F_{w,k} \in \mathbb{R}^{n_{Fw,k} \times s_{Fw,k}}$ e $F_{v,k} \in \mathbb{R}^{n_{Fv,k} \times s_{Fv,k}}$ são desconhecidas e limitadas, *i.e.*, $\forall k \in [0, N]$

$$F_{w,k}^T F_{w,k} \leq I, \quad F_{v,k}^T F_{v,k} \leq I. \quad (4.18)$$

4.4 MATRIZES DE COVARIÂNCIA NÃO-LINEARES

Nesta subsecção, busca-se encontrar as formulações das matrizes de covariância *a priori* e *a posteriori*. O primeiro passo consiste em obter a evolução temporal dos erros de predição e correção. Considere o sistema (4.5)-(4.6) e o filtro dado por (3.1)-(3.2). Se forem definidos os erros

$$e_{x,k+1|k} := x_{k+1} - \hat{x}_{k+1|k}, \quad (4.19)$$

$$e_{x,k|k} := x_k - \hat{x}_{k|k}, \quad (4.20)$$

então, podem-se obter erros de estimação normalizados como sendo

$$\begin{aligned} e_{x,k+1|k} - E\{e_{x,k+1|k}\} &= A_k (e_{x,k|k} - E\{e_{x,k|k}\}) + (e_{nf,k|k} - E\{e_{nf,k|k}\}) \\ &\quad + (B_{w,k} + \Delta B_{w,k})(w_k - \bar{w}_k) + (B_{v,k} + \Delta B_{v,k})(v_k - \bar{v}_k), \end{aligned} \quad (4.21)$$

$$\begin{aligned} e_{x,k|k} - E\{e_{x,k|k}\} &= -K_k (e_{nh,k|k-1} - E\{e_{nh,k|k-1}\}) (I - K_k C_k) (e_{x,k|k-1} - E\{e_{x,k|k-1}\}) \\ &\quad - K_k (D_{w,k} + \Delta D_{w,k})(w_k - \bar{w}_k) - K_k (D_{v,k} + \Delta D_{v,k})(v_k - \bar{v}_k), \end{aligned} \quad (4.22)$$

em que se definiram também

$$e_{nf,k|k} := f_k(x_k, u_k) - f_k(\hat{x}_{k|k}, u_k) - A_k(x_k - \hat{x}_{k|k}) - B_k u_k, \quad (4.23)$$

$$e_{nh,k|k-1} := h_k(x_k, u_k) - h_k(\hat{x}_{k|k-1}, u_k) - C_k(x_k - \hat{x}_{k|k-1}) - D_k u_k. \quad (4.24)$$

Estes resultados podem ser demonstrados a partir das próprias definições das variáveis. O erro (4.23) baseia-se na idéia de que a diferença $f_k(x_k, u_k) - f_k(\hat{x}_{k|k}, u_k)$ possa ser escrita como uma função de $x_{k+1} - \hat{x}_{k+1|k}$ e u_k , *e.g.*, por meio de uma representação em série de Taylor. Considerando o primeiro termo da sequência que depende de $x_{k+1} - \hat{x}_{k+1|k}$ e u_k , então o erro (4.23) irá depender apenas dos termos de ordem mais alta. Neste sentido, se $x_{k+1} - \hat{x}_{k+1|k}$ for pequeno, espera-se que o erro (4.23) seja menor ainda. O mesmo raciocínio pode ser aplicado para a diferença $h_k(x_k, u_k) - h_k(\hat{x}_{k|k-1}, u_k)$.

Nota 4.4.1 Em [64], é assumido que $E\{e_{x,k+1|k}\} = E\{e_{x,k|k}\} = 0$. ■

Os erros $e_{nf,k|k}$ e $e_{x,k|k}$ são correlacionados entre si e não-correlacionados com w_k e v_k . Utilizando a definição de erro mostrada em (4.23), as propriedades (4.13)-(4.14) e a definição de covariância, a matriz

de covariância *a priori* é dada por

$$P_{x,k+1|k} = \begin{bmatrix} A_k & I \end{bmatrix} \begin{bmatrix} P_{x,k|k} & P_{x,nf,k|k} \\ P_{nf,x,k|k} & P_{nf,k|k} \end{bmatrix} \begin{bmatrix} A_k^T \\ I \end{bmatrix} + (B_{w,k} + \Delta B_{w,k}) W_k (B_{w,k} + \Delta B_{w,k})^T + (B_{v,k} + \Delta B_{v,k}) V_k (B_{v,k} + \Delta B_{v,k})^T, \quad (4.25)$$

em que

$$P_{x,k+1|k} := \text{cov} \{e_{x,k+1|k}\}, \quad (4.26)$$

$$P_{x,k|k} := \text{cov} \{e_{x,k|k}\}, \quad (4.27)$$

$$P_{x,nf,k|k} := \text{cov} \{e_{x,k|k}, e_{nf,k|k}\}, \quad (4.28)$$

$$P_{nf,k|k} := \text{cov} \{e_{nf,k|k}\}. \quad (4.29)$$

Utilizando argumentos análogos, a covariância *a posteriori* é dada por

$$P_{x,k|k} = \begin{bmatrix} I - K_k C_k & K_k \end{bmatrix} \begin{bmatrix} P_{x,k|k-1} & P_{x,nh,k|k-1} \\ P_{nh,x,k|k-1} & P_{nh,k|k-1} \end{bmatrix} [\bullet]^T + (D_{w,k} + \Delta D_{w,k}) W_k (D_{w,k} + \Delta D_{w,k})^T + (D_{v,k} + \Delta D_{v,k}) V_k (D_{v,k} + \Delta D_{v,k})^T, \quad (4.30)$$

em que

$$P_{x,nh,k|k-1} := \text{cov} \{e_{x,k|k-1}, e_{nh,k|k-1}\}, \quad (4.31)$$

$$P_{nh,k|k-1} := \text{cov} \{e_{nh,k|k-1}\}, \quad (4.32)$$

4.5 MAJORANTES PARA AS COVARIÂNCIAS DAS FUNÇÕES NÃO-LINEARES

As matrizes de covariância (4.25) e (4.30) dependem de funções não-lineares $f_k(\bullet)$ e $h_k(\bullet)$. O objetivo desta subseção é encontrar limites superiores para estas matrizes não-lineares de covariância. O Lema 4.5.1 fornece os majorantes procurados.

Teorema 4.5.1 *Para todo $\alpha_{h,k} > 0$ e $\alpha_{f,k} > 0$, os limites superiores para as seqüências (4.25) e (4.30) são dados por*

$$\begin{aligned} \bar{P}_{k+1|k} &= \begin{bmatrix} A_k & I \end{bmatrix} \begin{bmatrix} P_{11,k|k} & 0 \\ 0 & P_{22,k|k} \end{bmatrix} \begin{bmatrix} A_k^T \\ I \end{bmatrix} \\ &+ (B_{w,k} + H_{Bw,k} F_{w,k} G_{w,k}) W_k (B_{w,k} + H_{Bw,k} F_{w,k} G_{w,k})^T \\ &+ (B_{v,k} + H_{Bv,k} F_{v,k} G_{v,k}) V_k (B_{v,k} + H_{Bv,k} F_{v,k} G_{v,k})^T, \end{aligned} \quad (4.33)$$

$$\begin{aligned} \bar{P}_{k|k} &= \begin{bmatrix} I - K_k C_k & K_k \end{bmatrix} \begin{bmatrix} P_{11,k|k-1} & 0 \\ 0 & P_{22,k|k-1} \end{bmatrix} [\bullet]^T \\ &+ (D_{w,k} + H_{Dw,k} F_{w,k} G_{w,k}) W_k (D_{w,k} + H_{Dw,k} F_{w,k} G_{w,k})^T \\ &+ (D_{v,k} + H_{Dv,k} F_{v,k} G_{v,k}) V_k (D_{v,k} + H_{Dv,k} F_{v,k} G_{v,k})^T, \end{aligned} \quad (4.34)$$

em que

$$P_{11,k|k} := (1 + \alpha_{f,k}) \bar{P}_{k|k}, \quad (4.35)$$

$$P_{22,k|k} := \left(1 + \frac{1}{\alpha_{f,k}}\right) \text{tr} \{ \Delta_{a,k} \bar{P}_{k|k} \Delta_{a,k}^T \} I, \quad (4.36)$$

$$P_{11,k|k-1} := (1 + \alpha_{h,k}) \bar{P}_{k|k-1}, \quad (4.37)$$

$$P_{22,k|k-1} := \left(1 + \frac{1}{\alpha_{h,k}}\right) \text{tr} (\Delta_{c,k} \bar{P}_{k|k-1} \Delta_{c,k}^T) I. \quad (4.38)$$

Prova Usando o resultado do Lema 4.5.2 e fazendo as identificações apropriadas, é correto afirmar que

$$0 \leq \begin{bmatrix} P_{x,k|k-1} & P_{x,nh,k|k-1} \\ P_{nh,x,k|k-1} & P_{nh,k|k-1} \end{bmatrix} \leq \begin{bmatrix} (1 + \alpha_{h,k}) P_{x,k|k-1} & 0 \\ 0 & \left(1 + \frac{1}{\alpha_{h,k}}\right) \text{tr} (P_{nh,k|k-1}) I \end{bmatrix}, \quad (4.39)$$

e

$$0 \leq \begin{bmatrix} P_{x,k|k} & P_{x,nf,k|k} \\ P_{nf,x,k|k} & P_{nf,k|k} \end{bmatrix} \leq \begin{bmatrix} (1 + \alpha_{f,k}) P_{x,k|k} & 0 \\ 0 & \left(1 + \frac{1}{\alpha_{f,k}}\right) \text{tr} (P_{nf,k|k}) I \end{bmatrix}. \quad (4.40)$$

Entretanto, os majorantes (4.39) e (4.40) ainda dependem das matrizes de covariância não-lineares $P_{nf,k|k}$ and $P_{nh,k|k-1}$. A fim de solucionar este problema, deve-se encontrar limites superiores para estas matrizes. Utilizando as definições de erros (4.23)-(4.24) e as hipóteses (4.7)-(4.8), conclui-se que

$$\|e_{nh,k|k-1} - E \{e_{nh,k|k-1}\}\| \leq \|\Delta_{c,k} (e_{x,k|k-1} - E \{e_{x,k|k-1}\})\|, \quad (4.41)$$

$$\|e_{nf,k|k} - E \{e_{nf,k|k}\}\| \leq \|\Delta_{a,k} (e_{x,k|k} - E \{e_{x,k|k}\})\|. \quad (4.42)$$

Por meio do Lema I.6.3 e das definições apresentadas em (4.27), (4.26), (4.29) e (4.32), é possível observar que

$$\text{tr} (P_{nh,k|k-1}) \leq \text{tr} (\Delta_{c,k} P_{x,k|k-1} \Delta_{c,k}^T), \quad (4.43)$$

$$\text{tr} (P_{nf,k|k}) \leq \text{tr} (\Delta_{a,k} P_{x,x,k|k} \Delta_{a,k}^T). \quad (4.44)$$

A expressão (4.33) é obtida por meio da substituição de (4.43) em (4.39) e substituindo o resultado em (4.30). Procedimento análogo é válido para (4.34). Utilizando o Lema 3.2 de [90], conclui-se que $P_{x,k+1|k} \leq \bar{P}_{k+1|k}$ e $P_{x,k|k} \leq \bar{P}_{k|k}$. ■

Lema 4.5.2 Se $P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \geq 0$, em que $P_{11} \in \mathbb{R}^{p \times p}$, $P_{12} \in \mathbb{R}^{p \times q}$, $P_{21} \in \mathbb{R}^{q \times p}$ e $P_{22} \in \mathbb{R}^{q \times q}$, então, para todo $\beta > 0$,

$$0 \leq \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \leq \begin{bmatrix} (1 + \beta) P_{11} & 0 \\ 0 & \left(1 + \frac{1}{\beta}\right) \text{tr} (P_{22}) I_q \end{bmatrix}. \quad (4.45)$$

Prova Como $P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \geq 0$, há matrizes C e D , tais que

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} C \\ D \end{bmatrix} \begin{bmatrix} C^T & D^T \end{bmatrix} = \begin{bmatrix} CC^T & CD^T \\ DC^T & DD^T \end{bmatrix}. \quad (4.46)$$

Assim, para quaisquer vetores x e y , é correto afirmar que

$$\begin{aligned}
0 \leq \begin{bmatrix} x^T & y^T \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} x^T & y^T \end{bmatrix} \begin{bmatrix} P_{11}x + P_{12}y \\ P_{21}x + P_{22}y \end{bmatrix} \\
&= x^T P_{11}x + x^T P_{12}y + y^T P_{21}x + y^T P_{22}y \\
&= x^T C C^T x + x^T C D^T y + y^T D C^T x + y^T D D^T y. \quad (4.47)
\end{aligned}$$

Um resultado conhecido na literatura é de que “produtos cruzados” são majorados por termos “quadrados perfeitos”, ou seja, $xy^T + yx^T \leq \beta xx^T + \frac{1}{\beta} yy^T$ para todo $\beta > 0$. Assim,

$$\begin{aligned}
\begin{bmatrix} x^T & y^T \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &\leq x^T C C^T x + \beta^2 x^T C C^T x + \frac{1}{\beta^2} y^T D D^T y + y^T D D^T y \\
&= (1 + \beta^2) x^T C C^T x + \left(1 + \frac{1}{\beta^2}\right) y^T D D^T y \\
&= (1 + \beta^2) x^T P_{11}x + \left(1 + \frac{1}{\beta^2}\right) y^T P_{22}y \\
&= \begin{bmatrix} x^T & y^T \end{bmatrix} \begin{bmatrix} (1 + \beta^2) P_{11} & 0 \\ 0 & \left(1 + \frac{1}{\beta^2}\right) P_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \quad (4.48)
\end{aligned}$$

Dado que a desigualdade

$$\begin{bmatrix} x^T & y^T \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} x^T & y^T \end{bmatrix} \begin{bmatrix} (1 + \beta^2) P_{11} & 0 \\ 0 & \left(1 + \frac{1}{\beta^2}\right) P_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (4.49)$$

vale para quaisquer vetores x e y , segue que

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \leq \begin{bmatrix} (1 + \beta^2) P_{11} & 0 \\ 0 & \left(1 + \frac{1}{\beta^2}\right) P_{22} \end{bmatrix}. \quad (4.50)$$

Sendo $0 \leq P_{22} \leq \text{tr}(P_{22}) I$, tem-se que, para todo $\beta \neq 0$, a desigualdade (4.45) é válida. ■

