



UNIVERSIDADE FEDERAL DO CEARÁ

PPGETI

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA
DOUTORADO EM ENGENHARIA DE TELEINFORMÁTICA

TIAGO DOS SANTOS FAÇANHA

**FILTRAGEM DE KALMAN DISTRIBUÍDA VIA PREDIÇÃO POR INTERAÇÃO:
CONTRIBUIÇÕES TEÓRICAS E APLICAÇÕES**

FORTALEZA

2023

TIAGO DOS SANTOS FAÇANHA

FILTRAGEM DE KALMAN DISTRIBUÍDA VIA PREDIÇÃO POR INTERAÇÃO:
CONTRIBUIÇÕES TEÓRICAS E APLICAÇÕES

Tese apresentada ao Curso de Doutorado em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do PPGETI da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Engenharia em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Orientador: Prof. Dr. Guilherme de Alencar Barreto

Co-Orientador: Prof. Dr. José Tarcísio Costa Filho

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

F123 Façanha, Tiago dos Santos.

Filtragem de Kalman Distribuída via Predição por Interação: Contribuições teóricas e Aplicações / Tiago dos Santos Façanha. – 2023.

105 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2023.

Orientação: Prof. Dr. Guilherme de Alencar Barreto.

Coorientação: Prof. Dr. José Tarcísio Costa Filho.

1. Filtragem distribuída. 2. Paralelização. 3. Predição por interação. 4. Equações de Riccati. I. Título.

CDD 621.38

TIAGO DOS SANTOS FAÇANHA

FILTRAGEM DE KALMAN DISTRIBUÍDA VIA PREDIÇÃO POR INTERAÇÃO:
CONTRIBUIÇÕES TEÓRICAS E APLICAÇÕES

Tese apresentada ao Curso de Doutorado
em Engenharia de Teleinformática do Pro-
grama de Pós-Graduação em Engenharia de
Teleinformática do PPGETI da Universidade
Federal do Ceará, como requisito parcial à
obtenção do título de doutor em Engenharia
em Engenharia de Teleinformática. Área de
Concentração: Sinais e Sistemas

Aprovada em: 27 de Janeiro de 2023

BANCA EXAMINADORA

Prof. Dr. Guilherme de Alencar Barreto (Orientador)
Universidade Federal do Ceará (PPGETI/UFC)

Prof. Dr. José Tarcísio Costa Filho (Co-Orientador)
Universidade Federal do Ceará (DETI/UFC)

Prof. PhD. Amit Bhaya
Universidade Federal do Rio de Janeiro (PEE/COPPE/UFRJ)

Prof. Dr. Adriano Almeida Gonçalves Siqueira
Universidade de São Paulo (PPGAEM/EESC-USP)

Prof. Dr. Bismark Claure Torrico
Universidade Federal do Ceará (PPGEE/UFC)

Prof. Dr. Charles Casimiro Cavalcante
Universidade Federal do Ceará (PPGETI/UFC)

À minha família, por sua capacidade de acreditar em mim e investir em mim. Mãe, seu cuidado e dedicação foi que deram, em alguns momentos, a esperança para seguir. Pai, sua presença significou segurança e certeza de que não estou sozinho nessa caminhada.

AGRADECIMENTOS

Agradeço,

A Deus por ter me abençoado e me dado forças para superar os momentos mais difíceis.

Aos meus pais, José Valter Façanha e Maria Celeste dos Santos Façanha, que me deram todas as condições necessárias para que pudesse chegar até aqui, pois nada seria possível sem o amor e a dedicação deles e de todos os outros familiares.

Aos Professores Dr. José Tarcísio Costa Filho e Dr. Guilherme Alencar Barreto pelo interesse, confiança, orientação, incentivo e paciência.

Aos outros professores e colegas de laboratório, LIT/IFCE e CENTAURO/UFC, pelo incentivo, companheirismo, amizade e aos momentos de descontração compartilhados durante este período de desenvolvimento pessoal, momentos especiais que certamente guardarei e lembrei com bastante orgulho.

Agradeço a todos os professores por me proporcionar o conhecimento não apenas racional, mas a manifestação do caráter e afetividade da educação no processo de formação profissional, por tanto que se dedicaram a mim, não somente por terem me ensinado, mas por terem me feito aprender.

A todos os funcionários do Programa de Pós-Graduação em Engenharia de Teleinformática que de forma direta ou indireta participaram do desenvolvimento deste trabalho

“O sonho é que leva a gente para frente. Se a gente for seguir a razão, fica aquietado, acomodado.”

(Ariano Suassuna)

RESUMO

O problema da filtragem de Kalman distribuída envolve o desafio técnico de migrar a bem estabelecida abordagem da filtragem de Kalman global aplicada a sistemas dinâmicos de larga escala com medições distribuídas para subsistemas mais simples, sem comprometer significativamente o seu desempenho. Nesta adaptação encontram-se problemas de convergência assintótica, definição dos critérios de parada e necessidade de um centro de fusão que não permitem que as aplicações sejam totalmente distribuídas. Assim, nesta tese são propostos dois algoritmos para a solução dos problemas de filtragem distribuída em tempo discreto. Estas abordagens utilizam uma metodologia de paralelização hierárquica que consiste em uma modificação realizada na estrutura do cálculo para diminuir a relação de dependência das equações do filtro e diminuir o tempo de convergência. Para isto, o algoritmo 1 explora a dualidade entre o filtro e o controle e aplica os conceitos de decomposição paramétrica via predição por interação. O algoritmo 2, também explora a dualidade já citada e aplica os princípios de otimização convexa para obter uma forma quadrática estática com dimensões reduzidas, equivalente ao lagrangiano associado a cada subsistema, através de procedimentos de regularização, via regiões de confiança, sem recorrer as equações de Riccati. Desse modo, é possível através do ajuste de um parâmetro escalar μ obter uma família de soluções eficientes e bem condicionadas para cada subsistema que pode amenizar os impactos dos ruídos de processo e medida no sistema linear dinâmico. Para ressaltar as vantagens dos dois algoritmos propostos, cinco exemplos do tipo provas de conceito e outros dois exemplos utilizados em estudos de benchmarking são apresentados e discutidos. Nestes exemplos é possível analisar o tempo de execução, condicionamento do sistema, obter uma família de ganhos para cada subsistema, verificar geometricamente a robustez do sistema de acordo com o conceito de contração da matriz e comparar através da média do erro quadrático, os algoritmos propostos com o estado da arte em algoritmos distribuídos baseados em consenso.

Palavras-chave: filtragem distribuída; paralelização; predição por interação; equações de Riccati.

ABSTRACT

The problem of distributed Kalman filtering involves the technical challenge of migrating the well-established global Kalman filtering approach applied to complex large-scale dynamic systems with distributed measurements to simpler subsystems, without significantly compromising their performance. In this adaptation there are problems of asymptotic convergence, definition of the stop criteria and the need for a fusion center that do not allow applications to be fully distributed. Thus, in this thesis two algorithms are proposed for the solution of distributed filtering problems in discrete time. These approaches use a hierarchical parallelization methodology that consists of a modification made in the structure of the calculation to decrease the dependency ratio of filter equations and decrease the convergence time. For this, algorithm 1 explores the duality between the filter and the control and applies the concepts of parametric decomposition via interaction prediction. Algorithm 2 also explores the aforementioned duality and applies the principles of convex optimization to obtain a static quadratic form with reduced dimensions, equivalent to the lagrangian associated with each subsystem, through regularization procedures, via regions of trust, without resorting to the Riccati equations. Thus, it is possible by adjusting a scale parameter μ to obtain a family of efficient and well-conditioned solutions for each subsystem that can mitigate the impacts of process and measurement noises in the dynamic linear system. To highlight the advantages of the two proposed algorithms, five proof-of-concept examples and two other examples used in benchmarking studies are presented and discussed. In these examples, it is possible to analyze the execution time, conditioning of the system, obtain a family of gains for each subsystem, geometrically verify the robustness of the system according to the concept of matrix contraction and compare, through the mean squared error, the algorithms proposed with the state of the art in consensus-based distributed algorithms.

Keywords: distributed filtering; parallelization; prediction by interaction Riccati equations.

LISTA DE FIGURAS

Figura 1 – Estrutura Geral do Filtro de Kalman Distribuído	19
Figura 2 – Diagrama de predição e atualização do filtro de Kalman	24
Figura 3 – Estrutura em dois níveis.	43
Figura 4 – Estrutura hierárquica para o método de coordenação por predição de interação.	46
Figura 5 – Estrutura em dois níveis do método predição por interação.	55
Figura 6 – Estrutura hierárquica em dois níveis.	61
Figura 7 – Procedimento de paralelização e implementação.	63
Figura 8 – Etapas de Partição : [a] Com coordenação e [b] Sem Coordenação	65
Figura 9 – Resposta do FKG (vermelho) e do método proposto FKDPI (preto) para a saída y_k^1 (azul).	66
Figura 10 – Resposta do FKG (vermelho) e do método proposto FKDPI (preto) para a saída y_k^2 (azul).	67
Figura 11 – Análise comparativa do erro quadrático entre FKD e FKDPI μ ($\mu = 0,5$) . .	82
Figura 12 – Efeito do parâmetro μ na contração da matriz H	83
Figura 13 – Saída y_1 : FKG e FKDPI μ ($\mu_1 = 1$ e $\mu_2 = -1$)	87
Figura 14 – Saída y_2 : FKG e FKDPI μ ($\mu_1 = 1$ e $\mu_2 = -1$)	88
Figura 15 – Visualização de $\ G\ $	106
Figura 16 – Visualização da Matriz Pequena	107

LISTA DE TABELAS

Tabela 1 – Equações do filtro de Kalman discreto	26
Tabela 2 – Equações do filtro de Kalman robusto (H_∞)	30
Tabela 3 – Filtro de Kalman (Informação) x LQR	36
Tabela 4 – Dualidade: Filtro de Kalman x LQR	37
Tabela 5 – Dualidade das restrições: filtro de Kalman e filtro de informação	52
Tabela 6 – Análise de desempenho algébrico e computacional	67
Tabela 7 – Análise de desempenho algébrico e computacional	68
Tabela 8 – Parâmetros iniciais do Exemplo 3.	69
Tabela 9 – Análise do erro quadrático médio (MSE) dos algoritmos avaliados.	70
Tabela 10 – Formulação do Filtro Independente das Equações de Riccati	76
Tabela 11 – Análise do desempenho do algoritmo proposto FDKPI μ para diferentes valores do parâmetro μ (Problema 1).	81
Tabela 12 – Análise do desempenho do algoritmo proposto FDKPI μ para diferentes valores do parâmetro μ (Problema 2).	85
Tabela 13 – Análise do erro quadrático médio (MSE) dos algoritmos	86
Tabela 14 – Impacto da variação dos parâmetros μ_1 e μ_2	88

LISTA DE ALGORITMOS

Algoritmo 1 – Filtro de Kalman Discreto	26
Algoritmo 2 – Nível 2	56
Algoritmo 3 – Nível 1	56
Algoritmo 4 – Subsistema 1	57
Algoritmo 5 – Subsistema 2	58
Algoritmo 6 – Coordenador	58
Algoritmo 7 – Subsistema 1	79
Algoritmo 8 – Subsistema 2	80

LISTA DE SÍMBOLOS

FKG	Filtro de Kalman Global
FKD	Filtro de Kalman Distribuído
$DBKF$	Filtro de Kalman- Bucy Distribuído
$IKCF$	Filtro de Kalman Consenso Ponderado em Informação
$MKCF$	Filtro de Kalman Consenso multiagente
$ADKF$	Filtro de Kalman Distribuído Aproximado
$SDKF$	Filtro de Kalman Distribuído Subótimo
$FKDPI$	Filtro de Kalman Distribuído via Predição por interação
$FKDPI\mu$	Filtro de Kalman Distribuído via Predição por Interação com parâmetro μ
LQR	Regulador linear quadrático
γ	Magnitude da Penalidade das perturbações (fator de ponderação)
H_2	Norma quadrática euclidiana
H_∞	Norma de uma matriz de transferência
I_n	Matriz identidade nxn
k	Variável de tempo discreto
μ	Parâmetro de condicionamento
$TPBVP$	Two point boundary values problem
A^T	Transposta de uma matriz A
$(C^T)^+$	Pseudoinversa da matriz C
KKT	Condições de Karush - Kun - Tucker
TRS	Subproblema via região de confiança
Q_w	Matriz de correlação do processo
R	Matriz de correlação de medida
HJB	Hamilton - Jacobi - Bellman
L	Lagrangiana
H	Hamiltoniano

S	Sistema dinâmico
N	Número de subsistemas
J	Função custo
β_k^i	Multiplicador de Lagrange
i	Dimensão do sistema
SMW	Sherman - Morrison - Woodbury

SUMÁRIO

1	INTRODUÇÃO	18
1.1	Objetivos da Tese	20
1.1.1	<i>Objetivo Geral</i>	20
1.1.2	<i>Objetivos Específicos</i>	20
1.2	Produções Relacionadas	21
1.3	Organização Geral desta Tese	21
2	FUNDAMENTOS	23
2.1	Filtro de Kalman Global	23
2.1.1	<i>Formulação do Problema e Resumo do Algoritmo do Filtro de Kalman</i>	24
2.2	Filtro de Kalman Robusto (H_∞)	26
2.2.1	<i>Formulação do Problema</i>	27
2.3	Algoritmos de Filtragem Distribuída baseados em Consenso	30
2.3.1	<i>Filtro de Kalman Distribuído</i>	31
2.3.2	<i>Filtro de Kalman Consenso Multiagente</i>	31
2.3.3	<i>Filtro de Kalman-Bucy Distribuído</i>	32
2.3.4	<i>Filtro de Kalman Consenso Ponderado em Informação</i>	33
2.3.5	<i>Filtro de Kalman Distribuído Aproximado</i>	34
2.4	Dualidade entre Filtro de Kalman e o Regulador Linear Quadrático	35
2.4.1	<i>Relação entre as Equações de Riccati discreta e contínua</i>	36
2.5	Formulação Primal-Dual	38
2.6	Forma Dual - Decomposição Paramétrica	40
2.7	Método de Decomposição numa estrutura de dois níveis	42
2.8	Método de Decomposição por Coordenação via Predição de Interação	44
2.8.1	<i>Solução dos Subproblemas de Otimização Independentes</i>	45
2.8.2	<i>Solução do Problema de Coordenação</i>	45
2.9	Comentário Final	47
3	NOVA METODOLOGIA PARA FILTRO DE KALMAN DISTRIBUÍDO	48
3.1	Decomposição do Problema	48
3.2	Formulação do Problema de Otimização Dinâmica	49
3.3	Particularidade da Formulação do Problema de Filtragem	50

3.4	Filtro de Kalman Distribuído: Método Predição por Interação	51
3.5	Descrição do Algoritmo Proposto	57
3.6	Estratégias de Programação Distribuída	59
3.6.1	<i>Problema de Partição</i>	59
3.6.2	<i>Problema de Distribuição</i>	60
3.6.3	<i>Problema de Comunicação e Sincronismo</i>	60
3.7	Metodologias de Otimização e Programação Distribuída	61
3.8	Paralelização do Algoritmo	62
3.8.1	<i>Primeira Etapa</i>	63
3.8.2	<i>Segunda Etapa</i>	64
3.9	Provas de Conceito	65
3.9.1	<i>Exemplo ilustrativo 1</i>	66
3.9.2	<i>Exemplo ilustrativo 2</i>	67
3.9.3	<i>Exemplo ilustrativo 3</i>	69
3.10	Comentário Final	70
4	FILTRO DE KALMAN DISTRIBUÍDO COM HORIZONTE DE TEMPO REDUZIDO	71
4.1	Formulação Dual via Programação Quadrática	71
4.2	Descrição do Algoritmo Proposto	78
4.3	Provas de Conceito	80
4.3.1	<i>Exemplo ilustrativo 1</i>	81
4.3.2	<i>Exemplo ilustrativo 2</i>	83
4.3.3	<i>Exemplo ilustrativo 3</i>	84
4.3.4	<i>Exemplo ilustrativo 4</i>	86
4.4	Comentário Final	89
5	CONCLUSÕES E TRABALHOS FUTUROS	90
5.1	Contribuições	90
5.2	Trabalhos Futuros	91
	REFERÊNCIAS	92
	APÊNDICES	101
	APÊNDICE A – Minimização da Forma Quadrática Estática	101
A.1	Formulação do Problema	101

A.2	Método das Regiões de Confiança	102
A.3	Procedimentos de Regularização	102
	APÊNDICE B – Contração de Matrizes	106

1 INTRODUÇÃO

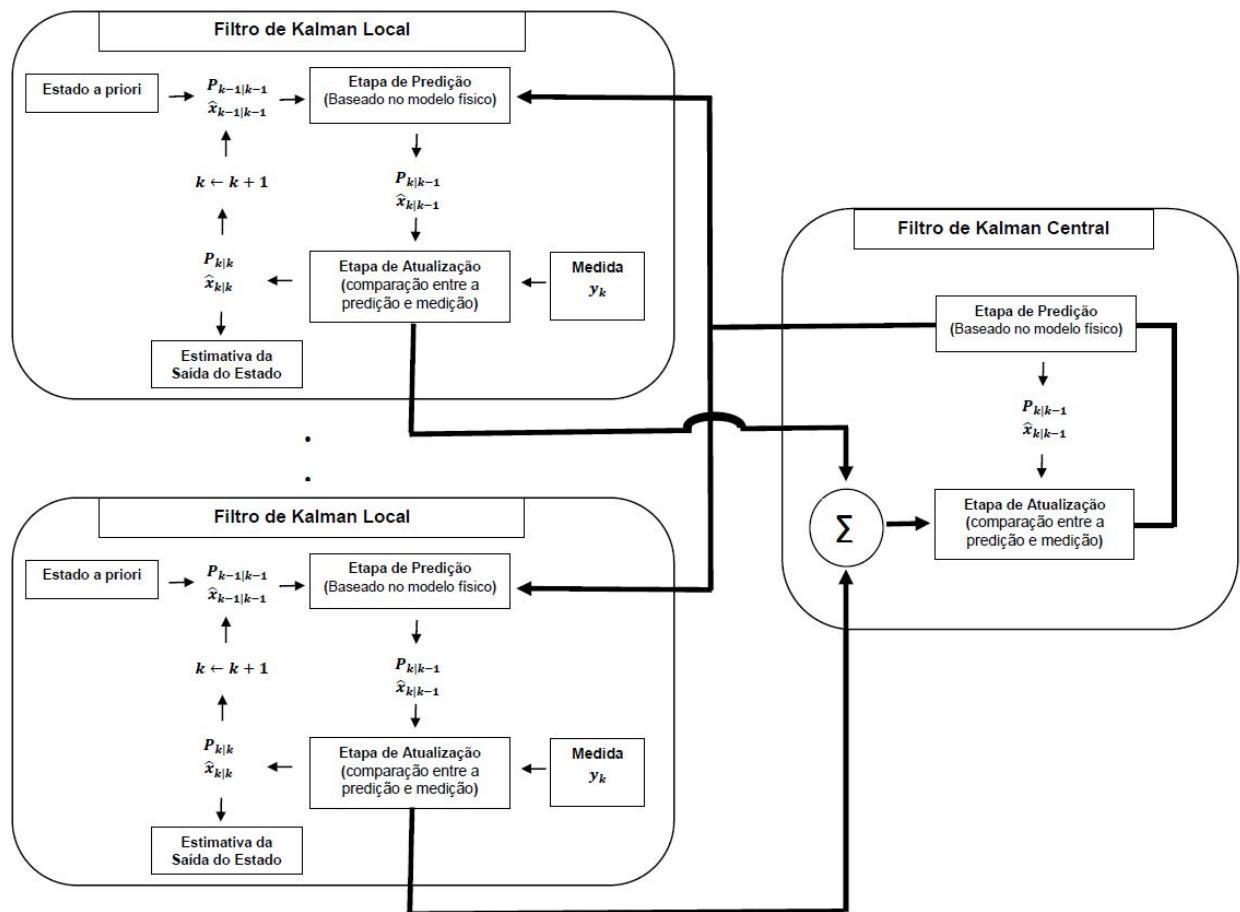
O estudo da solução do problema de filtragem (LEVINE, 2010) é de grande importância no campo da engenharia, como pode ser visto em uma gama de artigos encontrados na literatura com aplicações em robótica (DU *et al.*, 2016; ZHANG *et al.*, 2015), rastreamento de trajetórias e alvos (HE *et al.*, Tokyo, 2013; HU *et al.*, 2015), telecomunicações (MYRONCHUK *et al.*, Lviv-Slavsk, Ukraine, 2020; BELGACEM; LAMIR, 2018), odometria (ZHANG *et al.*, 2017), plantas industriais (CUMBO *et al.*, 2019) e sistemas de potência (SALINAS; LI, 2016; ZANNI *et al.*, 2017).

Desde meados do século vinte até então, vimos o desenvolvimento e o constante aprimoramento de técnicas para tratar os mais diversos casos, desde os problemas dinâmicos lineares, problemas de filtragem robusta com a incerteza da característica estocástica do ruído ou modelagem imprecisa do sistema, até aplicações de sistemas multiagentes que geralmente são modelados como um problema de estimação distribuída. Assim, o problema de filtragem Kalman distribuída (FKD) tem sido um tópico de pesquisa relevante por mais de uma década, em que o desafio técnico dessa abordagem envolve como migrar a bem estabelecida abordagem de filtragem de Kalman global (FKG) (ANDERSON; MOORE, 1979), (KALMAN, 1960) para complexos sistemas dinâmicos de larga escala com medições distribuídas em uma grande área geográfica (KHAN *et al.*, 2008), amplamente aplicados em redes de sensores (LI *et al.*, 2018; CHENG *et al.*, 2018), sensoriamento ambiental (WEI *et al.*, 2011), smartgrid (MOHAMMADI; PLATANIOTIS, 2018), dentre outras aplicações.

A arquitetura de um sistema distribuído, em geral, é muito poderosa, pois permite dividir o sistema dinâmico de grande escala em subsistemas mais simples, sem comprometer desempenho. Uma estrutura geral para o FKD pode ser observado na Figura 1. Algoritmos FKD disponíveis já são abundantes. Por exemplo, um algoritmo de previsão de um passo foi introduzido em Zhou (2013), um algoritmo de inversão de colapso iterado distribuído em conjunto com um gráfico de fusão de dados foi introduzido em Khan *et al.* (2008) para sistemas espacialmente esparsos. A estimativa de fusão distribuída foi proposta em Chen *et al.* (2014a), Chen *et al.* (2014b) e Chen *et al.* (2015). FKD usando informação quantizada foi estudada em Li *et al.* (2015), Ribeiro *et al.* (2006) e Song *et al.* (2014b). Um projeto FKD usando o protocolo gossip bem conhecido por propagar a informação local de forma eficiente e escalável, foi dado em Li *et al.* (2015). O FKD usando estratégias de difusão foi estudado em Cattivelli e Sayed (2010) e Hu *et al.* (2012). O FKD com medições por amostragem, fora de sequência foi tratado

em Shen *et al.* (2009), e há também os filtros FKD baseados no centro de fusão, com a utilização de um nó coordenador, foram mostrados em Song *et al.* (2014a) e Xu *et al.* (2012). Também pode-se encontrar diversas aplicações de FKD na revisão bibliográfica de Mahmoud e Khalid (2013). Trabalhos relacionados também incluem estimativas distribuída, baseadas no método Gauss-Newton (ZHAO; NEHORAI, 2007) e estimativas ponderadas por mínimos quadrados distribuídas (MARELLI; FU, 2015).

Figura 1 – Estrutura Geral do Filtro de Kalman Distribuído



Fonte: Mahmoud e Khalid (2013).

Uma abordagem muito comum para o FKD é usar uma estratégia de consenso (XIE *et al.*, 2018; OLFATI-SABER, New Orleans, LA, USA, textbf2007) em que é possível observar diversas aplicações para essa abordagem em Carli *et al.* (2008), Das e Moura (2015), Kar e Moura (2011), Song *et al.* (2014a), Xu *et al.* (2012) e Zhou (2013). As principais deficiências desta abordagem incluem as seguintes: (1) geralmente apenas a convergência assintótica é garantida, o que significa que um número infinito de iterações é requerido em teoria; (2) em aplicações práticas é difícil fornecer critérios de parada; (3) apenas estimativas subótimas são geralmente

dadas. Outro problema comum com muitos algoritmos FKD existentes é a necessidade de um centro de fusão, o que significa que eles não são totalmente distribuídos. Abordagens desse tipo incluem as introduzidas por Chen *et al.* (2014a), Chen *et al.* (2015), Chen *et al.* (2014b), Shen *et al.* (2009), Song *et al.* (2014b) e Xu *et al.* (2012). Também nota-se que o centro de fusão é comumente usado na chamada filtragem de Kalman paralela, em que medições locais são usadas para produzir estimativas locais que são então fundidas em um centro de fusão; ver, por exemplo, (HASHEMPOUR *et al.*, 1988). A necessidade de FKD totalmente distribuída com boas propriedades de otimização e baixas complexidades computacionais, comunicacionais e de armazenamento é urgente. Estas propriedades são essenciais para tornar o FKD escalável, principalmente para aplicações com sensores com características diferentes.

1.1 Objetivos da Tese

Esta tese trata especificamente de problemas de filtragem distribuída considerando sistemas lineares em tempo discreto representados por modelos em espaço de estados com os seguintes objetivos gerais e específicos.

1.1.1 *Objetivo Geral*

O objetivo geral desta tese é propor uma nova metodologia de paralelização que consiste numa modificação realizada na estrutura do cálculo para diminuir a relação de dependência das equações do filtro de Kalman global utilizando decomposição paramétrica via predição por interação.

1.1.2 *Objetivos Específicos*

Dentre os objetivos específicos desta tese, pode-se listar os seguintes aspectos:

- Formular uma metodologia de paralelização que permite uma modificação na estrutura do cálculo do filtro.
- Utilizar formas quadráticas com dimensões reduzidas, que possibilitam bom condicionamento e projetos alternativos a filtragem distribuída em aplicações de tempo real.
- Definir parâmetros livres (matrizes de ponderação) que podem ser fixados, sem a necessidade de um ajuste exaustivo. Isto permitirá que um parâmetro escalar seja utilizado como instrumento para adequar o sistema de filtragem distribuída às condições de desempenho e

estabilidade.

- Compor uma família de soluções de ganho para atender a necessidade do projeto de filtros distribuídos.
- Comparar via simulações por software, a abordagem proposta com abordagens tradicionais.

1.2 Produções Relacionadas

Durante o desenvolvimento desta tese, o seguinte trabalho foi produzido.

- T.S. Façanha, G. A. Barreto J. T. Costa Filho (2021). "A Novel Kalman Filter Formulation for Improving Tracking Performance of the Extended Kernel RLS". *Circuits Systems and Signal Processing*, v. 40, p. 1397-1419, 2021. DOI: <https://doi.org/10.1007/s00034-020-01533-4>.

1.3 Organização Geral desta Tese

Esta tese está distribuída em cinco capítulos e três apêndices. Um breve comentário acerca dos demais capítulos será feito a seguir.

- **Capítulo 2 - Fundamentos:**

Neste capítulo, apresentaremos os fundamentos teóricos a serem utilizados no decorrer da dissertação envolvendo filtros de Kalman, princípios de dualidade, regiões de confiança e métodos de decomposição numa estrutura de dois níveis.

- **Capítulo 3 - Nova Metodologia para Filtro de Kalman Distribuído:**

Neste capítulo tem como objetivo propor uma abordagem alternativa para problema de filtragem distribuída, explorando a dualidade entre o filtro e o controle, via o método predição por interação. A elaboração do algoritmo distribuído permitirá obter soluções escaláveis que amenizam a relação de dependência das equações do filtro e diminuem o tempo de execução para aplicações em tempo real.

- **Capítulo 4 - Filtro de Kalman Distribuído Com Horizonte de tempo Reduzido:**

Neste capítulo, estende-se a metodologia proposta no capítulo 3 para obtenção de uma família de ganhos para filtros distribuídos, baseado em equações estáticas mais simples para o cálculo do ganho e cálculo da matriz do erro de covariância.

- **Capítulo 5 - Conclusão e Trabalhos Futuros:**

Por fim, finaliza-se este trabalho apresentando algumas conclusões e discutindo algumas

extensões dos resultados obtidos.

- **Apêndice A:**

Apresenta-se os conceitos de Minimização da Forma Quadrática Estática.

- **Apêndice B:**

Apresenta-se o conceito de Contração de Matrizes.

- **Apêndice C:**

Artigos Submetidos.

2 FUNDAMENTOS

Neste capítulo apresentam-se os fundamentos teóricos a serem utilizados no decorrer da tese envolvendo os conceitos sobre filtro de Kalman, filtro de Kalman robusto e principais algoritmos de filtragem distribuída. Serão analisados os conceitos da dualidade entre o filtro e o controle para a solução do problema de filtragem utilizando a minimização quadrática via região de confiança e a metodologia de predição por interação aplicada ao controle.

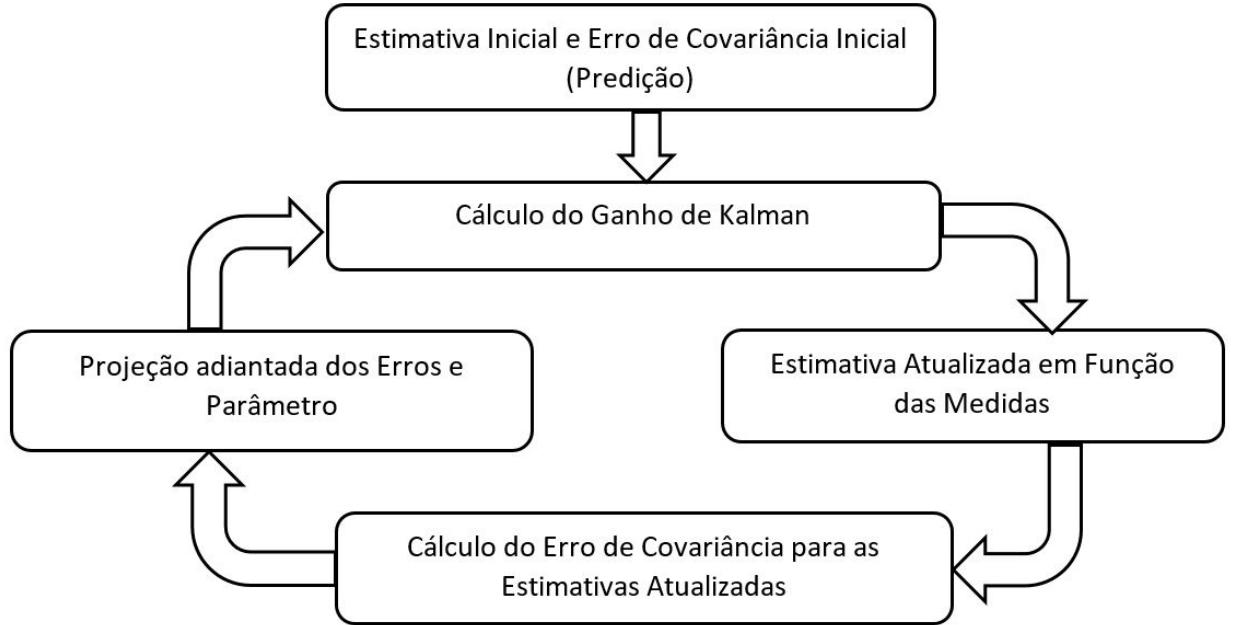
2.1 Filtro de Kalman Global

O filtro de Kalman (KALMAN, 1960) é considerado por muitos o grande avanço da teoria de estimação do século vinte. Muitas realizações desde sua introdução talvez não fossem possíveis sem ele. A filtragem de Kalman vem sendo aplicada em áreas diversas como, aeroespacial, navegação marítima, instrumentação de usinas nucleares, modelamento demográfico, astronomia, meteorologia, economia e indústria em geral. O filtro de Kalman é um estimador para o problema gaussiano linear quadrático. Tal problema consiste na estimação dos estados instantâneos de um sistema linear dinâmico com ruído gaussiano por meio de medições linearmente relacionadas aos estados e também contaminadas por ruído branco. Na prática, o filtro de Kalman é um conjunto de equações que oferece uma solução computacional eficiente para estimadores baseados no método dos mínimos quadrados.

Quando as características da perturbação e do ruído são condizentes com as suposições usadas no desenvolvimento do filtro de Kalman, ou seja, são variáveis aleatórias brancas e suas densidades espectrais são precisamente conhecidas, o filtro de Kalman consiste no processo ótimo de estimação de estados no sentido dos mínimos quadrados e da mínima variância. No caso estacionário, o filtro de Kalman produz erro mínimo em norma H_2 . Entretanto, é difícil ter informações a priori sobre as propriedades estocásticas das perturbações. Neste caso, a estimação de estados baseada no critério dos mínimos quadrados pode sofrer uma degradação por causa das incertezas estatísticas das perturbações (TAKABA, 1996).

