

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
DEZEMBRO DE 1980

IDENTIFICAÇÃO RECURSIVA
DE SISTEMAS MULTIVARIÁVEIS

43/80

Por : Wagner Caradori do Amaral
Orientador : Prof. Luis Gimeno Latre

Tese apresentada à Faculdade de Engenharia
FEC-UNICAMP como parte dos requisitos exi-
gidos para a obtenção do título de DOUTOR
EM ENGENHARIA

UNICAMP
BIBLIOTECA CENTRAL

Agradecimento

Desejo expressar meu agradecimento a todas as pessoas que direta ou indiretamente colaboraram na realização deste trabalho.

O agradecimento especial ao orientador Dr. Luís Gimeno Latre e ao Dr. Manuel de Jesus Mendes pelas idéias, discussões, incentivo, apoio e oportunidade oferecidas.

Ao Sidney, Taka e Miriam pelo apoio na edição das figuras e textos aqui apresentados.

RESUMO

Neste trabalho são analisados algoritmos de identificação para sistemas com múltiplas entradas e múltiplas saídas (MIMO). É feita uma generalização dos métodos da matriz estendida e dos mínimos quadrados generalizados para sistemas MIMO, considerando-se um modelo autoregressivo e média-móvel (ARMA) para o ruído. Utilizando a estrutura ARMA para o ruído obtem-se também o estimador dos mínimos quadrados linearizado para sistemas MIMO. Este estimador, para o caso média-móvel, corresponde a uma generalização do método de máxima verossimilhança apresentado por Soderstrom para sistemas monovariáveis. Em todos estes algoritmos são estudadas as condições necessárias para se obter, através de uma partição no problema de estimação, estimadores com tempo de cálculo reduzido. A partir do estimador da Matriz Estendida obtem-se um novo estimador que calcula separadamente os parâmetros do processo e do ruído, com uma redução substancial no tempo de cálculo. Finalmente, ilustram-se os resultados obtidos através de exemplos numéricos simulados e de uma aplicação prática para a previsão da altura de aço, no molde de uma máquina de lingotamento contínuo.

ÍNDICE

CAPÍTULO 1 - INTRODUÇÃO	1
1.1 - Identificação de sistemas. Objetivos do trabalho	1
1.2 - Modelagem de sistemas estocásticos lineares através de equações à diferenças..	6
1.3 - Formas canônicas para modelos em forma de equações à diferenças.....	11
1.4 - Conclusões	14
CAPÍTULO 2 - ALGORITMOS PARA ESTIMAÇÃO DE PARÂMETROS EM SISTEMAS MULTIVARIÁVEIS DESCRITOS POR EQUAÇÕES À DIFERENÇAS	15
2.1 - Introdução	15
2.2 - Obtenção de um modelo de regressão linear	16
2.3 - Estimador de Markov	17
2.4 - Algoritmo dos mínimos quadrados	20
2.5 - Algoritmo dos mínimos quadrados generalizado.....	22
2.6 - Algoritmo da matriz estendida	26
2.7 - Estimador de máxima verossimilhança	29
2.8 - Formulação do problema de estimação através do modelo erro previsto	31
2.8.1 - Modelo do erro previsto	32
2.8.2 - Definição do critério. Hipóteses simplificadoras	33
2.8.3 - Relação entre a metodologia de previsão do erro e o estimador de máxima verossimilhança	35
2.8.4 - Modelo linear para o erro de previsão	37
2.9 - Conclusões.....	38