4.6 MAJORANTES PARA OS RUÍDOS COM PROPRIEDADES INCERTAS

O próximo passo consiste em encontrar limites superiores para as matrizes de covariância (4.33)-(4.34). A motivação de incluir ruídos com propriedades estatísticas incertas surgiu a partir do estudo realizado para sistemas lineares apresentado neste capítulo. Assim, considerando as incertezas de norma limitada descritas em (4.17)-(4.18), pode-se utilizar o Lema 2 de [111] e o Lema 3.2 de [90] para se obter os seguintes majorantes:

$$\bar{P}_{k+1|k} = \begin{bmatrix} A_k & I \end{bmatrix} \begin{bmatrix} P_{11,k|k} & 0 \\ 0 & P_{22,k|k} \end{bmatrix} \begin{bmatrix} A_k^T \\ I \end{bmatrix} + \Psi_{B,k}, \quad (4.51)$$

$$\bar{P}_{k|k} = \begin{bmatrix} I - K_k C_k & K_k \end{bmatrix} \begin{bmatrix} P_{11,k|k-1} & 0 \\ 0 & P_{22,k|k-1} \end{bmatrix} [\bullet]^T + \Psi_{D,k}, \quad (4.52)$$

em que se define

$$\Psi_{B,k} := B_{w,k}W_{c,k}B_{w,k}^T + B_{v,k}V_{c,k}B_{v,k}^T + \Delta_{B,k}, \quad (4.53)$$

$$\Psi_{D,k} := D_{w,k}W_{c,k}D_{w,k}^T + D_{v,k}V_{c,k}D_{v,k}^T + \Delta_{D,k}, \quad (4.54)$$

$$W_{c,k} := W_k + W_k G_{w,k}^T \left(\alpha_{w,k}^{-1} I - G_{w,k} W_k G_{w,k}^T \right)^{-1} G_{w,k} W_k, \quad (4.55)$$

$$V_{c,k} := V_k + V_k G_{v,k}^T \left(\alpha_{v,k}^{-1} I - G_{v,k} V_k G_{v,k}^T \right)^{-1} G_{v,k} V_k, \quad (4.56)$$

$$\Delta_{B,k} := \alpha_{w,k}^{-1} H_{Bw,k} H_{Bw,k}^T + \alpha_{v,k}^{-1} H_{Bv,k} H_{Bv,k}^T, \quad (4.57)$$

$$\Delta_{D,k} := \alpha_{w,k}^{-1} H_{Dw,k} H_{Dw,k}^T + \alpha_{v,k}^{-1} H_{Dv,k} H_{Dv,k}^T, \quad (4.58)$$

de tal forma que os parâmetros escalares devam ser escolhidos para satisfazer

$$\alpha_{h,k} > 0, \quad (4.59)$$

$$\alpha_{w,k}^{-1} I - G_{w,k} W_k G_{w,k}^T > 0, \quad (4.60)$$

$$\alpha_{v,k}^{-1} I - G_{v,k} V_k G_{v,k}^T > 0. \quad (4.61)$$

4.7 MINIMIZAÇÃO DOS MAJORANTES E OBTENÇÃO DO FILTRO

Até este ponto, o filtro depende de cinco parâmetros, sendo quatro escalares e o ganho do filtro. O limite superior (4.51) é minimizado, calculando-se a primeira derivada parcial de $\bar{P}_{k+1|k}$ em relação a $\alpha_{f,k}$ e igualando o resultado a zero, *i.e.*, obtém-se o valor ótimo de $\alpha_{f,k} = \alpha_{f,k}^{opt}$, tal que

$$\alpha_{f,k}^{opt} = \left(\frac{tr \left\{ \Delta_{a,k} \bar{P}_{k|k} \Delta_{a,k}^T \right\}}{n_x \frac{tr \left\{ A_k \bar{P}_{k|k} A_k^T \right\}}{tr \left\{ A_k \bar{P}_{k|k} A_k^T \right\}}} \right)^{\frac{1}{2}}. \quad (4.62)$$

O lema 4.7.1 mostra a obtenção deste mínimo para uma expressão genérica. A fim de se alcançar o resultado 4.62, basta fazer as identificações necessárias e lembrar que $tr \{ I_{n_x} \} = n_x$.

Lema 4.7.1 *O mínimo de*

$$P = (1 + \alpha) \Psi_1 + \left(1 + \frac{1}{\alpha} \right) tr \{ \Psi_2 \} \Psi_3 + \Psi_4 \quad (4.63)$$

em relação à α , para $tr \{ \Psi_2 \} \Psi_3 \geq 0$, ocorre em

$$\alpha^* = \left(\frac{tr \{ \Psi_3 \} tr \{ \Psi_2 \}}{tr \{ \Psi_1 \}} \right)^{\frac{1}{2}}. \quad (4.64)$$

Prova Calculando a derivada de P em relação à α , tem-se que

$$\frac{\partial}{\partial \alpha} P = \Psi_1 - \frac{1}{\alpha^2} tr \{ \Psi_2 \} \Psi_3. \quad (4.65)$$

Igualando a derivada a zero,

$$\Psi_1 = \frac{1}{\alpha^2} tr \{ \Psi_2 \} \Psi_3 \Rightarrow \alpha^2 \Psi_1 = tr \{ \Psi_2 \} \Psi_3. \quad (4.66)$$

Lembrando que o traço de um escalar é o próprio escalar, tem-se que aplicando o operador de traço nos dois lados,

$$\text{tr} \{ \Psi_1 \} \alpha^2 = \text{tr} \{ \Psi_3 \} \text{tr} \{ \Psi_2 \} \Rightarrow \alpha^* = \left(\frac{\text{tr} \{ \Psi_3 \} \text{tr} \{ \Psi_2 \}}{\text{tr} \{ \Psi_1 \}} \right)^{\frac{1}{2}}. \quad (4.67)$$

Para se verificar que α é um ponto de mínimo, parte-se da derivada de (4.66) para mostrar

$$\frac{\partial^2}{\partial^2 \alpha} P = 2 \frac{1}{\alpha^3} \text{tr} \{ \Psi_2 \} \Psi_3 \geq 0 \quad (4.68)$$

■

A covariância (4.52) depende de $\alpha_{h,k}$ e K_k . Por hipótese, adotar-se-á $\alpha_{h,k} > 0$, independente de K_k . Portanto, a escolha do parâmetro $\alpha_{h,k}$ deve ser sintonizada para cada aplicação específica e vai depender do modelo não-linear, bem como do tipo de comportamento exigido para o sistema. A partir desta premissa, calcula-se a derivada parcial de primeira ordem de (4.52) em relação a K_k , igualando-a a zero, de maneira a se obter o valor ótimo $K_k = K_k^{opt}$, ou seja,

$$K_k^{opt} = (1 + \alpha_{h,k}) \bar{P}_{k|k-1} C_k^T R_k^{-1}. \quad (4.69)$$

A derivada parcial de segunda ordem indica que K_k^{opt} minimiza (4.52). O lema 4.7.2 mostra, por meio de expressões genéricas, como se chegar a este resultado.

Lema 4.7.2 *O mínimo de*

$$P_{k|k} = (I - KC) P (I - KC)^T + K \Psi K^T \quad (4.70)$$

em relação à K ocorre, para $\Psi \geq 0$ e $P \geq 0$, em

$$K^{opt} = PC^T (\Psi + CPC^T)^{-1}. \quad (4.71)$$

Prova Expandindo a expressão para $P_{k|k}$, chega-se a

$$P_{k|k} = P - KCP - PC^T K^T + K (CPC^T + \Psi) K^T. \quad (4.72)$$

Derivando o primeiro termo de $P_{k|k}$ em relação à K ,

$$\frac{\partial}{\partial K} P = 0. \quad (4.73)$$

A diferenciação do segundo e terceiro termos fornece

$$\frac{\partial}{\partial K} (KCP + PC^T K^T) = 2PC^T. \quad (4.74)$$

Derivando o último termo, tem-se que

$$\frac{\partial}{\partial K} (K (CPC^T + \Psi) K^T) = 2K (CPC^T + \Psi). \quad (4.75)$$

Portanto,

$$\frac{\partial}{\partial K} P_{k|k} = -2PC^T + 2K(CPC^T + \Psi). \quad (4.76)$$

Igualando a derivada com zero

$$K_k^{opt} (CPC^T + \Psi) = PC^T. \quad (4.77)$$

Supondo que a inversa existe,

$$K^{opt} = PC^T (\Psi + CPC^T)^{-1}, \quad (4.78)$$

Para se verificar se K é um ponto de mínimo, parte-se da derivada de (4.76)

$$\frac{\partial^2}{\partial^2 K} P_{k|k} = 2(CPC^T + \Psi) \geq 0 \quad (4.79)$$

■

Desta forma, a equação de Riccati pode ser escrita como

$$\bar{P}_{k|k} = (1 + \alpha_{h,k}) \left(I - K_k^{opt} C_k \right) \bar{P}_{k|k-1}. \quad (4.80)$$

Substituindo a expressão de K_k^{opt} em (4.80), obtém-se

$$\bar{P}_{k|k} = (1 + \alpha_{h,k}) \bar{P}_{k|k-1} - (1 + \alpha_{h,k})^2 \bar{P}_{k|k-1} C_k^T R_k^{-1} C_k \bar{P}_{k|k-1}. \quad (4.81)$$

Nota 4.7.3 Considerando o modelo sem entradas exógenas e com ruídos de média nula e propriedades estatísticas conhecidas, os parâmetros do filtro proposto, bem como a equação de Riccati, recaem no filtro apresentado em [64] que, por sua vez, recaem no FKE para sistemas tradicionais. ■

O algoritmo proposto para o RFKE é apresentado na Tabela 4.1.

Tabela 4.1: Algoritmo do Filtro Não-Linear Robusto.

Passo 0 (Condições iniciais): $\hat{x}_{0|-1} = \bar{x}_0$ e $\bar{P}_{0|-1} = X_0$.

Passo 1: Obter parâmetros escalares que satisfaçam

$$\begin{aligned} \alpha_{h,k} &> 0, \\ \alpha_{w,k}^{-1} I - G_{w,k} W_k G_{w,k}^T &> 0, \\ \alpha_{v,k}^{-1} I - G_{v,k} V_k G_{v,k}^T &> 0. \end{aligned}$$

e, em seguida, calcular

$$\begin{aligned} \Delta_{B,k} &= \alpha_{w,k}^{-1} H_{Bw,k} H_{Bw,k}^T + \alpha_{v,k}^{-1} H_{Bv,k} H_{Bv,k}^T, \\ \Delta_{D,k} &= \alpha_{w,k}^{-1} H_{Dw,k} H_{Dw,k}^T + \alpha_{v,k}^{-1} H_{Dv,k} H_{Dv,k}^T, \end{aligned}$$

Passo 2: Calcular as correções, devido à presença das incertezas

$$\begin{aligned} W_{c,k} &= W_k + W_k G_{w,k}^T \left(\alpha_{w,k}^{-1} I - G_{w,k} W_k G_{w,k}^T \right)^{-1} G_{w,k} W_k, \\ V_{c,k} &= V_k + V_k G_{v,k}^T \left(\alpha_{v,k}^{-1} I - G_{v,k} V_k G_{v,k}^T \right)^{-1} G_{v,k} V_k. \end{aligned}$$

Passo 3: Correção do filtro

$$\begin{aligned} K_k^{opt} &= (1 + \alpha_{h,k}) \bar{P}_{k|k-1} C_k^T R_k^{-1}, \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k^{opt} (y_k - h_k(\hat{x}_{k|k-1}, u_k)), \\ \bar{P}_{k|k} &= (1 + \alpha_{h,k}) \bar{P}_{k|k-1} - (1 + \alpha_{h,k})^2 \bar{P}_{k|k-1} C_k^T R_k^{-1} C_k \bar{P}_{k|k-1}, \end{aligned}$$

em que

$$\begin{aligned} R_k &:= \left(1 + \frac{1}{\alpha_{h,k}} \right) tr \left\{ \Delta_{c,k} \bar{P}_{k|k-1} \Delta_{c,k}^T \right\} I + (1 + \alpha_{h,k}) C_k \bar{P}_{k|k-1} C_k^T \\ &\quad + D_{w,k} W_{c,k} D_{w,k}^T + D_{v,k} V_{c,k} D_{v,k}^T + \Delta_{D,k}. \end{aligned}$$

Passo 4: Predição do filtro

$$\begin{aligned} \hat{x}_{k+1|k} &= f_k(\hat{x}_{k|k}, u_k), \\ \bar{P}_{k+1|k} &= \left(1 + \alpha_{f,k}^{opt} \right) A_k \bar{P}_{k|k} A_k^T + \left(1 + \frac{1}{\alpha_{f,k}^{opt}} \right) tr \left\{ \Delta_{a,k} \bar{P}_{k|k} \Delta_{a,k}^T \right\} \\ &\quad + B_{w,k} W_{c,k} B_{w,k}^T + B_{v,k} V_{c,k} B_{v,k}^T + \Delta_{B,k}, \end{aligned}$$

$$\text{em que } \alpha_{f,k}^{opt} = \left(n_x \frac{tr \{ \Delta_{a,k} \bar{P}_{k|k} \Delta_{a,k}^T \}}{tr \{ A_k \bar{P}_{k|k} A_k^T \}} \right)^{\frac{1}{2}}.$$

5 EXPERIMENTOS NUMÉRICOS

Nesta seção serão apresentados os critérios adotados para se avaliar o desempenho dos estimadores. Em seguida, há uma simulação com o objetivo de se visualizar a influência do parâmetro escalar do estimador na resposta do modelo. Por fim, apresentam-se os resultados numéricos dos algoritmos propostos.

5.1 INDICADORES DE DESEMPENHO

Uma das métricas desejadas para se quantificar o desempenho do estimador seria a variância exata do erro de estimação. Contudo, como este índice é difícil de se calcular a partir da resposta do modelo, a variância real do erro de estimação foi aproximada por meio da *ensemble-average* utilizada em trabalhos, tais como [70] e [68]. A *ensemble-average*, para sistemas com média não-nula, pode ser definida como

$$\text{var} \{e_{i,k}\} \approx \frac{1}{N} \sum_{j=1}^N \left(e_{i,k}^{(j)} - E \{e_{i,k}^{(j)}\} \right)^2, \quad (5.1)$$

$$E \{e_{i,k}^{(j)}\} \approx \frac{1}{N} \sum_{j=1}^N e_{i,k}^{(j)}, \quad (5.2)$$

em que $e_{i,k}^{(j)}$ é a i -ésima componente do vetor do erro de estimação $e_k^{(j)}$ do experimento j definido como

$$e_k^{(j)} := x_k^{(j)} - \hat{x}_{k|k-1}^{(j)}. \quad (5.3)$$

Uma outra maneira de se visualizar o desempenho do estimador é por meio da elipse de covariância. De acordo com [123], a elipse de covariância é uma maneira conveniente de se visualizar o tamanho relativo da matriz de covariância. Outros detalhes sobre elipses de covariância são apresentados no anexo I.4.

5.2 INFLUÊNCIA DO PARÂMETRO ESCALAR

Uma das questões em aberto para a execução do estimador é justamente a escolha dos parâmetros escalares. De fato, durante o desenvolvimento do algoritmo, somente são impostas condições de desigualdades e não foi encontrada na literatura uma expressão fechada para a escolha de tais parâmetros. No entanto, em [71], o parâmetro escalar foi incluído na minimização numérica obtida a partir de uma formulação LMI.

A partir desta motivação, deseja-se verificar a influência do parâmetro escalar - obtido a partir da incerteza na matriz de transição de estados - na variância aproximada do erro de estimação. Utilizando o exemplo apresentado em [110], cujo modelo é mostrado a seguir,

$$x_{k+1} = \begin{bmatrix} 0 & -0.5 \\ 1 & 1 + 0.3\delta_{1,k} \end{bmatrix} x_k + \begin{bmatrix} -6 \\ 1 + 0.2\delta_{2,k} \end{bmatrix} w_k + \begin{bmatrix} 0 & 0 \\ 1 + 0.1\delta_{3,k} & 0 \end{bmatrix} v_k, \quad (5.4)$$

$$y_k = \begin{bmatrix} -100 & 10 + 0.03\delta_{4,k} \end{bmatrix} x_k + (0.1 + 0.02\delta_{5,k}) w_k + \begin{bmatrix} 0.1 + 0.02\delta_{6,k} & 1 \end{bmatrix} v_k, \quad (5.5)$$

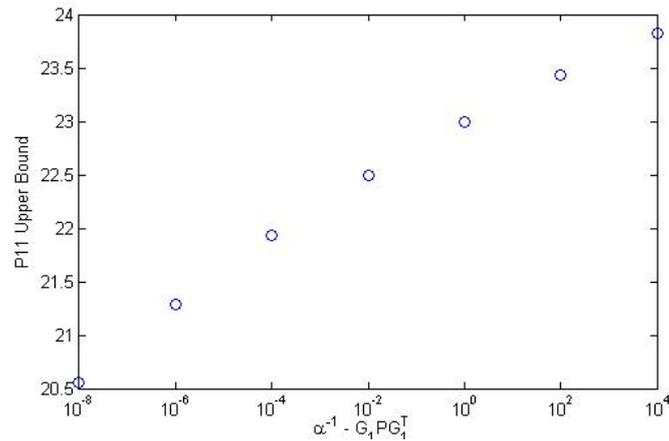


Figura 5.1: Influência do parâmetro escalar sobre a variância do erro de estimação do estado 1.

