

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA DE CAMPINAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

*Este exemplar corresponde a
revisão final da tese defendida
por Alice M.B.H. Tokarna e aprovada
pela comissão julgadora em 30/03/84*
Wagner

IDENTIFICAÇÃO DESCENTRALIZADA
DE SISTEMAS MULTIVARIÁVEIS

ALICE MARIA B.H. TOKARNIA

Orientador: WAGNER CARADORI DO AMARAL

012/33

Tese apresentada à Faculdade
de Engenharia da Universidade
de Estadual de Campinas - UNI
CAMP - como parte dos requi-
sitos exigidos para obtenção
do título de MESTRE EM CIÊNCIAS.

MARÇO 1984

UNICAMP
BIBLIOTECA CENTRAL

Pela orientação atenciosa ,
pelas idéias, discussões, in-
centivos e apoio oferecidos,
meu agradecimento ao Dr. Wag-
ner C. Amaral e ao Dr. Manuel
de Jesus Mendes.

Aos meus pais e irmãs

CONTEÚDO

RESUMO

CAPÍTULO 1 : INTRODUÇÃO	1
CAPÍTULO 2 : ALGORÍTMOS CENTRALIZADOS PARA IDENTIFICAÇÃO EM SISTEMAS MULTIVARIÁVEIS	19
CAPÍTULO 3 : ALGORÍTMOS PARA IDENTIFICAÇÃO DE SISTEMAS DES CENTRALIZADOS	48
CAPÍTULO 4 : ANÁLISE DOS ALGORÍTMOS ERRA1/1 e ERRB2/1 ATRA VÉS DE EXEMPLOS DE SIMULAÇÃO	107
CAPÍTULO 5 : CONCLUSÕES	125
APÊNDICE A : IDENTIDADES E DEFINIÇÕES MATEMÁTICAS	128
APÊNDICE B : ALGORÍTMO DE JANELA RETANGULAR	131
APÊNDICE C : RESULTADOS DE SIMULAÇÃO PARA O ALGORÍTMO ERRA1/1	136
APÊNDICE D : RESULTADOS DE SIMULAÇÃO PARA O ALGORÍTMO ERRB2/1	147
LISTA DE SÍMBOLOS E ABREVIATURAS	158
REFERÊNCIAS BIBLIOGRÁFICAS	164

RESUMO

Neste trabalho são desenvolvidos conceitos e algoritmos relativos à identificação recursiva em sistemas descentralizados. Os conceitos tornam precisa a noção de descentralização. Estabelece-se a diferenciação entre sistemas centralizados e os descentralizados, com base na troca de informações entre as diversas partes do sistema. Introduz-se o método de modelamento por decomposição e o conceito de erro de descentralização, este como uma maneira formal de exprimir a perda de informação devida a descentralização. Os algoritmos descentralizados são desenvolvidos para redes anel, estrela e com barramento, cada um deles considerando diferentes restrições na comunicação entre os processadores. Estes algoritmos empregam métodos de identificação centralizada adaptados a ambiente descentralizado, através do modelamento do erro de descentralização. Os métodos de identificação centralizada acima são as versões MIMO dos algoritmos recursivos dos mínimos quadrados, raiz quadrada, matriz estendida, mínimos quadrados particionado, janela exponencial e janela retangular. Apresentam-se resultados das simulações de sistemas constituídos de dois subsistemas com interações de intensidades diferentes. Simularam-se um algoritmo para rede anel e outro para rede com barramento. O desempenho destes algoritmos foi avaliado por índices de desempenho definidos no próprio trabalho.

CAPÍTULO 1

INTRODUÇÃO

- 1.1. Identificação de Sistemas e Novas Estruturas:
Hierárquica e Descentralizada
- 1.2. Modelos de Sistemas Dinâmicos Estocásticos
- 1.3. Noções sobre Redes de Computação Distribuída
- 1.4. Objetivos e Organização deste Trabalho

1.1. IDENTIFICAÇÃO DE SISTEMAS E NOVAS ESTRUTURAS : HIERÁRQUICA E DESCENTRALIZADA

Com o objetivo de se prever adequadamente o comportamento de processos, as técnicas de modelamento vêm merecendo especial atenção nos últimos anos. Além disto, o passo inicial para a aplicação das técnicas modernas é a construção de um modelo. Neste modelo devem estar representadas todas as características importantes do processo a ser controlado.

Os procedimentos envolvidos na elaboração de modelos matemáticos são usualmente classificados em [6]:

(a) análise físico-matemática: consiste numa importante ferramenta quando um modelo deve ser construído sem a tomada de medidas no sistema, como, por exemplo, durante a fase inicial do projeto de um sistema.

Considere-se o sistema formado por unidades mais simples, cada uma das quais é modelada com base nas leis de conservação de massa, energia e momento. Adicionando-se a estes modelos condições de contorno que descrevam as interações entre as diversas unidades, obtém-se o modelo do sistema.

(b) identificação experimental: largamente empregada quando se dispõe de medidas do sistema a ser controlado.

O modelo é construído a partir de medidas realizadas com o sistema em funcionamento. Quando uma determinada estrutura pode ser assumida o modelo é obtido através de métodos paramétricos de identificação. Caso contrário, métodos não-paramétricos são utilizados.

Em alguns casos, a identificação experimental é a fase final de um projeto, permitindo verificar se o desempenho do sistema construído corresponde ao previsto no modelo teórico.

Como o objetivo pretendido com a identificação é, em ge

ral, o controle das saídas do processo, o modelo obtido deve aproximar convenientemente o comportamento externo do sistema. Desse modo, os métodos de identificação tratados neste trabalho procuram minimizar algum tipo de função de erro entre as saídas do modelo e as saídas reais. Nenhum dos métodos analisados minimizam uma função da diferença entre os parâmetros estimados e os parâmetros verdadeiros. Como as técnicas existentes para controle de sistemas dinâmicos estocásticos são principalmente técnicas desenvolvidas para sistemas lineares, utilizam-se, neste trabalho, somente modelos lineares com parâmetros desconhecidos. Entretanto, como estes modelos permitem representar, no máximo, o comportamento do sistema real, em torno de um certo ponto de operação, utilizam-se métodos de identificação recursiva ("on-line") que possibilitam uma adaptação contínua do modelo às variações no ponto de operação do sistema [1].

A experiência acumulada no tratamento de diversos tipos de sistema permitiu a consolidação de alguns procedimentos de elaboração de modelos dinâmicos (variáveis de estado, funções de transferência, etc.), de descrição qualitativa (observabilidade, controlabilidade, estabilidade, etc.) e ainda de estratégias de controle (malhas de realimentação, controle ótimo, etc.). Estes procedimentos pressupõem a centralização das informações disponíveis a respeito do sistema, isto é, todas as medidas, modelos e cálculos referentes ao sistema devem estar num mesmo local. Esta hipótese de centralização dificulta a aplicação destes procedimentos em sistemas de grande porte. Para estes sistemas, a descentralização surge naturalmente tanto devido a falta de informações globais sobre o processo como à ausência de um local comum para o processamento eficiente destas informações. Exemplos destes sistemas são encontrados em complexos industriais, na geração e distribuição de energia elétrica, no tráfego urbano, nas redes de comunicação, nos complexos ecológicos e nas relações econômicas. Muitas situações críticas têm surgido nestes sistemas em decorrência da falta de um controle descentralizado que permita coordenar o funcionamento de suas diversas partes.

Sistemas de grande porte são constituídos de subsiste

mas interligados, que, em muitas situações práticas, não podem ser controlados de maneira centralizada. Para formalizar esta metodologia é útil distinguir dois tipos de estrutura: hierárquica e descentralizada [28].

Numa estrutura hierárquica, o modelamento e o controle são efetuados por algoritmos que operam em escalas de tempo diferentes. Como exemplo, uma estrutura hierárquica com três níveis está representada na figura 1.1. O vetor-controle $\underline{u}$ é determinado por um regulador de modo que o vetor-saída $\underline{y}$ acompanhe o valor de referência $\underline{y}^d$ num período de tempo T_1 . A referência $\underline{y}^d$ é calculada por um algoritmo de otimização para um período de tempo $T_2 \gg T_1$. Para o cálculo de $\underline{y}^d$ são fornecidos ao algoritmo de otimização, valores dos parâmetros $\hat{\theta}$ do modelo do sistema, computados por um algoritmo de identificação e atualizados ao fim de cada período T_2 . Quanto mais alto o nível, mais complexo deve ser o modelo utilizado e maior o período de tempo para atualização deste modelo.

Para uma estrutura descentralizada o vetor de controle $\underline{u}$ está dividido em subvetores. Cada um destes é obtido independentemente por um controlador local, com base num modelo e critério locais. Para que o processo como um todo seja eficientemente controlado, os modelos e critérios dos controladores locais devem ser ajustados de forma coordenada. A figura 1.2 representa um exemplo de um sistema descentralizado, no qual um controlador geral coordena os controladores locais.

Neste trabalho apresentam-se algoritmos de identificação para estruturas descentralizadas e também explora-se a atualização de modelos em períodos diferentes de tempo, conforme visto para redes hierárquicas.

A seguir apresenta-se uma revisão sobre modelamento de sistemas dinâmicos através de equações a diferenças bem como noções sobre redes de computação distribuída.

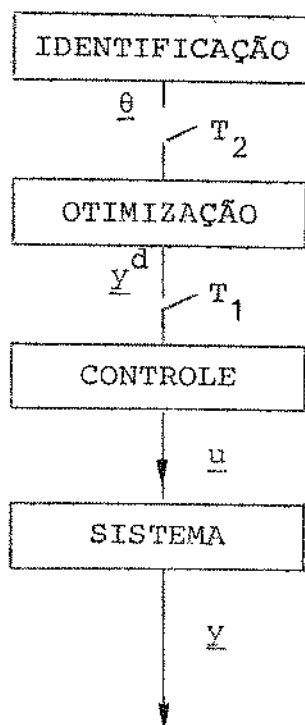


Figura 1.1 - Estrutura Hierárquica de Identificação e Controle

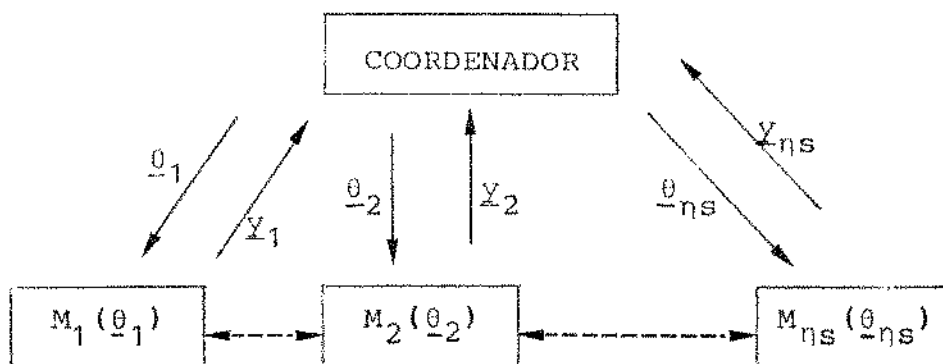


Figura 1.2 - Estrutura Descentralizada de Identificação e Controle

1.2. MODELOS DE SISTEMAS DINÂMICOS ESTOCÁSTICOS

Os métodos de identificação utilizam como modelos a representação de espaço de estado, de entrada-saída e de matriz de resposta a um vetor de impulsos. Como esta última representação tende a envolver um número elevado de parâmetros é frequentemente preterida em relação às duas primeiras.

A representação de espaço de estado consiste numa equação vetorial de primeira ordem, na qual o vetor-saída, num certo instante, é descrito a partir dos valores imediatamente anteriores do vetor-entrada e do vetor-estado. Na representação entrada-saída, uma equação vetorial descreve o vetor-saída num determinado instante como função de seus valores passados e também dos valores passados do vetor-entrada.

Nos algoritmos desenvolvidos neste trabalho emprega-se a representação entrada-saída. A terminologia adotada e algumas limitações dos problemas abordados são apresentados a seguir.

De acordo com a representação entrada-saída a equação (1.1) descreve a parcela determinística do vetor-saída y^D como função do vetor-entrada u .

$$A(q^{-1}) \cdot y_K^D = B(q^{-1}) \cdot u_K \quad (1.1)$$

onde:

$$y_K^D \in \mathbb{R}^n$$

$$u_K \in \mathbb{R}^l$$

q^{-1} é o operador de regressão temporal unitária, definido por:

$$q^{-1} \cdot y_K^D = y_{K-1}^D$$

$A(q^{-1})$ e $B(q^{-1})$ são as matrizes polinomiais:

$$A(q^{-1}) = I_n + A_1 q^{-1} + \dots + A_{N_A} q^{-N_A}$$

$$B(q^{-1}) = B_1 q^{-1} + \dots + B_{N_B} q^{-N_B}$$

com

$$A_i \in \mathbb{R}^{n \times n} \quad i = 1, \dots, N_A$$

e

$$B_j \in \mathbb{R}^{n \times l} \quad j = 1, \dots, N_B$$

Como na equação (1.1) diferentes pares de matrizes $\{A(q^{-1}), B(q^{-1})\}$ podem resultar na mesma matriz de transferência, é usual empregar pares de matrizes que sejam irredutíveis, isto é, que não apresentem fatores comuns a menos de matrizes unimodulares. Um par de matrizes $\{A(q^{-1}), B(q^{-1})\}$ que satisfaça esta condição é denominado Descrição por Matriz Partilhada (DMP).

Indicando por $\underline{\zeta}_K$ a parcela estocástica do vetor-saída $\underline{y}_K$, a equação (1.2) corresponde ao modelo de $\underline{y}_K$.

$$\underline{y}_K = \underline{y}_K^D + \underline{\zeta}_K \quad (1.2)$$

Um modelo de $\underline{\zeta}_K$ como função de uma sequência $\{\underline{w}_K\}$ de variáveis aleatórias n -dimensionais não correlatas, com média zero e matriz de covariância V pode ser obtido a partir do Teorema de Factorização Espectral. Uma sequência $\{\underline{w}_K\}$ com as propriedades mencionadas é referida como ruído branco.

Teorema de Factorização Espectral: Dada $\Gamma(q)$, $n \times n$, uma matriz densidade espectral racional discreta com característica n , então existe uma única matriz racional $H(q)$, $n \times n$ e uma única matriz simétrica V , tais que:

$$a) \Gamma(q) = H(q) \cdot V^{-1} \cdot H^T(q^{-1})$$

b) $H(q)$ é analítica no limite e fora do círculo unitário, isto é, para $|q| \geq 1$

c) $H(q^{-1})$ é analítica fora do círculo unitário, isto é, para $|q| > 1$

$$d) \lim_{q \rightarrow \infty} H(q) = I$$

A demonstração e alguns comentários interessantes podem ser encontrados nas referências bibliográficas [2],[7],[8],[21] , [27] e [31]. Para os propósitos deste trabalho é suficiente des tacar que em decorrência deste teorema, é possível representar um processo estocástico com média zero e densidade espectral $\Gamma(q)$ co mo a saída de um sistema linear cuja entrada é um ruído branco.

A representação de $\underline{z}_K$ como função de um ruído branco $\underline{w}_K$ portanto, tem a forma apresentada na equação (1.3).

$$\underline{z}_K = H(q^{-1}) \cdot \underline{w}_K \quad (1.3)$$

As equações (1.1), (1.2) e (1.3) constituem um modelo de entrada-saída de um sistema dinâmico estocástico.

Uma forma alternativa para o modelo descrito nas equa ções (1.1), (1.2) e (1.3) pode ser obtida decompondo-se a matriz $H(q^{-1})$ como:

$$H(q^{-1}) = D^{-1}(q^{-1}) \cdot C(q^{-1}) \quad (1.4)$$

onde:

$$C(q^{-1}) = I_n + C_1 \cdot q^{-1} + \dots + C_{N_C} \cdot q^{-N_C} \quad \text{com } C_i \in \mathbb{R}^{n \times n}$$

$$D(q^{-1}) = I_n + D_1 \cdot q^{-1} + \dots + D_{N_D} \cdot q^{-N_D} \quad \text{com } D_i \in \mathbb{R}^{n \times n}$$

Substituindo-se esta equação nas equações (1.1), (1.2) e (1.3) obtém-se o modelo de entrada-saída:

$$A(q^{-1}) \cdot \underline{y}_K = B(q^{-1}) \cdot \underline{u}_K + \underline{z}_K \quad (1.5)$$

$$D(q^{-1}) \cdot \underline{z}_K = C(q^{-1}) \cdot \underline{w}_K$$

Estas equações estão esquematizadas na figura 1.3, na qual se distinguem facilmente as parcelas determinística e estocás tica do modelo.

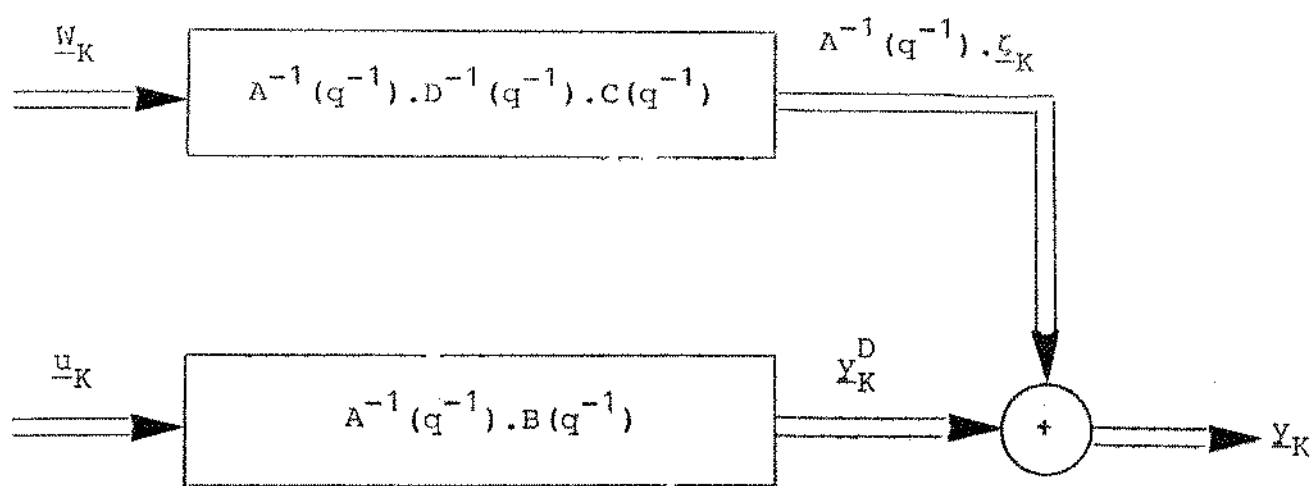


Figura 1.3 - Esquema de Estimação Recursiva

Em aplicações práticas, quando o ruído $\underline{z}_K$ tem média zero e memória finita, ou seja, para:

$$E\{\underline{z}_K\} = 0 \quad (1.6)$$

$$e \quad E\{\underline{z}_K \underline{z}_{K-j}^T\} = 0 \quad \forall j \geq m_0+1$$

é usual escrever o modelo (1.5) como;

$$F(q^{-1}) \underline{y}_K = R(q^{-1}) \underline{u}_K + C(q^{-1}) \underline{w}_K \quad (1.7)$$

onde:

$$F(q^{-1}) = D(q^{-1}) \cdot A(q^{-1})$$

$$R(q^{-1}) = D(q^{-1}) \cdot B(q^{-1})$$

Naturalmente, a ordem da equação a diferenças (1.7) em $\underline{y}_K$ e $\underline{u}_K$ é maior do que para a equação (1.5).

O modelo representado em (1.5) e (1.7) é caracterizado por parâmetros primários e secundários. Os parâmetros primários são N_A , N_B , N_C e N_D que indicam a ordem das equações e portanto, podem assumir apenas valores inteiros não-negativos. Para este trabalho, tomou-se sempre

$$N_A = N_B = N_C = N_D \stackrel{\Delta}{=} N_* \quad (1.8)$$

sendo o valor N_* referido como ordem do modelo. Esta simplificação não restringe a teoria desenvolvida nos próximos capítulos. Os parâmetros secundários são os coeficientes das matrizes presentes no modelo e da matriz de covariância V do ruído branco.

É útil ter em mente o grau de complexidades dos modelos representados na equação (1.7), pois ao ajustar-se um destes modelos a um conjunto de dados, os mais simples devem ser testados em primeiro lugar. Como as matrizes $F(q^{-1}) \cdot C(q^{-1})$ e a matriz de covariância V expressam as interconexões entre as saídas do modelo, o grau de complexidade dos modelos pode ser apresentado em termos

estruturas destas matrizes. Assim, começando com os modelos mais simples, tem-se [13]:

- (a) $F(q^{-1}) = \text{diagonal}$, $C(q^{-1}) = I_n$, $V = \text{diagonal}$
- (b) $F(q^{-1}) = \text{não diagonal}$, $C(q^{-1}) = I_n$, $V = \text{diagonal}$
- (c) $F(q^{-1}) = \text{não diagonal}$, $C(q^{-1}) = I_n$, $V = \text{não diagonal}$
- (d) $F(q^{-1}) = \text{não diagonal}$, $C(q^{-1}) = \text{diagonal}$, $V = \text{diagonal}$
- (e) $F(q^{-1}) = \text{não diagonal}$, $C(q^{-1}) = \text{diagonal}$, $V = \text{não diagonal}$
- (f) $F(q^{-1}) = \text{não diagonal}$, $C(q^{-1}) = \text{não diagonal}$, $V = \text{não diagonal}$

Os modelos (a)-(c) são autoregressivos (AR) enquanto os modelos (d)-(f) são autoregressivos e média-móvel (ARMA). O caso (a) corresponde a um modelo autoregressivo com saídas completamente independentes. A complexidade aumenta até o caso (c), o mais geral entre os modelos autoregressivos. Entre os modelos autoregressivos e média móveis os mais complexos estão no caso (f). Entretanto, é sempre possível, nas situações em que os modelos autoregressivos se revelam insuficientes, ajustar os dados disponíveis a um modelo do caso (d).

Naturalmente, como neste trabalho são estudados sistemas descentralizados, é necessário introduzir uma terminologia que permita especificar como é feita a divisão em subsistemas. Além disto, ficam faltando algumas considerações a respeito do grau de interação entre os diversos subsistemas. Estes itens são desenvolvidos no Capítulo 3.

1.3. NOÇÕES SOBRE REDES DE COMPUTAÇÃO DISTRIBUÍDA

A rigidez das especificações de desempenho de sistemas vem sendo satisfeita pelo desenvolvimento de dispositivos e metodologias de controle progressivamente sofisticadas. Os primeiros

dispositivos mecânicos e eletropneumáticos cederam lugar a dispositivos analógicos eletrônicos, que vem sendo gradualmente suplantados pela eficiência e flexibilidade dos computadores de processo ou de redes à base de microprocessadores.

Atualmente, processos de porte e complexidade crescentes dependem de métodos que permitam melhorar a qualidade e flexibilidade de seus produtos, mantendo a segurança e proteção ambiental, mesmo em áreas de tecnologia avançada. Estes requisitos devem ser confrontados com as possibilidades reais para a construção de sistemas de controle, ou seja, custos crescentes de peças mecânicas e condutores metálicos e custos notadamente decrescentes de circuitos integrados de larga escala, capazes de implementar funções digitais complexas. A necessidade de conciliação dos requisitos e disponibilidades mencionadas abre caminho para mais uma importante mudança qualitativa nos sistemas de controle: o aparecimento de novas estruturas e formas para descrever, programar, avaliar e otimizar o sistema de controle. Os métodos de identificação elaborados neste trabalho consideram algumas destas estruturas e uma breve exposição a respeito delas é o assunto desta seção.

Na verdade, os fornecedores de equipamentos para controle de processos já estão colocando no mercado sistemas de controle distribuído constituídos de microprocessadores interligados por uma rede de comunicações [5]. Estes sistemas vem sendo caracterizados por quatro aspectos, nem todos obrigatoriamente presentes:

- (a) Microprocessadores e controladores analógicos efetuam as funções básicas de controle.
- (b) Microcomputadores e computadores de grande porte se encarregam das funções de controle de alto nível com maior tolerância para tempo de resposta que as funções básicas.
- (c) Estrutura espacialmente distribuída implicando na existência de uma rede de comunicações entre computadores.
- (d) Interfaces para operação centralizada com acesso a

- Δ unidades de campo
- $\square$ unidades de processamento
- $\bigcirc$ unidades de gerenciamento de comunicações
- meios de transmissão

Figura 1.4 -- Componentes Básicos das Redes de Computação Distribuída

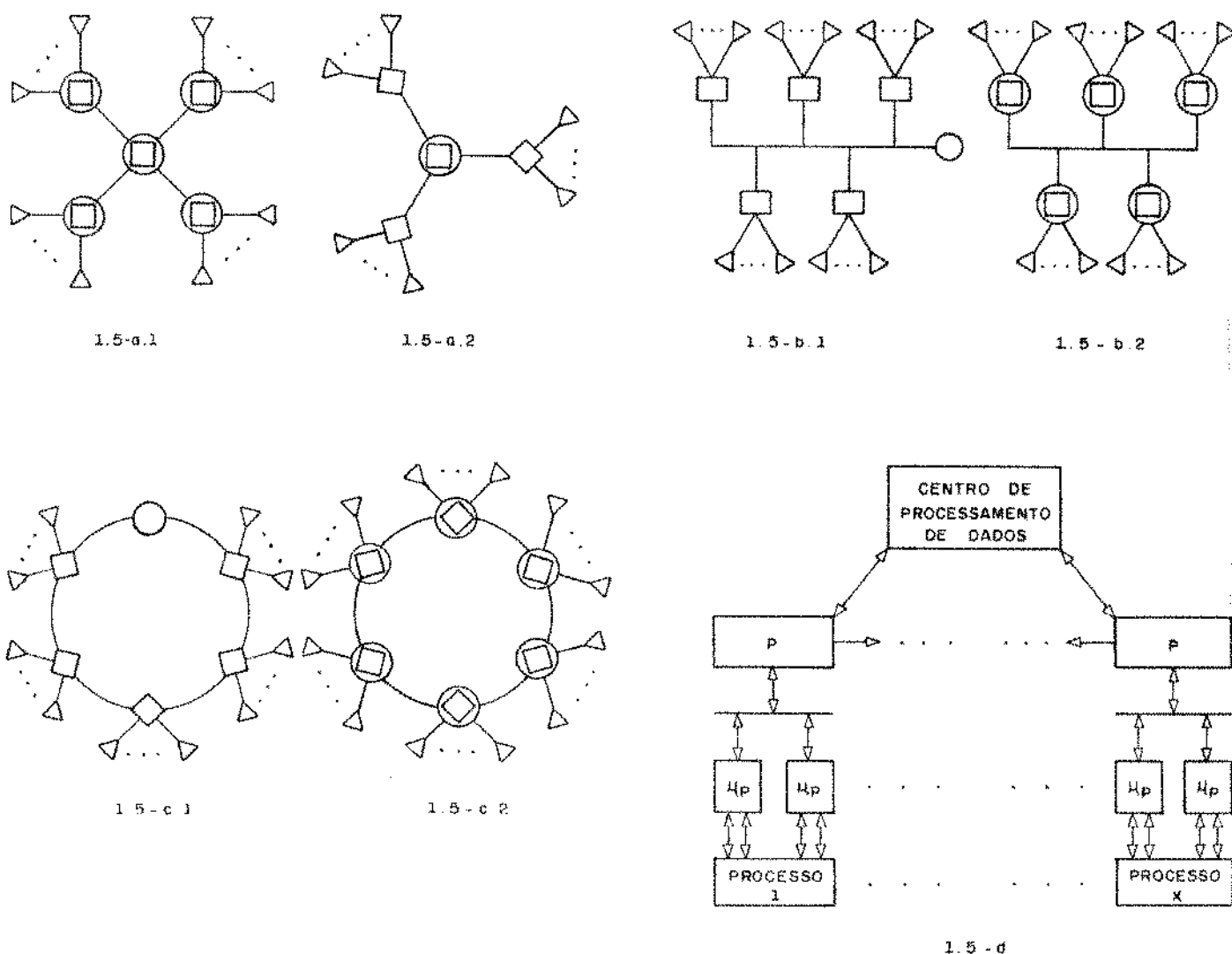


Figura 1.5 - Configurações Usuais de Redes de Computação Distribuída

todas as variações do processo, às variáveis de controle e aos indicadores de estado do sistema.

A gerência, os parâmetros e a estrutura da rede de comunicções constituem um ponto crítico para o bom desempenho do sistema distribuído. Assim sendo, deve se levar em conta a aplicação específica no projeto da configuração desta rede, das redundâncias necessárias e de regras de operação também denominadas de protocolo de comunicações. As configurações mais comuns, e mais intimamente relacionadas com os algoritmos desenvolvidos neste trabalho são a estrela, o barramento global, o anel e a rede hierárquica.

Para qualquer uma das configurações mencionadas, os componentes básicos das redes de computação distribuída se dividem em quatro grupos, representados simbolicamente na figura 1.4 [29]:

- (a) unidades de campo (sensores e atuadores)
- (b) unidades de processamento: controladores, computadores e suas interfaces com os meios de comunicação
- (c) unidades de gerenciamento da rede de comunicação (regulam o envio de mensagens através dos meios de comunicações, provenientes dos componentes dos outros dois grupos)
- (d) meios de transmissão (fios, linhas, até fibras ópticas, etc.)

Nas configurações estrela, mostradas na figura 1.5a, existe sempre um processador central de maior porte ligado a outros localizados nas "pontas" da estrela, os quais não mantêm entre si comunicação direta. Estas configurações têm como principais desvantagens a exigência de um computador de grande porte para a posição central, cuja disponibilidade pode se tornar crítica, e o alto custo dos meios de transmissão. As outras configurações procuram superar estes problemas de forma econômica.

Nas redes lineares, apresentadas na figura 1.5b, o canal de comunicação é denominado barramento global. O barramento opera

no modo de difusão múltipla, possibilitando a recepção de uma mensagem por todos os processadores. Por outro lado, como apenas uma mensagem pode ser transmitida de cada vez, o barramento deve en tão estar desocupado para aceitar a mensagem de um processador.

Para a rede na figura 1.5.b.1 não ocorrem problemas de acesso múltiplo, uma vez que o controle de comunicações é feito a partir de um único ponto. Esta rede, entretanto, tem baixa confi bilidade: pode ser totalmente desativada por uma falha no contro lador ou ter parte de seus processadores isolados por uma interrupção na rede. A solução deste problema é obtida com a rede da figu ra 1.5.b.2, pelo aumento do número de pontos de controle. Como re sultado são também necessários protocolos mais elaborados para re solver os problemas de acesso múltiplo. Constituem exemplos de tais protocolos:

ETHERNET[20], MITRIX[19], e HXDP[12].

As redes em anel são representadas esquematicamente nas figuras 1.5.c. Na rede da figura 1.5.c.1 o controle de comunica ções é realizado a partir de um único ponto. Ainda assim, esta re de apresenta maior confiabilidade que o barramento com um único pon to de controle, pois ocorrendo uma interrupção na rede, a comuni cação entre os processadores é mantida. A instalação de vários pon tos de controle como na rede da figura 1.5.c.2, torna os segmen tos rede eletricamente independentes, possibilitando transmissão e recepção simultânea de mensagens. Alguns exemplos de protocolos de comunicação para redes em anel são: New-hall [11],[17], Pierce [11],[17], Reames [11],[17].

Para a configuração hierárquica, genericamente represen tada na figura 1.5.d, existem vários níveis de complexidade pro gressiva. O nível inferior é ocupado por processadores simples (mi croprocessadores), responsáveis pela gerência de sensores e atua dores, aquisição de dados e algoritmos simples. No nível interme diário aparecem minicomputadores e microprocessadores de maior ca pacidade encarregados da supervisão e algoritmos mais complexos. No nível superior, computadores de maior porte executam a supervi são e o planejamento globais. Uma observação importante a respei

to das redes hierárquicas é que, enquanto os processadores do nível inferior devem ser especializados para um determinado processo, os processadores dos outros níveis podem ser compartilhados por vários processos em operação simultânea. Um estudo sobre redes hierárquicas encontra-se na referência bibliográfica [16]. É interessante destacar que existem redes hierárquicas nas quais os processadores de um ou mais níveis estão interligados por uma rede distribuída.

A partir das redes mencionadas pode-se construir outras com características híbridas. Alguns exemplos de tais redes são o anel duplo [18], o barramento duplo [18], o barramento e o anel associados [2] e estrelas em anel [29].

Além da configuração, outros critérios comuns para classificação de redes de computação são a distância e taxa de transmissão de informações entre os processadores. Tomando por base estes dados, as redes são agrupadas em geográficas, locais e de multi-processadores concentrados, conforme na tabela 1.1.

As redes de computação distribuída abrem um novo campo para pesquisas. O projeto adequado de tais redes pode proporcionar mais disponibilidade de recursos para processamento, confiabilidade, flexibilidade, facilidade de construção e manutenção, suporte para o desempenho eficiente de algoritmos distribuídos e redução de gastos com meios de transmissão [5]. Em particular, a maior eficiência de redes de controle de processos reduzirá a parcela de custo dos meios de transmissão e permitirá o desempenho adequado de algoritmos descentralizados de identificação e controle.

1.4. OBJETIVOS E ORGANIZAÇÃO DESTE TRABALHO

Os algoritmos desenvolvidos neste trabalho destinam-se a modelos dinâmicos, estocásticos, com parâmetros desconhecidos para sistemas multivariáveis, visando o emprego de técnicas de controle adaptativo. É explorado o problema de estimação de parâme

	DISTÂNCIA	TAXA DE TRANSMISSÃO
Redes Geográficas	$> 10\text{Km}$	$< 0,1\text{Mbits/seg}$
Redes Locais	$0,1\text{Km a } 10\text{Km}$	$0,1\text{Mbits /seg a } 10\text{Mbits/seg}$
Multiprocessadores Concentrados	$< 0,1\text{Km}$	$> 10\text{Mbits/seg}$

Tabela 1.1 - Classificação de Redes de Computação Distribuída segundo a Distância Máxima entre Processadores e a Taxa de Transmissão de Informações entre Processadores

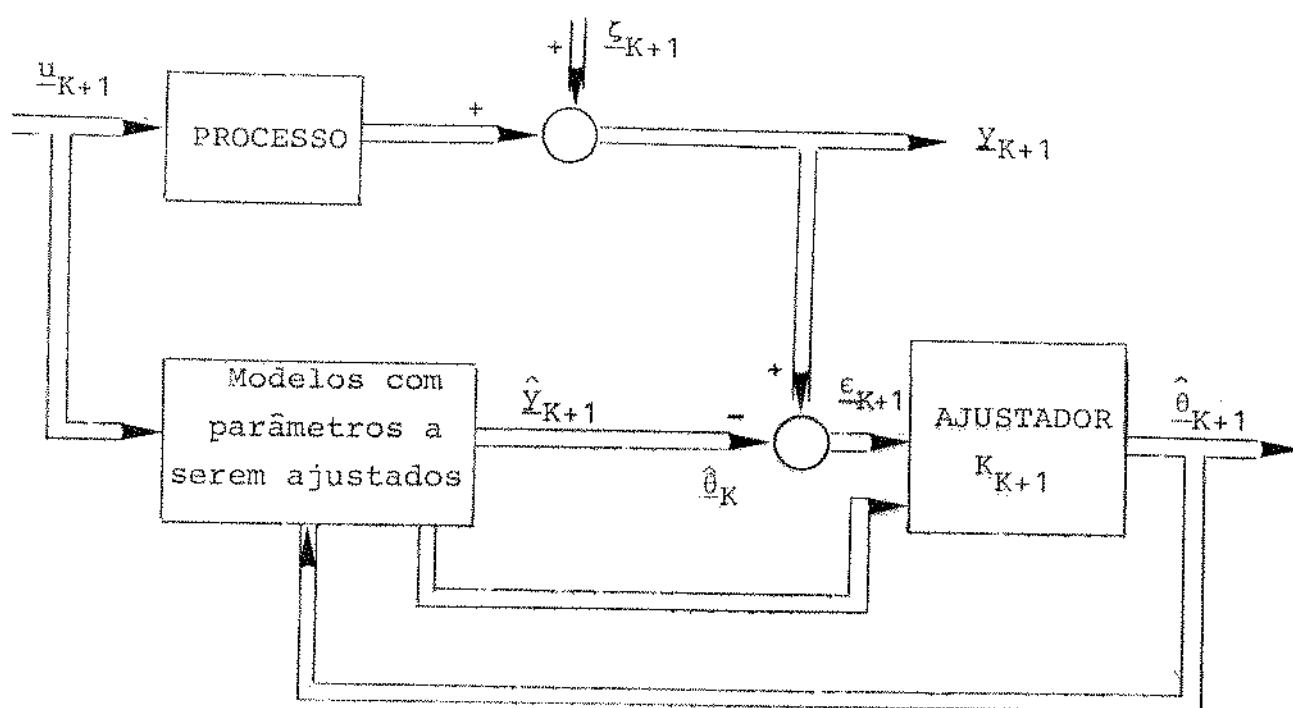


Figura 1.6 - Esquema de Estimação Recursiva

tros recursivamente, levando em conta a distribuição especial do sistema de modo a assegurar o desempenho eficiente dos algoritmos.

A estimação recursiva consiste na correção, a cada instante, do estimador obtido no instante anterior, através de um mecanismo de ajuste. Na figura 1.6 [1], mostra-se um esquema de identificação recursiva, onde a entrada $\underline{u}_K$, devido a sua mensurabilidade, pode ser dirigida para o processo e para o modelo.

Na literatura, o esquema mais comum de identificação recursiva é:

$$\hat{\underline{\theta}}_{K+1} = \hat{\underline{\theta}}_K + K_{K+1} \cdot \underline{\varepsilon}_{K+1} \quad (1.9)$$

onde:

$\hat{\underline{\theta}}_K$ é o vetor dos parâmetros estimados, no instante K

$\underline{\varepsilon}_{K+1}$ é o vetor de erro entre a saída do sistema e a saída do modelo

K_{K+1} é uma matriz, ajustada a cada instante, conforme as medidas realizadas

O Capítulo 2 descreve os métodos encontrados na literatura, para identificação de sistemas multivariáveis, que serviram de base no desenvolvimento dos algoritmos de identificação distribuída do Capítulo 3. A avaliação do desempenho destes algoritmos é realizada no Capítulo 4 e a partir desta avaliação, alguns comentários e sugestões de novos trabalhos são apresentados no Capítulo 5.

CAPÍTULO 2

ALGORÍTMOS CENTRALIZADOS PARA IDENTIFICAÇÃO EM SISTEMAS MULTIVARIÁVEIS

- 2.1. Introdução
- 2.2. Estimador de Máxima Verossimilhança
- 2.3. Métodos de Erro de Previsão
- 2.4. Estimador dos Mínimos Quadrados
- 2.5. Estimadores Recursivos
 - 2.5.1. Mínimos Quadrados Recursivo
 - 2.5.2. Estimador da Raiz Quadrada
 - 2.5.3. Estimadores de Parâmetros Variáveis
no Tempo
 - 2.5.3.1. Janela Exponencial
 - 2.5.3.2. Janela Retangular
 - 2.5.4. Matriz Estendida e Mínimos Quadrados
Particionado
- 2.6. Conclusões

2.1. INTRODUÇÃO

Neste capítulo resume-se os principais métodos para identificação centralizada em sistemas multivariáveis. As referências são alguns livros publicados nas últimas décadas, os congressos da IFAC em Identificação de Sistemas e Estimação de Parâmetros e artigos publicados pelo IEEE em Controle Automático.

Inicialmente, apresenta-se o estimador de máximas verossimilhança. Os demais métodos são introduzidos, através da formulação do erro de previsão, como alternativa caso a informação disponível seja insuficiente para o uso do estimador de máxima verossimilhança.

2.2. ESTIMADOR DE MÁXIMA VEROSSIMILHANÇA

A técnica de máxima verossimilhança é das mais abrangentes para a estimação de parâmetros, sendo aplicável sob diversas condições experimentais a vários tipos de modelos. Apesar de haver sido desenvolvida por Fisher (1912, 1921), algumas de suas idéias básicas são encontradas já nos trabalhos de Gauss (1809). Atualmente, seu uso é justificado por sua simplicidade conceitual e suas propriedades assintóticas [4].

2.2.1. Definição do Estimador de Máxima Verossimilhança

O estimador de máxima verossimilhança assume que os dados gerados pelo sistema possam ser representados por um dos membros de uma família paramétrica de densidades condicionais de probabilidade definida por: $\mathcal{P} = \{p(\cdot|\theta) : \theta \in \Theta\}$, onde $\theta \in \mathbb{R}$. Para uma certa matriz Y , com colunas iguais aos vetores-saída, $p(Y|\theta)$ é uma função de θ e Θ denominada função de máxima verossimilhança. O valor $\hat{\theta}$ que maximiza esta função é o estimador de máxima verossimilhança.

2.2.2. Propriedades do Estimador de Máxima Verossimilhança

Algumas das propriedades demonstradas para o estima
dor de máxima verossimilhança (MV) incluem o princípio de inva
riância e dados sobre um comportamento assintótico. O primeiro
estabelece que o estimador de máxima verossimilhança de uma fun
ção de parâmetros é igual ao valor desta função calculado para
o estimador de máxima verossimilhança dos parâmetros. Assintoti
camente, o estimador de máxima verossimilhança para uma sequên
cia independente de observações tem as seguintes características :

- (a) não-polaridade: converge para o valor real dos pa
râmetros
- (b) normalidade: indicando por N o número de observa
ções e por $\hat{\theta}_N$ o estimador de máxima verossimilhan
ça resultante destas observações

$$\sqrt{N} (\hat{\theta}_N - \theta) \xrightarrow[\text{em}]{\text{propabilidade}} \beta \quad (2.1)$$

$$\beta \sim N(0, \bar{M}_\theta^{-1})$$

onde θ é o valor real dos parâmetros e $\bar{M}_\theta$ é o valor médio da ma
triz de informação de Fisher dada por

$$\bar{M}_\theta = E_{Y|\theta} \left\{ \left[\frac{\partial \log p(Y|\theta)}{\partial \theta} \right]^T \left[\frac{\partial \log p(Y|\theta)}{\partial \theta} \right] \right\} \quad (2.2)$$

- (c) eficiência: numa classe de estimadores consisten
tes e assintoticamente uniformemente normais, o
estimador de máxima verossimilhança tem variância
mínima.