CAPÍTULO 3 - ESTIMAÇÃO RECURSIVA E PARTICIONADA DE SISTEMAS DESCRITOS POR EQUAÇÕES À DIFERENÇAS	40
3.1 - Introdução	40
3.2 - Estimador dos mínimos quadrados	41
3.3 - Estimador dos mínimos quadrados lineariz <u>ado</u>	49
3.3.1 - Algoritmo linearizado particiona <u>do</u> utilizando a forma pseudo-ca <u>nônica</u> I.....	59
3.3.2 - Algoritmo linearizado particiona <u>do</u> utilizando a forma canônica III	61
3.3.3 - Algoritmo linearizado particiona <u>do</u> utilizando a forma pseudo-ca <u>nônica</u> II	63
3.4 - Algoritmo da matriz estendida	64
3.5 - Algoritmo dos mínimos quadrados particio <u>nado</u>	67
3.6 - Algoritmo dos mínimos quadrados generaliz <u>ado</u>	73
3.6.1 - Algoritmo dos mínimos quadrados generalizado particionado utili <u>zando</u> a forma pseudo-canônica I..	77
3.6.2 - Algoritmo dos mínimos quadrados generalizado particionado utili <u>zando</u> a forma canônica III	80
3.6.3 - Algoritmo dos mínimos quadrados generalizado particionado utili <u>zando</u> a forma pseudo-canônica II.	81
3.7 - Conclusões	86
 CAPÍTULO 4 - ANÁLISE DOS ALGORITMOS ATRAVÉS DE EXEMPLOS DE SIMULAÇÃO	 88
4.1 - Introdução	88

4.2 - Critérios para comparação de desempenho dos algoritmos de estimação.....	89
4.3 - Comparação da eficiência de cálculo dos estimadores	91
4.3.1 - Tempo de cálculo dos estimadores.	91
4.3.2 - Memória	96
4.4 - Geração das medidas e inicialização	97
4.4.1 - Geração das medidas	97
4.4.2 - Inicialização	98
4.5 - Estimadores com modelos ARMA para o ruído	98
4.5.1 - Equivalentes exatos para o modelo ARMA	99
4.5.2 - Modelos ARMA sem equivalentes exatos	103
4.6 - Algoritmo MQL com forma canônica III de Kashyap	110
4.6.1 - Comparação do algoritmo MQL com os algoritmos ME e MQP	110
4.6.2 - Identificação de um sistema de lingotamento contínuo	112
CAPÍTULO 5 - CONCLUSÕES	116
REFERENCIAS BIBLIOGRÁFICAS	118

CAPÍTULO 1

1. INTRODUÇÃO

1.1. Identificação de sistemas. Objetivos do trabalho

Entende-se por *identificação* a determinação de um modelo do sistema representando os seus aspectos essenciais de uma forma adequada para a sua utilização [1].

A forma do modelo é determinada pelo tipo de sistema em consideração, por exemplo, estático, dinâmico, linear, não linear, estocástico, determinístico, etc.

A determinação do modelo no problema de identificação pode ser obtida através de duas linhas de ataque não necessariamente disjuntas:

- i) A *análise física matemática* ("model building") em que o modelo é obtido através da descrição do comportamento físico do sistema [2].
- ii) A *identificação experimental* [2, 3] em que se utiliza o fato do processo estar em funcionamento para, através de observações, construir um modelo paramétrico do sistema.

Através destes métodos obtém-se um modelo que represente os aspectos essenciais do sistema, tendo em vista a sua utilização. Não se pretende obter uma descrição matemática exata, mas um modelo adequado para uma determinada aplicação. Por exemplo, um modelo para a previsão e controle terá como requisito essencial uma representação suficientemente exata das características entrada/saída. Na prática usa-se a hipótese básica para construção de modelos, de que sistemas reais, em geral, não necessitam obrigatoriamente de modelos complicados.

No presente trabalho estuda-se o problema da identificação experimental tendo em vista que a aplicação em mente será sempre a de controle por realimentação de um processo dinâmico.

Os sistemas considerados são do tipo estocástico. Não se considera o problema de identificação determinística, válida apenas quando não existem perturbações ou estas podem ser desprezadas dada a aplicação futura do modelo. Considera-se assim que o problema de controle subjacente é um problema de controle estocástico.

Como o objetivo pretendido com a identificação é, em geral, o controle das saídas do processo, o modelo obtido deverá aproximar convenientemente o comportamento externo do sistema. Deste modo os métodos de identificação tratados neste trabalho procuram minimizar algum tipo de função do erro entre as saídas do modelo e as saídas reais. Nenhum dos métodos analisados minimiza 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 possibilitem uma adaptação contínua do modelo às variações no ponto de operação do sistema.