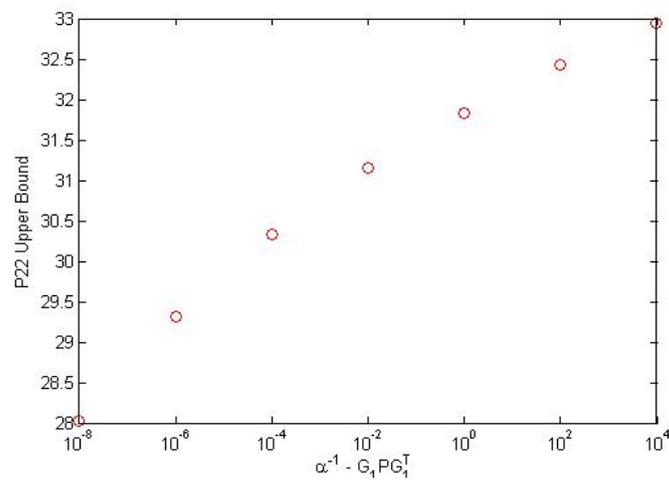


Figura 5.2: Influência do parâmetro escalar sobre a variância do erro de estimação do estado 2.

obtiveram-se os gráficos mostrados nas Figuras 5.1 e 5.2.

No eixo das abcissas, tem-se a distância do inverso do parâmetro escalar em relação ao seu limite inferior dado pela desigualdade. No eixo das ordenadas, apresenta-se a variância do erro de estimação para cada parâmetro adotado. Analisaram-se ambos os estados do sistema. Os resultados mostram que quanto mais perto do limite inferior estiver o inverso do parâmetro escalar, menor será a variância do erro de estimação - o que é um resultado desejável. Por outro lado, quanto menor esta distância, mais difícil de se obter as matrizes inversas do algoritmo, ocasionando erros numéricos para este cálculo.

5.3 ESTIMAÇÃO LINEAR

Nesta seção, compara-se o desempenho do preditor proposto neste trabalho com o preditor em [82]. Uma vez que o preditor em [82] fora desenvolvido para sistemas com ruídos não-correlacionados e restritos

a $\dim w_k = \dim v_k$, o seguinte modelo será adotado

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} 0 & -0.5 \\ 1 + \delta_{x,k} & 1 + 0.3\delta_{x,k} \end{bmatrix} x_k + \begin{bmatrix} -6 \\ 1 + 0.1\delta_{w,k} \end{bmatrix} w_k, \\ y_k &= \begin{bmatrix} -100 + 5\delta_{x,k} & 10 + 1.5\delta_{x,k} \end{bmatrix} x_k + 100\delta_{v,k}v_k, \end{aligned}$$

em que $\delta_{n,k}$ varia uniformemente a cada passo no intervalo unitário para $n = x, w, v$. Adotar-se-á, também, $\bar{w}_k = 0.1$, $\bar{v}_k = 0.9$, $W_k = 0.1$ e $V_k = 2$ com condições iniciais $\bar{x}_0 = [2 \ 1]^T$ e $X_0 = 0.1I$. As matrizes associadas às incertezas são dadas por

$$\begin{aligned} H_{A,k} &= \begin{bmatrix} 0 \\ 10 \end{bmatrix}, \quad H_{Bw,k} = \begin{bmatrix} 0 \\ 10 \end{bmatrix}, \quad H_{Bv,k} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ H_{C,k} &= 50, \quad H_{Dw,k} = 0, \quad H_{Dv,k} = 100, \\ G_{x,k} &= \begin{bmatrix} 0.1 & 0.03 \end{bmatrix}, \quad G_{w,k} = 0.01, \quad G_{v,k} = 1. \end{aligned} \quad (5.6)$$

Os parâmetros escalares $\alpha_{x,k}^{-1}$, $\alpha_{w,k}^{-1}$ and $\alpha_{v,k}^{-1}$ são calculados a cada passo da seguinte forma

$$\alpha_{x,k}^{-1} = \sigma_{max} \{ G_{x,k} \bar{P}_{k|k-1} G_{x,k}^T \} + \epsilon_x, \quad (5.7)$$

$$\alpha_{w,k}^{-1} = \sigma_{max} \{ G_{w,k} W_k G_{w,k}^T \} + \epsilon_w, \quad (5.8)$$

$$\alpha_{v,k}^{-1} = \sigma_{max} \{ G_{v,k} V_k G_{v,k}^T \} + \epsilon_v, \quad (5.9)$$

em que $\sigma_{max} \{ \bullet \}$ indica o valor singular máximo de uma matriz. Neste exemplo, observando o resultado da seção 5.2, escolheram-se $\epsilon_x = \epsilon_w = \epsilon_v = 0.1$ para ambos os preditores.

Os valores médios das estimativas e das matrizes de covariância, obtidos ao longo de 500 experimentos para $k = 500$, foram

$$\begin{aligned} E \{ \hat{x}_{new|500} \} &= \begin{bmatrix} 0.14 \\ -1.52 \end{bmatrix}, \quad E \{ \bar{P}_{new|500} \} = \begin{bmatrix} 14 & -23 \\ -23 & 76 \end{bmatrix}, \\ E \{ \hat{x}_{[2]|500} \} &= \begin{bmatrix} 0.72 \\ -2.81 \end{bmatrix}, \quad E \{ \bar{P}_{[2]|500} \} = \begin{bmatrix} 322 & -698 \\ -698 & 2181 \end{bmatrix}. \end{aligned}$$

As matrizes de covariância obtidas de ambos preditores são mostradas na Figura 5.3. Observa-se que a área da elipse relativa ao preditor desta dissertação é razoavelmente menor que a elipse resultante do preditor de [82]. Utilizando a *ensemble-average*, a Figura 5.4 apresenta a evolução da variância aproximada para o estado 2. A Figura 5.4 mostra que o preditor deste trabalho fornece um projeto menos conservador, com um menor limite superior e variância do erro de estimação. Na Figura 5.5 apresenta-se a média do estado real e suas estimativas ao longo de 500 experimentos. É perceptível que o preditor proposto aqui resulta em um menor erro médio de estimação também.

5.4 ESTIMAÇÃO NÃO-LINEAR

As simulações do desempenho do REKF serão realizadas, considerando a aplicação de um demodulador. Este é um caso em que, tipicamente, utiliza-se o FKE para demodular um sinal, especialmente

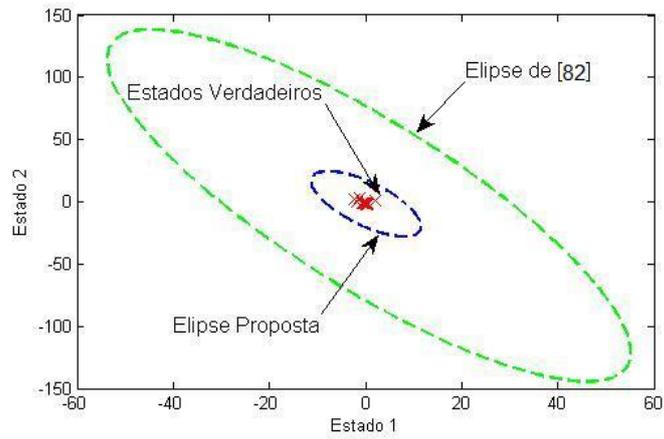


Figura 5.3: Média das elipses de covariância após 500 experimentos.

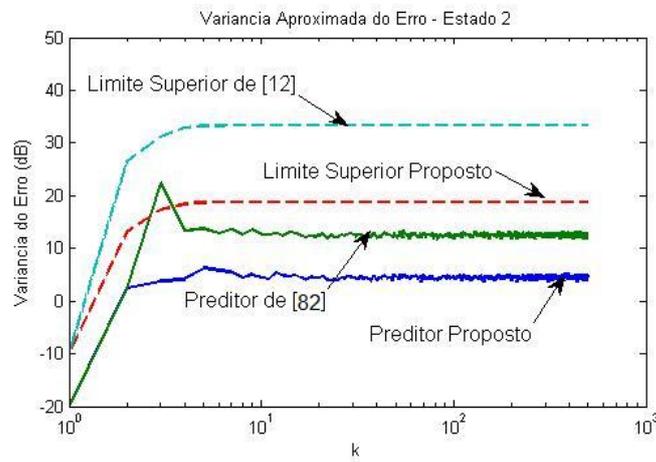


Figura 5.4: Variância aproximada do erro de estimação para simulação com ruídos não-correlacionados.

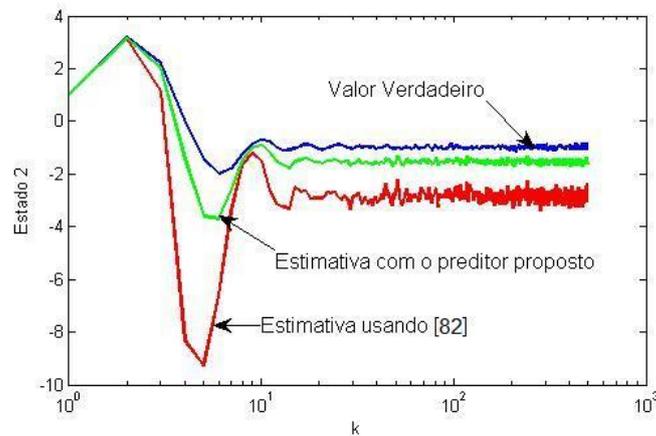


Figura 5.5: Estado real do sistema e suas estimativas, segundo os preditores analisados.

sinais com modulação FM [64]. Entretanto, nas próximas simulações, propõe-se estimar tanto a frequência quanto a amplitude do sinal. Conforme observado em [116], demoduladores geralmente são modelados de tal forma que as não-linearidades do sinal estejam presentes somente na equação de saída. No último exemplo numérico, verificou-se o desempenho do REFK com não-linearidades na equação de dinâmica presentes devido à variação não-linear da frequência ao longo do tempo. Os demoduladores estão sujeitos a ruídos, perturbações, atrasos e correlação entre os ruídos [124]. As próximas simulações contemplam as características citadas, exceto pelo atraso no tempo. Um modelo análogo, em que o autor estima a frequência e o deslocamento de fase, é apresentado em [125]. Uma referência completa sobre estimação aplicada a problema de demodulação pode ser encontrada em [124].

5.4.1 Demodulador sem Distúrbio de Fase

Considere o seguinte modelo de demodulador em que há correlação entre os ruídos de processo e de medida

$$x_{k+1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 0.1\delta_{w,k} \\ 0.0001\delta_{w,k} \end{bmatrix} w_k + \begin{bmatrix} \delta_{v,k} & \delta_{v,k} \\ \delta_{v,k} & 0.0001 + \delta_{v,k} \end{bmatrix} v_k, \quad (5.10)$$

$$y_k = x_{1,k} \cos\left(2\pi x_{2,k} \frac{k}{50}\right) + (0.01 + 0.1\delta_{w,k}) w_k + \begin{bmatrix} 0.1 + \delta_{v,k} & 0.1 + \delta_{v,k} \end{bmatrix} v_k, \quad (5.11)$$

em que as variáveis $\delta_{w,k}$ e $\delta_{v,k}$ são desconhecidas e variam uniformemente dentro do círculo unitário. As componentes do vetor de estado são, respectivamente, a amplitude do sinal $x_{1,k}$ e sua frequência $x_{2,k}$. Neste sentido, a estimação de estado do modelo (5.10)-(5.11) é equivalente à estimação da amplitude e da frequência de um sinal.

Os ruídos possuem média nula, ou seja, $\bar{w}_k = 0$ e $\bar{v}_k = [0 \ 0]^T$. As variâncias dos ruídos são unitárias da forma $W_k = 1$ e $V_k = \text{diag}\{1, 1\}$ com condições iniciais $\bar{x}_0 = [2 \ 0.2]^T$ e $X_0 = 0.1I$. As matrizes associadas às incertezas foram escolhidas da forma

$$\begin{aligned} H_{Bw,k} &= \begin{bmatrix} 0.1 \\ 0.0001 \end{bmatrix}, \quad H_{Bv,k} = \begin{bmatrix} 0.1 \\ 0.0001 \end{bmatrix}, \\ H_{Dw,k} &= 0.1, \quad H_{Dv,k} = 0.1, \\ G_{w,k} &= 1, \quad G_{v,k} = \begin{bmatrix} 1 & 1 \end{bmatrix}. \end{aligned} \quad (5.12)$$

Assim, para o projeto do filtro, admite-se que a jacobiana de $f_k(x_k)$ seja conhecida com exatidão e que a jacobiana de $h_k(x_k)$ possua uma incerteza associada. Matematicamente, este pensamento expressa-se como

$$\Delta_{a,k} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \Delta_{c,k} = \begin{bmatrix} 1 & 1 \end{bmatrix}. \quad (5.13)$$

Escolhendo $\epsilon_w = 0.1$ e $\epsilon_v = \text{diag}\{0.1, 0.1\}$, os parâmetros escalares foram escolhidos da forma

$$\alpha_{c,k} = 1, \quad (5.14)$$

$$\alpha_{w,k}^{-1} = \sigma_{\max}\{G_{w,k}W_kG_{w,k}^T\} + \epsilon_w, \quad (5.15)$$

$$\alpha_{v,k}^{-1} = \sigma_{\max}\{G_{v,k}V_kG_{v,k}^T\} + \epsilon_v. \quad (5.16)$$

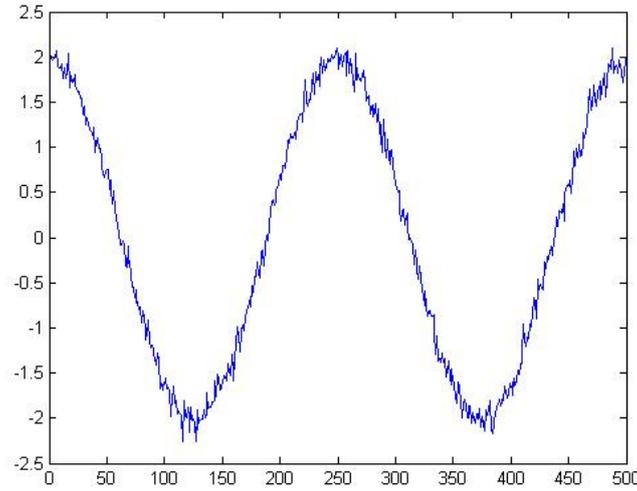


Figura 5.6: Exemplo de sinal de saída para o demodulador.

A Figura 5.6 ilustra a saída deste sistema para uma duração de 500 iterações. Nota-se que se trata de uma curva do tipo senóide com amplitude constante, sem deslocamento de fase e sujeita a ruídos.

O valor esperado das estimativas e das matrizes de covariância, obtidos ao longo de 500 experimentos e para $k = 500$, foram obtidos utilizando o filtro robusto proposto e o filtro clássico. Os resultados são mostrados nas Figuras 5.7 e 5.8.

Na Figura 5.7, observa-se que o filtro desta dissertação também resultou em um menor erro de estimação. Para a estimação da amplitude, o filtro tradicional convergiu para um valor diferente do valor do sistema, enquanto o filtro robusto apresentou um resultado mais próximo do valor esperado. Para a estimação da frequência, nota-se que o filtro robusto convergiu para o valor correto, enquanto o filtro de Kalman apresenta um comportamento aparentemente divergente.

Na Figura 5.8, observa-se que ambos os filtros apresentam uma variância semelhante. Entretanto, a variância real aproximada do filtro clássico está acima do valor fornecido por sua matriz de covariância, enquanto o resultado do filtro robusto encontra-se sempre abaixo do seu limite superior. Tal problema com o filtro tradicional já é conhecido na literatura e foi apresentado em [73], por exemplo. Para a estimação da frequência, o filtro robusto apresentou tanto uma menor variância quanto um limite superior menos conservador que o filtro de Kalman clássico.