O filtro de Kalman discreto é um algoritmo linear, recursivo e de variância mínima para a estimação dos estados desconhecidos de um sistema dinâmico a partir de medições discretas, corrompidas por ruído. Estado, neste contexto, se refere a qualquer quantidade de interesse envolvida na dinâmica do processo. Assim, esse filtro é um procedimento iterativo representado em várias etapas de processamento, conforme podemos observar na Figura 2.

Figura 2 – Diagrama de predição e atualização do filtro de Kalman



Fonte: Elaborada pelo autor.

Para o desenvolvimento do algoritmo devem ser consideradas algumas propriedades estatísticas das variáveis aleatórias gaussianas contaminadas com perturbações do sistema e ruídos de medida. Estas propriedades dão base matemática para o desenvolvimento das principais equações do filtro.

2.1.1 *Formulação do Problema e Resumo do Algoritmo do Filtro de Kalman*

Considere a descrição do sistema dinâmico invariante em tempo discreto dadas por sua equação de estado

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

e por sua equação de saída

$$y_k = C_k x_k + D_k v_k \quad k = 1, 2, 3, \dots \quad (2.2)$$

O vetor de estados a ser estimado é $x_k \in \mathbb{R}^n$, enquanto que $y_k \in \mathbb{R}^m$ é o vetor de medidas. É importante lembrar que o ruído de processo é usado para descrever tanto os erros ocorridos no processo quanto as incertezas no modelo de estados. As matrizes A_k , B_k , C_k e D_k possuem dimensões adequadas e a estimativa inicial x_0 é \bar{x}_0 , sem perda de generalidade.

Supõe-se que o ruído de processo, $w_k \in \mathbb{R}^n$, e o ruído de medida, $v_k \in \mathbb{R}^m$, são variáveis aleatórias independentes, brancas, de média zero e com distribuição de probabilidade

gaussiana. Suas matrizes de correlação são assumidas conhecidas e são representadas por Q e R , respectivamente.

Apesar de não ser trivial determinar Q e R , é importante que estas reflitam a precisão relativa existente entre o modelo dinâmico e o modelo de medição. Isto significa, por exemplo, que se as medições são precisas e o modelo dinâmico é impreciso, os elementos da matriz Q devem ser maiores que aqueles de R . Em particular, se as medições são perfeitas, $R = 0_{m \times m}$ e se o modelo dinâmico é perfeito, $Q = 0_{n \times n}$. Obviamente, isto não acontece na prática e mesmo se ocorresse, um filtro de Kalman não seria a ferramenta mais apropriada para resolver este problema.

A estabilidade do filtro de Kalman resulta da iteração da equação de Riccati que é dado pelo seguinte teorema.

Teorema: (Estabilidade Exponencial do Filtro de Kalman) O filtro de Kalman é nominalmente e exponencialmente estável devido o par (A, C) ser observável e $(A, Q^{\frac{1}{2}})$ ser estabilizável, $R > 0$ e $Q > 0$.

Prova. Ver referência Souza *et al.* (1986)

No filtro de Kalman, o erro de estimativa é causado pelas perturbações, constituídas pelas incertezas na estimativa inicial do vetor de estados e pelos ruídos de processo e de medida. O problema consiste em determinar a estimativa \hat{x}_k do vetor de estados através do seguinte processo de otimização:

$$\min_{x_0, w, v} J(x_0, w, v) \triangleq \|x_0 - \bar{x}_0\|_{\Pi^{-1}} + \sum_{k=0}^K \left(\|w_k\|^2 + \|v_k\|^2 \right), \quad (2.3)$$

sujeito às equações (2.1) e (2.2), onde Π uma matriz definida positiva que representa a incerteza do estado inicial x_0 , sendo \bar{x}_0 a estimativa ótima do estado e a norma quadrática definida por $\|v_k\|^2 = \|y_k - C_k x_k\|_{R^{-1}}^2$.

O filtro de Kalman estima o vetor de estados de um processo através de um controle com realimentação de estados. Por isso, suas equações são recursivas, o que torna o processo de estimativa de estados menos dispendioso em termos de esforço computacional. O filtro de Kalman tem sido vastamente estudado e aplicado em diversos problemas. Aos interessados no tema, recomenda-se a leitura do desenvolvimento matemático do processo de otimização detalhadamente descrito em Zarchan e Musoff (2010).

Na Tabela 1, encontra-se um resumo das equações do filtro de Kalman discreto.

Tabela 1 – Equações do filtro de Kalman discreto

Ação	Equações
Preditor (Estimativa a priori)	$\hat{x}_{k k-1} = A_{k-1}\hat{x}_{k-1 k-1}$
Recursão de Riccati	$P_{k k-1} = A_{k-1}P_{k-1 k-1}A_{k-1}^T + B_{k-1}QB_{k-1}^T$
Atualização do estado	$\hat{x}_{k k} = \hat{x}_{k k-1} + K_k(y_k - C_k\hat{x}_{k k-1})$
Atualização da covariância	$P_{k k} = (I - K_kC_k)P_{k k-1}$
Ganho de Kalman	$K_k = P_{k k-1}C_k^T(C_kP_{k k-1}C_k^T + R)^{-1}$

Fonte: Elaborada pelo autor.

Assim, pode-se estruturar os cálculos recursivos de Kalman, conforme descrito pelo Algoritmo 1.

Algoritmo 1: Filtro de Kalman Discreto

Inicialização:

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

Iteração para $i > 1$

Predição

$$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1|k-1}$$

$$\hat{y}_k = C_k\hat{x}_{k|k-1}$$

$$P_{k|k-1} = A_{k-1}P_{k-1|k-1}A_{k-1}^T + B_{k-1}QB_{k-1}^T$$

Atualização:

$$K_k = P_{k|k-1}C_k^T(C_kP_{k|k-1}C_k^T + R)^{-1}$$

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

$$P_{k|k} = (I - K_kC_k)P_{k|k-1}$$

fim

Na próxima seção, analisaremos o problema de filtragem quando não conhecemos as características estatísticas dos ruídos.

2.2 Filtro de Kalman Robusto (H_∞)

O filtro H_∞ pode ser definido como um problema de minimização da máxima energia do erro de estimativa devido a qualquer tipo de perturbação, ou seja, sem o conhecimento das características estatísticas dos ruídos. A otimalidade do filtro de Kalman está fundamentada em algumas aplicações particulares, onde há a necessidade de conhecimento das características do ruído de processo e de medida. No entanto, na maioria das aplicações, estas informações não estão disponíveis ou não estão bem modeladas. Assim, não podemos garantir um bom

desempenho para este filtro. Enquanto isso, o filtro H_∞ garante um bom desempenho numa determinada faixa de variação do erro de estimação. O problema de otimização H_∞ pode ser considerado como um problema de otimização minimax (SHAKED; THEODOR, 1992).

Assim, faz-se necessário obter um ganho que minimize o efeito da perturbação sobre a saída regulada. Nesta seção, discute-se uma apresentação das características do filtro H_∞ e apresenta-se as equações que definem o filtro robusto para sistemas dinâmicos em tempo discreto.

2.2.1 Formulação do Problema

Considera-se o seguinte sistema dinâmico discreto no tempo regido pelas seguintes equações:

$$x_{k+1} = A_k x_k + B_k w_k, \quad x_0 = 0 \quad (2.4)$$

$$y_k = C_k x_k + D_k v_k, \quad k = 1, 2, 3, \dots \quad (2.5)$$

onde $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^m$ são os vetores de estado e medição, respectivamente. Considera-se o ruído do processo, $w_k \in \mathbb{R}^n$, e o ruído da medida, $v_k \in \mathbb{R}^m$, como variáveis desconhecidas, visto que são sinais arbitrários em $L_2[0, k]$. O ruído do processo é utilizado para descrever tanto erros ocorridos no processo, como também para descrever incertezas do modelo de estados. Assumindo que D_k é não singular, então $R_k = D_k D_k^T > 0$ e que as matrizes A_k, B_k, C_k e D_k possuem dimensões apropriadas.

Este sistema não possui entrada controlada. Então, tem-se que estimar o vetor de estados, x_k , e também a saída regulada do sistema, z_k . Pode-se estimar $z_k \in \mathbb{R}^p$ por

$$z_k = L_k x_k, \quad (2.6)$$

onde L_k é uma matriz com dimensão apropriada. Considera-se \hat{z}_k como a estimativa de z_k baseada em y_0, \dots, y_k . Também supõe-se que a estimativa do estado inicial x_0 , a priori, é definida por \bar{x}_0 . A estimativa, \hat{z}_k , tenta minimizar o erro quadrático de estimação, $\sum_{k=0}^n \|z_k - \hat{z}_k\|^2$, enquanto se tenta maximizar o erro de estimativa ao quadrado das perturbações do sistema. A partir dos valores arbitrários de $\|w_k\|$, $\|v_k\|$ e $\|x_0\|$ que determinam o maior erro de estimação, pode-se definir a função custo, J , como

$$J_{H_\infty}(\hat{z}; x_0, w, v) = \sum_{k=0}^N \|z_k - \hat{z}_k\|^2 - \gamma^2 \left(\sum_{k=0}^N \|w_k\|^2 + \|v_k\|^2 + \|x_0 - \bar{x}_0\|_{\Pi^{-1}}^2 \right), \quad (2.7)$$

onde a constante positiva γ representa a magnitude de uma penalidade com relação à perturbação e Π é uma matriz definida positiva que representa a incerteza do estado inicial, x_0 .

Então, o problema da filtragem H_∞ é encontrar as estimativas \bar{x}_k e \bar{z}_k que satisfaçam a seguinte condição

$$\sup \frac{\sum_{k=0}^N \|z_k - \hat{z}_k\|^2}{\sum_{k=0}^N \|w_k\|^2 + \|v_k\|^2 + \|x_0 - \bar{x}_0\|_{\Pi^{-1}}^2} < \gamma^2, \quad (2.8)$$

Assim, este problema consiste em determinar as estimativas \bar{x}_k e \bar{z}_k que satisfaçam $J_{H_\infty}(\hat{z}; x_0, w, v) \leq 0$, sujeitos a $\sum_{k=0}^N \|w_k\|^2 + \|v_k\|^2 + \|x_0 - \bar{x}_0\|_{\Pi^{-1}}^2 \neq 0$.

O problema da filtragem H_∞ pode ser formulado como um problema minimax da função custo $J_{H_\infty}(\hat{z}; x_0, w, v)$. Ao se comparar a função custo na equação (2.7) com a função custo utilizada para o filtro de Kalman em (2.3), notamos que a equação (2.3) é parte da equação (2.7), onde a função custo do filtro H_∞ pode ser reescrita como

$$J_{H_\infty}(\hat{z}; x_0, w, v) = \sum_{k=0}^N \|z_k - \hat{z}_k\|^2 - \gamma^2 \{J_{H2}(\hat{z}; x_0, w, v)\}. \quad (2.9)$$

Desta forma, quanto maior é o valor do parâmetro γ em comparação com a energia do erro de estimativa, menor é a influência da minimização com respeito a \bar{z}_k no processo de estimativa. Portanto, para valores elevados do parâmetro γ , o filtro H_∞ tende a se aproximar do filtro de Kalman. Como este parâmetro indica o limite superior de energia dos erros de estimativa devido às perturbações e incertezas, ele pode ser usado para determinar o grau de robustez do filtro.

Continuando a formulação do problema do filtro H_∞ , com a utilização da equação (2.5), que relaciona o vetor de medida, y_k , ao ruído de medida, v_k , obtém-se

$$v_k = D_k^{-1}(y_k - C_k x_k). \quad (2.10)$$

Desenvolvendo a norma, $\|v_k\|^2$, chega-se ao seguinte resultado:

$$\begin{aligned} \|v_k\|^2 &= (y_k - C_k x_k)^T \cdot (D_k^T D_k)^{-1} \cdot (y_k - C_k x_k), \\ &= (y_k - C_k x_k)^T \cdot (R_k)^{-1} \cdot (y_k - C_k x_k), \\ &= \|y_k - C_k x_k\|_{R_k^{-1}}^2. \end{aligned} \quad (2.11)$$

Então, substituindo as equações (2.6) e (2.11) na equação (2.7) pode-se reescrever a função custo do filtro H_∞ , da seguinte forma:

$$J_{H_\infty}(\hat{z}; x_0, w, v) = \sum_{k=0}^N \|L_k x_k - \bar{z}_k\|^2$$

$$-\gamma^2 \left(\sum_{k=0}^N \|w_k\|^2 + \sum_{k=0}^N \|y_k - C_k x_k\|_{R_k^{-1}}^2 + \|x_0 - \bar{x}_0\|_{\Pi^{-1}}^2 \right). \quad (2.12)$$

Assim, o problema minimax entre \hat{z}_k e (x_0, w_k, v_k) passa a ser entre \hat{z}_k e (x_0, w_k, y_k) .

As estimativas ótimas são denotadas, respectivamente, por \hat{z}_k^* e (x_0^*, w_k^*, y_k^*) .

O ruído de processo ou perturbação do pior caso é representado por w_k^* . O ruído de medida do pior caso, v_k^* , corresponde à medida y_k^* . A solução ótima do problema minimax é definida por $(\hat{z}_k^*, x_0^*, w_k^*, y_k^*)$.

Para maximizar a equação (2.12) em relação x_0 e w_k , obtém-se a seguinte função hamiltoniana da equação (2.12):

$$H_k = \frac{1}{2} \gamma^{-2} \cdot \|L_k x_k - \hat{z}_k\|^2 - \frac{1}{2} \cdot \left\{ \|w_k\|^2 + \|C_k x_k\|_{R_k^{-1}}^2 \right\} - \frac{1}{2} \left\{ \|x_0 - \bar{x}_0\|_{\Pi^{-1}}^2 \right\} + \lambda_{k+1}^T (A_k x_k + B_k w_k), \quad (2.13)$$

onde o simbolo λ_k denota um vetor adjunto atuando no vetor de estados da função hamiltoniana.

Seja (x_k^*, λ_k^*) a trajetória ótima de (x_k, λ_k) que corresponde à perturbação do pior caso w_k^* . Assim, verificam-se as seguintes condições necessárias de otimalidade:

1. Condição de estado:

$$\frac{\partial H_k}{\partial \lambda_{k+1}} = x_{k+1}^* \implies x_{k+1}^* = A_k x_k^* + B_k w_k^*$$

2. Condição do hamiltoniano:

$$\frac{\partial H_k}{\partial w_k} = 0 \implies \frac{\partial}{\partial w_k} \left\{ -\frac{1}{2} w_k^T w_k + \lambda_{k+1}^T B_k w_k \right\} = 0 \implies w_k^* = B_k \lambda_{k+1}^*$$

3. Condição de co-estado: $\frac{\partial H_k}{\partial x_k} = \lambda_k^*$

$$\lambda_k^* = \gamma^{-2} L_k^T L_k x_k^* - \gamma^{-2} L_k^T \hat{z}_k + A_k^T \lambda_{k+1}^* + C_k^T R_k^{-1} y_k - C_k^T R_k^{-1} C_k x_k^*$$

4. Condição de contorno: $\frac{\partial H_0}{\partial x_0} = \lambda_0^*$

$$\lambda_0^* = \frac{\partial}{\partial x_0} \left\{ -\frac{1}{2} (x_0^T - \bar{x}_0^T) \Pi^{-1} (x_0 - \bar{x}_0) \right\}$$

assim, tem-se

$$x_0^* = \bar{x}_0 + \Pi \lambda_0^*.$$

As equações encontradas a partir da verificação das condições necessárias de otimalidade constituem um problema de TPBVP (*Two Point Boundary Values Problem*) (TAKABA, 1996), conforme pode-se observar a seguir:

$$\begin{bmatrix} I_n & -B_k B_k^T \\ 0 & A_k^T \end{bmatrix} \begin{bmatrix} x_{k+1}^* \\ \lambda_{k+1}^* \end{bmatrix} = \begin{bmatrix} A_k & 0 \\ C_k^T R_k^{-1} C_k - \gamma^{-2} L_k^T L_k & I_n \end{bmatrix} \begin{bmatrix} x_k^* \\ \lambda_k^* \end{bmatrix}$$

$$+ \begin{bmatrix} 0 \\ -C_k^T R_k^{-1} y_k + \gamma^{-2} L_k^T \hat{z}_k \end{bmatrix}, \quad (2.14)$$

para $x_0^* = \bar{x}_0 + \Pi \lambda_{k+1}^*$ e $\lambda_{k+1}^* = 0$.

A partir da análise destas condições de otimalidade no hamiltoniano da função custo do filtro robusto (2.13) e após algumas manipulações algébricas, obtém-se as equações do filtro, bastante conhecida na literatura (TAKABA, 1996) e apresentadas na Tabela 2.

Tabela 2 – Equações do filtro de Kalman robusto (H_∞)

Ação	Equações
Saída Regulada do Sistema	$z_k = L_k \hat{x}_{k k-1}$
Ganho de Kalman	$G_k = P_{k k-1} C_k^T (C_k P_{k k-1} C_k^T + R_k)^{-1}$
Atualização do estado	$\hat{x}_{k k} = \hat{x}_{k k-1} + G_k (y_k - C_k \hat{x}_{k k-1})$
Atualização da covariância	$P_{k+1} = A_k P_k E_k^{-1} A_k^T + B_k B_k^T, \quad P_0 = \Pi$ $E_k = I_n + (C_k^T R_k^{-1} C_k - \gamma^{-2} L_k^T L_k) P_k$

Fonte: Elaborada pelo autor.

Na seção a seguir, serão analisadas algumas variações do filtro de Kalman distribuído baseadas em consenso. Estes algoritmos farão parte da análise comparativa de desempenho com os algoritmos propostos nesta tese.

2.3 Algoritmos de Filtragem Distribuída baseados em Consenso

As redes de sensores se tornaram comuns nas últimas décadas devido à ampla variedade de aplicações, como por exemplo, em monitoramento de ambientes (OLIVEIRA; RODRIGUES, 2011), (HWANG *et al.*, 2010) e coleta de dados científicos de fontes distribuídas (LUO *et al.*, New York, NY, USA, 2009), (ZIMOS *et al.*, 2016). Neste contexto, sugeriram diversos algoritmos de consenso, ou seja, todos os nós de uma rede de sensores devem eventualmente chegar a um consenso sobre os valores de estimativas de cada nó. Em geral, a premissa para esta abordagem é considerar o estado alvo coletivamente observável por todos os sensores, ou seja, o par (A, C) com $C = [(C^1)^T, \dots, (C^N)^T]$ é observável.

Assim, pode-se citar os seguintes algoritmos distribuídos baseados em consenso (OLFATI-SABER; MURRAY, 2004): filtro consenso multi-agente (MKCF) (REN *et al.*, Portland, OR, USA, 2005), filtro de Kalman distribuído (DKF) (OLFATI-SABER, New Orleans, LA, USA, textbf2007), filtro de Kalman-Bucy distribuído (DKBF) (WU *et al.*, 2016), filtro de Kalman consenso ponderado em informação (IKCF) (JI *et al.*, 2017) e filtro de Kalman distribuído

aproximado (ADKF) (BATTILOTTI *et al.*, 2020). Na próxima subseção serão apresentadas algumas características de cada modelo supracitado.

2.3.1 Filtro de Kalman Distribuído

A estrutura proposta por Olfati-Saber (New Orleans, LA, USA, textbf2007) insere o conceito de consenso nas equações de cálculo do ganho, cálculo da equação recursiva de Riccati e cálculo da estimativa de atualização do estado, apresentados na Tabela 1, do filtro de Kalman global. Assim, cada nó executa um filtro de Kalman local de acordo com as seguintes equações:

$$P^i = AP^i + P^iA^T + BQB^T - P^i(C^T)^i(R^i)^{-1}C^iP^i, \quad (2.15)$$

$$K^i = P^i(C^T)^i(R^i)^{-1}, \quad (2.16)$$

$$\dot{\hat{x}}_k^i = A\hat{x}_k^i + K^i(y^i - C^i\hat{x}_k^i) + \gamma P^i \sum_{j \in N_i} (\hat{x}_k^j - \hat{x}_k^i), \quad (2.17)$$

onde $\gamma > 0$ e a covariância do erro da estimativa de estado distribuídos são as matrizes $P = E \left\{ (\hat{x}_k^i - x_k) (\hat{x}_k^i - x_k)^T \right\}$.

Ao final da equação (2.17) é adicionado o termo de consenso $\sum_{j \in N_i} (\hat{x}_k^j - \hat{x}_k^i)$ e leva em consideração as estimativas do estado dos nós sensores vizinhos. Se não houver ruído de medição ou de processo, então todos os nós alcançam as estimativas de consenso como que $\hat{x}_k^i \rightarrow x_k$. Nesta proposta o termo consenso é adicionado temporariamente fora da estrutura do filtro de Kalman ideal. Assim, o resultado das estimativas não podem mais ser garantidas como localmente ideais, principalmente quando há ruído na comunicação entre os sensores. Se nem um nó sensor, nem seus vizinhos imediatos podem medir a saída do nó de destino, chama-se esta situação de *nó-ingênuo*. A presença de nós ingênuos diminui o desempenho do filtro de Kalman consenso distribuído, conforme demonstrado por Kamal *et al.* (2013). Na próxima subseção apresentaremos algumas características do filtro de Kalman consenso multiagente.

2.3.2 Filtro de Kalman Consenso Multiagente

O filtro de Kalman consenso multiagente foi originalmente proposto em Ren *et al.* (Portland, OR, USA, 2005) e estendido em Ji *et al.* (2017). O algoritmo MKCF precisa trocar as estimativas e as matrizes P^i entre os nós vizinhos. Esta proposta pode ser aplicada em conjuntos

mais amplos da rede em relação ao *DKF*, mas é limitado pela suposição de um nó estático, ou seja, $C^i = I$ na equação (2.1). Dado o seguinte modelo para cada nó sensor:

$$y_t^i = C^i x_t + v^i = \begin{pmatrix} a^{i1}(\hat{x}_t^i + w^{i1}) \\ \vdots \\ a^{iN}(\hat{x}_t^N + w^{iN}) \end{pmatrix}, \quad (2.18)$$

onde numa rede de sensores uma coleção de N nós sensores são interconectados, numa topologia de interação em grafos, com $\mathfrak{g} = (v, \xi, \mathfrak{f})$ em que $v = \{1, 2, \dots, N\}$ é o conjunto de vértices representando os sensores, $\xi \subseteq v \times v$ é o conjunto de arestas que representam os canais de comunicação, e $\mathfrak{f} = [a^{ij}] \in \mathbb{R}^{N \times N}$. Neste contexto, w^{iN} são os ruídos de comunicação ou transmissão do canal, assim pode-se definir as equações do filtro MKCF, como

$$\dot{P}^i = P^i \left[\sum_{j=1}^N a^{ij} (P^j + \Pi^{ij})^{-1} \right] P^i + S, \quad (2.19)$$

$$K^{ij} = P^i (P^j + \Pi^{ij})^{-1}, \quad (2.20)$$

$$\dot{\hat{x}}_t^i = \sum_{j=1}^N a^{ij} K^{ij} \left[(\hat{x}_t^j + w^{ij}) - \hat{x}_t^i \right], \quad (2.21)$$

onde $\Pi^{ij} = E \{ w^{ij} (w^i)^T \}$ para todo $i = 1, 2, \dots, N, j \in N^i$.

Observa-se que tanto a equação (2.18) do modelo de medida como as equações (2.19) - (2.21) do filtro MKCF não contêm o estado alvo, assim não está claro que tipo de estimativa consenso é alcançada. Além disso, a demonstração apresentada por Ren *et al.* (Portland, OR, USA, 2005) assume que os ganhos do filtro são escalares e positivos, enquanto que K^{ij} é uma matriz que pode não ser definida positiva. Na próxima subseção serão apresentadas algumas características do filtro de Kalman-Bucy distribuído.

2.3.3 Filtro de Kalman-Bucy Distribuído

O filtro de Kalman-Bucy distribuído (WU *et al.*, 2016) é construído sobre a abordagem de Kim *et al.* (Las Vegas, NV, USA, 2016), onde foi mostrado que a precisão centralizada ideal pode ser melhorada se os ganhos de consenso forem escolhidos levando-se em consideração o erro da covariância da estimativa de cada nó. Como isso não é viável em uma configuração distribuída, isto resultou em uma escolha conservadora do ganho de consenso em Wu *et al.*

(2016) que propuseram as seguintes equações:

$$\dot{\hat{x}}_t^i = A\hat{x}_t^i + P^i(C^i)^T(R^i)^{-1}(y^i - C^i\hat{x}_t^i) + P^i \sum_{j \in N^i} [(P^j)^{-1}(\hat{x}_t^j - \hat{x}_t^i)] \quad (2.22)$$

e

$$\dot{P}^i = AP^i + P^iA^T + Q - P^i(C^i)^T(R^i)^{-1}C^iP^i - P^i \sum_{j \in N^i} [(P^j)^{-1} - (P^i)^{-1}] P^i. \quad (2.23)$$

Assim, o filtro DKBF garante a estabilidade ao calcular uma média do erro de estimativa sob uma suposição *ad hoc*, mas aumenta a complexidade quando se faz uma verificação da convergência do filtro. Esta metodologia é semelhante a abordagem em Olfati-Saber (New Orleans, LA, USA, textbf2007), porém o termo $(\hat{x}_t^j - \hat{x}_t^i)$ é ponderado por $P^i(P^j)^{-1}$. Isto significa, por exemplo, que $(\hat{x}_t^j - \hat{x}_t^i)$ sofre menos influência quando $P^i < P^j$, ou seja, se a estimativa \hat{x}_t^i for mais confiável que \hat{x}_t^j . Na próxima subseção serão apresentadas algumas características do filtro de Kalman consenso ponderado em informação.

2.3.4 Filtro de Kalman Consenso Ponderado em Informação

O algoritmo proposto por Ji *et al.* (2017) fornece um filtro de consenso distribuído ideal, cuja convergência possa ser formalmente comprovada. A abordagem consiste em impor a topologia do fluxo de informações antecipadamente, assumindo uma forma distribuída para as matrizes de medição dos nós sensores. No próximo passo, as equações do filtro de Kalman são formalmente aplicadas para desenvolver filtros de consenso distribuídos localmente ideais, em que os efeitos de estimativas ruins de vizinhos são descontados usando termos de ponderação de informação. O filtro IKCF modela explicitamente os ruídos de comunicação. Na ausência destes ruídos, tem-se

$$\dot{\hat{x}}_t^i = A\hat{x}_t^i + g^i P^i (C^i)^T (R^i)^{-1} (y^i - C^i \hat{x}_t^i) + P^i \sum_{j \in N^i} \left[(P^j)^{-1} (\hat{x}_t^j - \hat{x}_t^i) \right] \quad (2.24)$$

e

$$\begin{aligned} \dot{\hat{P}} = & (A + |\mathcal{N}^i|I)P^i + P^i(A + |\mathcal{N}^i|I)^T + Q - (g^i)^2 P^i (C^i)^T (R^i)^{-1} C^i P^i \\ & - P^i \sum_{j \in N^i} \left[(P^j)^{-1} - (P^i)^{-1} \right] P^i. \end{aligned} \quad (2.25)$$

As demonstrações e prova de convergência são apresentadas em Ji *et al.* (2017) e também uma análise crítica sobre problemas técnicos desta abordagem são citados em Cacace

(2019). Entretanto, pode-se citar as principais contribuições desta proposta. A topologia do fluxo de informações é formalmente capturada antecipadamente na formulação do filtro de Kalman, definindo um modelo de medição ponderada de informações distribuídas para cada nó. Os ruídos de processo e de comunicação de canal estão incluídos. As estimativas localmente ótimas são então derivadas com base no modelo de medição distribuída em cada nó, empregando formalmente as equações do filtro de Kalman. O resultado é um filtro de Kalman consenso totalmente distribuído que inclui ponderação de informações para descontar estimativas incorretas de nós vizinhos.

Na próxima subseção serão apresentadas algumas características do filtro de Kalman distribuído aproximado.

2.3.5 Filtro de Kalman Distribuído Aproximado

O algoritmo ADKF (BATTILOTTI *et al.*, 2020) estende o algoritmo DKF (OLFATI-SABER, New Orleans, LA, USA, textbf2007) para o caso de redes com nós que possuem capacidades de medição limitadas ou nulas e fornecem uma estimativa mais precisa. Esta abordagem consiste em um filtro para cada nó sensor da rede. Assim, a equação do nó sensor i é definida por

$$\dot{\hat{x}}_t^i = A\hat{x}_t^i + K^i(y_t^i - C^i\hat{x}_t^i) + \gamma P_\infty \sum_{j \in N^{(i)}} (\hat{x}_t^j - \hat{x}_t^i) \quad (2.26)$$

e

$$0 = AP_\infty + P_\infty A^T + Q - P_\infty C^T R^{-1} C P_\infty, \quad (2.27)$$

onde $K^i = NP_\infty(C^i)^T(R^i)^{-1}$ e a matriz do erro de covariância $P_t \in P_+(n)$ é limitada para todo $t \geq 0$ e $P_t \rightarrow +\infty$ com $P_t \in P_+(n)$, sendo a única solução da equação (2.27). Assim, cada y_t^i representa os dados disponíveis no nó i , $i = \{1, 2, \dots, N\}$, na rede. O objetivo é projetar um estimador de estado distribuído para modelo de processo com a topologia da rede em grafos $G = \{v, \varepsilon\}$ fornecida em que $v = \{1, 2, \dots, N\}$ e $\{\varepsilon \subseteq vxv\}$. Os estimadores distribuídos consistirão em N estimadores locais, um para cada nó, que trocam suas estimativas locais com os vizinhos. Seja $\gamma > 0$ um parâmetro escolhido de modo que a matriz $A_D(\gamma)$ considere o critério de estabilidade de Hurwitz. Assim, tem-se que

$$A_D(\gamma) := \text{diag}^i(A^i) - \gamma(\mathcal{E} \otimes P_\infty) \in \mathbb{R}^{nN \times nN} \quad (2.28)$$

$$A^i := A - K^i C^i, \quad i = 1, \dots, N. \quad (2.29)$$

onde \otimes é o produto de Kronecker (SCHACKE, 2004).

As demonstrações dos critérios e teoremas deste filtro são apresentadas em Battilotti *et al.* (2020). Caso se substituam os ganhos de K^i em ADKF por ganhos não ótimos, é possível obter um algoritmo denominado filtro de Kalman distribuído subótimo - SDKF (BATTILOTTI *et al.*, 2020), que chega a um consenso, mas não é mais ideal para $\gamma \rightarrow \infty$.

Na seção a seguir, serão abordados os conceitos de dualidade entre o filtro de Kalman e o regulador linear quadrático (LQR).

2.4 Dualidade entre Filtro de Kalman e o Regulador Linear Quadrático

O exemplo melhor conhecido da dualidade entre estimação e controle é a dualidade entre o filtro de Kalman e o regulador linear quadrático (NAIDU, 2002). Este resultado teórico torna possível a aplicação de algoritmos de estimação em problemas de controle e vice-versa. Neste trabalho, esta dualidade é considerada na dedução de uma nova metodologia de síntese do filtro de Kalman.

Seguindo a notação adotada para a formulação do problema de filtro de Kalman, apresenta-se sucintamente o seguinte problema de controle ótimo de sistema dinâmico invariante no tempo:

Sistema Dinâmico: $x_{k+1} = Ax_k + Hu_k + Bw_k$

Função Custo: $J(x_k, u_k, k) = \frac{1}{2}x_k^T Q x_k + \frac{1}{2}u_k^T R u_k$

Para a obtenção do controlador ótimo é também empregada a seguinte função quadrática:

$$v(x_k, k) = \frac{1}{2}x_k^T V_k x_k. \quad (2.30)$$

Definindo a equação de Bellman (BERTSEKAS; TSITSIKLIS, 1995), tem-se

$$\frac{1}{2}x_k^T V_k x_k = \min_u \left\{ \frac{1}{2}u_k^T R u_k + x_k^T Q x_k + \frac{1}{2}(Ax_k + Hu_k)^T V_{k+1} (Ax_k + Hu_k) \right\}. \quad (2.31)$$

De modo que, ao minimizar analiticamente o hamiltoniano, resulta em

$$u_k = -(R + H^T V_{k+1} H)^{-1} H^T V_{k+1} A x_k. \quad (2.32)$$

Substituindo a equação (2.32) na equação de Bellman, encontraremos a equação de Riccati em tempo discreto:

$$V_k = Q + A^T V_{k+1} A - A^T V_{k+1} H (R + H_k^T V_{k+1} H)^{-1} H_k^T V_{k+1} A, \quad (2.33)$$

onde pode-se observar que a equação (2.33) é idêntica a equação do cálculo de $P_{k/k-1}$, descrita na tabela 1.