Pode-se caracterizar a solução recursiva do problema de estimação como uma tarefa de correção do estimador obtido no instante anterior, designado por $\hat{\theta}_k$, através de algum mecanismo de ajuste, como mostrado na figura 1.1, onde nota-se que a entrada (\underline{u}_k) é dirigida simultaneamente para o processo e para o modelo com parâmetros ajustáveis.

O mecanismo de cálculo determina valores estimados dos parâmetros minimizando alguma função do erro entre a saída do modelo e a saída do processo.

O problema fundamental consiste na obtenção dos algoritmos de estimação (Estimador) que, a partir dos valores medidos, gerem os valores estimados, satisfazendo critérios especificados previamente.

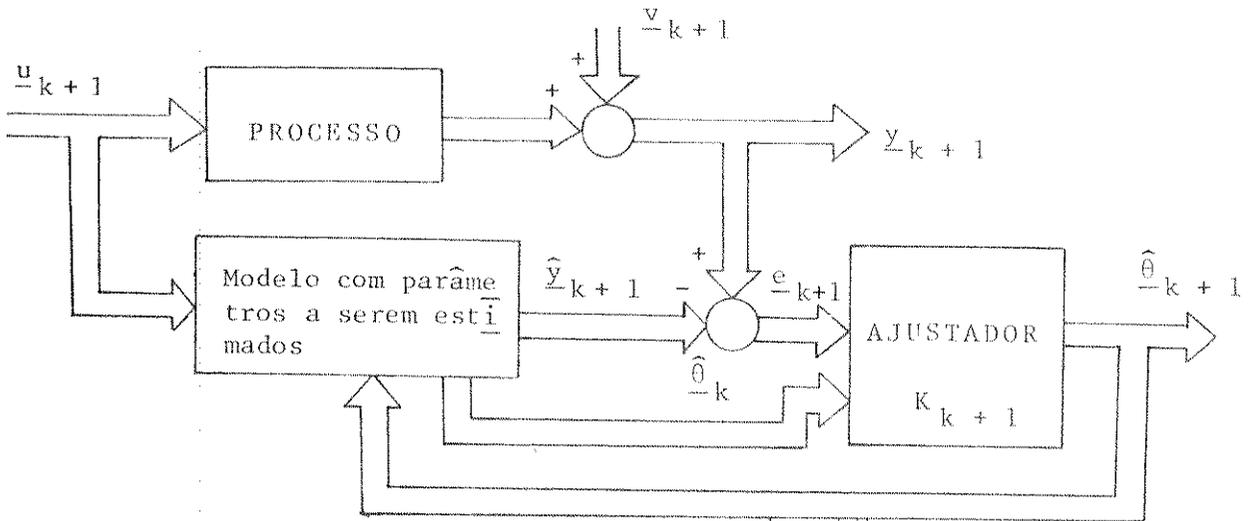


Figura 1.1 - Esquema de estimação recursiva

Na literatura considera-se a seguinte estrutura como sendo a mais utilizada para o estimador recursivo.

$$\hat{\underline{\theta}}_{k+1} = \hat{\underline{\theta}}_k + K_{k+1} \underline{e}_{k+1} \quad 1.1$$

onde

$\hat{\underline{\theta}}_k$ é o vetor dos valores estimados no instante k

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

K_{k+1} é uma matriz, ajustada em cada instante, em função das medidas realizadas.

Neste trabalho, estuda-se o desenvolvimento de métodos recursivos de identificação de parâmetros, para modelos lineares multivariáveis, descritos em forma de equações à diferenças estocásticas. O principal aspecto considerado é o da determinação de algoritmos de estimação de parâmetros, passíveis de implementação em pequenos processadores, com o objetivo de integrarem controladores adaptativos junto ao processo.