5.4.2 Demodulador com Distúrbio de Fase

O próximo sistema é similar ao apresentado anteriormente, exceto pelo fato de haver uma perturbação desconhecida que altera a fase do modelo. Este sistema pode ser representado como

$$x_{k+1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 0.1\delta_{w,k} \\ 0.0001\delta_{w,k} \end{bmatrix} w_k + \begin{bmatrix} \delta_{v,k} & \delta_{v,k} \\ \delta_{v,k} & 0.0001 + \delta_{v,k} \end{bmatrix} v_k, \quad (5.17)$$

$$y_k = x_{1,k} \cos\left(2\pi x_{2,k} \frac{k}{50} + \theta\right) + (0.01 + 0.1\delta_{w,k}) w_k + \begin{bmatrix} 0.1 + \delta_{v,k} & 0.1 + \delta_{v,k} \end{bmatrix} v_k, \quad (5.18)$$

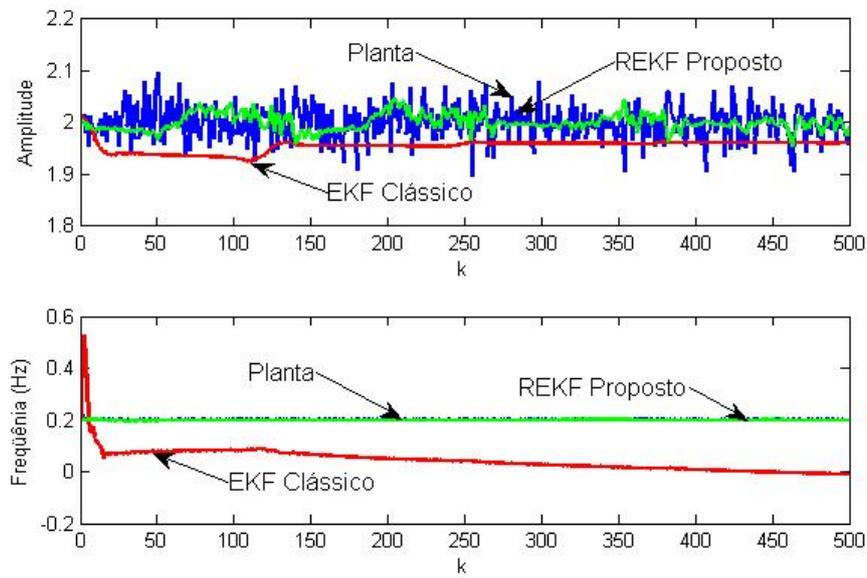


Figura 5.7: Média dos estados após 500 experimentos.

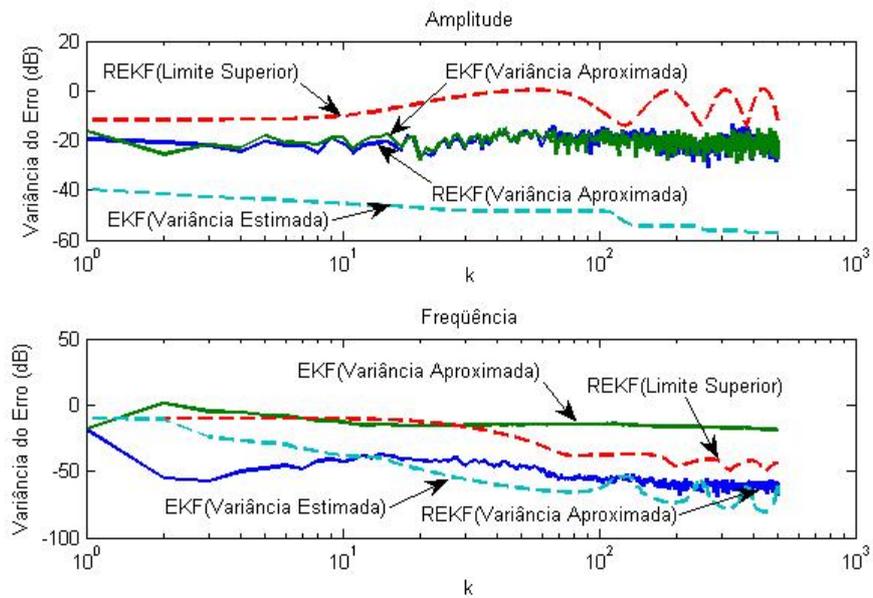


Figura 5.8: Variância aproximada do erro de estimação após 500 experimentos.

em que θ é uma variável desconhecida que varia uniformemente entre 0 e 25 graus. Os ruídos possuem média nula, ou seja, $\bar{w}_k = 0$ e $\bar{v}_k = [0 \ 0]^T$. As variâncias dos ruídos são unitárias da forma $W_k = 1$ e $V_k = \text{diag}\{1, 1\}$ com condições iniciais $\bar{x}_0 = [4 \ 0.2]^T$ e $X_0 = 0.1I$. As matrizes de incertezas foram escolhidas de acordo com (5.12). Apesar de a perturbação ser desconhecida, pode-se notar seu efeito máximo sobre a jacobiana¹. Assim,

$$\Delta_{a,k} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \Delta_{c,k} = \begin{bmatrix} \cos\left(25\frac{\pi}{180}\right) & 0.4\cos\left(25\frac{\pi}{180}\right) \\ 0 & 0 \end{bmatrix}. \quad (5.19)$$

Os resultados obtidos são apresentados nas Figuras 5.9 e 5.10. Nota-se que a estimativa robusta da amplitude apresenta uma oscilação próxima do o valor correto e que o filtro de Kalman tradicional convergiu para valores incorretos em ambos os estados. Os mesmos comentários acerca das variâncias da subseção anterior também são válidos para este exemplo.

5.4.3 Demodulador com Variação de Frequência e com Distúrbio de Fase

O exemplo desta subseção é um demodulador com um distúrbio de fase maior e com frequência variante no tempo. A variação da frequência é uma onda quadrada com amplitude mínima de 0.8 e amplitude máxima de 1.2. Considerando que θ seja uma variável desconhecida que varia uniformemente entre 0 e 25

¹Para saber o efeito máximo da perturbação, para este caso, basta calcular a jacobiana analiticamente.

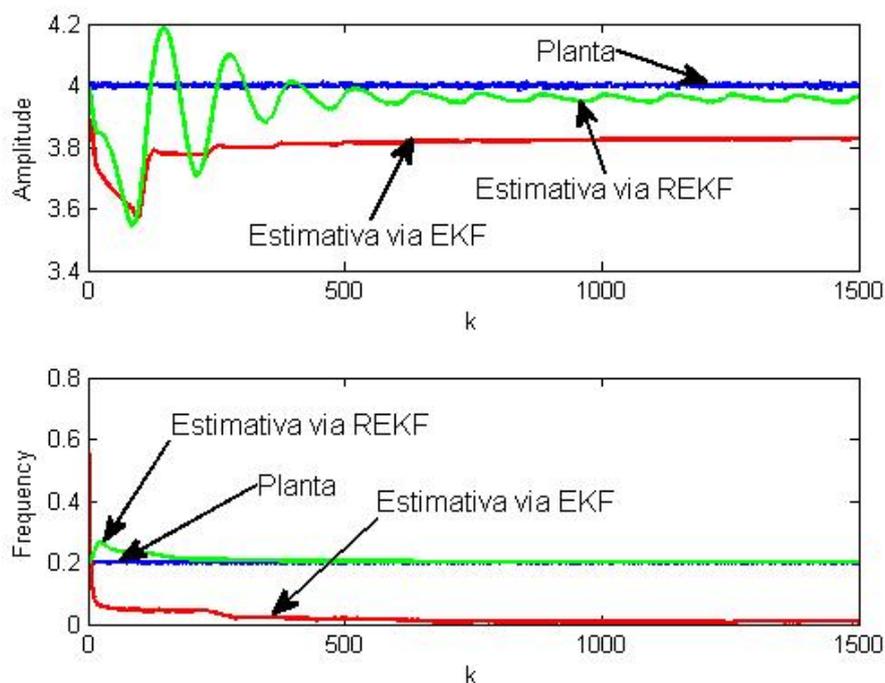


Figura 5.9: Média dos estados após 1500 experimentos.

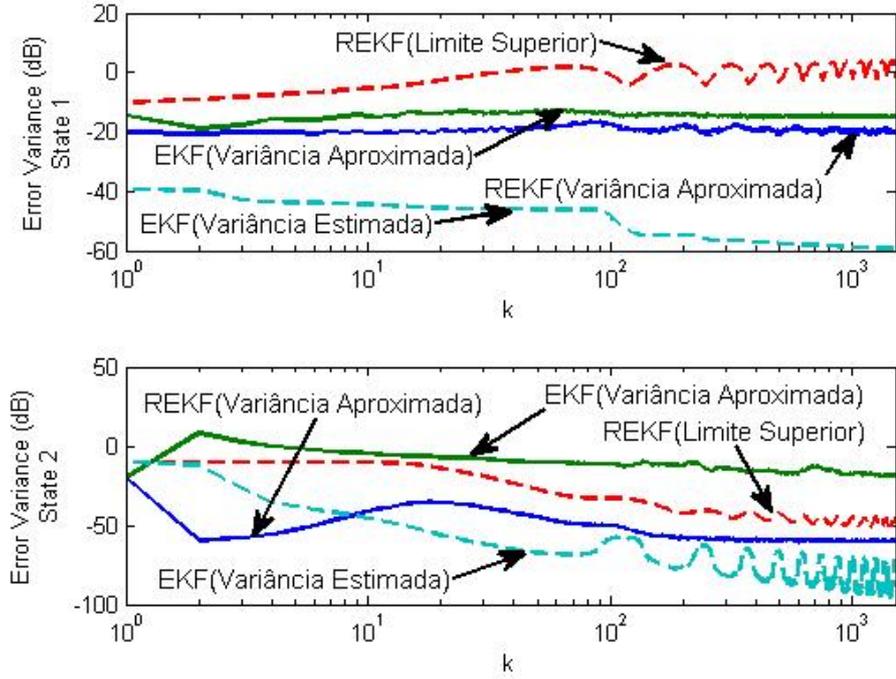


Figura 5.10: Variância aproximada do erro de estimação após 1500 experimentos.

graus, este sistema possui o seguinte comportamento

$$x_{k+1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} g_k(x_k) + \begin{bmatrix} 0.1\delta_{w,k} \\ 0.0001\delta_{w,k} \end{bmatrix} w_k + \begin{bmatrix} \delta_{v,k} & \delta_{v,k} \\ \delta_{v,k} & 0.0001 + \delta_{v,k} \end{bmatrix} v_k, \quad (5.20)$$

$$y_k = x_{1,k} \cos\left(2\pi x_{2,k} \frac{k}{50} + \theta\right) + (0.01 + 0.1\delta_{w,k}) w_k + \begin{bmatrix} 0.1 + \delta_{v,k} & 0.1 + \delta_{v,k} \end{bmatrix} v_k, \quad (5.21)$$

em que, sendo N o número total de iterações do experimento. A função $g_k(x_k)$ é uma onda quadrada para a frequência, de forma que $g_k(x_k) = x_k$, se $\sin\left(4\pi \frac{k}{N}\right) > 0$; caso contrário, tem-se que $g_k(x_k) = [x_{1,k} \ 1.1x_{2,k}]^T$. As propriedades estatísticas dos ruídos são as mesmas dos exemplos anteriores. As condições iniciais são dadas por $\bar{x}_0 = [1 \ 0.2]^T$ e $X_0 = 0.1I$. A saída deste sistema é mostrada na Figura 5.11.

As matrizes relativas às incertezas dos ruídos também são as mesmas escolhidas anteriormente. As matrizes associadas às incertezas das jacobianas foram escolhidas como sendo

$$\Delta_{a,k} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}, \quad \Delta_{c,k} = \begin{bmatrix} 2.4 & 2.4 \end{bmatrix}. \quad (5.22)$$

A escolha da matriz $\Delta_{a,k}$ reflete a incerteza de 10 por cento sobre o valor da frequência. Os parâmetros escalares foram escolhidos conforme (5.14)-(5.16).

Os resultados para 500 experimentos são apresentados nas Figuras 5.12 e 5.13. A análise da Figura 5.12 mostra que ambos os filtros não convergiram para o valor correto da amplitude. Entretanto, o filtro robusto aproximou-se mais do valor correto. Ambos os filtros também não foram capazes de rastrear corretamente a frequência, apesar de o filtro proposto apresentar um erro menor.

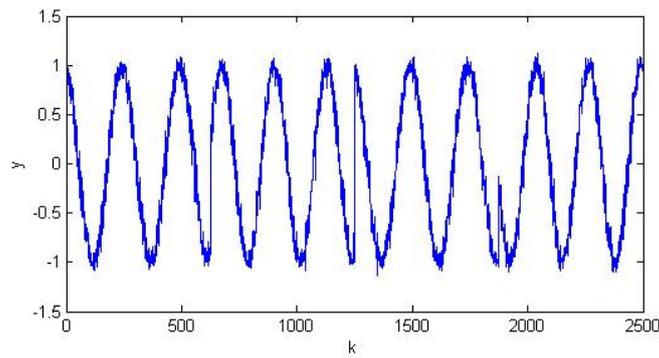


Figura 5.11: Exemplo de sinal de saída para o demodulador com frequência variante no tempo de distúrbio de fase de até 25 graus.

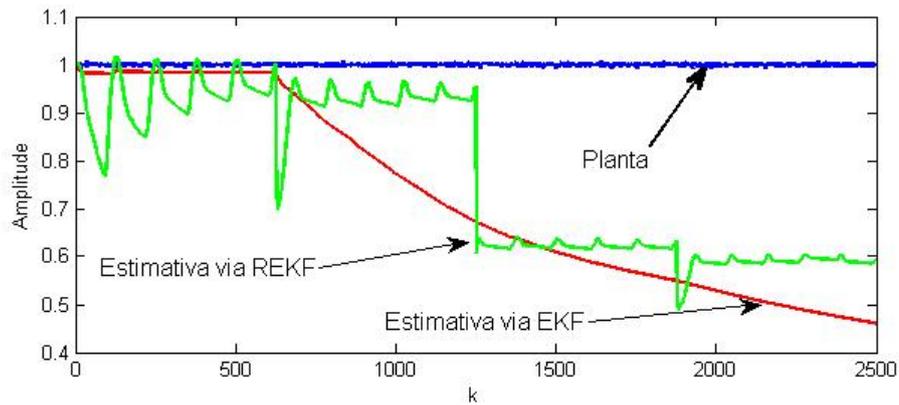


Figura 5.12: Média da amplitude após 500 experimentos.

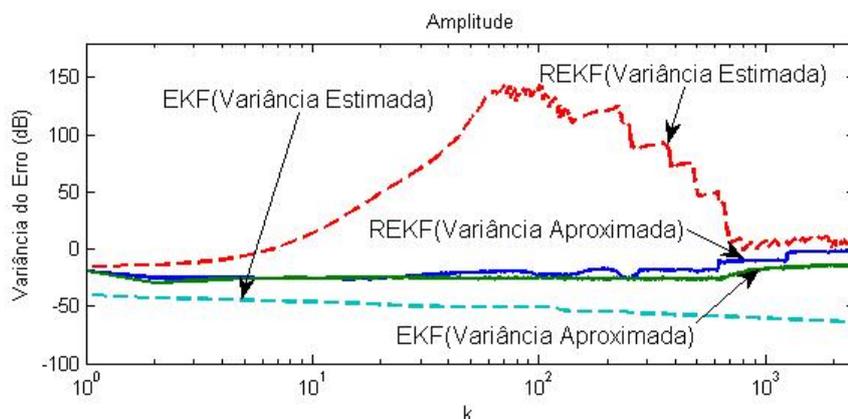


Figura 5.13: Variância aproximada do erro de estimação da amplitude após 500 experimentos.

6 CONCLUSÕES

O enfoque principal desta dissertação foi o projeto de estimadores robustos para sistemas lineares e não-lineares com incertezas limitadas em norma em seus parâmetros. Neste sentido, a abordagem por meio de equações de Riccati foi utilizada para o projeto dos estimadores robustos, desenvolvendo-se estimadores com desempenho garantido. Como contribuições deste estudo, pode-se destacar: um projeto menos conservador em relação a um preditor recentemente encontrado na literatura, permitir que os ruídos de processo e de medição possuam média e covariância-cruzada incertas, bem como dimensões distintas. Também foi desenvolvido o projeto de um filtro não-linear para ruídos com propriedades estatísticas incertas, apresentando um melhor desempenho quando comparado ao filtro de Kalman tradicional. Ainda foi mostrada uma condição mais geral para um lema conhecido na literatura para a obtenção de limites superiores de matrizes de covariância incertas. Foi apresentada uma maneira menos comum para a obtenção do filtro de Kalman tradicional.