Para o teste de hipóteses, a idéia básica é determi

nar o estimador de máxima verossimilhança sujeito às condições de uma hipótese "nula" bem como de uma hipótese "alternativa", a seguir, calcular o valor da razão de verossimilhança definida como:

$$\lambda(Y) = \frac{p(Y|\hat{\theta}_A)}{p(Y|\hat{\theta}_0)} \quad (2.3)$$

Naturalmente, se o valor desta razão é maior que 1, a hipótese "alternativa" é a que confere maior probabilidade às observações feitas. Caso contrário, a mais favorável é a hipótese "nula".

Em particular, se as observações feitas são independentes, a hipótese nula consiste nos parâmetros satisfazerem S restrições localmente independentes dadas por; $\ell_i(\theta) = 0$, $i = 1, \dots, S$, e a hipótese alternativa não estabelece nenhuma restrição, então a razão de verossimilhança $\lambda(Y)$ é tal que a variável aleatória $2\log \lambda(Y)$ converge em probabilidade para uma variável com distribuição qui-quadrado de S graus de liberdade.

Com base no resultado do parágrafo anterior, estabelece-se que a hipótese nula pode ser aceita num nível α de significância se $2\log \lambda(Y) < K_\alpha$, onde K_α é tal que 100(1- α)% da distribuição qui-quadrado com S graus de liberdade se encontra à esquerda de K_α .

Nesta seção mencionou-se de forma sucinta a definição e as propriedades mais importantes do estimador de máxima verossimilhança. Comentários mais detalhados podem ser encontrados nas referências [4],[21]. Algumas indicações das dificuldades encontradas nas aplicações do estimador de máxima verossimilhança são feitas a seguir.

2.3. MÉTODOS DO ERRO DE PREVISÃO

O estimador de máxima verossimilhança requer o acúmu

lo de informações a respeito do processo em estudo para o cálculo das densidades condicionais de probabilidade das saídas. Como em muitas situações, o esforço computacional envolvido é grande, recorre-se a reformulação do problema de estimação de modo a simplificar os dados estatísticos necessários. A seguir realiza-se esta aproximação para o estimador de máxima verossimilhança com observações sequenciais.

Para observações sequenciais, se o vetor-saída no instante t , $\underline{y}_t$ tem densidade condicional $p(\underline{y}_t | \underline{y}_{t-1}, \underline{\theta})$, onde $\underline{y}_{t-1}$ é a matriz cujas colunas são os vetores-saída $\underline{y}_1, \dots, \underline{y}_{t-1}$, então a densidade conjunta de N observações pode ser calculada aplicando-se a regra de Bayes:

$$p(\underline{Y}_N | \underline{\theta}) = p(\underline{y}_N | \underline{y}_{N-1}, \underline{\theta}) \dots p(\underline{y}_2 | \underline{y}_1, \underline{\theta}) \cdot p(\underline{y}_1 | \underline{\theta}) \quad (2.4)$$

Considerando ainda que as densidades condicionais são gaussianas e dadas por:

$$p(\underline{y}_t | \underline{y}_{t-1}, \underline{\theta}) = [(2\pi)^n \cdot \det V_t]^{-1/2} \exp\{-1/2(\underline{y}_t - \underline{m}_t)^T V_t^{-1} (\underline{y}_t - \underline{m}_t)\} \quad (2.5)$$

onde n é a dimensão de $\underline{y}_t$, o vetor $\underline{m}_t$ é a média condicional $\underline{m}_t = E\{\underline{y}_t | \underline{y}_{t-1}, \underline{\theta}\}$ e a matriz V_t é a covariância condicional $V_t = \text{Cov}\{\underline{y}_t | \underline{y}_{t-1}, \underline{\theta}\}$. A média condicional e a covariância condicional dependem do vetor de parâmetros desconhecidos $\underline{\theta}$.

Substituindo a equação (2.5) na equação (2.4) obtém-se a função de verossimilhança:

$$p(\underline{Y}_N | \underline{\theta}) = \prod_{i=1}^N \{ [(2\pi)^n \det V_i]^{-1/2} \cdot \exp[-1/2(\underline{y}_i - \underline{m}_i)^T V_i^{-1} (\underline{y}_i - \underline{m}_i)] \} \quad (2.6)$$

E, aplicando o logarítmo na função de máxima verossimilhança:

$$\begin{aligned}
 -\ell_n p(Y_N | \theta) &= \frac{1}{2} \cdot n \cdot N \ell_n(2\pi) + \\
 &+ \frac{1}{2} \sum_{i=1}^N \ln \det V_i + \frac{1}{2} \sum_{i=1}^N \underline{\varepsilon}_i^T V_i^{-1} \underline{\varepsilon}_i
 \end{aligned} \quad (2.7)$$

com

$$\underline{\varepsilon}_i = Y_i - \underline{m}_i \quad (2.8)$$

Portanto, no caso de observações sequenciais, com densidades condicionais de probabilidade gaussianas, a maximização da função de verossimilhança é equivalente à minimização da expressão em (2.7).

Se a covariância condicional V_i é conhecida, o problema reduz-se a minimização do critério:

$$J_1 = \frac{1}{2} \sum_{i=1}^N \underline{\varepsilon}_i^T V_i \underline{\varepsilon}_i \quad (2.9)$$

Se, ao contrário, V_i é também desconhecida, é necessário realizar a maximização da função de verossimilhança em relação a V_i e ao vetor de parâmetros desconhecidos.

A maximização é simplificada quando se considera o processo estacionário, ou seja, supõe-se a matriz de covariância constante e igual a V . Isto permite reescrever a equação (2.7) como:

$$\begin{aligned}
 -\ell_n p(Y_N | \theta) &= \frac{1}{2} \cdot n \cdot N \ell_n(2\pi) + \\
 &+ \frac{1}{2} N \cdot \ell_n \det V + \frac{1}{2} \sum_{i=1}^N \underline{\varepsilon}_i^T V^{-1} \underline{\varepsilon}_i
 \end{aligned} \quad (2.10)$$

que, em relação a V , apresenta um mínimo para:

$$V = \frac{1}{N} \sum_{i=1}^N \underline{\varepsilon}_i \underline{\varepsilon}_i^T \quad (2.11)$$

Substituindo este valor na equação (2.10), conclui-se que se a covariância é desconhecida, mas constante a maximização da função de verossimilhança é equivalente à minimização do critério:

$$J_2 = \frac{1}{2} \ell_n \det \left\{ \frac{1}{N} \sum_{i=1}^N \underline{\varepsilon}_i \underline{\varepsilon}_i^T \right\} \quad (2.12)$$

Nas equações anteriores, a média condicional $\underline{m}_t = E\{\underline{y}_t | \underline{y}_{t-1}, \theta\}$ pode ser interpretada como uma previsão do valor de $\underline{y}_t$ baseada nas saídas anteriores em $\underline{y}_{t-1}$. Seguindo o mesmo raciocínio $\underline{\varepsilon}_t$ é o erro de previsão e os critérios a minimizar J_1 e J_2 são funções quadráticas dos erros de previsão. Esta interpretação permite uma forma alternativa de resolver o problema de estimação.

Definido o critério a minimizar, J_1 ou J_2 , a estimação dos parâmetros desconhecidos torna-se um problema de otimização. Não são requeridas novas informações probabilísticas. Isto sugere uma técnica simples de dispensar as informações probabilísticas iniciais: é suficiente postular um critério, função do erro de previsão, e um modelo relacione o erro de previsão, os parâmetros e os dados.

Os métodos que traduzem o procedimento mencionado são denominados métodos do erro de previsão. Usualmente, quando a previsão da saída é feita para apenas um período de amostragem além dos dados obtidos, os critérios adotados são do mesmo tipo que J_1 ou J_2 . Para situações em que a previsão da saída deve abranger vários períodos de amostragem utiliza-se um critério como o descrito na equação (2.13).

$$J_3 = \frac{1}{2} \sum_{i=0}^N \sum_{K=0}^r \underline{\varepsilon}^T(t_{i+K} | t_{i-1}) \cdot V^{-1}(t_{i+K} | t_{i-1}) \cdot \underline{\varepsilon}(t_{i+K} | t_{i-1}) + 1/2 \sum_{i=0}^N \sum_{K=0}^r \ell_n \det V(t_{K+1} | t_{i-1}) \quad (2.13)$$

Os métodos do erro de previsão requerem o modelamento das relações entre o erro de previsão, os dados e os parâmetros. Estes são denominados modelos do erro previsto. Em se tratando de previsão da saída para um período de amostragem seguinte ao das medidas realizadas, os modelos de erro previsto tem a forma descrita na equação (2.14).

$$\underline{y}_K = \underline{f}(\underline{y}_{K-1}, \underline{u}_{K-1}, \underline{\theta}) + \underline{w}_K \quad (2.14)$$

onde: $\underline{y}_K$ é o vetor saída no instante t_K ; as matrizes $\underline{y}_{K-1}$ e $\underline{u}_{K-1}$ são dadas por $\underline{y}_{K-1} \triangleq \{\underline{y}_1, \dots, \underline{y}_{K-1}\}$ e $\underline{u}_{K-1} \triangleq \{\underline{u}_1, \dots, \underline{u}_{K-1}\}$, tendo colunas iguais às saídas e às entradas do sistema até t_{K-1} , respectivamente; $\underline{\theta}$ é o vetor que contém os parâmetros desconhecidos e $\{\underline{w}_K\}$ é uma sequência aleatória caracterizada por:

$$E\{\underline{w}_K / \underline{y}_{K-1}, \underline{u}_{K-1}\} = 0 \quad (2.15)$$

O processo $\{\underline{w}_K\}$ é denominado nova informação ou processo de inovação de $\{\underline{y}_K\}$ [2.1].

Para o modelo (2.14) a função $\underline{f}(\dots, \underline{\theta})$ representa a parte determinística no comportamento do processo, enquanto a parte estocástica corresponde à sequência $\{\underline{w}_K\}$.

Com este modelo, o erro de previsão no instante t_K é dado por

$$\underline{\varepsilon}_K \triangleq \underline{y}_K - \underline{f}(\underline{y}_{K-1}, \underline{u}_{K-1}, \hat{\underline{\theta}}_K) = \underline{y}_K - \hat{\underline{y}}_K \quad (2.16)$$

onde $\hat{\underline{\theta}}_K$ indica o valor estimado dos parâmetros em t_K e $\hat{\underline{y}}_K \triangleq \underline{f}(\underline{y}_{K-1}, \underline{u}_{K-1}, \hat{\underline{\theta}}_K)$ é o valor previsto da saída em t_K .

Uma das propriedades do modelo do erro de previsão consiste em colocar de forma explícita o valor previsto para a saída, o que é essencial na maioria dos algoritmos de controle estocástico [1].

Os modelos para sistemas estocásticos descritos no Ca

pítulo 1 podem ser reescritos como modelos de erro de previsão através de simples manipulações matemáticas. Estas transformações serão realizadas na medida em que se tornem necessárias para a exposição dos algoritmos.

Nas próximas seções são expostos vários algoritmos que adotam a formulação do erro previsto.

2.4. ESTIMADOR DOS MÍNIMOS QUADRADOS

Examinando o critério J_1 , descrito na equação (2.9) e reescrito a seguir:

$$J_1 = \frac{1}{2} \sum_{i=1}^N \epsilon_i^T V_i \epsilon_i \quad (2.17)$$

observa-se que se este critério for quadrático nos parâmetros desconhecidos, sua derivada será linear nestes parâmetros e uma equação explícita para o estimador poderá ser facilmente obtida. Para se obter estas condições, o erro de previsão deve ser uma função linear dos parâmetros.

Um modelo com as características acima mencionadas é empregado pelo estimador dos mínimos quadrados (M.Q.). Este modelo muitas vezes referido como modelo de mínimos quadrados é dado por:

$$A(q^{-1}) \cdot \underline{y}_K = B(q^{-1}) \underline{u}_K + \underline{w}_K \quad (2.18)$$

onde: $A(q^{-1}) = I + A_1 q^{-1} + \dots + A_{N_*} q^{-N_*}$

$$B(q^{-1}) = B_1 q^{-1} + \dots + B_{N_*} q^{-N_*}$$

$\underline{y}_K \in \mathbb{R}^n$ é o vetor-saída, $\underline{u}_K \in \mathbb{R}^m$ é o vetor-entrada e $\{\underline{w}_K\}$ é uma sequência branca de variáveis aleatórias de covariância V (constante).

O modelo descrito pela equação (2.18) também pode ser

reescrito de forma a agrupar num vetor $\underline{\theta}$ os coeficientes das matrizes $A(q^{-1})$ e $B(q^{-1})$ a serem estimadas, obtendo-se:

$$\underline{Y}_K = \psi_K \cdot \underline{\theta}^{A,B} + \underline{W}_K \quad (2.19)$$

onde $\psi_K \triangleq$ bloco diagonal $[\underline{\phi}_K^T, \dots, \underline{\phi}_K^T]$

$$\psi_K \in \mathbb{R}^{(n \times np)}$$

com $\underline{\phi}_K^T \triangleq [-\underline{y}_{K-1}^T \mid \dots \mid \underline{y}_{K-N_*}^T \mid \underline{u}_{K-1}^T \mid \dots \mid \underline{u}_{K-N_*}^T]$

$$\underline{\phi}_K \in \mathbb{R}^{(p)}$$

e $\underline{\theta}^{A,B} \triangleq [\underline{a}^T \mid \underline{b}^T \mid \dots \mid \underline{a}^T \mid \underline{b}^T]$

$$\underline{\theta} \in \mathbb{R}^{(np)}$$

com $\underline{a}^{iT} \triangleq$ i-ésima linha de $[A(q^{-1}) - I]$

$\underline{b}^{iT} \triangleq$ i-ésima linha de $B(q^{-1})$

$\rho \triangleq N_* (n + m)$

Usando o modelo (2.19) a equação do erro de previsão é:

$$\underline{\varepsilon}_K(\hat{\underline{\theta}}^{A,B}) = \underline{Y}_K - \psi_K \hat{\underline{\theta}}^{A,B} \quad (2.20)$$

Minimizando o critério J_1 com o valor de $\underline{\varepsilon}_K$, dado pela equação (2.20), obtêm-se o estimador de mínimos quadrados:

$$\hat{\underline{\theta}}^{A,B} = \left[\sum_{K=1}^N \psi_K^T \cdot V^{-1} \psi_K \right]^{-1} \left[\sum_{K=1}^N \psi_K^T V^{-1} \underline{Y}_K \right] \quad (2.21)$$

As origens deste estimador remontam ao século XIX em estudos independentes efetuados por Gauss e Legendre. Recentemente, mostrou-se [1] que é possível estimar separadamente os parâmetros $\underline{\theta}_i^{A,B} = [\underline{a}_i^T, \underline{b}_i^T]$ de cada linha do modelo (2.19). Esta prova é realizada utilizando-se as propriedades do produto de

Kronecker de matrizes [1] apresentadas no Apêndice A.

Empregando a identidade (A.2) pode-se reescrever a equação (2.20) como:

$$\hat{\underline{\theta}}^{A,B} = \left[\begin{array}{c} N \\ \Sigma \\ K=1 \end{array} V^{-1} \otimes I_P \psi_K^T \psi_K \right]^{-1} \left[\begin{array}{c} N \\ \Sigma \\ K=1 \end{array} V^{-1} \otimes I_P \psi_K^T y_K \right] \quad (2.22)$$

Como o ruído branco é estacionário, a matriz $V^{-1} \otimes I_P$ é constante e obtém-se:

$$\hat{\underline{\theta}}^{A,B} = \left[\begin{array}{c} N \\ \Sigma \\ K=1 \end{array} \psi_K^T \psi_K \right]^{-1} [V^{-1} \otimes I_P]^{-1} [V^{-1} \otimes I_P] \left[\begin{array}{c} N \\ \Sigma \\ K=1 \end{array} \psi_K^T y_K \right] \quad (2.23)$$

ou

$$\hat{\underline{\theta}}^{A,B} = \left[\begin{array}{c} N \\ \Sigma \\ K=1 \end{array} \psi_K^T \psi_K \right]^{-1} \left[\begin{array}{c} N \\ \Sigma \\ K=1 \end{array} \psi_K^T y_K \right] \quad (2.24)$$

Aplicando a identidade (A.4), visto que ψ_t tem estrutura bloco diagonal, a equação (2.24) pode ser separada em n equações independentes, uma para cada linha do modelo:

$$\hat{\underline{\theta}}_i^{A,B} = \left[\begin{array}{c} N \\ \Sigma \\ K=1 \end{array} \phi_K \phi_K^T \right]^{-1} \left[\begin{array}{c} N \\ \Sigma \\ K=1 \end{array} \phi_K y_K^i \right] \quad i = 1, \dots, n \quad (2.25)$$

onde y_t^i é a i -ésima componente do vetor-saída.

Dois aspectos importantes no caso de ruído estacionário são:

- i) o estimador é independente de V
- ii) o estimador descrito pela equação (2.25) simplifica bastante os cálculos tanto devido à redução da ordem da matriz a ser invertida como devido à possibilidade de cálculo paralelo dos vetores $\hat{\underline{\theta}}_i^{A,B}$.

A matriz de covariância do estimador de mínimos quadrados é dada por:

$$\text{Cov}(\hat{\underline{\theta}}^{A,B}) = E\{(\hat{\underline{\theta}}^{A,B} - \underline{\theta}^{A,B})(\hat{\underline{\theta}}^{A,B} - \underline{\theta}^{A,B})^T\} \quad (2.26)$$

onde $\underline{\theta}^{A,B}$ é o valor verdadeiro dos parâmetros.

Substituindo nesta equação, as equações (2.19) e (2.26), tem-se:

$$\begin{aligned} \text{Cov}(\hat{\underline{\theta}}^{A,B}) &= \left[\begin{array}{cc} N & \\ \sum_{K=1} & \psi_K^T \psi_K \end{array} \right]^{-1} \cdot \left[\begin{array}{cc} N & \\ \sum_{K=1} & \psi_K^T V \psi_K \end{array} \right] \cdot \\ &\quad \cdot \left[\begin{array}{cc} N & \\ \sum_{t=1} & \psi_K^T \psi_K \end{array} \right]^{-1} \end{aligned} \quad (2.27)$$

Na equação (2.27) da covariância do estimador aparece a covariância V do ruído estacionário. Isto está de acordo com o sentimento intuitivo de que a precisão do estimador está relacionada com a "amplitude" do ruído.

Auxiliando durante longo tempo a identificação de sistemas, o estimador de mínimos quadrados já foi objeto de muitas pesquisas e inúmeras propriedades importantes deste estimador foram obtidas. Em particular, para um sistema descrito pelo modelo (2.18) com $\{\underline{w}_K\}$ uma sequência branca e estacionária de covariância V , o estimador de mínimos quadrados é não polarizado e tem variância mínima entre os estimadores lineares não polarizados. Entretanto, quando o ruído presente não é uma sequência branca o estimador de mínimos quadrados não converge para os valores verdadeiros dos parâmetros. Este problema é resolvido modelando-se as perturbações a partir de uma sequência branca. Estimadores que utilizam esta técnica serão apresentados após a descrição do estimador recursivo dos mínimos quadrados.

2.5. ESTIMADORES RECURSIVOS

Os estimadores apresentados nesta seção podem ser computados em tempo real, devido à redução dos cálculos em cada e

tapa. Esta redução viabiliza a aplicação de pequenos computadores para estimação em tempo real de parâmetros variáveis no tempo, necessária nos esquemas de controle adaptativo.

2.5.1. Mínimos Quadrados Recursivo

Conforme estabelecido na seção anterior, o estimador dos parâmetros do modelo:

$$Y_K = \psi_K \cdot \underline{\theta}^{A,B} + \underline{W}_K \quad (2.28)$$

escolhido de forma a minimizar o critério:

$$J_1 = \frac{1}{2} \sum_{K=1}^N (Y_K - \psi_K \cdot \underline{\theta}^{A,B})^T \cdot V^{-1} (Y_K - \psi_K \cdot \underline{\theta}^{A,B}) \quad (2.29)$$

é o estimador de mínimos quadrados dado por:

$$\hat{\underline{\theta}}_{-N}^{A,B} = \left[\begin{array}{c} N \\ \sum \\ K=1 \end{array} \psi_K^T \cdot \psi_K \right]^{-1} \cdot \left[\begin{array}{c} N \\ \sum \\ K=1 \end{array} \psi_K^T \cdot Y_K \right] \quad (2.30)$$

onde o sub-índice N indica o número de observações utilizadas no cálculo do estimador.

Para se obter a versão recursiva, considera-se a situação em que mais uma observação, entrada-saída, é realizado no instante t_{N+1} . Definindo-se:

$$P_N = \left[\begin{array}{c} N \\ \sum \\ K=1 \end{array} \psi_K^T \cdot \psi_K \right]^{-1} \quad (2.31)$$

tem-se:

$$P_{N+1} = \left[\begin{array}{c} N+1 \\ \sum \\ K=1 \end{array} \psi_K^T \cdot \psi_K \right]^{-1} = [P_N^{-1} + \psi_{N+1}^T \cdot \psi_{N+1}]^{-1} \quad (2.32)$$

Aplicando-se a esta equação o lema de inversão matricial (Identidade A.1 - Apêndice A), tem-se:

$$P_{N+1} = P_N - P_N \psi_{N+1}^T [I + \psi_{N+1} P_N \psi_{N+1}^T]^{-1} \psi_{N+1} \cdot P_N \quad (2.33)$$

E, substituindo a equação (2.33) na equação (2.30):

$$\hat{\theta}_{N+1}^{A,B} = \hat{\theta}_N^{A,B} + K_{N+1} \cdot \{y_{N+1} - \psi_{N+1} \cdot \hat{\theta}_N^{A,B}\} \quad (2.34)$$

onde

$$K_{N+1} = P_{N+1} \cdot \psi_{N+1}^T \quad (2.35)$$

As equações (2.33), (2.34) e (2.35) constituem o estimador recursivo de mínimos quadrados. Na literatura, a matriz inicial P_0 é obtida a partir de

$$\sum_{K=-(N_0-1)}^0 \psi_K^T \psi_K$$

ou, se estes valores não são conhecidos, assume-se que $P_0 = I/\delta$, com δ igual a $(10^{-1} - 10^{-3})$.

É fácil ver como a matriz P_N está relacionada com a covariância do estimador. Da equação (2.27) tem-se:

$$\text{Cov}(\hat{\theta}_N^{A,B}) = \left[\sum_{K=1}^N \psi_K^T \psi_K \right]^{-1} \cdot \left[\sum_{K=1}^N \psi_K^T \cdot v \psi_K \right] \cdot \left[\sum_{K=1}^N \psi_K^T \psi_K \right]$$

que, com o uso das identidades A.2 e A.6 do Apêndice A pode ser reescrita como:

$$\text{Cov}(\hat{\theta}_N^{A,B}) = \left[\sum_{K=1}^N \psi_K^T \psi_K \right]^{-1} (V \otimes I_p) \cdot \left[\sum_{K=1}^N \psi_K^T \psi_K \right] \left[\sum_{K=1}^N \psi_K^T \psi_K \right]^{-1}$$

$$\text{ou} \quad \text{Cov}(\hat{\theta}_N^{A,B}) = \left[\sum_{K=1}^N \psi_K^T \psi_K \right]^{-1} (V \otimes I_p) \quad (2.36)$$

daí

$$\text{Cov}(\hat{\theta}_{-N}^{A,B}) = (V \otimes I_p) \begin{bmatrix} N & & \\ \Sigma & \psi_K^T & \psi_K \\ K=1 & & \end{bmatrix}^{-1} = (V \otimes I_p) P_N \quad (2.37)$$

Portanto, a matriz inicial P_0 está relacionada com a incerteza dos valores iniciais $\hat{\theta}_0^{A,B}$ atribuídos aos parâmetros.

Analogamente, a forma recursiva da versão (2.25):

$$\hat{\theta}_{-i}^{A,B} = \begin{bmatrix} N \\ \Sigma \\ K=1 \end{bmatrix} \phi_K \cdot \phi_K^T \begin{bmatrix} K \\ \Sigma \\ K=1 \end{bmatrix} \phi_K \cdot y_K^i \quad i = 1, \dots, N$$

que calcula separadamente os valores estimados dos parâmetros de cada linha do modelo (2.18) é dada por:

$$P_{i,N+1} = P_{i,N} - \frac{P_{i,N} \phi_{N+1} \cdot \phi_{N+1}^T \cdot P_{i,N}}{1 + \phi_{N+1}^T P_{i,N+1} \phi_{N+1}} \quad (2.38)$$

$$\hat{\theta}_{-i,N+1}^{A,B} = \hat{\theta}_{i,N}^{A,B} + K_{i,N+1} \left\{ y_{N+1}^i - \phi_{N+1}^T \hat{\theta}_{-i,N}^{A,B} \right\} \quad (2.39)$$

$$K_{i,N+1} = P_{i,N+1} \cdot \phi_{N+1}^T \quad i = 1, \dots, n \quad (2.40)$$

Esta forma, simplifica os cálculos ao substituir a in versão matricial por uma divisão escalar. Maior simplificação ainda é conseguida se matrizes iniciais $P_{i,0}$ são as mesmas para $i = 1, \dots, n$, pois neste caso as equações (2.38) e (2.40) necessitam ser calculadas apenas uma vez para atualização do algoritmo.

A relação de $P_{i,N}$ com a matriz de covariância pode ser facilmente deduzida. Observando-se que:

$$P_N = \text{bloco diagonal} \left[P_{1,N} \mid \dots \mid P_{n,N} \right] \quad (2.41)$$

e substituindo-se na equação (2.37), obtém-se:

$$\text{Cov}(\hat{\theta}_{-N}^{A,B}) = (V \otimes I_p) \cdot \text{bloco diagonal}[P_{1,N} \mid \dots \mid P_{n,N}] \quad (2.42)$$

2.5.2. Estimador da Raiz Quadrada

Existe ainda outra forma recursiva, mais rápida e numericamente mais robusta, para as equações (2.25). Esta forma é usualmente referida como estimador da raiz quadrada, sendo dada por:

$$\hat{\theta}_{i,N+1}^{A,B} = \mathcal{B}_{i,N+1}^{-1} \cdot \underline{\Omega}_{i,N+1} \quad (2.43)$$

onde $\mathcal{B}_{N+1}^i$, uma matriz triangular superior, e $\underline{\Omega}_{N+1} \in \mathbb{R}^p$, são obtidos de:

$$\begin{bmatrix} \mathcal{B}_{i,N+1} \\ 0 \end{bmatrix} = X_N \cdot \begin{bmatrix} \mathcal{B}_{i,N} \\ \phi_{N+1}^T \end{bmatrix} \quad (2.44)$$

sendo X_N uma matriz ortogonal escrita como:

$$X_N = Q_p \cdot Q_{p-1} \cdots Q_1 \quad (2.45)$$

onde

$$Q_j = \left[\begin{array}{c|cc|c} I_{j-1} & & & 0 \\ \hline & c_j & 0 & s_j \\ \hline 0 & 0 & I_{p-j} & 0 \\ \hline & s_j & 0 & c_j \end{array} \right] \quad j = 1, \dots, p \quad (2.46)$$

$$\text{com} \quad c_j = (\mathcal{B}_{i,N})_{jj} / [(\mathcal{B}_{i,N})_{jj}^2 + (\phi_{N+1})_j^2]^{1/2} \quad (2.47)$$

$$s_j = (\phi_{N+1})_j / [(\mathcal{B}_{i,N})_{jj}^2 + (\phi_{N+1})_j^2]^{1/2} \quad (2.48)$$

$$\text{e} \quad \begin{bmatrix} \underline{\Omega}_{N+1} \\ h_{N+1} \end{bmatrix} = X_N \begin{bmatrix} \underline{\Omega}_N \\ y_{N+1}^i \end{bmatrix} \quad (2.49)$$

Nesta forma, se as matrizes iniciais $\mathbb{B}_0^i$ são escolhidas iguais para todas as linhas, então as equações (2.44), (2.45), (2.46), (2.47) e (2.48) só precisam ser calculadas uma vez na atualização do algoritmo.

A relação entre esta forma recursiva e a do estimador dos mínimos quadrados é estabelecida através de:

$$P_{i,N} = (\mathbb{B}_{i,N}^T \cdot \mathbb{B}_{i,N})^{-1} \quad (2.50)$$

2.5.3. Estimadores de Parâmetros Variáveis no Tempo

Para a estimação de parâmetros variáveis no tempo os critérios considerados devem dar maior importância às observações mais recentes. Naturalmente, isto pode ser feito de várias formas. A seguir são descritos dois algoritmos. Ambos empregam o modelo (2.19) e estão relacionados com o método dos mínimos quadrados.

2.5.3.1. Janela Exponencial

Neste caso, a função a ser minimizada pondera exponencialmente o erro de previsão, sendo descrita por:

$$J_4 = \sum_{K=1}^N \delta^{N-K} (\underline{y}_K - \psi_K \cdot \underline{\theta}^{A,B})^T V^{-1} (\underline{y}_K - \psi_K \cdot \underline{\theta}^{A,B}) \quad (2.51)$$

As equações recursivas para o estimador satisfazendo este critério são:

$$P_{N+1} = 1/\delta P_N - 1/\delta P_N \psi_{N+1}^T [\delta I + \psi_{N+1} P_N \psi_{N+1}^T]^{-1} \psi_{N+1} P_N \quad (2.52)$$

$$K_{N+1} = P_{N+1} \psi_{N+1}^T \quad (2.53)$$

$$\hat{\theta}_{-N+1}^{A,B} = \hat{\theta}_{-N}^{A,B} + K_{N+1} (\underline{y}_{N+1} - \psi_{N+1} \hat{\theta}_{-N}^{A,B}) \quad (2.54)$$

Para estimação linha por linha do modelo (2.19), as equações recursivas tomam a forma seguinte:

$$P_{i,N+1} = 1/\delta P_{i,N} - 1/\delta \frac{P_{i,N} \phi_{N+1} \phi_{N+1}^T P_{i,N}}{\delta + \phi_{N+1}^T P_{i,N} \phi_{N+1}} \quad (2.55)$$

$$K_{i,N+1} = P_{i,N+1} \cdot \phi_{N+1} \quad (2.56)$$

$$\hat{\theta}_{-i,N+1}^{A,B} = \hat{\theta}_{-i,N}^{A,B} + K_{i,N+1} \cdot (\underline{y}_{N+1}^i - \phi_{N+1}^T \hat{\theta}_{-i,N}^{A,B}) \quad i = 1, \dots, n \quad (2.57)$$

2.5.3.2. Janela Retangular

A característica essencial dos algoritmos do tipo janela retangular é que a estimação, num certo instante, é realizada somente com um número limitado de observações. As demais observações são descartadas [21].

Para uma janela retangular de comprimento fixo N, toda vez que uma nova observação é realizada, a observação mais antiga é descartada, de modo a não alterar o comprimento da janela. No instante $n+N$ o procedimento a ser efetuado consta de dois passos. Em primeiro lugar, a observação $\underline{y}_{r+N}$, $\underline{u}_{r+N}$ é acrescentada através da atualização, realizada pelas equações (2.33), (2.34) e (2.35), ou seja:

$$P_{r+N}^r = P_{r+N-1}^r - P_{r+N-1}^r \cdot \psi_{r+N}^T [I + \psi_{r+N} \cdot P_{r+N-1}^r \cdot \psi_{r+N}^T]^{-1} \psi_{r+N-1} \cdot P_{r+N-1}^r \quad (2.58)$$

$$\hat{\theta}_{-r,N}^{A,B} = \hat{\theta}_{-r+N-1}^{A,B} + K_{r+N}^r [\underline{y}_{r+N} - \psi_{r+N} \hat{\theta}_{-r+N-1}^{A,B}] \quad (2.59)$$

$$\overset{r}{K}_{r+N} = \overset{r}{P}_{r+N} \cdot \psi_{r+N}^T \quad (2.60)$$

onde o índice r sobre $\hat{\theta}$, P e K indica que estes valores são calculados tomando por base observações feitas entre r e $r+N$.

No segundo passo, a observação y_r , u_r realizada no instante r é descartada usando as equações:

$$\overset{r+1}{P}_{r+N} = \overset{r}{P}_{r+N} + \overset{r}{P}_{r+N} \psi_r^T [I + \psi_r \overset{r}{P}_{r+N} \cdot \psi_r^T]^{-1} \psi_r \cdot \overset{r}{P}_{r+N} \quad (2.61)$$

$$\overset{r+1}{\hat{\theta}}_{r+N}^{A,B} = \overset{r}{\hat{\theta}}_{r+N}^{A,B} - K_{r+N}^{r+1} [y_r - \psi_r \cdot \overset{r}{\hat{\theta}}_{r+N}^{A,B}] \quad (2.62)$$

$$\overset{r+1}{K}_{r+N} = \overset{r+1}{P}_{r+N} \cdot \psi_r^T \quad (2.63)$$

onde o índice $r+1$ sobre $\hat{\theta}$, P e K , no lado esquerdo das equações (2.61), (2.62) e (2.63) indica que são consideradas somente as informações entre $r+1$ e $r+N$.

Existe para o algoritmo acima, uma versão que considera isoladamente os parâmetros de cada uma das linhas do modelo (2.19). Nesta versão tem-se:

- primeiro passo:

$$\overset{r}{P}_{i,r+N} = \overset{r}{P}_{i,r+N-1} - \frac{\overset{r}{P}_{i,r+N-1} \cdot \phi_{r+N}^T \phi_{r+N} \overset{r}{P}_{i,r+N-1}}{1 + \phi_{r+N}^T \overset{r}{P}_{i,r+N-1} \cdot \phi_{r+N}} \quad (2.64)$$

$$\overset{r}{\hat{\theta}}_{i,r+N}^{A,B} = \overset{r}{\hat{\theta}}_{i,r+N-1}^{A,B} + K_{i,r+N} [y_{r+N}^i - \phi_{r+N}^T \cdot \overset{r}{\hat{\theta}}_{i,r+N-1}^{A,B}] \quad (2.65)$$

$$\overset{r}{K}_{i,r+N} = \overset{r}{P}_{i,r+N} \cdot \phi_{r+N} \quad (2.66)$$

- segundo passo:

$${}^{r+1}P_{i,r+N} = {}^rP_{i,r+N} + \frac{{}^rP_{i,r+N} \phi_r \phi_r^T {}^rP_{i,r+N}}{1 - \phi_r^T {}^rP_{i,r+N} \phi_r} \quad (2.67)$$

$$\hat{\theta}_{i,r+N}^{A,B} = \hat{\theta}_{i,r+N}^r - K_{r+N}^{r+1} [y_r^i - \phi_r^T \cdot \hat{\theta}_{i,r+N}^{A,B}] \quad (2.68)$$

$${}^{r+1}K_{i,r+N} = {}^{r+1}P_{i,r+N} \cdot \phi_r \quad (2.69)$$

$i = 1, \dots, n$

As equações acima podem ser deduzidas usando o lema de inversão matricial do Apêndice A.

No algoritmo acima os últimos N vetores ϕ_k devem ser armazenados. Isto pode constituir um problema para valores elevados de N . Uma opção consiste em utilizar-se um algoritmo de memória oscilante que descarta as observações mais antigas em blocos de N valores. A memória deste algoritmo varia entre N e $2N$. Para adicionar uma nova observação as equações (2.58), (2.59) e (2.60) são suficientes. Quando o número de observações em uso atinge $2N$, as N observações mais antigas são descartadas através de:

$$\hat{\theta}_{-r+2N}^{A,B} = P_{r+2N}^{r+N+1} \left[P_{r+2N}^{r+1} \hat{\theta}_{-r+2N}^{A,B} - P_{r+N}^{r+1} \hat{\theta}_{r+N}^{A,B} \right] \quad (2.70)$$

$$P_{r+N+1}^{r+2N} = P_{r+2N}^{r+1} - P_{r+N}^{r+1} \quad (2.71)$$

Caso se deseje considerar separadamente cada uma das linhas do modelo (2.19), as equações (2.70) e (2.71) são ainda válidas, mas a adição de novas observações deve ser realizada pelas equações (2.64), (2.65) e (2.66).

Das últimas equações segue que para descartar um bloco de N observações basta armazenar o vetor

$$\begin{matrix} r+N \\ \hat{\theta}_{A,B} \\ -r+1 \end{matrix}$$

e a matriz

$$\begin{matrix} r+N \\ P \\ r+1 \end{matrix}.$$

A dedução detalhada das equações desta seção é feita no Apêndice B.

2.5.4. Matriz Estendida e Mínimos Quadrados Particionado

Nos algoritmos apresentados, existe a hipótese básica de que a parcela aleatória da saída pode ser modelada através do ruído branco. Quando esta suposição não é válida, estes algoritmos conduzem a estimações polarizadas dos parâmetros do sistema.

Vários procedimentos vêm sendo elaborados para superar esta dificuldade [1],[21]. Dois métodos que utilizam um modelo auto-regressivo e média móvel para o ruído na saída do sistema são descritos nesta seção.

O modelo utilizado nestes algoritmos é descrito por:

$$A(q^{-1}) \cdot Y_K = B(q^{-1}) \cdot u_K + \zeta_K \quad (2.72)$$

$$D(q^{-1}) \cdot L_K = C(q^{-1}) \cdot W_K$$

A partir deste modelo, o erro de previsão pode ser descrito como:

$$\varepsilon_K(\hat{\theta}) = \hat{C}^{-1}(q^{-1}) \cdot \hat{D}(q^{-1}) \{ \hat{A}(q^{-1}) \cdot Y_K - \hat{B}(q^{-1}) \cdot u_K \} \quad (2.73)$$

Como o erro de previsão é uma função não linear dos parâmetros desconhecidos, é conveniente a elaboração de hipóteses que permitam obter uma equação explícita para os estimadores.

Da equação (2.73) tem-se:

$$\hat{C}(q^{-1}) \cdot \varepsilon(\hat{\theta}) = \hat{D}(q^{-1}) \cdot \{\hat{A}(q^{-1}) \cdot \underline{y}_K - \hat{B}(q^{-1}) \cdot \underline{u}_K\} \quad (2.74)$$

Definindo-se:

$$\begin{aligned} A^*(q^{-1}) &\triangleq A(q^{-1}) - I \\ C^*(q^{-1}) &\triangleq C(q^{-1}) - I \\ D^*(q^{-1}) &\triangleq D(q^{-1}) - I \end{aligned} \quad (2.75)$$

torna-se, então, possível reescrever (2.74) como:

$$\underline{\varepsilon}_K(\cdot) = \hat{A}(q^{-1}) \cdot \underline{y}_K - \hat{B}(q^{-1}) \cdot \underline{u}_K + \hat{D}^*(q^{-1}) \cdot \underline{\varepsilon}_K - \hat{C}^*(q^{-1}) \cdot \underline{\varepsilon}_K(\cdot) \quad (2.76)$$

Supondo-se que $\underline{\varepsilon}_K$ e $\underline{\varepsilon}_K(\cdot)$ são mensuráveis, o erro de previsão em (2.76) é uma função linear dos parâmetros desconhecidos e um procedimento análogo ao do algoritmo de mínimos quadrados pode ser utilizado.

Assim, o modelo (2.76) do erro previsto pode ser escrito como:

$$\underline{y}_K = \psi_K \cdot \underline{\theta}^{A,B,C,D} + \underline{\varepsilon}_K \quad (2.77)$$

com:

$$\underline{\theta}^{A,B,C,D} = [\underline{\theta}_1^{A,B,C,D} \mid \dots \mid \underline{\theta}_n^{A,B,C,D}]$$

$$\underline{\theta}_i^{A,B,C,D} = [\underline{a}_i^T \mid \underline{b}_i^T \mid \underline{c}_i^T \mid \underline{d}_i^T] \quad i = 1, \dots, n$$

$\underline{a}_i^T$ = elementos da i-ésima linha de $A^*(q^{-1})$

$\underline{b}_i^T$ = elementos da i-ésima linha de $B(q^{-1})$

$\underline{c}_i^T$ = elementos da i-ésima linha de $C^*(q^{-1})$

$$\hat{\underline{\zeta}}_K(\hat{\underline{\theta}}) = \underline{y}_K - \psi_K^{A,B} \hat{\underline{\theta}}_{K-1}^{A,B} \quad (2.82)$$

$$\underline{\varepsilon}_K(\hat{\underline{\theta}}) = \underline{y}_K - \psi_K \cdot \hat{\underline{\theta}}_{K-1}^{A,B,C,D} \quad (2.83)$$

com

$$\psi_K^{A,B} = \text{bloco diagonal} [\underbrace{\phi_K^{T_{A,B}}}_{1}, \dots, \underbrace{\phi_K^{T_{A,B}}}_{n}] \quad (2.84)$$

$$\phi_K^{T_{A,B}} = [-\underline{y}_{K-1}^T \mid \dots \mid -\underline{y}_{K-N^*}^T \mid \underline{u}_{K-1}^T \mid \dots \mid \underline{u}_{K-N^*}^T]$$

Os parâmetros do modelo empregado pelo algoritmo de matriz estendida podem ser estimados por um algoritmo mais simples, pelo uso de uma partição que separa os parâmetros do processo, correspondentes às matrizes A e B, dos parâmetros do ruído, correspondentes às matrizes C e D.