2.4.1 Relação entre as Equações de Riccati discreta e contínua

Para o regulador linear quadrático, a equação de Riccati em tempo contínuo pode ser definida utilizando as equações de Hamilton-Jacobi-Bellman (HJB) (BARDI; CAPUZZO-DOLCETTA, 2008):

$$-\dot{V} = Q + A^T V + V A - V H R^{-1} H^T V. \quad (2.34)$$

Para o filtro de Kalman-Bucy (MORTENSEN, 1968), a equação de Riccati em tempo contínuo é dada como

$$\dot{S} = C^T (D D^T)^{-1} C - A^T S - S A - S B R^{-1} B^T S, \quad (2.35)$$

onde S é a inversa da matriz de covariância.

Comparando a equação de Riccati para o regulador linear quadrático (2.34) e filtro de Kalman (2.35), obtém-se a dualidade de Kalman em tempo contínuo mostrada na Tabela 3:

Tabela 3 – Filtro de Kalman (Informação) x LQR

Regulador Linear Quadrático	Filtro de Kalman
V	P^{-1}
A	$-A$
$H R^{-1} H^T$	$B R^{-1} B^T$
Q	$C^T (D D^T)^{-1} C$

Fonte: Elaborada pelo autor.

Porem, é possível obter uma nova forma de dualidade entre controle e estimação. Seja a inversa da matriz de covariância $S = P^{-1}$. Esta correspondência resulta numa equivalência entre V e P^{-1} (TODOROV, 2006). Pode-se explorar o fato de que, se uma matriz simétrica é definida positiva, a derivada no tempo de sua inversa é dada por

$$\frac{d}{dt} [S^{-1}] = -S^{-1} \dot{S} S^{-1}. \quad (2.36)$$

Substituindo $S = P^{-1}$ na equação (2.35), obtém-se

$$\dot{P} = -P^{-1}\dot{S}P^{-1}. \quad (2.37)$$

Na equação de Riccati (2.35) substituindo S^{-1} por P , tem-se

$$\dot{S} = C^T(DD^T)^{-1}C - A^T P^{-1} - P^{-1}A - P^{-1}BR^{-1}B^T P^{-1}. \quad (2.38)$$

Substituindo a equação (2.38) na (2.37) e organizando os termos, obteremos a seguinte equação de Riccati para filtros:

$$\begin{aligned} \dot{P} &= -P[C^T(DD^T)^{-1}C - A^T P^{-1} - P^{-1}A - P^{-1}BB^T P^{-1}]P, \\ &= BR^{-1}B^T + AP + PA^T - PC^T(DD^T)^{-1}CP. \end{aligned} \quad (2.39)$$

Assim, comparando esta nova forma da equação de Riccati (2.39) com a equação (2.34), chega-se a equivalência mostrada na Tabela 4.

Tabela 4 – Dualidade: Filtro de Kalman x LQR

Regulador Linear Quadrático	Filtro de Kalman
V	P
A	A^T
H	C^T
R	DD^T
Q	$BR^{-1}B^T$

Fonte: Elaborada pelo autor.

Esta dualidade não pode ser generalizada, é válida especificamente para filtros e controle LQR gaussiano (NAIDU, 2002).

Nesta seção, observou-se a dualidade entre o filtro e o controle LQR em tempo contínuo mas também se pode considerar esta mesma relação em tempo discreto, através de uma aproximação de Euler, dada por

$$x_k \leftarrow x(k\Delta), A \leftarrow (I + \Delta B), R \leftarrow \Delta R, Q \leftarrow \Delta Q, \quad (2.40)$$

onde Δ é o intervalo de tempo, utilizado na aproximação de Euler para mapear um sistema contínuo em um sistema discreto, ou seja, obter uma solução exata para o controle LQR em tempo discreto.

Será analisada uma relação entre as equações de Riccati para controle LQR em tempo contínuo e discreto, conforme mostrado na Proposição 2.1 (FACANHA, 2010). Desta relação fica evidente que a equação de Riccati em tempo discreto (2.34) possui mais termos do que a sua versão em tempo contínuo mostrada na equação (2.35).

Proposição 2.1. As equações de Riccati em tempo contínuo e tempo discreto descritas pela equação (2.41) são idênticas no limite $\Delta \rightarrow 0$.

$$\frac{V_k - V_{k-1}}{\Delta} = Q + A^T V_{k+1} + V_{k+1} A - V_{k-1} H (R + \Delta H_k^T V_{k+1} H)^{-1} H_k^T V_{k+1} + \frac{o(\Delta^2)}{\Delta} \quad (2.41)$$

onde $o(\Delta^2)$ são os termos de segunda ordem.

Prova: Substituir as matrizes da equação (2.33) com seus análogos em tempo contínuo (2.40), conforme pode-se observar em Facanha (2010).

Na próxima seção, será apresentada a teoria de um sistema hierárquico de dois níveis utilizando o método de decomposição por coordenação via predição por interação aplicado ao controle (COSTA, 1988),(GOMIDE, 1979). Estes conceitos de controle são importantes para estender sua aplicação aos filtros distribuídos explorando sua dualidade.

2.5 Formulação Primal-Dual

Seja o problema de otimização dinâmica e a função custo aditivamente separável. O problema pode ser posto na seguinte forma:

$$\min_{u_k} J(x_k, u_k) = \min_{u_k} \sum_{i=1}^N J^i(x_k^i, u_k^i) \quad (2.42)$$

sujeito a

$$x_{k+1}^i = g^i(x_k^i, u_k^i, z_k^i), \quad (2.43)$$

$$z_k^i = \sum_{i \neq j}^N g^{ij}(x_k^j, u_k^j), \quad (2.44)$$

$$x_0^i = x^{i0}. \quad (2.45)$$

O lagrangiano para este problema é dado por

$$L(x_k^i, u_k^i, \beta_k^i, z_k^i) = \sum_{i=1}^N \left\{ \sum_{k=0}^{K-1} \left[f^i(x_k^i, u_k^i) + (\beta_k^i)^T \left(z_k^i - \sum_{i \neq j}^N g^{ij}(x_k^j, u_k^j) \right) + (\lambda_k^i)^T (-x_k^i + g^i(x_k^i, u_k^i, z_k^i)) \right] \right\},$$

(2.46)

onde β_k^i são parâmetros de Lagrange e λ_{k+1}^i são as variáveis adjuntas (coestado) do sistema. Para que se possa decompor o lagrangiano em sublagrangianos é necessário fazer a separação do termo

$$\sum_{i=1}^N (\beta_k^i)^T \sum_{i \neq j}^N g^{ij}(x_k^j, u_k^j) = \sum_{i=1}^N \sum_{i \neq j}^N (\beta_k^i)^T g^{ij}(x_k^j, u_k^j), \quad (2.47)$$

pois é uma função das variáveis dos subsistemas j . Isto pode ser realizado pela introdução de um nível superior de coordenação que impõe o valor de certas variáveis do primeiro nível, onde são tratados como constantes. O tipo de decomposição por coordenação obtido depende da natureza destas variáveis ou da relação destas variáveis com a variável de coordenação. No entanto, a variável dual β_k como a variável primal z_k são parâmetros de coordenação (método primal - dual, método misto).

Utilizando a relação (2.46), a equação (2.47) pode ser reescrita como

$$L(.) = \sum_{i=1}^N \left\{ \sum_{k=0}^{K-1} f^i(x_k^j, u_k^j) + (\beta_k^i)^T z_k^i - \sum_{j \neq i}^N (\beta_k^i)^T g^{ji}(x_k^j, u_k^j) + \lambda_{k+1}^i (-x_{k+1}^i + g^i(x_k^j, u_k^j, z_k^i)) \right\} \quad (2.48)$$

ou

$$L(.) = \sum_{i=1}^N L^i(.),$$

onde

$$L^i(.) = \sum_{k=0}^{K-1} \left\{ f_i(x_k^j, u_k^j) + (\beta_k^i)^T z_k^i + \sum_{j \neq i}^N (\beta_k^i)^T g^{ji}(x_k^j, u_k^j) + (\lambda_{k+1}^i)^T (-x_{k+1}^i + g^i(x_k^j, u_k^j, z_k^i)) \right\} \quad (2.49)$$

Como β_k e z_k são fixados pelo coordenador, o lagrangiano é aditivamente separável e podemos associar para cada subsistema i o i -ésimo sublagrangiano. Portanto, o problema do lagrangiano global, pode ser resolvido através de cada subproblema do lagrangiano independentemente, que são mais fáceis de serem resolvidos, pois são de dimensão reduzida.

Consequentemente, a função objetivo do problema dual será dada por

$$D(\beta_k) = \min_{x^i, u^i} L(x_k^i, u_k^i, \beta_k^i, z_k^i), \quad (2.50)$$

e o problema dual por

$$\max_{\beta_k} D(\beta_k). \quad (2.51)$$

Os subproblemas do lagrangiano são, portanto, dados por

$$\min_{x^i, u^i} \sum_{k=0}^{K-1} \left\{ f^i(x_k^i, u_k^i) + (\beta_k^i)^T z_k^i + \sum_{j \neq i}^N (\beta_k^i)^T g^{ji}(x_k^j, u_k^j) \right\} \quad (2.52)$$

Sujeito a

$$x_{k+1}^i = g^i(x_k^i, u_k^i, z_k^i), \quad (2.53)$$

$$x_0^i = x^{i0} \quad i = 1, \dots, N. \quad (2.54)$$

Podemos ver pela equação (2.52) que a função objetivo ou o função custo associado ao i -ésimo subproblema é sensivelmente modificada. Por outro lado, desde que as variáveis de interconexão são fixadas, indiretamente as variáveis de estado e controle também são fixadas. Portanto, trata-se também de um método de coordenação pelo modelo, daí então, a razão para o nome *método misto*.

O problema fundamental no estabelecimento de condições de coordenação é como modificar problemas de nível inferior para tornar o sistema coordenado. Essa abordagem é um exemplo clássico dessas modificações.

As equações de otimalidade para a solução dos subproblemas do nível inferior podem ser obtidas pela aplicação de técnicas como a programação dinâmica (BRYSON; HO, 2018) ou ainda pela combinação de várias técnicas dependendo da natureza de cada subproblema.

2.6 Forma Dual - Decomposição Paramétrica

Seja o problema

$$\min_u J(x_k, u_k) = \phi(x_k) + \sum_{k=0}^{K-1} f(x_k, u_k) \quad (2.55)$$

sujeito a

$$x_{k+1} = g(x_k, u_k), \quad (2.56)$$

$$x_0 = x^0. \quad (2.57)$$

Assim, o lagrangiano é dado por

$$L(x_k, u_k, \lambda_k) = \phi(x_k) - \lambda_{k-1}^T x_k + \sum_{k=0}^{K-1} \{ f(x_k, u_k) - \lambda_{k-1}^T (x_{k+1} - g(x_k, u_k)) \}. \quad (2.58)$$

Consequentemente a função objetivo do problema dual será escrita como

$$D(\lambda_k) = \min_{x,u} L(x_k, u_k, \lambda_k). \quad (2.59)$$

Reescrevendo o lagrangiano, obtém-se

$$L(\cdot) = \phi(x_k) - \lambda_{k-1}^T x_k + \sum_{k=0}^{K-1} \{ f(x_k, u_k) - \lambda_{k-1}^T x_{k+1} + \lambda_{k-1}^T g(x_k, u_k) \}. \quad (2.60)$$

Definindo o hamiltoniano.

$$H(x_k, u_k, \lambda_k) = f(x_k, u_k) + \lambda_k^T g(x_k, u_k), \quad (2.61)$$

o lagrangiano pode ser escrito, manipulando o hamiltoniano (2.61), como

$$L(\cdot) = \sum_{k=0}^{K-1} \{ H(x_k, u_k, \lambda_k) - \lambda_{k-1}^T x_k \} + \phi(x_k) - \lambda_{k-1}^T x_k. \quad (2.62)$$

Podemos observar, pela equação (2.62), que as variáveis do sistema estão relacionadas sempre no mesmo estágio $k, k = 0, \dots, K$. Pode-se então definir $k+1$ subsistemas, uma vez que as variáveis são os parâmetros de coordenação e, portanto, fixadas pelo nível superior, permitindo com isto, a decomposição do lagrangiano, pois está agora na forma separável aditiva.

Os problemas associados a cada subsistema se definem como, para $\lambda_k = \lambda_k^*$ fixado pelo coordenador.

Para $k = 0$:

$$\min_{u_0} H(x_0, u_0, \lambda_k^*) \quad (2.63)$$

sujeito a $x_0 = x^0$

Para $k = 1, 2, \dots, K-1$

$$\min_{x_k, u_k} \{ H(x_k, u_k, \lambda_k^*) - (\lambda_{k-1}^*)^T x_k \} \quad (2.64)$$

Para $k = K$:

$$\min_{x_k} \{ \phi(x_k) - (\lambda_{k-1}^*)^T x_k \} \quad (2.65)$$

tal que

$$\lambda_{-1} = 0.$$

$$\lambda_k = 0, k > K.$$

2.7 Método de Decomposição numa estrutura de dois níveis

A teoria dos sistemas hierárquico multiníveis tem sido utilizada na elaboração de técnicas para a solução de problemas de otimização dinâmica. Estas técnicas possuem características diferentes e vários graus de potencialidades e aplicabilidade. Nesta subseção apresentaremos o método denominado predição por interação proposto por Singh *et al.* (1976). Assim, seja um sistema dinâmico descrito pela seguinte equação:

$$x_{k+1} = g(x_k, u_k), \quad (2.66)$$

que, por hipótese, é constituído de um conjunto de N subsistemas isolados. Tem-se assim que

$$x_{k+1}^i = g^i(x_k^i, u_k^i, z_k^i), \quad i = 1, 2, \dots, N. \quad (2.67)$$

interconectados pela seguinte relação:

$$z_k^i = \sum_{j \neq i}^N g^{ij}(x_k^j, u_k^j) \quad (2.68)$$

onde

$$x_k \in E^n, \quad x_k^i \in E^{n_i}$$

$$u_k \in E^m, \quad u_k^i \in E^{m_i}$$

$$x_k = [x_k^1, \dots, x_k^N]^T$$

$$u_k = [u_k^1, \dots, u_k^N]^T$$

O problema de otimização é determinar a sequência de vetores de controle u_k tal que u_k minimize a seguinte função custo:

$$J(x_k, u_k) = \sum_{k=0}^{K-1} f(x_k, u_k). \quad (2.69)$$

Assumindo que a função custo seja aditivamente separável, o problema global pode ser reformulado, permitindo que o problema original seja reescrito como

$$\min_{u_k^i} \sum_{i=1}^N J(x_k^i, u_k^i) \quad (2.70)$$

sujeito a

$$x_{k+1}^i = g^i(x_k^i, u_k^i, z_k^i), \quad (2.71)$$

$$z_k^i = \sum_{i \neq j}^N g^{ij}(x_k^j, u_k^j), \quad (2.72)$$

$$x_0^i = x^{i0}. \quad (2.73)$$

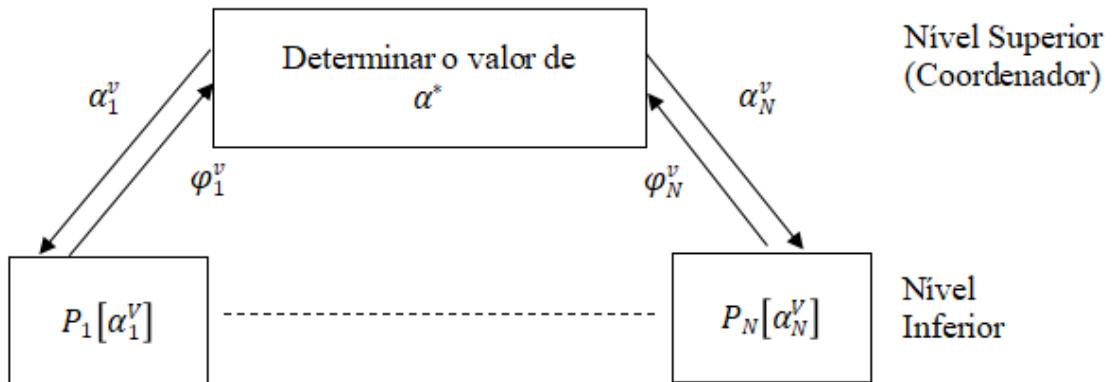
O problema global é transformado num conjunto de subproblemas pela introdução da relação de interconexão mostrada em (2.72). O fato de que o problema global não é uma coleção de subproblemas independentes é refletido pela restrição de interconexão. Este procedimento constitui o primeiro passo na construção de estruturas de dois níveis.

O passo seguinte consiste na composição do problema global na forma descrita pelas equações (2.70) a equação (2.73) em uma família de N subproblemas independentes parametrizados, de estrutura mais simples e de dimensão reduzida.

Sejam $\alpha = [\alpha_1, \dots, \alpha_N]^T$ os parâmetros e os subproblemas $P_1[\alpha_1], \dots, P_N[\alpha_N]$. A decomposição deve ser tal que gere uma família de subproblemas, onde o problema global está contido, no sentido de que existe um parâmetro α^* para o qual a solução de $P_1[\alpha^*], \dots, P_N[\alpha_N^*]$ nos fornece a solução do problema global. Este conceito é definido em [Takahara e Mesarovic \(1969\)](#) e [Gromov \(2010\)](#).

Com o parâmetro α^* , que é a variável de coordenação, é inicialmente desconhecido, o método que será discutido propõe um procedimento iterativo para a sua obtenção, que de uma maneira geral pode ser representado pelo algoritmo conceitual apresentado na Figura 3.

Figura 3 – Estrutura em dois níveis.



Fonte: Elaborada pelo autor.

A estrutura em dois níveis tem a seguinte configuração:

- Passo 1 – O nível superior (coordenador) envia o conjunto de parâmetros α_i^v , $i = 1, \dots, N$ aos subsistemas que compõe o nível inferior, onde v é o passo de iteração.
- Passo 2 - No nível inferior, a família de subproblemas $P_i[\alpha_i^v]$ é resolvida independentemente pelos subsistemas.
- Passo 3 – Os subsistemas enviam ao coordenador alguma função φ_i^v de suas soluções $i = 1, \dots, N$.
- Passo 4 – Se os valores de α_i^v foram obtidos, interrompe-se o algoritmo. A solução ótima para o problema foi obtida. Caso contrário seguir para o passo 5.
- Passo 5 – O coordenador calcula um novo valor para α^{v+1} baseados nos valores de φ_i^v e α_i^v , $i = 1, \dots, N$. Seguir para o passo 1.

2.8 Método de Decomposição por Coordenação via Predição de Interação

Seja o problema de otimização na forma primal-dual. O lagrangiano para o problema é dado pela seguinte expressão:

$$L(x_k, u_k^i, \beta_k^i, z_k^i) = \sum_{i=1}^N \left\{ \sum_{k=0}^{K-1} \left(f^i(x_k^j, u_k^j) + (\beta_k^i)^T z_k^i \right. \right. \\ \left. \left. - \sum_{j \neq i}^N g^{ji}(x_k^j, u_k^j) + \lambda_{k+1}^i (-x_{k+1}^i + g^i(x_k^j, u_k^j, z_k^i)) \right) \right\} \quad (2.74)$$

ou

$$L(.) = \sum_{i=1}^N L^i(.)$$

Os subproblemas do lagrangiano são descritos como

$$D(\beta_k) = \min_{x^i, u^i} \sum_{k=0}^{K-1} \left\{ f^i(x_k^j, u_k^j) + (\beta_k^i)^T z_k^i + \sum_{j \neq i}^N (\beta_k^i)^T g^{ji}(x_k^j, u_k^j) \right\}, \quad (2.75)$$

sujeito a

$$x_{k+1}^i = g^i(x_k^j, u_k^j, z_k^i), \quad (2.76)$$

$$x_0^i = x^{i0}, i = 1, \dots, N \quad (2.77)$$

e o problema dual é representado como

$$\max_{\beta_k} D(\beta_k). \quad (2.78)$$

O problema global foi decomposto em uma família de N subproblemas independentes. A razão é que desde que a relação de interconexão proporciona o único acoplamento entre os subsistemas e já foi incluída no lagrangiano, a propriedade de separabilidade aditiva, para trajetórias de β_k^i , z_k^i , que são os parâmetros de coordenação, fornecidos pelo coordenador, justifica a validade da decomposição. O nível inferior resolve os subproblemas do lagrangiano dado pelas equações (2.75) a (2.77) enquanto que o coordenador resolve o problema de maximização da função dual, dada por (2.78), possibilitando uma divisão de carga computacional entre os dois níveis.

2.8.1 Solução dos Subproblemas de Otimização Independentes

As condições de otimalidade para os subproblemas definidos pelas equações (2.75) a (2.77) são obtidas pela aplicação de técnicas de programação dinâmica. Assim, define-se o seguinte hamiltoniano:

$$H(\cdot) = f_i(x_k^j, u_k^j) + (\beta_k^i)^T z_k^i - \sum_{j \neq i}^N (\beta_k^i)^T g^{ji}(x_k^j, u_k^j) + \lambda_{k+1}^i g^i(x_k^j, u_k^j, z_k^i). \quad (2.79)$$

As condições necessárias de primeira ordem, que são as equações a serem resolvidas pelos subsistemas do nível inferior são dadas por

$$\frac{\partial H^i}{\partial \lambda_{k+1}^i} = x_{k+1}^i = g^i(x_k^i, u_k^i, z_k^i), \quad (2.80)$$

$$\frac{\partial H^i}{\partial x_k^i} = \lambda_k^i = \left(\frac{\partial f^i}{\partial x_k^i} \right)^T - \sum_{j \neq i}^N \left(\frac{\partial g^{ji}(x_k^i, u_k^i)}{\partial x_k^i} \right)^T \beta_k^j + \left(\frac{\partial g^{ji}(x_k^i, u_k^i, z_k^i)}{\partial x_k^i} \right)^T \lambda_{k+1}^i, \quad (2.81)$$

$$\frac{\partial H^i}{\partial u_k^i} = 0 \therefore \left(\frac{\partial f^i}{\partial u_k^i} \right)^T + \left(\frac{\partial g^i(\cdot)}{\partial u_k^i} \right)^T \lambda_{k+1}^i - \sum_{j \neq i}^N \left(\frac{\partial g^{ji}(\cdot)}{\partial u_k^i} \right)^T \beta_k^j = 0, \quad (2.82)$$

com as seguintes condições de contorno: $x_0^i = x^0$ e $\lambda_k^i = 0$.

A determinação das trajetórias ótimas $(x_k^i)^*$, $(u_k^i)^*$ e $(\lambda_k^i)^*$ provém da solução do TPBVP (*Two Point Boundary Values Problem*) (TAKABA, 1996), descrito pelas equações (2.80) a (2.82).

2.8.2 Solução do Problema de Coordenação

Uma vez obtidas as soluções ótimas para os subproblemas no nível inferior, o problema torna-se um melhoramento de z_k^i e β_k^i de maneira tal que a função dual seja maximizada.

As condições necessárias de primeira ordem para a maximização são as seguintes:

$$\frac{\partial L(.)}{\partial z_k^i} = 0 \quad \text{e} \quad \frac{\partial L(.)}{\partial \beta_k^i} = 0, \quad (2.83)$$

de onde se obtém

$$\beta_k^i = - \left(\frac{\partial g^i(.)}{\partial z_k^i} \right)^T \lambda_{k+1}^i, \quad (2.84)$$

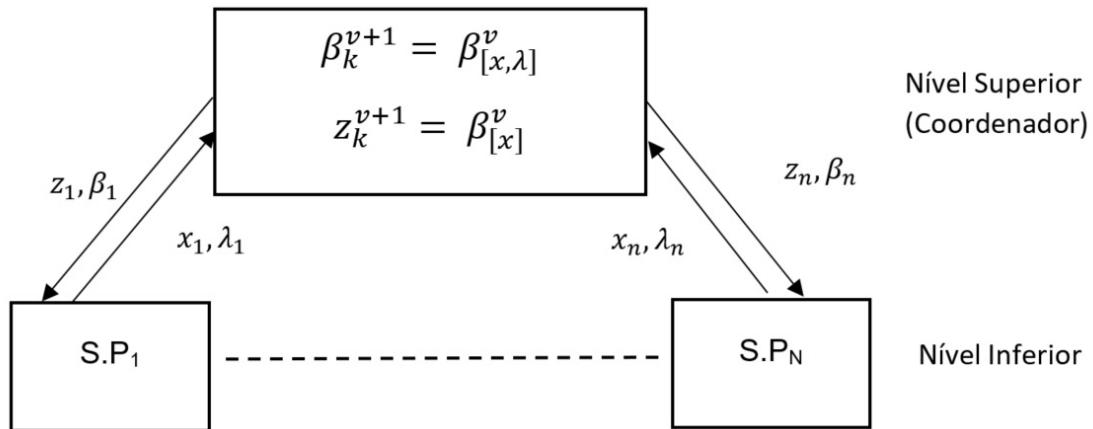
$$z_k^i = \sum_{j \neq i}^N g^{ij}(x_k^j, u_k^j). \quad (2.85)$$

Estas equações que fornecem uma relação explícita das variáveis ou parâmetros de coordenação em função das variáveis do nível inferior. A regra de coordenação, usando o princípio de predição de interação, do passo de iteração v para $(v+1)$, é dada por

$$\begin{bmatrix} \beta_k^i \\ z_k^i \end{bmatrix}^{v+1} = \begin{bmatrix} - \left(\frac{\partial g^i(.)}{\partial z_k^i} \right)^T \lambda_{k+1}^i \\ \sum_{j \neq i}^N g^{ij}(x_k^j, u_k^j) \end{bmatrix}^v. \quad (2.86)$$

Os valores de x_k^i , u_k^i e λ_{k+1}^i que são obtidos das minimizações independentes parametrizados em z_k^i e β_k^i , são substituídos nas equações (2.84) e (2.85) para a obtenção de z_k^i e β_k^i , $i = 1, \dots, N$, os quais são usados subsequentemente como uma nova predição de z_k^i e β_k^i . Este método foi proposto por Takahara e Mesarovic (1969) que provaram sua convergência.

Figura 4 – Estrutura hierárquica para o método de coordenação por predição de interação.



Fonte: Elaborada pelo autor.

A Figura 4 representa a estrutura com dois níveis e troca de informações entre nível superior e inferior.

2.9 Comentário Final

Neste capítulo, foram analizados os aspectos mais importantes que serão empregados na estruturação da nova metodologia para filtro de Kalman distribuído, referentes a sistemas dinâmicos lineares em tempo discreto, dualidade, decomposição de filtros distribuídos e estrutura de paralelização.

No capítulo a seguir, será feita uma descrição detalhada do novo método para a solução do problema de filtragem distribuída aplicando dualidade, com decomposição paramétrica via predição por interação, e será apresentada uma análise comparativa entre o método proposto e os métodos convencionais para filtros distribuídos baseados em consenso.

3 NOVA METODOLOGIA PARA FILTRO DE KALMAN DISTRIBUÍDO

Este capítulo tem como objetivo propor uma abordagem alternativa para o problema de filtragem distribuída, explorando a dualidade entre o filtro e o controle, via o método de predição por interação. A elaboração do algoritmo distribuído permitirá soluções escaláveis que amenizam a relação de dependência das equações do filtro e diminuem o tempo de execução. No cenário analisado, desconsiderar-se-á o tempo de comunicação e sincronismo.

Ao final do capítulo, serão apresentadas simulações de três provas de conceitos, na forma de exemplos didáticos para ilustrar as principais vantagens do método proposto, como por exemplo, consistência algébrica (boas propriedades de estabilidade numérica e matrizes bem condicionadas) e também analisar o seu desempenho computacional.

3.1 Decomposição do Problema

Vamos considerar que o sistema S pode ser decomposto em N subsistemas interconectados S^i pela partição do vetor de estado x e do vetor de medida y , para $i = 1, \dots, N$. Assim, tem-se que

$$S^i = \begin{cases} x_{k+1}^i = A_k^i x_k^i + B_k^i w_k^i + z_k^i \\ y_k^i = C_k^i x_k^i + D_k^i v_k^i \end{cases} \quad (3.1)$$

onde $x_k^i \in \mathbb{R}^{n_i}$ é o vetor de estado i -ésimo subsistema, tal que $\mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \dots \times \mathbb{R}^{n_i}$. $y_k^i \in \mathbb{R}^{m_i}$ é o vetor de medida i -ésimo subsistema, tal que $\mathbb{R}^m = \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \dots \times \mathbb{R}^{m_i}$. Definimos z_k^i como a função que representa a interconexão do i -ésimo subsistema dentro do sistema global S . A divisão do sistema S em N subsistemas pode ser realizada de forma matemática.

Assim, para garantir um desempenho satisfatório do Sistema S vamos supor que a função custo seja aditivamente separável e representado por

$$\min J = \sum_{i=1}^N \min J^i. \quad (3.2)$$

Nas próximas seções, serão apresentadas a formulação do problema de filtragem e a aplicação do método predição por interação para partição do algoritmo.

3.2 Formulação do Problema de Otimização Dinâmica

A partir da notação do problema de filtragem adotado no capítulo 2, considere a seguinte equação da função custo do filtro de Kalman global:

$$J(x_0, w_k, v_k) \triangleq \|x_0 - \bar{x}_0\|_{\Pi^{-1}}^2 + \sum_{k=0}^n (\|w_k\|^2 + \|v_k\|^2). \quad (3.3)$$

Continuando a formulação do problema do filtro de Kalman, com a utilização da equação de medida, que relaciona o vetor de medida, y_k , ao ruído de medida, v_k , obtém-se

$$v_k = D_k^{-1}(y_k - C_k x_k). \quad (3.4)$$

Desenvolvendo a norma quadrática, $\|v_k\|^2$, tem-se

$$\begin{aligned} \|v_k\|^2 &= v_k^T v_k = (y_k - C_k x_k)^T \underbrace{(D_k^T D_k)^{-1}}_{R_k^i} (y_k - C_k x_k), \\ &= (y_k - C_k x_k)^T (R_k)^{-1} (y_k - C_k x_k), \\ &= \|y_k - C_k x_k\|_{R_k^{-1}}^2. \end{aligned} \quad (3.5)$$

Então, substituindo a equação (3.5) na função de medida podemos reescrever a equação da função custo do filtro da seguinte forma:

$$J_{H_2}(x_0, w_k, y_k) \triangleq \|x_0 - \bar{x}_0\|_{\Pi^{-1}}^2 + \sum_{k=0}^n (\|w_k\|_{R_k}^2 + \|y_k - C_k x_k\|_{R_k^{-1}}^2). \quad (3.6)$$

Assim, o problema de minimização da função custo do filtro H_2 passa a ser determinada em relação a (x_0, w_k, y_k) e não mais em (x_0, w_k, v_k) . Para minimizar J_{H_2} em relação a x_0 e w_k , forma-se o seguinte hamiltoniano:

$$H_k = \|x_0 - \bar{x}_0\|_{\Pi^{-1}}^2 + \sum_{k=0}^n (\|w_k\|_{R_k}^2 + \|y_k - C_k x_k\|_{R_k^{-1}}^2) + \lambda_{k+1}^T (A_k x_k + B_k w_k). \quad (3.7)$$

Deste modo, pode-se verificar a seguir as condições necessárias de otimalidade.

Condição de estado:

$$\frac{\partial H_k}{\partial \lambda_{k+1}} = x_{k+1}^* \quad \therefore \quad x_{k+1}^* = A_k x_k^* + B_k w_k^* \quad (3.8)$$

Condição hamiltoniano:

$$\frac{\partial H_k}{\partial w_k} = 0 \quad \therefore \quad \frac{\partial}{\partial w_k} \left\{ \frac{1}{2} w_k^T R_k w_k + \lambda_{k+1}^T B_k w_k \right\} = 0 \quad \therefore \quad w_k^* = -R_k^{-1} B_k^T \lambda_{k+1} \quad (3.9)$$

Condição co-estado:

$$\begin{aligned} \frac{\partial H_k}{\partial x_k} = \lambda_k^* \quad \therefore \quad \lambda_k^* &= \frac{\partial}{\partial x_k} \left\{ -\frac{1}{2} (y_k^T - x_k^T C_k^T) (R_k)^{-1} (y_k - C_k x_k) + \lambda_{k+1}^T A_k x_k \right\} \\ &\therefore \quad \lambda_k = C_k^T R_k^{-1} C_k x_k + A_k \lambda_{k+1} \end{aligned} \quad (3.10)$$

Condição contorno:

$$\frac{\partial H_k}{\partial w_k} = \lambda_0^* \quad \therefore \quad \lambda_0^* = \frac{\partial}{\partial x_0} \left\{ -\frac{1}{2} (x_0^T - \bar{x}_0^T) \Pi^{-1} (x_0 - \bar{x}_0) \right\} \quad (3.11)$$

Visto que o problema é convexo, estas condições são necessárias e suficientes. Assim, devido à inexistência de procedimentos de recorrência, os cálculos de w_k e x_k em (3.9) e (3.10), dependem somente de λ_{k+1} . Assim,

$$w_k^* = -R_k^{-1} B_k^T \lambda_{k+1}, \quad (3.12)$$

$$x_k = Q_k^{-1} (\lambda_k^* - A_k^T \lambda_k), \quad (3.13)$$

$$x_N = Q_k^{-1} \lambda_N^*. \quad (3.14)$$

3.3 Particularidade da Formulação do Problema de Filtragem

Considerando que \hat{x}_k seja uma estimativa de x_k , baseado na medição de $\{y_0, \dots, y_k\}$, podemos reescrever a equação (3.5) como

$$v_k = D_k^{-1} C_k (x_k - \hat{x}_k). \quad (3.15)$$

Definindo $\tilde{x}_k = x_k - \hat{x}_k$ como sendo o vetor dos resíduos entre os estados reais e os estimados.