Durante a realização deste estudo, os seguintes artigos foram publicados, ou aceitos para publicação, em congressos internacionais: [126], [110] e [127]. O seguinte artigo foi aceito para publicação em revista internacional: [128]. Em se tratando de congressos nacionais, publicou-se o artigo e [129].

Trabalhos futuros podem ser realizados no sentido de se projetarem estimadores robustos para sistemas descritores a tempo discreto. Extensões para o caso de horizonte de tempo infinito também podem ser desenvolvidas. No caso do preditor não-linear, pode-se investigar técnicas de otimização numérica para o cálculo dos parâmetros do estimador robusto.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] XIONG, K.; ZHANG, H.; LIU, L. Adaptive robust extended Kalman filter for nonlinear stochastic systems. *IET Control Theory and Applications*, v. 2, n. 3, p. 239–250, 2008.
- [2] WANG, F.; BALAKRISHNAN, V. Robust steady-state filtering for systems with deterministic and stochastic uncertainties. *IEEE Transactions on Signal Processing*, v. 51, n. 10, p. 2550–2558, Outubro 2003.
- [3] WANG, F.; BALAKRISHNAN, V. Robust Kalman filters for linear time varying systems with stochastic parametric uncertainties. *IEEE Transactions on Signal Processing*, v. 50, n. 4, p. 803–813, April 2002.
- [4] TARIGHAT, A.; SUBRAMANIAN, A.; SAYED, A. H. An uplink ds-cdma receiver using a robust post-correlation Kalman structure. *IEEE Transactions on Wireless Communications*, v. 4, n. 3, p. 847–854, May 2005.
- [5] SUBRAMANIAN, A.; SAYED, A. H. Joint rate and power control algorithms for wireless networks. *IEEE Transactions on Signal Processing*, v. 53, n. 11, p. 4204–4214, November 2005.
- [6] KOSANAM, S. *Kalman Filtering for Uncertain Noise Covariances*. Tese (Doutorado) — Departamento de Engenharia Elétrica, Universidade do Estado de Cleveland, 2000.
- [7] SIMON, D.; SIMON, D. L. Aircraft turbofan engine health estimation using constrained Kalman filtering. *ASME Journal of Engineering for Gas Turbines and Power*, v. 127, n. 2, p. 323–328, April 2005.
- [8] SIMON, D.; SIMON, D. L. Kalman filter constraint switching for turbofan engine health estimation. *European Journal of Control*, v. 12, n. 3, p. 331–343, May 2006.
- [9] STRATTON, D. A.; STENGEL, R. E. Robust Kalman filter design for predictive wind shear detection. *IEEE Transactions on Aerospace and Electronic Systems*, v. 39, n. 4, p. 1185–1194, October 1993.
- [10] CONG, S.; HONG, L.; LAYNE, J. Iterative robust filtering for ground target tracking. *IET Control Theory Appl*, v. 1, n. 1, p. 2550–2558, January 2007.
- [11] KAPPL, J. J. Nonlinear estimation via Kalman filtering. *IEEE Transactions on Aerospace and Electronic Systems*, v. 7, n. 1, p. 79–84, 1971.
- [12] CAMPBELL, M. E.; BRUNKE, S. Nonlinear estimation of aircraft models for on-line control customization. In: *Proceedings of the IEEE Aerospace Conference*. [S.l.: s.n.], 2001. v. 2, p. 621–628.
- [13] JAZWINSKI, A. H. *Stochastic Processes and Filtering Theory*. [S.l.]: Academic Press, 1970.
- [14] CAMPBELL, J. K.; SYNNOTT, S. P.; BIERMAN, G. J. Voyager orbit determination at jupiter. *IEEE Transactions on Automatic Control*, v. 28, n. 3, p. 256–269, 1983.
- [15] LEFFERTS, E. J.; MARKLEY, F. L.; SHUSTER, M. D. Kalman filtering for spacecraft attitude estimation. *Journal of Guidance*, v. 5, n. 5, p. 417–429, 1982.

- [16] PSIAKI, M. L.; HUANG, L.; FOX, S. M. Ground tests of magnetometer-based autonomous navigation (magnav) for low-earthorbiting spacecraft. *Journal of Guidance, Control and Dynamics*, v. 16, n. 1, p. 206–214, 1993.
- [17] PAYNTER, S. J.; BISHOP, R. H. Adaptive nonlinear attitude control and momentum management of spacecraft. *Journal of Guidance, Control and Dynamics*, v. 20, n. 5, p. 1025–1032, 1997.
- [18] WU, A. SBIRS high payload LOS attitude determination and calibration. In: *Proceedings of the IEEE Aerospace Conference*. [S.l.: s.n.], 1998. v. 05, p. 243–253.
- [19] MISU, T.; HASHIMOTO, T.; NINOMIYA, K. Optical guidance for autonomous landing of spacecraft. *IEEE Transactions on Aerospace and Electronic Systems*, v. 35, n. 2, p. 459–473, 1999.
- [20] DELGADO, E.; BARREIRO, A. Sonar-based robot navigation using nonlinear robust observers. *Automatica*, v. 39, p. 1195–1203, 2003.
- [21] FUNG, P. T.-K.; GRIMBLE, M. J. Dynamic ship positioning using a self-tuning Kalman filter. *IEEE Transactions on Automatic Control*, v. 28, n. 3, p. 256–269, 1983.
- [22] FOSSEN, T. I.; SAGATUN, S. I.; SORENSEN, A. J. Identification of dynamically positioned ships. *Control Engineering Practice*, v. 04, n. 3, p. 369–376, 1996.
- [23] CACCIA, M.; BRUZZONE, G.; VERUGGIO, G. Active sonar-based bottom-following for unmanned underwater vehicles. *Control Engineering Practice*, n. 7, p. 459–468, 1999.
- [24] HOHMAN, D. et al. GPS roadside integrated precision positioning system. In: *Proceedings of the IEEE position location and navigation symposium*. [S.l.: s.n.], 2000. p. 221–230.
- [25] RAY, L. R. Nonlinear tire force estimation and road friction identification: Simulation and experiments. *Automatica*, v. 33, n. 10, p. 1819–1833, 1997.
- [26] THAM, Y. K.; WANG, H.; TEOH, E. K. Multi-sensor fusion for steerable four-wheeled industrial vehicles. *Control Engineering Practice*, n. 7, p. 1233–1248, 1999.
- [27] LEONARD, J. J.; DURRANT-WHYTE, H. F. Mobile robot localization by tracking geometric beacons. *IEEE Transactions on Robotics and Automation*, v. 7, n. 3, p. 376–382, 1991.
- [28] MUN-LI, H.; KLEEMAN, L. Ultrasonic classification and location of 3D room features using maximum likelihood estimation part i. *Robotica*, v. 15, p. 483–491, 1997.
- [29] NILSSON, B. et al. Control of flexible mobile manipulators: Positioning and vibration reduction using an eye-in-hand range camera. *Control Engineering Practice*, n. 7, p. 741–751, 1999.
- [30] SPIRITO, M. A. Further results on gsm mobile station location. *Electronics Letters*, v. 35, n. 11, p. 867–869, 1999.
- [31] RIVALS, I.; PERSONNAZ, L. A recursive algorithm based on the extended Kalman filter for the training of feedforward neural models. *Neurocomputing*, n. 20, p. 279–294, 1998.
- [32] CHEE, Y. K.; SOH, Y. C. Robust Kalman filter design. In: *Proceedings of 39th IEEE Conference on Decision and Control*. [S.l.: s.n.], 2000. v. 3, p. 3813–3818.

- [33] MILLS, S.; PRIDMORE, T.; HILLS, M. Tracking in a hough space with the Extended Kalman Filter. In: *Proceedings of the British Machine Vision Conference*. [S.l.: s.n.], 2003. p. 173–182.
- [34] CZYZEWSKI, A.; DALKA, P. Examining Kalman filters applied to tracking objects in motion. *2008 Ninth International Workshop on Image Analysis for Multimedia Interactive*, IEEE Computer Society, v. 7-9, p. 175–178, 2008.
- [35] ZHOU, J.; BISCHOFY, W. F.; CAELLI, T. Robust and efficient road tracking in aerial images. In: *CMRT Conference*. [S.l.: s.n.], 2005. v. 36, p. 35–40.
- [36] MASOUD, O.; PAPANIKOLOPOULOS, N. Robust pedestrian tracking using a model-based approach. In: *IEEE Conference on Intelligent Transportation System*. [S.l.: s.n.], 1997. p. 338–343.
- [37] DONEY, K. L. *Leak Detection in Pipelines using the Extended Kalman Filter and Extended Boundary Approach*. Tese (Doutorado) — Department of Mechanical Engineering, University of Saskatchewan, 2007.
- [38] LI, Q.; MARK, R. G.; CLIFFORD, G. D. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. *Physiological Measurement*, v. 29, p. 15–32, 2008.
- [39] HOTEIT, I.; PHAM, D. *An Adaptively Reduced Order Extended Kalman Filter for Data Assimilation in the Tropical Pacific*. 655, avenue de l Europe, 38330 Montbonnot St Martin - France, 2001.
- [40] BAHMANI, O.; BROWN, F. Kalman filter approach to estimate the demand for international reserves. *Applied Economics*, v. 36, n. 15, p. 1655–1668, 2004.
- [41] CHEUNG, Y. Exchange rate risk premiums. *Journal of International Money and Finance*, v. 12, p. 182–194, 1993.
- [42] BHAR, R.; CHIARELLA, C. Interest rate futures: Estimation of volatility parameters in an arbitrage-free framework. In: *Proceedings of the IEEE conference on computational intelligence for financial engineering*. [S.l.: s.n.], 1996. p. 168–182.
- [43] HANSEN, L. P.; SARGENT, T. J. *Robust Control and Filtering for Macroeconomics*. [S.l.: s.n.], 2002.
- [44] LIN, Y.; JIN, X.; XUE, A. Robust multisensor fusion estimation for multisensor. In: *Proceedings of the International Conference on Sensing, Computing and Automation*. [S.l.: s.n.], 2006. p. 412–1889.
- [45] BELAVKIN, V. P. Optimization of quantum observation and control. In: *Proceedings of the 9th IFIP Conference on Optimization Techniques*. [S.l.: s.n.], 1979.
- [46] BELAVKIN, V. P. Theory of the control of observable quantum systems. *Automatica and Remote Control*, v. 44, n. 2, p. 178–188, 1983.
- [47] BELAVKIN, V. P. Non-demolition measurement and control in quantum dynamical systems. In: *Information Complexity and Control in Quantum Systems*. [S.l.: s.n.], 1987. p. 311–329.

- [48] BELAVKIN, V. P. Non-demolition measurements, nonlinear filtering and dynamic programming of quantum stochastic processes. In: *Lecture Notes in Control and Inform Sciences 121*. [S.l.: s.n.], 1989. p. 245–265.
- [49] BELAVKIN, V. P. Non-demolition stochastic calculus in fock space and nonlinear filtering and control in quantum systems. In: *Stochastic Methods in Mathematics and Physics*. [S.l.: s.n.], 1989. p. 310–324.
- [50] STOCKTON, J. K. et al. Robust quantum parameter estimation: Coherent magnetometry with feedback. *Physical Review A*, v. 69, n. 3, p. 032109, 2004.
- [51] GEREMIA, J. et al. Quantum Kalman filtering and the Heisenberg Limit in atomic magnetometry. *Phys. Review Letters*, v. 91, n. 25, p. 250801, 2003.
- [52] BOUTEN, L.; HANDEL, R. van; JAMES, M. An introduction to quantum filtering. *SIAM J. Control Optim.*, v. 46, p. 2199–2241, 2007.
- [53] GALILEI, G. *Dialogo Sopra i due Sestemi del Mondo: Tolemaico e Copernicano*. [S.l.]: Landini (Translation: University of California Press, Berkeley, CA, 1957), 1632.
- [54] GAUSS, C. F. *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientum*. [S.l.]: (Translation: Dover, New York, 1963), 1809.
- [55] MARTIN, C. J.; MINTZ, M. Methode des moindres quarres, pour trouver le milieu le plus probable entre les resultats de diffrentes observations. *Mem. Inst. France*, p. 149–154, 1810.
- [56] ADRIAN, R. Research concerning the probabilities of the errors which happen in making observations. *Analyst*, v. 1, p. 193–209, 1808.
- [57] STRANG, G. *Introduction to Linear Algebra*. [S.l.]: Wellesley-Cambridge Press, 1993.
- [58] KOLMOGOROV, A. N. Sur l'interpolation et extrapolation des suites stationairres. *C. R. Acad. Sci.*, p. 208–2043, 1939.
- [59] KREIN, M. G. On a generalization of some investigations of G. Szego, W. M. Smirnov and A. N. Kolmogorov. *Dokl. Akad. Nauk SSSR*, v. 46, p. 208–2043, 1945.
- [60] WIENER, N. *Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*. [S.l.]: Technology Press and Wiley, New York, NY, 1949.
- [61] HASSIBI, B. *Indefinite Metric Spaces in Estimation, Control and Adaptative Filtering*. Tese (Doutorado) — Department of Electrical Engineering, Stanford University, 1996.
- [62] ASTROM, K. J. *Introduction to Stochastic Control Theory*. [S.l.]: Academic Press, 1970.
- [63] KALMAN, R. E. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering, Transactions of the ASME, Series D*, v. 82, n. 1, p. 35–45, 1960.
- [64] ANDERSON, B. D. O.; MOORE, J. B. *Optimal Filtering*. [S.l.]: Prentice-Hall, 1979.

- [65] XIE, L.; SOUZA, C. E. de; FU, M. H_∞ estimation for discrete-time uncertain systems. *International Journal Robust and Nonlinear Control*, v. 1, n. 1, p. 111–123, 1991.
- [66] LI, H.; FU, M. A linear matrix inequality approach to robust H_∞ filtering. *IEEE Transactions on Signal Processing*, v. 45, n. 9, p. 2338–2350, September 1997.
- [67] ZHU, X.; SOH, Y. C.; XIE, L. Robust Kalman filter design. In: *Proceedings of 39th IEEE Conference on Decision and Control*. [S.l.: s.n.], 2000. p. 3813–3818.
- [68] SAYED, A. H. A framework for state space estimation with uncertain models. *IEEE Transactions on Automatic Control*, v. 46, n. 7, p. 998–1013, July 2001.
- [69] SUBRAMANIAN, A.; SAYED, A. H. A robust minimum-variance filter for time varying uncertain discrete-time systems. In: *Proceedings of the American Control Conference*. [S.l.: s.n.], 2003. p. 1885–1889.
- [70] ISHIHARA, J. Y.; TERRA, M. H.; CAMPOS, J. C. T. Robust Kalman filter for descriptor systems. *IEEE Transactions on Automatic Control*, v. 51, n. 8, p. 1354–1358, August 2006.
- [71] FU, M.; SOUZA, C. E. de; LUO, Z.-Q. Finite-horizon robust Kalman filter design. *IEEE Transactions on Signal Processing*, v. 49, n. 9, p. 2103–2112, September 2001.
- [72] TERRA, M. H.; SAYED, A. H. Design of robust filters with improved robustness margins via parameter scaling. In: *Proceedings of the 40th IEEE Conference on Decision and Control*. [S.l.: s.n.], 2001. p. 664–669.
- [73] GHAOUI, L. E.; CLAFIORE, G. Robust filtering for discrete-time systems with bounded noise and parametric uncertainty. *IEEE Transactions on Automatic Control*, v. 46, n. 7, p. 1084–1089, July 2001.
- [74] WANG, Z.; QIAO, H. Robust filtering for bilinear uncertain stochastic discrete-time systems. *IEEE Transactions on Signal Processing*, v. 50, n. 3, p. 560–567, March 2002.
- [75] SUBRAMANIAN, A.; SAYED, A. H. Regularized robust estimators for time varying uncertain discrete-time systems. In: *Proceedings of the 13th IFAC Symposium on System Identification*. [S.l.: s.n.], 2003.
- [76] SUBRAMANIAN, A.; SAYED, A. H. Regularized robust filters for time-varying uncertain discrete-time systems. *IEEE Transactions on Automatic Control*, v. 49, n. 6, p. 970–976, June 2004.
- [77] SUBRAMANIAN, A.; SAYED, A. H. Multiobjective filter design for uncertain stochastic time-delay systems. *IEEE Transactions on Automatic Control*, v. 49, n. 1, p. 149–154, January 2004.
- [78] SUN, K.; PACKARD, A. Robust H_2 and H_∞ filters for uncertain lft systems. *IEEE Transactions on Automatic Control*, v. 50, n. 5, p. 715–720, May 2005.
- [79] SOUZA, C. E. de; TROFINO, A. Advances in linear matrix inequality methods in control. In: SIAM (Ed.). *A Linear Matrix Inequality Approach to the Design of Robust H_2 Filters*. [S.l.]: SIAM, 2005. p. 175–185.
- [80] GEROMEL, J. C.; BERNUSSOU, J.; GARCIA, G. Siam j. control optim. H_2 and H_∞ Robust Filtering for Discrete-Time Linear Systems, v. 38, n. 5, p. 175–185, May 2005.