A teoria de identificação de sistemas lineares dinâmicos pode ser hoje considerada como fechada para sistemas SISO (Single Input Single Output). O maior problema encontrado na extensão das técnicas de identificação de sistemas SISO para sistemas MIMO (Multiple Input Multiple Output) tem sido o aumento do tempo de cálculo envolvido, devido ao elevado número de parâmetros a serem estimados neste último caso.

Este aumento ocorre porque, como será visto nos capítulos seguintes, para se obter a matriz de ganho K_k , na equação do estimador recursivo (equação 1.1), é necessário obter uma matriz quadrada (P_k) de ordem proporcional ao número de parâmetros p a serem estimados, onde p é obtido através da equação 1.2.

$$p = 2 N_o \cdot n^2 + N_o mn \quad 1.2$$

sendo n o número de saídas do sistema, m o número de entradas do sistema e N_o a ordem do modelo considerado.

Para diminuir o tempo de cálculo destes algoritmos Morf et al. [4] e Ljung [5] desenvolveram métodos rápidos de cálculo ("fast-algorithms"), utilizando a propriedade de que as matrizes envolvidas no cálculo de K_k são meramente deslocadas quando se realiza uma nova medida. Entretanto, Robbins [6] mostrou que a técnica introduzida por Morf ocasiona uma sensível redução no tempo de computação somente quando a ordem do modelo é muito maior que o número de saídas do modelo. É comum verificar-se que, na área de controle de processos, os modelos normalmente utilizados são de ordem baixa, razão pela qual os algoritmos rápidos não representam uma real melhoria no tempo de cálculo dos mesmos.

Tendo em vista este fato procura-se, neste trabalho, reduzir o tempo de cálculo envolvido, através de métodos que permitam utilizar técnicas que explorem o paralelismo nos algoritmos analisados. Para tal, decompõe-se o problema de identificação em subproblemas que podem ser resolvidos paralelamente. Divide-se o vetor de parâmetros a estimar em vetores de menor dimensão, cada um contendo o conjunto de parâmetros de uma linha do modelo,

que são estimados separadamente. Mostra-se, neste trabalho, que essa separação depende essencialmente de três fatores:

- i - do algoritmo de identificação utilizado;
- ii - da matriz de covariância das componentes do ruído;
- iii - da estrutura do modelo.

A apresentação do trabalho está organizada da seguinte forma: neste primeiro capítulo resumem-se os principais resultados em modelagem de sistemas discretos estocásticos a serem utilizados.

O capítulo 2 apresenta um resumo dos algoritmos de estimação para sistemas multivariáveis existentes na literatura. O problema de identificação de parâmetros é reformulado, através do modelo do erro previsto introduzido por Ljung, que será utilizado no capítulo 3.

No capítulo 3 estuda-se o problema da decomposição do algoritmo de identificação. Mostra-se que a extensão do método dos mínimos quadrados (MQ) para sistemas MIMO, frequentemente utilizada na literatura [1,4,7,8,9,10,11,19], fornece diretamente uma decomposição em subproblemas. Esta partição continua válida, quando as componentes do vetor de perturbação são correlatas entre si, apesar de utilizar-se uma função de custo que envolve a matriz de correlação.

Quando o vetor de ruídos não é branco, são bem conhecidas as limitações do algoritmo anterior e torna-se necessária a utilização de outros métodos.

Mostra-se que o algoritmo da matriz estendida (ME) [1,11,12,13,14,15,16] permite decompor diretamente o problema de estimação em vários subproblemas e que esta decomposição independe da matriz de correlação.

O método dos mínimos quadrados generalizado (MQG) é deduzido de uma nova forma, diferente da apresentada por Payne [17] e Keviczky [10], que permite determinar as simplificações necessárias no modelo utilizado, de forma a obter-se uma estrutura de cálculo paralelo. Mostra-se, que se consegue este resulta

do através da forma pseudo-canônica I, sem necessidade de postular simplificações para o modelo proposto como em Morf [4].