Como estimador não-recursivo dos parâmetros do modelo (2.77) que minimiza a função J_5 (2.78) é dado para cada uma das linhas deste modelo por:

$$\hat{\underline{\theta}}_{-1}^{A,B,C,D} = \left[\begin{array}{c} N \\ \Sigma \\ \sum_{K=1} \phi_K \phi_K^T \end{array} \right]^{-1} \left[\begin{array}{c} N \\ \Sigma \\ \sum_{K=1} \phi_K y_K^1 \end{array} \right] \quad i = 1, \dots, n \quad (2.85)$$

Definindo:

$$\phi_K^{P^T} \triangleq [-\underline{y}_{K-1}^T \mid \dots \mid -\underline{y}_{K-N^*}^T \mid \underline{u}_{K-1}^T \mid \dots \mid \underline{u}_{K-N^*}^T] \quad (2.86)$$

$$\phi_K^{R^T} \triangleq [\underline{\varepsilon}_{K-1}^T \mid \dots \mid \underline{\varepsilon}_{K-N^*}^T \mid -\underline{\zeta}_{K-1}^T \mid \dots \mid -\underline{\zeta}_{K-N^*}^T]$$

tem-se:

$$\phi_K^T = [\phi_K^{P^T} \mid \phi_K^{R^T}] \quad (2.87)$$

Utilizando a partição de ϕ_K^T em (2.87), para reescrever a equação (2.85), obtém-se:

$$\begin{bmatrix} \hat{\underline{\theta}}_i^P \\ \hat{\underline{\theta}}_i^R \end{bmatrix} = \begin{bmatrix} \hat{\underline{\theta}}_i^{A,B} \\ \hat{\underline{\theta}}_i^{C,D} \end{bmatrix} = \begin{bmatrix} \sum_{K=1}^N \underline{\phi}_K^P \underline{\phi}_K^{P^T} & \sum_{K=1}^N \underline{\phi}_K^P \underline{\phi}_K^{R^T} \\ \sum_{K=1}^N \underline{\phi}_K^R \underline{\phi}_K^{P^T} & \sum_{K=1}^N \underline{\phi}_K^R \underline{\phi}_K^{R^T} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum_{K=1}^N \underline{\phi}_K^P y_K^i \\ \sum_{K=1}^N \underline{\phi}_K^R y_K^i \end{bmatrix} \quad (2.88)$$

A aplicação do lema da inversa particionada, na identidade 6 do Apêndice A, para encontrar a matriz inversa desta última equação resulta em:

$$\hat{\underline{\theta}}_i^P = \begin{bmatrix} \sum_{K=1}^N \underline{\phi}_K^P \underline{\phi}_K^{P^T} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{K=1}^N \underline{\phi}_K^P y_K^i \end{bmatrix} - \begin{bmatrix} \sum_{K=1}^N \underline{\phi}_K^P \underline{\phi}_K^{P^T} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{K=1}^N \underline{\phi}_K^P \underline{\phi}_K^{R^T} \cdot \hat{\underline{\theta}}_i^R \end{bmatrix} \quad (2.89)$$

Nesta equação é importante observar que o estimador do processo depende de duas parcelas. A primeira corresponde ao estimador de mínimos quadrados e a segunda é uma correção, efetuada devido à polarização no estimador de mínimos quadrados na presença de um ruído não-branco. Para concretizar estas idéias, a equação (2.89) pode ser descrita como:

$$\hat{\underline{\theta}}_i^P = \hat{\underline{\theta}}_i^{M.Q.} - \hat{\underline{\theta}}_i^C \quad (2.90)$$

onde $\hat{\underline{\theta}}_i^{M.Q.}$ é o estimador de mínimos quadrados para os parâmetros das matrizes $A(q^{-1})$ e $B(q^{-1})$ e $\hat{\underline{\theta}}_i^C$ é o termo corretivo dado por:

$$\hat{\underline{\theta}}_i^C = \begin{bmatrix} \sum_{K=1}^N \underline{\phi}_K^P \underline{\phi}_K^{P^T} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{K=1}^N \underline{\phi}_K^P \underline{\phi}_K^{R^T} \end{bmatrix} \hat{\underline{\theta}}_i^R \quad (2.91)$$

A equação (2.90) mostra que é necessário obter um estimador dos parâmetros do ruído para se estimar os parâmetros do processo. Para tanto, a equação do erro previsto é reescrita de modo a envolver somente os parâmetros do ruído. Assim:

$$\underline{\varepsilon}_K(\cdot) = \underline{z}_K(\cdot) - \underline{\psi}_K^R \cdot \underline{\theta}^R \quad (2.92)$$

com

$$\underline{\psi}_K^R = \text{bloco diagonal} [\underline{\phi}^{RT} | \dots | \underline{\phi}^{RT}]$$

$$\underline{\theta}^{RT} = [\underline{\theta}_1^{RT} | \dots | \underline{\theta}_n^{RT}]$$

e

$$\underline{\theta}_i^R = [\underline{c}_i^T | \underline{d}_i^T]$$

Para minimizar a função:

$$J_6 = \frac{1}{2} \sum_{K=1}^N (\underline{z}_K - \underline{\psi}_K^R \underline{\theta}^R)^T V^{-1} (\underline{z}_K - \underline{\psi}_K^R \underline{\theta}^R) \quad (2.93)$$

o estimador dos parâmetros em $\underline{\theta}^R$ deve satisfazer a:

$$\underline{\theta}_i^R = \left[\sum_{K=1}^N \underline{\phi}_K^R \underline{\phi}_K^{RT} \right]^{-1} \left[\sum_{K=1}^N \underline{\phi}_K^R \underline{z}_K^i \right] \quad i = 1, \dots, n \quad (2.94)$$

onde a estimação em separado de cada uma das linhas do modelo (2.92) é possível devido a estrutura bloco diagonal de $\underline{\psi}_K^R$.

A versão recursiva para o algoritmo particionado é obtida usando-se o lema de inversão matricial nas equações do estimador de mínimos quadrados, do estimador dos parâmetros do ruído e do termo corretivo. As equações a seguir descrevem o estimador obtido:

$$\hat{\underline{\theta}}_{i,N+1}^{M.Q.} = \hat{\underline{\theta}}_{i,N}^{M.Q.} + K_{i,N+1}^P \{Y_{N+1}^i - \underline{\phi}_{N+1}^{PT} \cdot \hat{\underline{\theta}}_{i,N}^{M.Q.}\}$$

$$P_{i,N+1}^P = P_{i,N}^P - \frac{P_{i,N}^P \underline{\phi}_{N+1}^P \underline{\phi}_{N+1}^{PT} P_{i,N}^P}{1 + \underline{\phi}_{N+1}^{PT} P_{i,N}^P \underline{\phi}_{N+1}^P}$$

$$K_{i,N+1}^P = P_{i,N}^P \cdot \phi_{N+1}^P$$

$$\hat{\theta}_{i,N+1}^R = \hat{\theta}_{i,N}^R + K_{i,N+1}^R \{ \hat{z}_{N+1}^i - \phi_{N+1}^{RT} \hat{\theta}_{i,N}^R \}$$

$$P_{i,N+1}^R = P_{i,N}^R - \frac{P_{i,N}^R \cdot \phi_{N+1}^R \phi_{N+1}^{RT} P_{i,N}^R}{1 + \phi_{N+1}^{RT} P_{i,N}^R \cdot \phi_{N+1}^R}$$

$$K_{i,N+1}^R = P_{i,N}^R \cdot \phi_{N+1}^R$$

$$\hat{\theta}_{i,N+1}^C = \hat{\theta}_{i,N}^C + K_{i,N+1}^P \left\{ \phi_{N+1}^{RT} \cdot \hat{\theta}_{i,N+1}^R - \phi_{N+1}^{PT} \cdot \hat{\theta}_{i,N}^C \right\}$$

$$\hat{\theta}_{i,N+1}^P = \hat{\theta}_{i,N+1}^{M.Q.} - \hat{\theta}_{i,N+1}^C \quad i = 1, \dots, n \quad (2.95)$$

A simplificação em relação ao estimador de matriz estendida consiste no fato de que enquanto este estimador requer a atualização de uma matriz $p \times p$, com $p = (3n+m)N_*$, o estimador particionado atualiza uma matriz $p_1 \times p_1$ e e outra $p_2 \times p_2$, com $p_1 \triangleq (n+m).N_*$ e $p_2 \triangleq 2n.N_*$.

Os estimadores descritos nas seções 2.5.1 e 2.5.4 são apresentados na referência [1]. Exceção é feita ao estimador da raiz quadrada que consiste numa aplicação das idéias desenvolvidas nesta referência a um método numérico em uso para sistemas com uma entrada e uma saída [21]. O item 2.5.3, por sua vez, é uma aplicação original da estimação linha por linha encontrada na referência [1] para modelos com parâmetros variáveis no tempo.

2.6. CONCLUSÕES

Foram expostos, neste capítulo, algoritmos que se destinam à identificação de sistemas multivariáveis. Estes algoritmos baseiam-se na hipótese de centralização, isto é, todas as informações sobre o sistema devem estar reunidas num mesmo local onde são processadas.

O algoritmo de máxima verossimilhança, apresentado no início do capítulo, tem propriedades assintóticas adequadas. Entretanto, seu uso é condicionado a obtenção de informações sobre o comportamento estatístico das saídas do sistema. Como alternativa, existem os métodos do erro de previsão que são estruturados a partir de um modelo, que relaciona o erro de previsão com as saídas e os parâmetros do sistema, e um critério para seleção do estimador ótimo.

O primeiro método de erro de previsão apresentado foi o estimador de mínimos quadrados.

Atualmente, vem recebendo especial atenção os algoritmos recursivos de estimação. Este fato é, em parte, devido a dois fatores:

- i) redução do tempo de cálculo e do volume de informações requeridas a cada passo nestes estimadores permitindo em consequência o uso de computadores de pequeno porte e
- ii) possibilidade de estimação recursiva de parâmetros variáveis no tempo, indispensável em sistemas de controle adaptativo.

Os algoritmos de "janela exponencial" e "janela retangular", destinados a identificação de sistemas multivariáveis com parâmetros variáveis no tempo fazem uso de técnicas desenvolvidas para sistemas com uma entrada e uma saída.

Quando as perturbações atuando no sistema são correlatas, elas devem ser modeladas a partir de ruídos brancos. O algoritmo de matriz estendida e sua forma particionada adotam este procedimento.

CAPÍTULO 3

ALGORÍTMOS PARA IDENTIFICAÇÃO DE SISTEMAS DESCENTRALIZADOS

- 3.1. Introdução
- 3.2. Formas Canônicas para Sistemas Multivariáveis
- 3.3. Sistemas Descentralizados: Subsistemas, Índices de Distribuição e Causalidade
- 3.4. Algoritmos para Identificação Descentralizada
- 3.5. Estimadores Recursivos para Redes Estrela (ERRE)
 - 3.5.1. Algoritmos ERRE1
 - 3.5.2. Algoritmos ERRE2
- 3.6. Estimadores Recursivos para Redes Anel (ERRA)
 - 3.6.1. Algoritmos ERRA1
 - 3.6.2. Algoritmos ERRA2
- 3.7. Estimadores Recursivos para Redes com Barramento (ERRB)
 - 3.7.1. Algoritmos ERRB1
 - 3.7.2. Algoritmos ERRB2
- 3.8. Conclusões

3.1. INTRODUÇÃO

Neste capítulo realizam-se adaptações nos algoritmos a apresentados no Capítulo 2, de modo a torná-los aplicáveis a sistemas, cujos recursos computacionais estejam organizados numa rede descentralizada ou hierárquica.

As três primeiras seções dedicam-se à construção de uma terminologia para o tratamento de sistemas descentralizados e à montagem de uma base comum para o desenvolvimento dos algoritmos.

Em primeiro lugar são apresentadas formas canônicas, cujo uso é frequente em algoritmos para sistemas multivariáveis. Em seguida, formaliza-se a noção de sistema descentralizado, com a elaboração da terminologia necessária e a definição de relações de causalidade entre os vários subsistemas. São introduzidos índices de distribuição que guardam certa analogia com o índice de concentração de renda de Andres Garcia Perez [26].

É proposta uma forma de construção de modelo do sistema descentralizado: a decomposição. No modelamento deste sistema, emprega-se o conceito de nível de descentralização que ajuda a especificar as restrições que diferenciam um sistema descentralizado e um sistema centralizado. Para cada uma destas restrições são mencionados modelos adequados.

Durante este capítulo, fica clara a existência de fortes vínculos entre a distribuição física do sistema em estudo, a configuração da rede de computação disponível e o algoritmo descentralizado a ser utilizado. Estes vínculos não existem no caso de algoritmos centralizados, para os quais as relações com o sistema computacional limitou-se ao tamanho da memória (quantidade), e à velocidade de processamento (qualidade), sem importar a configuração básica.

São apresentados:

- estimadores recursivos para redes estrela (ERRE)
- estimadores recursivos para redes anel (ERRA)

- estimadores recursivos para redes com barramento (ERRB)

Para cada um destes são analisados dois tipos de restrição quanto a troca de informações entre os processadores da rede. Modelos com parâmetros fixos e variáveis no tempo são considerados. Tabelas e diagramas são utilizados para condensar algumas informações importantes.

3.2. FORMAS CANÔNICAS PARA SISTEMAS MULTIVARIÁVEIS

Algumas das mais eficientes metodologias adotadas recentemente na identificação de sistemas lineares multivariáveis discretos são baseadas em modelos canônicos de entrada-saída e na determinação em separado da estrutura (identificação estrutural) e dos parâmetros (identificação paramétrica).

Esta seção apresenta formas canônicas frequentes na representação de processos. O uso de formas canônicas permite assegurar a unicidade das equações da representação.

O modelo geral de um sistema estocástico linear multivariável discreto descrito nas equações (1.5) e (1.7), é obtido de modo que as saídas do modelo e do processo sejam vetores com a mesma densidade espectral. No modelo do ruído $\underline{z}_K$ a partir do ruído branco $\underline{w}_K$ a unicidade da matriz de transferência $H(q^{-1})$ é garantida pelo teorema da representação espectral. Entretanto, como a descrição por matriz partilhada não é única torna-se necessária a definição de critérios para escolha das matrizes $C(q^{-1})$ e $D(q^{-1})$. Estes critérios são o objetivo das formas canônicas.

Dados os modelos descritos pelas equações (1.5) e (1.7) e reescritos a seguir:

$$A(q^{-1}) \cdot \underline{y}_K = B(q^{-1}) \cdot \underline{u}_K + \underline{z}_K \quad (3.1)$$

$$D(q^{-1}) \cdot \underline{z}_K = C(q^{-1}) \cdot \underline{w}_K$$

$$e \quad F(q^{-1}) \underline{Y}_K = R(q^{-1}) \underline{u}_K + C(q^{-1}) \cdot \underline{W}_K \quad (3.2)$$

$$\text{com} \quad F(q^{-1}) = D(q^{-1}) \cdot A(q^{-1})$$

$$R(q^{-1}) = D(q^{-1}) \cdot B(q^{-1})$$

$$E\{\underline{W}_K \cdot \underline{W}_K^T\} = V$$

Para melhor entendimento dos propósitos envolvidos na definição de formas canônicas considera-se o conjunto $\mathcal{D}^*$ formado por elementos H_i , constituídos pelas quatro matrizes $F_i(q^{-1})$, $R_i(q^{-1})$, $C_i(q^{-1})$ e V_i :

$$\mathcal{D}^* = \{F_i(q^{-1}) \cdot R_i(q^{-1}) \cdot C_i(q^{-1}), V_i\} \quad (3.3)$$

Entre os elementos H_i de $\mathcal{D}^*$ define-se a relação de equivalência (E) através de: $H_i (E) H_j$ se H_i e H_j fornecem modelos com a mesma densidade espectral de saída S_{YY} e densidade espectral cruzada entrada-saída S_{Yu} . Esta relação de equivalência divide $\mathcal{D}^*$ em classes de equivalência $\mathcal{D}_\alpha^*$ para cada par $\{S_{YY}^\alpha, S_{Yu}^\alpha\}$. É possível provar que elementos H_i e H_j pertencentes a mesma classe de equivalência $\mathcal{D}_\alpha^*$ satisfazem a:

$$F_i = U_\alpha \cdot F_j, \quad R_i = U_\alpha \cdot R_j, \quad C_i = U_\alpha \cdot C_j \cdot S_\alpha^{-1}, \quad V_i = S_\alpha \cdot V_j \cdot S_\alpha^T \quad (3.4)$$

A definição de uma forma canônica C_α^i consiste em estabelecer um critério T_β que permita a escolha de um único elemento dentro de cada classe $\mathcal{D}_\alpha^*$. Estes elementos formam o subconjunto canônico $C_\beta \subset \mathcal{D}^*$. Modificando-se o critério, T_β de acordo com a aplicação pretendida, diversas formas canônicas podem ser definidas.

Para que a representação de um processo exista e seja única exige-se numa forma canônica que qualquer mudança nos valores dos parâmetros de uma forma canônica resulte em mudança no par $\{S_{YY}^\alpha, S_{Yu}^\alpha\}$ de densidades espectrais. Matematicamente:

- (a) se H_1 e H_2 pertencem a C_1 e $H_1 \oplus H_2$ então $H_1 = H_2$
- (b) para qualquer elemento H de $\mathcal{H}^*$ existe um único elemento H_1 em C_β tal que $H \oplus H_1$

A condição (a) é indispensável na estimação de parâmetros visto que impossibilita os algoritmos de oscilarem entre duas soluções equivalentes. A condição (b) é a que torna a forma canônica geral, isto é, capaz de gerar qualquer par de densidades espectrais.

Naturalmente, se a estrutura do modelo de um processo for conhecida a priori, em decorrência de estudos já efetuados na estimação de parâmetros pode-se empregar esta estrutura e ignorar a condição (b).

São apresentadas a seguir, três formas canônicas existentes na literatura [1] e referentes ao modelo na equação (3.2):

Forma canônica I: caracterizada por $F(0) = I$, $F(q^{-1})$, triangular $R(q^{-1})$, $C(q^{-1})$ e V arbitrários. Esta forma inclui o menor número de parâmetros.

Forma canônica II: caracterizada por $F(0) = I$, $F(q^{-1})$, $R(q^{-1})$ e V arbitrários, $C(q^{-1})$ diagonal.

Forma canônica III: caracterizado por $F(0)$ ser uma matriz triangular com elementos unitários na diagonal, $F(q^{-1})$ e $R(q^{-1})$ arbitrários, $C(q^{-1})$ e V diagonais. Nesta forma canônica, cada componente do ruído atua independentemente em cada saída do modelo.

3.3. SISTEMAS DESCENTRALIZADOS: SUBSISTEMAS, ÍNDICES DE DISTRIBUIÇÃO E CAUSALIDADE

O destaque que vem recebendo a idéia de descentralização na última década está intimamente ligado às dificuldades de aplicação da teoria de identificação e controle no tratamento de sistemas de grande porte. Entretanto, a idéia em si não constitui

uma novidade pois faz parte de uma abordagem largamente empregada na análise de sistemas complexos. Um procedimento adequado para modelar um sistema complexo é dividi-lo em vários subsistemas, elaborar um modelo para cada um deles e, a seguir, investigar suas interligações. Os modelos, obtidos através deste procedimento, compõem-se de equações e condições de fronteira.

As tentativas em lidar com sistemas que envolvam muitas variáveis ou ocupem áreas expressivas sem considerar como fundamentais estes aspectos vem encontrando imensas barreiras tanto a nível teórico quanto nas possibilidades de aplicação. Um exemplo interessante são os projetos de sistemas de controle baseados na realimentação de estado onde em muitos casos o número ideal de malhas de realimentação pode ser impossível de implementação prática [28].

Apesar das reformulações que exige, a idéia de descentralização aparece como uma contribuição promissora para simplificar a identificação de parâmetros primários e secundários e os projetos de controle de sistemas de grande porte.

Os desenvolvimentos em microeletrônica são os responsáveis pela oportunidade de passar a descentralização do plano teórico, como passo inicial na solução de problemas, para o plano prático, como configuração final dos algoritmos que regem um sistema. Contudo, há pouco tempo esta oportunidade parecia tão irreal que o esforço para elaboração de uma base teórica sobre o comportamento de algoritmos descentralizados remonta a poucos anos. Assim, muito da terminologia, definições e suposições encontradas neste trabalho são emprestadas de outras áreas de conhecimento.

As diversas partes de um sistema são definidas como subsistemas. O número total de subsistemas que compõe um sistema é indicado por n_s . Os subsistemas são individualmente referidos por S_i , $i = 1, \dots, n_s$. A cada subsistema corresponde um conjunto de entradas, saídas e interconexões. As saídas, as entradas e as interconexões são exclusivas para cada subsistema. As interconexões são definidas como sendo as interações físicas do subsistema com

outros subsistemas. Nem sempre existem interconexões.

Na verdade, a característica básica da descentralização é a restrição na troca de informações entre alguns subsistemas. Quando a taxa de troca de informações deve ser mantida em valores abaixo do necessário para implementação de algoritmos centralizados ocorre a descentralização parcial.

Fica claro, portanto, que a descentralização dos algoritmos de identificação e controle elaborados para um sistema tanto pode ser uma decorrência da configuração física do sistema como pode tratar-se de uma forma conveniente de dividir um sistema, que mesmo ocupando uma área pequena, englobe muitas variáveis.

A descentralização assume uma distribuição das saídas entre os subsistemas. Esta distribuição juntamente com o número de subsistemas é um dado importante para avaliar o quanto o sistema original foi partilhado. Esta avaliação não pode ser realizada apenas a partir do número de subsistemas. Como ilustração, considere-se dois sistemas de nove saídas divididas em três subsistemas, sendo que, no primeiro caso, cada um dos subsistemas têm três saídas e, no segundo, dois subsistemas possuem uma saída e o outro subsistema as sete saídas restantes. Neste exemplo, a distribuição é maior no primeiro caso, enquanto no segundo, muito da configuração original do sistema foi preservada.

Para quantificar a distribuição das saídas entre os subsistemas, define-se o índice de distribuição D como o quociente normalizado da soma do número acumulado de saídas pela soma do número acumulado de subsistemas correspondentes. Ou seja:

$$D = \frac{ns}{n} \frac{\sum n_a}{\sum ns_a} \quad (3.5)$$

com: ns = número total de subsistemas

n = número total de saídas

ns_a = número acumulado de subsistemas

n_a = número acumulado de saídas correspondentes em ordem crescente

O valor máximo do índice de distribuição é igual a unidade, sendo atingido quando as saídas estão igualmente distribuídas entre os subsistemas. Se a maior parte das saídas pertence a um único subsistema, o valor de D tende a:

$$\frac{ns(ns-1) + 2n}{(ns+1)n} \quad (3.6)$$

Este último valor depende do número total de saídas e subsistemas.

Para o exemplo citado acima, os valores do índice de distribuição são 1,00 e 0,67 para o primeiro e segundo casos, respectivamente. Na tabela 3.1 estão resumidos as etapas e o valor final do cálculo de D para cinco sistemas de 200 saídas distribuídas entre 20 subsistemas.

Assim, o conhecimento do índice de distribuição e do número de subsistemas é eficiente para avaliar o quanto o sistema original foi partilhado. A utilidade do índice torna-se evidente quando grande número de saídas e subsistemas estão envolvidos como nos exemplos da tabela 3.1.

Uma ressalva a respeito do índice de distribuição descrito pela eq. (3.5) é que a sua definição não considera as diferenças associadas às saídas dos subsistemas. Caso existam diferenças significativas, pode ser preferível utilizar um índice que expresse a distribuição de potência entre os diversos subsistemas. Tal índice pode ser definido como:

$$D^* = \frac{ns}{p} \frac{\sum p_a}{\sum ns_a} \quad (3.7)$$

onde: ns = número total de subsistemas

p = potência total das saídas

ns_a = número acumulado de subsistemas

Σns_a	A	B	C	D	E
	n_a	n_a	n_a	n_a	n_a
1	1	3	5	9	10
2	2	5	10	17	20
3	3	10	15	25	30
4	4	14	20	31	40
5	5	18	25	37	50
6	6	20	30	40	60
7	7	26	35	42	70
8	8	31	40	45	80
9	9	33	45	51	90
10	10	35	50	57	100
11	11	39	55	60	110
12	12	41	60	65	120
13	13	45	65	79	130
14	14	47	70	95	140
15	15	49	75	110	150
16	16	50	80	120	160
17	17	54	85	135	170
18	18	57	90	150	180
19	19	60	100	170	190
20	200	200	200	200	200
$\Sigma ns_a = 210$	$\Sigma n_a = 390$	$\Sigma n_a = 837$	$\Sigma n_a = 1155$	$\Sigma n_a = 1538$	$\Sigma n_a = 2100$
	$D = 0,186$	$D = 0,399$	$D = 0,550$	$D = 0,732$	$D = 1.000$

Tabela 3.1 - Cálculo do Índice de Distribuição para Sistemas de 20 Saídas Distribuídas em 20 Subsistemas

p_a = potência acumulada das saídas correspondentes

A subdivisão do sistema é o ponto de partida para investigar as influências exercidas entre os subsistemas, isto é, avaliar qual a influência exercida sobre as saídas de um subsistema pelas entradas e saídas passadas dos demais subsistemas.

O conceito de causalidade, frequentemente aplicado em ciências experimentais e modificado por Grauger (1963) [13], para o uso em problemas de identificação pode ser estendido também a sistemas descentralizados. Diz-se que um subsistema S_i é causal para outro subsistema S_j , se a precisão das previsões das saídas de S_j obtidas utilizando as saídas passadas de S_i é maior que a precisão obtida usando apenas as saídas de S_j . Formalmente, esta definição fica:

Seja, S_i e S_j dois subsistemas de um sistema descentralizado. S_i é causal para S_j se é válida a desigualdade:

$$\text{cov}[y_t^{j,\ell} | y_{t-r}^j, y_{t-r}^i, r \geq 1] < \text{cov}[y_t^{j,\ell} | y_{t-r}^j, r \geq 1] \quad (3.8)$$

para $\ell = 1, \dots, n_j$

onde y_t^j = vetor-saída do subsistema S_j

y_t^i = vetor-saída do subsistema S_i

$y_t^{j,\ell}$ = ℓ -ésima componente do vetor-saída y_t^j de S_j

n_j = número de componentes do vetor-saída y_t^j

Quando a desigualdade (3.8) é satisfeita apenas para algumas das componentes da saída de S_j , diz-se que S_i é causal apenas para estas componentes de S_j .

Em sistemas descentralizados nem sempre as saídas passadas de um subsistema podem ser transmitidas a outro subsistema e utilizadas nas previsões das saídas deste último. Somente quando

isto pode ser feito é que se justifica um estudo das relações de causalidade entre subsistemas. Em fase de projeto de um sistema de controle, o estudo das relações de causalidade pode ajudar a decidir quais interconexões devem ser construídas.

Se o subsistema S_i é causal para o subsistema S_j e, S_j não é causal para S_i , o vetor $\underline{y}^i$ nas equações das saídas do subsistema é denominado exógeno (entrada).

Para verificar se o subsistema S_i é causal para o subsistema S_j utiliza-se o seguinte algoritmo:

(a) constroem-se, inicialmente, os modelos:

$$\underline{y}_t^{j,l} = (\underline{\phi}_t^j)^T \underline{\theta}^{j,l} + \underline{w}_t^{j,l} \quad \text{para } l = 1, \dots, n_j \quad (3.9)$$

$$\underline{y}_t^i = \underline{\psi}_t^i \cdot \underline{\theta}^i + \underline{w}_t^i \quad (3.10)$$

nos quais aparecem apenas dados de um mesmo subsistema. O vetor $\underline{\phi}_t^j$ contém valores passados da saída $\underline{y}^j$ do subsistema S_j . Da mesma forma, a matriz $\underline{\psi}_t^i$ é constituída por valores passados da saída $\underline{y}^i$ do subsistema S_i .

(b) se S_i for causal para S_j então existe algum vetor $\underline{\alpha}_K$ não nulo tal que:

$$\underline{w}_t^{j,l} = \sum_{K=1}^{n_\alpha} \underline{\alpha}_K^T \underline{w}_{t-K}^i + \eta_t^l \quad \text{para cada } l = 1, \dots, n_j \quad (3.11)$$

onde $\{\eta_t^l\}$ é uma sequência de variáveis aleatórias independentes uniformemente distribuída e com média zero.

Para provar que $\underline{\alpha}_K$ é significativamente diferente do vetor nulo, vários testes estatísticos podem ser utilizados. Dois destes, o teste de hipótese e o teste modificado de máxima verossimilhança, são apresentados a seguir.

Seja H_0^l a hipótese $\underline{\alpha}_K = \underline{0}$, para $K = 1, \dots, n_\alpha$.

Teste de Hipótese: Tendo sido feitas observações de

ω_t^i e $\omega_t^{j,l}$ para $t = 1, \dots, N$, seja p_1 o valor estimado para média do quadrado de $\omega_t^{j,l}$ e p_2 o valor estimado para a média do quadrado do resíduo n_l quando H_0^l é falsa. Ou seja:

$$p_1 = \frac{1}{N} \sum_{t=1}^N (\omega_t^{j,l})^2 \quad (3.12)$$

$$p_2 = \frac{1}{N-n_\alpha} \sum_{t=n_\alpha+1}^N \left(\omega_t^{j,l} - \sum_{K=1}^{n_\alpha} \hat{\alpha}_K^T \omega_{t-K}^i \right)^2 \quad (3.13)$$

onde $\hat{\alpha}_K$ é o valor estimado para α_K , obtido utilizando-se as observações $\omega_t^{j,l}$, ω_t^i em $t = 1, \dots, N$.

Se H_0^l for verdadeira, então a variável X , definida por:

$$X = \left(\frac{p_1 - p_2}{p_2} \right) \cdot \left(\frac{N - n_\alpha - 1}{n_\alpha} \right) \quad (3.14)$$

tem a seguinte distribuição de probabilidade:

$$X \sim F(n_\alpha, N - n_\alpha - 1) \quad (3.15)$$

Escolhendo um nível conveniente de significância e indicando por x o valor correspondente numa tabela da distribuição F , tem-se:

$$X \leq x \rightarrow H_0^l \text{ deve ser aceita}$$

$$X > x \rightarrow H_0^l \text{ deve ser rejeitada}$$

Teste Modificado de Máxima Verossimilhança: Sejam J_1 e J_2 os custos associados ao estimador de máxima verossimilhança, descrito pela equação (2.12) supondo H_0^l verdadeira ou falsa respectivamente:

$$J_1 = \frac{N}{2} \ln p_1 \quad J_2 = \frac{N-n_\alpha}{2} \ln p_2 + n_\alpha \quad (3.16)$$

A suposição a qual corresponde o menor valor (J_1 ou J_2) deve ser escolhida.

3.4. ALGORÍTMOS PARA IDENTIFICAÇÃO DESCENTRALIZADA

Nesta seção apresentam-se os modelos utilizados pelos algoritmos de identificação descentralizada e, também, a discussão de funções que estabelecem os objetivos a serem atingidos por estes algoritmos. A configuração descentralizada possibilita a introdução de métodos novos para construção de modelos e alguns detalhes levam à reformulação dos objetivos a serem satisfeitos pelos estimadores.

As idéias desenvolvidas, em sua maior parte, não se referem a uma configuração particular, mas levam em conta as características básicas de descentralização: a divisão em subsistemas e a redução na taxa de troca de informações entre os subsistemas. A aplicação dos métodos de modelamento e critérios de estimação às redes com configuração estrela, anel e com barramento é objeto das próximas seções.

Como, em cada situação, a descentralização pode assumir uma forma particular, antes de prosseguir é necessário avaliar os tipos de restrição que podem limitar o fluxo de informação. A figura 3.1 resume estas restrições e, além disto, sugere que um sistema descentralizado pode ser visto como resultante de um sistema centralizado no qual ocorreram falhas na transmissão de informações entre as várias partes do sistema. Talvez um pouco menos evidente seja o fato de que cada uma das falhas representadas na figura 3.1 apresenta diversos níveis. Estes são definidos como níveis de degradação sendo o nível máximo correspondente a suspensão da troca de informações entre os subsistemas. Seguindo esta linha de pensamento, os algoritmos descentralizados podem auxiliar nas previsões da saída do sistema quando ocorrer falhas no fluxo de informações. Para tanto, são necessários um mecanismo de alarme e procedimentos que conduzam ao chaveamento do algoritmo

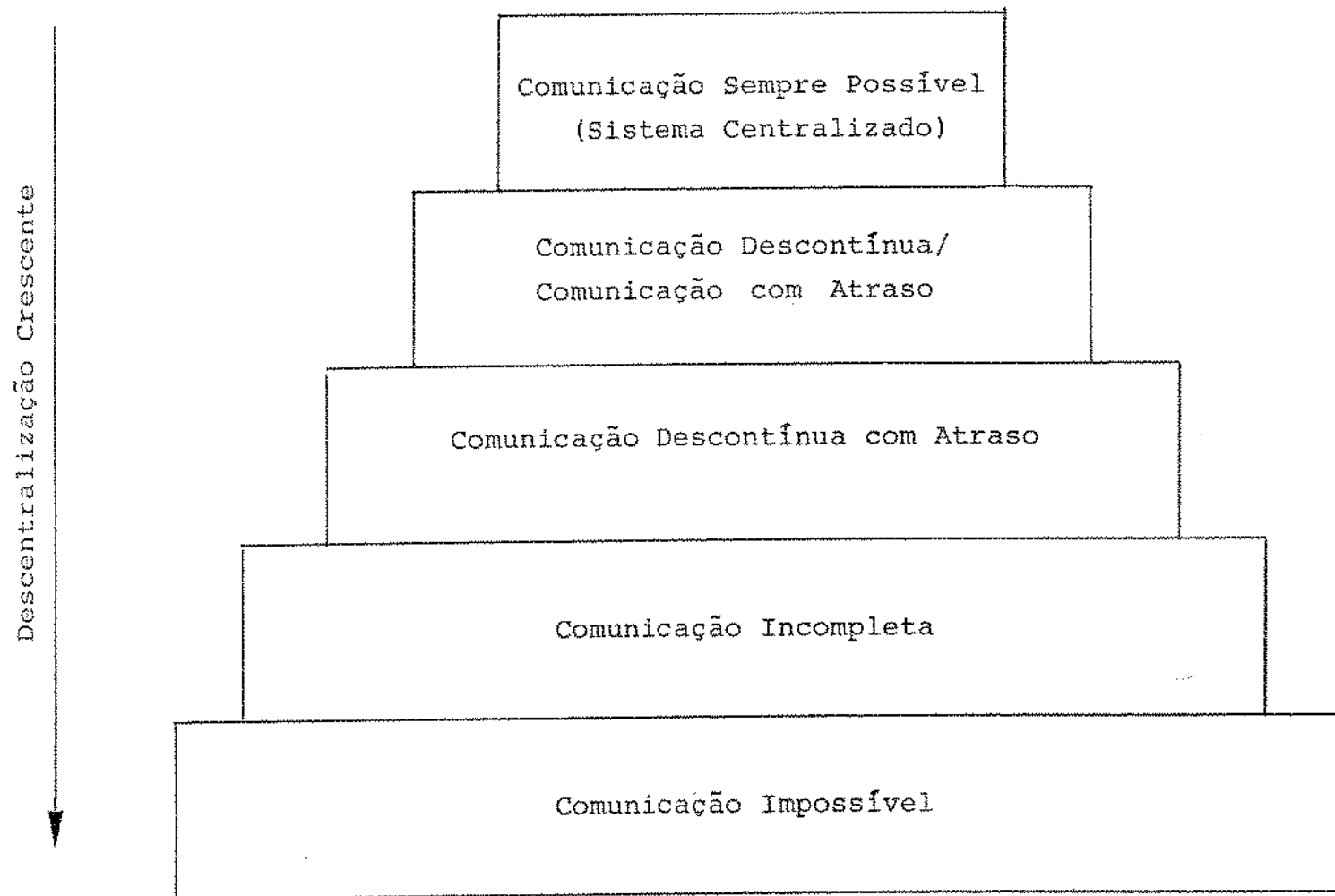


Figura 3.1 - Tipos de Falhas no Fluxo de Informações entre Subsistemas

centralizado para outro, descentralizado. Uma vez superadas as falhas, a volta aos esquemas centralizados deve também seguir rotinas predeterminadas.

Na figura 3.1, centralização máxima corresponde a possibilidade de transmitir todas as informações sobre o sistema a um ou mais locais de processamento para o desenvolvimento de algoritmos. Se existem vários processadores e é possível a cada um deles receber informações completas sobre o sistema, os algoritmos utilizados não são descentralizados, pois não há restrições no fluxo de informações. Estes algoritmos estão apenas distribuídos em vários locais de processamento. Entretanto, quando entre duas transmissões de dados ao local de processamento corresponde a um subsistema, existe um intervalo maior que o período de amostragem das medidas no próprio subsistema, a comunicação é denominada descontínua. O nível de degradação $\mathcal{N}_D$, no caso de comunicação descontínua, é o número de amostragens realizadas no subsistema até que novas informações sobre o resto do sistema estejam disponíveis. Na comunicação com atraso, o subsistema recebe informações do resto do sistema com atraso e o número de intervalos de amostragem correspondente a este atraso é o nível de degradação $\mathcal{N}_A$. A comunicação descontínua com atraso reúne as duas restrições mencionadas e seu nível de degradação é caracterizado pelo par $(\mathcal{N}_D, \mathcal{N}_A)$.

Quando um subsistema não recebe informações a respeito de alguns dos demais subsistemas, a comunicação é considerada incompleta. Neste caso, o nível de degradação $\mathcal{N}_I$ é definido como a razão entre o número de subsistemas que não enviam informações e o número total de subsistemas menos um.

A seguir são apresentados alguns modelos que se adaptam às situações da figura 3.1. Estes modelos são lineares, pois ainda que as previsões possam ter diversas utilizações, uma das mais importantes está associada aos algoritmos de controle em tempo real.

A técnica utilizada para construção de modelos de sistemas descentralizados neste trabalho é denominada método de decomposição. Neste método obtém-se um modelo global para o sistema

no qual figuram apenas as entradas e saídas. A partir deste modelo global, são elaborados modelos para os subsistemas e suas interconexões.

Para se obter, por decomposição, o modelo de um sistema descentralizado, o ponto de partida é um modelo global como o descrito na equação (3.17).

$$A(q^{-1}) \cdot \underline{y}_K = B(q^{-1}) \cdot \underline{u}_K + \underline{\delta}_K \quad (3.17)$$

onde os vetores $\underline{u}_K$ e $\underline{y}_K$ contêm as entradas e as saídas de todos os subsistemas.

Reescrevendo a equação (3.17) de modo a tornar explícitos os termos referentes a cada subsistema, tem-se:

$$\begin{bmatrix} A^{1,1}(q^{-1}) & \dots & A^{1,i}(q^{-1}) & \dots & A^{1,ns}(q^{-1}) \\ \vdots & & \vdots & & \vdots \\ A^{i,1}(q^{-1}) & \dots & A^{i,i}(q^{-1}) & \dots & A^{i,ns}(q^{-1}) \\ \vdots & & \vdots & & \vdots \\ A^{ns,1}(q^{-1}) & \dots & A^{ns,i}(q^{-1}) & \dots & A^{ns,ns}(q^{-1}) \end{bmatrix} \begin{bmatrix} \underline{y}_K^1 \\ \vdots \\ \underline{y}_K^i \\ \vdots \\ \underline{y}_K^{ns} \end{bmatrix} = \begin{bmatrix} B^{1,1}(q^{-1}) & \dots & B^{1,i}(q^{-1}) & \dots & B^{1,ns}(q^{-1}) \\ \vdots & & \vdots & & \vdots \\ B^{i,1}(q^{-1}) & \dots & B^{i,i}(q^{-1}) & \dots & B^{i,ns}(q^{-1}) \\ \vdots & & \vdots & & \vdots \\ B^{ns,1}(q^{-1}) & \dots & B^{ns,i}(q^{-1}) & \dots & B^{ns,ns}(q^{-1}) \end{bmatrix} \begin{bmatrix} \underline{u}_K^1 \\ \vdots \\ \underline{u}_K^i \\ \vdots \\ \underline{u}_K^{ns} \end{bmatrix} + \begin{bmatrix} \underline{\delta}_K^1 \\ \vdots \\ \underline{\delta}_K^i \\ \vdots \\ \underline{\delta}_K^{ns} \end{bmatrix} \quad (3.18)$$

sendo:

$\underline{u}_K^i$ o vetor entrada do subsistema S_i com dimensão m^i

$\underline{y}_K^i$ o vetor saída do subsistema S_i com dimensão n^i

$\underline{\delta}_K^i$ o vetor ruído presente em S_i com dimensão n^i

É importante observar que a equação (3.18) não é obrigatoriamente uma versão particionada da equação (3.17). Isto ocorre somente quando não há componentes comuns nas entradas dos subsistemas. No caso geral, componentes do vetor-entrada de um subsistema podem estar presentes também no vetor-entrada de outros subsistemas e, conseqüentemente, o vetor $[\underline{u}_K^1 \dots \underline{u}_K^i \dots \underline{u}_K^{ns}]^T$ pode ter elementos repetidos e as matrizes $B^{i,j}(q^{-1})$ podem não constituir uma partição de $B(q^{-1})$. Entretanto, as matrizes $B^{i,j}(q^{-1})$ são constituídas de elementos da matriz $B(q^{-1})$ e, uma regra para escolha destas matrizes é maximizar os coeficientes das matrizes localiza

das na diagonal e reduzir os coeficientes das matrizes fora da diagonal. Esta regra é justificável, visto que medidas que podem ser realizadas no próprio subsistema não devem ser obtidas fora deste.