- [81] WANG, Z. et al. Robust finite-horizon filtering for stochastic systems with missing measurements. *IEEE Signal Processing Letters*, v. 12, n. 6, p. 437–440, June 2005.
- [82] DONG, Z.; YOU, Z. Finite horizon robust Kalman filtering for uncertain discrete time-varying systems with uncertain-covariance white noises. *IEEE Signal Processing Letters*, v. 13, n. 8, p. 493–496, August 2006.
- [83] SOUZA, C. E. de; BARBOSA, K. A.; TROFINO, A. Robust filtering for linear systems with convex-bounded uncertain time-varying parameters. *IEEE Transactions on Automatic Control*, v. 52, n. 6, p. 1132–1138, June 2007.
- [84] SKOGESTAD, S.; POSTLETHWAITE, I. *Multivariable Feedback Control - Analysis and design*. [S.l.]: John Wiley and Sons, 2001.
- [85] BARBOSA, K. A. *Projeto de Filtros Robustos para Sistemas Lineares e Não Lineares*. Tese (Doutorado) — Departamento de Engenharia Elétrica, Universidade Federal de Santa Catarina, 2003.
- [86] XIE, L. On robust H_2 estimation. *Automatica Sinica*, v. 31, n. 1, p. 1–12, January 2005.
- [87] XING, Z. *Analysis and Design of Discrete Time Robust Filters*. Tese (Doutorado) — School of Electrical and Electronic Engineering, Nanyang Technological University, 2001.
- [88] XIE, L.; SOH, Y. C.; SOUZA, C. E. de. Robust Kalman filtering for uncertain discrete-time systems. *IEEE Transactions on Automatic Control*, v. 39, n. 6, p. 1310–1314, June 1994.
- [89] ZHU, X.; SOH, Y. C.; XIE, L. Design and analysis of discrete-time robust Kalman filters. *Automatica*, v. 38, n. 6, p. 1069–1077, March 2002.
- [90] THEODOR, Y.; SHAKED, U. Robust discrete-time minimum variance filtering. *IEEE Transactions on Signal Processing*, v. 44, n. 2, p. 181–189, February 1996.
- [91] GEROMEL, J. C.; OLIVEIRA, M. C. de. H_2 and H_∞ robust filtering for convex bounded uncertain systems. In: *Proceedings of the 37th IEEE Conference on Decision and Control*. [S.l.: s.n.], 1998. p. 146–151.
- [92] GEROMEL, J. C. Optimal linear filtering under parameter uncertainty. *IEEE Transactions on Signal Processing*, v. 47, n. 1, p. 168–175, January 1999.
- [93] SHAKED, U.; XIE, L.; SOH, Y. C. New approaches to robust minimum variance filter design. *IEEE Transactions on Signal Processing*, v. 49, p. 2620–2629, 2001.
- [94] GEROMEL, J. C.; KOROGUI, R. H. H_2 robust filter design with performance certificate via convex programming. *Automatica*, v. 44, n. 4, p. 937–948, 2008.
- [95] ELДАР, Y. C.; MERHAV, N. Robust linear estimation under a minimax mse-ratio criterion. *IEEE Transactions on Signal Processing*, v. 53, n. 4, p. 1335–1347, 2005.
- [96] MARTIN, C. J.; MINTZ, M. Robust filtering and prediction for linear systems with uncertain dynamics: A game - theoretic approach. *IEEE Transactions on Automatic Control*, v. 28, n. 9, p. 888–896, 1983.

- [97] OHRN, K.; AHLEN, A.; STERNAD, M. A probabilistic approach to multivariable robust filtering, prediction and smoothing. In: *Proceedings of the 32th IEEE Conference on Decision and Control*. [S.l.: s.n.], 1993. p. 1227–1232.
- [98] ZAMES, G. Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms, and approximate inverses. *IEEE Transactions on Automatic Control*, v. 26, p. 301–320, 1981.
- [99] FU, M.; SOUZA, C. E. de; XIE, L. H_∞ estimation for uncertain systems. *International Journal Robust and Nonlinear Control*, v. 2, n. 2, p. 87–105, 1992.
- [100] TEIXEIRA, B. Ask the experts - Present and Future. *IEEE Control Systems Magazine*, IEEE Computer Society, v. 4, p. 16–18, 2008.
- [101] FEYNMAN, R.; LEIGHTON, R.; SANDS, M. *The Feynman Lectures on Physics: Volume 1*. [S.l.]: Addison-Wesley, Reading, Massachusetts, 1963.
- [102] ASTROM, K. J.; MURRAY, R. M. *Feedback Systems: An Introduction for Scientists and Engineers*. [S.l.]: Princeton University Press, 2008.
- [103] SIMON, D. *Optimal State Estimation*. [S.l.]: John Wiley and Sons, 2006.
- [104] KAILATH, T.; SAYED, A. H.; HASSIBI, B. *Linear Estimation*. [S.l.]: Prentice-Hall, 2000.
- [105] DAUM, F. Nonlinear filters: Beyond the Kalman filter. *IEEE Transactions on Aerospace and Electronic Systems*, v. 20, n. 8, p. 57–69, August 2005.
- [106] REED, M.; SIMON, B. *Methods of modern mathematical physics*. [S.l.]: Academic Press, 1972.
- [107] BARBOSA, K. A.; SOUZA, C. E. de; TROFINO, A. Robust H_2 filtering for uncertain linear systems: LMI based methods with parametric lyapunov functions. *Syst. Contr. Lett.*, v. 54, n. 3, p. 251–262, 2005.
- [108] SOUZA, C. E. de; BARBOSA, K. A.; FU, M. Robust filtering for uncertain linear discrete-time descriptor systems. *Automatica*, v. 44, n. 3, p. 792–798, 2008.
- [109] PETERSEN, I. R.; MCFARLANE, D. C. Robust state estimation for uncertain linear systems. In: *Proceedings of the 30th IEEE Conference on Decision and Control*. [S.l.: s.n.], 1991. p. 2630–2631.
- [110] SOUTO, R. F.; ISHIHARA, J. Y. Robust Kalman filter for discrete-time systems with correlated noises. In: *Proceedings of the 16th IEEE Mediterranean Conference on Control and Automation*. [S.l.: s.n.], 2008. p. 1658–1662.
- [111] WANG, Z.; ZHU, J.; UNBEHAUEN, H. Robust filter design with time-varying parameter uncertainty and error variance constraints. *International Journal of Control*, v. 72, n. 1, p. 30–38, 1999.
- [112] PETERSEN, I. Robust guaranteed cost state estimation for nonlinear stochastic uncertain systems via an IQC approach. In: *Proceedings of the Conference on Information Decision and Control*. [S.l.: s.n.], 2007. p. 17–22.
- [113] HOANG, N. T. et al. Robust filtering for discrete nonlinear fractional transformation systems. *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS*, v. 51, n. 11, p. 520–526, 2004.

- [114] PERALA, T.; PICHE, R. Robust extended Kalman filtering in hybrid positioning applications. In: *4th WORKSHOP ON POSITIONING, NAVIGATION AND COMMUNICATION*. [S.l.: s.n.], 2007. p. 55–63.
- [115] YU, M.-J.; LEE, S. W. A robust extended filter design for SDINS in-flight alignment. *International Journal of Control, Automation, and Systems*, v. 1, n. 4, p. 520–526, 2003.
- [116] EINICKE, G. A.; WHITE, L. B. Robust extended Kalman filtering. *IEEE Transactions on Signal Processing*, v. 47, p. 2596–2599, 1999.
- [117] ZHANG, Y. et al. A svd-based extended Kalman filter and applications to aircraft. In: *Proceedings of American Control Conference*. [S.l.: s.n.], 1994. p. 1809–1813.
- [118] YU, M. J.; LEE, S. W. A robust extended filter design for sdins in-flight alignment. *International Journal of Control, Automation and Systems*, v. 1, p. 520–526, 2003.
- [119] PATHIRANA, A. V. S. P. N.; JHA, S. Robust extended Kalman filter applied to location tracking and trajectory prediction for pcs networks. In: *Proceedings of the IEEE International Conference on Control Applications*. [S.l.: s.n.], 2004. p. 63–68.
- [120] WU, Z.; SHI, H.; LIU, B. Ins/gps integrated navigation uncertain system state estimation based on minimal variance robust filtering. In: *Multiconference on Computational Engineering in Systems Applications*. [S.l.: s.n.], 2006. p. 631–634.
- [121] GILMAN., A.; RHODES, I. Cone-bounded nonlinearities and mean-square bounds - Estimation upper bound. *IEEE Transactions on Automatic Control*, v. 18, n. 3, p. 260–260, 1973.
- [122] GILMAN., A.; RHODES, I. Cone-bounded nonlinearities and mean-square bounds - Estimation lower bound. *IEEE Transactions on Automatic Control*, v. 20, n. 5, p. 632–642, 1975.
- [123] CHEN, L.; ARAMBEL, P. O.; MEHRA, R. K. Estimation under unknown correlation: Covariance intersection revisited. *IEEE Transactions on Automatic Control*, v. 47, p. 1879–1882, 2002.
- [124] TREES, H. L. V. *Detection, Estimation, and Modulation Theory: Radar-Sonar Signal Processing and Gaussian Signals in Noise*. Melbourne, FL, USA: Krieger Publishing Co., Inc., 1992. ISBN 0894647482.
- [125] KOOTSOOKOS, P. J. *A Review of the Frequency Estimation and Tracking Problems*. 1991.
- [126] SOUTO, R. F.; ISHIHARA, J. Y. Enhanced robust Kalman predictor for discrete-time systems with uncertain correlated noises. In: *Proceedings of the 47th IEEE Conference on Decision and Control*. [S.l.: s.n.], 2008. p. 1660–1665.
- [127] SOUTO, R. F.; ISHIHARA, J. Y. A bounded nonlinear filter with uncertain correlated noises. In: *Proceedings of the American Control Conference*. [S.l.: s.n.], 2009. p. To be published.
- [128] SOUTO, R. F.; ISHIHARA, J. Y. Comments on “finite-horizon robust Kalman filtering for uncertain discrete time-varying systems with uncertain-covariance white noises”. *IEEE Signal Processing Letters*, n. To be published.

- [129] SOUTO, R. F.; ISHIHARA, J. Y. Filtro de Kalman robusto para horizonte-finito. In: *Proceedings of the XVII Congresso Brasileiro de Automática*. [S.l.: s.n.], 2008.
- [130] DUNCAN, W. Some devices for the solution of large sets of simultaneous linear equations (with an appendix on the reciprocation of partitioned matrices). *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, Seventh Series*, v. 35, p. 660–670, 1944.
- [131] HOUSEHOLDER, A. *Optimal Filtering*. [S.l.]: Prentice-Hall, 1953.
- [132] HENDERSON, H.; SEARLE, S. On deriving the inverse of a sum of matrices. *SIAM Review*, v. 23, n. 1, p. 53–60, January 1981.
- [133] BOYD, S.; VANDENBERGHE, L. *Convex Optimization*. [S.l.]: Cambridge University Press, 2008.
- [134] CAMPBELL, S. L.; MEYER, C. D. *Generalized Inverses of Linear Transformations*. [S.l.]: Dover Publications, 1991.
- [135] RAO, C. R.; MITRA, S. K. *Generalized Inverse of Matrices and its Applications*. [S.l.]: John Wiley & Sons New Yorks, 1971.
- [136] SONTAG, E. D. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*. [S.l.]: Springer-Verlag, 1998.
- [137] TAY, T. T.; MAREELS, I. M. Y.; MOORE, J. B. *High Performance Control*. [S.l.]: Birkhauser Boston, 1997.
- [138] IOANNOU, P. A.; SUN, J. *Robust Adaptive Control*. [S.l.]: Prentice Hall, 1996.

I. RESULTADOS COMPLEMENTARES

Neste anexo, encontram-se conceitos e resultados complementares para a leitura da dissertação.

I.1 LEMA DE INVERSÃO DE MATRIZES

Nesta seção, será apresentado o *Lema de Inversão de Matrizes*, também conhecido como *Identidade de Woodbury*, ou *Fórmula de Sherman-Morrison*, ou *Fórmula de Sherman-Woodbury-Morrison*. Este lema, utilizado diversas vezes ao longo desta dissertação, é especialmente útil em diversas áreas, tais como controle, processamento de sinais e na própria teoria de estimação [103]. Uma das primeiras apresentações deste lema é atribuída a William Duncan em 1944 [130], e identidades similares foram obtidas por Alston Householder em 1953 [131]. Kailath introduziu a sua utilização em problemas de estimação em 1960 e Ho, em 1963, atribuiu-lhe a celebrada denominação de Lema de Inversão de Matrizes. Maiores detalhes sobre as origens, extensões e variações deste lema podem ser encontradas em [132]. Este lema é particularmente útil para se reduzir o esforço para se calcular a inversão matricial [133]. A seguir, apresenta-se o Lema de Inversão de Matrizes.

Lema I.1.1 (*Lema de Inversão de Matrizes*). *Considere as matrizes $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{m \times n}$ e $D \in \mathbb{R}^{m \times m}$, em que A e D são matrizes simétricas e definidas positivas. Então, segue que*

$$(A + BD^{-1}C)^{-1} = A^{-1} - AB(D + CAB)^{-1}CA, \quad (\text{I.1})$$

$$(A - BD^{-1}C)^{-1} = A^{-1} + AB(D - CAB)^{-1}CA. \quad (\text{I.2})$$

Prova ¹Considere

$$E := D - CA^{-1}B, \quad (\text{I.3})$$

$$F := A - BD^{-1}C, \quad (\text{I.4})$$

e que E seja invertível. Efetuando as devidas operações, pode-se observar que

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} A^{-1} + A^{-1}BE^{-1}CA^{-1} & -A^{-1}BE^{-1} \\ -E^{-1}CA^{-1} & E^{-1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \quad (\text{I.5})$$

Assume-se agora que F seja invertível. Após as devidas manipulações algébricas, mostra-se que

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} F^{-1} & -A^{-1}BE^{-1} \\ -D^{-1}CF^{-1} & E^{-1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \quad (\text{I.6})$$

Assim, (I.5) e (I.6) são duas expressões para a inversa de $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$. Como a inversa é única², conclui-se que estas duas expressões devam ser iguais. Olhando para a matriz na posição 1×1 de ambas expressões,

¹Retirada de [103].