A seguir apresenta-se o *algoritmo linearizado*, obtido a partir do método desenvolvido por Soderstrom [18] para sistemas SISO. O algoritmo é obtido a partir de uma linearização do erro previsto. Mostra-se que a forma canônica III de Kashyap e a forma pseudo-canônica I permitem decompor o problema num conjunto de subproblemas com um certo grau de paralelismo.

A seguir apresenta-se o *algoritmo particionado dos mínimos quadrados*, que consegue uma sensível diminuição no tempo de cálculo envolvido, utilizando conjuntamente a decomposição em subproblemas e uma nova partição da matriz de ganho K_k .

Todos os métodos propostos neste capítulo são desenvolvidos a partir de modelos autoregressivo e média-móvel (ARMA) para o ruído. É comum nos trabalhos nesta área utilizar-se modelos particulares (autoregressivo (AR) ou média-móvel (MA)) dependendo do algoritmo analisado. Como será mostrado na seção 1.2, um modelo ARMA representará em geral uma redução no número de parâmetros a serem estimados.

O capítulo 4 contém um estudo comparativo dos algoritmos desenvolvidos através de exemplos simulados. Analisa-se, basicamente, a evolução dos erros nos parâmetros estimados e o desempenho do modelo como previsor. A comparação é realizada através de simulações de Monte Carlo.

As necessidades, em tempo de cálculo e memória, para os diferentes algoritmos são comparadas. Utiliza-se, para tanto, o tempo de cálculo e memória de subrotinas escritas em linguagem de máquina para operações matriciais [58].

1.2. Modelagem de sistemas estocásticos lineares através de equações à diferenças

Coexistem, na literatura, métodos de identificação em que se utilizam modelos de estado, modelos de entrada/saída e

modelos em forma de matriz de resposta a um vetor de impulsos [20].

O modelo de espaço de estado é um sistema de equações às diferenças de 1ª ordem para o sistema. O modelo ARMA (modelos de entrada/saída) é uma equação às diferenças linear que relaciona a saída do sistema, num certo instante, com a entrada do sistema, neste instante, e nos instantes anteriores e também os valores da saída nos instantes anteriores.

Os diferentes usuários advocam vantagens e inconvenientes para cada uma destas representações e em geral o tipo de modelo é escolhido de forma a facilitar a implementação do algoritmo de controle que se pretende utilizar. Os tipos de modelos mais utilizados são os dois primeiros porque o último exige um número elevado de parâmetros e há necessidade de efetuar um truncamento.

A determinação da estrutura do modelo, no caso monovariável, limita-se à determinação da ordem do modelo e existem várias soluções para o problema:

- i) seja introduzindo a ordem do modelo utilizado no critério minimizado [21,1,22] ou
- ii) através de testes sobre a previsão de um passo [23], ou ainda testes de cancelamento de polos e zeros [23,24].

As técnicas de determinação da ordem do modelo tentam, de algum modo, reduzir ao mínimo o número de parâmetros desconhecidos, isto é, aplicam o princípio da "parsimonia" na construção do modelo.

Quando se analisam sistemas MIMO, a determinação da estrutura do modelo é mais complexa. Deseja-se, como no caso SISO, que a parametrização do modelo tenha as seguintes características:

- i) que seja universal, para aplicá-la a todos os sistemas da classe considerada;
- ii) que contenha o menor número possível de parâmetros;
- iii) que os parâmetros sejam identificáveis [26,27].

A determinação da estrutura do modelo está intimamente ligada com a obtenção de formas canônicas, isto é, a determinação de modelos com características especiais e que envolvam de alguma forma um número mínimo de parâmetros.

O cálculo de formas canônicas é um problema importante, exhaustivamente estudado nestes últimos anos [25,28,29,30]. O problema diminui de importância quando se considera a aplicação de técnicas de identificação em controle, onde não se pretende obter um modelo exato mas apenas uma aproximação, válida para os objetivos de controle, desde que se assegure a identificabilidade de todos os parâmetros [32].