Como as entradas $\underline{u}_K^i$ $i = 1, \dots, ns$ podem ser geradas por mecanismos independentes das saídas torna-se difícil imaginar a identificação de um modelo entrada-saída, realizada nos subsistemas, sem que haja informação de todas as componentes importantes da entrada para cada subsistema. Assim sendo, nos modelos a seguir, as matrizes $B^{i,j}$ com $i \neq j$ não são significativamente diferentes da matriz nula.

Isolando a i -ésima linha do modelo descrito pela equação (3.18), obtêm-se :

$$A^{i,1}(q^{-1}) \cdot \underline{y}_K^1 + \dots + A^{i,i}(q^{-1}) \cdot \underline{y}_K^i + \dots + A^{i,ns}(q^{-1}) \cdot \underline{y}_K^{ns} = B^i(q^{-1}) \cdot \underline{u}_K^i + \delta_K^i \quad (3.19)$$

onde é clara a presença de termos contendo saídas de vários subsistemas, cujas medidas podem não estar disponíveis no subsistema S_i . Para eliminar estes termos, reescreve-se a equação (3.19), como segue:

$$A^i(q^{-1}) \cdot \underline{y}_K^i = B^i(q^{-1}) \cdot \underline{u}_K^i + \underline{\gamma}_K^i + \delta_K^i \quad (3.20)$$

sendo $\underline{\gamma}_K^i$ denominado erro de descentralização, uma vez que seu aparecimento é uma consequência da configuração descentralizada do sistema.

O modelo a ser utilizado na estimação do erro de descentralização deve ser equacionado a partir das medidas disponíveis em cada subsistema. A cada uma das situações da figura 3.1 corresponde um modelo diferente. Em particular, alguns modelos adequados são:

3.4.1. Comunicação sempre Possível

$$\underline{y}_K^i = \underline{A}^{i,1}(q^{-1})\underline{g}(\underline{y}_{K-1}^1) + \dots + \underline{A}^{i,i-1}(q^{-1})\underline{g}(\underline{y}_{K-1}^{i-1}) + \underline{A}^{i,i+1}(q^{-1})\underline{g}(\underline{y}_{K-1}^{i+1}) + \dots + \underline{A}^{i,ns}(q^{-1})\underline{g}(\underline{y}_{K-1}^{ns}) + \underline{\omega}_{Y_K}^i \quad (3.21)$$

onde $\underline{g}(\cdot)$ é uma função das saídas passadas de cada subsistema, sendo transmitida ao subsistema S_i . No caso mais simples, $\underline{g}(\cdot)$ é a função identidade e as matrizes $\underline{A}^{i,j}$ são iguais as matrizes $A^{i,j}|q^{-1}$ da equação (3.19). De qualquer forma, a situação (a) corresponde a algoritmos centralizados e não apresenta maiores dificuldades, visto que medidas de todos os subsistemas podem ser transmitidas, sempre que necessário.

3.4.2. Comunicação Descontínua

Se o nível de degradação $\mathcal{N}_D$ é pequeno, o modelo descrito pela equação (3.21) pode ser utilizado supondo-se $\underline{y}_K^i$ constante até que novas medidas estejam disponíveis para uma nova atualização. Com o aumento de $\mathcal{N}_D$ o modelo descrito pela equação (3.21) pode tornar-se irreal. Neste caso, propõe-se a utilização do seguinte modelo:

$$\underline{y}_K^i = \underline{A}^{i,1}(q^{-1})\underline{g}(\underline{y}_{K-1}^1) + \dots + \underline{A}^{i,i-1}(q^{-1})\underline{g}(\underline{y}_{K-1}^{i-1}) + \underline{A}^{i,i+1}(q^{-1})\underline{g}(\underline{y}_{K-1}^{i+1}) + \dots + \underline{A}^{i,ns}(q^{-1})\underline{g}(\underline{y}_{K-1}^{ns}) + \mathcal{I}^1 \underline{y}_{K-N^*-1}^i + \dots + \mathcal{I}^\alpha \underline{y}_{K-N^*-\alpha}^i + \underline{\omega}_{Y_K}^i \quad (3.22)$$

Em relação ao modelo (3.21), a equação (3.22) inclui novos termos com saídas de S_i em instantes de tempo anteriores àqueles do modelo (3.20). Uma modificação equivalente a esta seria aumentar a ordem do modelo (3.20) de N^* para $N^*+\alpha$. Naturalmente, num sistema de configuração descentralizada fixa não existe distinção entre estas alterações. Entretanto, quando ocorrem falhas num sistema, aumentando o nível de degradação $\mathcal{N}_D$, o modelo (3.22) explica a parcela do erro de descentralização devida a estas falhas.

Esta parcela pode ser omitida do modelo logo que a falha seja eliminada.

3.4.3. Comunicação com Atraso

Propõe-se a utilização do seguinte modelo incluindo toda a informação disponível:

$$\begin{aligned} \underline{y}_K^i = & \mathcal{A}^{i,1}(q^{-1})\underline{y}_{K-r}^1 + \dots + \mathcal{A}^{i,i-1}(q^{-1})\underline{y}_{K-r}^{i-1} + \\ & + \mathcal{A}^{i,i+1}(q^{-1})\underline{y}_{K-r}^{i+1} + \dots + \mathcal{A}^{i,ns}(q^{-1})\underline{y}_{K-r}^{ns} + \underline{\omega}_K^i \end{aligned} \quad (3.23)$$

onde r é igual ao índice de degradação $\mathcal{N}_A$.

3.4.4. Comunicação Descontínua com Atraso

Se o nível de degradação $\mathcal{N}_D$ é pequeno, o modelo (3.23) pode ser usado, fazendo a atualização de $\underline{y}_K^i$ sempre que novas medidas estejam disponíveis. Outro modelo é necessário, se $\mathcal{N}_D$ aumenta. Um modelo possível consiste em:

$$\begin{aligned} \underline{y}_K^i = & \mathcal{A}^{i,1}(q^{-1})\underline{y}_{K-r}^1 + \dots + \mathcal{A}^{i,i-1}(q^{-1})\underline{y}_{K-r}^{i-1} + \\ & + \mathcal{A}^{i,i+1}(q^{-1})\underline{y}_{K-r}^{i+1} + \dots + \mathcal{A}^{i,ns}(q^{-1})\underline{y}_{K-r}^{ns} + \\ & + \mathcal{I}^1 \underline{y}_{K-N^*-1}^i + \dots + \mathcal{I}^\alpha \underline{y}_{K-N^*-\alpha}^i + \underline{\omega}_K^i \end{aligned} \quad (3.24)$$

que pode ser parcialmente atualizado com medidas do subsistema S_i .

3.4.5. Comunicação Incompleta

Neste caso, o modelo de $\underline{y}_K^i$ deve ser constituído pelas informações dos subsistemas que apresentam interconexões com o subsistema S_i . Assim, representando estes subsistemas por S_a , S_b e S_c , um modelo apropriado é:

$$\begin{aligned} \underline{y}_K^i = & \underline{a}^{i,a}(q^{-1})\underline{y}_{K-1}^a + \underline{a}^{i,b}(q^{-1})\underline{y}_{K-1}^b + \underline{a}^{i,c}(q^{-1})\underline{y}_{K-1}^c + \\ & + \underline{I}^1 \underline{y}_{K-N^*-1}^i + \dots + \underline{I}^\alpha \underline{y}_{K-N^*-\alpha}^i + \underline{\omega}_{Y_K}^i \end{aligned} \quad (3.25)$$

3.4.6. Comunicação Impossível

Esta situação corresponde a descentralização completa resultante de falhas na comunicação entre subsistemas. Neste caso, o erro de descentralização pode ser modelado somente com medidas do próprio subsistema, isto é:

$$\underline{y}_K^i = \underline{I}^1 \underline{y}_{K-N^*-1}^i + \dots + \underline{I}^\alpha \underline{y}_{K-N^*-\alpha}^i + \underline{\omega}_{Y_K}^i \quad (3.26)$$

Este modelo é apenas uma solução momentânea, mas pode ser o suficiente até que as falhas sejam superadas.

Em todas as situações descritas, o modelo do ruído $\underline{\zeta}_K^i$ pode ser elaborado, com base no teorema da factorização espectral (1.3), a partir do ruído branco, ou seja:

$$D(q^{-1}) \cdot \underline{\zeta}_K^i = C(q^{-1}) \cdot \underline{\omega}_K^i \quad (3.27)$$

onde

$$E\{\omega_j^i \cdot \omega_K^i\} = \begin{cases} 0 & \text{para } j \neq K \\ V & \text{para } j = K \end{cases}$$

Com a finalidade de avaliar o desempenho de algoritmos em sistemas descentralizados, é necessária a definição de outros critérios além dos usuais para sistemas centralizados. Para evitar que pontos importantes fossem mascarados pela definição de um índice único, optou-se pela consideração de quatro índices distintos.

São definidos a seguir os índices de desempenho utilizados:

(a) Custo convencional:

$$J_{D1}^i = \sum_{K=1}^N (\text{erro de previsão})^2 \quad (3.28)$$

(b) Valor médio do número de informações de outros subsistemas recebidas pelo subsistema S_i por intervalo de amostragem indicado por

$$J_{D2}^i \quad (3.29)$$

(c) Número total de parâmetros dos modelos (modelo de S_i e modelo da interconexão) definido como

$$J_{D3}^i \quad (3.30)$$

(d) Número total das interconexões físicas no subsistema S_i (quando existentes):

$$J_{D4}^i \quad (3.31)$$

O índice J_{D1}^i é usualmente empregado para avaliar o desempenho de algoritmos centralizados. Os índices J_{D2}^i e J_{D4}^i refletem o custo das comunicações necessárias para manter em funcionamento o algoritmo e a possibilidade de ocorrência de erros. O índice J_{D3}^i mostra a capacidade local de armazenamento exigida e a complexidade do modelo.

As variações nestes índices, associadas aos níveis de degradação mencionados, serão utilizadas para caracterizar o comportamento dos algoritmos quando ocorrem falhas na comunicação entre os subsistemas.

Para o desenvolvimento de algoritmos descentralizados nas próximas três seções foram escolhidas como base as configurações estrela, anel e a rede com barramento. A configuração estre

la corresponde a muitos sistemas atualmente em funcionamento. A configuração anel tem aparecido em sistemas mais recentes e as redes com barramento despontam como uma configuração de largo uso nos próximos anos. Além disto, a existência de várias escalas de tempo nos algoritmos desenvolvidos constitui uma referência implícita a configuração hierárquica.

3.5. ESTIMADORES RECURSIVOS PARA REDES ESTRELA (ERRE)

Nas configurações estrela para redes de computação (figura 1.5.(a)) vários processadores se encontram ligados a um processador central de maior ou igual porte. Não existe comunicação direta entre os processadores localizados nas extremidades da estrela.

A seguir são desenvolvidos algoritmos de identificação descentralizada para configurações deste tipo. Naturalmente, existem diversas maneiras de localizar os subsistemas nestas redes. A possibilidade explorada nos próximos algoritmos corresponde à situação mais comum, na qual os subsistemas estão nas extremidades da estrela. O processador central recebe informações de todos os subsistemas e pode estar dedicado à identificação ou ser partilhado para outras aplicações. Outra forma de localização é considerar todas as posições, inclusive a central, ocupadas por subsistemas. A escolha do subsistema que ocupará o centro deve levar em conta as relações de causalidade entre os subsistemas e a possibilidade de manter as comunicações entre os subsistemas nas extremidades da estrela. Esta segunda forma de localização não faz parte dos algoritmos neste trabalho.

Dois esquemas de comunicação entre os processadores da rede são propostos nos algoritmos ERRE. O primeiro admite o envio de medidas pelos processadores junto aos subsistemas e o recebimento de informações processadas. Neste caso, um subsistema não recebe valores de medidas realizados em outros subsistemas, pois o processador central envia somente os resultados calculados com estas medidas. No segundo esquema não existe esta restrição. Além

de informações processadas, os subsistemas recebem medidas defasadas provenientes de outros subsistemas.

Os algoritmos desenvolvidos para o primeiro e o segundo esquema de comunicação são denominados ERRE1 e ERRE2, respectivamente. Na descrição destes algoritmos, os processadores junto aos subsistemas são referidos como processadores locais e a estimação de parâmetros feita nestes processadores é denominada estimação local. O processador no centro da rede é definido como processador central e a estimação nele realizada é dita global, pois baseia-se num modelo com todas as entradas e saídas do sistema.

3.5.1. Algoritmos ERRE1

Nestes algoritmos os processadores locais utilizam modelos da seguinte forma:

$$A^i(q^{-1}) \cdot \underline{y}_K^i = B^i(q^{-1}) \cdot \underline{u}_K^i + \underline{y}_K^i + \underline{\delta}_K^i \quad (3.32)$$

onde o índice i representa o subsistema S_i .

Nestes modelos, o erro de descentralização $\underline{y}_K^i$ é assumido constante e igual ao último valor enviado pelo processador central, sendo os aumentos excessivos do erro de previsão evitados através da escolha adequada da frequência de recepção de valores do processador central. A escolha desta frequência é discutida mais adiante. Os coeficientes das matrizes $A^i(q^{-1})$ e $B^i(q^{-1})$ são atualizados por estimação local e caso o ruído não possa ser assumido branco, deve-se modelá-lo como:

$$\underline{\delta}_K^i = C^i(q^{-1}) \cdot \underline{\omega}_K^i \quad (3.33)$$

onde a matriz $C^i(q^{-1})$ é diagonal, conforme descrito nas formas canônicas II e III (seção 3.2).

Para que no processador central calcule-se os erros de descentralização $\underline{y}_K^i$ a serem enviados a cada subsistema, as medidas

de entrada e saída de cada subsistema devem ser enviadas ao processador central. Com estas informações realiza-se, no processador central, uma estimação global de parâmetros do sistema usando o modelo:

$$A(q^{-1}) \cdot \underline{y}_K = B(q^{-1}) \cdot \underline{u}_K + \underline{\delta}_K \quad (3.34)$$

onde o ruído $\underline{\delta}_K$ pode ser ou não branco. No caso de ruído colorido utilizando o modelo das formas canônicas II ou III, isto é:

$$\underline{\delta}_K = C(q^{-1}) \cdot \underline{\omega}_K \quad (3.35)$$

para $C(q^{-1})$ diagonal. Esta forma canônica evita que as interferências entre os subsistemas atuem na parte do modelo referente ao ruído. Deste modo, as interferências existem apenas na matriz $A(q^{-1})$, pois a matriz $B(q^{-1})$ é bloco diagonal conforme descrito na seção 3.4.

A partir dos parâmetros estimados dos modelos descritos pelas equações (3.34) e (3.35) calcula-se, no processador central o erro de descentralização $\underline{y}_K^i$ por:

$$\begin{aligned} \underline{y}_K^i = & -(\hat{A}^{i,1}(q^{-1}) \cdot \underline{y}_K^1 + \dots + \hat{A}^{i,i-1}(q^{-1}) \cdot \underline{y}_K^{i-1} + \\ & + \hat{A}^{i,i+1}(q^{-1}) \cdot \underline{y}_K^{i+1} + \dots + \hat{A}^{i,ns}(q^{-1}) \cdot \underline{y}_K^{ns}) \end{aligned} \quad (3.36)$$

para $i = 1, \dots, ns$. As matrizes $\hat{A}^{i,j}(q^{-1})$ são obtidas da matriz $A(q^{-1})$ como mostrado na equação (3.19).

Os valores calculados a partir da equação (3.36) são enviados a cada processador local para a identificação do modelo de cada subsistema.

Como o cálculo dos erros de descentralização de cada subsistema envolve medidas dos demais subsistemas, deve existir um sequenciamento na amostragem e envio de medidas ao processador central, para que elas estejam disponíveis para a estimação global.

Parte desta ordem é, implicitamente, estabelecida na equação (3.36) que assume períodos de amostragem iguais para todos os subsistemas. Além disto, o envio de medidas ao processador central depende dos meios físicos existentes e da frequência desejada para atualização dos valores de y_K^i nos subsistemas. Estas afirmações podem ser ilustradas através de algumas situações. Apresentando por Δn o número de intervalos de amostragem durante os quais é possível manter constantes os erros de descentralização temos:

- i) quando Δn é desprezível torna-se necessário enviar as medidas ao processador central a cada amostragem ou em pequenos grupos, exigindo um certo sincronismo na tomada e envio de medidas dos subsistemas.
- ii) quando Δn assume valores elevados, grupos maiores de medidas podem ser enviados de uma só vez e maiores são os intervalos em que a via não é utilizada pelo algoritmo de identificação.

Se o sistema tem parâmetros variáveis no tempo, a utilização de valores elevados para Δn favorecem o emprego de algoritmos do tipo janela retangular para estimação global. Isto ocorre porque se o comprimento da janela se torna igual ao do grupo de medidas enviadas basta manter no processador central, os valores necessários para reinicialização do algoritmo cada vez que um novo grupo de medidas é recebido. Deste modo, os recursos do processador central ficam livres para outras aplicações. Entretanto, é conveniente lembrar que aumentar Δn provoca uma redução no desempenho dos algoritmos de identificação nos subsistemas.

O valor do número de intervalos de amostragens até a atualização de y_K^i , no subsistema S_i , é indicado por Δn^i e merece algumas considerações adicionais. Baseando-se em estudos preliminares, a escolha destes valores pode ser feita "off-line" determinando valores fixos que podem ser iguais ou não para todos os subsistemas. No caso de valores iguais estes são representados por Δn , como nas situações citadas. Outra alternativa é um procedimento "on-line", que atualiza o valor de Δn^i durante a estimação recursiva local como função dos erros de previsão, ou seja:

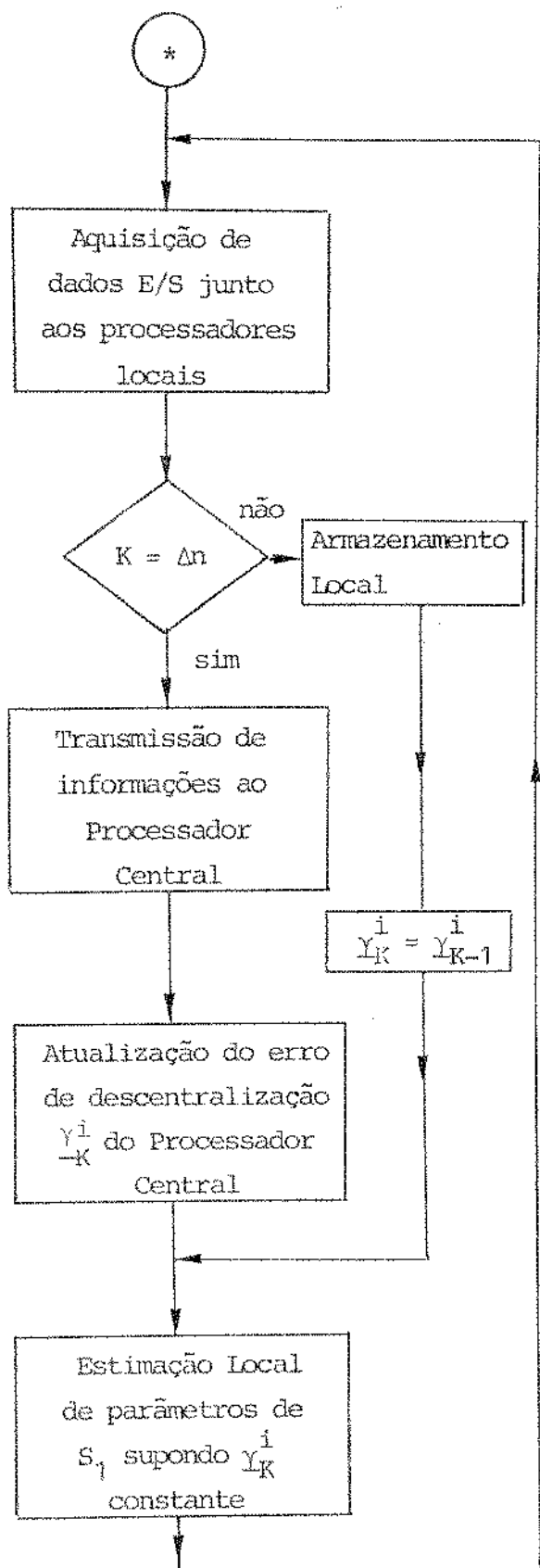
$$\Delta n_K^i = (\varepsilon_K^i, \dots, \varepsilon_{K-N}^i)_{\varepsilon^i} \quad (3.37)$$

onde $\varepsilon_K^i, \dots, \varepsilon_{K-N}^i$ são os erros locais de previsão.

Até agora discutiu-se para o algoritmo ERRE1, a escolha dos modelos a serem utilizados nos processadores, o envio de informações ao processador central e a atualização dos valores do erro de descentralização. A coordenação entre estas etapas para realização de um ciclo completo dos algoritmos é mostrada nas figuras 3.2(a) e 3.2(b). Nestas figuras o símbolo $\odot$ representa a inicialização do algoritmo.

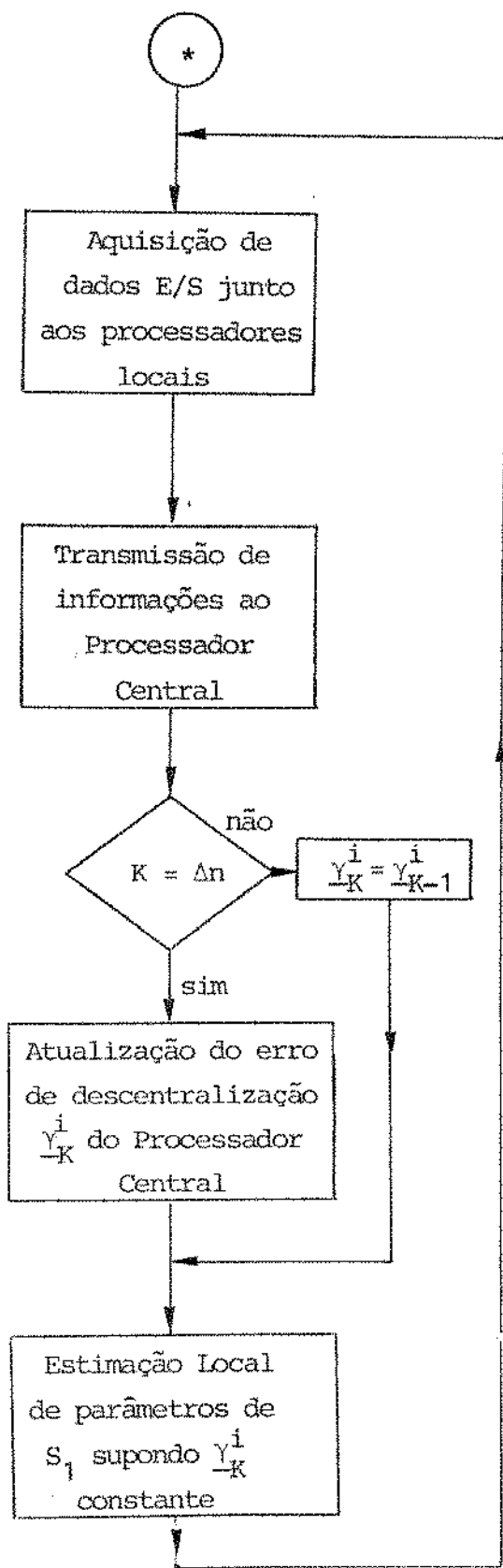
No algoritmo representado na figura 3.2(a) pode-se armazenar nos processadores locais todas as informações do subsistema. Estas informações são enviadas ao processador central somente quando se atualiza os valores de y_K^i . Portanto, no processador central armazena-se apenas os valores necessários para reinicializar a estimação quando novas informações são recebidas. Neste caso, é conveniente a adoção de um único valor Δn para todos os subsistemas, pois se apenas o subsistema S_i , requer a atualização de y_K^i , o processador central envia mensagens aos demais processadores pedindo os dados necessários. Estes dados ficam armazenados até que cada um dos subsistemas utilize-os para renovar seus erros de descentralização. Portanto, valores diferentes para Δn^i anulariam a vantagem citada de diminuir a memória permanentemente ocupada pela estimação no processador central.

Na figura 3.2(b) as medidas realizadas em cada subsistema são imediatamente enviadas ao processador central. Nos processadores locais ficam armazenadas apenas as medidas necessárias para as estimações locais. Este esquema permite a elaboração ("off-line" ou "on-line") de critérios para determinação de Δn^i em cada subsistema e relaxa a exigência de velocidade do processador central. Além disto, falhas em qualquer das linhas de comunicação podem ser facilmente detetadas pelo processador central e a perda de informações é a menor possível. Como desvantagem, a probabilidade de erros de transmissão é maior devido ao envio constante de medidas provenientes dos subsistemas, que podem estar longe do centro.



(a)

Com armazenamento Local
de medidas



(b)

Sem armazenamento Local
de medidas

Figura 3.2 - Etapas de um Algoritmo ERRE1 no Subsistema S_1

Para completar a descrição dos algoritmos ERRE1 a tabela 3.2 indica os métodos de estimação usados nos processadores locais e no processador central. São mostrados, tanto o caso de parámetros fixos como de parâmetros variáveis. Para evitar duplicidade, assumiu-se que o mesmo modelo para o ruído é usado nas estimações locais e na estimação global.

Nestes algoritmos, utilizam-se os métodos para identificação centralizada descritos no Capítulo 2, adaptando-os para identificação descentralizada pela atualização dos erros de descentralização dos subsistemas.

3.5.2. Algoritmos ERRE2

Para os algoritmos ERRE2 o esquema de comunicação admite a possibilidade dos subsistemas receberem medidas, realizadas em outros subsistemas. Estas medidas estão atrasadas em relação às medidas realizadas no próprio subsistema, devido a comunicação ser feita via processador central.

Nestes algoritmos os modelos mantidos pelos processadores locais têm a forma:

$$A^i(q^{-1}) \cdot \underline{y}_K^i = B^i(q^{-1}) \cdot \underline{u}_K^i + \underline{y}_K^i + \underline{\delta}_K^i \quad (3.38)$$

onde o índice i representa o subsistema S_i e o ruído $\underline{\delta}_K^i$ pode ser branco ou colorido. Se o ruído é colorido, torna-se necessário um modelo adicional dado como o descrito na equação (3.33).

Os parâmetros deste modelo são atualizados em cada amostragem das entradas e saídas do subsistema, fazendo uso do valor corrente do erro de descentralização $\underline{y}_K^i$.

Uma previsão para $\underline{y}_K^i$ é efetuada a cada amostragem no subsistema S_i utilizando medidas atrasadas de outros subsistemas através de:

		Processador Central	Processadores Locais	Algoritmo
Parâmetros Fixos		Mínimos Quadrados Recursivo	Mínimos Quadrados (Recursivo ou não -Recursivo)	ERRE1/1
		Matriz Estendida ou Mínimos Quadrados Particionado	Matriz Estendida ou Mínimos Quadrados Particionado	ERRE1/2
Parâmetros Variáveis no Tempo		Janela Exponencial	Janela Exponencial ou Retangular ou Mínimos Quadrados não-Recursivo	ERRE1/3
		Janela Retangular	Janela Retangular ou Mínimos Quadrados não-Recursivo	ERRE1/4

Tabela 3.2 - Métodos de Estimação de Parâmetros utilizados nos Algoritmos ERRE1

$$\begin{aligned} \hat{Y}_K^i = & \hat{A}^{i,1}(q^{-1}) \cdot \underline{Y}_{K-r}^1 + \dots + \hat{A}^{i,i-1}(q^{-1}) \underline{Y}_{K-r}^{i-1} + \\ & + \hat{A}^{i,i+1}(q^{-1}) \underline{Y}_{K-r}^{i+1} + \dots + \hat{A}^{i,ns}(q^{-1}) \underline{Y}_{K-r}^{ns} \end{aligned} \quad (3.39)$$

onde a estimação dos coeficientes das matrizes $\hat{A}^{i,j}$ é obtida a partir do seguinte modelo:

$$\begin{aligned} \underline{Y}_K^i = & A^{i,1}(q^{-1}) \cdot \underline{Y}_{K-r}^1 + \dots + A^{i,i-1}(q^{-1}) \cdot \underline{Y}_{K-r}^{i-1} + A^{i,i+1}(q^{-1}) \cdot \underline{Y}_{K-r}^{i+1} + \\ & \dots + A^{i,ns}(q^{-1}) \underline{Y}_{K-r}^{ns} + \omega_{Y_K}^i \end{aligned} \quad (3.40)$$

e deve ser atualizada por um algoritmo local sempre que novos valores para $\underline{Y}_K^i$ são fornecidos pelo processador central.

No processador central o cálculo dos erros de descentralização para os subsistemas é feito usando os parâmetros do modelo do sistema visto como um todo, a saber:

$$A(q^{-1}) \cdot \underline{Y}_K = B(q^{-1}) \cdot \underline{u}_K + \delta_K \quad (3.41)$$

onde o ruído δ_K pode ser ou não branco. No caso de ruído colorido é conveniente usar o modelo das formas canônicas II ou III descrito na equação (3.35). Da mesma forma que para os algoritmos ERRE1 o objetivo do uso destas formas canônicas é evitar a existência de interferências entre os subsistemas na parte do modelo referente ao ruído.

O cálculo do erro de descentralização realizado no processador central é:

$$\begin{aligned} \underline{Y}_K^i = & (\hat{A}^{i,1}(q^{-1}) \cdot \underline{Y}_K^1 + \dots + \hat{A}^{i,i-1}(q^{-1}) \cdot \underline{Y}_K^{i-1} + \hat{A}^{i,i+1}(q^{-1}) \cdot \underline{Y}_K^{i+1} + \\ & \dots + \hat{A}^{i,ns}(q^{-1}) \underline{Y}_K^{ns}) \end{aligned} \quad (3.42)$$

onde as matrizes $\hat{A}^{i,j}(q^{-1})$ são obtidas das matrizes $\hat{A}(q^{-1})$ da for

ma mostrada na equação (3.19) e a omissão das matrizes $B^{i,j}(q^{-1})$ é devida a suposição feita neste trabalho de que estas matrizes são nulas (equação (3.20)).

A sequência de eventos realizada nos algoritmos ERRE2 nos processadores locais está representada na figura 3.3. Ao contrário dos algoritmos ERRE1 não faz sentido supor que existe armazenamento local de medidas, pois estas devem ser continuamente trocadas, via processador central, entre os processadores locais. Nesta figura Δn^i representa o número de intervalos de amostragem até a atualização dos erros de descentralização γ_K^i pelo processador central e pode ser obtido "off-line" ou "on-line".

A tabela 3.3 apresenta os métodos usados nos processadores locais e no processador central para estimação de parâmetros nos algoritmos ERRE2. Devido ao intervalo para obtenção de um novo valor de γ_K^i do processador central conter vários intervalos de amostragem do subsistema não se aconselha o uso do algoritmo do tipo janela retangular, na estimação dos parâmetros locais do erro de descentralização.

Como os algoritmos ERRE2 obtêm nos processadores locais novas previsões para o erro de descentralização, o uso do processador central para cálculos é menor que nos algoritmos ERRE1. Por outro lado, cresce o número de transferências de valores realizadas via processador central. Além disto, nos algoritmos ERRE2 é maior o espaço de memória exigido nos processadores locais, devido às previsões locais para γ_K^i e a atualização dos parâmetros na equação (3.40).

Para que os algoritmos ERRE2 atuem eficientemente é necessário que, no caso de parâmetros fixos, a convergência dos parâmetros em $A^{i,1}(q^{-1}), \dots, A^{i,i-1}(q^{-1}), A^{i,i+1}(q^{-1}), \dots, A^{i,ns}(q^{-1})$ seja mais rápida que a convergência dos parâmetros em $A^i(q^{-1})$ e $B^i(q^{-1})$. No caso de parâmetros variáveis, os parâmetros correspondentes ao modelo local do erro de descentralização devem variar mais lentamente que os parâmetros do modelo entrada-saída do subsistema.

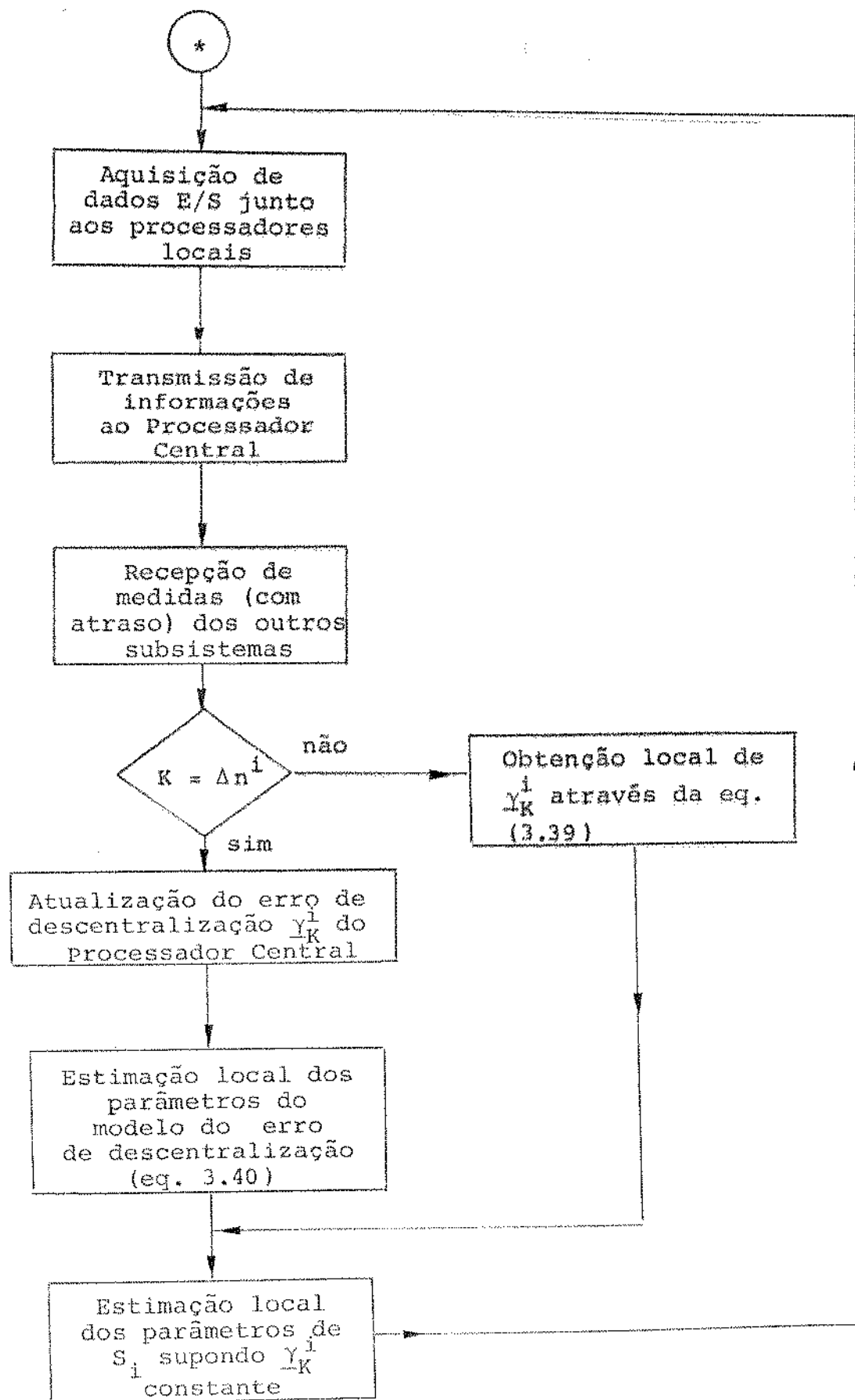


Figura 3.3 - Etapas de um Algoritmo ERRE2 no Subsistema S_i

Parâmetros Fixos	Processador Local Parâmetros do Sistema - Eq. (3.38)	Processador Local Parâmetros do Erro de Descentralização - eq. (3.40)	Processador Central Parâmetros do Sistema como um todo - eq. (3.41)	Algoritmo
	Mínimos Quadrados Recurso	Mínimos Quadrados Recurso	Mínimos Quadrados (Recurso ou não-Recurso)	ERRE2/1
Parâmetros Variáveis no Tempo	Matriz Estendida ou Mínimos Quadrados Particionado	Mínimos Quadrados Recurso	Matriz Estendida ou Mínimos Quadrados Particionado	ERRE2/2
	Janela Exponencial	Janela Exponencial	Janela Exponencial ou Retangular ou Mínimos Quadrados não-Recurso	ERRE2/3
	Janela Retangular	Janela Exponencial	Janela Retangular ou Mínimos Quadrados não-Recurso	ERRE2/4

Tabela 3.3 - Métodos de Estimacão de Parâmetros utilizados nos
Algoritmos ERRE2

Devido à configuração da rede estrela, que apresenta um processador central, as trocas de informações e os cálculos nos algoritmos ERRE deixam a impressão de não abandonar completamente o esquema centralizado de estimação, embora representando um progresso em direção a descentralização. Na próxima seção os algoritmos são desenvolvidos para uma rede anel, na qual não existe distinção posicional entre os processadores e, a descentralização é mais acentuada.

3.6. ESTIMADORES RECURSIVOS PARA REDES ANEL (ERRA)

Nas redes de computação distribuída de configuração anel cada processador está ligado a dois outros, usualmente referidos como seus vizinhos. Nestas redes, mostradas na figura 1.5.c, existe uma equivalência posicional entre os processadores, que é entendida pelos algoritmos desenvolvidos nesta seção. Os modelos, as trocas de informações e os métodos de estimação utilizados em cada um dos processadores para identificação descentralizada do sistema são semelhantes. Não há modelos centralizados e como resultado as exigências de sincronismo são menos rígidas do que nos algoritmos ERRE. Além disto, se forem trocadas apenas medidas entre os processadores vizinhos a falha de uma das linhas de comunicação interfere no desempenho do algoritmo de estimação em no máximo dois subsistemas. Entretanto, caso os processadores troquem também previsões, existe a possibilidade de haver propagação de erros por toda rede.

Neste trabalho, dois esquemas de comunicação entre os processadores são utilizados para o desenvolvimento dos algoritmos ERRA. No primeiro, existe troca de informações somente entre cada um dos processadores e seus vizinhos. Para o segundo esquema, admite-se que os processadores recebam informações defasadas de subsistemas não vizinhos. Em qualquer dos esquemas, as linhas de comunicação da rede são bidirecionais e, portanto, os dois vizinhos de um sistema tem a mesma facilidade de envio e recepção de informações com relação a este subsistema.

Os algoritmos são denominados ERRA1 ou ERRA2 conforme u tilizem o primeiro ou o segundo esquema de comunicação.

Resta observar que a inexistência de um processador central nas redes anel, dispensa o uso de um modelo para estimação global. Como consequência, a utilização de formas canônicas não é obrigatória, pois não há necessidade de garantir a concordância entre os modelos estimados pelos processadores locais e o modelo estimado no processador central.

3.6.1. Algoritmos ERRA1

Neste caso, em cada um dos processadores, o modelo utilizizado para o subsistema tem a forma:

$$A^i(q^{-1}) \cdot \underline{y}_K^i = B^i(q^{-1}) \cdot \underline{u}_K^i + \underline{y}_K^i + \underline{\delta}_K^i \quad (3.43)$$

onde o índice i representa o subsistema S_i e o ruído $\underline{\delta}_K^i$ pode ser branco ou colorido. Para ruído colorido, emprega-se o seguinte modelo:

$$D^i(q^{-1}) \cdot \underline{\delta}_K^i = C^i(q^{-1}) \cdot \underline{w}_K^i \quad (3.44)$$

onde $\underline{w}_K^i$ é um ruído branco.

Os parâmetros, neste modelo, são atualizados com base nas medidas de entrada e saída realizadas no subsistema S_i , supondo o erro de descentralização $\underline{y}_K^i$ constante.

O erro de descentralização é calculado a partir das informações recebidas dos subsistemas vizinhos S_{i-1} e S_{i+1} ⁽¹⁾. Estas

(1) Em consonância com a estrutura da rede, define-se que para o subsistema S_1 tem-se $S_{i-1} = S_{n_S}$ e, para S_{n_S} tem-se $S_{i+1} = S_1$.

informações são as parcelas $\underline{h}_{K-1}^{i,i-1}$ e $\underline{h}_{K-1}^{i,i+1}$ adicionadas à saída de S_i por influência destes subsistemas. Assim, o erro de descentralização pode ser escrito como:

$$\underline{y}_K^i = \underline{h}_{K-1}^{i,i-1} + \underline{h}_{K-1}^{i,i+1} \quad (3.45)$$

As parcelas $\underline{h}^{i,i-1}$ e $\underline{h}^{i,i+1}$ são denominadas interconexões.

Na equação (3.45) foram desprezadas as parcelas não-mensuráveis devidas aos subsistemas não vizinhos de S_i . À primeira vista, pode parecer que isto torne os algoritmos ERRA1 inadequados para identificação em sistemas nos quais não seja desprezível a influência de um subsistema em outro não-vizinho. Entretanto, as influências dos demais subsistemas estão representadas nas equações (3.43) e (3.44) pelo aumento na ordem das matrizes polinomiais.... $A(q^{-1})$, $B(q^{-1})$, $C(q^{-1})$ e $D(q^{-1})$ em relação aos modelos utilizados em algoritmos centralizados. A ordem destas matrizes é determinada através de um pré-tratamento das informações disponíveis nos subsistemas, pois com informações sobre as interconexões $\underline{h}^{i,i-1}$ e ... $\underline{h}^{i,i+1}$ o erro de descentralização descrito na equação (3.45) torna-se mensurável.

Se não ocorrem falhas de comunicação, os modelos citados são suficientes para identificação. Ainda que ocorram descontinuidades na comunicação, para valores pequenos do nível de degradação $\mathcal{N}_D$, a simples estratégia de supor $\underline{y}^i$ constante até que novos valores para $\underline{h}^{i,i-1}$ e $\underline{h}^{i,i+1}$ estejam disponíveis mantém o desempenho do estimador. Quando $\mathcal{N}_D$ aumenta excessivamente, outras estratégias para modelar $\underline{y}^i$ devem ser aplicadas. Duas alternativas são propostas a seguir.