²consultar anexo I.2

tem-se que

$$\begin{aligned} F^{-1} &= A^{-1} + A^{-1}BE^{-1}CA^{-1} \\ \Rightarrow (A - BD^{-1}C)^{-1} &= A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}. \end{aligned} \quad (\text{I.7})$$

■

Prova ³Considere o sistema linear

$$(A + BD^{-1}C)x = b, \quad (\text{I.8})$$

em que $A \in \mathbb{R}^{n \times n}$ é não singular, $B \in \mathbb{R}^{n \times p}$ e $C \in \mathbb{R}^{p \times n}$. Introdz-se uma nova expressão $Dy = Cx$ e reescrevem-se as equações como

$$Ax + By = b, \quad (\text{I.9})$$

$$y = D^{-1}Cx. \quad (\text{I.10})$$

Reescrevendo em formato matricial, tem-se que

$$\begin{bmatrix} A & B \\ D^{-1}C & -I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}. \quad (\text{I.11})$$

Nota-se que o coeficiente inicial em (I.8), $A + BD^{-1}C$, é o complemento de Schur de $-I$ na matriz aumentada em (I.11). Em alguns casos, é mais eficiente resolver o sistema com mais equações (I.11) do que o sistema original (I.8). Este seria o caso, por exemplo, em que A , B , C e D^{-1} são relativamente esparsas, embora a matriz $A + BD^{-1}C$ esteja longe de o ser. A partir de (I.9), conclui-se que $x = A^{-1}(b - By)$. Substituindo esta relação em (I.9), tem-se que

$$(D + CA^{-1}B)y = CA^{-1}b, \quad (\text{I.12})$$

de forma que

$$y = (D + CA^{-1}B)^{-1}CA^{-1}b. \quad (\text{I.13})$$

Utilizando novamente que $x = A^{-1}(b - By)$, obtém-se

$$x = \left(A^{-1} - A^{-1}B(D + CA^{-1}B)^{-1}CA^{-1} \right) b. \quad (\text{I.14})$$

Sendo b arbitrário, pode-se afirmar que (I.1) é válida. ■

I.2 MATRIZ INVERSA GENERALIZADA

Nesta seção, será apresentado o conceito de matriz inversa generalizada. Este já é um resultado bastante conhecido e pode ser encontrado em diversos livros de álgebra linear, tais como [134, 135]. As definições a seguir serão importantes nos resultados que se seguirão sobre matrizes inversas generalizadas.

³Extensão do resultado obtido em [133], em que o autor considerou $D = I$.

Definição I.2.1 Considere a matriz $A = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix}$, em que $A \in \mathfrak{C}^{n \times m}$.

- Se existirem escalares não todos nulos λ_i , $i = 1 \dots m$, tais que $\sum_{i=1}^m \lambda_i a_i = 0$, então diz-se que os vetores a_i , $i = 1 \dots m$ são linearmente dependentes (LD).
- Se $\sum_{i=1}^m \lambda_i a_i = 0$ implicar que $\lambda_i = 0$, $\forall i$, então diz-se que os vetores a_i , $i = 1 \dots m$ são linearmente independentes (LI).

Definição I.2.2 O posto de uma matriz A é o número de linhas (e colunas) LI de X . A notação utilizada é $\text{posto}(A)$.

Nota I.2.3 O posto de uma matriz A nula é zero.

Considere o seguinte sistema linear de equações

$$AX = C, \quad (\text{I.15})$$

em que $A \in \mathfrak{C}^{n \times m}$ e $C \in \mathfrak{C}^{n \times q}$ são matrizes conhecidas e $X \in \mathfrak{C}^{m \times q}$ é uma matriz desconhecida. O posto de A é r , tal que $r \leq \min(n, m)$. O problema para se encontrar a matriz X pode ser resolvido de duas formas. A primeira consiste em encontrar uma matriz $L \in \mathfrak{C}^{m \times n}$, tal que $LA = I_m$. Para a outra solução, deseja-se encontrar $R \in \mathfrak{C}^{m \times n}$, de maneira que $AR = I_n$. Portanto, empregando as duas abordagens, tem-se que

$$i. \quad AX = C \Rightarrow LAX = LC \Rightarrow I_m X = LC \Rightarrow X = LC. \quad (\text{I.16})$$

$$ii. \quad AX = C \Rightarrow I_n AX = I_n C \Rightarrow AX = ARC \Rightarrow X = RC. \quad (\text{I.17})$$

É neste contexto que então surge o estudo das inversas de uma matriz. Considere as seguintes definições sobre matrizes inversas de A :

Definição I.2.4 Seja $A \in \mathfrak{C}^{n \times m}$. Uma matriz $L \in \mathfrak{C}^{m \times n}$ é dita inversa à esquerda de A , se $LA = I_m$.

Definição I.2.5 Seja $A \in \mathfrak{C}^{n \times m}$. Chama-se inversa à direita de A uma matriz $R \in \mathfrak{C}^{m \times n}$, se $AR = I_n$.

Daí, segue que

Lema I.2.6 Se $A \in \mathfrak{C}^{n \times m}$ possui uma matriz inversa à esquerda $L \in \mathfrak{C}^{m \times n}$ e outra matriz inversa à direita $R \in \mathfrak{C}^{m \times n}$, então L e R são iguais e únicas.

Prova Partindo das definições

$$LA = I_m \Rightarrow (LA)R = I_m R \Rightarrow L(AR) = R \Rightarrow LI_n = R \Rightarrow L = R. \quad (\text{I.18})$$

Para provar a unicidade, assuma que exista uma matriz inversa à esquerda dada por $L' \in \mathfrak{C}^{m \times n}$. Portanto, adotando procedimento análogo, encontra-se que $L' = R$. Porém, sabe-se também que $L = R$. Logo, $L' = L = R$. Consequentemente, a inversa à direita R também é única.

Nota 1 - É possível que, para uma certa matriz, não exista um dos tipos de inversa ou mesmo ambos. No primeiro caso, a unicidade não é garantida. ■

Apresentam-se, a seguir, condições necessárias e suficientes para a existência de inversas à esquerda e à direita.

Lema I.2.7 *As seguintes propriedades são válidas para uma matriz $A \in \mathfrak{C}^{n \times m}$*

1. *A matriz A possui inversa à esquerda se, e somente se, $\text{posto}(A) = m$.*
2. *A matriz A possui inversa à direita se, e somente se, $\text{posto}(A) = n$.*
3. *A matriz A possui inversa à esquerda e à direita se, e somente se, $\text{posto}(A) = n = m$.*

Prova Para a prova da propriedade 1, seja $L \in \mathfrak{C}^{m \times n}$ uma matriz inversa à esquerda de A . Desta forma,

$$AX = 0 \Rightarrow LAX = 0 \Rightarrow X = 0. \quad (\text{I.19})$$

Portanto, $\text{posto}(A) = m$. Se $\text{posto}(A) = m$, então a forma escalonada de A é dada por:

$$FA = \begin{bmatrix} I_m \\ 0_{(n-m) \times n} \end{bmatrix}. \quad (\text{I.20})$$

Segue que

$$(A^T F^T F) A = (A^T F^T) (FA) = (FA)^T (FA) = \begin{bmatrix} I_m & 0_{(n-m) \times n} \end{bmatrix} \begin{bmatrix} I_m \\ 0_{(n-m) \times n} \end{bmatrix} = I_m. \quad (\text{I.21})$$

Logo, $L = A^T F^T F$ é uma inversa à esquerda de A . A demonstração de 2 é análoga à de 1. A terceira propriedade é decorrência das outras duas. ■

Corolário I.2.8 *Uma matriz $A \in \mathfrak{C}^{n \times m}$ não possui inversa à esquerda nem à direita, se, e se somente se, $\text{posto}(A) < \min\{n, m\}$.*

Corolário I.2.9 *Uma matriz que tem inversa à direita e à esquerda é quadrada.*

Corolário I.2.10 *Uma matriz quadrada tem inversa à esquerda se, e somente se, tiver inversa à direita.*

Ainda no sentido de se resolver a equação (I.15), pode-se definir um tipo de inversa que generalize os resultados anteriores e que existe para qualquer matriz. Considere a definição de inversa generalizada dada a seguir.

Definição I.2.11 *Considere uma matriz $A \in \mathfrak{C}^{n \times m}$. Uma inversa generalizada de A é uma matriz $A^- \in \mathfrak{C}^{m \times n}$ que satisfaz a relação*

$$AA^-A = A, \forall A. \quad (\text{I.22})$$

Definição I.2.12 *Da definição I.2.11, segue-se que*

$$(I - AA^-)A = 0. \quad (\text{I.23})$$

$$A(I - A^-A) = 0. \quad (\text{I.24})$$

A existência da inversa generalizada pode ser mostrada por meio do cálculo da SVD da matriz A . Seja $A \in \mathfrak{C}^{n \times m}$ e uma SVD dada por

$$A = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^T, \quad (\text{I.25})$$

em que $UU^T = I_n$, $VV^T = I_m$ e $D \in \mathfrak{C}^{r \times r}$ é uma matriz diagonal que contém os valores singulares não-nulos de A . Portanto, uma expressão para a inversa generalizada de A é dada por

$$A^- = V \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^T. \quad (\text{I.26})$$

O próximo lema mostra que a matriz inversa generalizada é a solução de (I.17).

Lema I.2.13 *Se a equação $AX = C$ possui solução, então esta é dada por A^-C .*

Prova Seja X uma solução da equação. Logo,

$$\begin{aligned} AX &= C \\ (AA^-)AX &= (AA^-)C \\ AA^-AX &= A(A^-C) \\ AX &= A(A^-C) \\ C &= A(A^-C). \end{aligned} \quad (\text{I.27})$$

■

O Lema I.2.13 pode ser generalizado, conforme mostrado a seguir.

Lema I.2.14 *Sejam matrizes $A \in \mathfrak{C}^{n \times m}$, $B \in \mathfrak{C}^{p \times q}$ e $C \in \mathfrak{C}^{n \times q}$. Portanto, a equação*

$$AXB = C \quad (\text{I.28})$$

possui uma solução se, e somente se,

$$AA^-CB^-B = C. \quad (\text{I.29})$$

Esta condição também é conhecida como condição de consistência. Se esta condição for satisfeita, o conjunto de soluções do sistema não é vazio e pertencerá ao conjunto

$$S' = \{A^-CB^- + Z - A^-AZBB^- | Z \in \mathfrak{C}^{m \times p}\}, \quad (\text{I.30})$$

em que Z é uma matriz arbitrária.

Prova Seja X uma solução da equação. Logo,

$$\begin{aligned} AXB &= C \\ (AA^-)AXB(B^-B) &= (AA^-)C(B^-B) \\ AA^-AXB(B^-B) &= A(A^-CB^-)B \\ AXB &= A(A^-CB^-)B \\ C &= A(A^-CB^-)B. \end{aligned} \quad (\text{I.31})$$

Por conseguinte, A^-CB^- é uma solução de (I.28). Ainda, se X satisfizer (I.28), então

$$\begin{aligned} C &= AXB \\ C &= AA^-AXBB^-B \\ C &= AA^-ACBB^-B. \end{aligned} \tag{I.32}$$

Seja S o conjunto de soluções de (I.28). Para $X \in S$, escolhe-se $Z = X$. Desta forma,

$$\begin{aligned} &A^-CB^- + Z - A^-AZBB^- \\ &= A^-CB^- + X - A^-AXBB^- \\ &= A^-CB^- + X - A^-CB^- \\ &= X. \end{aligned} \tag{I.33}$$

Portanto, $X \in S'$. Por outro lado, se $X \in S'$, ou seja, $X = A^-CB^- + Z - A^-AZBB^-$, para $Z \in \mathfrak{C}^{m \times p}$, tem-se que

$$\begin{aligned} &AXB \\ &= A(A^-CB^- + Z - A^-AZBB^-)B \\ &= AA^-CB^-B + AZB - AA^-AZBB^-B \\ &= AA^-CB^-B. \end{aligned} \tag{I.34}$$

Como a condição de consistência é satisfeita, o conjunto de soluções não é vazio, ou seja, $S \neq \{0\}$. Tomando uma outra solução $X_2 \in S'$, segue que, partindo de (I.34),

$$\begin{aligned} AXB &= AA^-CB^-B \\ AXB &= AA^-AX_2BB^-B \\ AXB &= AX_2B. \end{aligned} \tag{I.35}$$

Logo, $AX_2B = C$, o que implica que $X \in S$. Desta forma, pode-se concluir que $S' = S$. ■

Considere a definição:

Definição I.2.15 *Se uma matriz quadrada A tem inversa (à esquerda ou à direita), ela é dita invertível. Neste caso, tal inversa, que é única, é denotada por A^{-1} .*

Nota I.2.16 *Dadas matrizes A e B invertíveis, tem-se que*

1. $(AB)^{-1} = B^{-1}A^{-1}$.
2. $(A^{-1})^{-1} = A$.
3. $(A^T)^{-1} = (A^{-1})^T$.

No caso de a matriz $A \in \mathfrak{C}^{n \times m}$ possuir todas as colunas LI, então uma substituição direta mostra que

$$V \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^T = (A^T A)^{-1} A^T. \tag{I.36}$$

Quando as colunas de A são LD, $A^T A$ não é invertível e se utiliza a definição de inversa generalizada. A matriz dada por (I.36) também é conhecida como pseudo-inversa ou inversa de Moore-Penrose e será representada por A^\dagger . Se a matriz pseudo-inversa existe, ela é única e satisfaz as condições

$$i. \quad AA^\dagger A = A. \quad (\text{I.37})$$

$$ii. \quad A^\dagger AA^\dagger = A^\dagger. \quad (\text{I.38})$$

$$iii. \quad (AA^\dagger)^T = AA^\dagger. \quad (\text{I.39})$$

$$iv. \quad (A^\dagger A)^T = A^\dagger A. \quad (\text{I.40})$$

Considere agora uma matriz quadrada $A \in \mathfrak{C}^{n \times n}$. Se A possuir posto pleno, ou seja, $\text{posto}(A) = n$, então existe uma matriz única denominada de inversa, tal que

$$AA^{-1} = A^{-1}A = I_n, \quad (\text{I.41})$$

em que a notação A^{-1} indica a matriz inversa de A . Nota-se também que, quando A for não-singular, a matriz inversa A^{-1} e a pseudo-inversa A^\dagger são iguais.

As possibilidades para o sistema mostrado em (I.15) podem ser definidas a partir das relações entre $\text{posto}(A)$ e $\text{posto}\left(\begin{bmatrix} A & C \end{bmatrix}\right)$, enquadrando-se em um, e somente um, dos seguintes casos:

- $\text{posto}(A) < \text{posto}\left(\begin{bmatrix} A & C \end{bmatrix}\right) \Rightarrow$ solução inconsistente.
- $\text{posto}(A) = \text{posto}\left(\begin{bmatrix} A & C \end{bmatrix}\right) = m \Rightarrow$ solução única.
- $\text{posto}(A) = \text{posto}\left(\begin{bmatrix} A & C \end{bmatrix}\right) < m \Rightarrow$ infinitas soluções.

Uma outra maneira de se relacionar os possíveis casos para o sistema (I.15) é por meio da existência de inversas para A . Deste modo,

- $\text{posto}(A) = n \leq m \Rightarrow$ existe uma inversa à direita $R \in \mathfrak{C}^{m \times n}$.
 - Se $n = m$, então há uma solução única dada por RC .
 - Se $n < m$, então há infinitas soluções; uma delas é dada por RC .
- $\text{posto}(A) = m \leq n \Rightarrow$ existe uma inversa à esquerda $L \in \mathfrak{C}^{m \times n}$.
 - Se $\text{posto}\left(\begin{bmatrix} A & C \end{bmatrix}\right) = m$, então há uma solução única dada por LC .
 - Se $\text{posto}\left(\begin{bmatrix} A & C \end{bmatrix}\right) = m + 1$, então a solução é inconsistente.
- $\text{posto}(A) = r < \min\{n, m\} \Rightarrow$ não existem inversa à esquerda ou à direita, mas apenas a inversa generalizada.
 - Se $\text{posto}\left(\begin{bmatrix} A & C \end{bmatrix}\right) = r$, então há infinitas soluções, uma delas é A^-C .
 - Se $\text{posto}\left(\begin{bmatrix} A & C \end{bmatrix}\right) = r + 1$, então a solução é inconsistente.

Nota I.2.17 Se A for uma matriz quadrada, i.e., $n = m$, e não singular ($r = n = m$), em outras palavras, de posto pleno, então existe uma solução única para X dada por $X = A^{-1}C$.