Logo, podemos reescrever a função custo (3.6) do filtro como

$$J(\tilde{x}_0, w_k, y_k) \triangleq \|x_0 - \bar{x}_0\|_{\Pi^{-1}}^2 + \sum_{k=0}^n \|\tilde{x}_k\|_{Q_k}^2 + \sum_{k=0}^n \|w_k\|_{R_k}^2, \quad (3.16)$$

onde $Q_k = C_k^T R_k^{-1} C_k$.

Empregando os resultados descritos pelas tabelas (3) e (4), podemos redefinir $Q_k = B_k R_k^{-1} B_k^T$ e a função lagrangiana do sistema como

$$L(\tilde{x}_k, w_k, \lambda_{k+1}, k) = \|x_0 - \bar{x}_0\|_{\Pi^{-1}}^2 + \sum_{k=0}^n \|\tilde{x}_k\|_{Q_k}^2 + \sum_{k=0}^n \|w_k\|_{R_k}^2 + \lambda_{k+1}^T (A_k^T \tilde{x}_k + C_k^T w_k - \tilde{x}_{k+1}). \quad (3.17)$$

3.4 Filtro de Kalman Distribuído: Método Predição por Interação

Considere a descrição do sistema dinâmico linear e discreto exposto nas seguintes equações:

$$x_{k+1}^i = A_k^i x_k^i + B_k^i w_k^i + z_k^i, \quad (3.18)$$

$$y_k^i = C_k^i x_k^i + D_k^i v_k^i, \quad (3.19)$$

onde

$$z_k^i = \sum_{j \neq i}^N A_k^{ij} x_k^j. \quad (3.20)$$

Vamos, primeiramente, introduzir um conjunto de multiplicadores Lagrange, β_k^i , de dimensão n e o vetor adjunto, λ_{k+1}^i , de dimensão n para acrescentar a restrição de igualdade (3.20) e a restrição dinâmica (3.18) à função custo (3.16). O hamiltoniano para o subsistema apresentado nas equações (3.18) e (3.19) é definido por

$$H_k^i = \|x_0 - \bar{x}_0\|_{\Pi^{-1}}^2 + \|\tilde{x}_k^i\|_{(R_k^i)^{-1}}^2 + \|w_k^i\|_{R_k^i}^2 + (\beta_k^i)^T z_k^i - \sum_{j \neq i}^N [(\beta_k^j)^T A_k^{ij} x_k^j] + (\lambda_{k+1}^i)^T (A_k^i x_k^i + B_k^i w_k^i + z_k^i). \quad (3.21)$$

O procedimento de solução adotado é resolver as condições necessárias de otimalidade iterativamente numa estrutura em dois níveis. As condições necessárias para otimalidade são descritas a seguir.

Condição de estado:

$$\frac{\partial H_k^i}{\partial \lambda_{k+1}^i} = 0 \quad \therefore \quad x_{k+1}^i = A_k^i x_k^i + B_k^i w_k^i + z_k^i. \quad (3.22)$$

Condição Hamiltoniano:

$$\frac{\partial H_k^i}{\partial w_k^i} = 0 \quad \therefore \quad \frac{\partial}{\partial w_k^i} \left\{ \frac{1}{2} (w_k^i)^T R_k^i w_k^i + (\lambda_{k+1}^i)^T B_k^i w_k^i \right\} = 0 \quad \therefore \quad w_k^i = -(R_k^i)^{-1} (B_k^i)^T \lambda_{k+1}^i \quad (3.23)$$

Condição Co-Estado:

$$\frac{\partial H_k^i}{\partial x_k^i} = \lambda_k^i,$$

$$\lambda_k^i = \frac{\partial}{\partial x_k^i} \left\{ -\frac{1}{2} [(y_k^i)^T - (x_k^i)^T (C_k^i)^T] (R_k^i)^{-1} (y_k^i - C_k^i x_k^i) + (\lambda_{k+1}^i)^T A_k^i x_k^i - \sum_{j \neq i}^N [(\beta_k^j)^T A_k^{ij} x_k^i] \right\},$$

$$\begin{aligned} \lambda_k^i &= \underbrace{(C_k^i)^T (R_k^i)^{-1} C_k^i}_{Q_k^i} x_k^i + A_k^i \lambda_{k+1}^i - \sum_{j \neq i}^N [(A_k^{ij})^T \beta_k^j], \\ &= (Q_k^i) x_k^i + A_k^i \lambda_{k+1}^i - \sum_{j \neq i}^N [(A_k^{ij})^T \beta_k^j]. \end{aligned} \quad (3.24)$$

Isolando o termo x_k^i da equação (3.24), encontra-se:

$$x_k^i = (Q_k^i)^{-1} \left\{ \lambda_k^i - A_k^i \lambda_{k+1}^i + \sum_{j \neq i}^N [(A_k^{ij})^T \beta_k^j] \right\}. \quad (3.25)$$

Aplicando a dualidade entre o filtro de Kalman e o filtro de informação, para as equações (3.22), (3.23) e (3.25), tem-se a equivalência mostrada na Tabela (5).

Tabela 5 – Dualidade das restrições: filtro de Kalman e filtro de informação

Variáveis	Filtro de Kalman	Filtro de Informação
x_{k+1}^i	$A_k^i x_k^i + B_k^i w_k^i + z_k^i$	$(A_k^i)^T x_k^i + (C_k^i)^T w_k^i + z_k^i$
w_k^i	$-(R_k^i)^{-1} (B_k^i)^T \lambda_{k+1}^i$	$-(R_k^i)^{-1} C_k^i \lambda_{k+1}^i$
x_k^i	$(Q_k^i)^{-1} \left\{ \lambda_k^i - A_k^i \lambda_{k+1}^i + \sum_{j \neq i}^N [(A_k^{ij})^T \beta_k^j] \right\}$	$(Q_k^i)^{-1} \left\{ \lambda_k^i - (A_k^i)^T \lambda_{k+1}^i + \sum_{j \neq i}^N [(A_k^{ij})^T \beta_k^j] \right\}$

Fonte: Elaborada pelo autor.

Observe que as equações (3.22) e (3.25) representam um problema de dupla condição de contorno. Estas equações podem ser desacopladas pela definição do vetor adjunto λ_k^i . Explorando a dualidade entre o filtro e o regulador linear quadrático, pode-se definir

$$\lambda_k^i = P_k^i x_k^i + s_k^i, \quad (3.26)$$

onde s_k^i é o termo de malha aberta e substituindo a equação (3.26) na equação (3.24), chega-se ao seguinte resultado:

$$P_k^i x_k^i + s_k^i = Q_k^i x_k^i + A_k^i P_{k+1}^i x_{k+1}^i + A_k^i s_{k+1}^i - \sum_{j \neq i}^N [A_k^{ij} \beta_k^j]. \quad (3.27)$$

O próximo passo é substituir a equação (3.26) na equação (3.22), assim

$$\begin{aligned}
x_{k+1}^i &= (A_k^i)^T x_k^i + (C_k^i)^T w_k^i + z_k^i, \\
&= (A_k^i)^T x_k^i + (C_k^i)^T \underbrace{[-(R_k^i)^{-1} C_k^i \lambda_{k+1}^i]}_{w_k^i} + z_k^i, \\
&= (A_k^i)^T x_k^i - (C_k^i)^T (R_k^i)^{-1} C_k^i \underbrace{[P_{k+1}^i x_{k+1}^i + s_{k+1}^i]}_{\lambda_{k+1}^i} + z_k^i, \\
&= (A_k^i)^T x_k^i - (C_k^i)^T (R_k^i)^{-1} C_k^i P_{k+1}^i x_{k+1}^i - (C_k^i)^T (R_k^i)^{-1} C_k^i s_{k+1}^i + z_k^i, \\
&= \underbrace{[I_k^i + (C_k^i)^T (R_k^i)^{-1} C_k^i P_{k+1}^i]}_{\Psi_k^i}^{-1} [(A_k^i)^T x_k^i - (C_k^i)^T (R_k^i)^{-1} C_k^i s_{k+1}^i + z_k^i], \\
&= \Psi_k^i [(A_k^i)^T x_k^i - (C_k^i)^T (R_k^i)^{-1} C_k^i s_{k+1}^i + z_k^i], \tag{3.28}
\end{aligned}$$

onde

$$\Psi_k^i = [I_k^i + (C_k^i)^T (R_k^i)^{-1} C_k^i P_{k+1}^i]^{-1}. \tag{3.29}$$

Observa-se que há um grande custo computacional para encontrar o resultado da equação (3.29). Assim, para amenizar o efeito deste cálculo aplica-se a generalização da fórmula de Sherman-Morrison-Woodbury (XU, 2017) e chega-se ao seguinte resultado:

$$\begin{aligned}
\Psi_k^i &= [I_k^i + (C_k^i)^T (R_k^i)^{-1} C_k^i P_{k+1}^i]^{-1} \quad \therefore \quad [P_{k+1}^i (P_{k+1}^i)^{-1} + (C_k^i)^T (R_k^i)^{-1} C_k^i P_{k+1}^i]^{-1}, \\
&= [(P_{k+1}^i)^{-1} + (C_k^i)^T (R_k^i)^{-1} C_k^i P_{k+1}^i] P_{k+1}^i]^{-1} \quad \therefore \quad (P_{k+1}^i)^{-1} [P_{k+1}^i]^{-1} + (C_k^i)^T (R_k^i)^{-1} C_k^i]^{-1}, \\
&= (P_{k+1}^i)^{-1} [P_{k+1}^i - P_{k+1}^i (C_k^i)^T [C_k^i P_{k+1}^i (C_k^i)^T + R_k^i]^{-1} C_k^i P_{k+1}^i]^{-1}, \\
&= I_k^i - (C_k^i)^T [R_k^i + C_k^i P_{k+1}^i (C_k^i)^T]^{-1} C_k^i P_{k+1}^i. \tag{3.30}
\end{aligned}$$

Assim, eliminando x_{k+1}^i , por substituir a equação (3.28) na equação (3.27), obtém-se

$$P_k^i x_k^i + s_k^i = Q_k^i x_k^i + A_k^i P_{k+1}^i \underbrace{\Psi_{k+1}^i [(A_k^i)^T x_k^i - (C_k^i)^T (R_k^i)^{-1} C_k^i s_{k+1}^i + z_k^i]}_{x_{k+1}^i} + A_k^i s_{k+1}^i - \sum_{j \neq i}^N [A_k^{ij} \beta_k^j],$$

$$\begin{aligned}
&= \underbrace{[Q_k^i + A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T] x_k^i}_{P_k^i} \\
&\quad + \underbrace{[-A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T (C_k^i)^T (R_k^i)^{-1} C_k^i s_{k+1}^i + A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T z_k^i + A_k^i s_{k+1}^i - \sum_{j \neq i}^N [A_k^{ij} \beta_k^j]]}_{s_k^i}.
\end{aligned}$$

Tem-se então

$$P_k^i = Q_k^i + A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T, \quad (3.31)$$

$$\begin{aligned}
s_k^i &= -A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T (C_k^i)^T (R_k^i)^{-1} C_k^i s_{k+1}^i \\
&\quad + A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T z_k^i + A_k^i s_{k+1}^i - \sum_{j \neq i}^N [A_k^{ij} \beta_k^j],
\end{aligned} \quad (3.32)$$

onde P_k^i é a matriz $N \times N$ simétrica solução da equação discreta de Riccati. As condições finais para as equações (3.31) e (3.32) podem ser determinada comparando a equação (3.26) para $k = T$ com a equação (3.27). Deste modo,

$$P_T^i = \bar{P}^i, \quad (3.33)$$

$$s_T^i = 0. \quad (3.34)$$

Seguindo esta formulação, define-se

$$w_k^i = \Theta_{k+1}^i (A_k^i)^T x_{k+1}^i - [\Theta_{k+1}^i - (R_k^i)^{-1} C_k^i] s_{k+1}^i - \Theta_{k+1}^i z_k^i, \quad (3.35)$$

onde

$$\Theta_{k+1}^i = [R_k^i + C_k^i P_{k+1}^i (C_k^i)^T]^{-1} C_k^i P_{k+1}^i. \quad (3.36)$$

A escolha da estratégia de otimização não pode atingir um desempenho ótimo a menos que todos os acoplamentos estejam ausentes ($z_k^i = 0, 1, \dots, N$). A solução deste problema está na formação de um problema de segundo nível que essencialmente prediz novos vetores de coordenação z_k e β_k . Para esta proposta, a função lagrangiana $L(x_k, w_k, z_k, \beta_k, \lambda_{k+1})$ torna-se aditivamente separável

$$L[\cdot] = \sum_{i=1}^N L^i[\cdot], \quad (3.37)$$

onde

$$L^i[\cdot] = \|x_0^i - \bar{x}_0^i\|_{\Pi^{-1}}^2 + \sum_{k=0}^{t-1} \|\tilde{x}_k^i\|_{Q_k^i}^2 + \sum_{k=0}^{t-1} \|w_k^i\|_{R_k^i}^2 + (\beta_k^i)^T z_k^i - \sum_{j \neq i}^N [(\beta_k^j)^T (A_k^{ij})^T x_k^i] \\ + (\lambda_{k+1}^i)^T [-\tilde{x}_{k+1}^i + (A_k^i)^T \tilde{x}_k^i + (C_k^i)^T w_k^i + z_k^i]. \quad (3.38)$$

Os valores de z_k^i e β_k^i podem ser obtidos por

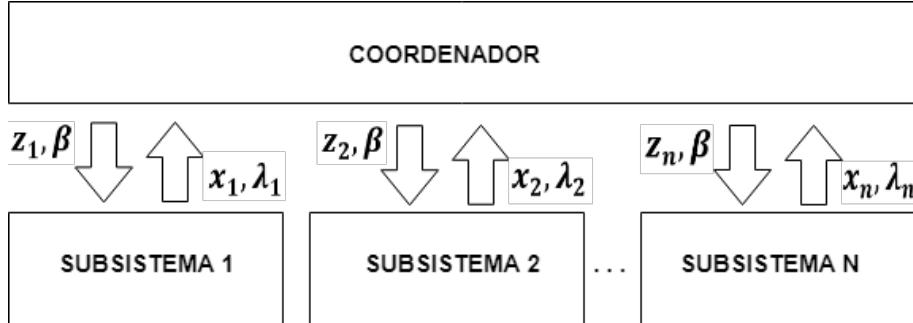
$$\frac{\partial L^i[\cdot]}{\partial \beta_k^i} = 0 \quad \therefore \quad z_k^i - \sum_{j \neq i}^N [(A_k^{ij})^T x_k^i] = 0, \quad (3.39)$$

e

$$\frac{\partial L^i[\cdot]}{\partial z_k^i} = 0 \quad \therefore \quad \beta_k^i + \lambda_{k+1}^i = 0 \quad i = 1, \dots, N. \quad (3.40)$$

Este procedimento de otimização resulta numa estrutura de dois níveis, tal como ilustrado na Figura 5.

Figura 5 – Estrutura em dois níveis do método predição por interação.



Fonte: Elaborada pelo autor.

A seguir, detalha-se o passo a passo do funcionamento da estrutura de dois níveis.

- Passo 1: Inicializar os parâmetros z_0^i e β_0^i , para $k = 0, \dots, t-1$ e enviá-los para o nível 2 (subsistemas). Fazer $i = 1$.
- Passo 2: No nível 2, solucionar os N subsistemas de otimização independentes:
 - Resolver a equação: $P_k^i = Q_k^i + A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T$ com condição final $P_t^i = \Pi^i$
Para obter P_k^i e armazenar, somente uma vez, $i = 1, \dots, N$;
 - Obter s_k^i para condição final $s_k^i = 0$ em $s_k^i = -A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T (C_k^i)^T (R_k^i)^{-1} C_k^i s_{k+1}^i + A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T z_k^i + A_k^i s_{k+1}^i - \sum_{j \neq i}^N [A_k^{ij} \beta_k^i]$;
 - Obter x_k^i em $x_k^i = \Psi_k^i [(A_k^i)^T x_k^i - (C_k^i)^T (R_k^i)^{-1} C_k^i s_{k+1}^i + z_k^i]$;
 - Obter λ_{k+1}^i em $\lambda_{k+1}^i = (Q_k^i) x_k^i + A_k^i \lambda_{k+1}^i - \sum_{j \neq i}^N [(A_k^{ij})^T \beta_k^i]$;
 - Enviar x_k^i e λ_{k+1}^i para Nível 1, $i = 1, \dots, N$;

- Passo 3: No Nível 1, os vetores x_k^i e λ_{k+1}^i são substituídos em $z_k^i = \sum_{j \neq i}^N [(A_k^{ij})^T x_k^i]$ e $\beta_k^i = -\lambda_{k+1}^i$, respectivamente e obter z_k^i e β_k^i .

Este processo de otimização resulta numa estrutura de dois níveis, cujos pseudoscódigo são apresentados a seguir nos Algoritmos 2 e 3.

Algoritmo 2: Nível 2

Entrada: $A_k^i, B_k^i, C_k^i, D_k^i, P_k^i, Q_k^i, R_k^i, x_0^i, z_0^i$ e β_0^i

início

```

para  $i = 1:N$  faça
   $\Psi_k^i = I_k^i - (C_k^i)^T [R_k^i + C_k^i P_{k+1}^i (C_k^i)^T]^{-1} C_k^i P_{k+1}^i$ 
   $P_k^i = Q_k^i + A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T$ 
   $s_k^i = -A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T (C_k^i)^T (R_k^i)^{-1} C_k^i s_{k+1}^i$ 
    +  $A_k^i P_{k+1}^i \Psi_{k+1}^i (A_k^i)^T z_k^i + A_k^i s_{k+1}^i - \sum_{j \neq i}^N [A_k^{ij} \beta_K^j]$ 
   $x_{k+1}^i = \Psi_k^i [(A_k^i)^T x_k^i - (C_k^i)^T (R_k^i)^{-1} C_k^i s_{k+1}^i + z_k^i]$ 
   $\Theta_{k+1}^i = [R_k^i + C_k^i P_{k+1}^i (C_k^i)^T]^{-1} C_k^i P_{k+1}^i$ 
   $w_k^i = \Theta_{k+1}^i (A_k^i)^T x_{k+1}^i - [\Theta_{k+1}^i - (R_k^i)^{-1} C_k^i] s_{k+1}^i - \Theta_{k+1}^i z_k^i$ 
   $\lambda_k^i = (Q_k^i) x_k^i + A_k^i \lambda_{k+1}^i - \sum_{j \neq i}^N [A_k^{ij} \beta_k^j]$ 
fim

```

fim

Algoritmo 3: Nível 1

Entrada: $A_k^i, B_k^i, C_k^i, D_k^i, x_0^i, P_k^i, Q_k^i$ e R_k^i, x_1, λ_1

início

```

   $z_k^i = \sum_{j \neq i}^N A_k^{ij} x_k^j$ 
   $\beta_k^i = -\lambda_{k+1}^i$ 

```

fim

Nos Algoritmos 2 e 3 é possível ter uma visão geral das equações demonstradas no capítulo. Na próxima seção, será apresentada uma descrição do algoritmo proposto.

3.5 Descrição do Algoritmo Proposto

Para ilustrar detalhadamente o algoritmo proposto, vamos considerar, sem perda de generalidade, a decomposição do sistema dinâmico linear em dois subsistemas, conforme descrito pelas seguintes equações:

$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + \begin{bmatrix} B_k^1 \\ B_k^2 \end{bmatrix} \begin{bmatrix} w_k^1 \\ w_k^2 \end{bmatrix}, \quad (3.41)$$

$$\begin{bmatrix} y_k^1 \\ y_k^2 \end{bmatrix} = \begin{bmatrix} C_{11} & 0 \\ 0 & C_{12} \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + D_k \begin{bmatrix} v_k^1 \\ v_k^2 \end{bmatrix}. \quad (3.42)$$

Assim, o sistema linear expresso nas Equações (3.41) e (3.42) é particionado em dois subsistemas. Segue abaixo o Subsistema 1, apresentado no Algoritmo 4.

Algoritmo 4: Subsistema 1

Entrada: $A_k^i, B_k^i, C_k^i, P_k^i, Q_k^i, R_k^i, \beta_k^2, z_k^1$ e yv

$A_k^1 = A_{11}, B_k^1 = B_{11}, C_k^1 = C_{12}, Q_k^1 = Q_{11}, R_k^1 = R_{11}, P_k^1 = P_{11}$ e $x_k^1 = x_{11}$.

início

```

para  $k = 1:N$  faça
   $\Psi_k^1 = I_k - (C_k^1)^T [R_k^1 + C_k^1 P_{k+1}^1 (C_k^1)^T]^{-1} C_k^1 P_{k+1}^1$ 
   $P_k^1 = Q_k^1 + A_k^1 P_{k+1}^1 \Psi_{k+1}^1 (A_k^1)^T$ 
   $K_k^1 = [R_k^1 + C_k^1 P_{k+1}^1 (C_k^1)^T]^{-1} C_k^1 P_{k+1}^1$ 
fim

para  $k = T-1:-1:1$  faça
   $\Gamma^1 = A_{2,1}^T \beta_k^2$ 
   $s_k^1 = - A_k^1 P_{k+1}^1 \Psi_{k+1}^1 (A_k^1)^T (C_k^1)^T (R_k^1)^{-1} C_k^1 s_{k+1}^1$ 
     $+ A_k^1 P_{k+1}^1 \Psi_{k+1}^1 (A_k^1)^T z_k^1 + A_k^1 s_{k+1}^1 - \Gamma^1$ 
fim

para  $k = 1:T-1$  faça
   $x_{k+1}^1 = \Psi_k^1 [(A_k^1)^T x_k^1 - (C_k^1)^T (R_k^1)^{-1} C_k^1 s_{k+1}^1 + z_k^1]$ 
   $x_{k+1}^1 = x_{k+1}^1 + K_k^1 (yv - C_k^1 x_k^1)$ 
fim

para  $k = T-1:-1:1$  faça
   $\lambda_k^1 = (Q_k^1) x_k^1 + A_{1,k}^1 \lambda_{k+1}^i - \Gamma^1$ 
fim
fim

```

O Subsistema 2 é detalhado no Algoritmo 5.

Algoritmo 5: Subsistema 2

Entrada: $A_k^i, B_k^i, C_k^i, P_k^i, Q_k^i, R_k^i, \beta_k^1, z_k^2$ e yv

$A_k^2 = A_{22}, B_k^2 = B_{21}, C_k^2 = C_{12}, Q_k^2 = Q_{22}, R_k^2 = R_{22}, P_k^2 = P_{22}$ e $x_k^2 = x_{21}$.

início

para $k = 1:N$ **faça**

$$\Psi_k^2 = I_k - (C_k^2)^T [R_k^2 + C_k^2 P_{k+1}^2 (C_k^2)^T]^{-1} C_k^2 P_{k+1}^2$$

$$P_k^2 = Q_k^2 + A_k^2 P_{k+1}^2 \Psi_{k+1}^2 (A_k^2)^T$$

$$K_k^2 = [R_k^2 + C_k^2 P_{k+1}^2 (C_k^2)^T]^{-1} C_k^2 P_{k+1}^2$$

fim

para $k = T-1:-1:1$ **faça**

$$\Gamma^2 = A_{1,2}^T \beta_k^1$$

$$s_k^2 = -A_k^2 P_{k+1}^2 \Psi_{k+1}^2 (A_k^2)^T (C_k^2)^T (R_k^2)^{-1} C_k^2 s_{k+1}^2 \\ + A_k^2 P_{k+1}^2 \Psi_{k+1}^2 (A_k^2)^T z_k^2 + A_k^2 s_{k+1}^2 - \Gamma^2$$

fim

para $k = 1:T-1$ **faça**

$$x_{k+1}^2 = \Psi_k^2 [(A_k^2)^T x_k^2 - (C_k^2)^T (R_k^2)^{-1} C_k^2 s_{k+1}^2 + z_k^2]$$

$$x_{k+1}^2 = x_{k+1}^2 + K_k^2 (yv - C_k^2 x_k^2)$$

fim

para $k = T-1:-1:1$ **faça**

$$\lambda_k^2 = (Q_k^2) x_k^2 + A_k^2 \lambda_{k+1}^2 - \Gamma^2$$

fim

fim

O coordenador do sistema paralelizado pode ser definido no Algoritmo 6.

Algoritmo 6: Coordenador

Entrada: $A_k, x_k^1, x_k^2, \lambda^1$ e λ^2

início

$$z_k^1 = A_{12} x_k^2;$$

$$z_k^2 = A_{21} x_k^1;$$

$$\beta_k^1 = -\lambda_{k+1}^1$$

$$\beta_k^2 = -\lambda_{k+1}^2$$

fim

Na próxima subseção, serão apresentadas algumas estratégias de programação distribuída para aplicação do algoritmo proposto.

3.6 Estratégias de Programação Distribuída

A elaboração de algoritmos paralelos torna-se tão importante quanto o desenvolvimento de arquiteturas para processamento paralelo. A introdução de paralelismos em projeto de sistema de computação requer o desenvolvimento de algoritmos que devem ser processados eficientemente sobre tais arquiteturas.

Em boa parte dos algoritmos sequenciais há algum grau de paralelismo, e cabe ao programador analisar e avaliar o grau de concorrência resultante. Quando utilizamos sistemas de múltiplos processadores, há um limite para o grau de paralelismo que pode ser obtido com um mínimo de sobrecarga com a comunicação e a sincronização. Visto que o sistema operacional é monoprocessador e a linguagem empregada é inherentemente sequencial. Assim, não é possível implementar com eficiência o paralelismo de pequenas operações de rápida execução devido à existência de muitos pontos de sincronização e uma grande sobrecarga computacional resultante.

Uma vez que a implementação e o desempenho computacional de algoritmos paralelos dependem significativamente do sistema de computação, há portanto, necessidade de especificá-lo quando consideramos o grau de paralelismo que pode ser obtido e eficientemente implementado. Para a estratégia de paralelização de algoritmos sequenciais é conveniente considerarmos os seguintes problemas:

- Problema de Partição;
- Problema de Distribuição;
- Problema de Comunicação e de Sincronização;

Estes problemas são discutidos nas próximas seções.

3.6.1 *Problema de Partição*

O critério para a escolha da partição depende dos seguintes elementos:

1. ***Dimensão do problema:*** particionar um problema de grande dimensão em subsistemas de dimensões menores que podem ser resolvidos em paralelo, levando em consideração a dependência de dados entre os subsistemas, não é fácil. A tarefas relativas aos subsistemas refletem quando ou não o algoritmo paralelizado apresenta uma comunicação mais intensa. Neste sentido, maior será a quantidade de interações entre as tarefas.
2. ***Número de unidades de processamento utilizadas para a execução de um algoritmo paralelo:*** para tentarmos particionar um algoritmo num certo número de tarefas que podem

ser executadas em paralelo é importante que saibamos com antecedência o número de unidades de processamento disponíveis.

3. ***Relação de dependência entre as equações presentes no algoritmo:*** quando se realiza a paralelização de um algoritmo sequencial deve-se ter o cuidado de garantir que o algoritmo paralelizado seja equivalente ao algoritmo sequencial no que se refere ao conjunto de equações e aos resultados da solução de uma aplicação. Neste sentido, uma análise das relações de dependência oferece condições de especificarmos um conjunto de equações que podem ser executadas simultaneamente. O ganho resultante na precisão da análise das relações de dependência é importante para a qualidade da paralelização do algoritmo.

3.6.2 Problema de Distribuição

Trata-se do problema de determinar a cada tarefa uma ou mais unidades de processamento e vice-versa para a execução do processamento global.

Considera-se apenas um modo de distribuição das tarefas: o modo estático. Neste modo de distribuição, cada unidade de processamento executará exatamente a mesma tarefa durante o processamento global. Devido à distribuição no modo estático, o algoritmo paralelo não convergirá a menos que todas as unidades de processamento estejam operando. Assim, é importante que a distribuição das tarefas determine um bom balanceamento de carga, isto é, que uma das unidades de processamento em paralelo não fique esperando muito tempo que outros processadores terminem suas atividades para iniciar uma nova execução.

3.6.3 Problema de Comunicação e Sincronismo

Trata-se do problema de assegurar a integridade de dados compartilhados pelas tarefas e especificar uma ordem de execução que garanta resultados corretos.

A frequência com que as tarefas se comunicam e a quantidade de informações que deve ser comunicada influenciam a velocidade com a qual o algoritmo paralelo pode ser executado. As implementações podem ser divididas em duas partes lógicas:

1. ***Cálculo:*** É a parte da tarefa que pode ser executada sem interação com as outras tarefas;
2. ***Ponto de Comunicação:*** É a parte da tarefa que ocorre ao final de cada passo de cálculo, permitindo a troca de dados com outras tarefas;

A análise desses problemas permitirá obter e implementar o paralelismo que pode ser explorado num algoritmo sequencial.

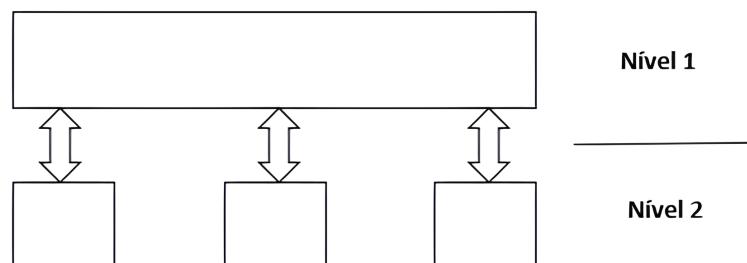
3.7 Metodologias de Otimização e Programação Distribuída

A otimização de sistemas dinâmicos com vetores de estado de grandes dimensões é possível, em princípio, usando métodos bem estabelecidos tais como a Programação Dinâmica e o Princípio de Pontryagin (LEMOS, 2001). A aplicação destes métodos na solução de um problema global de filtragem, envolvendo um grande número de variáveis, pode tornar a carga e o tempo computacional para a otimização excessivos. Isto nos leva a considerar a decomposição do problema de otimização.

Considerando que cada subsistema tem seus próprios vetores de estado e de medida, as soluções dos problemas de otimização dos subsistemas podem ser determinadas em paralelo e combinadas para a obtenção de uma solução ótima para o problema global.

A paralelização hierárquica de multinível é uma técnica de decomposição e coordenação para solucionar problemas de otimização dinâmica de sistemas com grande carga computacional. Este método de paralelização consiste na decomposição de um problema global P , otimizando o critério de desempenho associado ao sistema, em subsistemas independentes P_i ($i = 1, \dots, N$) que se mantêm interligados via uma unidade de coordenação separada, conforme podemos observar na Figura 6.

Figura 6 – Estrutura hierárquica em dois níveis.



Fonte: Elaborada pelo autor.

As principais vantagens do ponto de vista computacional para solucionar problemas de filtragem com paralelização hierárquica multinível envolve uma *redução significativa do tempo de execução computacional*, especialmente se um sistema de múltiplos processadores for utilizado para explorar o paralelismo natural e intrínseco da estrutura multinível. Esta característica é extremamente importante para permitir uma investigação da aplicabilidade dos algoritmos de filtragem ótima multinível. Também se destaca que a *decomposição do problema de filtragem* em subproblemas independentes permite, apesar do aumento do número de variáveis envolvidas no procedimento de otimização, uma distribuição da memória total do cálculo por

várias unidades de processamento de um sistema de múltiplos núcleos ou processadores, uma vez que as unidades de processamento locais têm de resolver somente um problema de dimensão reduzida. Outra vantagem importante é que a *decomposição do problema global* consiste em vários módulos separados, que podem ser implementados e testados individualmente de forma mais conveniente do que o problema global.