Neste trabalho expressa-se a parte determinística das saídas em termos do passado das saídas e entradas, através do sistema de equações à diferenças

$$A(z^{-1})\bar{y}_k = B(z^{-1})u_k \quad 1.3$$

onde $\bar{y}_k \in R^n$ é a componente determinística do vetor de saída,

$u_k \in R^m$ o vetor de entrada e $A(z^{-1})$ e $B(z^{-1})$ são matrizes polinomiais definidas pelas equações

$$A(z^{-1}) = I + A_1 z^{-1} + \dots + A_{N_A} z^{-N_A}$$
$$B(z^{-1}) = B_1 z^{-1} + \dots + B_{N_B} z^{-N_B}$$

em que A_i, B_i são matrizes $(n \times n)$ e $(n \times m)$ respectivamente.

Como vários modelos, descrito pela equação 1.3, podem gerar a mesma matriz de transferência, procura-se trabalhar com

matrizes $A(z^{-1})$ e $B(z^{-1})$ que apresentem estrutura simplificada e sejam irredutíveis [33], isto é, tais que as matrizes $A(z^{-1})$ e $B(z^{-1})$ não contenham fatores comuns, a menos de matrizes unimodulares, o par de matrizes $\{A(z^{-1}), B(z^{-1})\}$ é definido como descrição por matriz partilhada (DMP) ("Matrix Fraction Description") [33].

Para se modelar a parte estocástica do processo utiliza-se o teorema da fatorização espectral.

Teorema 1 (Fatorização espectral) [34,35] Seja $\Gamma(Z)$, $n \times m$, uma matriz densidade espectral racional com característica n : então existe uma única matriz racional $H(Z)$, $n \times n$, e uma única matriz simétrica V , tais que :

- i) $\Gamma(Z) = H(Z) V^{-1} H^T(Z^{-1})$
- ii) $H(Z)$ é analítica fora do círculo unitário
- iii) $H^{-1}(Z)$ é analítica fora do círculo unitário
- iv) $\lim_{Z \rightarrow \infty} H(Z) = I$

A partir do teorema 1 pode-se representar um processo estocástico, com média zero e densidade espectral $\Gamma(Z)$ como a saída de um sistema linear, em que a entrada é uma sequência branca de variáveis aleatórias.

Das considerações anteriores resulta o seguinte modelo do sistema com perturbações estocásticas:

$$A(Z^{-1}) \underline{y}_k = B(Z^{-1}) \underline{u}_k + \underline{v}_k \tag{1.4}$$

$$\underline{v}_k = H(Z^{-1}) \underline{e}_k \tag{1.5}$$

em que $\{\underline{e}_k\}$ é uma sequência branca de variáveis aleatórias, com média zero e matriz de covariância V e $\underline{e}_k \in R^n$. Como no caso de determinístico utiliza-se a seguinte decomposição:

$$H(Z^{-1}) = D^{-1}(Z^{-1}) C(Z^{-1})$$

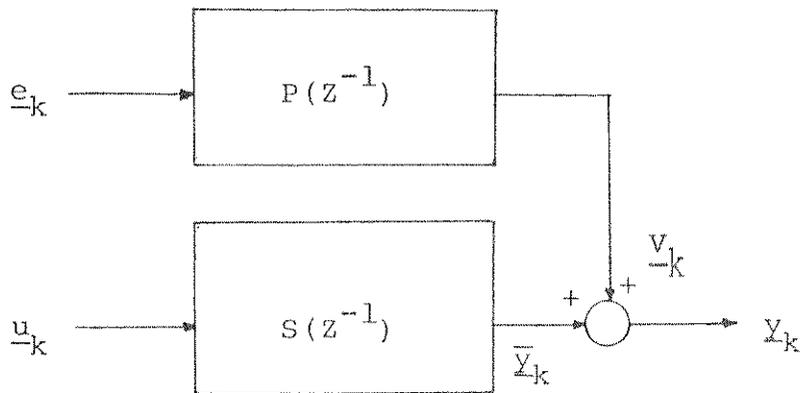
1.6

onde

$$C(Z^{-1}) = C_0 + C_1 Z^{-1} + \dots + C_{N_C} Z^{-N_C}, \quad C_i (n \times n)$$

$$D(Z^{-1}) = D_0 + D_1 Z^{-1} + \dots + D_{N_D} Z^{-N_D}, \quad D_i (n \times n)$$