A primeira proposta consiste em modelar $\underline{h}^{i,i-1}$ e $\underline{h}^{i,i+1}$ como descrito nas equações (3.46) e (3.47):

$$\begin{aligned} \underline{h}_{K-1}^{i,i-1} = & \mathcal{D}_-^1 \underline{h}_{K-2}^{i,i-1} - \dots - \mathcal{D}_-^{N_h} \underline{h}_{K-1-N_h}^{i,i-1} + \mathcal{G}_-^1 \underline{y}_{K-N^*-1}^i + \dots \\ & \dots + \mathcal{G}_-^{N_-} \underline{y}_{K-N^*-N_-}^i + \omega_{-K}^- \end{aligned} \quad (3.46)$$

$$\begin{aligned} \underline{h}_{K-1}^{i,i+1} = & -\mathcal{H}_+^1 \underline{h}_{K-2}^{i,i+1} - \dots - \mathcal{H}_+^{N_h} \underline{h}_{K-1-N_h}^{i,i+1} + \mathcal{G}_+^1 \underline{y}_{K-N^*-1}^i + \\ & \dots + \mathcal{G}_+^N \underline{y}_{K-N^*-N}^i + \underline{\omega}_K^+ \end{aligned} \quad (3.47)$$

onde $\underline{\omega}_K^-$ e $\underline{\omega}_K^+$ são ruídos brancos. Estes modelos são utilizados nos instantes em que não se dispõe de medidas das interconexões $\underline{h}^{i,i-1}$ e $\underline{h}^{i,i+1}$. Seus parâmetros devem ser atualizados sempre que novas medidas estão disponíveis.

Com os modelos (3.46) e (3.47) é possível atualizar separadamente os parâmetros associados a $\underline{h}^{i,i-1}$ e $\underline{h}^{i,i+1}$, caso estes valores não sejam recebidos simultaneamente. Além disto, se a comunicação com um dos subsistemas se tornar impossível o modelo da interconexão correspondente ao outro subsistema pode continuar sendo atualizado.

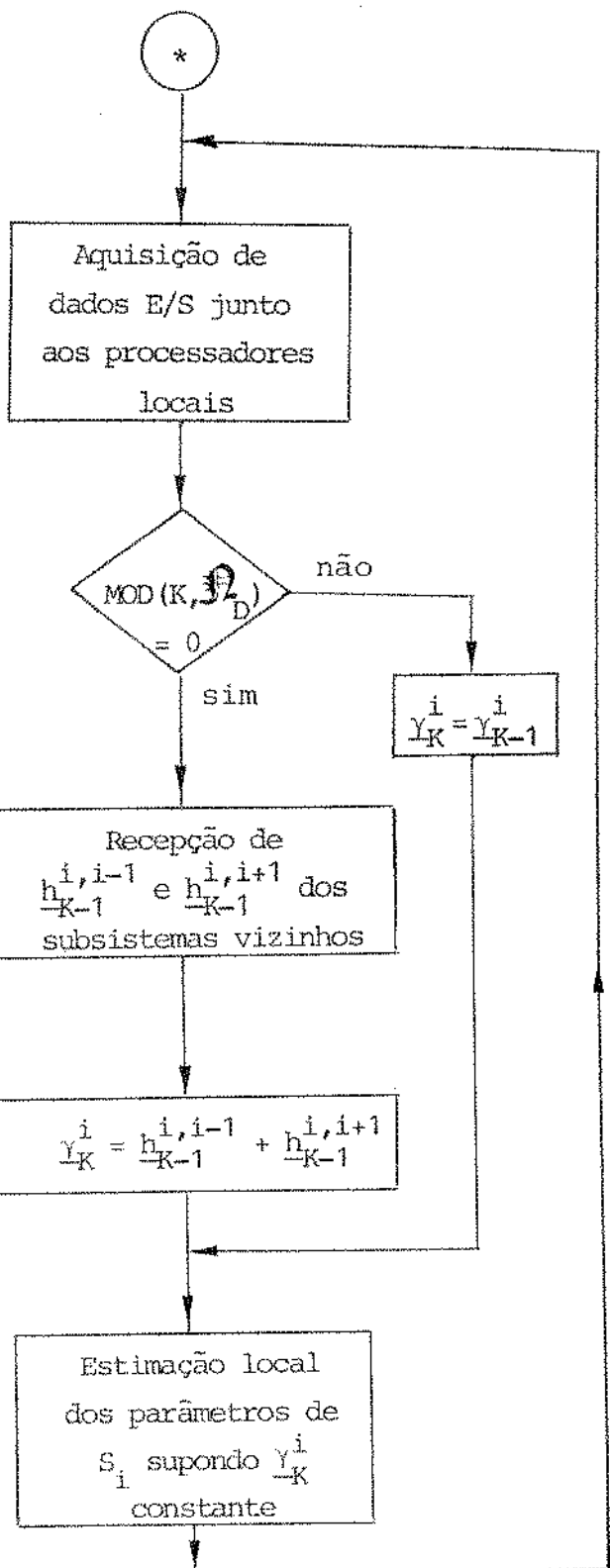
O modelo do erro de interconexão da segunda proposta é descrito na equação (3.48).

$$\begin{aligned} \underline{y}_K^i = & -\mathcal{F}^1 \underline{y}_{K-1}^i - \dots - \mathcal{F}^N \underline{y}_{K-N}^i + \mathcal{L}^1 \underline{y}_{K-N^*-1}^i + \dots \\ & \dots + \mathcal{L}^\alpha \underline{y}_{K-N^*-\alpha}^i + \underline{\omega}_{y_K}^i \end{aligned} \quad (3.48)$$

onde $\underline{\omega}_{y_K}^i$ é um ruído branco. Este modelo contém um menor número de parâmetros que o proposto anteriormente, mas sua atualização exige medidas simultâneas das interconexões $\underline{h}_{K-1}^{i,i-1}$ e $\underline{h}_{K+1}^{i,i+1}$.

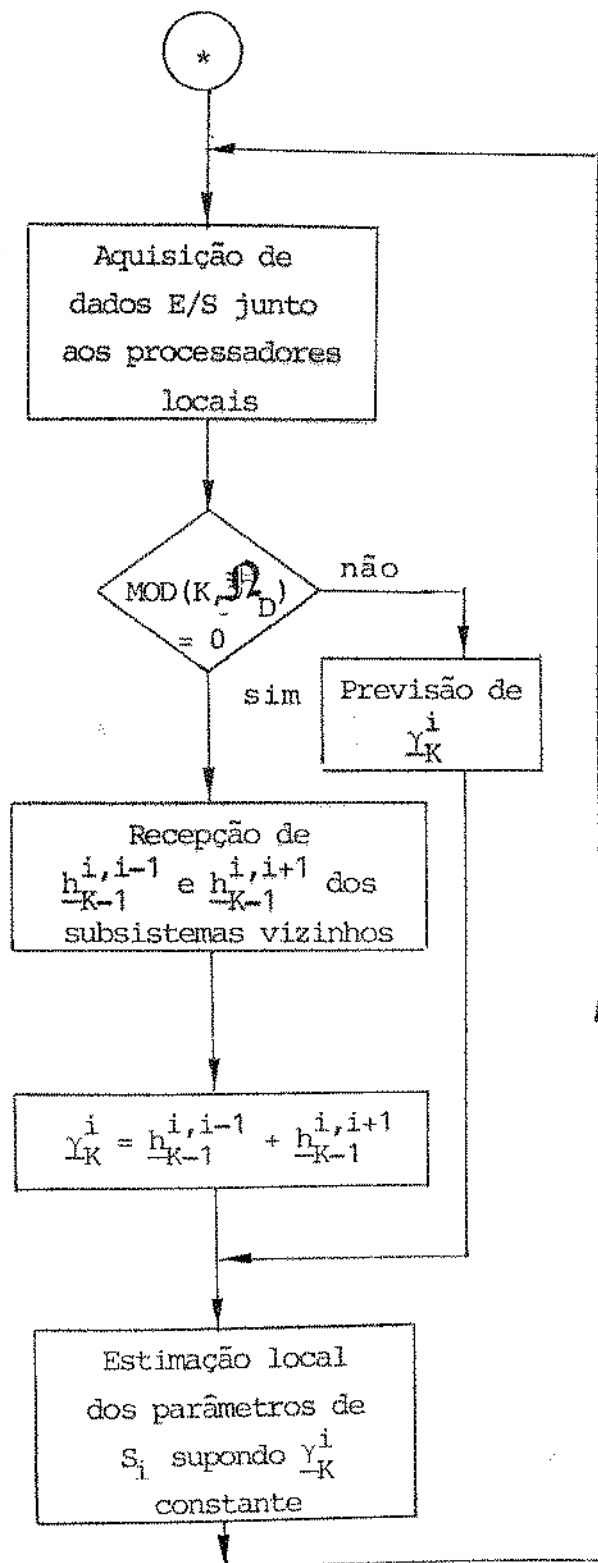
Na atualização dos modelos descritos pelas equações (3.46) e (3.47) ou pela equação (3.48) os valores correspondentes aos instantes em que não são recebidas informações dos subsistemas são preenchidos com previsões anteriormente calculadas.

O esquema de execução para o algoritmo ERRA1 está representado na figura 3.4. Para a figura 3.4(a) o valor do nível de degradação $\mathcal{D}_D$ é suficientemente pequeno para manter constante o erro de descentralização até que novas medidas de $\underline{h}^{i,i-1}$ e $\underline{h}^{i,i+1}$ sejam recebidas dos subsistemas vizinhos. Na figura 3.4(b), ao contrário, utiliza-se um modelo que permite previsões do erro de des



(a)

Sem modelo local para o erro de descentralização



(b)

Com modelo local para o erro de descentralização

Figura 3.4 - Etapas de um Algoritmo ERRAT no Subsistema S_1

centralização quando não há medidas disponíveis para $\underline{h}^{i,i-1}$ e $\underline{h}^{i,i+1}$. Estas previsões podem ser realizadas através dos modelos (3.46) e (3.47) para cada uma das parcelas $\underline{h}^{i,i-1}$ e $\underline{h}^{i,i+1}$ ou através do modelo (3.48). Naturalmente, ainda que sejam utilizados estes modelos, o valor de $\mathcal{P}_D$ pode crescer a ponto de comprometer o desempenho do estimador.

Os estimadores empregados nos algoritmos ERRA1 estão na tabela 3.4. Esta tabela abrange as situações de parâmetros fixos e variáveis no tempo, e inclui uma coluna indicando os estimadores a serem usados para $\underline{h}_K^{i,i-1}$ e $\underline{h}_K^{i,i+1}$ ou $\underline{y}_K^i$ quando houver necessidade. Não foi utilizado o algoritmo de janela retangular para estimação nesta última coluna, pois o número de medidas das interconexões numa janela é variável e pode tornar-se muito pequeno.

Como o algoritmo em cada subsistema é independente dos algoritmos nos demais subsistemas, é possível a coexistência de vários tipos de algoritmos ERRA1 numa mesma rede anel.

3.6.2. Algoritmos ERRA2

Estes algoritmos são aplicáveis às redes anel nas quais cada subsistema pode receber informações defasadas de todos os demais subsistemas. Em situações reais, se o subsistema S_i pode receber informações com atraso "a" de seus vizinhos S_{i-1} e S_{i+1} , as informações provenientes de S_{i-2} e S_{i+2} chegam a S_i com atraso maior que "a" devido a maior distância e assim sucessivamente.

O modelo utilizado para S_i no processador situado junto a este subsistema tem a seguinte forma:

$$A^i(q^{-1}) \cdot \underline{y}_K^i = B^i(q^{-1}) \cdot \underline{u}_K^i + \underline{y}_K^i + \underline{\delta}_K^i \quad (3.49)$$

onde o ruído $\underline{\delta}_K^i$ pode ser ou não branco. No caso de ruído colorido, o modelamento é feito através da equação (3.44).

A atualização dos parâmetros do modelo (3.49) é realiza

	Parâmetros do Sistema	Parâmetros de $\underline{h}_{K-1}^{i,i-1}$ e $\underline{h}_{K-1}^{i,i+1}$ ou de γ_K^i (caso necessário)	Algoritmo
Parâmetros Fixos	Mínimos Quadrados Recurso	Mínimos Quadrados Recurso	ERRA1/1
	Matriz Estendida ou Mínimos Quadrados Particionado	Mínimos Quadrados Recurso	ERRA1/2
Parâmetros Variáveis no Tempo	Janela Exponencial	Janela Exponencial	ERRA1/3
	Janela Retangular	Janela Exponencial	ERRA1/4

Tabela 3.4 - Métodos de Estimação de Parâmetros utilizados
nos Algoritmos ERRA1

da usando os medidas locais de entrada e saída e supondo $\underline{y}_K^i$ constante.

O erro de descentralização $\underline{y}_K^i$ é calculado a partir das informações enviadas pelos demais subsistemas. Cada subsistema envia a S_i a parcela correspondente à sua influência na saída de S_i . Na verdade, esta informação vem acompanhada de dois valores.

O primeiro valor identifica o subsistema que envia a informação e é utilizado para calcular o atraso, devido a distância, na recepção, por S_i . O segundo valor indica quantos intervalos de amostragem após a medida no subsistema de origem, a influência atuará na saída de S_i . A partir destes valores, é calculado, em S_i , o instante no qual a influência estará adicionada à sua saída.

No subsistema S_i , a informação sobre a influência de outro subsistema num certo instante é armazenada até que sejam recebidas informações dos demais subsistemas correspondentes ao mesmo instante. Logo que as informações provenientes de todos os subsistemas estejam disponíveis calcula-se o erro de descentralização conforme segue:

$$\underline{y}_{K-K_a}^i = \underline{h}^{i,1}_{K-K_a-1} + \dots + \underline{h}^{i,i-1}_{K-K_a-1} + \underline{h}^{i,i+1}_{K-K_a-1} + \dots + \underline{h}^{i,ns}_{K-K_a-1} \quad (3.50)$$

onde $\underline{h}^{i,j}$ é a parcela devida a influência de S_j sobre S_i . O índice K_a representa o atraso do valor calculado para o erro de descentralização em relação a amostragem feita em S_i . Em outras palavras, quando são medidas a entrada $\underline{u}_K^i$ e a saída $\underline{y}_K^i$ em S_i os dados recebidos dos demais subsistemas são suficientes para o cálculo de $\underline{y}_{K-K_a}^i$.

Não existe uma relação óbvia entre o valor de K_a e a extensão do sistema no qual é realizada a identificação descentralizada, pois ainda que as informações sobre subsistemas mais afastados demorem mais para chegar ao subsistema S_i , as influências exercidas sobre S_i podem ser medidas nesses subsistemas com maior antecedência.

Se o meio de comunicação é suficientemente rápido para assegurar valores não positivos para K_a , as equações (3.49) e (3.50)

constituem um modelo adequado para identificação em cada subsistema. Não há necessidade de previsões para $\underline{y}_K^i$ que é calculado a cada passo através da equação (3.50). Esta situação é esquematizada na figura 3.5(a).

Por outro lado, valores positivos de K_a podem tornar necessário a elaboração de modelos para previsão do erro de descentralização. Propõe-se a utilização do modelo para $\underline{y}_{K-K_a}^i$ descrito por:

$$\begin{aligned} \underline{y}_{K-K_a}^i = & -\mathcal{F}^1 \cdot \underline{y}_{K-K_a-1}^i - \dots - \mathcal{F}^{N_Y} \cdot \underline{y}_{K-K_a-N_Y}^i + \mathcal{V}^1 \\ & \underline{y}_{K-K_a-N^*-1}^i + \dots + \mathcal{V}^\alpha \underline{y}_{K-K_a-N^*-\alpha}^i + \omega_{K-K_a}^i \end{aligned} \quad (3.51)$$

onde ω_{K-K}^i é um ruído branco. Os parâmetros deste modelo são atualizados sempre que novos valores estejam disponíveis para o cálculo da equação (3.50).

Os parâmetros do modelo (3.51) são usados para o cálculo da previsão:

$$\begin{aligned} \hat{\underline{y}}_K^i = & -\mathcal{F}^1 \cdot \hat{\underline{y}}_{K-1}^i - \dots - \mathcal{F}^{K_a-1} \cdot \hat{\underline{y}}_{K-K_a+1}^i - \mathcal{F}^{K_a} \cdot \underline{y}_{K-K_a}^i - \\ & \dots - \mathcal{F}^{N_Y} \cdot \underline{y}_{K-N_Y}^i + \mathcal{V}^1 \cdot \underline{y}_{K-N^*-1}^i + \dots + \mathcal{V}^\alpha \underline{y}_{K-N^*-\alpha}^i \end{aligned} \quad (3.52)$$

onde no lugar dos valores de $\underline{y}^i$ com atrasos menores que K_a são utilizadas previsões calculadas anteriormente.

À medida que diminuem o número de previsões do lado direito da equação (3.52) diminui também o número de intervalos de amostragem entre o instante K de previsão de $\underline{y}^i$ e o instante $K-K_a$ da última estimação dos coeficientes das matrizes $\mathcal{F}$ e $\mathcal{V}$ e, conseqüentemente melhora-se a qualidade do preditor. Além disto, o aparecimento de previsões no lado direito de (3.52) restringe esta equação ao caso de parâmetros fixos ou parâmetros de variação suficientemente lenta.

A figura 3.5(b) mostra as etapas da identificação quando a equação (3.52) é utilizada para previsão do erro de descentralização.

Na tabela 3.5 apresenta-se os métodos usados nos algoritmos ERRA2 para estimação dos parâmetros das equações (3.49) e (3.51). Tanto o caso de parâmetros fixos como de parâmetros variáveis no tempo são mostrados nesta tabela.

Um tipo de sistema, ao qual parecem particularmente aplicáveis as hipóteses feitas na elaboração dos algoritmos ERRA2, são aqueles em que ocorre transporte de material entre os subsistemas. Além disto, apesar da equação (3.50) representar o erro de descentralização em S_i como função dos demais subsistemas é perfeitamente aceitável que somente alguns dentre os subsistemas contribuam para $\underline{y}^i$. Para aplicação citada, isto corresponderia ao fato de um subsistema receber material de apenas alguns dos demais subsistemas.

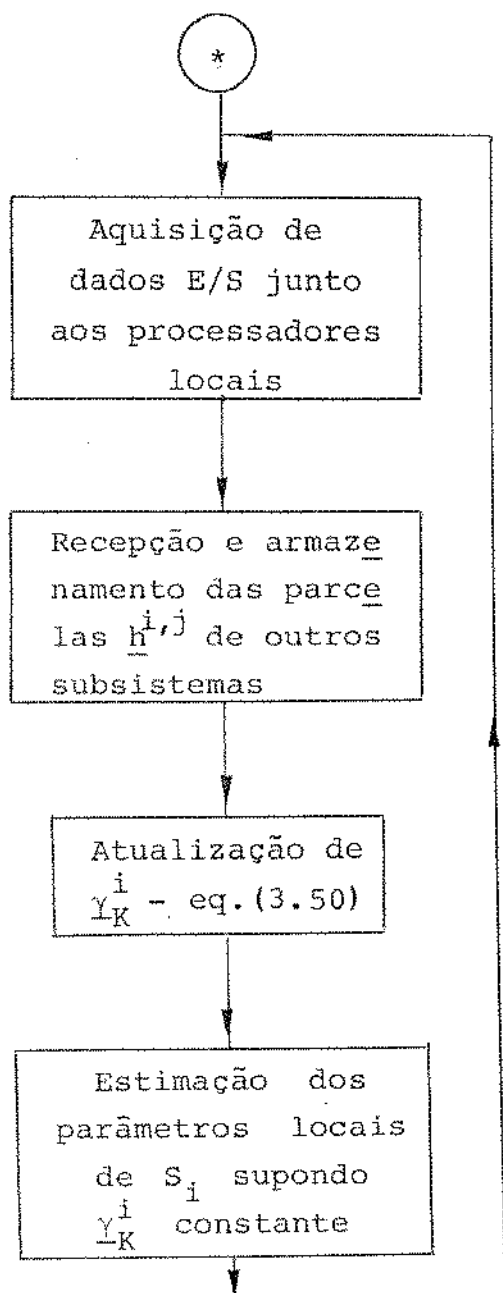
Como última ressalva, é conveniente enfatizar que por causa da troca de informações entre os subsistemas, seus intervalos para amostragem de entradas e saídas devem ser iguais.

3.7. ESTIMADORES RECURSIVOS PARA REDES COM BARRAMENTO (ERRB)

Este tipo de rede vem recebendo atenção especial nos últimos anos devido a simplicidade para construir, manter e ampliar suas instalações. Como resultado, alguns dos protocolos de comunicação mais divulgados incluindo CSMA, CSMA/CD, ALHOA, "TOKEN PASSING" destinam-se a redes com barramento.

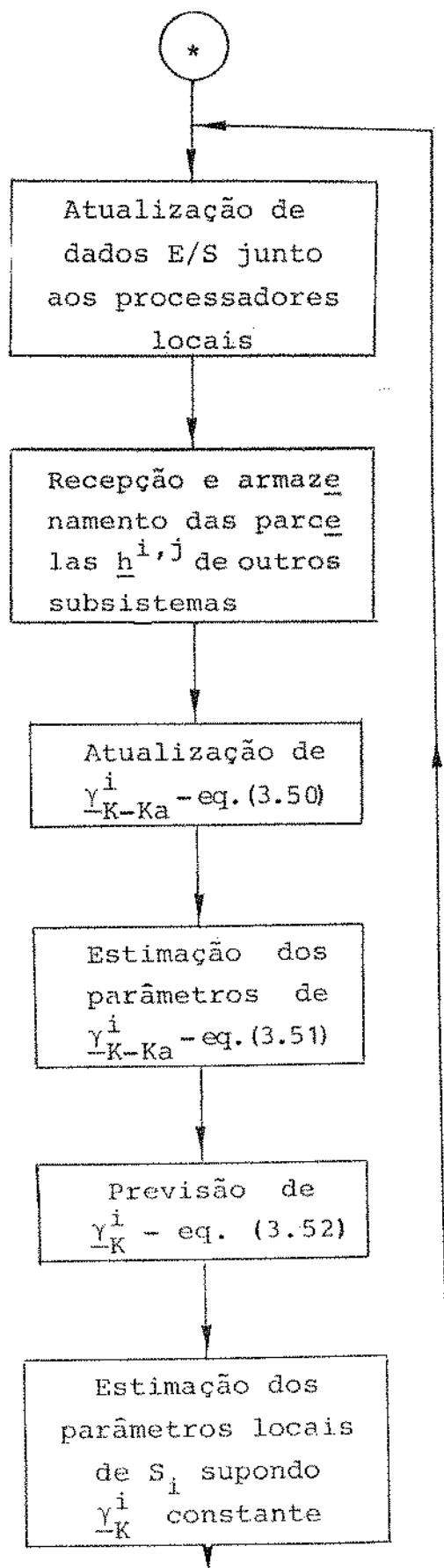
A forma mais usual do barramento é a linear representada na figura 1.5(b), mas existe também o barramento circular mostrado na figura 3.6. A primeira vista, este barramento tem certa semelhança com a rede anel, mas, nesta rede, os processadores não interrompem o meio de comunicação.

Comparando-se a rede estrela e o barramento circular, observa-se que o elemento central inteligente na rede estrela foi



(a)

Sem a utilização de preditor para $\underline{\gamma}_K^i$



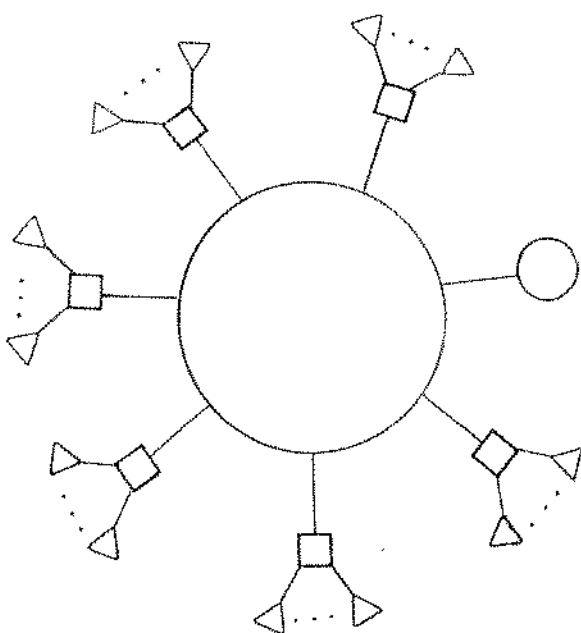
(b) Com a utilização do preditor para $\underline{\gamma}_K^i$

Figura 3.5 - Etapas de um Algoritmo ERRAS no Subsistema S_i

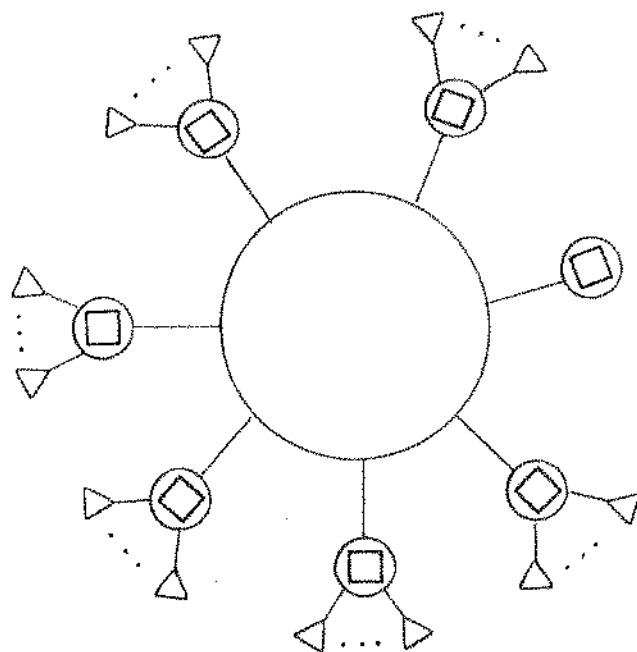
Parâmetros Fixos	Parâmetros do Subistema	Parâmetros do modelo de $y_{K-K_a}^i$ (caso necessário)	Algoritmo
	Mínimos Quadrados Recursivo	Mínimos Quadrados Recursivo	ERRA2/1
	Matriz Estendida ou Mínimos Quadrados Particionado	Mínimos Quadrados Recursivo	ERRA2/2
Parâmetros Variáveis no Tempo	Janela Exponencial	Janela Exponencial*	ERRA2/3
	Janela Retangular	Janela Retangular*	ERRA2/4

* somente parâmetros com variação suficientemente lenta

Tabela 3.5.- Métodos de Estimação de Parâmetros utilizados nos Algoritmos ERRA2



(a) Controle Centralizado
das Comunicações na
Rede



(b) Controle Distribuído das
Comunicações na Rede

Figura 3.6 - Barramento Circular

* substituído por um elemento passivo de comunicação - o barramento .

Aproveitando a operação em difusão múltipla que possibilita a recepção de uma mensagem por todos processadores ligados ao barramento e, respeitando o limite de transmissão de uma única mensagem de cada vez, diversos esquemas de comunicação são apropriados para identificação recursiva descentralizada.

Num destes esquemas, cada subsistema coloca suas medidas no barramento para que os demais as recebam. Supondo que, em cada subsistema, exista memória e no intervalo entre duas estimações sejam recebidas informações de todos os subsistemas, é possível que cada subsistema disponha de um modelo incluindo entradas e saídas de todos os subsistemas.

Os algoritmos ERRB1 foram desenvolvidos para uma situação apenas um pouco diferente da anterior. É mantida a recepção por todos os subsistemas das medidas que cada um coloca no barramento, mas o intervalo entre estimações não é suficiente para o envio de medidas por todos os subsistemas. Isto ocorre quando a transmissão no barramento é lenta em relação à capacidade de amostragem e a necessidade de atualização dos modelos nos subsistemas.

Em outro tipo de esquema de comunicação estão baseados os algoritmos ERRB2 que não ocupam um espaço grande na memória de todos os subsistemas. A idéia básica é que num dos subsistemas exista um processador de maior porte responsável pela recepção, processamento e envio de informações aos demais. Se o intervalo entre estimações for suficiente para o envio de informações para atualização dos modelos locais mantidos nos demais subsistemas, os algoritmos ERRE1 podem ser empregados. Entretanto, para alguns subsistemas pode ser conveniente atualizar rapidamente seus modelos locais, utilizando para tanto intervalos entre estimações reduzidos que podem tornar impossível o envio de informações para todos os subsistemas. Esta é a situação tratada nos algoritmos ERRB2.

Nos estimadores para redes com barramento, a capacidade de difusão múltipla é explorada resultando na exigência de sincronismo entre os estimadores nos subsistemas. Isto ocorre porque: i) a troca de informações entre os subsistemas não é útil para estimação

de parâmetros se os intervalos entre amostragens e estimações não são iguais; ii) o atraso de um subsistema em enviar suas medidas pode comprometer a estimacão nos demais. Para evitar a transmissão de informações inúteis, deve-se estabelecer um limite máximo de tempo para o envio das medidas pelos subsistemas. Após este limite, os subsistemas devem aguardar a oportunidade para o envio das medidas obtidas na próxima amostragem.

3.7.1. Algoritmos ERB1

Neste algoritmo supõe-se que o intervalo entre estimacões é demasiado pequeno para que todos os subsistemas enviem suas medidas ao barramento. Cada subsistema deve colocar no barramento, um conjunto de medidas que pode incluir todas as medidas realizadas desde a última vez que este subsistema teve acesso ao barramento ou apenas as medidas mais recentes. O número de medidas entrada/saída, a ser enviado, é determinado de acordo com os modelos utilizados nos demais subsistemas.

A seleção de quais subsistemas devem colocar no barramento suas medidas num determinado intervalo de amostragem pode ser realizada de forma estática ou dinâmica e o modelo de cada subsistema deve considerar o esquema de seleção adotado.

Para seleção estática, isto é, fixada com base em dados experimentais anteriores, a equação (3.53) representa o modelo local do subsistema S_i .

$$A(q^{-1}) \cdot \underline{y}_K^{i*} = B(q^{-1}) \cdot \underline{u}_K^{i*} + \underline{\delta}_K^{i*} \quad (3.53)$$

onde:

$$\underline{y}_K^{i*} = \begin{bmatrix} y_{K-r}^{i1} \\ \vdots \\ y_K^i \\ \vdots \\ y_{K-r}^{ins} \end{bmatrix} \quad \underline{u}_K^{i*} = \begin{bmatrix} u_{K-r}^{i1} \\ \vdots \\ u_K^i \\ \vdots \\ u_{K-r}^{ins} \end{bmatrix} \quad \underline{\delta}_K^{i*} = \begin{bmatrix} \delta_{K-r}^{i1} \\ \vdots \\ \delta_K^i \\ \vdots \\ \delta_{K-r}^{ins} \end{bmatrix}$$

O retardo r^i nas entradas, saídas e ruídos dos demais subsistemas tem por finalidade permitir que todas as medidas requeridas em cada passo da atualização recursiva dos parâmetros deste modelo estejam disponíveis no subsistema. A determinação deste retardo está ligada a sequência na qual os subsistemas enviam suas medidas.

Caso o ruído δ_K^* não seja branco, deve ser modelado pela equação (3.44).

A figura 3.7.a sintetiza o procedimento seguido em cada subsistema para estimação de parâmetros com seleção estática dos subsistemas que colocam no barramento suas medidas.

A seleção dinâmica, realizada enquanto a estimação está em curso, implica na inexistência de uma sequência fixa para envio de medidas pelos subsistemas. A seleção é feita de acordo com os erros de descentralização em cada subsistema. Se o controle do acesso ao barramento for centralizado, cada subsistema deve enviar ao controlador central a componente de maior magnitude de seu erro de descentralização. Para redes com barramento, em que cada subsistema pode tomar a iniciativa de acesso, a estratégia de seleção dinâmica é diferente. É necessário elaborar duas sequências para acesso: a sequência de acesso opcional e a sequência de acesso obrigatório. Cada uma destas sequências é iniciada quando a rede começa a operar e repetida ciclicamente. Cada vez que uma das sequências é empregada, ela é retomada a partir do ponto atingido na última vez. A sequência de acesso opcional consiste de pequenos intervalos nos quais cada um dos subsistemas pode iniciar o envio de dados de acordo com o valor de seus erros de descentralização. Quando um número mínimo de subsistemas não toma iniciativa de enviar suas medidas, entra em vigor a sequência de acesso obrigatório que consiste de intervalos nos quais cada subsistema deve enviar seus dados. A função da sequência de acesso obrigatório é impedir que o barramento fique ocioso no intervalo entre estimações. Tanto a sequência de acesso opcional como a sequência de acesso obrigatório incluem todos os subsistemas.

Quando é adotada uma seleção dinâmica dos subsistemas que acessam o barramento para envio de dados, o uso do modelo (3.53) não

garante que todas as medidas necessárias num passo de estimação es-
tejam disponíveis. Esta deficiência é compensada pelo modelo (3.54)
para o erro de descentralização.

$$\underline{y}_K^i = -\underline{F}^1 \cdot \underline{y}_{K-1}^i - \dots - \underline{F}^{N_Y} \cdot \underline{y}_{K-N_Y}^i + \underline{U}^1 \cdot \underline{y}_{K-N^*-1}^i + \dots + \underline{U}^\alpha \cdot \underline{y}_{K-N^*-\alpha}^i + \frac{\omega^i}{\gamma_K} \quad (3.54)$$

Este modelo é usado na previsão do erro de descentrali-
zação, quando não estão disponíveis no subsistema S_i as medidas de
outros subsistemas. Neste caso, para atualização do modelo (3.53),
propõe-se a utilização do modelo local descrito pela eq. (3.55),

$$\underline{y}_K^i = -\underline{A}^i(q^{-1}) \cdot \underline{y}_{K-1}^i + \underline{B}^i(q^{-1}) \cdot \underline{u}_{K-1}^i + \underline{y}_K^i + \underline{\delta}_K^i \quad (3.55)$$

Na atualização dos parâmetros estimados para o modelo (3.
54) o cálculo de $\underline{y}_K^i$ é realizado através da equação (3.56) toda vez
que as medidas dos demais subsistemas estão disponíveis em S_i . Os
parâmetros $\hat{A}^{i,j}(q^{-1})$ são os mesmos estimados para o modelo (3.53).

$$\underline{y}_K^i = -(\hat{A}^{i,1}(q^{-1}) \cdot \underline{y}_{K-r}^1 + \dots + \hat{A}^{i,i-1}(q^{-1}) \cdot \underline{y}_{K-r}^{i-1} + \dots + \hat{A}^{i,i+1}(q^{-1}) \cdot \underline{y}_{K-r}^{i+1} + \dots + \hat{A}^{i,ns}(q^{-1}) \cdot \underline{y}_{K-r}^{ns}) \quad (3.56)$$

O procedimento seguido na estimação de parâmetros com se-
leção dinâmica dos subsistemas que colocam suas medidas no barramen-
to é resumido na figura 3.7(b).

A tabela 3.6 mostra os métodos utilizados nos algoritmos
ERRB1 para estimação de parâmetros.

3.7.2. Algoritmos ERB2

Nos algoritmos ERB2, o espaço de memória dedicado à i-
dentificação em cada subsistema é menor que nos algoritmos ERB1.
A exceção é o subsistema no qual está um processador de maior por-
te que recebe medidas e envia os valores dos erros de descentrali-

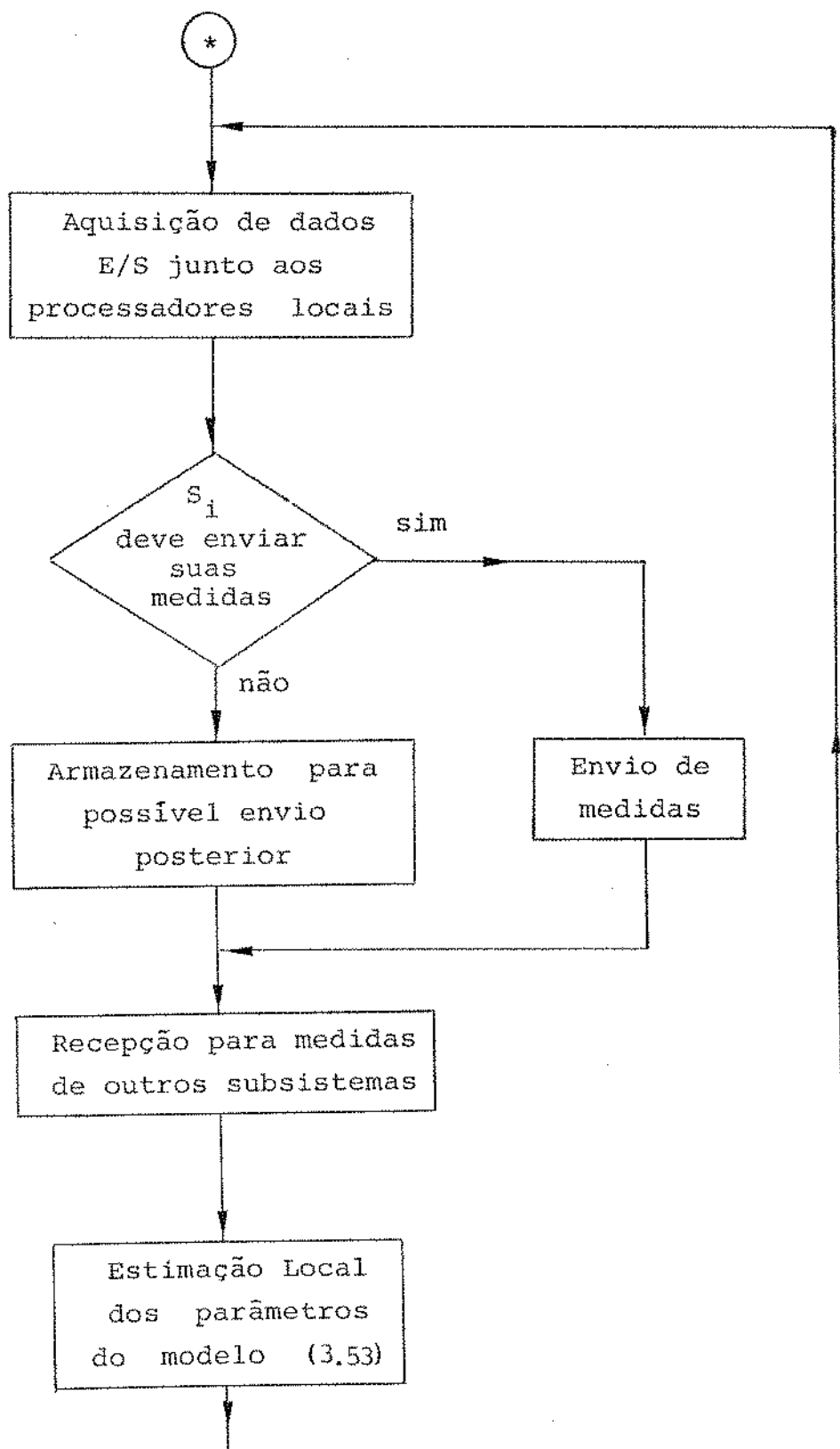


Figura 3.7(a) - Procedimento no Subsistema S_1 para Estimação de Parâmetros com Seleção Estática dos Subsistemas que acessam o Barramento para Envio de Medidas

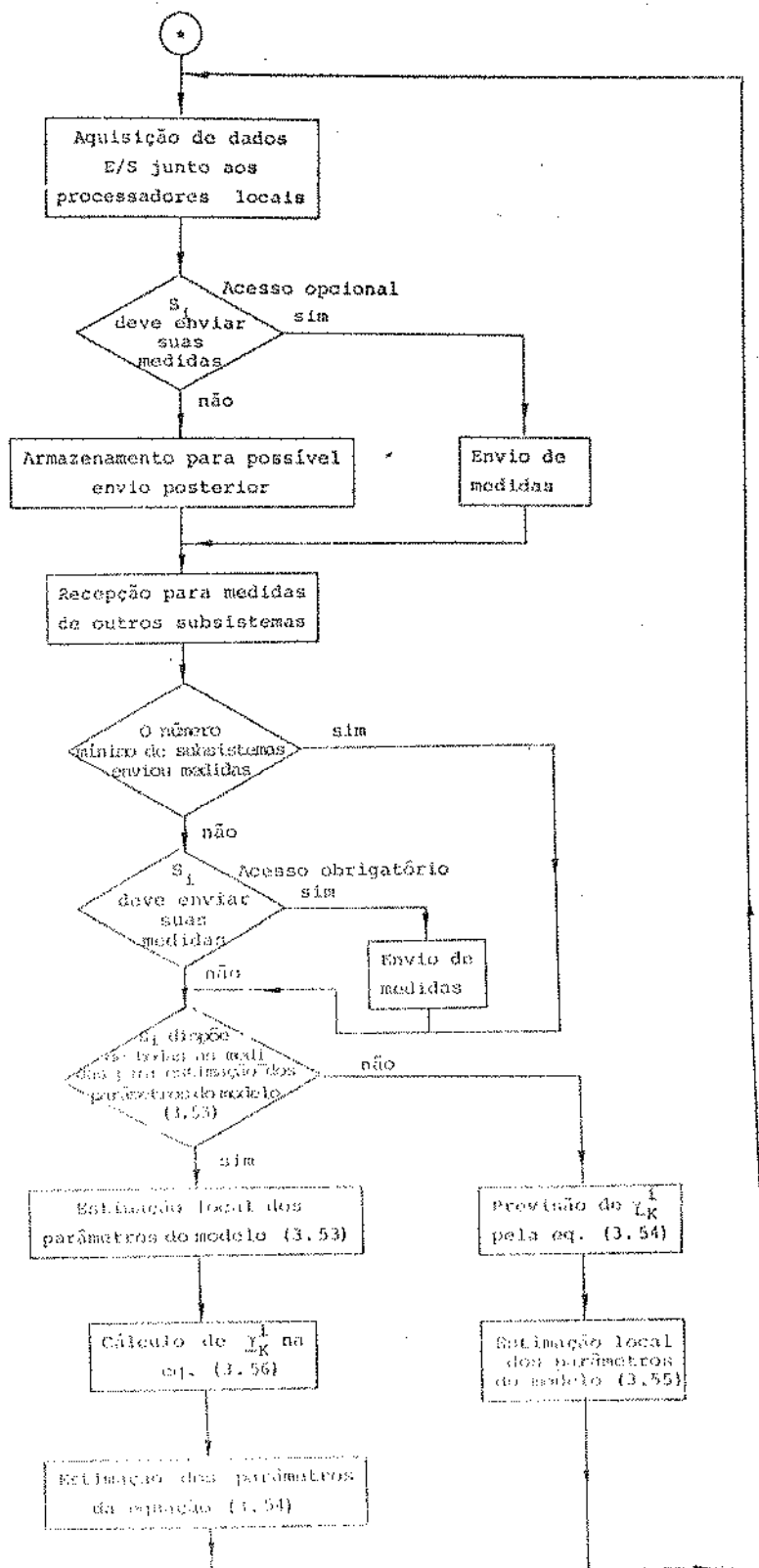


Figura 3.7(b) - Procedimento no subsistema S_1 para estimação de parâmetros com seleção dinâmica dos subsistemas que acessam o barramento para envio de medidas

		Parâmetros do Sistema	Parâmetros do modelo de γ_K^i (selação dinâmica)	Algoritmo
Parâmetros Fixos		Mínimos Quadrados Recursivo	Mínimos Quadrados Recursivo	ERRB1/1
		Matriz Estendida ou Mínimos Quadrados Recursivo	Mínimos Quadrados Recursivo	ERRA1/2
Parâmetros Variáveis no Tempo		Janela Exponencial	Janela Exponencial	ERRB1/3
		Janela Retangular	Janela Exponencial	ERRB1/4

Tabela 3.6 - Métodos de Estimação de Parâmetros usados nos Algoritmos ERB1

zação aos demais subsistemas.