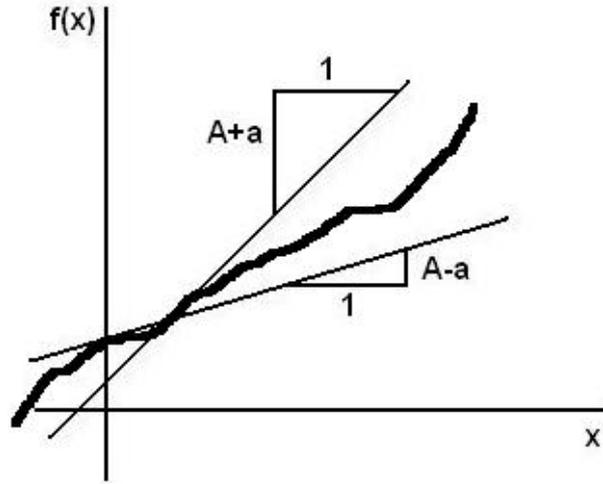


Figura I.1: Exemplo de uma função não-linear limitada em cone.

I.3 FUNÇÕES DE LIPSCHITZ

Seja um sistema discreto não-linear descrito por

$$x_{k+1} = f_k(x_k) + B_{w,k}w_k, \quad (\text{I.42})$$

$$y_k = h_k(x_k) + D_{v,k}v_k, \quad (\text{I.43})$$

em que x_k representa o vetor de estados, y_k denota o vetor de saída e w_k e v_k denotam os ruídos de processo e de medida, respectivamente. Admitem-se funções não-lineares $f_k\{\bullet\}$ e $h_k(\bullet)$. O subscrito em $f_k\{\bullet\}$ é incluído para representar uma possível dependência do tempo. Além do mais, $B_{w,k}$ é independente de x_k .

De acordo com [121], as funções $f_k(x_k)$ e $h_k(x_k)$ são restritas a uma subclasse de funções uniformes segundo Lipschitz em x_k . De maneira geral, pode-se escrever que a matriz de transição de estados depende de um vetor de entradas exógenas, ou seja, $f_k(x_k, u_k)$. Considerando que $f_k(x_k, u_k)$ seja uniforme segundo Lipschitz em x_k e u_k , o autor afirma que ⁴

$$\|f_k(x_k + \delta_k, u_k + \epsilon_k) - f_k(x_k, u_k) - A_k\delta_k - B_{u,k}\epsilon_k\| \leq \Delta_{a,k} \|\delta_k\| + \Delta_{u,k} \|\epsilon_k\|, \quad (\text{I.44})$$

$$\|h_k(x_k + \delta_k) - h_k(x_k) - C_k\delta_k\| \leq \Delta_{c,k} \|\delta_k\|, \quad (\text{I.45})$$

em que A_k , $B_{u,k}$ e C_k parametrizam um modelo linear nominal associado a (I.42)-(I.43):

$$x_{k+1} = A_k x_k + B_{u,k} u_k + B_{w,k} w_k, \quad (\text{I.46})$$

$$y_k = C_k x_k + D_{v,k} v_k, \quad (\text{I.47})$$

e $\Delta_{a,k}$, $\Delta_{u,k}$ e $\Delta_{c,k}$ limitam o afastamento em relação à dinâmica, controle e observações nominais, respectivamente. A Figura I.1 ilustra um exemplo de uma função escalar limitada por um cone.

Nota I.3.1 Para o caso considerado em [64], em que $f_k(x_k, u_k)$ não depende de u_k , tem-se que

$$\|f_k(x_k + \delta_k) - f_k(x_k) - A_k\delta_k\| \leq \Delta_{a,k} \|\delta_k\|. \quad (\text{I.48})$$

⁴Este resultado não é provado em [121] nem nas referências consultadas. Tal resultado será considerado como uma conjectura e será argumentado mais adiante nesta seção.

Nota I.3.2 Funções exponenciais e polinomiais não satisfazem (I.44)-(I.45). Modelos senoidais e inversa da tangente satisfazem. Portanto, muitos modelos podem não ser bem representados por (I.46)-(I.47).

Por fim, as chamadas funções Lipschitz podem ser definidas como

Definição I.3.3 [136] *Um função é contínua segundo Lipschitz se, e somente se, for absolutamente contínua e possuir suas derivadas limitadas.*

Definição I.3.4 [137] *Uma função $f : \mathbb{N} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ é localmente limitada e localmente contínua em x (uniforme em k) segundo Lipschitz, se, para cada subconjunto compacto $D \subset \mathbb{R}^n$, existir constantes positivas F e λ , possivelmente dependentes de D , mas independentes de k , que satisfaçam*

1. f é local e uniformemente limitada, ou seja,

$$F > 0, \forall x \in D, \forall k : \|f_k(x)\| \leq F. \quad (\text{I.49})$$

2. f é local e uniformemente contínua segundo Lipschitz, i.e.,

$$\lambda : \forall x, y \in D, \forall k : \|f_k(x) - f_k(y)\| \leq \lambda \|x - y\|. \quad (\text{I.50})$$

Para o caso unidimensional, pode-se adotar a seguinte definição:

Definição I.3.5 [138] *Uma função $f : [a, b] \rightarrow \mathbb{R}$ é do tipo Lipschitz em $[a, b]$, se*

$$|f(x_1) - f(x_2)| \leq k|x_1 - x_2|, \forall x_1, x_2 \in [a, b], \quad (\text{I.51})$$

em que $k \geq 0$ é uma constante conhecida como constante de Lipschitz.

Conjectura I.3.6 *Considere uma função $f : \mathbb{R} \rightarrow \mathbb{R}$ diferenciável. Se, para algum ponto $x \in \mathbb{R}$, a derivada de f estiver no intervalo $[A - \Delta, A + \Delta]$ para algum $A \in \mathbb{R}$ e $\Delta \in \mathbb{R}^+$, então existirá uma região $D \subset \mathbb{R}$ em que é válido*

$$|f(x + \delta) - f(x) - A\delta| \leq \Delta|\delta|, \quad (\text{I.52})$$

para todo $\delta \in D$.

Prova Partindo da definição, tem-se que a aproximação da derivada de uma função é dada por

$$\frac{f(x + \delta) - f(x)}{\delta}, \quad (\text{I.53})$$

em que δ é suficientemente pequeno. Para um δ escolhido de maneira satisfatória, a derivada de f é representada por A , tal que, com uma boa aproximação,

$$\lim_{\delta \rightarrow 0} \frac{f(x + \delta) - f(x)}{\delta} = A. \quad (\text{I.54})$$

Suponha agora que o valor verdadeiro da derivada esteja dentro de um intervalo conhecido em torno de seu valor nominal A , ou seja, $[A - \Delta, A + \Delta]$. De outra forma,

$$A - \Delta \leq \frac{f(x + \delta) - f(x)}{\delta} \leq A + \Delta \Rightarrow -\Delta \leq \frac{f(x + \delta) - f(x)}{\delta} - A \leq \Delta. \quad (\text{I.55})$$

Supondo $\delta > 0$, segue que

$$-\Delta\delta \leq f(x + \delta) - f(x) - A\delta \leq \Delta\delta. \quad (\text{I.56})$$

Como $\Delta > 0$, pode-se concluir que

$$|f(x + \delta) - f(x) - A\delta| \leq \Delta |\delta|. \quad (\text{I.57})$$

Supondo $\delta < 0$,

$$\Delta\delta \geq f(x - \delta) - f(x) + A\delta \geq -\Delta\delta, \quad (\text{I.58})$$

Como $\Delta > 0$, tem-se que

$$|f(x - \delta) - f(x) + A\delta| \leq \Delta |\delta|. \quad (\text{I.59})$$

Assim, (I.52) é válida para todo $\delta \in D$.

Este desenvolvimento é válido então para funções $f : \mathbb{R} \rightarrow \mathbb{R}$. Neste sentido, pode-se fazer a conjectura de que esta afirmativa também seja válida para \mathbb{R}^n . ■

I.4 ELIPSES DE CONFIANÇA

Considere o caso de uma distribuição gaussiana bivariada caracterizada por

$$x := \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad E\{x\} := \begin{bmatrix} E\{x_1\} \\ E\{x_2\} \end{bmatrix}, \quad P_x := \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}, \quad (\text{I.60})$$

em que $\sigma_{12} = \rho\sigma_1\sigma_2$, e o parâmetro ρ denota o coeficiente de correlação. Como P_x é simétrica, ela pode ser decomposta em $P_x = e_1v_1v_1^T + e_2v_2v_2^T$ sendo e_1, e_2 os autovalores de P_x e v_1, v_2 seus autovetores correspondentes. A função de densidade de probabilidade conjunta de x é dada por

$$P_{x_1x_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} e^{-\frac{1}{2}q}, \quad (\text{I.61})$$

em que

$$\begin{aligned} q &= (x - E\{x\})^T P_x^{-1} (x - E\{x\}), \\ &= \frac{1}{(1-\rho^2)} \left\{ \left(\frac{x_1 - E\{x_1\}}{\sigma_1} \right)^2 - 2\rho \left(\frac{x_1 - E\{x_1\}}{\sigma_1} \right) \left(\frac{x_2 - E\{x_2\}}{\sigma_2} \right) + \left(\frac{x_2 - E\{x_2\}}{\sigma_2} \right)^2 \right\}. \end{aligned} \quad (\text{I.62})$$

Utilizando a notação

$$u_1 := \frac{x_1 - E\{x_1\}}{\sigma_1}, \quad (\text{I.63})$$

$$u_2 := \frac{x_2 - E\{x_2\}}{\sigma_2}, \quad (\text{I.64})$$

e definindo $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ e $P_u = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, conclui-se que $q = u^T P_u^{-1} u$. De outra forma,

$$q = \frac{1}{(1 - \rho^2)} \{u_1^2 - 2\rho u_1 u_2 + u_2^2\}. \quad (\text{I.65})$$

Deseja-se encontrar uma maneira gráfica de demonstrar a distribuição de x . Uma delas é por meio de curvas de densidade constante $P_x(x) = c$. Neste caso, para se obter $P_x(x)$ constante, deve-se impor $q = c$, em que c é uma constante. Desta forma,

$$q = \frac{1}{(1 - \rho^2)} \{u_1^2 - 2\rho u_1 u_2 + u_2^2\} = c \quad (\text{I.66})$$

$$\Rightarrow u_1^2 - 2\rho u_1 u_2 + u_2^2 = (1 - \rho^2) c. \quad (\text{I.67})$$

Retornando com as variáveis originais do problema, obtém-se a equação de uma elipse com os seguintes parâmetros

- Centro: $E\{x\}$.
- Ângulo em relação a x_1 : $\theta = \frac{1}{2} \arctan\left(\frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}\right)$.
- Eixos principais em que r_1 é o eixo menor e r_2 , o maior. Ademais, r_1 é representado na direção v_1

$$\begin{aligned} r_1^2 &= c \frac{\sigma_1^2 \sigma_2^2 (1 - \rho^2)}{\sigma_2^2 \cos^2(\theta) - 2\rho\sigma_1\sigma_2 \cos(\theta) \sin(\theta) + \sigma_1^2 \sin^2(\theta)} \\ &= c \frac{\sigma_1^2 \sigma_2^2 (1 - \rho^2)}{(\sigma_2 \cos(\theta) - \sigma_1 \sin(\theta))^2}, \end{aligned} \quad (\text{I.68})$$

$$\begin{aligned} r_2^2 &= c \frac{\sigma_1^2 \sigma_2^2 (1 - \rho^2)}{\sigma_2^2 \cos^2(\theta) + 2\rho\sigma_1\sigma_2 \cos(\theta) \sin(\theta) + \sigma_1^2 \sin^2(\theta)} \\ &= c \frac{\sigma_1^2 \sigma_2^2 (1 - \rho^2)}{(\sigma_2 \cos(\theta) + \sigma_1 \sin(\theta))^2}. \end{aligned} \quad (\text{I.69})$$

I.5 CÁLCULO MATRICIAL

Algumas expressões úteis para derivadas de primeira ordem

$$\frac{\partial}{\partial X} A^T X B = A B^T, \quad (\text{I.70})$$

$$\frac{\partial}{\partial X} A^T X^T B = B^T A, \quad (\text{I.71})$$

$$\frac{\partial}{\partial X} A^T X A = \frac{\partial}{\partial X} A^T X^T A = A A^T, \quad (\text{I.72})$$

concluindo-se que

$$\frac{\partial}{\partial X} X B = B^T, \quad (\text{I.73})$$

$$\frac{\partial}{\partial X} A^T X^T = A^T, \quad (\text{I.74})$$

$$\frac{\partial}{\partial X} X C^T = \frac{\partial}{\partial X} C X^T = C. \quad (\text{I.75})$$

Para derivadas de segunda ordem,

$$\frac{\partial}{\partial X} X^T B X = (B + B^T)X. \quad (\text{I.76})$$

Se B for simétrica,

$$\frac{\partial}{\partial X} X^T B X = 2BX, \quad (\text{I.77})$$

$$\frac{\partial}{\partial X} X B X^T = 2XB. \quad (\text{I.78})$$

I.6 RESULTADOS AUXILIARES

Lema I.6.1 *Sejam matrizes A e B com mesmas dimensões, tais que*

$$A \geq B \geq 0, \quad (\text{I.79})$$

então, para um vetor qualquer com dimensões apropriadas, é correto afirmar que

$$x^T A x \geq x^T B x \geq 0, \forall x. \quad (\text{I.80})$$

Prova Se (I.79) é satisfeita,

$$A \geq B \geq 0 \Rightarrow A - B \geq 0. \quad (\text{I.81})$$

Assim, $(A - B)$ também é semidefinida positiva e, portanto, valem as relações

$$x^T (A - B) x \geq 0 \Rightarrow x^T A x - x^T B x \geq 0 \Rightarrow x^T A x \geq x^T B x. \quad (\text{I.82})$$

■

Lema I.6.2 *Sejam matrizes A e B com mesmas dimensões, tais que*

$$A \leq B. \quad (\text{I.83})$$

Portanto, para um vetor qualquer com dimensões apropriadas, é correto afirmar que

$$\text{tr} \{A\} \leq \text{tr} \{B\}. \quad (\text{I.84})$$

Prova Se (I.83) é satisfeita,

$$A \leq B \Rightarrow A - B \leq 0. \quad (\text{I.85})$$

Assim, $(A - B)$ também é negativa definida e, portanto, valem as relações

$$\text{tr} (A - B) \leq 0 \Rightarrow \text{tr} \{A\} - \text{tr} \{B\} \leq 0 \Rightarrow \text{tr} \{A\} \leq \text{tr} \{B\}. \quad (\text{I.86})$$

■

Lema I.6.3 Considerando a definição

$$P := E \left\{ (e - E \{e\}) (e - E \{e\})^T \right\}, \quad (\text{I.87})$$

é correto afirmar que

$$E \left\{ \|e - E \{e\}\|^2 \right\} = \text{tr} (P). \quad (\text{I.88})$$

Prova Inicialmente, tem-se que

$$E \left\{ \|e - E \{e\}\|^2 \right\} = E \left\{ (e - E \{e\})^T (e - E \{e\}) \right\}. \quad (\text{I.89})$$

Como se trata de um produto interno, o valor acima é um escalar e, portanto, igual ao valor do seu traço. Logo,

$$\begin{aligned} E \left\{ (e - E \{e\})^T (e - E \{e\}) \right\} &= \text{tr} \left(E \left\{ (e - E \{e\})^T (e - E \{e\}) \right\} \right) \\ &= E \left\{ \text{tr} \left((e - E \{e\})^T (e - E \{e\}) \right) \right\}. \end{aligned} \quad (\text{I.90})$$

Utilizando a propriedade do traço, $\text{tr} (AB) = \text{tr} (BA)$, conclui-se que

$$E \left\{ \text{tr} \left((e - E \{e\})^T (e - E \{e\}) \right) \right\} = E \left\{ \text{tr} \left((e - E \{e\}) (e - E \{e\})^T \right) \right\}. \quad (\text{I.91})$$

Como o operador de traço é linear,

$$E \left\{ \text{tr} \left((e - E \{e\}) (e - E \{e\})^T \right) \right\} = \text{tr} \left(E \left\{ (e - E \{e\}) (e - E \{e\})^T \right\} \right) = \text{tr} (P). \quad (\text{I.92})$$

■

II. DESCRIÇÃO DO CONTEÚDO DO CD

O material contido no CD, apenso, tem por objetivo fornecer uma compreensão mais bem detalhada desta dissertação.

No CD, encontram-se os seguintes diretórios:

- Artigos publicados a partir desta dissertação.
- Dissertação: contém os arquivos necessários para gerar este documento.