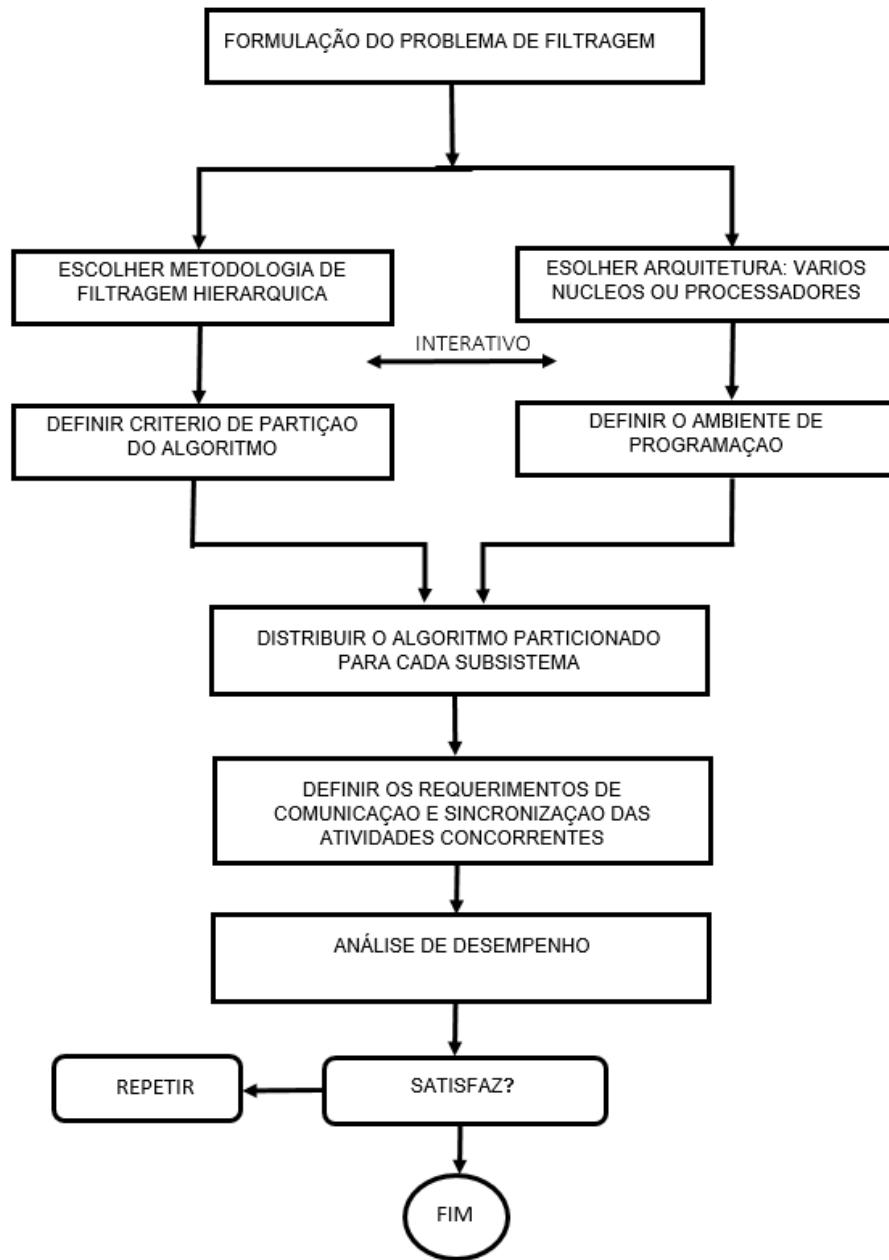
Estas vantagens podem apresentar um maior ou menor grau de utilização eficiente de um sistema de múltiplos núcleos ou processadores quando um método de decomposição e coordenação é aplicado a um problema de filtragem ótimo particular. Os principais fatores que afetam a eficiência da solução do problema de filtragem ótima são a estrutura e a dimensão do problema; a simplicidade dos subproblemas, uma vez que os métodos iterativos envolvem a resolução dos subproblemas várias vezes e as características dos sistemas de múltiplos núcleos ou processadores.

Um desempenho computacional superior pode ser obtido através da investigação do paralelismo subjacente na estrutura do cálculo. Esta atitude implica na modificação do algoritmo para se atingir um maior grau de paralelismo, mantendo uma dependência com a arquitetura de computação utilizada. A partir deste contexto, adota-se o seguinte procedimento para combinar a escolha da metodologia de filtragem hierárquica e a estratégia de paralelização para o processamento sobre sistemas de múltiplos núcleos ou processadores, conforme pode-se observar na Figura 7.

3.8 Paralelização do Algoritmo

A paralelização do algoritmo para processamento paralelo ou distribuído numa arquitetura de múltiplos processadores é determinada pelo critério adotado para a escolha da partição do algoritmo. Para o estudo da partição do algoritmo, definem-se dois critérios. O primeiro critério explora apenas o paralelismo natural resultando da estrutura hierárquica em dois níveis. O segundo critério implica na modificação da estrutura de cálculo do algoritmo para a obtenção de um grau maior de paralelismo. Sob este aspecto determinam-se duas etapas de implementação que são detalhadas nas próximas subseções.

Figura 7 – Procedimento de paralelização e implementação.



Fonte: Elaborada pelo autor.

3.8.1 Primeira Etapa

A partir da decomposição do problema de otimização, algumas características importantes podem ser observadas. Por exemplo, os subsistemas apresentam o mesmo conjunto de equações, ou seja, os subproblemas são semelhantes. Também, há a possibilidade de se ter um número arbitrário de subsistemas de dimensões variadas. Outra particularidade importante é que os subsistemas são assíncronos, ou seja, a execução de cálculos de otimização não envolve interação com outros subproblemas.

Através das características apresentadas, pode-se especificar para os N subsistemas

de otimização do nível inferior da estrutura em dois níveis N tarefas tal que os subsistemas tenham a mesma dimensão. Considerando a disponibilidade de $(N + 1)$ unidades de processamento idênticas será mais eficiente para o processamento paralelo que haja N subsistemas de mesma dimensão. Se isto não ocorrer, o sistema de múltiplos processadores de velocidades de processamento diferentes e por conseguinte uma consideração especial deve ser dada à alocação adicional de partes do algoritmo para os processadores mais rápidos, o que complica consideravelmente a implementação do algoritmo. Sob este aspecto, faz-se necessário especificar as seguintes tarefas:

TAREFA $i, i = 1, \dots, N$

- Coletar z_k^i e β_k^i ;
- Resolver P_k^i , somente uma vez, e s_k^i ;
- Calcular x_k^i e β_k^i .
- Enviar para a Tarefa $(N + 1)$: $\begin{bmatrix} x_k^i \\ \beta_k^i \end{bmatrix}$

TAREFA $j, j = 1, \dots, N$

- Predizer z_k^i e β_k^i iniciais, somente uma vez;
- Enviar z_k^i e β_k^i a cada TAREFA i ;
- Coletar x_k^i e β_k^i ;
- Calcular z_k^i ;
- Testar: $\begin{bmatrix} z_k^{i+1} - z_k^i \\ \beta_k^{i+1} - \beta_k^i \end{bmatrix} < \varepsilon$, onde ε é um valor muito pequeno.
- Realização do cálculo de: β_k^i e λ_{k+1}^i

Assim, é realizado por cada TAREFA $i = 1, \dots, N$, reduzindo o cálculo da tarefa de coordenação.

3.8.2 Segunda Etapa

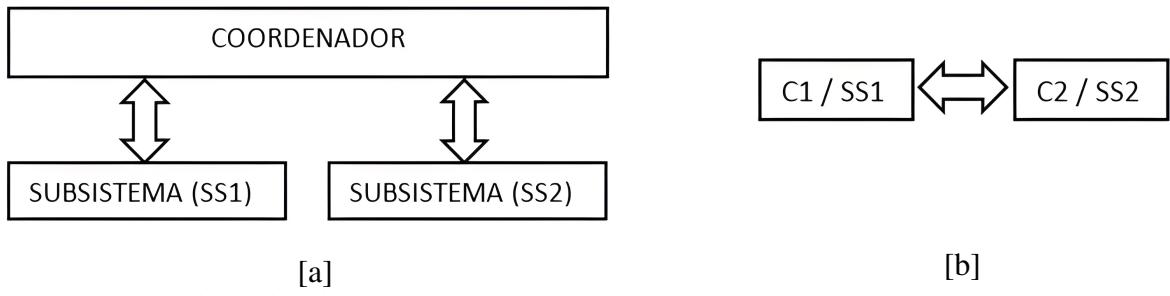
A paralelização do algoritmo obtida a partir do paralelismo natural apresenta um bom desempenho computacional quando comparado com a execução sequencial do algoritmo. Entretanto, pode-se verificar o seguinte aspecto: a tarefa de coordenação tende a limitar a potência de processamento global do algoritmo. A TAREFA $(N + 1)$ manipula variáveis globais, ou seja, $z_k^i \in \mathbb{R}^n$ e $\beta_k^i \in \mathbb{R}^n$ independente da ordem dos subsistemas. Embora o tempo de execução de cada TAREFA i cresça rapidamente com a ordem do subsistema, o tempo de execução da TAREFA $(N + 1)$ praticamente permanece o mesmo.

Em adição, é possível adicionar mais unidades de processamento para cada TAREFA i de forma que o tempo total de execução tenda para o tempo de execução da TAREFA ($N + 1$). Assim, pode-se propor um segundo critério que explora o paralelismo na tarefa de coordenação quanto a partição do algoritmo e distribuição das tarefas.

Particionamento

Esta implementação se caracteriza pela subdivisão da tarefa de coordenação em N subtarefas. O vetor de coordenação é formado pelo vetor de interconexão e pelos multiplicadores de Lagrange. Como não há nenhuma relação explícita entre o cálculo de z_k^i e β_k^i é possível partitionar o vetor de coordenação em N subvetores, ou seja, $\begin{bmatrix} z_k^i \\ \beta_k^i \end{bmatrix}$ tal que: $\sum_{i=1}^N n_i = n$, conforme pode-se observar na Figura 8. O cálculo de $z_k^i (i = 1, \dots, N)$ depende da variável de estado

Figura 8 – Etapas de Partição : [a] Com coordenação e [b] Sem Coordenação



Fonte: Elaborada pelo autor.

$x_k^j (j = 1, \dots, N; j \neq i)$ e o cálculo de s_k^i e de λ_{k+1}^i dependem de β_k^j . Quando associamos os N subvetores de coordenação aos N subsistemas de otimização, as relações de dependência mencionadas determinam o grau de concorrência resultante deste novo critério de partição da estrutura de cálculo.

Através da partição adotada pode-se verificar que o número de tarefas é igual ao número de subsistemas. Estes subsistemas apresentam o mesmo conjunto de equações e estão restritas a um só nível computacional.

3.9 Provas de Conceito

Nesta seção, serão apresentados três exemplos com o intuito de ilustrar a metodologia proposta de paralelização para a solução do problema de filtragem. Nos dois primeiros exemplos analisaremos os resultados comparativos entre a abordagem do filtro de Kalman global (FKG) e o algoritmo distribuído proposto (FKDPI).

No Exemplo 3 é realizada uma análise comparativa, considerando como critério o parâmetro de erro quadrático médio, entre o algoritmo proposto e outros algoritmos distribuídos (ADKF, SDKF, DKF, MKCF, DKBF e IKCF). Os exemplos são simulados com o software Matlab 2017b, na versão estudante.

3.9.1 Exemplo ilustrativo 1

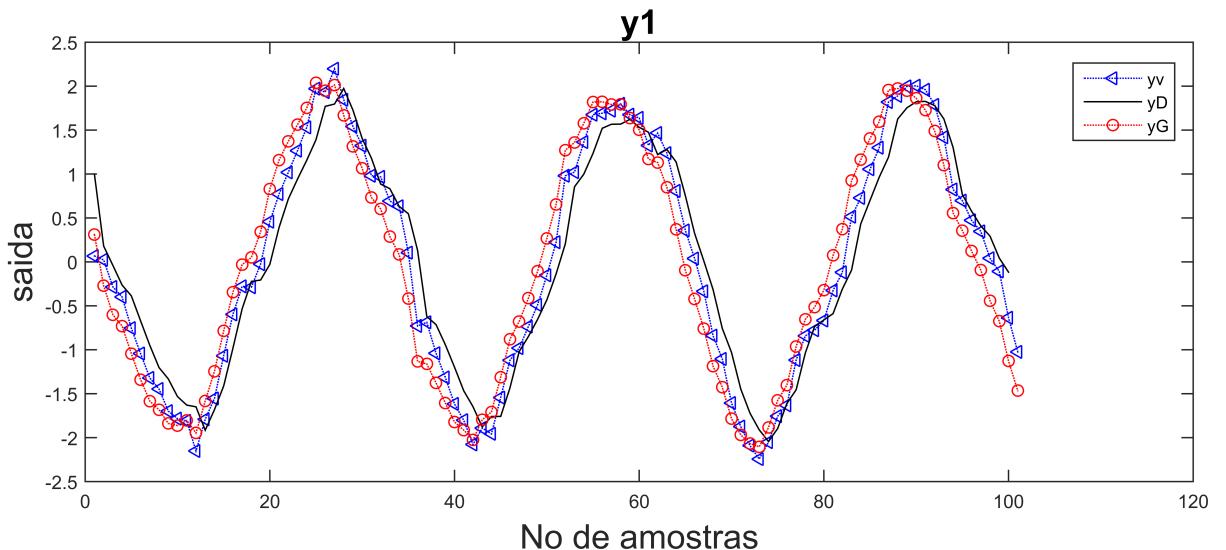
Considera-se inicialmente o seguinte sistema dinâmico:

$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \end{bmatrix} = \begin{bmatrix} 1,1269 & -0,4940 \\ 1,0000 & 0 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + \begin{bmatrix} -0,3832 \\ 0,5919 \end{bmatrix} \begin{bmatrix} w_k^1 \\ w_k^2 \end{bmatrix}, \quad (3.43)$$

$$\begin{bmatrix} y_k^1 \\ y_k^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + \begin{bmatrix} v_k^1 \\ v_k^2 \end{bmatrix}. \quad (3.44)$$

Neste caso, consideramos as matrizes de correlação: $Q = 0,03$ e $R = 0,01$. O ruído de processo e de medição são vistos como gaussianos e com média zero.

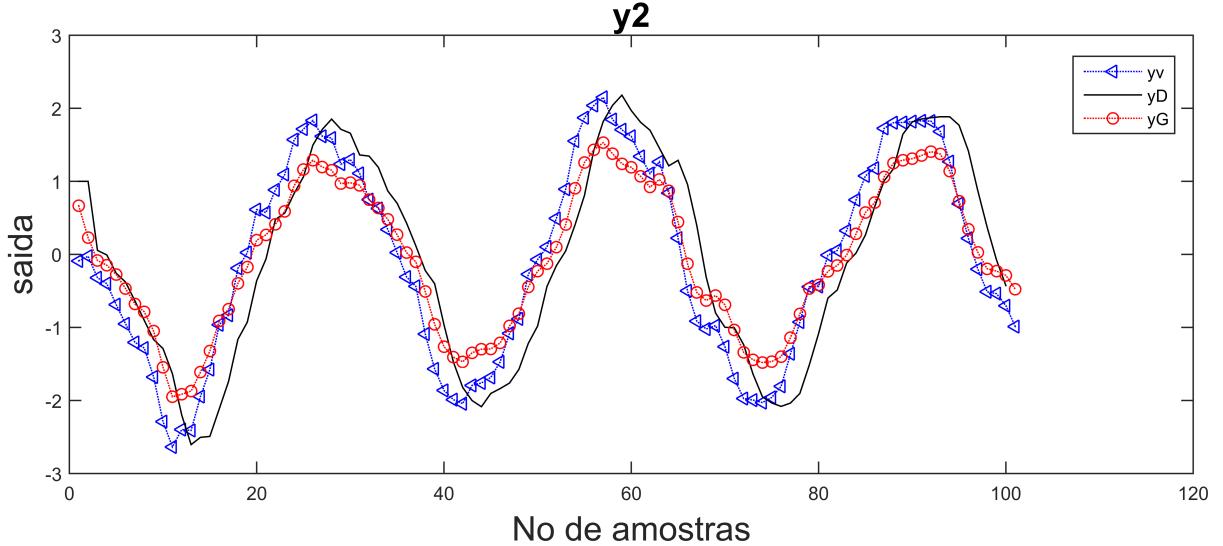
Figura 9 – Resposta do FKG (vermelho) e do método proposto FKDP (preto) para a saída y_k^1 (azul).



Fonte: Elaborada pelo autor.

Nas Figuras 9 e 10 são apresentados um gráfico comparativo entre o erro de estimativa do filtro FKG (vermelho) e FKDP (preto) para as saídas y_k^1 e y_k^2 , respectivamente. Observamos que para y_k^1 o filtro FKG alcançou um erro de estimativa menor, enquanto que para y_k^2 o filtro

Figura 10 – Resposta do FKG (vermelho) e do método proposto FKDPI (preto) para a saída y_k^2 (azul).



Fonte: Elaborada pelo autor.

FKDPI obteve melhor resultado. Segue na Tabela 6, a análise de desempenho algébrico e computacional utilizando o software Matlab com pacote *Parallel Computing*, para simular um ambiente de um sistema paralelo com o objetivo de eliminar o tempo de sincronismo.

Tabela 6 – Análise de desempenho algébrico e computacional

Filtro	Ganho	Autovalores	Tempo de Execução 1(s)	Tempo de Execução 2(s)
FKG	$\begin{bmatrix} 0,6196 \\ 0,0057 \end{bmatrix}$	$\begin{bmatrix} 0,4537 \\ 0,2537 \end{bmatrix}$	0,002073	–
FKDPI	$\begin{bmatrix} 0,8007 \\ 0,0000 \end{bmatrix}$	$\begin{bmatrix} 0,2631 \\ 0,1631 \end{bmatrix}$	0,007037	0,003738

Fonte: Elaborada pelo autor.

A coluna Tempo de Execução 1 apresenta a marcação de tempo sem a utilização do pacote de *Parallel Computing*. Observa-se que particionar um problema de filtragem com dimensão pequena não é computacionalmente vantajoso, apesar de o algoritmo proposto permitir que um conjunto de equações seja executado simultaneamente.

3.9.2 Exemplo ilustrativo 2

Considera-se como segundo exemplo ilustrativo o seguinte sistema dinâmico, retirado de Costa (1988).

$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \\ x_{k+1}^3 \\ x_{k+1}^4 \\ x_{k+1}^5 \\ x_{k+1}^6 \end{bmatrix} = \begin{bmatrix} 0,800 & 0,200 & 0,000 & 0,000 & 0,00 & 0,00 \\ 0,200 & 0,500 & 0,089 & 0,000 & 0,00 & 0,00 \\ 0,000 & 0,089 & -0,250 & 0,250 & 0,00 & 0,00 \\ 0,000 & 0,000 & 0,250 & -0,250 & 0,13 & 0,00 \\ 0,000 & 0,000 & 0,000 & 0,045 & -0,60 & 0,35 \\ 0,000 & 0,000 & 0,000 & 0,000 & 0,35 & -0,35 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ x_k^4 \\ x_k^5 \\ x_k^6 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_k^1 \\ w_k^2 \\ w_k^3 \end{bmatrix},$$

$$\begin{bmatrix} y_k^1 \\ y_k^2 \\ y_k^3 \\ y_k^4 \\ y_k^5 \\ y_k^6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ x_k^4 \\ x_k^5 \\ x_k^6 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} v_k^1 \\ v_k^2 \\ v_k^3 \end{bmatrix}.$$

Tabela 7 – Análise de desempenho algébrico e computacional

Filtro	Ganho	Autovalores	Tempo de Execução 1(s)	Tempo de Execução 2(s)
FKS		$\begin{bmatrix} 0,1030 \\ 0,0476 \\ 0,0130 \\ 0,0283 \\ -0,0687 \\ 0,0451 \end{bmatrix} \begin{bmatrix} 0,7912 \\ 0,4088 \\ -0,7764 \\ 0,4879 \\ 0,0032 \\ -0,1103 \end{bmatrix}$	0,026101	–
FKDPI		$\begin{bmatrix} 0,0765 \\ 0,0332 \\ 0,0115 \\ 0,0160 \\ 0,0416 \\ 0,0226 \end{bmatrix} \begin{bmatrix} 0,8268 \\ 0,8994 \\ -0,4099 \\ -0,5001 \\ 0,0060 \\ -0,1108 \end{bmatrix}$	0,008977	0,006281

Fonte: Elaborada pelo autor.

Neste Exemplo 2, o sistema dinâmico linear foi dividido em três subsistemas. E notase que na tabela 7, que o algoritmo proposto FKDPI apresentou um desempenho computacional satisfatório, conforme descrito nas colunas Tempo de Execução 1 e 2. Isto ocorre quando se

minimiza a relação de dependência das equações do filtro e aumenta a dimensão do problema. Nesta simulação com o software Matlab é desconsiderado o atraso de comunicação e sincronismo.

3.9.3 Exemplo ilustrativo 3

Considerando o terceiro exemplo, o modelo apresentado no artigo *Asymptotically Optimal Consensus-based Distributed Filtering of Continuous-time Systems* (BATTILOTTI *et al.*, 2020), tem-se o seguinte sistema dinâmico, definido pelas seguintes equações:

$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \\ x_{k+1}^3 \\ x_{k+1}^4 \end{bmatrix} = \begin{bmatrix} -0,1 & 0 & 0 & 0 \\ 0,5 & -0,5 & 0 & 0 \\ 1,5 & 0 & -0,2 & 0 \\ -1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ x_k^4 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_k^1 \\ w_k^2 \\ w_k^3 \\ w_k^4 \end{bmatrix}, \quad (3.45)$$

$$\begin{bmatrix} y_k^1 \\ y_k^2 \\ y_k^3 \\ y_k^4 \end{bmatrix} = \begin{bmatrix} C_k^i \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ x_k^4 \end{bmatrix} + \begin{bmatrix} v_k^1 \\ v_k^2 \\ v_k^3 \\ v_k^4 \end{bmatrix}. \quad (3.46)$$

Tabela 8 – Parâmetros iniciais do Exemplo 3.

Parâmetros	Valores
Iterações	100
Intervalo	$t \in [0, 50]$
Passo	$dt = 2 \cdot 10^{-3}$
Matriz de Covariância dos Ruídos de Processo	$Q = q_f I_4$, onde $q_f = 0,3$
Matriz de Covariância dos Ruídos de Medida	$R = r_f I_i$, onde $r_f = 0,6$
ADKF, SDKF, DKF	$\gamma = 100$

Fonte: Elaborada pelo autor.

Consideram-se para este exemplo dois cenários. Seja $N = 5$ com os nós conectados em uma cadeia, ou seja, as bordas são $(i, i+1)$, tem-se assim os seguintes cenários:

- **Cenário 1: Informações completas.** Informações completas. O estado do sistema está disponível para todos os nós, ou seja, $C^i = I_4$.

- **Cenário 2: Observabilidade Local.** A saída está disponível para todos os nós, $C^i = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T$. $(C^i A^T)$ é observável.

O algoritmo DKF pode ser usado apenas quando há observabilidade local. No entanto, em nós com nenhuma medição do ganho, o P^i pode assumir qualquer valor. Os demais parâmetros são definidos na Tabela 8.

Tabela 9 – Análise do erro quadrático médio (MSE) dos algoritmos avaliados.

Cenários		tr(P_∞)	CKBF	ADKF	SDKF	tr(S_L)	DKF	MKCF	DKBF	IKCF	FKDPI
1	mse	0,319	0,310	0,382	0,502	0,405	0,408	7,374	0,553	1,544	0,469
2	mse	0,797	0,742	0,781	1,971	1,912	0,820	2,539	0,898	$>10^3$	0,840

Fonte: Elaborada pelo autor.

Na Tabela 9, nos dois cenários analisados observa-se que o filtro proposto FKDPI apresenta um erro quadrático menor que os algoritmos SDKF, MKCF, DKBF e IKCF. No entanto, os algoritmos CKBF, ADKF e DKF obtiveram um resultado melhor que o algoritmo proposto. Ressalta-se a necessidade de analisar a inclusão de uma formulação algébrica que substitua as equações de Riccati e que permita obter um parâmetro de ajuste do ganho do filtro para aprimorar o resultado do algoritmo proposto.

3.10 Comentário Final

Neste capítulo, foi introduzida uma metodologia alternativa distribuída para se obter uma solução do problema de filtragem. A paralelização permitiu analisar a modificação realizada na estrutura do cálculo para diminuir a relação de dependência das equações do filtro de Kalman tradicional, via Riccati. Foram apresentados exemplos e simulações de cunho didático com o objetivo de comparar o algoritmo proposto, FKDPI, com o filtro de Kalman global e algoritmos alternativos e verificar o seu desempenho computacional e algébrico.

No próximo capítulo será reescrita esta metodologia proposta, sem a utilização das equações de Riccati, com o objetivo de garantir não somente o desempenho e estabilidade, mas também uma flexibilização no sentido de se obter uma família de soluções eficientes através do ajuste de um parâmetro escalar.

4 FILTRO DE KALMAN DISTRIBUÍDO COM HORIZONTE DE TEMPO REDUZIDO

O problema de filtragem ótima discreta envolve uma solução iterativa de uma sequência de sistemas resultantes da discretização de um processo contínuo, representado por equações diferenciais parciais que descrevem seu comportamento dinâmico. De modo geral, intervalos de tempo menores tornam mais precisa a aproximação discreta, porém aumentam o número de variáveis de estado. Para obtenção de soluções eficientes, há a necessidade de explorar estruturas matriciais especiais.

Este capítulo tem como objetivo propor um novo algoritmo de filtragem distribuída ($\text{FKDPI}\mu$), sem utilizar a solução recursiva de Riccati, baseado em equações estáticas mais simples para o cálculo do ganho e para o cálculo da matriz do erro de covariância apresentados no Capítulo 3, nos Algoritmos 4 e 5. Esta metodologia, com consistência algébrica (boas propriedades de estabilidade numérica e matrizes bem condicionadas) subjacentes, possibilita-nos obter uma família de resultados subótimos para cada subsistema, em alguns casos, até melhores que as obtidas via solução recursiva de Riccati, através do ajuste um parâmetro escalar.

Ao final do capítulo, serão apresentados quatro provas de conceitos para ilustrar as principais vantagens do método proposto. No terceiro exemplo é realizada uma análise comparativa, considerando como figura de mérito o erro quadrático médio, entre o algoritmo proposto ($\text{FKDPI}\mu$) e outros algoritmos distribuídos (CKBF, ADKF e DKF).

4.1 Formulação Dual via Programação Quadrática

Considera-se o seguinte problema dual:

$$\begin{aligned}
 & \max_{\lambda_{k+1}} \min_{\tilde{x}_k, w_k} L(\tilde{x}_k, w_k, \lambda_{k+1}, k) \\
 & s.a \quad \lambda_{k+1} \in \mathbb{R}^n \\
 & \quad \tilde{x}_0 = \xi \\
 & \quad x_N = Q_k^{-1} \lambda_N^* \\
 & \quad w_k^* = -R_k^{-1} C_k \lambda_{k+1} \\
 & \quad \tilde{x}_k = Q_k^{-1} (\lambda_k^* - A_k \lambda_{k+1})
 \end{aligned} \tag{4.1}$$

onde $L(\tilde{x}_k, w_k, \lambda_{k+1}, k)$ é definido pela Equação (3.17). A demonstração da equação do modelo dual em (4.1) é apresentada em Facanha (2010).

A solução dual pode ser obtida numericamente por técnicas do tipo gradiente. Entretanto, tal gradiente é computacionalmente dispendioso, pois envolve uma grande quantidade de cálculos. Para reduzir os custos computacionais e simplificarmos o processo, dada a concavidade estrita da função $\nabla L(\lambda_{k+1})$, propõe-se uma reformulação do problema dual dinâmico de modo a se obter uma forma quadrática estática definida nas três proposições apresentadas a seguir.

Proposição 1: Dada a concavidade da função dual e as restrições dadas pela Equação (4.1), o problema dual dinâmico pode ser resolvido através da seguinte forma quadrática estática:

$$\min_{\bar{\lambda}} M(\bar{\lambda}) = \frac{1}{2} \bar{\lambda}^T \tilde{H} \bar{\lambda} + \bar{\lambda}^T b + c \quad (4.2)$$

sujeito a $\bar{\lambda} \in \mathbb{R}^{nN}$,

onde $\tilde{H}_{nN \times nN}$ é uma matriz simétrica por bloco tridiagonal, definida positiva e formada por N^2 blocos de dimensões $n \times n$, b é um vetor de dimensões $nN \times 1$, $c \in \mathbb{R}^n$ e $\bar{\lambda} = [\lambda_1 \cdots \lambda_N]^T \in \mathbb{R}^{nN}$.

Tem-se que

$$H = \begin{bmatrix} -C^T R^{-1} C - Q^{-1} & Q^{-1} A & 0 \\ (Q^{-1} A)^T & V_1 - A^T Q^{-1} A & Q^{-1} A \\ 0 & (Q^{-1} A)^T & V_1 - A^T Q^{-1} A \end{bmatrix}, \quad b = \begin{bmatrix} A^T \tilde{x}_0 \\ 0 \\ 0 \end{bmatrix}, \quad c = \frac{1}{2} \tilde{x}_0^T Q \tilde{x}_0 \quad (4.3)$$

onde $V_1 = -C^T R^{-1} C - Q^{-1}$.

Demonstração: Para obter a forma quadrática estática da Equação (4.2), realiza-se uma substituição das restrições dadas na Equação (4.1) na lagrangiana, $L(\tilde{x}_k, w_k, \lambda_{k+1}, k)$; em seguida, expandimos os termos do somatório para os valores de $k = 0, 1, \dots, N-1$, conforme podemos observar na seguinte equação:

$$\begin{aligned} L(\bar{\lambda}) = & -\frac{1}{2} \bar{\lambda}_N^T Q^{-1} \bar{\lambda}_N + \left[\frac{1}{2} \tilde{x}_0^T Q \tilde{x}_0 - \frac{1}{2} \bar{\lambda}_1^T C_0^T R^{-1} C_0 \bar{\lambda}_1 + \bar{\lambda}_1^T A_0^T X_0 \right] \\ & + \sum_{k=1}^{N-1} \left[-\frac{1}{2} \bar{\lambda}_k^T Q^{-1} \bar{\lambda}_k + \frac{1}{2} \bar{\lambda}_{k+1}^T A_k^T Q^{-1} \bar{\lambda}_k + \frac{1}{2} \bar{\lambda}_k^T Q^{-1} A_k^T \bar{\lambda}_{k+1} \right. \\ & \quad \left. - \frac{1}{2} \bar{\lambda}_{k+1}^T (C_k^T R^{-1} C_k + A_k^T Q^{-1} A_k) \bar{\lambda}_{k+1} \right]. \end{aligned}$$

Agrupando os termos segundo os vetores adjuntos, obtém-se a matriz H definida na Equação (4.3). Portanto, a forma quadrática estática está definida em

$$\max_{\bar{\lambda}} L(\bar{\lambda}) = \frac{1}{2} \bar{\lambda}^T H \bar{\lambda} + \bar{\lambda}^T b + c. \quad (4.4)$$

Nota-se que maximizar $L(\bar{\lambda})$ é equivalente a minimizar $-L(\bar{\lambda})$ na mesma região do espaço. Assim, considera-se $M(\bar{\lambda}) = -L(\bar{\lambda})$ e define a seguinte forma quadrática estática:

$$\min_{\bar{\lambda}} M(\bar{\lambda}) = \frac{1}{2} \bar{\lambda}^T \tilde{H} \bar{\lambda} + \bar{\lambda} \tilde{b} + \tilde{c}, \quad (4.5)$$

sujeito a $\bar{\lambda} \in \mathbb{R}^{nN}$,

onde $\tilde{H} = -H$, $\tilde{b} = -b$ e $\tilde{c} = -c$.

Isto completa a prova. \square

Proposição 2: Para definir uma nova metodologia para síntese do filtro de Kalman, apresenta-se uma forma equivalente à solução do problema de filtragem via equação recursiva de Riccati, explorando a dualidade entre a estimativa e o controle (TODOROV, 2006) (TODOROV, 2008). Considere o problema de filtragem descrito no Capítulo 2, em que existe uma matriz P_N que pode ser obtida a partir da seguinte equação:

$$P_N = A \tilde{H}_{11} A^T + Q, \quad (4.6)$$

onde a função custo (2.3) tem valor mínimo dado por

$$J^*(\tilde{x}_k, w_k, k) = \min_{w_k} J(\tilde{x}_k, w_k, k) = \frac{1}{2} \tilde{x}_0^T P_N \tilde{x}_0. \quad (4.7)$$

Demonstração: Calculando o seguinte valor ótimo de $\bar{\lambda}^*$ dado pela Equação (4.2), tem-se

$$\Delta_{\lambda} M(\bar{\lambda}) = \tilde{H} \bar{\lambda}^* - b = 0 \implies \bar{\lambda}^* = \tilde{H}^{-1} b. \quad (4.8)$$

Organizando os termos, a formulação matricial apresentada na Equação (4.6) pode ser definida como

$$\begin{bmatrix} \lambda_1^* \\ \lambda_2^* \\ \vdots \\ \lambda_N^* \end{bmatrix} = \begin{bmatrix} \tilde{H}_{11} & \tilde{H}_{12} & \cdots & \tilde{H}_{1N} \\ \tilde{H}_{21} & \tilde{H}_{22} & \cdots & \tilde{H}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{H}_{N1} & \tilde{H}_{N2} & \cdots & \tilde{H}_{NN} \end{bmatrix} \begin{bmatrix} A^T \tilde{x}_0 \\ 0_{n \times 1} \\ \vdots \\ 0_{n \times 1} \end{bmatrix},$$

onde $\lambda_1^* = \tilde{H}_{11} A^T \tilde{x}_0, \dots, \lambda_N^* = \tilde{H}_{N1} A^T \tilde{x}_0$.