Nestes algoritmos supõem-se que o intervalo entre estimações não é suficiente para que todos os subsistemas recebam valores atualizados de seus erros de descentralização.

Os fatores considerados na escolha do intervalo entre estimações são a capacidade de amostragem e a taxa mínima aceitável para atualizar o erro de descentralização de cada subsistema. Em particular, a duração máxima admissível para o intervalo entre amostragens é determinada pela taxa mínima de atualização do erro de descentralização em todos os subsistemas.

A seleção dos subsistemas que devem receber valores atualizados para seus erros de descentralização num determinado intervalo entre estimações pode ser estática ou dinâmica.

A seleção estática obedece a uma sequência preestabelecida.

Para a seleção dinâmica, os subsistemas enviam ao processador de maior porte, além de suas medidas, a componente de maior valor absoluto de seu erro local de previsão. A seleção feita leva em conta os erros locais de previsão e do número de intervalos consecutivos durante os quais cada subsistema não recebeu um valor atualizado de seu erro de descentralização.

O modelo local de cada subsistema independe da forma de seleção adotada.

Para cada subsistema, exceto aquele onde se encontra o processador de maior porte, o modelo local é descrito pela equação (3.57).

$$A^i(q^{-1}) \cdot \underline{y}_K^i = B^i(q^{-1}) \cdot \underline{u}_K^i + \underline{y}_K^i + \underline{\delta}_K^i \quad (3.57)$$

Nesta equação, se o ruído $\underline{\delta}_K^i$ for colorido deve ser modelado pela equação (3.33). Nesta equação a matriz $C^i(q^{-1})$ é diagonal permitindo a concordância entre os modelos locais para o ruído e o modelo utilizado no processador de maior porte.

O erro de descentralização $\underline{y}_K^i$ usado na estimação dos pa

parâmetros do modelo (3.57) pode ser fornecido pelo processador de maior porte ou estimado localmente. Para estimações locais é utilizado do modelo descrito na equação (3.58) a seguir:

$$\begin{aligned} \underline{y}_K^i = & -\mathcal{F}^1 \cdot \underline{y}_{K-1}^i - \dots - \mathcal{F}^{N_Y} \cdot \underline{y}_{K-N_Y}^i + \\ & + \mathcal{L}^1 \cdot \underline{y}_{K-N^*-1}^i + \dots + \mathcal{L}^\alpha \cdot \underline{y}_{K-N^*-\alpha}^i + \frac{\omega}{\gamma_K} \underline{y}_K^i \end{aligned} \quad (3.58)$$

Os parâmetros deste modelo são atualizados quando são recebidos do processador de maior porte valores do erro de descentralização.

No subsistema com um processador de maior porte existe um modelo completo do sistema na forma:

$$A(q^{-1}) \cdot \underline{y}_K = B(q^{-1}) \cdot \underline{u}_K + \underline{\hat{\epsilon}}_K \quad (3.59)$$

sendo $\underline{\hat{\epsilon}}_K$ o ruído. Caso se trate de um ruído colorido, emprega-se o modelo descrito na equação (3.35).

Quando novas medidas são recebidas dos demais subsistemas, os parâmetros do modelo descrito pela equação (3.59) são atualizados e os erros de descentralização de cada subsistema (S_i) são calculados através da equação (3.60)

$$\begin{aligned} \underline{y}_{K+1}^i = & (\hat{A}^{i,1}(q^{-1}) \cdot \underline{y}_K^i + \dots + \hat{A}^{i,i-1}(q^{-1}) \cdot \underline{y}_K^{i-1}(q^{-1}) + \hat{A}^{i,i+1}(q^{-1}) \cdot \\ & \cdot \underline{y}_K^{i+1} + \dots + \hat{A}^{i,ns}(q^{-1}) \cdot \underline{y}_K^{ns}) \end{aligned} \quad (3.60)$$

Embora apenas alguns destes erros de descentralização sejam enviados aos subsistemas correspondentes, todos são calculados e armazenados no processador de maior porte. Estes valores são utilizados na estimação de parâmetros caso alguns subsistemas não enviem suas medidas ao processador de maior porte. Neste caso, são atualizados os parâmetros correspondentes aos subsistemas que enviam medidas, isto é, as matrizes na diagonal da matriz $A(q^{-1})$ do modelo descrito pela equação (3.59). Nesta atualização, o modelo usado para o subsistema S_i tem a forma da equação (3.57) que corresponde a i -ésima linha da equação (3.59). Além disto, os erros de

descentralização enviados pelo processador de maior porte aos sub sistemas, são iguais aos valores calculados na última vez na qual as medidas de todos os subsistemas estavam disponíveis.

A figura 3.8 apresenta o procedimento seguido em cada um dos subsistemas para a estimação de parâmetros nos algoritmos ERRE2.

Os métodos de estimação usados são mostrados na tabela 3.7. São considerados tanto o caso de parâmetros fixos como o de parâmetros variáveis no tempo. Assume-se ainda que o mesmo modelo para o ruído é usado nas estimações realizadas em todos os proces sadores.

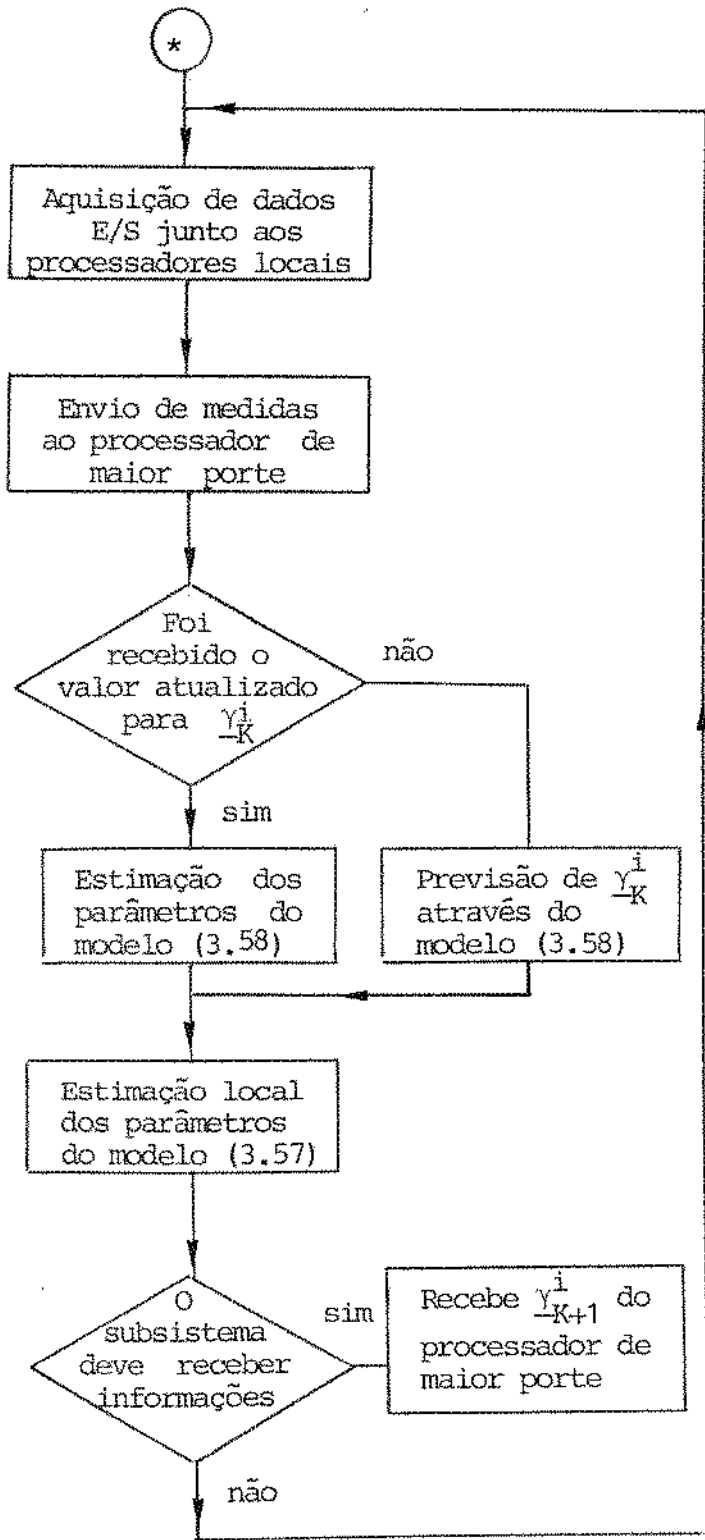
Os algoritmos propostos não esgotam os esquemas de comu nicação entre subsistemas ligados a um barramento. Para um sistema particular, um estudo das relações de causalidade entre os subsiste mas pode conduzir a esquemas de comunicação adequados para identi ficação descentralizada utilizando barramento.

3.8. CONCLUSÕES

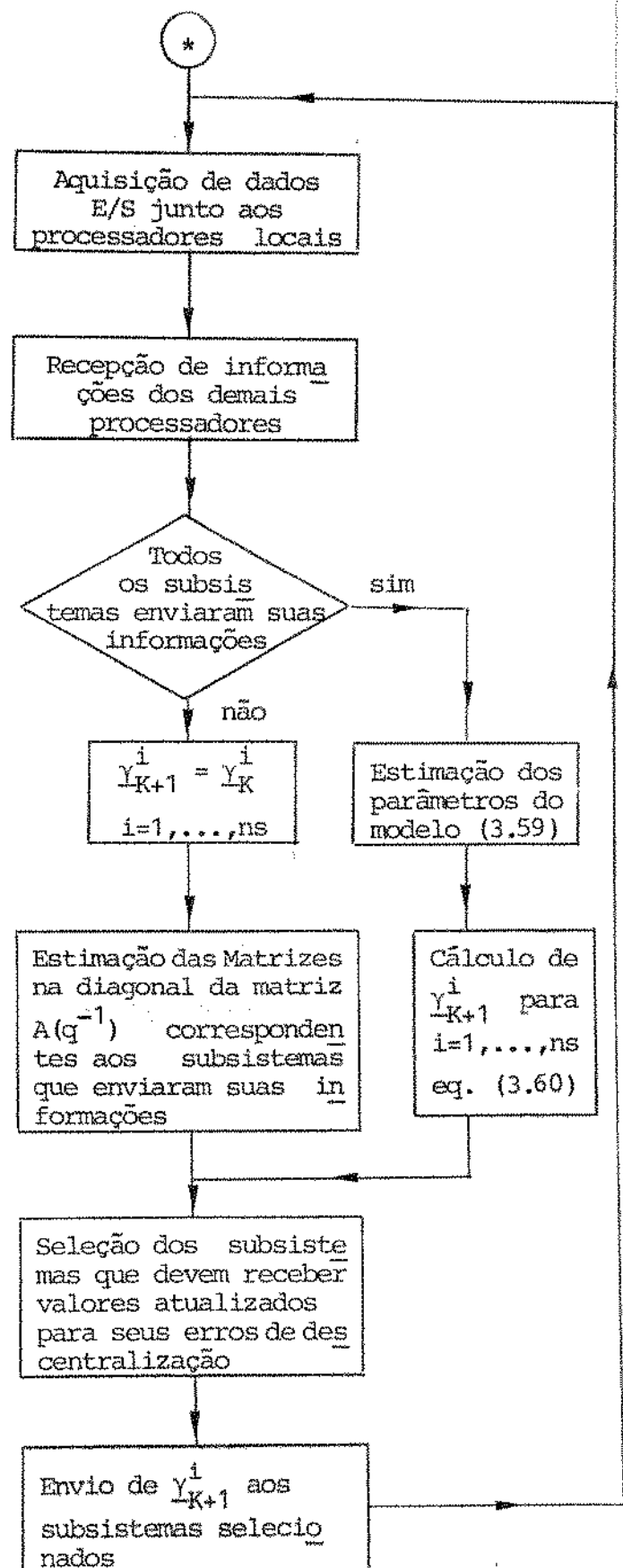
Neste capítulo foram elaborados algoritmos de estimação para sistemas descentralizados com rede de computação do tipo es trela, anel ou barramento.

Para descrição de tais algoritmos introduziu-se para sis temas descentralizados os conceitos de causalidade entre subsiste mas e de níveis de descentralização. O método de modelamento por decomposição é a base comum na elaboração de algoritmos descentra lizados adequados para diferentes tipos de rede.

Os algoritmos para rede estrela são subdivididos em algo ritmos ERRE1 e ERRE2. Nos algoritmos ERRE1 supõem-se que os proces sadores, situados junto aos subsistemas, enviam suas medidas ao processador central e recebem deste apenas os resultados de opera ções realizadas sobre estas medidas. Nos algoritmos ERRE2 o proces sador central pode também servir como caminho para troca de infor mações entre subsistemas, de modo que um subsistema pode receber ,



(a) S_i é um Subsistema com Processador de Pequeno Porte



(b) S_i é o Subsistema com Processador de maior porte

Figura 3.8 - Etapas de um Algoritmo ERB2 no Subsistema S_i

		Processador de maior porte	Demais processadores	Algoritmo
Parâmetros Fixos		Mínimos Quadrados Recursivo	Mínimos Quadrados Recursivo	ERRB2/1
		Mínimos Quadrados Particionado ou Matriz Estendida	Mínimos Quadrados Particionado ou Matriz Estendida	ERRB2/2
Parâmetros Variáveis no Tempo		Janela Exponencial	Janela Exponencial	ERRB2/3
		Janela Retangular	Janela Retangular	ERRB2/4

Tabela 3.7 - Métodos de Estimação de Parâmetros utilizados nos Algoritmos ERB2

com certo atraso, medidas realizadas em outro subsistema.

As redes estrela podem também ser classificadas como redes hierárquicas de dois níveis. O processador central ocupa o nível superior e os processadores nas "pontas" da estrela ocupam o nível inferior.

Os algoritmos para a rede anel são classificados como ERRA1 ou ERRA2, conforme a troca de informações com um subsistema seja limitada a seus vizinhos ou possa incluir também outros subsistemas mais distantes.

Os algoritmos ERRB1 para redes com barramento exploram a operação em difusão múltipla que permite a recepção de uma mensagem por todos os processadores. Os algoritmos ERRB2 são aplicáveis quando o espaço de memória disponível para identificação é reduzido e adotam uma organização hierárquica de dois níveis com um processador de maior porte no nível superior e os demais processadores no nível inferior.

A possibilidade de processamento descentralizado de informações num sistema constitui uma promessa de simplificação para o tratamento de alguns problemas ligados a identificação, como o número de parâmetros envolvidos em cada etapa de modelamento e a determinação da ordem adequada para o modelo de cada subsistema. Entretanto, a determinação das informações a serem trocadas pelos subsistemas exige um modelamento elaborado.

Em todos os algoritmos descentralizados o método de mínimos quadrados pode ser substituído pelo estimador da raiz quadrada apresentado na seção 2.5.2.

No próximo capítulo são analisados os desempenhos dos algoritmos ERRA1/1 e ERRB2/1 através de exemplos de simulação.

CAPÍTULO 4

ANÁLISE DOS ALGORÍTMOS ERRA1/1 e ERRB2/1 ATRAVÉS DE EXEMPLOS DE SIMULAÇÃO

- 4.1. Introdução
- 4.2. Critérios para Avaliação do Desempenho de Algoritmos de Estimação Descentralizada
- 4.3. Geração das Medidas e Inicialização
 - 4.3.1. Geração das Medidas
 - 4.3.2. Inicialização
- 4.4. Descrição por Equações à Diferenças dos Sistemas Simulados
- 4.5. Algoritmo ERRA1/1
- 4.6. Algoritmo ERRB2/1

4.1. INTRODUÇÃO

Neste capítulo apresentam-se alguns resultados obtidos na simulação de exemplos que ilustram o comportamento dos algoritmos ERRA1/1 e ERRB 2/1.

A seção 4.2. contém uma apresentação dos critérios utilizados para avaliação do desempenho dos algoritmos. A geração de medidas e a inicialização dos algoritmos é descrita na seção 4.3.

Os resultados obtidos para o algoritmo ERRA1/1 estão na seção 4.5.

As simulações realizadas incluem a ocorrência de descon^{tinu}idades de comunicação entre os subsistemas e utilizam os seguintes modelos para as interconexões:

- i) as interconexões são supostas constantes até que novas medidas estejam disponíveis
- ii) modelos ARMA descritos nas equações (3.46) e (3.47)

O desempenho dos modelos (i) e (ii) é comparado para diferentes níveis de degradação de descontinuidade.

Na seção 4.6 são descritos os resultados da simulação do algoritmo ERRB 2/1.

Os algoritmos de identificação descentralizada ERRA1 / 1 e ERRB 2/1 são comparados ao algoritmo centralizado de mínimos quadrados.

Os sistemas simulados consistem de dois subsistemas sendo a interação entre eles variável.

4.2. CRITÉRIOS PARA AVALIAÇÃO DOS ALGORÍTMOS DE ESTIMAÇÃO DESCENTRALIZADA

Com a finalidade de comparar o desempenho experimental dos algoritmos de estimação descentralizada são utilizados os índices apresentados na seção 3.4, que são:

a) Custo convencional:

$$J_{D1}^i = \sum_{j=1}^{N^i} (\text{erro de previsão})^2 \quad (4.1)$$

b) Valor médio do número de informações de outros subsistemas recebidas pelo subsistema S_i por intervalo de amostragem indicado por

$$J_{D2}^i \quad i = 1, \dots, ns \quad (4.2)$$

c) Número total de parâmetros dos modelos (modelo local do subsistema e modelo de interconexão) utilizados no subsistema S_i de nominado

$$J_{D3}^i \quad i = 1, \dots, ns \quad (4.3)$$

d) Número total de interconexões físicas no subsistema S_i definido como

$$J_{D4}^i \quad i = 1, \dots, ns \quad (4.4)$$

Os algoritmos são também analisados quanto à convergência para os valores nominais dos parâmetros do sistema através de resultados em tabelas e gráficos e do cálculo da distância paramétrica normalizada do processo definida por [1]:

$$\delta_{K,np}^2 = \sum_{i=1}^{np} \left[\frac{\Delta \theta_K^i}{\theta_K^i} \right]^2 \quad (4.5)$$

onde $\Delta \theta_K^i = \theta_K^i - \hat{\theta}_K^i$ é o erro no parâmetro estimado $\hat{\theta}^i$ no instante K e np é o número de parâmetros estimados do processo (vetor $\hat{\theta}^{A,B}$ formado pelas matrizes polinomiais $A(q^{-1})$ e $B(q^{-1})$ do processo).

4.3. GERAÇÃO DAS MEDIDAS E INICIALIZAÇÃO

4.3.1. Geração das Medidas

Para se obter a saída y_K é utilizado um modelo multivariável em que a entrada u_K e a perturbação w_K são geradas de forma independente. O ruído w_K é uma sequência branca de variáveis aleatórias gaussianas com média zero e matriz de covariância igual a unidade. Como entrada u_K utiliza-se uma sequência de variáveis aleatórias gaussianas e matriz de covariância diagonal.

A variância de cada componente da entrada é escolhida de modo que a relação sinal/ruído em cada uma das componentes $y_{i,K}$ da saída tenha o valor 5,0 (cinco). Esta relação sinal/ruído é definida como a raiz quadrada da relação entre a variância da entrada observada $u_{i,K}^*$ e a variância da correspondente componente do ruído (neste caso igual a 1,0), onde:

$$u_{i,K}^* = b_i(q^{-1}) \cdot u_K \quad (4.6)$$

onde $b_i(q^{-1})$ é a i -ésima linha da matriz polinomial $B(q^{-1})$ definida no Capítulo 3.

4.3.2. Inicialização

Os algoritmos de estimação recursiva centralizada de parâmetros devem ser inicializados com valores dos parâmetros $\hat{\theta}(0)$ e da matriz $P(0)$. A matriz $P(0)$ reflete o grau de conhecimento "a priori" dos valores iniciais $\hat{\theta}(0)$.

Da mesma forma, os algoritmos de estimação recursiva descentralizada propostos exigem também uma inicialização. E quando, a cada passo, podem ocorrer falhas na comunicação, é necessário um conhecimento inicial melhor do que no caso dos algoritmos centralizados.

Para o algoritmo ERRA1/1 as inicializações requeridas variam de acordo com os modelos locais para as interconexões:

- a) se as interconexões são supostas iguais aos últimos valores disponíveis, o procedimento de inicialização consiste em:
- i) utilizar as medidas de entrada/saída do sistema para estimação centralizada através do algoritmo recursivo de mínimos quadrados inicializado com valores nulos para os componentes de $\hat{\theta}(0)$ e a matriz $P(0) = 10I$.
 - ii) inicializar os algoritmos locais a partir dos últimos valores obtidos com o algoritmo em (i)

O número de estimações centralizadas recursivas realizadas é fornecido junto a cada exemplo simulado.

- b) se as interconexões têm modelos ARMA, a inicialização é feita em três etapas:

- i) utilização das medidas das interconexões para estimação centralizada através do algoritmo de mínimos quadrados visando a seleção de um modelo ARMA do tipo descrito no Capítulo 3 e a inicialização do algoritmo local de mínimos quadrados recursivo que atualiza este modelo quando novos valores das interconexões estão disponíveis.
- ii) utilização das medidas de entrada/saída do sistema para estimação centralizada através do algoritmo recursivo de mínimos quadrados inicializado com valores nulos para os componentes de $\hat{\theta}(0)$ e a matriz $P(0) = 10I$.

O número de estimações centralizadas recursivas é fornecido junto a cada exemplo simulado. Para permitir comparações, utilizou-se em cada caso um número igual ao das estimações centralizadas realizadas para a situação (a).

- iii) inicialização do algoritmo local para estimação dos parâmetros do processo a partir dos resultados da etapa (ii) e inicialização do algoritmo local para estimação dos parâmetros das interconexões a partir dos resultados da etapa (i).

Para o algoritmo ERRB2/1 a inicialização pode ser realizada na própria rede com barramento e consiste de um número inicial de estimações recursivas realizadas no processador de maior porte. Em cada um dos exemplos simulados, os valores iniciais para estas séries de estimações centralizadas são $\hat{\theta}(0) = \underline{0}$ e $P(0) = 10I$.

O número de estimações recursivas realizadas na inicialização é fornecido para cada um dos exemplos simulados.

4.4. DESCRIÇÃO POR EQUAÇÕES À DIFERENÇAS DOS SISTEMAS SIMULADOS

Os sistemas simulados para analisar o desempenho dos algoritmos de estimação descentralizada ERRA1/1 e ERRB2/1 tem a forma representada na equação (4.7):

$$\begin{bmatrix} y^1 \\ y^2 \end{bmatrix}_K = \begin{bmatrix} 0,4 & -\lambda \\ \lambda & -0,9 \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \end{bmatrix}_{K-1} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u^1 \\ u^2 \end{bmatrix}_{K-1} + \begin{bmatrix} -0,3 & 0 \\ 0 & 0,4 \end{bmatrix} \begin{bmatrix} u^1 \\ u^2 \end{bmatrix}_{K-2} + \begin{bmatrix} e^1 \\ e^2 \end{bmatrix}_K \quad (4.7)$$

onde λ assume diversos valores nas simulações realizadas.

As componentes da entrada, da saída e da perturbação dos sistemas representados na equação (4.7) estão divididos em dois subsistemas conforme indicado pelos supra-índices e pela linha tracejada.

Nesta equação pode-se observar que a interação entre os subsistemas é variável e aumenta com o valor de λ . No subsistema S_1 , a influência do subsistema S_2 é representada pelo termo λy_{K-1}^2 . Da mesma forma, a contribuição do subsistema S_1 na saída do subsistema S_2 é dada pelo termo λy_{K-1}^1 .

A equação (4.7) representa um sistema estável para diversos valores de λ , permitindo avaliar o desempenho dos algoritmos à medida que a interação entre os subsistemas varia. A estabi

lidade para alguns valores de λ pode ser confirmada na tabela 4.1. A primeira coluna apresenta o valor absoluto dos zeros de $\det[A(z^{-1})]$ para as matrizes $A(z^{-1})$ dos sistemas da equação (4.7).

$$A(z^{-1}) = \begin{bmatrix} 1 + 0,4z^{-1} & -\lambda z^{-1} \\ \lambda z^{-1} & 1 - 0,9z^{-1} \end{bmatrix} \quad (4.8)$$

Na segunda coluna estão os zeros de $\det[B(z^{-1})]$ para a matriz $B(z^{-1})$ dos sistemas na equação (4.7).

$$B(z^{-1}) = \begin{bmatrix} 1 - 0,3z^{-1} & 0 \\ 0 & 1 + 0,4z^{-1} \end{bmatrix} \quad (4.9)$$

Exceto para $\lambda = 1,17$, os sistemas correspondentes aos valores de λ na tabela são estáveis, pois os zeros de $\det[A(z^{-1})]$ e $\det[B(z^{-1})]$ estão fora do círculo unitário e $A(z^{-1})$ é não-singular.

Os sistemas simulados têm λ igual a 0,3 , 0,5 e 0,8.

4.5. ALGORÍTMO ERRA1/1

Os resultados das simulações do algoritmo ERRA1/1 para os sistemas descritos na seção 4.4 são apresentados através das tabelas e gráficos do Apêndice C.

Nestas tabelas e gráficos são considerados dois modelos distintos para a interconexão de cada subsistema. As interconexões estão representadas na figura 4.1 e os modelos utilizados são:

(a) A interconexão é suposta constante até que novas medidas sejam recebidas. Ou seja:

$$\text{no subsistema } S_1 : h_K^{1,2} = \begin{cases} h_{K-1}^{1,2}, & \text{se não foi recebida} \\ \text{uma nova medida de } h^{1,2}. \\ \text{a medida recebida.} \end{cases} \quad (4.10)$$

λ	$ z^{-1} $ para	
	$\det[A(z^{-1})] = 0$	$\det[B(z^{-1})] = 0$
0,0	2,50 1,11	3,33 e 2,50
0,1	2,55 1,12	
0,3	3,06 1,21	
0,5	6,05 1,50	
0,6	∞	
0,7	2,77 (raízes complexas)	
0,8	1,89 (raízes complexas)	
0,9	1,49 (raízes complexas)	
1,0	1,25 (raízes complexas)	
1,17	1,00 (raízes complexas)	

Tabela 4.1 - Valor Absoluto dos Zeros de $\det[A(z^{-1})]$ e $\det[B(z^{-1})]$ para os Sistemas Representados na equação (4.7)

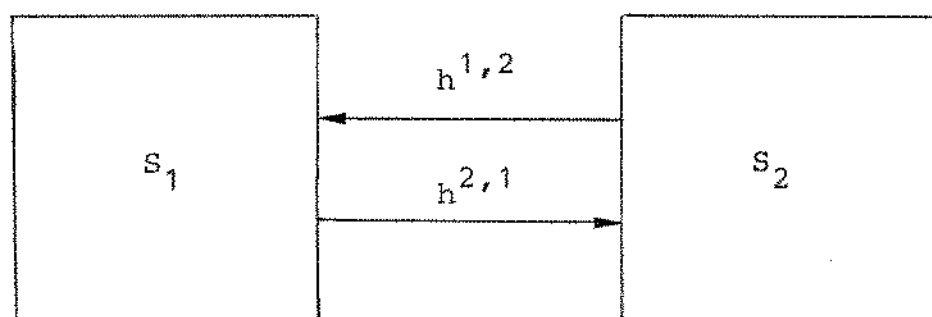


Figura 4.1 - Interconexões entre os Subsistemas no Algoritmo ERRAl/1

$$- \text{ no subsistema } S_2 : h_K^{2,1} = \begin{cases} h_{K-1}^{2,1}, & \text{se não foi recebida} \\ \text{uma nova medida de } h^{2,1} \\ \text{a medida recebida} \end{cases} \quad (4.11)$$

(b) A interconexão tem um modelo local ARMA. Este modelo é descrito por:

- no subsistema S_1 :

$$h_{K-1}^{1,2} = \mathcal{H}_+^1 h_{K-2}^{1,2} + \dots + \mathcal{H}_+^{N_h} h_{K-N_h-1}^{1,2} + \mathcal{G}_+^1 y_{K-2}^1 + \dots + \mathcal{G}_+^N y_{K-N-1}^1 + \omega_K^+ \quad (4.12)$$

- no subsistema S_2 :

$$h_{K-1}^{2,1} = \mathcal{H}_-^1 h_{K-2}^{2,1} + \dots + \mathcal{H}_-^{N_h} h_{K-N_h-1}^{2,1} + \mathcal{G}_-^1 y_{K-2}^2 + \dots + \mathcal{G}_-^N y_{K-N-1}^2 + \omega_K^- \quad (4.13)$$

Como existem apenas dois subsistemas, a equação (3.45), que fornece a relação entre o erro de descentralização e as interconexões pode ser escrita como:

$$\underline{Y}_K^1 = \underline{h}_{K-1}^{1,2} \quad (\text{para o subsistema } S_1) \quad (4.14)$$

$$\text{e } \underline{Y}_K^2 = \underline{h}_{K-1}^{2,1} \quad (\text{para o subsistema } S_2) \quad (4.15)$$

As tabelas C.1, C.2 e C.3 e os gráficos das figuras C.1 e C.2 do Apêndice C referem-se aos resultados das simulações, para os diversos valores de λ , no subsistema S_1 .

Nas tabelas C.1, foi utilizado o modelo (4.10) para a interconexão. Os custos e o comportamento dos parâmetros do sub

sistema são apresentados para vários valores do nível de descontinuidade de comunicação $\mathcal{R}_D$. O valor deste nível fornece o número de previsões locais sucessivas para a interconexão. $\mathcal{R}_D$ é constante para as simulações correspondentes às linhas da tabelas C.1. Os parâmetros do subsistema tem dois tipos de comportamento:

- i) "valor nominal" indicando que os parâmetros convergem para os valores nominais: $A_1^1 = 0,4$, $B_1^1 = 1$ e $B_2^1 = -0,3$.
- ii) "convergente" quando os parâmetros convergem para valores diferentes dos nominais.

Os parâmetros convergem para seus valores nominais para valores maiores do nível de degradação $\mathcal{R}_D$ quando a interação entre os subsistemas é mais fraca, isto é, para valores menores de λ .

As tabelas C.2 apresentam os resultados das estimações realizadas na escolha de um modelo do tipo (4.12) para a interconexão do subsistema S_1 . Cada tabela corresponde a um dos valores de λ e suas colunas mostram as ordens do modelo, os parâmetros estimados e o custo. Os modelos escolhidos para simulação do algorítmo ERRA1/1 estão assinalados por setas. A escolha dos modelos com ordens $N_{h+} = 2$ e $N_+ = 2$ para $\lambda = 0,3$ e $\lambda = 0,5$ é justificada pelo fato deste modelo armazenar mais informações relativas às interconexões e saídas passadas do subsistema podendo, talvez, fornecer previsões melhores para a interconexão.

Os resultados da simulação de ERRA1/1, utilizando para a interconexão os modelos assinalados nas tabelas C.2 são mostrados nas nas tabelas C.3. Em cada uma destas tabelas estão indicados o valor de λ , a ordem do modelo ARMA para a interconexão, a matriz P utilizada na inicialização do algoritmo de mínimos quadrados que atualiza o modelo da interconexão. Os valores iniciais dos parâmetros do modelo da interconexão são os obtidos na identificação representada nas tabelas C.2.

Os melhores resultados, isto é, menores valores para os índices de desempenho J_{D1} , J_{D2} , J_{D3} e J_{D4} . são obtidos utilizando para a interconexão o modelo de ordens $N_{h+} = 1$ e $N_+ = 1$. Estes rere

sultados estão representados nas tabelas C.3(a), (d) e (f).

A comparação entre os desempenhos do algoritmo $ERRA1/1$ supondo interconexão igual ao último valor disponível e utilizando o modelo ARMA com ordens $N_{h+} = 1$ e $N_{+} = 1$ é feita nos gráficos da figura C.1. Estes gráficos representam o custo J_{D1} como função do nível de degradação $\mathcal{R}_D$, para cada um dos valores de λ .

A informação sobre as saídas passadas do subsistema, incluídas no modelo ARMA, é responsável por este desempenho mais eficiente.

A utilização do modelo ARMA para a interconexão resulta em valores menores para o custo J_{D1} , além de reduzir o crescimento deste custo para os maiores valores de $\mathcal{R}_D$ simulados.

Nos gráficos da figura C.2 o custo J_{D1} é função de λ para diversos valores de $\mathcal{R}_D$. Na figura C.2(a), a interconexão é suposta constante até a recepção de nova medida enquanto na figura C.2(b) é utilizado o modelo ARMA com $N_{h+} = 1$ e $N_{+} = 1$. Quando a interação entre os subsistemas é mais forte, isto é, para valores maiores de λ , são observados desvios maiores em relação ao algoritmo (centralizado) de mínimos quadrados. Este algoritmo corresponde a $\mathcal{R}_D = 0$, situação na qual as medidas da interconexão são obtidas em todos os intervalos de amostragem. O gráfico para $\mathcal{R}_D = 0$ está na figura C.2(a).

Nas tabelas C.3(b), (c) e (e) o modelo para a interconexão tem a forma:

$$h_{K-1}^{1,2} = -\mathcal{H}_{+}^1 h_{K-2}^{1,2} - \mathcal{H}_{+}^2 h_{K-3}^{1,2} + \mathcal{G}_{+}^1 y_{K-2}^1 + \mathcal{G}_{+}^2 y_{K-3}^1 + \omega_K^+ \quad (4.16)$$

Os resultados obtidos mostram que estes modelos, apesar de utilizarem mais informações, conduzem a previsões piores que os modelos de ordens $N_{h+} = 1$ e $N_{+} = 1$ levando à redução do desempenho do algoritmo $ERRA1/1$. A explicação está na acumulação de erros pelo emprego de previsões para o cálculo de novas previsões. Este problema torna-se crítico com o aumento na ordem da parte autoregressiva do modelo e também com o aumento do nível de degradação $\mathcal{R}_D$.

Na tabela C.3(e) a acumulação dos erros de previsão da interconexão foi reduzida atualizando o valor da última previsão, quando uma nova medida se encontra disponível. Portanto, o procedimento adotado após a aquisição de uma nova medida consiste em dois passos: i) estimação dos parâmetros do modelo descrito pela equação (4.16) e ii) atualização da previsão feita no instante anterior utilizando os novos valores estimados dos parâmetros $\hat{H}_+^1$, $\hat{H}_+^2$, $\hat{G}_+^1$ e $\hat{G}_+^2$. Entretanto, mesmo com este procedimento não se obteve melhores resultados do que o algoritmo que utilizou o modelo de ordens $N_{h+} = 1$ e $N_{g+} = 1$.

Os resultados obtidos no subsistema S_2 para os diversos valores de λ são mostrados nas tabelas C.4, C.5 e C.6 e nos gráficos das figuras C.3 e C.4 do Apêndice C.

Estas tabelas e estes gráficos mostram relações entre as mesmas variáveis que as tabelas e os gráficos descritos para o subsistema S_1 . Além disto, as comparações do desempenho do algoritmo nos diversos casos conduzem aos mesmos resultados e às mesmas explicações.

Os valores nominais dos parâmetros do subsistema S_2 são $A_1^2 = -0,9$, $B_1^2 = 1$ e $B_2^2 = 0,4$.

Em situações reais, o nível de degradação $\mathcal{R}_D$ é variável durante a identificação descentralizada. Devido a isto, os modelos elaborados para as interconexões no algoritmo ERRA1/1 não pressupõem um valor fixo para $\mathcal{R}_D$.

É possível fazer uma análise dos modelos para as interconexões dos subsistemas S_1 e S_2 , obtidos nas tabelas C.2 e C.5, respectivamente. As duas linhas da equação (4.7) são:

$$y_K^1 = -0,4 y_{K-1}^1 + \lambda y_{K-1}^2 + u_{K-1}^1 - 0,3 u_{K-1}^1 + e_K^1 \quad (4.17)$$

$$e \quad y_K^2 = -\lambda y_{K-1}^1 + 0,9 y_{K-1}^2 + u_{K-1}^2 + 0,4 u_{K-2}^2 + e_K^2 \quad (4.18)$$

Examinando estas equações, a interconexão de cada um dos subsistemas pode ser escrita como:

$$y_K^1 = h_{K-1}^{1,2} = \lambda y_{K-1}^2 \quad (4.19)$$

$$e \quad y_K^2 = h_{K-1}^{2,1} = -\lambda y_{K-1}^1 \quad (4.20)$$

Substituindo as saídas y_{K-1}^1 e y_{K-1}^2 pelas entradas e saídas passadas de cada subsistema, obtém-se:

$$y_K^1 = h_{K-1}^{1,2} = \lambda [-\lambda y_{K-2}^1 + 0,9 y_{K-2}^2 + u_{K-2}^2 + 0,4 u_{K-3}^2 + e_{K-1}^2] \quad (4.21)$$

$$y_K^1 = h_{K-1}^{1,2} = +0,9 h_{K-2}^{1,2} - (\lambda)^2 y_{K-2}^1 + \lambda [u_{K-2}^2 + 0,4 u_{K-3}^2 + e_{K-1}^2] \quad (4.22)$$

$$e \quad y_K^2 = h_{K-1}^{2,1} = -\lambda [-0,4 y_{K-2}^1 + \lambda y_{K-2}^2 + u_{K-2}^1 - 0,3 u_{K-3}^1 + e_{K-1}^1] \quad (4.23)$$

$$y_K^2 = h_{K-1}^{2,1} = -0,4 h_{K-2}^{2,1} - (\lambda)^2 y_{K-2}^2 - \lambda [u_{K-2}^1 - 0,3 u_{K-3}^1 - e_{K-1}^1] \quad (4.24)$$

As equações (4.22) e (4.24) apresentam a mesma forma do modelo ARMA com ordens $N_h = 1$ e $N = 1$ utilizado para a interconexão de cada subsistema. Nestas equações, tem-se:

$$H_+^1 = -0,9 \quad G_+^1 = -(\lambda)^2 \quad (4.25)$$

$$e \quad H_-^1 = 0,4 \quad G_-^1 = -(\lambda)^2 \quad (4.26)$$

que são valores próximos dos obtidos nas tabelas C.2 e C.5 para os modelos de ordens $N_h = 1$ e $N = 1$. As diferenças podem ser explicadas pela presença nas equações (4.22) e (4.24) de um termo dependente das entradas passadas do outro subsistema. Como este termo é proporcional a λ , os resultados das tabelas C.2 e C.5 estão mais próximos das equações (4.25) e (4.26) para valores menores de λ .

A dedução das equações (4.22) e (4.24) a partir das equações (4.17) e (4.18) mostra como valores iniciais para os parâmetros dos modelos das interconexões de cada subsistema no exemplo simulado, podem ser obtidos dos parâmetros da representação

entrada/saída do sistema. Além disso, nessas equações $N_h = 1$ e $N=1$ podem ser a inicialização de um método de procura da ordem do modelo ARMA da interconexão, que fornece o melhor desempenho do algoritmo de identificação.

4.6. ALGORÍTMO ERRB2/1

Os resultados das simulações do algoritmo ERRB2/1 para os sistemas descritos na seção 4.4 são apresentados através das tabelas e gráficos do Apêndice D.

As simulações são realizadas para uma versão simplificada do algoritmo ERRB2/1 descrito na seção 3.7.

Conforme mostra a figura 4.2, o processador de maior porte não está localizado num dos subsistemas, mas está conectado ao mesmo barramento que os processadores dos subsistemas.

Se os recursos do processador de maior porte não forem integralmente utilizados pela rede com barramento, representada na figura 4.2, ele pode ser conectado a outra rede e/ou empregado em outras tarefas.