Considerando $N = 2$, sem perda de generalidade, obtém-se

$$\begin{bmatrix} \lambda_1^* \\ \lambda_2^* \end{bmatrix} = \begin{bmatrix} \tilde{H}_{11} & \tilde{H}_{12} \\ \tilde{H}_{21} & \tilde{H}_{22} \end{bmatrix} \begin{bmatrix} A^T \tilde{x}_0 \\ 0_{n \times 1} \end{bmatrix}. \quad (4.9)$$

Portanto, tem-se a seguinte identidade

$$\begin{bmatrix} \bar{V}_1 & W \\ W^T & \bar{V}_2 \end{bmatrix} \begin{bmatrix} \bar{H}_{11} & \bar{H}_{12} \\ \bar{H}_{21} & \bar{H}_{22} \end{bmatrix} = \begin{bmatrix} I_{n \times n} & 0_{n \times n} \\ 0_{n \times n} & I_{n \times n} \end{bmatrix}. \quad (4.10)$$

A Equação (3.16) da função custo do filtro de Kalman pode ser reescrita para $N = 2, k = 0, 1, 2$. como

$$J(x_k, w_k, k) = \frac{1}{2} x_0^T Q x_0 + \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ w_0 \\ w_1 \end{bmatrix}^T \begin{bmatrix} Q & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & R & 0 \\ 0 & 0 & 0 & R \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ w_0 \\ w_1 \end{bmatrix}. \quad (4.11)$$

Inserindo as restrições da equação (4.1) na equação (4.11), chega-se a

$$J(x_k, w_k, k) = \frac{1}{2} x_0^T Q x_0 + \begin{bmatrix} Q^{-1}(\lambda_1^* - A \lambda_2^*) \\ Q^{-1} \lambda_2^* \\ -R^{-1} C \lambda_1^* \\ -R^{-1} C \lambda_2^* \end{bmatrix}^T \begin{bmatrix} Q & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & R & 0 \\ 0 & 0 & 0 & R \end{bmatrix} \begin{bmatrix} Q^{-1}(\lambda_1^* - A \lambda_2^*) \\ Q^{-1} \lambda_2^* \\ -R^{-1} C \lambda_1^* \\ -R^{-1} C \lambda_2^* \end{bmatrix}. \quad (4.12)$$

Substituindo, os termos $\lambda_1^* = \bar{H}_{11} A^T \tilde{x}_0$ e $\lambda_2^* = \bar{H}_{21} A^T \tilde{x}_0$ na equação (4.12), tem-se

$$\begin{aligned} J^*(\tilde{x}_k, w_k, k) = & \frac{1}{2} \tilde{x}_0^T Q \tilde{x}_0 + \frac{1}{2} \tilde{x}_0^T A \bar{H}_{11} (\bar{V}_1 \bar{H}_{11} - W \bar{H}_{21}) A^T \tilde{x}_0 \\ & + \frac{1}{2} \tilde{x}_0^T A \bar{H}_{21}^T (\bar{V}_2 \bar{H}_{21} + W^T \bar{H}_{11}) A^T \tilde{x}_0. \end{aligned} \quad (4.13)$$

Assim, ao considerar a matriz identidade $(\bar{V}_1 \bar{H}_{11} - W \bar{H}_{21} = I_{n \times n})$ e $(\bar{V}_2 \bar{H}_{21} - W^T \bar{H}_{11} = 0_{n \times n})$ e substitui-los na Equação (4.13), chega-se ao seguinte resultado

$$J^*(\tilde{x}_k, w_k, k) = \frac{1}{2} \tilde{x}_0^T (A \bar{H}_{11} A^T + Q) \tilde{x}_0. \quad (4.14)$$

Então, para $k = 0, 1, \dots, N-1$, a estrutura citada na Equação (4.14) permanece inalterada. Portanto, conclui-se que a partir da Equação (4.7), a matriz fornece o custo mínimo, quando está definida na forma

$$P_N = A \bar{H}_{11} A^T + Q. \quad (4.15)$$

Isto completa a prova. \square

Proposição 3: Considerando as propriedades do filtro de Kalman (GREWAL; ANDREWS, 2014), a equação de erro é definida como

$$\tilde{x}_{k+1} = \underbrace{(A + \hat{F}CA)}_{\Theta} \tilde{x}_k - \hat{F}v_{k+1} + (I + \hat{F}C)w_k. \quad (4.16)$$

É possível definir o ganho \hat{F}^T através da seguinte expressão

$$\hat{F}^T = [(CA)^T]^+ (\Theta^T - A^T), \quad (4.17)$$

onde $\Theta = Q^{-1}(\bar{H}_{11}A^T - A\bar{H}_{21}A^T)$. A operação $[(C^TA)^T]^+$ é definida pela inversa de Moore-Penrose, com $[(CA)^T]^+ = [(CA)(CA)^T]^{-1}(CA)$.

Demonstração: Usando novamente o sistema (4.9), sem perda de generalidade, tem-se

$$\begin{bmatrix} \lambda_1^* \\ \lambda_2^* \end{bmatrix} = \begin{bmatrix} \bar{H}_{11} & \bar{H}_{12} \\ \bar{H}_{21} & \bar{H}_{22} \end{bmatrix} \begin{bmatrix} A^T x_0 \\ 0_{n \times 1} \end{bmatrix}. \quad (4.18)$$

Ou seja, a exemplo de (4.18), temos $\lambda_1^* = \bar{H}_{11}A^T\tilde{x}_0$ e $\lambda_2^* = \bar{H}_{21}A^T\tilde{x}_0$. Portanto, ao considerar a equação de restrição (4.1), obtém-se

$$\begin{aligned} \tilde{x}_1 &= Q^{-1}(\lambda_1^* - A\lambda_2^*), \\ &= Q^{-1}(\bar{H}_{11}A^T - A\bar{H}_{21}A^T)\tilde{x}_0. \end{aligned} \quad (4.19)$$

Ou seja, $x_k = (\Theta)^k x_0$. Portanto, $x_1 = (\Theta)^k x_0$ e $A_{\hat{F}}^T = \Theta$, para que possamos manipular Θ para isolar o ganho \hat{F} e obter a expressão definida na Proposição 3.

Para garantir a existência da matriz $(CA)^T$ que satisfaça a Equação (4.17). Existirá um \hat{F}^T , se somente se

$$U_1^T [(CA)^T \hat{F}^T] = 0_{(n-m) \times n}, \quad (4.20)$$

onde a matriz $U_1 \in \mathbb{R}^{n \times (n-m)}$ é obtida através da decomposição QR de $(CA)^T$, tem-se

$$(CA)^T = \begin{bmatrix} U_0 & U_1 \end{bmatrix} \begin{bmatrix} Z_{m \times m} \\ 0_{(n-m) \times m} \end{bmatrix}. \quad (4.21)$$

Visto que a matriz $U_0 \in \mathbb{R}^{n \times m}$, $U = [U_0 \ U_1]$ é ortogonal e Z é não singular. Então, \hat{F}^T é dado explicitamente por

$$\hat{F}^T = Z^{-1}U_0^T(\Theta^T - A^T). \quad (4.22)$$

A Equação (4.22) implica que \hat{F}^T existe, se e somente se, $\mathfrak{R}\{\hat{F}^T - A^T\} \subset \mathfrak{R}\{(CA)^T\} \equiv \mathfrak{R}\{U_0\}$, onde $\mathfrak{R}\{\cdot\}$ é o espaço *range*. Considerando que a matriz C tem posto completo, isto implica na existência da decomposição em (4.21). Assim, \hat{F} deve satisfazer

$$(CA)^T \hat{F}^T = \Theta^T - A^T. \quad (4.23)$$

Multiplicando as Equações (4.22) e (4.23) por $U^T = [U_0 \ U_1]^T$, obtém-se as seguintes equações:

$$Z\hat{F}^T = U_0^T(\Theta^T - A^T)$$

e

$$0_{(n-m) \times n} = U_1^T(\Theta^T - A^T). \quad (4.24)$$

Pelo qual as Equações (4.20) e (4.21) seguem diretamente.

Isto completa a prova. \square

Uma alternativa para se calcular o ganho \hat{F} é realizar uma alocação de polos, com o objetivo de incluir todos os autovalores do sistema em malha fechada em uma sub-região do plano convexo, através da determinação de uma lei de estimativa conveniente. A localização dos polos do sistema está diretamente relacionada aos índices de desempenho do sistema. Para alocar arbitrariamente todos os polos do sistema é necessário que todas as variáveis de estado possam ser medidas ou observadas com sucesso, ou seja, o sistema deve ser completamente observável. Existem funções¹ do Matlab para calcular este ganho, a partir da informação de $(A^T, C^T A, P)$ e da localização desejada para os polos.

As três proposições descritas acima permitem uma formulação mais simples para o cálculo do ganho e para o cálculo da matriz do erro de covariância para cada subsistema do algoritmo proposto, conforme podemos observar na Tabela 10.

Tabela 10 – Formulação do Filtro Independente das Equações de Riccati

	Ganho de Kalman	Matriz de Erro de Covariância
Riccati	$K_k = P_{k/k-1} C_k^T (C_k P_{k/k-1} C_k^T + R)^{-1}$	$P = APA^T - APC^T [C^T PC + R]^{-1} CPA^T + Q$
Prog. Quadrática	$\hat{F} = (C_k^T A_k)^+ [Q_k^{-1} (\bar{H}_{11} A_k^T - A_k \bar{H}_{21} A_k^T) - A_k^T]$	$P = A \bar{H}_{11} A^T + q$

Fonte: Elaborada pelo autor.

Embora a formulação proposta para o cálculo do ganho e matriz do erro de covariância necessite da inversa da matriz hessiana (\tilde{H}), conforme definido na Equação (4.8), tem-se que

¹ função place

esta matriz inversa pode ser calculada off-line utilizando diversos métodos numéricos tais como a fatorização de Householder (ADADE FILHO, 2003), rotação de Givens (ADADE FILHO, 2003), fatorização de Cholesky (CAMPOS, 2008) e decomposição em valores singulares (REGIS FILHO, 2004).

A aplicação direta de algoritmos de otimização não linear estática em problemas dinâmicos pode resultar em métodos de otimização com problemas de estabilidade numérica. Garante-se um resultado mais significativo quando se considera as condições do sistema e a possibilidade de resolver o problema apresentado pela Equação (4.5) com dimensões reduzidas da matriz hessiana (\tilde{H}). Assim, os requisitos de desempenho computacional para fins de aplicação do algoritmo distribuído em tempo real seriam satisfeitos, considerando-se as relevantes questões de instabilidade numérica e soluções imprecisas concernentes a um sistema mal condicionado. Para amenizar o problema de um sistema mal condicionado, pode-se conferir uma análise de casos envolvendo o número de condição do sistema exposto em Facanha (2010).

Neste contexto, propõe-se uma estratégia para obter um problema equivalente ao original de tal forma que o mesmo seja resolvido através de subproblemas de TRS, para os quais a matriz hessiana \tilde{H} tenha dominância bloco diagonal (OLIVEIRA, 1999), (GARDENGHI; SANTOS, 2012) e se consiga trabalhar com dimensões reduzidas da matriz \tilde{H} . Portanto, define-se o seguinte problema de programação não linear baseado em regiões de confiança:

$$\begin{aligned} \min_{\bar{\lambda}} M(\bar{\lambda}) &= \frac{1}{2} \bar{\lambda}^T \tilde{H} \bar{\lambda} + \bar{\lambda}^T \tilde{b} + \tilde{c} \\ s.a \|\bar{\lambda}\| &\leq \Delta, \Delta > 0, \end{aligned} \quad (4.25)$$

onde $\tilde{H} = -H$, $\tilde{b} = -b$, $\tilde{c} = -c$. É garantido que Δ não será alto se o problema (4.25) estiver bem modelado.

A estrutura especial do problema (4.25) proporciona características dos minimizadores muito mais poderosas que no caso geral de minimização irrestrita (LAUB, 1991). Assim, ao se resolver sucessivas vezes o problema (4.25), isto resulta na aproximação da solução citada na Equação (4.5). Neste contexto, considera-se necessário garantir a existência de uma solução $\bar{\lambda}^*$.

Através dos experimentos realizados, verifica-se que para N pequeno, o erro do resíduo $\|\Delta_{\lambda} M(\bar{\lambda})\|$ é reduzido, tornando a matriz \tilde{H} melhor condicionada. Neste sentido, é possível uma redução significativa da dimensão da matriz \tilde{H} , de nN para nd , onde $d \ll N$, o que possibilita diminuir seu número de condição e ratifica o fato de que é possível resolver este problema de otimização convexa (4.5) por intervalos (LAUB, 1991), ao se resolver sucessivas

vezes o problema (4.25).

Por este critério, considera-se que $d = 2$ com $\bar{\lambda}^* = [\lambda_1 \ \lambda_2]^T$ e tem-se o problema (4.25) aproximado do problema (4.5), assim

$$[\hat{H} + \mu I_{2n \times 2n}] \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} A^T x_0 \\ 0_{n \times n} \end{bmatrix}, \quad (4.26)$$

onde

$$\hat{H} = \begin{bmatrix} C^T R^{-1} C + Q^{-1} & -Q^{-1} A \\ -A^T Q^{-1} & C^T R^{-1} C + Q^{-1} + A^T Q^{-1} A \end{bmatrix}. \quad (4.27)$$

Observa-se que quando $\mu \rightarrow \infty$, ou mais especificamente, $\mu < \max\{0, -\sigma_1\}$, a matriz $(\hat{H} + \mu I)$ tende para

$$\begin{bmatrix} \Omega + \mu I_{n \times n} & 0_{n \times n} \\ 0_{n \times n} & \Omega + A^T Q^{-1} A + \mu I_{n \times n} \end{bmatrix}, \quad (4.28)$$

onde $\Omega = C^T R^{-1} C + Q^{-1}$. Portanto, com as perturbações causadas pelo parâmetro μ , é possível situar os autovalores da matriz \bar{H} , dentro de uma região de confiança

$$\bar{H} = (\hat{H} + \mu I)^{-1}. \quad (4.29)$$

Assim, torna-se possível reduzir a dimensão do sistema e garantir soluções estáveis e bem condicionadas. A partir de $\Theta = Q^{-1}(\bar{H}_{11} - A\bar{H}_{21})$ é possível através de ajustes em μ situar os autovalores da matriz Θ numa determinada região. Na próxima seção, é apresentada uma descrição do algoritmo proposto.

4.2 Descrição do Algoritmo Proposto

Para ilustrar detalhadamente o algoritmo proposto com estrutura mais simples para o cálculo do ganho e para o cálculo da matriz do erro de covariância para cada subsistema, vamos considerar, sem perda de generalidade, a decomposição do sistema dinâmico linear em dois subsistemas, conforme descrito pelas seguintes equações:

$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + \begin{bmatrix} B_k^1 \\ B_k^2 \end{bmatrix} \begin{bmatrix} w_k^1 \\ w_k^2 \end{bmatrix} \quad (4.30)$$

e

$$\begin{bmatrix} y_k^1 \\ y_k^2 \end{bmatrix} = \begin{bmatrix} C_{11} & 0 \\ 0 & C_{12} \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + D_k^1 \begin{bmatrix} v_k^1 \\ v_k^2 \end{bmatrix}. \quad (4.31)$$

O coordenador do sistema paralelizado continua o mesmo, definido no Algoritmo 6. Assim, o sistema linear expresso nas equações (4.30) e (4.31) são particionados em dois subsistemas. Segue abaixo o Subsistema 1 apresentado no Algoritmo 7.

Algoritmo 7: Subsistema 1

Entrada: $A_k^i, B_k^i, C_k^i, P_k^i, Q_k^i, R_k^i, \beta_k^2, z_k^1, \mu$ e yv

$$A_k^1 = A_{11}, B_k^1 = B_{11}, C_k^1 = C_{12}, Q_k^1 = Q_{11}, R_k^1 = R_{11}, P_k^1 = P_{11},$$

$$V_1 = -(C_k^1)^T (R_k^1)^{-1} C_k^1 - (Q_k^1)^{-1} \text{ e } x_k^1 = x_{11}.$$

início

$$H = \begin{bmatrix} -(C_k^1)^T (R_k^1)^{-1} C_k^1 - (Q_k^1)^{-1} & (Q_k^1)^{-1} A_k^1 & 0 \\ [(Q_k^1)^{-1} A_k^1]^T & V_1 - (A_k^1)^T (Q_k^1)^{-1} A_k^1 & (Q_k^1)^{-1} A_k^1 \\ 0 & [(Q_k^1)^{-1} A_k^1]^T & V_1 - (A_k^1)^T (Q_k^1)^{-1} A_k^1 \end{bmatrix}$$

$$\bar{H} = (H + \mu I)^{-1}$$

para $k = 1:N$ **faça**

$$\Psi_k^1 = I_k - (C_k^1)^T [R_k^1 + C_k^1 P_{k+1}^1 (C_k^1)^T]^{-1} C_k^1 P_{k+1}^1$$

$$P_{k+1}^1 = A_k^1 \bar{H}_{11} (A_k^1)^T + Q_k^1$$

$$K_k^1 = [(C_k^1)^T A_k^1]^+ [(Q_k^1)^{-1} (\bar{H}_{11} (A_k^1)^T - A_k^1 \bar{H}_{21} (A_k^1)^T) - (A_k^1)^T]$$

fim

para $k = T-1:-1:1$ **faça**

$$\Gamma^1 = A_{2,1}^T \beta_k^2$$

$$s_k^1 = -A_k^1 P_{k+1}^1 \Psi_{k+1}^1 (A_k^1)^T (C_k^1)^T (R_k^1)^{-1} C_k^1 s_{k+1}^1$$

$$+ A_k^1 P_{k+1}^1 \Psi_{k+1}^1 (A_k^1)^T z_k^1 + A_k^1 s_{k+1}^1 - \Gamma^1$$

fim

para $k = 1:T-1$ **faça**

$$x_{k+1}^1 = \Psi_k^1 [(A_k^1)^T x_k^1 - (C_k^1)^T (R_k^1)^{-1} C_k^1 s_{k+1}^1 + z_k^1]$$

$$x_{k+1}^1 = x_{k+1}^1 + K_k^1 (yv - C_k^1 x_k^1)$$

fim

para $k = T-1:-1:1$ **faça**

$$\lambda_k^1 = (Q_k^1) x_k^1 + A_{1,k} \lambda_{k+1}^i - \Gamma^1$$

fim

fim

Segue o Subsistema 2 apresentado no Algoritmo 8.

Algoritmo 8: Subsistema 2

Entrada: $A_k^i, B_k^i, C_k^i, P_k^i, Q_k^i, R_k^i, \beta_k^1, z_k^2, \mu$ e yv

$$A_k^2 = A_{22}, B_k^2 = B_{21}, C_k^2 = C_{12}, Q_k^2 = Q_{22}, R_k^2 = R_{22}, P_k^2 = P_{22} \text{ e } x_k^2 = x_{21}.$$

$$V_1 = -(C_k^2)^T (R_k^2)^{-1} C_k^2 - (Q_k^2)^{-1} \text{ e } x_k^2 = x_{21}.$$

início

$$H = \begin{bmatrix} -(C_k^2)^T (R_k^2)^{-1} C_k^2 - (Q_k^2)^{-1} & (Q_k^2)^{-1} A_k^2 & 0 \\ [(Q_k^2)^{-1} A_k^2]^T & V_1 - (A_k^2)^T (Q_k^2)^{-1} A_k^2 & (Q_k^2)^{-1} A_k^2 \\ 0 & [(Q_k^2)^{-1} A_k^2]^T & V_1 - (A_k^2)^T (Q_k^2)^{-1} A_k^2 \end{bmatrix}$$

$$\bar{H} = (H + \mu I)^{-1}$$

para $k = 1:N$ **faça**

$$\Psi_k^2 = I_k - (C_k^2)^T [R_k^2 + C_k^2 P_{k+1}^2 (C_k^2)^T]^{-1} C_k^2 P_{k+1}^2$$

$$P_{k+1}^2 = A_k^2 \bar{H}_{11} (A_k^2)^T + Q_k^2$$

$$K_k^2 = [(C_k^2)^T A_k^2]^+ [(Q_k^2)^{-1} (\bar{H}_{11} (A_k^2)^T - A_k^2 \bar{H}_{21} (A_k^2)^T) - (A_k^2)^T]$$

fim

para $k = T-1:-1:1$ **faça**

$$\Gamma^2 = A_{1,2}^T \beta_k^1$$

$$s_k^2 = -A_k^2 P_{k+1}^2 \Psi_{k+1}^2 (A_k^2)^T (C_k^2)^T (R_k^2)^{-1} C_k^2 s_{k+1}^2$$

$$+ A_k^2 P_{k+1}^2 \Psi_{k+1}^2 (A_k^2)^T z_k^2 + A_k^2 s_{k+1}^2 - \Gamma^2$$

fim

para $k = 1:T-1$ **faça**

$$x_{k+1}^2 = \Psi_k^2 [(A_k^2)^T x_k^2 - (C_k^2)^T (R_k^2)^{-1} C_k^2 s_{k+1}^2 + z_k^2]$$

$$x_{k+1}^2 = x_{k+1}^2 + K_k^2 (yv - C_k^2 x_k^2)$$

fim

para $k = T-1:-1:1$ **faça**

$$\lambda_k^2 = (Q_k^2) x_k^2 + A_k^2 \lambda_{k+1}^2 - \Gamma^2$$

fim

fim

4.3 Provas de Conceito

Nesta seção, serão apresentados quatro exemplos para ilustrar a nova metodologia de paralelização para a solução do problema de filtragem com horizonte de tempo reduzido, via subproblemas de região de confiança possibilitando obter uma família de resultados subótimos para cada filtro em seus respectivos subsistemas. Neste contexto, serão analisados os resultados comparativos entre a abordagem do filtro de Kalman distribuído (FKD) e o algoritmo distribuído

com parâmetro μ (FKDPI μ). Os exemplos são simulados com o software Matlab 2017b, na versão estudante.

4.3.1 Exemplo ilustrativo 1

Como primeiro exemplo ilustrativo, considere o seguinte sistema dinâmico:

$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \end{bmatrix} = \begin{bmatrix} 1,1269 & -0,4940 \\ 1,0000 & 0 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + \begin{bmatrix} -0,3832 \\ 0,5919 \end{bmatrix} \begin{bmatrix} w_k^1 \\ w_k^2 \end{bmatrix} \quad (4.32)$$

e

$$\begin{bmatrix} y_k^1 \\ y_k^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + \begin{bmatrix} v_k^1 \\ v_k^2 \end{bmatrix}. \quad (4.33)$$

Neste caso, utilizam-se as matrizes de correlação: $Q = 0,03$ e $R = 0,01$. O ruído de processo e de medição são vistos como gaussianos e com média zero.

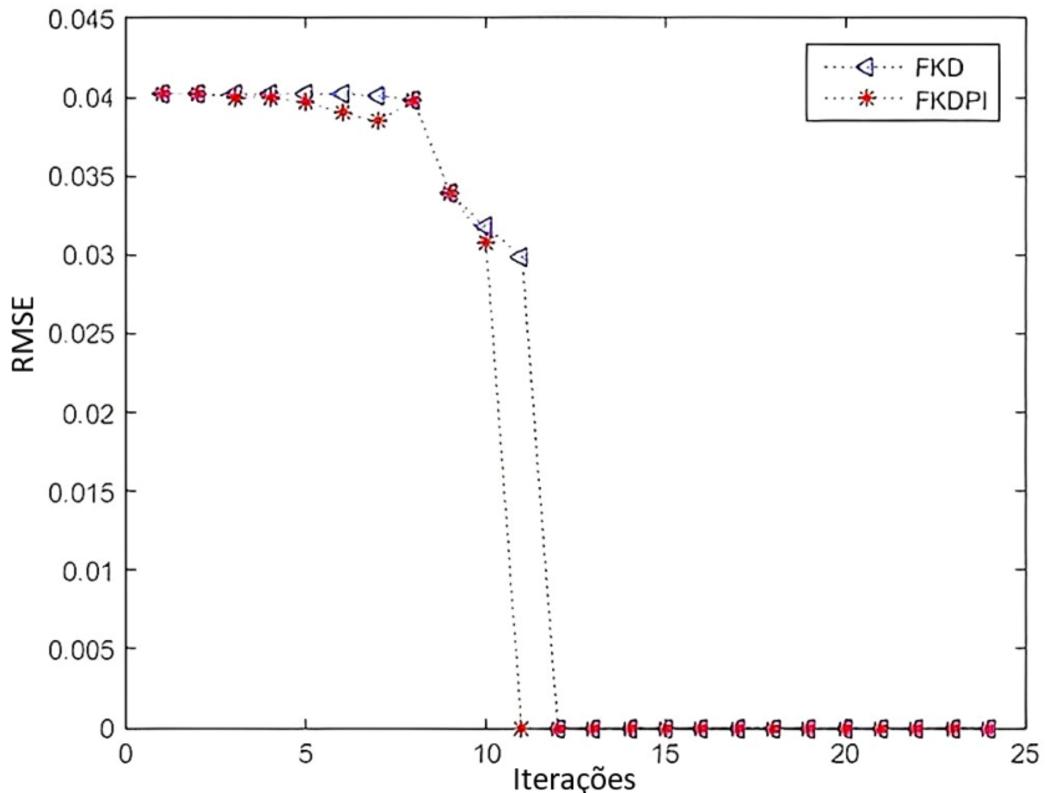
Tabela 11 – Análise do desempenho do algoritmo proposto FDKPI μ para diferentes valores do parâmetro μ (Problema 1).

μ	$\begin{bmatrix} \lambda_{max} \\ \lambda_{min} \end{bmatrix}$	Condicionamento (H)	Ganho - FDKPI μ
0,5	$\begin{bmatrix} 0,4537 \\ 0,2537 \end{bmatrix}$	1,7742	$\begin{bmatrix} 0,8994 \\ 0 \end{bmatrix}$
1	$\begin{bmatrix} 0,2630 \\ 0,1130 \end{bmatrix}$	1,7673	$\begin{bmatrix} 0,9008 \\ 0 \end{bmatrix}$
5	$\begin{bmatrix} 0,2502 \\ 0,1102 \end{bmatrix}$	1,7409	$\begin{bmatrix} 0,9065 \\ 0 \end{bmatrix}$
10	$\begin{bmatrix} 0,2668 \\ 0,1068 \end{bmatrix}$	1,7104	$\begin{bmatrix} 0,9133 \\ 0 \end{bmatrix}$
20	$\begin{bmatrix} 0,2307 \\ 0,1007 \end{bmatrix}$	1,6563	$\begin{bmatrix} 0,9254 \\ 0 \end{bmatrix}$
30	$\begin{bmatrix} 0,1553 \\ 0,0953 \end{bmatrix}$	1,6099	$\begin{bmatrix} 0,9362 \\ 0 \end{bmatrix}$
50	$\begin{bmatrix} 0,1562 \\ 0,0862 \end{bmatrix}$	1,5343	$\begin{bmatrix} 0,9546 \\ 0 \end{bmatrix}$

Fonte: Elaborada pelo autor.

Na Tabela 11, o parâmetro escalar μ é considerado o mesmo para cada subsistema e observa-se que a variação do parâmetro μ influência no condicionamento na matriz H , ao situar os autovalores da matriz dentro de uma região de confiança e permitir a obtenção de uma família de ganhos semelhante ou maior ganho via FKD tradicional, definido em $\begin{bmatrix} 0,6196 & 0,0057 \end{bmatrix}^T$.

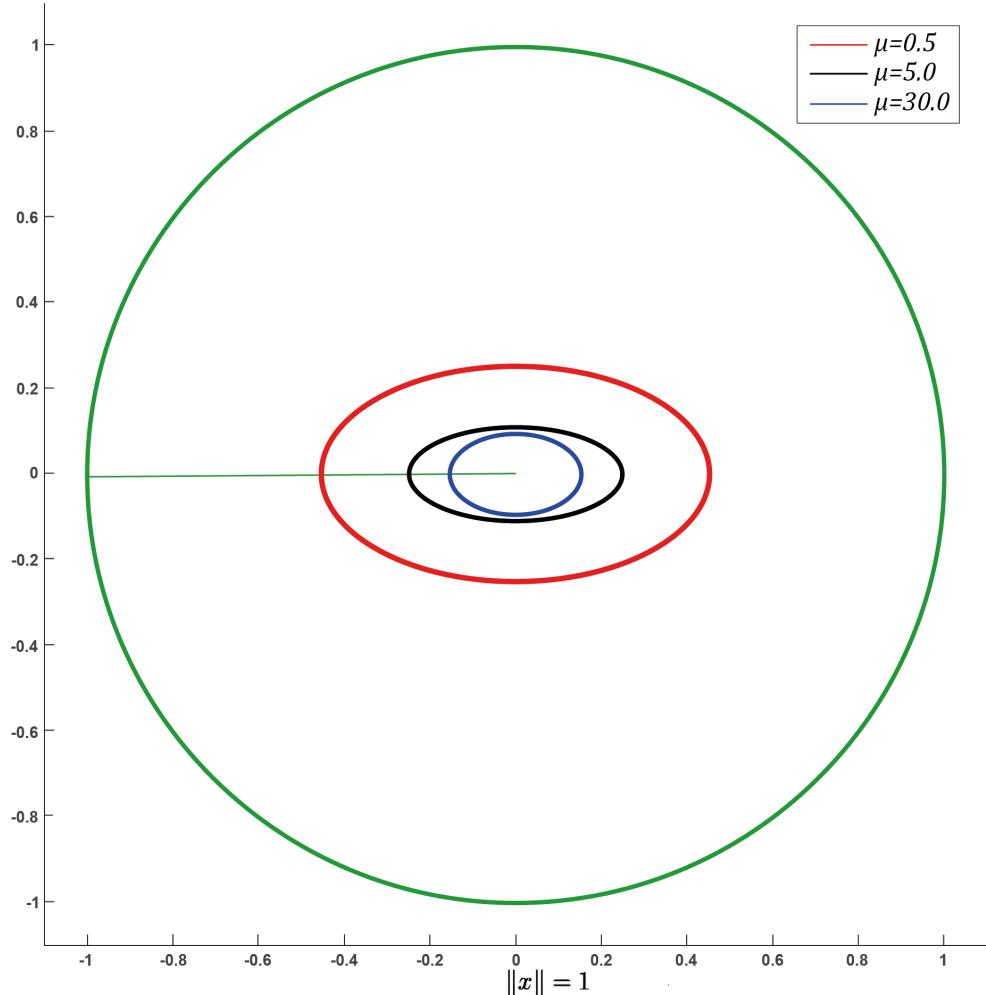
Figura 11 – Análise comparativa do erro quadrático entre FKD e FKDPI
 $\mu(\mu = 0,5)$



Fonte: Elaborada pelo autor.

No algoritmo FDKPI μ é possível obter o mesmo erro quadrático com um menor número de iterações, conforme observamos no Figura 11.

Na Figura 12, apresenta-se a análise geométrica da robustez do sistema de acordo com o conceito de contração da matriz, apresentada no Apêndice (B), quando define-se que uma matriz H é pequena, se e somente se $\sigma_M(H) << 1$, em que σ_M é o maior valor singular da matriz H .

Figura 12 – Efeito do parâmetro μ na contração da matriz H .

Fonte: Elaborada pelo autor.

4.3.2 Exemplo ilustrativo 2

Considera-se como segundo exemplo ilustrativo o seguinte sistema dinâmico, retirado de Costa (1988):

$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \\ x_{k+1}^3 \\ x_{k+1}^4 \\ x_{k+1}^5 \\ x_{k+1}^6 \end{bmatrix} = \begin{bmatrix} 0,800 & 0,200 & 0,000 & 0,000 & 0,00 & 0,00 \\ 0,200 & 0,500 & 0,089 & 0,000 & 0,00 & 0,00 \\ 0,000 & 0,089 & -0,250 & 0,250 & 0,00 & 0,00 \\ 0,000 & 0,000 & 0,250 & -0,250 & 0,13 & 0,00 \\ 0,000 & 0,000 & 0,000 & 0,045 & -0,60 & 0,35 \\ 0,000 & 0,000 & 0,000 & 0,000 & 0,35 & -0,35 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ x_k^4 \\ x_k^5 \\ x_k^6 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_k^1 \\ w_k^2 \\ w_k^3 \end{bmatrix}$$

e

$$\begin{bmatrix} y_k^1 \\ y_k^2 \\ y_k^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ x_k^4 \\ x_k^5 \\ x_k^6 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_k^1 \\ v_k^2 \\ v_k^3 \end{bmatrix}. \quad (4.34)$$

O sistema apresentado neste exemplo é dividido em três subsistemas de dimensão 2. Neste caso, consideram-se as matrizes de correlação: $Q = 0,1I_{6 \times 6}$ e $R = 10I_{3 \times 3}$. O ruído de processo e de medição são vistos como gaussianos, com média zero e variância unitária. O parâmetro escalar μ é considerado o mesmo para cada subsistema.

Na Tabela 12, observa-se que a variação do parâmetro μ entre 0,5 e 20 influencia no condicionamento da matriz H , visto que o valor mínimo encontrado ainda ficou acima de 1, situa os autovalores da matriz dentro de uma região de confiança e permitir a obtenção de uma família de ganhos semelhante ao ou maior do que o ganho obtido via FKD tradicional, definido em $\begin{bmatrix} 0,1030 & 0,0476 & -0,0130 & 0,0283 & -0,0687 & 0,0451 \end{bmatrix}^T$.

4.3.3 Exemplo ilustrativo 3

Como terceiro exemplo, iremos considerar o modelo apresentado no artigo *Asymptotically Optimal Consensus-based Distributed Filtering of Continuous-time Systems* em Battilotti *et al.* (2020), que é descrito pelas seguintes equações:

$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \\ x_{k+1}^3 \\ x_{k+1}^4 \end{bmatrix} = \begin{bmatrix} -0,1 & 0 & 0 & 0 \\ 0,5 & -0,5 & 0 & 0 \\ 1,5 & 0 & -0,2 & 0 \\ -1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ x_k^4 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_k^1 \\ w_k^2 \\ w_k^3 \\ w_k^4 \end{bmatrix} \quad (4.35)$$

Tabela 12 – Análise do desempenho do algoritmo proposto FDKPI μ para diferentes valores do parâmetro μ (Problema 2).

μ	Autovalores	Condicionamento (H)	Ganho - FDKPI μ
0,5	$\begin{bmatrix} 0,7912 \\ 0,4088 \\ -0,7764 \\ -0,4879 \\ 0,0032 \\ -0,1103 \end{bmatrix}$	5,6336	$\begin{bmatrix} 0,0137 \\ 0,0046 \\ -0,0031 \\ 0,0031 \\ -0,0099 \\ 0,0062 \end{bmatrix}$
1	$\begin{bmatrix} 0,7749 \\ 0,3995 \\ -0,7556 \\ -0,4792 \\ 0,1102 \\ -0,0165 \end{bmatrix}$	4,7527	$\begin{bmatrix} 0,1241 \\ 0,0398 \\ -0,0297 \\ 0,0297 \\ -0,0903 \\ 0,0564 \end{bmatrix}$
5	$\begin{bmatrix} 0,6263 \\ 0,5378 \\ -0,1835 \\ -0,1835 \\ 0,3179 \\ -0,1079 \end{bmatrix}$	3,1370	$\begin{bmatrix} 0,3453 \\ 0,1006 \\ -0,0933 \\ 0,0933 \\ -0,2547 \\ 0,1548 \end{bmatrix}$
10	$\begin{bmatrix} 0,6023 \\ 0,1251 \\ -0,1251 \\ -0,1683 \\ 0,1009 \\ -0,4849 \end{bmatrix}$	2,3842	$\begin{bmatrix} 0,4639 \\ 0,1277 \\ -0,1330 \\ 0,1330 \\ -0,3444 \\ 0,2061 \end{bmatrix}$
20	$\begin{bmatrix} 0,1014 \\ 0,1014 \\ -0,4691 \\ -0,5927 \\ 0,0789 \\ -0,1241 \end{bmatrix}$	1,8137	$\begin{bmatrix} 0,5716 \\ 0,1500 \\ -0,1713 \\ 0,1713 \\ -0,4265 \\ 0,2519 \end{bmatrix}$

Fonte: Elaborada pelo autor.

e

$$\begin{bmatrix} y_k^1 \\ y_k^2 \\ y_k^3 \\ y_k^4 \end{bmatrix} = \begin{bmatrix} C_k^i \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ x_k^4 \end{bmatrix} + \begin{bmatrix} v_k^1 \\ v_k^2 \\ v_k^3 \\ v_k^4 \end{bmatrix}. \quad (4.36)$$

Seja $N = 5$, o número de nós conectados em uma cadeia, ou seja, as bordas são $(i, i + 1)$; assim, tem-se os seguintes cenários:

- **Cenário 1: Informações completas.** O estado do sistema está disponível para todos os nós, ou seja, $C^i = I_4$.
- **Cenário 2: Observabilidade Local.** A saída está disponível para todos os nós, $C^i = [1, 1, 1, 1]$. $(C^i A^T)$ é observável.

O algoritmo de DKF pode ser usado apenas quando há observabilidade local. No entanto, em nós com nenhuma medição do ganho, o P^i pode assumir qualquer valor. Os demais parâmetros são definidos na Tabela 8, apresentado no Capítulo 3.

Tabela 13 – Análise do erro quadrático médio (MSE) dos algoritmos

Cenários		CKBF	ADKF	DKF	FKDPI	FKDPI ($\mu_1 = 5, \mu_2 = 0$)	FKDPI ($\mu_1 = 7, \mu_2 = -2$)
1	mse	0,310	0,382	0,408	0,469	0,394	0,351
2	mse	0,742	0,781	0,820	0,840	0,799	0,762

Fonte: Elaborada pelo autor.

Na Tabela 9, apresentada no exemplo três do Capítulo 3, nos dois cenários analisados observa-se que o filtro proposto FKDPI apresenta um erro quadrático menor que os algoritmos SDKF, MKCF, DKBF e IKCF. No entanto, os algoritmos CKBF, ADKF e DKF obtiveram um resultado melhor que o algoritmo proposto.

Na Tabela 13 é possível observar que o ajuste do parâmetro escalar μ , onde $\mu_1 \neq \mu_2$, consegue diminuir o erro quadrático médio. Assim, o algoritmo proposto FKDPI μ apresenta um resultado melhor que os algoritmos DKF e ADKF. A definição dos valores dos parâmetros μ_1 e μ_2 respeita os limites definidos pelo maior autovalor da matriz H_1 e H_2 , respectivamente.

4.3.4 Exemplo ilustrativo 4

Como quarto exemplo ilustrativo tem-se uma variação fortemente acoplada do modelo apresentado no artigo *Asymptotically Optimal Consensus-based Distributed Filtering of Continuous-time Systems* (BATTILOTTI *et al.*, 2020), sendo definido pelas seguintes equações:

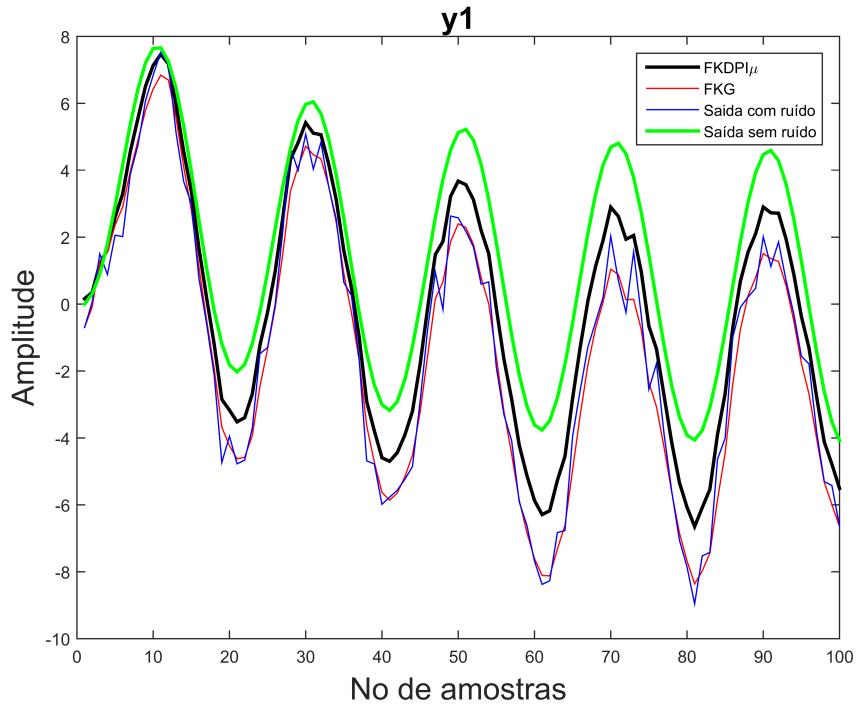
$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \\ x_{k+1}^3 \\ x_{k+1}^4 \end{bmatrix} = \begin{bmatrix} 0,800 & 0,200 & 0,200 & 0,100 \\ 0,200 & 0,500 & 0,089 & 0,100 \\ 0,200 & 0,089 & -0,250 & 0,250 \\ 0,100 & 0,100 & 0,250 & -0,250 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ x_k^4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} w_k^1 \\ w_k^2 \\ w_k^3 \\ w_k^4 \end{bmatrix}$$

e

$$\begin{bmatrix} y_k^1 \\ y_k^2 \\ y_k^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ x_k^4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} v_k^1 \\ v_k^2 \\ v_k^3 \end{bmatrix}.$$

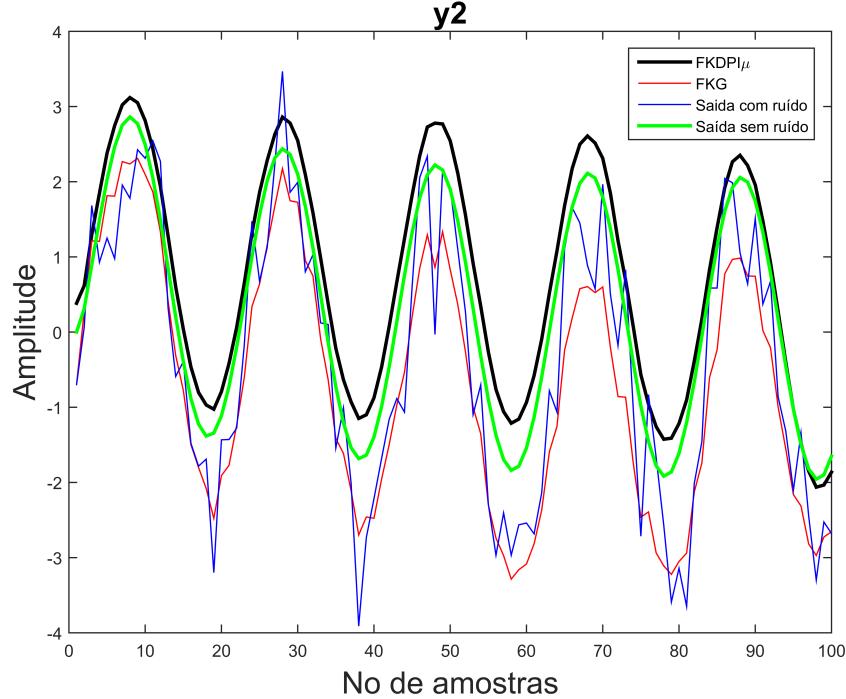
O sistema apresentado neste exemplo é dividido em dois subsistemas de dimensão 2. Neste caso, considera-se as matrizes de correlação: $Q = 0,09I_{6 \times 6}$ e $R = 0,36I_{3 \times 3}$. O ruído de processo e de medição são vistos como gaussianos, com média zero e variância unitária. Os demais parâmetros são definidos na Tabela 8, apresentada no Capítulo 3.

Figura 13 – Saída y_1 : FKG e FKDPI μ ($\mu_1 = 1$ e $\mu_2 = -1$)



Fonte: Elaborada pelo autor.

Nas Figuras 13 e 14, os parâmetros escalares são definidos em $\mu_1 = 1$ e $\mu_2 = -1$. Assim, nota-se que filtro FKG (vermelho) tende a ser mais influenciado pelo ruído, pois este se aproxima da saída com ruído (azul). Também se observa que o algoritmo proposto FKDPI μ (preto) possui um padrão de robustez que ameniza os efeitos do ruído no sinal. Isto corre devido ao ajuste de convergência do filtro através da variação dos parâmetros escalares μ_1 e μ_2 .

Figura 14 – Saída y2 : FKG e FKDPI μ ($\mu_1 = 1$ e $\mu_2 = -1$)

Fonte: Elaborada pelo autor.

Tabela 14 – Impacto da variação dos parâmetros μ_1 e μ_2

μ_1	μ_2	Condicionamento de (H_1)	Condicionamento de (H_2)	rmse (FK)	rmse (FKDPI μ)
0	0	4,66	2,56	[0,5419 0,6505]	[0,7442 0,8914]
1	0	4,15	2,56	[0,4779 0,5982]	[0,5559 0,7872]
5	0	3,02	2,56	[0,5361 0,5828]	[0,3408 0,7867]
10	0	2,41	2,56	[0,4762 0,6053]	[0,1909 0,8083]
20	0	1,86	2,56	[0,4571 0,5330]	[0,0676 0,7282]
20	1	1,86	2,38	[0,4825 0,5519]	[0,0650 0,7803]
20	5	1,86	1,95	[0,5357 0,5812]	[0,0751 0,9325]
20	-5	1,86	5,39	[0,5064 0,5905]	[0,0757 0,4876]
20	-5,5	1,86	6,37	[0,4320 0,5217]	[0,0622 0,3597]

Fonte: Elaborada pelo autor.

Na Tabela 14, mostra-se a influência da variação dos parâmetros escalares μ_1 e μ_2 no condicionamento da matriz H e no RMSE. Por exemplo, para $\mu_1 = 0$ e $\mu_2 = 0$, o filtro FKG apresenta desempenho melhor, pois o RMSE é menor do que o filtro proposto. No entanto, ao variar os parâmetros para $\mu_1 = 20$ e $\mu_2 = -5$, o filtro proposto têm uma melhora

no seu desempenho, tornando-o melhor do que o filtro FKG. Isto ocorre porque o ajuste destes parâmetros gera uma família de ganhos que permite compensar os impactos dos ruídos inerente ao tipo de sensor utilizado numa aplicação em tempo real.

4.4 Comentário Final

Neste capítulo, foi apresentado um aprimoramento da metodologia para síntese do filtro de Kalman distribuído em tempo discreto, mencionado no Capítulo 3. Este aprimoramento ocorre no controle de desempenho e estabilidade por meio de um parâmetro escalar, μ . Este parâmetro pode ser diferente para cada subsistema e respeita os limites definidos pelos autovalores das matrizes H_1 e H_2 .

Este algoritmo provê ao projetista uma flexibilidade no sentido de se poder obter uma família de soluções eficientes e bem condicionadas que podem amenizar os impactos dos ruídos de processo e medida no sistema linear dinâmico. Isto é observado através de quatro experimentos de simulação computacional realizados neste capítulo. No exemplo três foi possível observar as vantagens do algoritmo proposto quando comparado com outros algoritmos distribuídos baseados em consenso.

5 CONCLUSÕES E TRABALHOS FUTUROS

Na tese foram apresentados dois novos algoritmos FKDPI e FKDPI μ para solução dos problemas de filtragem distribuída em tempo discreto. O algoritmo FKDPI explora a dualidade entre o filtro e o controle via o método predição por interação. Neste método é acrescentado no vetor adjunto λ_k^i um termo de malha aberta para solucionar o problema de dupla condição de contorno e que redefinem as equações do filtro. Este algoritmo trata sobre a paralelização hierárquica em que se modifica a estrutura do cálculo do filtro com o objetivo de amenizar a relação de dependência de suas equações e diminuir o tempo de convergência.

A metodologia de paralelização realizada pelo algoritmo proposto FKDPI μ não utiliza as equações recursivas de Riccati e também diminui a relação de dependência das equações do filtro. Através do ajuste de um parâmetro escalar μ é possível obter uma família de soluções eficientes e bem condicionadas para cada subsistema que podem amenizar os impactos dos ruídos de processo e medida no sistema linear dinâmico. Para este propósito, utilizam-se os princípios de otimização convexa e dualidade para obter uma forma quadrática estática com dimensões reduzidas, equivalente ao lagrangiano associado a cada subsistema, através de procedimentos de regularização via regiões de confiança.

Para ressaltar as vantagens dos dois algoritmos propostos, foram examinados duas provas de conceito, um exemplo do artigo de Battilotti *et al.* (2020) e um outro exemplo da variação fortemente acoplado do mesmo artigo. Nestes exemplos foi possível analisar o tempo de execução e o condicionamento do sistema, obter uma família de ganhos para cada subsistema, verificar geometricamente a robustez do sistema de acordo com o conceito de contração da matriz e comparar, através da média do erro quadrático, o algoritmo proposto com outros algoritmos distribuídos baseados em consenso, como por exemplo, DKF, SDKF, ADKF, MKBF e IKCF.

5.1 Contribuições

Neste sentido, destacam-se as seguintes contribuições como as principais desta tese.

- A modificação realizada na estrutura do cálculo para diminuir a relação de dependência das equações do filtro de Kalman.
- A utilização de formas quadráticas com dimensões reduzidas, que possibilitam bom condicionamento e projetos alternativos a filtragem distribuída em aplicações de tempo real.

- Com esta metodologia, os parâmetros livres; ou seja, as matrizes de ponderação, podem ser fixados, sem a necessidade de ajuste exaustivo, pois o parâmetro escalar μ pode ser usado como instrumento para adequar o sistema de filtragem às condições de desempenho e estabilidade de cada subsistema.
- O ajuste dos parâmetros μ_1 e μ_2 garante um certo padrão de robustez do sistema, ao amenizar os impactos dos ruídos de processo e de medida no sistema linear dinâmico.
- Os algoritmos propostos FKDPI e FKDPI μ apresentaram robustez geométrica e convergem rapidamente inclusive para exemplos fortemente acoplados.

5.2 Trabalhos Futuros

Algumas possíveis extensões dos resultados da presente tese são listadas a seguir.

- Propor a formulação e solução de problemas distribuídos com base em jogos dinâmicos (YI; ZORZI, 2020), (MAESTRE *et al.*, Delft, Netherlands, 2011).
- Estudar a aplicação dos algoritmos FKDPI e FKDPI μ em sistemas embarcados, como por exemplo, *cluster de raspberry* , GPU, *jetson*, etc.
- Estender a nova metodologia para solução do problema de filtragem distribuída mista e robusta.
- Estender a nova metodologia de formulação quadrática estática via região de confiança para os algoritmos DKF, SDKF, ADKF, MKBF e IKCF.
- Estender esta nova metodologia para solução de problemas de filtragem distribuída para casos em que os ruídos não são gaussianos.
- Analisar a aplicação da metodologia proposta em filtros robustos H_∞ a *outliers* (FU *et al.*, 2020) em sistemas de tempo discreto.

REFERÊNCIAS

- ADADE FILHO, A. **Análise de Sistemas Dinâmicos**. São José dos Campos: ITA, **2003**.
- ANDERSON, B. D.; MOORE, J. B. **Optimal filtering**. Englewood Cliffs, v. 21, p. 22–95, **1979**.
- BARDI, M.; CAPUZZO-DOLCETTA, I. **Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations**. [S.l.]: Springer Science & Business Media, **2008**.
- BATTILOTTI, S.; CACACE, F.; D'ANGELO, M.; GERMANI, A. **Asymptotically optimal consensus-based distributed filtering of continuous-time linear systems**. Automatica, Elsevier, v. 122, p. 109189, **2020**. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0005109820303873>. Acesso em: 25 fev. 2021.
- BELGACEM, B.; LAMIR, S. **Optimal distributed power control in wireless cellular network based on mixed Kalman/ H_∞ filtering**. AEU-International Journal of Electronics and Communications, Elsevier, v. 90, p. 103–109, **2018**. Disponível em: <https://doi.org/10.1016/j.aeue.2018.04.016>. Acesso em: 13 jan. 2021.
- BERTSEKAS, D. P.; TSITSIKLIS, J. N. **Neuro-dynamic programming**: an overview. In: IEEE PUBL. PISCATAWAY, NJ. Proceedings of the 34th IEEE Conference on Decision and Control. [S.l.], **1995**. v. 1, p. 560–564. Disponível em: <https://ieeexplore.ieee.org/document/478953>. Acesso em: 23 fev. 2022.
- BOYD, S.; GHAOUI, L. E.; FERON, E.; BALAKRISHNAN, V. **Linear matrix inequalities in system and control theory**. [S.l.]: Siam, **1994**. v. 15.
- BOYD, S.; VANDENBERGHE, L. **Convex optimization**. [S.l.]: Cambridge university press, **2004**.
- BRYSON, A. E.; HO, Y.-C. **Applied optimal control: optimization, estimation, and control**. [S.l.]: Routledge, **2018**. Disponível em: <https://doi.org/10.1201/9781315137667>. Acesso em: 23 fev. 2022.
- CACACE, F. **Comments on “Distributed information-weighted Kalman consensus filter for sensor networks”**. Automatica, Elsevier, v. 109, p. 108552, **2019**. Disponível em: <https://ieeexplore.ieee.org/document/10382338>. Acesso em: 10 jan. 2020.
- CAMPOS, L. E. S. **Um estudo sobre fatorações de matrizes e a resolução de sistemas lineares**. Dissertação (Mestrado em Matemática) — Instituto de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas (UNICAMP), Campinas, **2008**. Disponível em: <http://www.unicamp.br/dissertacoes/2008/fatoracoes-matrizes>. Acesso em: 15 jan. 2022.
- CARLI, R.; CHIUSO, A.; SCHENATO, L.; ZAMPIERI, S. **Distributed Kalman filtering based on consensus strategies**. IEEE Journal on Selected Areas in Communications, IEEE, v. 26, n. 4, **2008**. Disponível em: <https://ieeexplore.ieee.org/abstract/document/4497788>. Acesso em: 10 jun. 2021.

CATTIVELLI, F. S.; SAYED, A. H. **Diffusion strategies for distributed Kalman filtering and smoothing**. IEEE Transactions on Automatic Control, IEEE, v. 55, n. 9, p. 2069–2084, 2010. Disponível em: <https://ieeexplore.ieee.org/abstract/document/5411741>. Acesso em: 16 abr. 2021.

CHEN, B.; ZHANG, W.-A.; YU, L. **Distributed finite-horizon fusion Kalman filtering for bandwidth and energy constrained wireless sensor networks**. IEEE Transactions on Signal Processing, IEEE, v. 62, n. 4, p. 797–812, 2014. Disponível em: <https://ieeexplore.ieee.org/abstract/document/6680714>. Acesso em: 23 fev. 2022.

CHEN, B.; ZHANG, W.-A.; YU, L. **Distributed fusion estimation with missing measurements, random transmission delays and packet dropouts**. IEEE Transactions on Automatic Control, IEEE, v. 59, n. 7, p. 1961–1967, 2014. Disponível em: <https://ieeexplore.ieee.org/abstract/document/6701193>. Acesso em: 23 fev. 2022.

CHEN, B.; ZHANG, W.-A.; YU, L.; HU, G.; SONG, H. **Distributed fusion estimation with communication bandwidth constraints**. IEEE Transactions on Automatic Control, IEEE, v. 60, n. 5, p. 1398–1403, 2015. Disponível em: <https://ieeexplore.ieee.org/abstract/document/6897987>. Acesso em: 02 mar. 2021.

CHENG, L.; FENG, L.; WANG, Y. **A Residual Analysis-Based Improved Particle Filter in Mobile Localization for Wireless Sensor Networks**. Sensors, Multidisciplinary Digital Publishing Institute, v. 18, n. 9, p. 2945, 2018. Disponível em: <https://doi.org/10.3390/s18092945>. Acesso em: 07 fev. 2022.

COSTA, J. T. **Processamento paralelo de algoritmos de controle hierárquico**. Dissertação (Dissertação (Mestrado em Engenharia Elétrica)) — Faculdade de Engenharia Elétrica, Universidade Estadual de Campinas (UNICAMP), Campinas, jul. 1988. Disponível em: <http://www.unicamp.br/dissertacoes/1988/controle-hierarquico>. Acesso em: 10 jan. 2020.

COUTINHO, M.; COUTINHO, D.; TROFINO, A.; BARBOSA, K. **A new strategy to the multi-objective control of linear systems**. In: IEEE. Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on. [S.l.], 2005. p. 3741–3746. Disponível em: <https://ieeexplore.ieee.org/document/1582744>. Acesso em: 23 fev. 2022.

CRUZ, J. J. D. **Controle Robusto Multivariável: O Método LGQ/LTR**. [S.l.]: EdUSP, 1996, Vol. 05.

CUMBO, R.; TAMAROZZI, T.; JANSSENS, K.; DESMET, W. **Kalman-based load identification and full-field estimation analysis on industrial test case**. Mechanical Systems and Signal Processing, Elsevier, v. 117, p. 771–785, 2019. Disponível em: <https://doi.org/10.1016/j.ymssp.2018.08.045>. Acesso em: 12 fev. 2022.

DAS, S.; MOURA, J. M. **Distributed Kalman filtering with dynamic observations consensus**. IEEE Transactions on Signal Processing, IEEE, v. 63, n. 17, p. 4458–4473, 2015. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7088659>. Acesso em: 27 jan. 2022.

DOYLE, J. **Guaranteed margins for LQG regulators**. IEEE Transactions on Automatic Control, IEEE, v. 23, n. 4, p. 756–757, 1978. Disponível em: <https://ieeexplore.ieee.org/document/1101812>. Acesso em: 23 fev. 2022.

DU, G.; LEI, Y.; SHAO, H.; XIE, Z.; ZHANG, P. **A human–robot interface using particle filter, Kalman filter, and over-damping method.** Intelligent Service Robotics, Springer, v. 9, n. 4, p. 323–332, 2016. Disponível em: <https://doi.org/10.1007/s11370-016-0202-9>. Acesso em 05 fev. 2022.

FACANHA, T. S. **Uma nova metodologia para síntese de filtros H_2 e H_∞ .** Dissertação (Mestrado em Engenharia de Teleinformática) — Universidade Federal do Ceará (UFC), Fortaleza, ago. 2010. Disponível em: <http://www.ufc.br/dissertacoes/2010/metodologia-filtros>. Acesso em: 17 jun. 2019.

FU, H.; DONG, H.; HAN, F.; SHEN, Y.; HOU, N. **Outlier-resistant H_∞ filtering for a class of networked systems under Round-Robin proto-col.** Neurocomputing, Elsevier, v. 403, p. 133–142, 2020. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0925231220306251>. Acesso em: 17 jun. 2019.

GARDENGHI, J. L. C.; SANTOS, S. A. **Minimização irrestrita usando gradientes conjugados e regiões de confiança.** In: Instituto de Matemática, Estatística e Computação Científica. Relatório de Pesquisa. [S.I.], 2012. Disponível em: <https://www.ime.unicamp.br/pesquisa/relatorios/2012/4>. Acesso em: 23 fev. 2022.

GOLUB, G. H.; LOAN, C. F. V. **Matrix computations.** [S.I.]: JHU press, 2013. v. 3.

GOMIDE, F. A. C. **Análise e Implementação de Algoritmos de Controle Hierárquico de Sistemas Dinâmicos.** Dissertação (Mestrado) — Faculdade de Engenharia Elétrica da Universidade Estadual de Campinas - UNICAMP, Campinas, SP, mar. 1979. Disponível em: <https://hdl.handle.net/20.500.12733/1580116>. Acesso em: 10 jan. 2020.

GREWAL, M. S.; ANDREWS, A. P. **Kalman filtering: Theory and Practice with MATLAB.** [S.I.]: John Wiley & Sons, 2014.

GROMOV, D. **Analysis of hierarchical structures for hybrid control systems.** Tese (Doutorado) — Electrical Engineering and Computer Science of the Technical University of Berlin, ago. 2010. Disponível em: <https://www.semanticscholar.org/paper/Analysis-of-Hierarchical-Structures-for-Hybrid-Gromov/cbc7c30d351372210e41c643df0bba47969aea>. Acesso em: 20 set. 2022. Disponível em: <https://depositonce.tu-berlin.de/bitstream/11303/2861/1/Dokument_29.pdf>.

HASHEMIPOUR, H. R.; ROY, S.; LAUB, A. J. **Decentralized structures for parallel Kalman filtering.** IEEE Transactions on Automatic Control, IEEE, v. 33, n. 1, p. 88–94, 1988. Disponível em: <https://ieeexplore.ieee.org/document/364>. Acesso em: 27 jan. 2022.

HE, W.; FANG, Y.; ZHANG, X. **Prediction-based interception control strategy design with a specified approach angle constraint for wheeled service robots.** In: IEEE. Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on. [S.I.], Tokyo, 2013. p. 2725–2730. Doi: 10.1109/IROS.2013.6696741.

HU, J.; XIE, L.; ZHANG, C. **Diffusion Kalman filtering based on covariance intersection.** IEEE Transactions on Signal Processing, IEEE, v. 60, n. 2, p. 891–902, 2012. Disponível em: <https://ieeexplore.ieee.org/document/6072310>. Acesso em: 23 fev. 2022.

- HU, X.; BAO, M.; ZHANG, X.-P.; GUAN, L.; HU, Y. H. **Generalized Iterated Kalman Filter and its Performance Evaluation.** IEEE Transactions on Signal Processing, v. 63, n. 12, p. 3204–3217, **2015**. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7086336>. Acesso em: 23 fev. 2022.
- HWANG, J.; SHIN, C.; YOE, H. **Study on an agricultural environment monitoring server system using wireless sensor networks.** Sensors, Molecular Diversity Preservation International, v. 10, n. 12, p. 11189–11211, **2010**. Disponível em: <https://doi.org/10.3390/s101211189v>. Acesso em: 27 abr. 2021.
- JI, H.; LEWIS, F. L.; HOU, Z.; MIKULSKI, D. **Distributed information-weighted Kalman consensus filter for sensor networks.** Automatica, Elsevier, v. 77, p. 18–30, **2017**. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0005109816304496>. Acesso em: 25 fev. 2021.
- KALMAN, R. E. **A new approach to linear filtering and prediction problems.** Journal of basic Engineering, American Society of Mechanical Engineers, v. 82, n. 1, p. 35–45, **1960**. Disponível em: <https://ieeexplore.ieee.org/abstract/document/5311910>. Acesso em: 13 fev. 2021.
- KAMAL, A. T.; FARRELL, J. A.; ROY-CHOWDHURY, A. K. **Information weighted consensus filters and their application in distributed camera networks.** IEEE Transactions on Automatic Control, IEEE, v. 58, n. 12, p. 3112–3125, **2013**. Disponível em: <https://ieeexplore.ieee.org/document/6576197>. Acesso em: 25 fev. 2021.
- KAR, S.; MOURA, J. M. **Gossip and distributed Kalman filtering:** weak consensus under weak detectability. IEEE Transactions on Signal Processing, IEEE, v. 59, n. 4, p. 1766–1784, **2011**. Disponível em: <https://ieeexplore.ieee.org/abstract/document/5671497>. Acesso em: 27 jan. 2022.
- KASSAM, S. A.; POOR, H. V. **Robust techniques for signal processing:** a survey. Proceedings of the IEEE, IEEE, v. 73, n. 3, p. 433–481, **1985**. Disponível em: <https://ieeexplore.ieee.org/document/1457435>. Acesso em: 23 fev. 2022.
- KHAN, U. A.; MOURA, J. M. *et al.* **Distributing the Kalman filter for large-scale systems.** IEEE Transactions on Signal Processing, v. 56, n. 10-1, p. 4919–4935, **2008**. Disponível em: <https://ieeexplore.ieee.org/abstract/document/4547458>. Acesso em: 14 fev. 2022.
- KIM, J.; SHIM, H.; WU, J. **On distributed optimal Kalman-Bucy filtering by averaging dynamics of heterogeneous agents.** In: IEEE. 2016 IEEE 55th Conference on Decision and Control (CDC). [S.I.], Las Vegas, NV, USA, **2016**. p. 6309–6314. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7799240>. Acesso em: 10 jan. 2020.
- LANCASTER, P.; RODMAN, L. **Algebraic Riccati equations.** [S.I.]: Clarendon press, **1995**.
- LAUB, A. J. **Invariant subspace methods for the numerical solution of Riccati equations.** In: The Riccati Equation. [S.I.]: Springer, **1991**. p. 163–196.
- LEMOS, J. M. **Introdução ao Controlo Óptimo.** Lisboa Instituto Superior Técnico, SSC, **2001**. Disponível em: <http://www.dem.ist.utl.pt/poliveira/Courses/COA1011/docs/COpt1.pdf>. Acesso em: 20 set. 2022.