As medidas de entrada/saída e os erros locais de previsão são enviados ao processador central a cada intervalo de amostragem. Caso o quadrado do erro local de previsão de um dos subsistemas ultrapasse o valor EQ_{MIN} , o processador central envia aos dois subsistemas novos valores para os erros de descentralização. Estes valores são calculados pela equação (3.60).

Assim, enquanto para o algoritmo da seção 3.7 considera-se a seleção de quais subsistemas devam receber valores atualizados para seus erros de descentralização e em qual intervalo, nestes exemplos realiza-se apenas a seleção dinâmica dos intervalos para atualização.

Não foram consideradas durante as simulações falhas no envio de informações ao processador de maior porte nem a utiliza

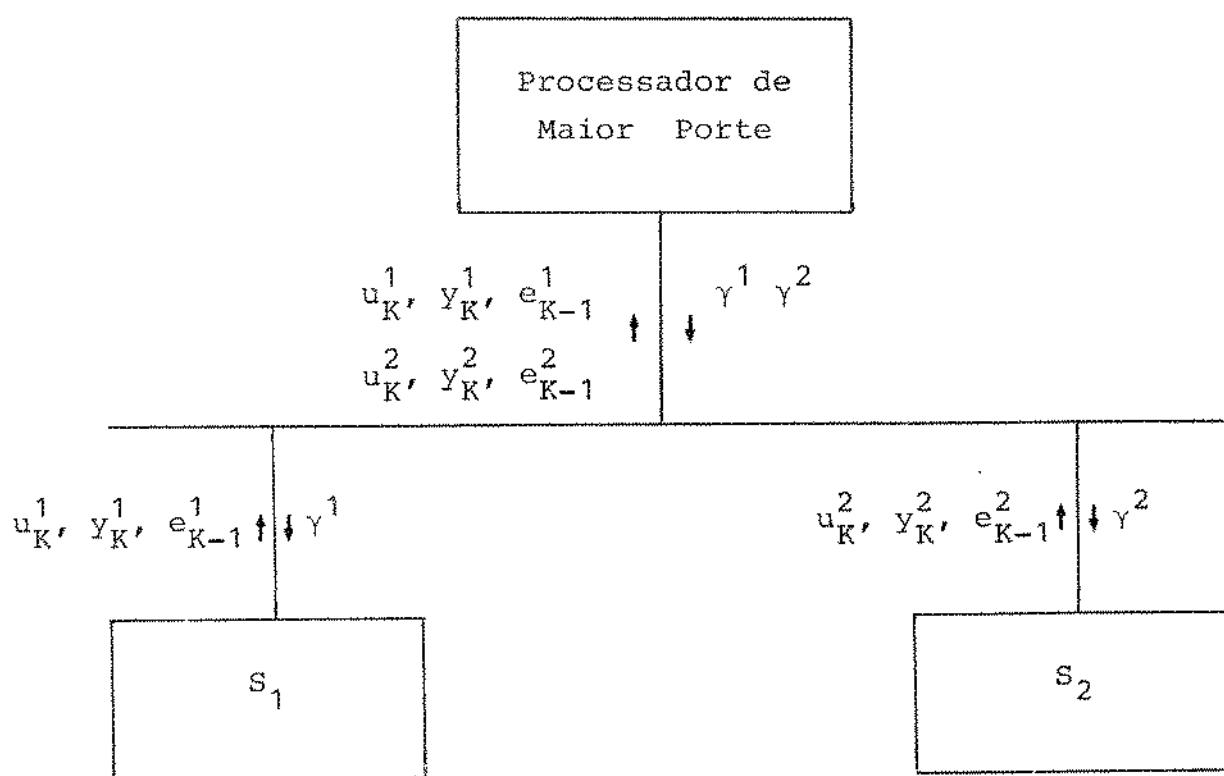


Figura 4.2 - Troca de Informações entre os Subsistemas e o Processador de Maior Porte no Algoritmo ERRB2/1

ção de modelos ARMA para o erro de descentralização.

Em cada subsistema, o erro de descentralização é suposto constante até que novos valores, calculados no processador de maior porte, sejam recebidos. Ou seja, os modelos utilizados para o erro de descentralização tem a forma:

$$\text{ - no subsistema } S_1 : \gamma_K^1 = \begin{cases} \gamma_{K-1}^1, & \text{se não foi recebido um} \\ \text{novo valor.} & \\ \text{o valor recebido.} & \end{cases} \quad (4.27)$$

$$\text{ - no subsistema } S_2 : \gamma_K^2 = \begin{cases} \gamma_{K-1}^2, & \text{se não foi recebido um} \\ \text{novo valor.} & \\ \text{o valor recebido.} & \end{cases} \quad (4.28)$$

Nas tabelas e gráficos do Apêndice D, o índice $K.J_{D2}$ representa o produto do número de intervalos de amostragem K pelo índice J_{D2} . Ou seja, $K.J_{D2}$ é igual ao número de informações externas recebidas em cada subsistema até o instante considerado. Como informação externa, cada subsistema recebe apenas o valor do seu erro de descentralização calculado no processador de maior porte e consequentemente $K.J_{D2}$ é igual ao número de vezes que o erro local de previsão de um dos subsistemas ultrapassou o valor EQ_{MIN} .

Na tabela D.1 são apresentados os valores EQ_{MIN} utilizados para cada valor de λ . Nesta tabela estão também os índices J_{D1} e $K.J_{D2}$ obtidos em 500 estimações. Pode-se observar que com o aumento de EQ_{MIN} o crescimento do índice J_{D1} é mais rápido para subsistemas com interação mais forte, isto é, para valores maiores de λ . Para $EQ_{MIN} = 0$, o processador de maior porte envia valores atualizados dos erros de descentralização aos subsistemas a cada intervalo de amostragem. Os resultados obtidos nesta situação equivalem aos do algoritmo de mínimos quadrados.

Os gráficos das figuras D.1, D.2 e D.3 representam a evolução do índice $K.J_{D2}$ com o número de intervalos de amostragem K para cada um dos valores de λ . A reta $K.J_{D2} = K$ corresponde a $EQ_{MIN} = 0$

e abaixo dela estão as curvas para os demais valores de EQ_{MIN} . Os pontos comuns $K.J_{D2} = K$, $0 \leq K \leq 20$ correspondem a inicialização do algoritmo ERB2/1.

As tabelas D.2, D.3 e D.4 mostram as distâncias paramétricas em $K = 100, 200, 300, 400$ e 500 no subsistema S_1 , para diferentes valores de λ e EQ_{MIN} . A distância paramétrica aumenta com os valores de EQ_{MIN} e λ .

Nas tabelas D.5, D.6 e D.7 são apresentadas as distâncias paramétricas para o subsistema S_2 .

Em todas as simulações realizadas, o índice J_{D3} é igual a 3 e o índice J_{D4} é igual a 1 (devido a ligação ao barramento).

As simulações apresentadas envolvem o número mínimo de dois subsistemas, mas servem como ilustração das etapas necessárias na implementação e do desempenho dos algoritmos ERRA1/1 e ERB2/1.

CAPÍTULO 5

CONCLUSÕES

As contribuições deste trabalho consistem no desenvolvimento de um tratamento matemático do problema de identificação descentralizada e na apresentação de algoritmos adaptados às redes de computação distribuída atualmente utilizadas.

A definição dos índices de distribuição e a extensão do conceito de causalidade quantificam as idéias de distribuição das saídas entre os subsistemas de um sistema descentralizado e de influências exercidas entre estes subsistemas.

A abordagem do problema de comunicação entre os subsistemas com a definição de vários níveis de degradação introduz um vínculo entre os sistemas centralizados e descentralizados. Como resultado, os algoritmos de identificação e controle descentralizados podem ser vistos como versões de algoritmos centralizados, modificados para operar quando as trocas de informações entre as partes do sistema são, temporária ou permanentemente, deficientes. Quando as comunicações são restabelecidas, o desempenho dos algoritmos descentralizados deve igualar o desempenho dos estimadores centralizados.

O conceito de erro de descentralização, apresentado junto com o método de decomposição para modelamento de sistemas descentralizados, mostra os limites impostos ao modelo de cada subsistema, em decorrência das falhas de comunicação. Este erro é também uma forma conveniente de relacionar sistemas centralizados e descentralizados e é o seu modelamento adequado que simplifica a implementação de algoritmos aplicáveis a diversos níveis de degradação.

Os algoritmos para redes anel, estrela e com barramento empregam métodos de identificação centralizada, adaptando-os ao ambiente descentralizado através do modelamento do erro de descentralização. Entre os métodos centralizados, foram desenvolvidas as versões MIMO dos estimadores raiz quadrada, janela exponencial e janela retangular.

Para avaliação dos algoritmos descentralizados, foram definidos índices de modo a comparar as taxas de troca de informações, os meios físicos utilizados e o número de parâmetros envolvidos.

Versões industriais de redes estrela, anel e com barramento, estão surgindo depois de um período de pesquisas em laboratórios de indústrias e universidades. Alguns modelos empregam apenas uma configuração e outros são híbridos, procurando uma melhor adaptação às variações de distância e taxa de transmissão nos diversos pontos da rede. Os meios físicos de transmissão incluem pares trançados, cabos coaxiais e fibras ópticas. A duplicação de processadores e de trechos do meio de transmissão, assim como a monitoração do funcionamento dos processadores, são algumas das estratégias para aumentar a confiabilidade destas redes.

A utilização de algoritmos descentralizados de identificação e controle permitirá um emprego eficiente destas redes.

Em sequência, sugere-se algumas linhas de pesquisa para o desenvolvimento de outros trabalhos:

i) Simulações dos demais algoritmos deste trabalho. Para avaliação do desempenho destes algoritmos, pode-se começar com exemplos simples como o descrito no Capítulo 4 e a seguir utilizar alguns exemplos com estrutura mais complexa, como sistemas de maior ordem e/ou constituídos por um número maior de subsistemas.

ii) Desenvolvimento de uma metodologia de composição para elaboração de modelos de sistemas descentralizados. Este método consistiria de duas etapas: a primeira para o desenvolvimento de modelos entrada/saída para cada um dos subsistemas; a segunda para determinar as restrições impostas a estes modelos, devido a vizinhança entre os subsistemas.

iii) Elaboração de algoritmos descentralizados, baseados no modelamento por decomposição, utilizando modelos não-lineares para o erro de descentralização. As equações (3.21)-(3.26), que representam modelos do erro de descentralização em função da comunicação entre os subsistemas, não se restringem a modelos lineares.

iv) Extensão do tratamento matemático apresentado no Capítulo 3 para incluir os conceitos precisos de convergência e estabilidade de algoritmos descentralizados.

APÊNDICE A

IDENTIDADES E DEFINIÇÕES MATEMÁTICAS

Identidade 1: Lema de Inversão Matricial

Sejam A, B e C matrizes quadradas de ordem n, sendo A e C inversíveis, então:

$$[A + B^T C B]^{-1} = A^{-1} - A^{-1} B^T [C^{-1} + B A^{-1} B^T]^{-1} B A^{-1} \quad (A.1)$$

Esta identidade é útil na dedução de algoritmos recursivos.

Definição 1: Produto de Kronecker

Seja $A = [a_{ij}]$ uma matriz $m \times n$ e $B = [b_{ij}]$ uma matriz de dimensão $s \times t$ o produto de Kronecker $A \otimes B$ é definido como a matriz $ms \times nt$ dada por

$$A \otimes B \triangleq [a_{ij} \cdot B] \quad (A.2)$$

Definição 2: Vec de uma Matriz

Seja $A = [a_{ij}]$ uma matriz $m \times n$, denomina-se Vec. A o vetor de dimensão $m \cdot n$ formado pelas colunas de A, isto é:

$$\text{Vec A} \triangleq \begin{bmatrix} a_{.1} \\ \vdots \\ a_{.n} \end{bmatrix} \quad (A.3)$$

onde $a_{.j}$ indica a j-ésima coluna de A.

Estas duas definições são importantes na manipulação das expressões de estimadores para sistemas multivariáveis. Algumas propriedades destas operações são apresentadas a seguir.

Identidade 2: Para uma matriz $A = \text{bloco diagonal } [\underline{a}, \dots, \underline{a}]$ de dimensão $\ell m \times m$, onde $\underline{a}$ é um vetor de dimensão ℓ e uma matriz simétrica B de ordem m , tem-se:

$$A \cdot B = (B \otimes I_\ell) \cdot A \quad (\text{A.4})$$

onde I_ℓ é a matriz identidade de ordem ℓ .

Identidade 3: Para uma matriz $A = \text{bloco diagonal } [\underline{a}_1, \dots, \underline{a}_m]$ de dimensão $\ell m \times m$, onde $\underline{a}_i$ ($i = 1, \dots, m$) é um vetor de dimensão ℓ e uma matriz B diagonal de ordem m , tem-se:

$$A \cdot B = (B \otimes I_\ell) \cdot A \quad (\text{A.5})$$

onde I_ℓ é a matriz identidade de ordem ℓ .

Identidade 4: Para uma matriz $A = \text{bloco diagonal } [\underline{a}_1, \dots, \underline{a}_m]$ de dimensão $\ell m \times m$, onde $\underline{a}_i$ ($i = 1, \dots, m$) é um vetor de dimensão ℓ e um vetor $\underline{b}$ de dimensão m tal que $\underline{b}^T = [b^1, \dots, b^m]$, tem-se:

$$[A \cdot A^T]^{-1} \cdot [A \cdot \underline{b}] = \underline{c} \quad (\text{A.6})$$

onde o vetor $\underline{c}$ é dado por:

$$\underline{c}^T = [\underline{c}_1^T, \dots, \underline{c}_p^T]$$

com
$$\underline{c}_i = (\underline{a}_i \underline{a}_i^T)^{-1} \underline{a}_i \cdot b^i$$

Identidade 5: Sejam A uma matriz $m \times m$ e B uma matriz $n \times n$, então:

$$A \otimes B = (A \otimes I_n) \cdot (I_m \otimes B) = (I_m \otimes B) \cdot (A \otimes I_n) \quad (\text{A.7})$$

Identidade 6: Lema de Inversa Particionada

Sejam A , B , C e D matrizes quadradas inversíveis e seja

$$F = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]$$

então a inversa F^{-1} é dada por

$$F^{-1} = \left[\begin{array}{c|c} A^{-1} (I + B E^{-1} C A)^{-1} & -A^{-1} B E^{-1} \\ \hline -E^{-1} C A^{-1} & E^{-1} \end{array} \right] \quad (A.8)$$

onde E é a matriz dada por

$$E = D - C A^{-1} B$$

APÊNDICE B

ALGORÍTMO DE JANELA RETANGULAR

Este algoritmo recursivo é utilizado na estimação de parâmetros variáveis no tempo. Num certo instante, a estimação é realizada somente com um número fixo de observações. As demais são descartadas.

Para uma janela retangular de comprimento fixo N , o procedimento seguido no instante $r+N$ consta de dois passos:

i) A observação $\underline{y}_{r+N}$, $\underline{u}_{r+N}$ é acrescentada na atualização de parâmetros usando o algoritmo de mínimos quadrados recursivo, ou seja:

$$\begin{aligned} \underline{P}_{r+N}^r &= \underline{P}_{r+N-1}^r - \underline{P}_{r+N-1}^r \cdot \underline{\psi}_{r+N}^T [I + \underline{\psi}_{r+N} \cdot \underline{P}_{r+N-1}^r \cdot \underline{\psi}_{r+N}^T]^{-1} \cdot \\ &\quad \cdot \underline{\psi}_{r+N-1} \underline{P}_{r+N-1}^r \end{aligned} \quad (B.1)$$

$$\underline{K}_{r+N}^r = \underline{P}_{r+N}^r \cdot \underline{\psi}_{r+N}^T \quad (B.2)$$

$$\underline{\hat{\theta}}_{r+N}^{r,A,B} = \underline{\hat{\theta}}_{r+N}^{r,A,B} + \underline{K}_{r+N}^r [\underline{y}_{r+N} - \underline{\psi}_{r+N} \underline{\hat{\theta}}_{r+N-1}^{r,A,B}] \quad (B.3)$$

onde o índice r sobre $\underline{\hat{\theta}}$, $\underline{P}$ e $\underline{K}$ indica que estes valores são calculados tomando por base as observações realizadas entre r e $r+N$.

ii) A observação $\underline{y}_r$, $\underline{u}_r$ realizada no instante r é descartada.

Pode-se escrever o estimador baseado nas observações entre $r+1$ e $r+N$, como:

$$\underline{\hat{\theta}}_{r+N}^{r+1,A,B} = \left[\begin{array}{c} r+N \\ \Sigma \\ K=r+1 \end{array} \underline{\psi}_K^T \cdot \underline{\psi}_K \right]^{-1} \left[\begin{array}{c} r+N \\ \Sigma \\ K=r+1 \end{array} \underline{\psi}_K^T \cdot \underline{y}_K \right] \quad (B.4)$$

onde

$${}^{r+1}P_{r+N} = \left[\begin{array}{c} r+N \\ \Sigma \\ K=r+1 \end{array} \psi_K^T \cdot \psi_K \right]^{-1} \quad (B.5)$$

Reescrevendo esta equação e utilizando o lema de inversão matricial (equação (A.1)), tem-se:

$${}^{r+1}P_{r+N} = [P_{r+N}^r - \psi_r^T \cdot \psi_r]^{-1} \quad (B.6)$$

$${}^{r+1}P_{r+N} = P_{r+N}^r + P_{r+N}^r \cdot \psi_r^T \cdot [I - \psi_r \cdot P_{r+N}^r \cdot \psi_r^T]^{-1} \cdot \psi_r \cdot P_{r+N}^r \quad (B.7)$$

Substituindo a equação (B.5) na equação (B.4), obtêm-se:

$$\hat{\theta}_{r+N}^{r+1} A, B = {}^{r+1}P_{r+N} \cdot \left[\begin{array}{c} r+N \\ \Sigma \\ K=r \end{array} \psi_K^T \cdot \underline{y}_K - \psi_r^T \underline{y}_r \right] \quad (B.8)$$

$$\hat{\theta}_{r+N}^{r+1} A, B = {}^{r+1}P_{r+N} \left[\begin{array}{c} r+N \\ \Sigma \\ K=r \end{array} \psi_K^T \cdot \underline{y}_K \right] - {}^{r+1}P_{r+N} \cdot \psi_r^T \cdot \underline{y}_r \quad (B.9)$$

E substituindo a eq. (B.7) no primeiro termo do lado direito desta última equação:

$$\begin{aligned} \hat{\theta}_{r+N}^{r+1} A, B &= P_{r+N}^r + P_{r+N}^r \psi_r^T [I - \psi_r P_{r+N}^r \psi_r^T]^{-1} \cdot \psi_r \cdot P_{r+N}^r \\ &\quad \left[\begin{array}{c} r+N \\ \Sigma \\ K=r \end{array} \psi_K^T \cdot \underline{y}_K \right] - {}^{r+1}P_{r+N} \cdot \psi_r^T \cdot \underline{y}_r \end{aligned} \quad (B.10)$$

ou

$$\begin{aligned} \hat{\theta}_{r+N}^{r+1} A, B &= \hat{\theta}_{r+N}^r A, B + P_{r+N}^r \psi_r^T [I - \psi_r P_{r+N}^r \psi_r^T]^{-1} \cdot \psi_r \cdot \hat{\theta}_{r+N}^r A, B - \\ &\quad {}^{r+1}P_{r+N} \cdot \psi_r^T \cdot \underline{y}_r \end{aligned} \quad (B.11)$$

Multiplicando a equação (B.7) por ψ_r^T , tem-se:

$$\begin{aligned}
 {}^{r+1}P_{r+N} \cdot \psi_r^T &= {}^rP_{r+N} \cdot \psi_r^T + {}^rP_{r+N} \cdot \psi_r^T [I - \psi_r \cdot {}^rP_{r+N} \cdot \psi_r^T]^{-1} \cdot \\
 &\quad \cdot \psi_r \cdot {}^rP_{r+N} \cdot \psi_r^T
 \end{aligned} \tag{B.12}$$

$${}^{r+1}P_{r+N} \cdot \psi_r^T = {}^rP_{r+N} \cdot \psi_r^T \{ I + [I - \psi_r {}^rP_{r+N} \psi_r^T]^{-1} \psi_r \cdot {}^rP_{r+N} \cdot \psi_r^T \} \tag{B.13}$$

$$\begin{aligned}
 {}^{r+1}P_{r+N} \cdot \psi_r^T &= {}^rP_{r+N} \cdot \psi_r^T [I - \psi_r {}^rP_{r+N} \psi_r^T]^{-1} \{ [I - \psi_r {}^rP_{r+N} \psi_r^T] + \\
 &\quad + \psi_r \cdot {}^rP_{r+N} \cdot \psi_r^T \}
 \end{aligned} \tag{B.14}$$

$${}^{r+1}P_{r+N} \cdot \psi_r^T = {}^rP_{r+N} \cdot \psi_r^T [I - \psi_r {}^rP_{r+N} \psi_r^T]^{-1} \tag{B.15}$$

Definindo:

$${}^{n+1}_K = {}^{r+1}P_{r+N} \cdot \psi_r^T = {}^rP_{r+N} \psi_r^T [I - \psi_r {}^rP_{r+N} \psi_r^T]^{-1} \tag{B.16}$$

a equação (B.11) pode ser escrita como:

$$\hat{\theta}_{r+N}^{r+1} A, B = \hat{\theta}_{r+N}^r A, B - {}^{r+1}_K \{ \underline{y}_r - \psi_r \cdot \hat{\theta}_{r+N}^r A, B \} \tag{B.17}$$

As equações (B.7), (B.16) e (B.17) constituem o segundo passo do algoritmo janela retangular.

A dedução da versão linha por linha do algoritmo janela retangular é análoga à versão linha por linha do estimador de mínimos quadrados, eq. (2.24).

Uma variante do algoritmo de janela retangular é o algoritmo de memória oscilante. A memória deste algoritmo varia entre N e 2N. Para acrescentar uma nova observação, são utilizadas as

equações (B.1), (B.2) e (B.3). Quando já foram acrescentadas N novas observações e o número de observações em uso atinge $2N$, as N observações mais antigas são descartadas.

Como:

$$\hat{\theta}_{r+2N}^{A,B} = P_{r+2N}^{r+N+1} \cdot \begin{bmatrix} r+2N \\ \Sigma \\ K=r+N+1 \end{bmatrix} \psi_K^T Y_K \quad (B.18)$$

$$P_{r+2N}^{r+N+1} = \begin{bmatrix} r+2N \\ \Sigma \\ K=r+N+1 \end{bmatrix}^{-1} \quad (B.19)$$

tem-se:

$$P_{r+2N}^{r+N+1} = \begin{bmatrix} r+2N \\ \Sigma \\ K=r+1 \end{bmatrix}^{-1} \psi_K^T \psi_K - \begin{bmatrix} r+N \\ \Sigma \\ K=r+1 \end{bmatrix}^{-1} \psi_K^T \psi_K \quad (B.20)$$

$$P_{r+2N}^{r+N+1} = P_{r+2N}^{r+1} - P_{r+N}^{r+1} \quad (B.21)$$

A equação (B.18) pode ser escrita na forma:

$$\hat{\theta}_{r+2N}^{A,B} = P_{r+2N}^{r+N+1} \begin{bmatrix} r+2N \\ \Sigma \\ K=r+1 \end{bmatrix} \psi_K^T Y_K - \begin{bmatrix} r+N \\ \Sigma \\ K=r+1 \end{bmatrix} \psi_K^T Y_K \quad (B.22)$$

$$\hat{\theta}_{r+2N}^{A,B} = P_{r+2N}^{r+N+1} \begin{bmatrix} r+1 \\ P_{r+2N}^{-1} & P_{r+2N}^{r+1} & \Sigma \\ K=r+1 \end{bmatrix} \psi_K^T Y_K - \begin{bmatrix} r+1 \\ P_{r+N}^{-1} & P_{r+N}^{r+1} & \Sigma \\ K=r+1 \end{bmatrix} \psi_K^T Y_K \quad (B.23)$$

$$\hat{\theta}_{r+2N}^{A,B} = P_{r+2N}^{r+N+1} \begin{bmatrix} r+1 & r+1 & r+1 & r+1 \\ P_{r+2N}^{-1} & \hat{\theta}_{r+2N}^{A,B} & P_{r+N}^{-1} & \hat{\theta}_{r+N}^{A,B} \end{bmatrix} \quad (B.24)$$

Através das eqs. (B.19) e (B.24) as N observações mais antigas são descartadas. Para realizar esta operação, os únicos dos antigos a serem armazenados são:

$$\begin{array}{c} r+N \\ \hat{\theta}^{A,B} \\ -r+1 \end{array}$$

e

$$\begin{array}{c} r+N \\ P_{r+1} \end{array} \cdot$$

APÊNDICE C

RESULTADOS DE SIMULAÇÃO PARA O ALGORÍTMO ERRA1/1

Este Apêndice contém tabelas e gráficos com os resultados da simulação do algoritmo ERRA1/1 para o sistema descrito na seção 4.4.

A seção 4.5 apresenta os comentários referentes às tabelas e aos gráficos deste Apêndice.

a) $\lambda = 0,3$

$\mathcal{R}_D$	J_{D1}	J_{D2}	J_{D3}	J_{D4}	Parâmetros
0	132	1	3	2	Valor Nominal
1	298	1/2	3	2	Valor Nominal
2	590	1/3	3	2	Valor Nominal
3	779	1/4	3	2	Convergente

b) $\lambda = 0,5$

$\mathcal{R}_D$	J_{D1}	J_{D2}	J_{D3}	J_{D4}	Parâmetros
0	137	1	3	2	Valor Nominal
1	743	1/2	3	2	Convergente
2	1622	1/3	3	2	Convergente

c) $\lambda = 0,8$

$\mathcal{R}_D$	J_{D1}	J_{D2}	J_{D3}	J_{D4}	Parâmetros
0	148	1	3	2	Valor Nominal
1	3158	1/2	3	2	Convergente
2	6714	1/3	3	2	Convergente

Tabela C.1 - Algoritmo ERRA1/1 para o Subsistema S_1

Interconexões supostas constantes até a recepção de novas medidas

Número de estimações centralizadas para inicialização : 10

Número de estimações descentralizadas : 500

a) $\lambda = 0,3$

Ordens		Parâmetros Estimados	Custos
Nh ₊	N ₊		
0	1	$G_+^1 = 0,119$	1659
0	2	$G_+^1 = 0,184$	1500
		$G_+^2 = 0,193$	
1	0	$H_+^1 = -0,894$	358
1	1	$H_+^1 = -0,977$	311
		$G_+^1 = -0,107$	
1	2	$H_+^1 = -0,999$	320
		$G_+^1 = -0,122$	
		$G_+^2 = -0,028$	
2	0	$H_+^1 = -1,137$	337
		$H_+^2 = 0,273$	
2	2	$H_+^1 = -1,267$	299
		$H_+^2 = 0,349$	
		$G_+^1 = -0,095$	
		$G_+^2 = 0,003$	

b) $\lambda = 0,5$

Ordens		Parâmetros Estimados	Custos
Nh ₊	N ₊		
0	1	$G_+^1 = 0,113$	3277
0	2	$G_+^1 = 0,135$	3244
		$G_+^2 = 0,148$	
1	0	$H_+^1 = -0,798$	1235
1	1	$H_+^1 = -1,038$	841
		$G_+^1 = -0,303$	
1	2	$H_+^1 = -1,058$	861
		$G_+^1 = 0,317$	
		$G_+^2 = -0,040$	
2	0	$H_+^1 = -0,993$	1178
		$H_+^2 = 0,245$	
2	2	$H_+^1 = -1,273$	816
		$H_+^2 = 0,366$	
		$G_+^1 = -0,257$	
		$G_+^2 = 0,091$	

c) $\lambda = 0,8$

Ordens		Parâmetros Estimados	Custos
Nh ₊	N ₊		
0	1	$G_+^1 = -0,105$	7206
0	2	$G_+^1 = -0,094$	7181
		$G_+^2 = -0,158$	
1	0	$H_+^1 = -0,564$	4964
1	1	$H_+^1 = -1,114$	2063
		$G_+^1 = -0,754$	
1	2	$H_+^1 = -1,115$	2109
		$G_+^1 = -0,755$	
		$G_+^2 = 0,005$	
2	0	$H_+^1 = -0,788$	4249
		$H_+^2 = 0,397$	
2	2	$H_+^1 = -1,277$	2049
		$H_+^2 = 0,367$	
		$G_+^1 = -0,653$	
		$G_+^2 = 0,228$	

Tabelas C.2 - Modelos para a Interconexão do Subsistema S₁ na forma:

$$h_{K-1}^{1,2} = (H_+^1 h_{K-2}^{1,2} + \dots + H_+^{N_h} h_{K-N_h-1}^{1,2}) + G_+^1 y_{K-2}^1 + \dots + G_+^{N_g} y_{K-N_g-1}^1 + \omega_K^+$$

Número de medidas utilizadas na elaboração do modelo : 500

Método de estimação : mínimos quadrados recursivo

a) $\lambda = 0,3$

Interconexão $N_{h+}=1, N_{+}=1$

$P(0) = 0,01I \quad J_{D3} = 5$

N_D	J_{D1}	J_{D2}	Parâmetros
1	263	1/2	Valor Nominal
2	543	1/3	Valor Nominal
3	690	1/4	Convergente
4	739	1/5	Convergente

b) $\lambda = 0,3$

Interconexão $N_{h+}=2, N_{+}=2$

$P(0) = 0,01I \quad J_{D3} = 7$

N_D	J_{D1}	J_{D2}	Parâmetros
1	-	-	-
2	701	1/3	Convergente
3	1090	1/4	Convergente
4	1067	1/5	Convergente

c) $\lambda = 0,3$

Interconexão $N_{h+}=2, N_{+}=2$

$P(0) = 10^{-4}I \quad J_{D3} = 7$

N_D	J_{D1}	J_{D2}	Parâmetros
1	257	1/2	Valor Nominal
2	583	1/3	Valor Nominal
3	800	1/4	Convergente
4	933	1/5	Convergente

d) $\lambda = 0,5$

Interconexão $N_{h+}=1, N_{+}=1$

$P(0) = 0,01I \quad J_{D3} = 5$

N_D	J_{D1}	J_{D2}	Parâmetros
1	481	1/2	Valor Nominal
2	1244	1/3	Convergente
3	1556	1/4	Convergente
4	1532	1/5	Convergente

e) $\lambda = 0,5$

Interconexão* $N_{h+}=2, N_{+}=2$

$P(0) = 0,01I \quad J_{D3} = 7$

N_D	J_{D1}	J_{D2}	Parâmetros
1	479	1/2	Valor Nominal
2	1609	1/3	Convergente
3	2742	1/4	Convergente
4	2165	1/5	Convergente

f) $\lambda = 0,8$

Interconexão $N_{h+}=1, N_{+}=1$

$P(0) = 0,01I \quad J_{D3} = 5$

N_D	J_{D1}	J_{D2}	Parâmetros
1	3792	1/2	Valor Nominal
2	5930	1/3	Convergente
3	6535	1/4	Convergente
-	-	-	-

* O valor da última previsão feita para interconexão é atualizado antes de sua utilização para o cálculo de uma nova previsão.

Tabelas C.3 - Algoritmo ERRA1/1 para o Subsistema S_1

As ordens dos modelos utilizados para a interconexão e o valor inicial da matriz P para o algoritmo de mínimos quadrados que atualiza os parâmetros destes modelos estão indicados acima de cada tabela.

Número de estimações centralizadas para inicialização : 10

Número de Estimções descentralizadas : 500

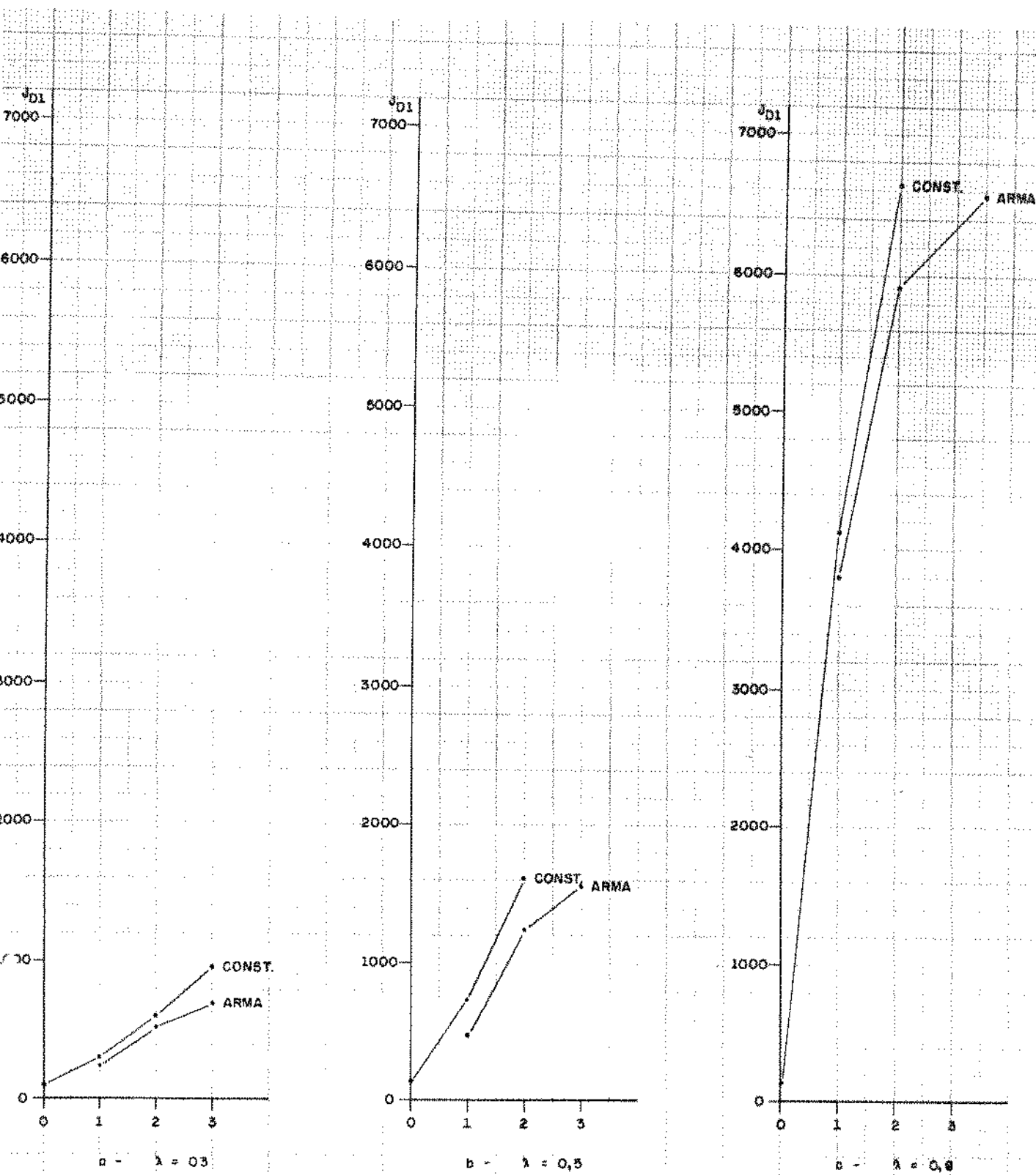


Figura C.1 - Custo $J_{D1} \times n_D$ para o Algoritmo ERRA1/1 no Subsistema S_1
 CONST.: interconexão suposta constante até a obtenção de um novo valor
 ARMA : modelo ARMA para interconexão

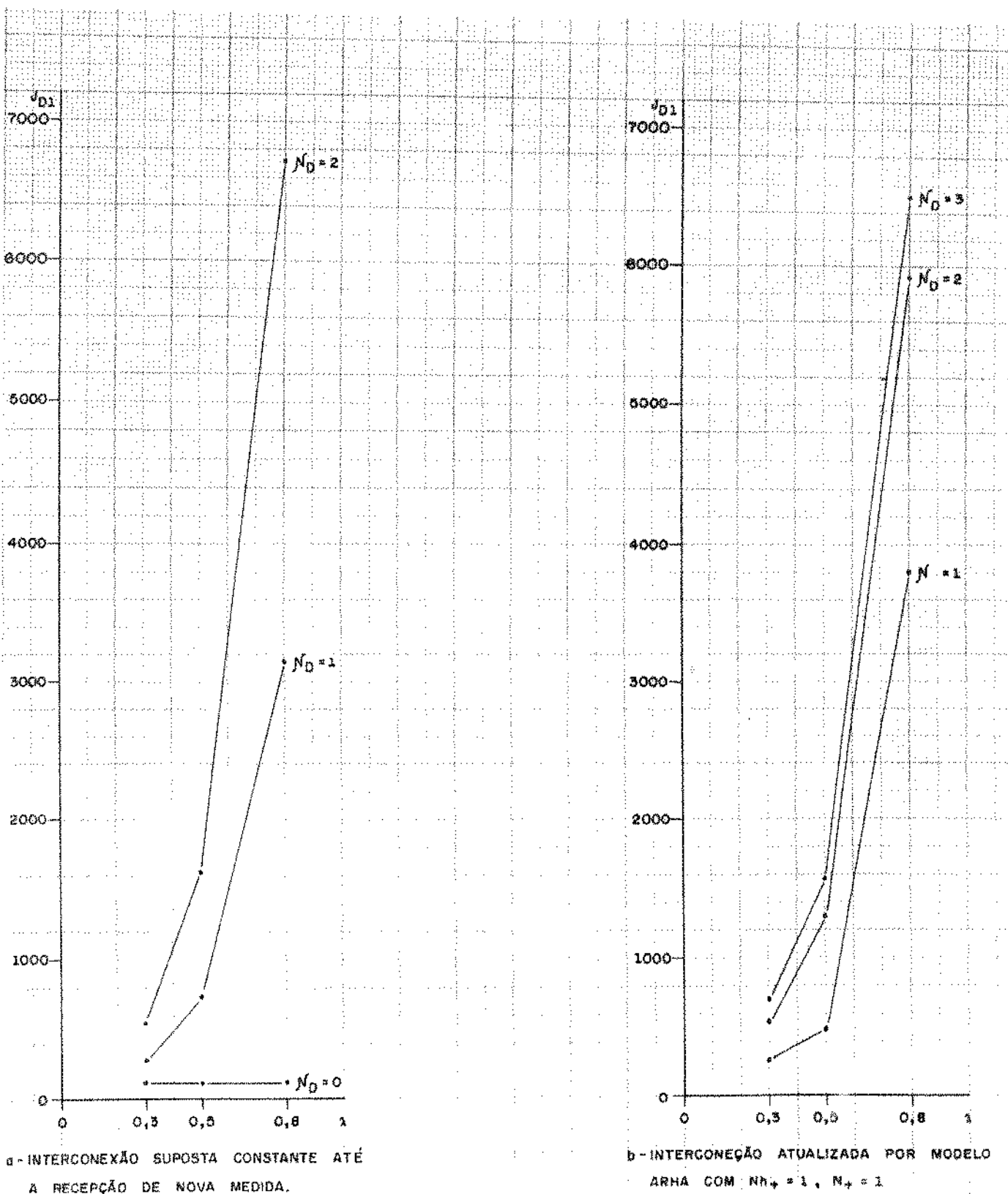


Figura C.2 - Custo J_{D1} x λ para o Algoritmo ERRA1/1 no Sub sistema S_1

$\lambda = 0,3$

N_D	J_{D1}	J_{D2}	J_{D3}	J_{D4}	Parâmetros
0	160	1	3	2	Valor Nominal
1	813	1/2	3	2	Valor Nominal
2	865	1/3	3	2	Valor Nominal
3	941	1/4	3	2	Valor Nominal

$\lambda = 0,5$

N_D	J_{D1}	J_{D2}	J_{D3}	J_{D4}	Parâmetros
0	156	1	3	2	Valor Nominal
1	1899	1/2	3	2	Convergente
2	2293	1/3	3	2	Convergente

$\lambda = 0,8$

N_D	J_{D1}	J_{D2}	J_{D3}	J_{D4}	Parâmetros
0	157	1	3	2	Valor Nominal
1	4123	1/2	3	2	Convergente
2	6640	1/3	3	2	Convergente

Tabelas C.4 - Algoritmo ERRA1/1 para o Subsistema S_2
 Interconexões supostas constantes até a recepção
 de novas medidas
 Número de estimações centralizadas para iniciali-
 zação : 10
 Número de estimações descentralizadas : 500

$\lambda = 0,3$

Ordens		Parâmetros	Custos
Nh_	N_	Estimados	
0	1	$G_-^1 = -0,060$	465
0	2	$G_-^1 = -0,066$	478
		$G_-^2 = 0,007$	
1	0	$H_-^1 = 0,337$	469
1	1	$H_-^1 = 0,599$	313
		$G_-^1 = -0,102$	
1	2	$H_-^1 = 0,600$	321
		$G_-^1 = -0,115$	
		$G_-^2 = 0,014$	
2	0	$H_-^1 = 0,305$	470
		$H_-^2 = -0,094$	
2	2	$H_-^1 = 0,724$	319
		$H_-^2 = 0,245$	
		$G_-^1 = -0,093$	
		$G_-^2 = -0,030$	

 $\lambda = 0,5$

Ordens		Parâmetros	Custos
Nh_	N_	Estimados	
0	1	$G_-^1 = -0,170$	1285
0	2	$G_-^1 = -0,198$	1311
		$G_-^2 = 0,035$	
1	0	$H_-^1 = 0,153$	1627
1	1	$H_-^1 = 0,622$	855
		$G_-^1 = -0,293$	
1	2	$H_-^1 = 0,620$	875
		$G_-^1 = -0,305$	
		$G_-^2 = 0,015$	
2	0	$H_-^1 = 0,137$	1636
		$H_-^2 = -0,101$	
2	2	$H_-^1 = 0,728$	877
		$H_-^2 = 0,260$	
		$G_-^1 = -0,250$	
		$G_-^2 = -0,098$	

 $\lambda = 0,8$

Ordens		Parâmetros	Custos
Nh_	N_	Estimados	
0	1	$G_-^1 = -0,464$	3405
0	2	$G_-^1 = -0,544$	3338
		$G_-^2 = 0,141$	
1	0	$H_-^1 = -0,071$	5816
1	1	$H_-^1 = 0,612$	2176
		$G_-^1 = -0,750$	
1	2	$H_-^1 = 0,645$	2218
		$G_-^1 = -0,738$	
		$G_-^2 = -0,048$	
2	0	$H_-^1 = -0,084$	5740
		$H_-^2 = 0,175$	
2	2	$H_-^1 = 0,729$	2217
		$H_-^2 = 0,261$	
		$G_-^1 = -0,631$	
		$G_-^2 = -0,270$	

Tabelas C.5 - Modelos para a Interconexão do Subsistema S_2 na forma

$$h_{K-1}^{2,1} = (H_-^1 h_{K-2}^{2,1} + \dots + H_-^{N_h} h_{K-N_h-1}^{2,1}) + G_-^1 y_{K-2}^2 + \dots + G_-^{N_-} y_{K-N_-1}^2 + \omega_K^-$$

Número de medidas utilizadas na elaboração do modelo : 500

Método de estimação : mínimos quadrados recursivo

$$\lambda = 0,3$$

Interconexão $N_h = 1$, $N = 1$

$$P(0) = 0,01I \quad J_{D3} = 5$$

N_D	J_{D1}	J_{D2}	Parâmetros
1	304	1/2	Valor Nominal
2	414	1/3	Valor Nominal
3	478	1/4	Valor Nominal
4	519	1/5	Valor Nominal

$$\lambda = 0,3$$

Interconexão $N_h = 2$, $N = 2$

$$P(0) = 10^{-4}I \quad J_{D3} = 7$$

N_D	J_{D1}	J_{D2}	Parâmetros
1	304	1/2	Valor Nominal
2	396	1/3	Valor Nominal
3	473	1/4	Valor Nominal
4	514	1/5	Valor Nominal

$$\lambda = 0,5$$

*Interconexão $N_h = 2$, $N = 2$

$$P(0) = 0,01I \quad J_{D3} = 7$$

N_D	J_{D1}	J_{D2}	Parâmetros
1	584	1/2	Valor Nominal
2	900	1/3	Valor Nominal
3	1050	1/4	Valor Nominal
4	1158	1/5	Convergente

$$\lambda = 0,3$$

Interconexão $N_h = 2$, $N = 2$

$$P(0) = 0,01I \quad J_{D3} = 7$$

N_D	J_{D1}	J_{D2}	Parâmetros
1	-	-	-
2	425	1/3	Valor Nominal
3	487	1/4	Valor Nominal
4	526	1/5	Valor Nominal

$$\lambda = 0,5$$

Interconexão $N_h = 1$, $N = 1$

$$P(0) = 0,01I \quad J_{D3} = 5$$

N_D	J_{D1}	J_{D2}	Parâmetros
1	562	1/2	Valor Nominal
2	864	1/3	Valor Nominal
3	1022	1/4	Valor Nominal
4	1103	1/5	Valor Nominal

$$\lambda = 5$$

*Interconexão $N_h = 1$, $N = 1$

$$P(0) = 0,01I \quad J_{D3} = 5$$

N_D	J_{D1}	J_{D2}	Parâmetros
1	954	1/2	Valor Nominal
2	3112	1/3	Convergente
3	5201	1/4	Convergente
4	-	-	-

* O valor da última previsão feita para a interconexão é atualizado antes de sua utilização para o cálculo de uma nova previsão.