LEVENBERG, K. A method for the solution of certain non-linear problems in least squares. **Quarterly of applied mathematics**, v. 2, n. 2, p. 164–168, **1944**.

LEVINE, W. S. **The control handbook**: Control system fundamentals. [S.l.]: CRC press, **2010**.

LI, D.; KAR, S.; MOURA, J. M.; POOR, H. V.; CUI, S. **Distributed Kalman filtering over massive data sets**: analysis through large deviations of random Riccati equations. *IEEE Transactions on Information Theory*, IEEE, v. 61, n. 3, p. 1351–1372, **2015**. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7005535>. Acesso em: 23 fev. 2022.

LI, Y.; WU, J.; CHEN, T. **Transmit power control and remote state estimation with sensor networks**: a Bayesian inference approach. *Automatica*, Elsevier, v. 97, p. 292–300, **2018**. Disponível em: <https://doi.org/10.1016/j.automatica.2018.01.023>. Acesso em: 23 fev. 2022.

LIU, G. P.; YANG, J.-B.; WHIDBORNE, J. F. **Multiobjective optimisation and control**. [S.l.]: Research Studies Press,, **2003**.

LUO, C.; WU, F.; SUN, J.; CHEN, C. W. **Compressive data gathering for large-scale wireless sensor networks**. In: *Proceedings of the 15th annual international conference on Mobile computing and networking*. [S.l.: s.n.], New York, NY, USA, **2009**. p. 145–156. Disponível em: <https://doi.org/10.1145/1614320.1614337>. Acesso em: 27 abr. 2021.

MAESTRE, J. M.; PENA, D. M. de la; CAMACHO, E. F. **Wireless sensor network analysis through a coalitional game: Application to a distributed Kalman filter**. In: *IEEE. 2011 International Conference on Networking, Sensing and Control*. [S.l.], Delft, Netherlands, **2011**. p. 1–6. Disponível em: <https://ieeexplore.ieee.org/document/5874953>. Acesso em: 23 fev. 2022.

MAHMOUD, M. S.; KHALID, H. M. **Distributed Kalman filtering: a bibliographic review**. *IET Control Theory & Applications*, IET, v. 7, n. 4, p. 483–501, **2013**. Disponível em: <https://doi.org/10.1049/iet-cta.2012.0732>. Acesso em: 10 jun. 2021.

MARELLI, D. E.; FU, M. **Distributed weighted least-squares estimation with fast convergence for large-scale systems**. *Automatica*, Elsevier, v. 51, p. 27–39, **2015**. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0005109814004580>. Acesso em: 23 fev. 2022.

MOHAMMADI, A.; PLATANIOTIS, K. N. **Non-Circular Attacks on Phasor Measurement Units for State Estimation in Smart Grid**. *IEEE Journal of Selected Topics in Signal Processing*, IEEE, **2018**. Disponível em: <https://ieeexplore.ieee.org/abstract/document/8365680>. Acesso em: 02 fev. 2021.

MORTENSEN, R. **Maximum-likelihood recursive nonlinear filtering**. *Journal of Optimization Theory and Applications*, Springer, v. 2, n. 6, p. 386–394, **1968**.

MYRONCHUK, O.; SHPYLKA, O.; ZHUK, S. **Algorithm of channel frequency response estimation in orthogonal frequency division multiplexing systems based on Kalman filter**. In: *IEEE. 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*. [S.l.], Lviv–Slavske, Ukraine, **2020**. p. 31–34. Disponível em: <https://ieeexplore.ieee.org/document/9088672>. Acesso em: 23 fev. 2022.

NAIDU, D. S. **Optimal control systems**. [S.l.]: CRC press, **2002**.

NAMARVAR, H. H.; BERGER, T. W. **An efficient training algorithm for dynamic synapse neural networks using trust region methods.** Neural Networks, Elsevier, v. 16, n. 5-6, p. 585–591, 2003. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0893608003000911>. Acesso em: 14 ago. 2022.

NASCIMENTO, V. A. **Uma Contribuição ao Controle Robusto de Sistemas Dinâmicos.** Dissertação (Mestrado) — Universidade Federal do Ceará - UFC, Programa de Pós-Graduação em Engenharia de Teleinformática, set, 2007.

OGATA, K.; OGATA, K. **Designing linear control systems with MATLAB.** [S.l.]: Prentice Hall Englewood Cliffs, NJ, 1994. v. 994.

OLFATI-SABER, R. **Distributed Kalman filtering for sensor networks.** In: IEEE. Decision and Control, 2007 46th IEEE Conference on. [S.l.], New Orleans, LA, USA, textbf2007. p. 5492–5498. Disponível em: ieeexplore.ieee.org/abstract/document/4434303. Acesso em: 23 jan. 2021.

OLFATI-SABER, R.; MURRAY, R. M. **Consensus problems in networks of agents with switching topology and time-delays.** IEEE Transactions on Automatic Control, IEEE, v. 49, n. 9, p. 1520–1533, 2004. Disponível em: <https://ieeexplore.ieee.org/document/1333204>. Acesso em: 23 fev. 2022.

OLIVEIRA, L. M.; RODRIGUES, J. J. **Wireless Sensor Networks: a survey on environmental monitoring.** Journal of Communications, v. 6, n. 2, p. 143–151, 2011. Disponível em: https://www.researchgate.net/publication/220520415_Wireless_Sensor_Networks_A_Survey_on_Environmental_Monitoring. Acesso em: 27 abr. 2021.

OLIVEIRA, M. Controle de sistemas lineares baseado nas desigualdades matriciais lineares. Tese (Doutorado) — Faculdade de Engenharia Elétrica da Universidade Estadual de Campinas - UNICAMP, Departamento de Máquinas, Componentes e Sistemas Inteligentes, 4 1999. Disponível em: <https://repositorio.unicamp.br/acervo/detalhe/175738>. Acesso em: 23 fev. 2022.

PEDREGAL, P. **Introduction to optimization.** [S.l.]: Springer Science & Business Media, 2006. v. 46.

REGIS FILHO, L. A. V. **Filtragem Ótima Robusta em Sistemas Dinâmicos.** Dissertação (Mestrado) — Universidade Estadual de Campinas - UNICAMP, Campinas, SP, nov. 2004. Disponível em: <https://hdl.handle.net/20.500.12733/1597346>. Acesso em: 17 jun. 2019.

REN, W.; BEARD, R. W.; KINGSTON, D. B. **Multi-agent Kalman consensus with relative uncertainty.** In: IEEE. Proceedings of the 2005, American Control Conference, 2005. [S.l.], Portland, OR, USA, 2005. p. 1865–1870. Disponível em: <https://ieeexplore.ieee.org/document/1470240>. Acesso em: 25 fev. 2021.

RIBEIRO, A.; GIANNAKIS, G. B.; ROUMELIOTIS, S. I. **SOI-KF: distributed Kalman filtering with low-cost communications using the sign of innovations.** IEEE Transactions on Signal Processing, IEEE, v. 54, n. 12, p. 4782–4795, 2006. Disponível em: <https://ieeexplore.ieee.org/abstract/document/4014362>. Acesso em: 23 fev. 2022.

SALINAS, S. A.; LI, P. **Privacy-preserving energy theft detection in microgrids:** a state estimation approach. IEEE Transactions on Power Systems, IEEE, v. 31, n. 2, p. 883–894, 2016. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7087399>. Acesso em: 15 fev. 2022.

SCHACKE, K. **On the kronecker product.** Master's thesis, University of Waterloo, **2004**. Disponível em: <https://www.math.uwaterloo.ca/hwolkowi/henry/reports/kronthesisschaecke04.pdf>. Acesso em: 14 ago. 2022.

SHAKED, U.; THEODOR, Y. **H_{sub infinity}-optimal estimation:** a tutorial. In: **IEEE Decision and Control, 1992., Proceedings of the 31st IEEE Conference on.** [S.l.], **1992**. p. 2278–2286. Disponível em: <https://ieeexplore.ieee.org/document/371384>. Acesso em: 14 ago. 2022.

SHEN, X.; SONG, E.; ZHU, Y.; LUO, Y. **Globally optimal distributed Kalman fusion with local out-of-sequence-measurement updates.** IEEE Transactions on Automatic Control, IEEE, v. 54, n. 8, p. 1928–1934, **2009**. Disponível em: <https://ieeexplore.ieee.org/document/5175320>. Acesso em: 23 fev. 2022.

SHULTZ, G. A.; SCHNABEL, R. B.; BYRD, R. H. **A family of trust-region-based algorithms for unconstrained minimization with strong global convergence properties.** SIAM Journal on Numerical Analysis, SIAM, v. 22, n. 1, p. 47–67, **1985**. Disponível em: <https://doi.org/10.1137/0722003>. Acesso em: 14 ago. 2022.

SINGH, M. G.; HASSAN, M. F.; TITLI, A. **Multilevel feedback control for interconnected dynamical systems using the prediction principle.** IEEE Transactions on Systems, Man, and Cybernetics, IEEE, n. 4, p. 233–239, **1976**. Disponível em: <https://ieeexplore.ieee.org/document/5408774>. Acesso em: 17 jun. 2019.

SONG, E.; XU, J.; ZHU, Y. **Optimal distributed Kalman filtering fusion with singular covariances of filtering errors and measurement noises.** IEEE Transactions on Automatic Control, IEEE, v. 59, n. 5, p. 1271–1282, **2014**. Disponível em: <https://ieeexplore.ieee.org/abstract/document/6758348>. Acesso em: 10 jun. 2021.

SONG, H.; YU, L.; ZHANG, W.-A. **Distributed consensus-based Kalman filtering in sensor networks with quantised communications and random sensor failures.** IET Signal Processing, IET, v. 8, n. 2, p. 107–118, **2014**. Disponível em: <https://doi.org/10.1049/iet-spr.2012.0274>. Acesso em: 16 abr. 2021.

SOUZA, C.; GEVERS, M.; GOODWIN, G. **Riccati equations in optimal filtering of nonstabilizable systems having singular state transition matrices.** IEEE Transactions on Automatic Control, IEEE, v. 31, n. 9, p. 831–838, **1986**. Disponível em: <https://ieeexplore.ieee.org/document/1104415>. Acesso em: 23 fev. 2022.

TAKABA, K. **Studies on H_∞ filtering problems for linear discrete-time systems.** Tese (Doutorado) — Applied Mathematics and Physics Kyoto University, 1 **1996**. Disponível em: <https://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/77827/1/D_Takaba_Kiyotsugu.pdf>.

TAKAHARA, Y.; MESAROVIC, M. **Coordinability of dynamic systems.** IEEE Transactions on Automatic Control, IEEE, v. 14, n. 6, p. 688–698, **1969**. Disponível em: <https://ieeexplore.ieee.org/document/1099337>. Acesso em: 20 set. 2022.

TODOROV, E. **Optimal control theory.** In: **Bayesian brain: probabilistic approaches to neural coding.** [S.l.]: Massachusetts: MIT Press Cambridge, **2006**. p. 269–298.

- TODOROV, E. **General duality between optimal control and estimation**. In: **2008 47th IEEE Conference on Decision and Control**. [S.l.]: IEEE, 2008. p. 4286–4292. Disponível em: <https://ieeexplore.ieee.org/document/4739438>. Acesso em: 14 ago. 2022.
- WEI, G.; LING, Y.; GUO, B.; XIAO, B.; VASILAKOS, A. V. **Prediction-based data aggregation in wireless sensor networks**: combining grey model and Kalman filter. Computer Communications, Elsevier, v. 34, n. 6, p. 793–802, 2011. Disponível em: <https://doi.org/10.1016/j.comcom.2010.10.003>. Acesso em: 02 mar. 2021.
- WU, J.; ELSER, A.; ZENG, S.; ALLGÖWER, F. **Consensus-based distributed Kalman-Bucy filter for continuous-time systems**. IFAC-PapersOnLine, Elsevier, v. 49, n. 22, p. 321–326, 2016. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2405896316320043>. Acesso em: 25 fev. 2021.
- XIE, K.; CAI, Q.; ZHANG, Z.; FU, M. **A fast convergent distributed algorithm for weighted average consensus**. IEEE Transactions on Automatic Control, 2018.
- XU, J.; SONG, E.; LUO, Y.; ZHU, Y. **Optimal distributed Kalman filtering fusion algorithm without invertibility of estimation error and sensor noise covariances**. IEEE Signal Processing Letters, IEEE, v. 19, n. 1, p. 55–58, 2012. Disponível em: <https://ieeexplore.ieee.org/document/6087271>. Acesso em: 27 jan. 2022.
- XU, X. **Generalization of the Sherman–Morrison–Woodbury formula involving the Schur complement**. Applied Mathematics and Computation, Elsevier, v. 309, p. 183–191, 2017. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0096300317302230>. Acesso em: 20 set. 2022.
- YI, S.; ZORZI, M. **Low-rank Kalman filtering under model uncertainty**. In: **IEEE. 2020 59th IEEE Conference on Decision and Control (CDC)**. [S.l.], 2020. p. 2930–2935. Disponível em: <https://ieeexplore.ieee.org/document/9303925>. Acesso em: 14 ago. 2022.
- ZANNI, L.; BOUDEC, J.-Y. L.; CHERKAOUI, R.; PAOLONE, M. **A prediction-error covariance estimator for adaptive Kalman filtering in step-varying processes**: application to power-system state estimation. IEEE Transactions on Control Systems Technology, Institute of Electrical and Electronics Engineers, v. 25, n. EPFL-ARTICLE-224583, p. 1683–1697, 2017. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7763847>. Acesso em: 28 fev. 2022.
- ZARCHAN, P.; MUSOFF, H. Fundamentals of kalman filtering: A practical approach. **Virginia, Published by the American Institute of Aeronautics and Astronautics**, p. 129–139, 2010.
- ZHANG, L.; SHI, P.; BOUKAS, E.-K.; WANG, C. **H model reduction for uncertain switched linear discrete-time systems**. Automatica, Elsevier, v. 44, n. 11, p. 2944–2949, 2008. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0005109808002458>. Acesso em: 23 fev. 2022.
- ZHANG, X.; FANG, Y.; SUN, N. **Visual servoing of mobile robots for posture stabilization**: from theory to experiments. International journal of robust and nonlinear control, Wiley Online Library, v. 25, n. 1, p. 1–15, 2015. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/rnc.3067>. Acesso em: 15 fev. 2022.
- ZHANG, X.-b.; WANG, C.-y.; FANG, Y.-c.; XING, K.-x. **An Extended Kalman Filter-Based Robot Pose Estimation Approach with Vision and Odometry**. In: Wearable Sensors and

Robots. [S.I.]: Springer, 2017. p. 539–552. Disponível em: https://doi.org/10.1007/978-981-10-2404-7_41. Acesso em: 01 fev. 2022.

ZHAO, T.; NEHORAI, A. **Information-driven distributed maximum likelihood estimation based on Gauss-Newton method in wireless sensor networks**. IEEE Transactions on Signal Processing, IEEE, v. 55, n. 9, p. 4669–4682, 2007. Disponível em: <https://ieeexplore.ieee.org/abstract/document/4291880>. Acesso em: 10 jun. 2021.

ZHOU, T. **Coordinated one-step optimal distributed state prediction for a networked dynamical system**. IEEE Transactions on Automatic Control, IEEE, v. 58, n. 11, p. 2756–2771, 2013. Disponível em: <https://ieeexplore.ieee.org/document/6545320>. Acesso em: 27 jan. 2022.

ZIMOS, E.; TOUMPAKARIS, D.; MUNTEANU, A.; DELIGIANNIS, N. **Multiterminal source coding with copula regression for wireless sensor networks gathering diverse data**. IEEE Sensors Journal, IEEE, v. 17, n. 1, p. 139–150, 2016. Disponível em: <https://ieeexplore.ieee.org/document/7499830>. Acesso em: 23 fev. 2022.

APÊNDICE A – MINIMIZAÇÃO DA FORMA QUADRÁTICA ESTÁTICA

Uma vez que os problemas de filtragem serão resolvidos via forma quadrática estática conforme definido em (A.1), serão apresentados a seguir os fundamentos teóricos envolvendo minimização de quadráticas.

A.1 Formulação do Problema

Dada a matriz simétrica $H \in \mathbb{R}^{n \times n}$, o vetor $b \in \mathbb{R}^n$ e a constante $c \in \mathbb{R}$, consideraremos o seguinte problema de otimização

$$\min f(s) = \frac{1}{2}s^T H s + b^T s + c. \quad (\text{A.1})$$

Se $f(s) = \frac{1}{2}s^T H s + b^T s + c$, então $\nabla f(s) = H$ para todo $s \in \mathbb{R}^n$. Portanto, de acordo com os conceitos de otimalidade de segunda ordem e convexidade, podemos classificar facilmente os pontos estacionários de (A.1). Com efeito, se s^* é um ponto minimizador local, necessariamente teremos $H = \nabla^2 f(s^*) \geq 0$. Por outro lado, se $H \geq 0$, temos que a Hessiana $\nabla^2 f(s)$ é semi-definida positiva para todo $s \in \mathbb{R}^n$ e, em consequência, f é uma função convexa. Portanto, se $H > 0$ e s^* um ponto estacionário, necessariamente será um minimizador global.

Lema 2.1 – Se $H > 0$ e s^* um ponto estacionário de (A.1), então é um minimizador global da equação (A.1).

Para resolver o problema (A.1) temos à disposição uma variedade de métodos tanto diretos como, por exemplo, a decomposição espectral ou fatoração de Cholesky, quanto iterativos a exemplo do método dos gradientes conjugados. No entanto, o uso de tais métodos incorre, freqüentemente, em um custo operacional intolerável ou problemas de instabilidade numérica (LAUB, 1991), além de que, a solução de (A.1) via gradientes conjugados podem ser intoleravelmente lenta em problemas de grande porte, se o número de condição da matriz Hessiana (H) é grande.

Nesses casos, o método baseado em região de confiança vem sendo recentemente utilizado para “regularizar” o problema (A.1), de modo que a matriz Hessiana (H) passe a ter um número de condição menor. A estratégia, de modo geral, consiste em se construir um problema (A.2) equivalente ao original (A.1), porém bem condicionado e apresentando soluções estáveis, conforme detalharemos na seção a seguir.

A.2 Método das Regiões de Confiança

As primeiras idéias envolvendo regiões de confiança surgiram em um artigo de Levenberg (KASSAM; POOR, 1985), e posteriormente foram retomadas por Marquardt (LANCASTER; RODMAN, 1995) que apresentou, pela primeira vez, a estrutura de subproblemas de regiões de confiança, que permitiram a Powell (NAMARVAR; BERGER, 2003) desenvolver um algoritmo de regiões de confiança para a resolução de problemas de minimização sem restrições. Os métodos de regiões de confiança vêm sendo utilizados na implementação de redes neurais (LIU *et al.*, 2003), reconhecimento de imagens (LEVENBERG, 1944) e, principalmente, como subproblemas auxiliares para resolução de problemas mal condicionados (LAUB, 1991), (OGATA; OGATA, 1994), como aqui é o caso. Há uma vasta teoria que trata da convergência dos métodos de regiões de confiança (ZHANG *et al.*, 2008), (NAMARVAR; BERGER, 2003), (NASCIMENTO, 2007), atribuindo-lhes um fundamento algébrico bastante intuitivo com fortes propriedades de convergência e comprovada eficiência numérica (DOYLE, 1978), (LAUB, 1991) e (SHULTZ *et al.*, 1985).

A.3 Procedimentos de Regularização

Nesta seção, considera-se o seguinte problema

$$\begin{aligned} \min f(s) &= \frac{1}{2} s^T H s + b^T s + c. \\ \|s\| &\leq \Delta \end{aligned} \tag{A.2}$$

onde $H = h^T \in \mathbb{R}^{nxn}$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$, $\Delta > 0$ e $\|\cdot\| = \|\cdot\|_2$ convenção adotada daqui em diante.

Recentemente, alguns autores utilizaram (A.2) como uma maneira de “regularizar” o problema de minimizar uma quadrática irrestrita, pois a estrutura muito especial do problema (A.2) proporciona caracterizações dos minimizadores mais poderosos que no caso geral de minimização restrita (LAUB, 1991). A idéia é que, quando H é mal condicionada, a solução exata de (A.1) carece de sentido, por ser extremamente sensível aos erros dos dados, ou ao arredondamento. Por outro lado, o problema (A.2) é bem condicionado se Δ não é grande. Portanto, substituir (A.1) por (A.2) assegura uma maior estabilidade numérica uma vez que (A.2) é bem condicionada para pequenos valores do raio Δ . Para obter um subproblema (A.2) que forneça uma solução para (A.1) apresentando matrizes bem condicionadas e convergência

garantida, considera-se os seguintes teoremas a seguir.

Teorema D.1 Seja o problema (A.2), se o vetor $\bar{s}^* \in \beta[0, \Delta]$ é a solução para (A.2) então \bar{s}^* é uma solução para a equação da forma

$$(H + \mu I)\bar{s}^* = -b. \quad (\text{A.3})$$

onde $(H + \mu I) \geq 0, \mu \geq 0, \mu(\Delta - \|s^*\|) = 0$.

Prova: O problema (A.2) é equivalente a:

$$\begin{aligned} \min \quad & f(\bar{s}) \\ \text{s.t.} \quad & \bar{s}^T \bar{s} \leq \Delta^2 \end{aligned} \quad (\text{A.4})$$

Como \bar{s}^* é a solução de (A.2), \bar{s}^* satisfaz as condições de KKT (Karush - Kun-Tucker) para (A.3), ou seja, há um $\mu \geq 0$ tal que:

$$\begin{aligned} & H\bar{s}^* + b + \mu\bar{s}^* \\ & \mu(\bar{s}^{*T}\bar{s}^* - \Delta) = 0 \end{aligned} \quad (\text{A.5})$$

Portanto, \bar{s}^* e μ verificam (A.3). Para que $(H + \mu I) \geq 0$, consideraremos inicialmente que $\bar{s}^* \neq 0$. Como \bar{s}^* é a solução de (A.2), \bar{s}^* também é minimizador global de $f(\bar{s})$ sujeito a $\|\bar{s}\| = \|\bar{s}^*\|$. Então:

$$(\bar{s}) \geq f(\bar{s})^* \text{ para todo } \bar{s} \text{ tal que } \|\bar{s}\| = \|\bar{s}^*\| \quad (\text{A.6})$$

Substituindo a equação (A.3) em (A.6), tem-se

$$\frac{1}{2}\bar{s}^T H\bar{s} - (H + \mu I)\bar{s} = \frac{1}{2}\bar{s}^* H\bar{s}^* - \bar{s}^*(H + \mu I)\bar{s}^*. \quad (\text{A.7})$$

Reorganizando a equação (A.7), obtém-se a seguinte equação

$$\frac{1}{2}(\bar{s} - \bar{s}^*)^T (H + \mu I)(\bar{s} - \bar{s}^*) \geq 0 \quad (\text{A.8})$$

para todo \bar{s} tal que $\|\bar{s}\| = \|\bar{s}^*\|$. Como $\bar{s} \neq 0$, as direções $\bar{s} - \bar{s}^*$ tais que $\|\bar{s}\| = \|\bar{s}^*\|$ envolvem todas as direções do espaço exceto os ortogonais a \bar{s}^* . Agora, qualquer vetor ortogonal a \bar{s}^* é o limite de uma seqüência de vetores v_k para as quais $v_k^T (H + \mu I) v_k \geq 0$. Assim, passando ao limite, a expressão $v^T (H + \mu I) v \geq 0$ vale também para os vetores v ortogonais a \bar{s}^* . Portanto, $(H + \mu I)v \geq 0$.

Se $\bar{s}^* = 0$, através da equação (A.3) tem-se que $b = 0$. Então $\bar{s}^* = 0$ é minimizador local de

$$\min \frac{1}{2} \bar{s}^T H s + c \quad (A.9)$$

s.a $\| s \| \leq \Delta$

uma vez que $H \geq 0$ e $v^T (H + \mu I)v \geq 0$ vale para todo $v \in \mathbb{R}^{nN}$ com $\mu = 0$.

No teorema acima, mostra-se que, num minimizador global (A.2), a Hessiana do lagrangiano deve ser semidefinida positiva globalmente, e não apenas restrita a um subespaço (PEDREGAL, 2006), (BOYD *et al.*, 1994).

O próximo teorema fornece as condições suficientes que garantem que \bar{s}^* é a solução do problema de minimização apresentado na equação (A.2).

Teorema 2.5 - Sejam $u \in \mathbb{R}$ e \bar{s}^* satisfazendo (A.3) tais que

$$(H + \mu I)\bar{s}^* = -b \text{ com } (H + \mu I) \geq 0 \quad (A.10)$$

com H semi-definida positiva. Assim, observa-se os seguintes aspectos.

1. Se $\mu = 0$ e $\| \bar{s}^* \| \leq \Delta$, então s^* é a solução da equação (A.3).
2. Se $\| \bar{s}^* \| = \Delta$, então s^* é a solução de

$$\min_{\bar{s} \in \mathbb{R}^n} M(\bar{s}) = \frac{1}{2} \bar{s}^T H \bar{s} + c$$

s.a $\| \bar{s} \| = \Delta$

3. Se $\mu \geq 0$ e $\| \bar{s}^* \| = \Delta$, então s^* é a solução de (A.3). Além disso, se $(H + \mu I)$ é definida positiva, s^* é único em (1), (2) e (3).

Prova: Se μ e s^* satisfazem (A.10), s^* é minimizador da quadrática de

$$M(\bar{s}) = \frac{1}{2} \bar{s}^T (H + \mu I) \bar{s} + \bar{s}^T b + \tilde{c}.$$

Logo, tem-se que

$$\frac{1}{2} \bar{s}^T (H + \mu I) \bar{s} + \bar{s}^T b + \tilde{c} \geq \frac{1}{2} \bar{s}^* (H + \mu I) \bar{s}^* + \bar{s}^* b + \tilde{c}, \quad (A.11)$$

para todo $\bar{s} \in \mathbb{R}^n$. De (A.11) segue-se

$$(\bar{s})(\bar{s}^* + \frac{\mu}{2}(\bar{s}^{*T} \bar{s}^* - \bar{s} \cdot T \bar{s})). \quad (A.12)$$

Para todo $\bar{s} \in \mathbb{R}^n$. As afirmações i), ii) e iii) são consequências imediatas da equação (A.12) A unicidade segue de (A.11), pois se $(H + \mu I) \geq 0$, a desigualdade é estrita para $\bar{s} \neq \bar{s}^*$.

Os teoremas D.1 e D.2 mostram que, se há uma solução \bar{s}^* para o problema apresentado na equação (A.2) situado na fronteira da bola, $\beta[0, \Delta]$ esta deve satisfazer com seu multiplicador correspondente μ , as seguintes equações

$$(H + \mu I)\bar{s}^* = -b, \|\bar{s}^*\| = \Delta, \quad (\text{A.13})$$

com $\mu \geq 0$ e $(H + \mu I) \geq 0$.

Se $\lambda_1 \leq \dots \leq \lambda_n$ são os autovalores de H , a condição $H + \mu I \geq 0$ é equivalente a $\mu \geq -\lambda_1$. Assim, as limitações sobre o multiplicador λ , para detectarmos as soluções na fronteira, se resumem em

$$\mu \geq \max\{0, -\lambda_1\}. \quad (\text{A.14})$$

Portanto, para encontrar as soluções do problema de minimização apresentado na equação (A.2) na superfície da bola basta perceber que $\mu \geq 0$ uma vez que H é definida positiva, portanto, o sistema $(H + \mu I)\bar{s}^* = -b$ tem como solução única $\bar{s}^* = -(H + \mu I)^{-1}b$ uma vez que, neste caso, $(H + \mu I)$ é não singular. Então, encontrar $\mu \geq 0$ satisfazendo (A.13) é equivalente a resolver

$$\|(H + \mu I)^{-1}b\| = \Delta. \quad (\text{A.15})$$

Observação : Alguns algoritmos para obtenção dos valores de μ podem ser encontrados em (BOYD; VANDENBERGHE, 2004), (COUTINHO *et al.*, 2005), (LEVENBERG, 1944) e (LIU *et al.*, 2003).

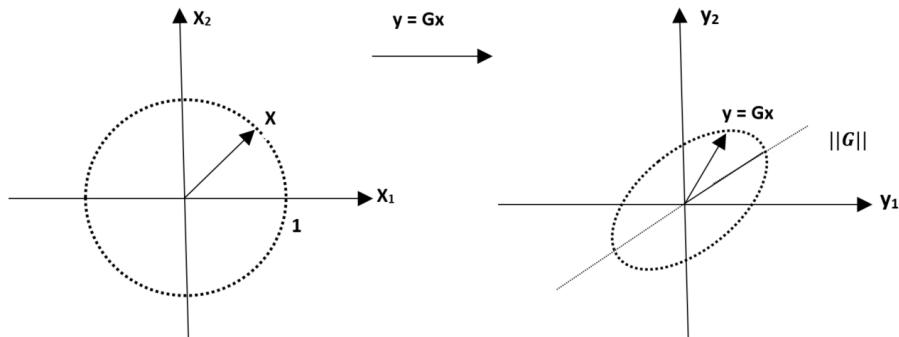
APÊNDICE B – CONTRAÇÃO DE MATRIZES

A ideia de contração de matrizes trata de se determinar modelos de ordem mais baixa que se aproximem do modelo original segundo algum critério pre-definido. Neste contexto, uma ideia gráfica para a redução proposta pode ser obtida por intermédio do que se entende por uma matriz pequena. Assim, defini-se a norma espectral de uma matriz G como a norma induzida pela norma euclidiana, ou seja:

$$\|G\| = \max_{\|x\|=1} \|Gx\| \quad (\text{B.1})$$

O que representa uma generalização do conceito de ganho para sistemas multivariáveis (CRUZ, 1996, Vol. 05). Graficamente, pode-se visualizá-la, conforme apresentado na figura 15.

Figura 15 – Visualização de $\|G\|$



Fonte: Cruz(1995).

Do quociente de Rayleigh (GOLUB; LOAN, 2013), resulta de imediato a seguinte propriedade

$$\|G\| = \sqrt{\rho_M(G^T G)}. \quad (\text{B.2})$$

Em que G^T é a transposta conjugada de G e $\rho_M(G^T G)$ representa o máximo autovalor de $(G^T G)$.

Ao considerarmos os valores singulares de

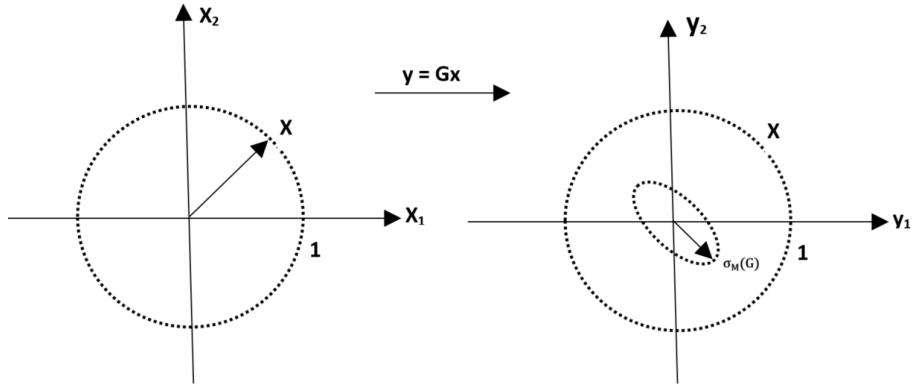
$$\sigma_i(G) = \sqrt{\rho_M(G^T G)} \quad (i = 1, 2, \dots, m). \quad (\text{B.3})$$

Onde $\rho_i(\cdot)$ representa o i -ésimo autovalor de $(G^T G)$, denotamos o valor singular por $\sigma_M(G)$, de modo que $\sigma_M(G) = \sigma_m(G)$ para

$$\sigma_m(G) \geq \dots \geq \sigma_2(G) \geq \sigma_1(G) \geq 0. \quad (\text{B.4})$$

Tendo-se a matriz como uma transformação que opera sobre um conjunto de vetores complexos, produzindo um outro conjunto de vetores complexos diz-se que uma matriz é grande quando ela produz uma amplificação grande para todas as direções dos vetores sobre os quais opera. De forma semelhante diz-se que uma matriz é pequena quando ela produz uma grande atenuação para todas as direções dos vetores sobre os quais opera, conforme a figura 16.

Figura 16 – Visualização da Matriz Pequena



Fonte: Cruz(1995).

Assim, diz-se que G é pequena, se e somente se, $\sigma_M(G) \ll 1$. Neste contexto, entende-se indistintamente as ideias de atenuação e contração de matrizes.