Tabelas C.6 - Algoritmo ERRA1/1 para o Subsistema S_2

As ordens os modelos utilizados para a interconexão e o valor inicial da matriz P para o algoritmo de mínimos quadrados que atualiza os parâmetros destes modelos estão indicados acima de cada tabela.

Número de estimações centralizadas para inicialização : 40

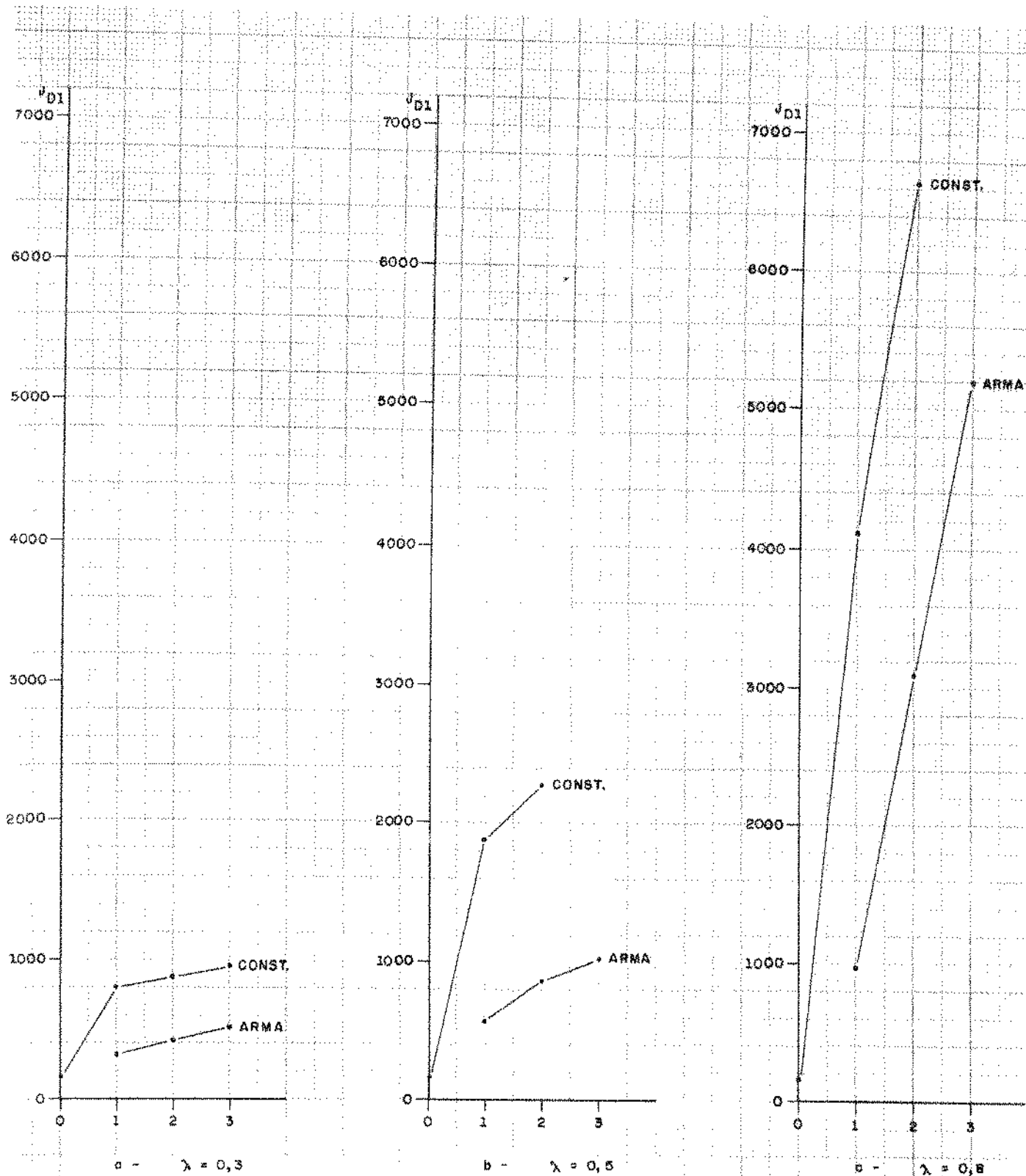


Figura C.3 - Custo $J_{D1} \times n_D$ para o Algoritmo ERRAl/1 no Subsis_{tema} S_2

CONST: interconexão suposta constante até a obtenção de um novo valor

ARMA : modelo ARMA para interconexão

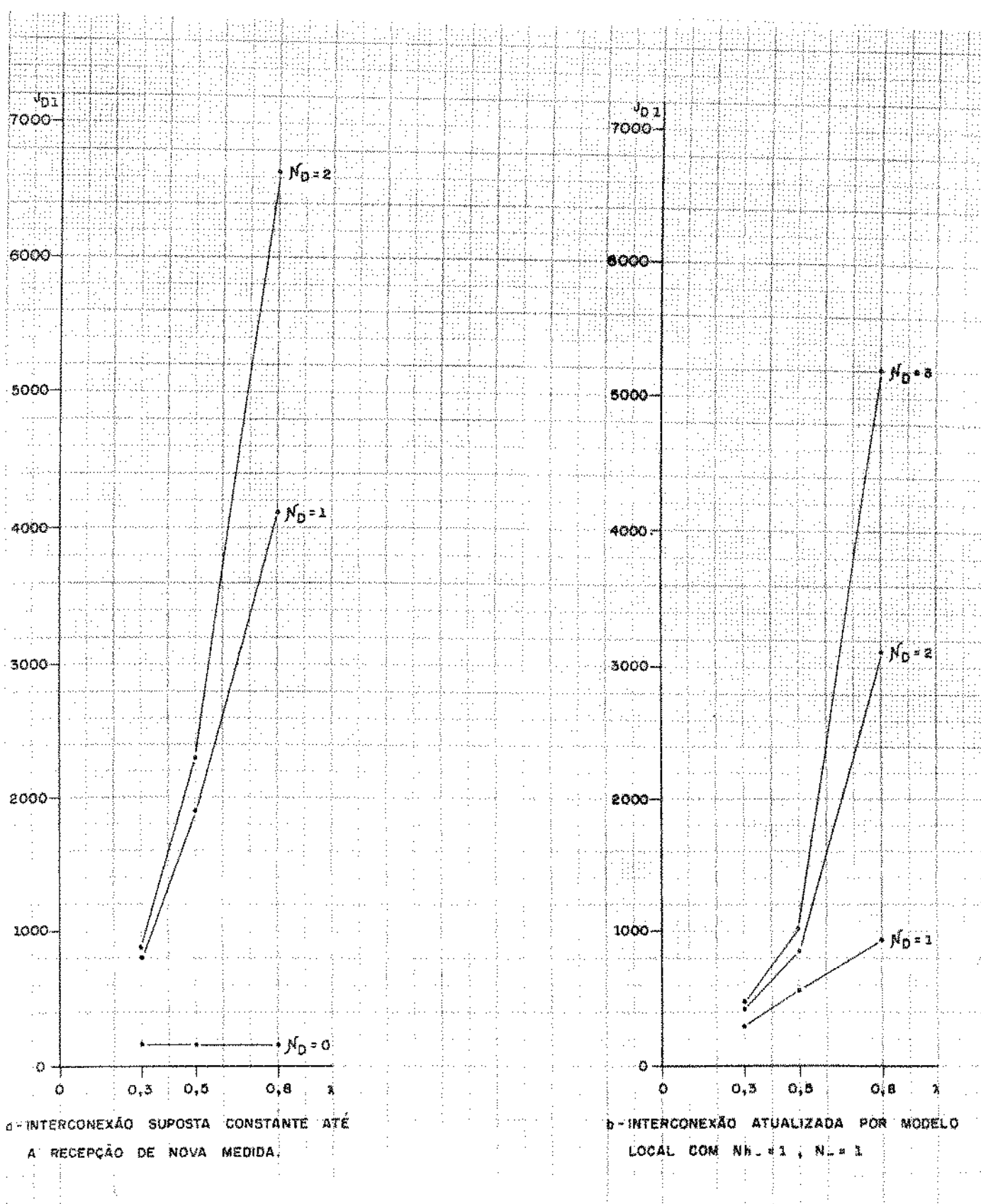


Figura C.4 - Custo $J_{D1} \times \lambda$ para o Algoritmo ERR1/1 no Subsistema S_2

APÊNDICE D

RESULTADOS DE SIMULAÇÃO PARA O ALGORÍTMO ERRB2/1

Neste Apêndice estão as tabelas e gráficos com os resultados da simulação do algoritmo ERRB2/1 para o sistema descrito na seção 4.4.

Os comentários referentes a estes gráficos e tabelas são apresentados na seção 4.6.

λ	EQ_{MIN}	$K.J_{D2}$ em 500 estimações	J_{D1} em 500 estimações	
			s_1	s_2
0,3	0,0	500	135	161
	0,2	314	191	394
	1,0	172	336	693
	5,0	92	791	794
	10,0	45	1204	816
0,5	0,0	500	138	157
	0,1	384	260	518
	0,2	290	357	797
	0,5	279	451	1078
0,8	0,0	500	145	155
	0,05	440	356	640
	0,1	412	456	893

Tabela D.1 - Índices J_{D1} e $K.J_{D2}$ para os valores de λ e EQ_{MIN} utilizados nas Simulações do Algoritmo ERB2/1

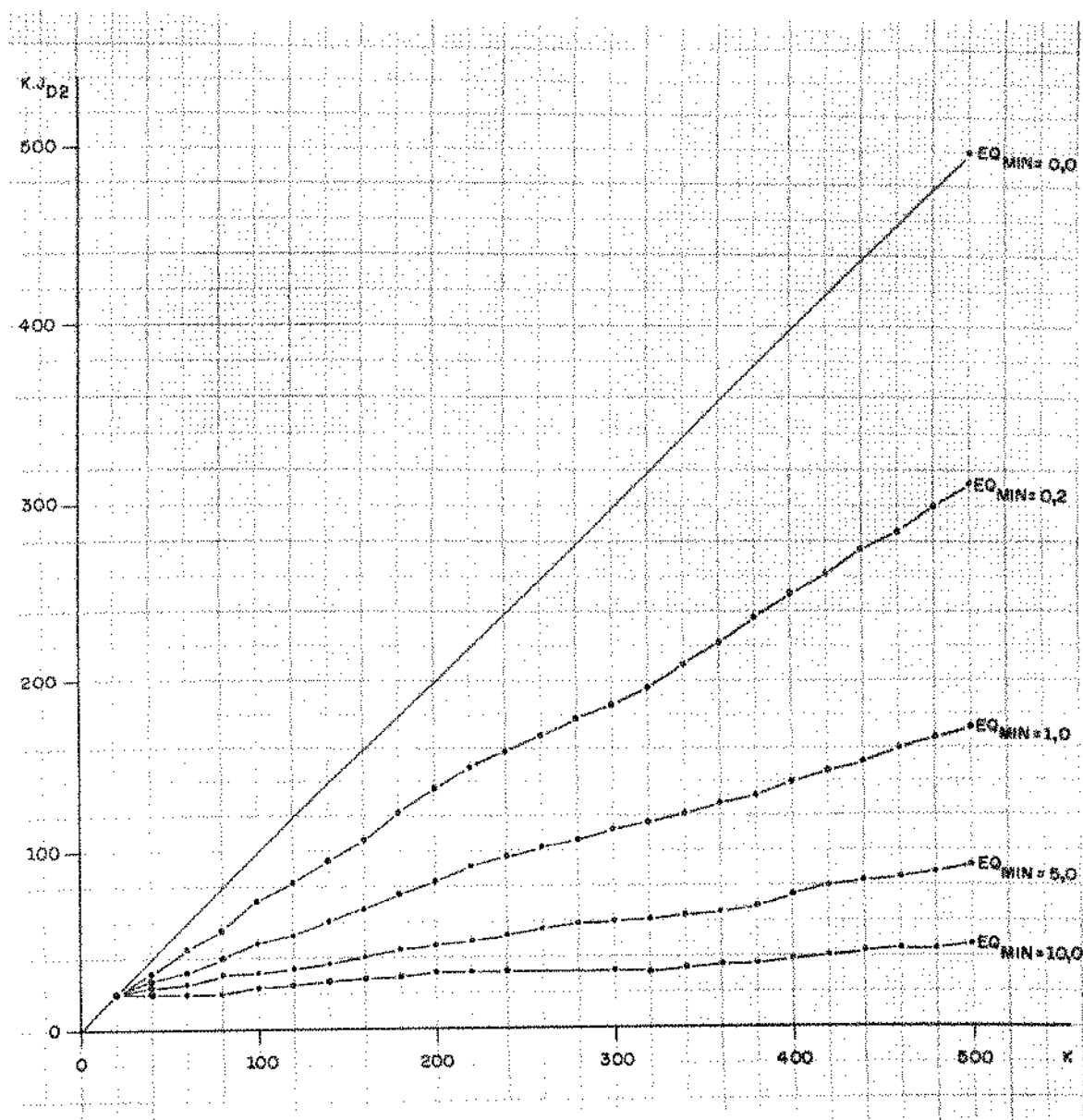


Figura D.1 - Índice $K.J_{D2}$ x Número de Estimações para $\lambda = 0,3$
Algoritmo ERB2/1

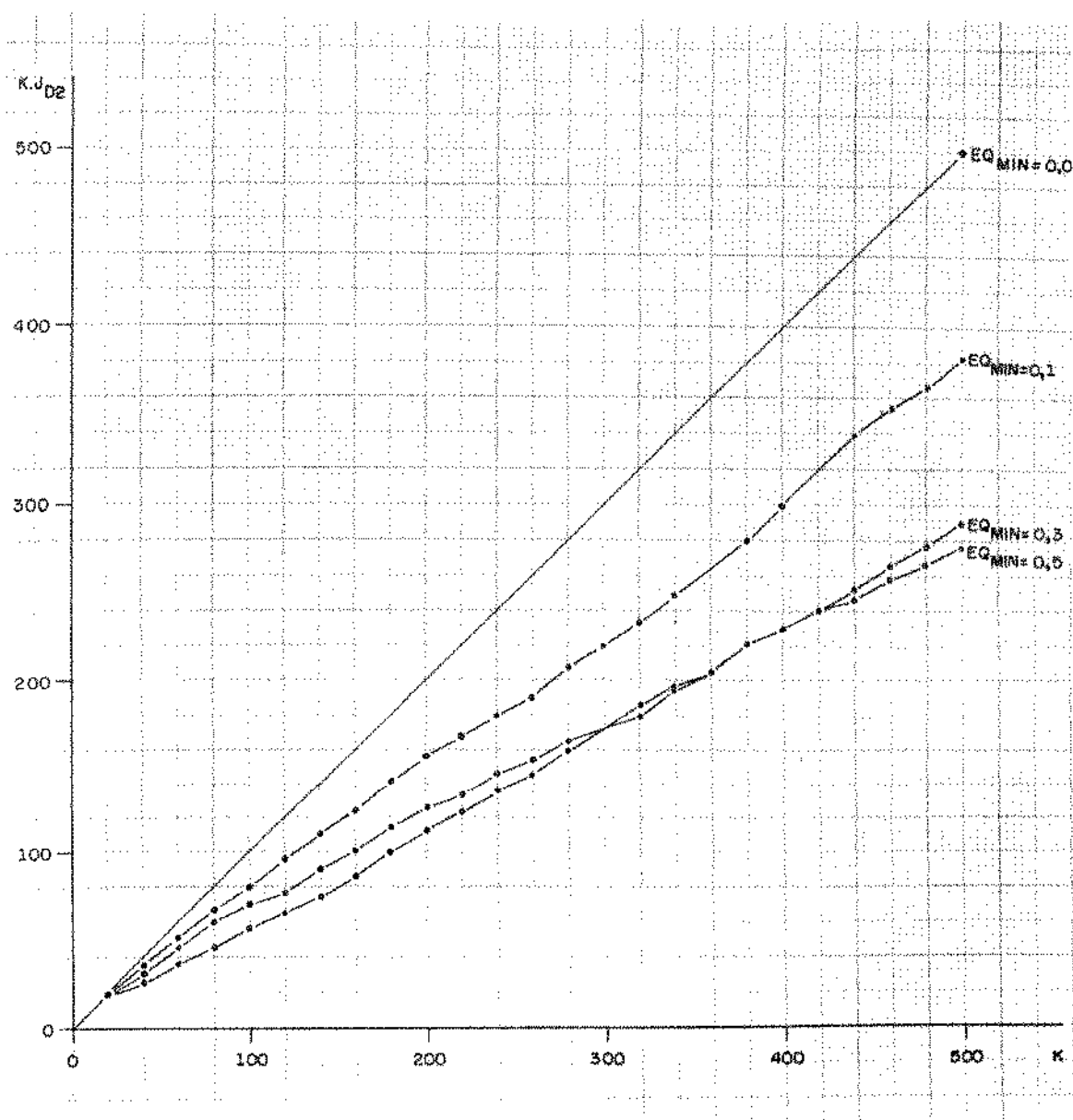


Figura D.2 - Índice $K.J_{D2}$ x Número de Estimações para $\lambda = 0,5$
Algoritmo ERB2/1

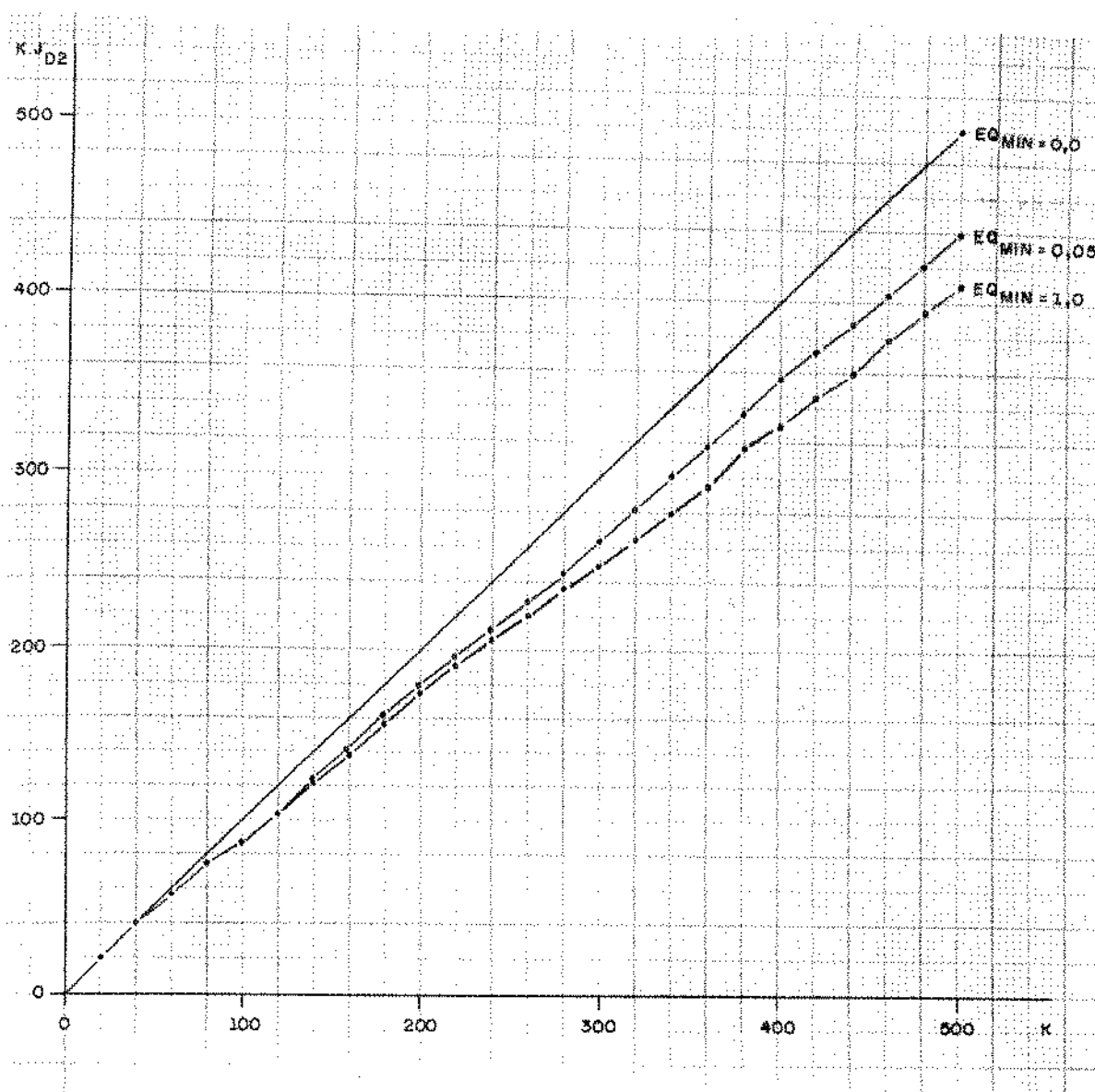


Figura D.3 - Índice $K.J_{D2}$ x Número de Estimções para $\lambda = 0,8$
Algoritmo ERB2/1

K	$\delta_{K,3}$
100	0,0741
200	0,0510
300	0,334
400	0,0282
500	0,0124

K	$\delta_{K,3}$
100	0,1602
200	0,1753
300	0,0292
400	0,0376
500	0,0394

K	$\delta_{K,3}$
100	0,8211
200	0,2412
300	0,0821
400	0,1220
500	0,1291

K	$\delta_{K,3}$
100	1,1856
200	0,5840
300	1,2406
400	0,4175
500	0,5713

K	$\delta_{K,3}$
100	0,4494
200	0,1385
300	1,5075
400	0,5643
500	0,7321

EQ_{MIN}

0,0

0,2

1,0

5,0

10,0

Tabelas D.2 - Distâncias Paramétricas Normalizadas para o Algoritmo ERRB2/1 no Subsistema S_1 com $\lambda = 0,3$

K	$\delta_{K,3}$
100	0,1475
200	0,0849
300	0,0526
400	0,0223
500	0,0142

K	$\delta_{K,3}$
100	0,2025
200	0,1537
300	0,0466
400	0,0955
500	0,1938

K	$\delta_{K,3}$
100	0,3912
200	0,1695
300	0,3722
400	0,2171
500	0,0428

K	$\delta_{K,3}$
100	0,3594
200	0,6500
300	0,5536
400	0,3801
500	0,4468

EQ_{MIN}

0,0

0,1

0,3

0,5

Tabelas D.3 - Distâncias Paramétricas Normalizadas para o Algoritmo ERB2/1 no Subsistema S_1 com $\lambda = 0,5$

K	$\delta_{K,3}$
100	0,0582
200	0,0349
300	0,0330
400	0,0823
500	0,0513

K	$\delta_{K,3}$
100	0,0712
200	0,0617
300	0,0795
400	0,1666
500	0,0659

K	$\delta_{K,3}$
100	0,1565
200	0,1068
300	0,1583
400	0,0341
500	0,1532

EO_{MIN}

0,0

0,05

0,1

Tabelas D.4 - Distâncias Paramétricas Normalizadas para o Algoritmo ERB2/1 no Subsistema S_1 com $\lambda = 0,8$

K	$\delta_{K,3}$
100	0,1098
200	0,0448
300	0,0270
400	0,0318
500	0,0319

K	$\delta_{K,3}$
100	0,0464
200	0,0529
300	0,0800
400	0,0890
500	0,0826

K	$\delta_{K,3}$
100	0,0673
200	0,1607
300	0,0225
400	0,2595
500	0,2377

K	$\delta_{K,3}$
100	0,1638
200	0,1717
300	0,0708
400	0,0344
500	0,1640

K	$\delta_{K,3}$
100	0,1983
200	0,1615
300	0,0704
400	0,2236
500	0,1198

EQ_{MIN}

0,0

0,2

1,0

5,0

10,0

Tabelas D.5 - Distâncias Paramétricas Normalizadas para o Algoritmo ERRB2/1 no Subistema S_2 com $\lambda = 0,3$

K	$\delta_{K,3}$
100	0,110
200	0,0207
300	0,0641
400	0,0776
500	0,0816

K	$\delta_{K,3}$
100	0,116
200	0,0731
300	0,0866
400	0,0512
500	0,0640

K	$\delta_{K,3}$
100	0,1467
200	0,0426
300	0,1265
400	0,2846
500	0,2815

K	$\delta_{K,3}$
100	0,2424
200	0,6474
300	0,5793
400	0,4513
500	0,2810

EQ_{MIN}

0,0

0,1

0,3

0,5

Tabelas D.6 - Distâncias Paramétricas Normalizadas para o Algoritmo ERRB2/1 no Subistema S_2 com $\lambda = 0,5$

K	$\delta_{K,3}$
100	0,0464
200	0,0243
300	0,0799
400	0,0991
500	0,0775

K	$\delta_{K,3}$
100	0,1576
200	0,2805
300	0,2312
400	0,1096
500	0,3257

K	$\delta_{K,3}$
100	0,5123
200	0,2055
300	0,4206
400	0,4363
500	0,5666

EQ_{MIN}

0,0

0,05

0,1

Tabelas D.7 - Distâncias Paramétricas Normalizadas para o Algoritmo ERRB2/1 no Subsistema S_2 com $\lambda = 0,8$

LISTA DE SÍMBOLOS E ABREVIATURAS

A lista seguinte reúne os símbolos e abreviaturas presentes neste trabalho:

- $\underline{a}, a$: vetor-coluna da matriz $A(q^{-1})$ - elemento desta matriz
- $\underline{b}, b$: vetor-coluna da matriz $B(q^{-1})$ - elemento desta matriz
- $\underline{c}, c$: vetor-coluna da matriz $C(q^{-1})$ - elemento desta matriz
- $\underline{d}, d$: vetor-coluna da matriz $D(q^{-1})$ - elemento desta matriz
- $\underline{f}$: função vetorial, indica o comportamento determinístico de um processo
- $\underline{g}(\cdot)$: função das saídas dos demais subsistemas transmitida ao subsistema S_i
- $\underline{h}_K^{i,j}$: parcela da saída, no instante K , do subsistema S_i devida a influência do subsistema S_j
- $\begin{Bmatrix} i \\ j \end{Bmatrix}$: índices inteiros
- K_a : atraso das informações recebidas pelo subsistema S_i em relação à amostragem neste subsistema
- K : índice usado para indicar um número de intervalos de amostragem nos algoritmos recursivos
- ℓ : índice inteiro
- m_t : média condicional de um processo no qual são feitas observações sequenciais
- m : dimensão do vetor entrada de um sistema
- m^i : dimensão do vetor entrada só subsistema S_i
- n : dimensão do vetor saída de um sistema
- n^i : dimensão do vetor saída do subsistema S_i
- n_a : número acumulado de saídas dos subsistemas

- ns_a : número acumulado de subsistemas de um sistema descentralizado
- np : número de parâmetros a serem estimados
- $p(\cdot)$: densidade de probabilidade
- p : potência total das saídas de um sistema
- p_a : potência acumulada das saídas dos subsistemas
- q : operador de avanço temporal
- r : índice inteiro
- t : instante de tempo
- $\underline{u}, u$: vetor-entrada de um sistema - elemento deste vetor
- $\underline{u}^i, u^{i,\ell}$: vetor-entrada do subsistema S_i , ℓ -ésima componente deste vetor
- $\underline{\omega}, \omega$: vetor-ruído branco - elemento deste vetor
- $\underline{\omega}^i$: vetor-ruído branco no modelo do subsistema S_i
- $\underline{\omega}^i_{-Y}$: ruído branco no modelo local do erro de descentralização
- $\underline{\omega}^-_{-K}$: ruído branco no modelo de $\underline{h}^{i,j}_{-K-1}$ nos algoritmos ERRA1
- $\underline{\omega}^+_{-K}$: ruído branco no modelo de $\underline{h}^{i,j}_{-K-1}$ nos algoritmos ERRA1
- x : valor numa tabela de distribuição de probabilidade
- $\underline{y}, y$: vetor-saída de um sistema - elemento deste vetor
- $\hat{\underline{y}}$: valor estimado do vetor-saída
- $\underline{y}^d$: valor desejado do vetor-saída
- $\underline{y}^i, y^{i,\ell}$: vetor-saída do subsistema S_i . ℓ -ésima componente deste vetor
- $\left\{ \begin{array}{l} A(q^{-1}) \\ B(q^{-1}) \\ C(q^{-1}) \\ D(q^{-1}) \end{array} \right.$: matrizes polinomiais utilizadas na descrição do modelo do processo
- D : índice de distribuição das saídas entre os subsistemas

- D* : índice de distribuição das potências entre os sub sistemas
- (E) : relação de equivalência
- E(.) : esperança de uma variável aleatória
- EQ_{MIN} : valor mínimo dos erros locais de previsão dos sub sistemas para atualização dos erros de descentra lização na simulação do algoritmo ERB2/1
- F(q⁻¹) : matriz polinomial utilizada na descrição do mode lo do processo
- F(.,.) : distribuição de probabilidade
- H(q⁻¹) : matriz racional
- I : matriz identidade
- J : critério a ser minimizado com a escolha dos parâ metros
- $\left\{ \begin{array}{l} J_{D1}^i \\ J_{D2}^i \\ J_{D3}^i \\ J_{D4}^i \end{array} \right\}$: índices para avaliação do desempenho de algoritmos descentralizados em cada subsistema S_i (i = 1, ..., ns)
- K : ganho nos estimadores recursivos
- M : modelo de cada subsistema
- M₀ : matriz de informação de Fisher
- N_A : ordem da matriz polinomial A(q⁻¹)
- N_B : ordem da matriz polinomial B(q⁻¹)
- N_C : ordem da matriz polinomial C(q⁻¹)
- N_D : ordem da matriz polinomial D(q⁻¹)
- N_R : ordem da matriz polinomial R(q⁻¹)
- N* : ordem do modelo do processo quando N_A=N_B=N_C=N_D=N_R $\hat{=}$ N*
- N : número de medidas feitas no processo
- P : matriz propagada nos estimadores recursivos
- Q : matriz no estimador raiz quadrada
- R(q⁻¹) : matriz polinomial utilizada na descrição do modelo do processo

S_i : i-ésimo subsistema de um sistema descentralizado

T : intervalo de tempo

U : matriz cujas colunas são vetores-entrada

X : matriz ortonormal no estimador raiz quadrada

Y : matriz cujas colunas são vetores-saída

$A^{i, \cdot} (q^{-1})$: matriz polinomial usada na descrição do modelo de $\underline{y}^i$

B : matriz propagada no estimador da raiz quadrada

D^* : conjunto usado na definição das formas canônicas

$F^i(q^{-1})$: matriz usada no modelo de $\underline{y}_K^i$ nos algoritmos ERRA

$\left\{ \begin{array}{l} G^i \\ H^i \end{array} \right.$: matrizes usadas nos modelos de $\underline{h}^{i,i-1}$ e $\underline{h}^{i,i+1}$ nos algoritmos ERRA1

$I^i(q^{-1})$: matriz usada no modelo de $\underline{y}^i$ nos algoritmos ERRA

N_D : nível de degradação num sistema descentralizado com descontinuidade na troca de informações

N_A : nível de degradação num sistema descentralizado com atraso na troca de informações

N_I : nível de degradação num sistema descentralizado com interrupções na troca de informações

P : família de densidades condicionais de probabilidade

R : conjunto dos números reais

X : variável aleatória

C_s : escalares usadas no estimador raiz quadrada

α : índice

p : variável aleatória

$\underline{y}_K^i$: vetor erro de descentralização

$\underline{\delta}_K$: vetor ruído (branco ou colorido) na saída de um sub sistema

$\delta_{K,np}^2$: distância paramétrica normalizada de um processo com nq parâmetros estimados no instante K

$\hat{\underline{\sigma}}_K^i$: vetor ruído (branco ou colorido) na saída do sub
sistema S_i

$\underline{\varepsilon}$: vetor erro de previsão

$\underline{\zeta}_K$: vetor ruído colorido na saída de um subsistema

$\underline{\zeta}_K^i$: vetor ruído colorido na saída do subsistema S_i

$\underline{\theta}$: vetor de parâmetros a serem estimados

λ : razão de verossimilhança

μ_p : microprocessador

π : 3,14159...

$\underline{\phi}$: vetor que compõe a matriz ψ

ψ : matriz que reúne as medidas feitas, utilizadas no
cálculo do estimador

Γ : matriz densidade espectral racional

Δ : variação

$\left\{ \begin{array}{l} \Delta n \\ \Delta n^i \end{array} \right.$: número de intervalos de amostragem até a renovação
do erro de descentralização pelo processador cen
tral nos algoritmos ERRE

Δ : definição

Θ : conjunto dos valores de θ

Π : produto

Σ : soma

$\underline{\Omega}$: vetor usado no estimador raiz quadrada

ETHERNET : rede produzida pela Xerox Corporation

ALOHA

"TOKEN PASSING"

HXDP

: protocolos para redes de computação distribuída

HITRIX

CSMA

CSMA/CD

DMP : descrição por matriz partilhada

ERRA : estimador recursivo para redes anel

ERRE : estimador recursivo para redes estrela
ERRB : estimador recursivo para redes com barramento
COV : covariância
det : determinante
exp : exponencial de base $e = 2,7182...$
 $\ln$: logarítmo neperiano
max(.) : valor máximo da variável entre parênteses
vec : vetor das colunas de uma matriz
 $\otimes$: produtos de Kronecker de matrizes
 $\hat{\cdot}$: valor estimado

OBS: Esta lista não inclui o material dos Apêndices.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Amaral, W.C., "Identificação Recursiva de Sistemas Multivariáveis", Tese de Doutorado, Faculdade de Engenharia Elétrica, UNICAMP, 1980.
- [2] Anderson, G.A., "Interconnecting Distributed Processor System for Avionics", 1^d Annual Symposium on Computer Architecture Proceedings, 1973.
- [3] Aström, K.J., Introduction to Stochastic Control Theory, Academic Press, New York, 1970.
- [4] Aström, K.J., "Maximum Likelihood and Prediction Error Methods", Tutorials on System Identification, 5th. IFAC - Symposium on Identification and System Parameter Estimation, Darmstadt, September, 1979.
- [5] Buchner, M.R. and Lefkowitz, I., "Distributed Computer Control for Industrial Process Systems: Characteristics, Attributes and an Experimental Facility", IEEE Control Systems Magazine, March 1982.
- [6] Fasol, K.H. and Jorge, H.P., "Modelling and Identification", Tutorials on System Identification: 5th. IFAC - Symposium on Identification and System Parameter Estimation, Darmstadt, September, 1979.
- [7] Gohberg, I.C. and Krein, M.C., "On the Factorization of Operators in Hilbert Spaces", Amer. Math. Soc. Trans. 51, 1966.
- [8] Hannan, E.J., "The Identification Problem for Multiple Equation Systems with Moving Average Errors", Econometrics 39, 1971.
- [9] Isermann, R., "Practical Aspects of Process Identification", Tutorials on System Identification, 5th. IFAC - Symposium on Identification and System Parameter Estimation, Darmstadt, September, 1979.

- [10] Isermann, R., Digital Control Systems, Springer-Verlag, 1981.
- [11] Jafari, H. et al., "Simulation at a Class of Ring Structured Networks", IEEE Trans. on Computers, Vol. 29, Nº 5, May 1980.
- [12] Jensen, E.D., "The Honeywell Experimental Distributed Processor: An Overview", IEEE Computer Magazine, Jan 1978.
- [13] Kashyap, R.L. and Rao, A.R., Dynamic Stochastic Models from Empirical Data, Academic Press, 1976.
- [14] Knuth, D.E., The Art of Computer Programming, Vol. 2: Semi-numerical Algorithms, Addison-Wesley Publishing Co., 1969.
- [15] Kuo, B.C., Automatic Control Systems, Prentice-Hall, Inc., 1975.
- [16] Lages, N.A.L., "Comunicação em Redes Locais de Computadores para Controle Distribuído de Processos", Tese de Doutorado, Faculdade de Engenharia Elétrica, UNICAMP, 1981.
- [17] Liu, M.T., "Distributed Loop Computer Networks", Advances in Computers, Vol. 17, pp. 163/221, Academic Press, 1978.
- [18] Liu, M.T. et al., "System Design of the Distributed Double-Loop Computer Network (DOLCN)", Proceedings of the 1st International Conference on Distributed Computing Systems, Huntsville, Alabama, October, 1979.
- [19] Luzak, E.C., "Global Bus Computer Communication Techniques", Proceedings of the Computer Networking Symposium, Dez. 1978.
- [20] Metcalfe, R.H., "Ethernet: Distributed Packet Switching for Local Computer Networks", Communications of the ACM, Vol. 19, Nº 7, July 1976.
- [21] Payne, R.L. and Goodwin, G.C., Dynamic System Identification: Experimental Design and Data Analysis, Academic Press, 1977.

- [22] PDP-11 FORTRAN-77, Language Reference Manual, Digital Equipment Corporation, May 1982.
- [23] PDP-11 FORTRAN-77 User's Guide, Digital Equipment Corporation, May 1982.
- [24] PDP-11 FORTRAN-77 Object Time System Reference Manual, Digital Equipment Corporation, May 1982.
- [25] PDP-11 FORTRAN-77 Installation Guide/Release Notes, Digital Equipment Corporation, May 1982.
- [26] Rossetti, J.P., Introdução à Economia, Editora Atlas S.A., 1974 (3ª edição).
- [27] Rozanov, Y.A., Stationary Random Process, Holden-Day, San Francisco, 1967.
- [28] Sandell Jr., N.R. et al., "Survey of Decentralized Control Methods for Large Scale Systems", IEEE Trans. on Automatic Control, Special Issue on Large-Scale Systems and Dezentralized Control, Vol. AC-23, Nº 2, April 1978.
- [29] Steusloff, H., "Structures of Automatic Control Systems and Consequences for Programming Languages", Process Automation, 1979.
- [30] Strejc, V., "Least Squares Parameter Estimation", Tutorials on System Identification, 5th. IFAC - Symposium on Identification and System Parameter Estimation, Darmstadt, Setp. 1979.
- [31] Youla, D.C. "On the Factorization of Rational Matrices", IEEE Trans. on Inform. Theory, IT-7, pp. 172-189, 1961.