A saída do modelo linear descrito pela equação 1.4, pode ser representada esquematicamente através da figura 1.2



$$P(Z^{-1}) = A^{-1}(Z^{-1}) D^{-1}(Z^{-1}) C(Z^{-1})$$

$$S(Z^{-1}) = A^{-1}(Z^{-1}) B(Z^{-1})$$

Figura 1.2 - Representação de um sistema linear estocástico através de relações entrada-saída

A matriz V de covariância do ruído branco será, no caso geral uma matriz cheia. No capítulo 3, consideram-se casos em que esta matriz é diagonal.

Das equações 1.3, 1.4, 1.5 e 1.6 obtem-se

$$A(Z^{-1}) y_k = B(Z^{-1}) u_k + D^{-1}(Z^{-1}) C(Z^{-1}) e_k \quad 1.7$$

ou

$$D(Z^{-1}) A(Z^{-1}) y_k = D(Z^{-1}) B(Z^{-1}) u_k + C(Z^{-1}) e_k \quad 1.8$$

finalmente

$$F(Z^{-1}) \underline{y}_k = G(Z^{-1}) \underline{u}_k + C(Z^{-1}) \underline{e}_k \quad 1.9$$

onde

$$F(Z^{-1}) = D(Z^{-1}) A(Z^{-1})$$

$$G(Z^{-1}) = D(Z^{-1}) B(Z^{-1})$$

A equação 1.9 mostra, em relação à equação 1.7, que um modelo linear do sistema estocástico, pode ser sempre colocado na forma de um modelo média-móvel para o ruído. Contudo a ordem das equações às diferenças em \underline{y}_k e \underline{u}_k do sistema é maior que no caso precedente.

1.3. Formas canônicas para modelos em forma de equações às diferenças [31]

O modelo geral obtido no item anterior e descrito pelas equações 1.7 ou 1.9 é obtido de forma a gerar um vetor de saídas \underline{y}_k com densidade espectral igual à densidade espectral da saída real. O teorema da representação espectral fornece, para tanto, uma matriz de transferência $H(Z^{-1})$ única mas, como foi visto anteriormente, a descrição por matriz partilhada não é única. Chega-se assim à conclusão que podem existir várias representações paramétricas de $C(Z^{-1})$ e $D(Z^{-1})$ fornecendo a mesma densidade espectral. O problema de definir-se uma única representação de $C(Z^{-1})$ e $D(Z^{-1})$ é o problema típico da definição de uma forma canônica.

Seja o modelo geral descrito pela equação 1.9, isto é:

$$F(Z^{-1}) \underline{y}_k = G(Z^{-1}) \underline{u}_k + C(Z^{-1}) \underline{e}_k \quad \text{e} \quad E\{\underline{e}_k \underline{e}_k^T\} = V$$

Seja P o conjunto formado por elementos H_i constituídos pelas quatro matrizes $F_i(Z^{-1})$, $G_i(Z^{-1})$, $C_i(Z^{-1})$ e V_i

$$P = \{ F_i(z^{-1}), G_i(z^{-1}), C_i(z^{-1}), V_i \}$$

Define-se uma relação de equivalência \textcircled{E} entre os elementos H_i de P por : $H_i \textcircled{E} H_j$ se H_i e H_j fornecem a mesma densidade espectral da saída S_{yy} e a mesma densidade espectral cruzada S_{yu} . A relação de equivalência divide P em classes de equivalência P_α para cada par $\{ S_{yy}^\alpha, S_{yu}^\alpha \}$ com elementos equivalentes. Prova-se [31] que, dentro de uma classe P_α de equivalência, um elemento pode ser obtido a partir de outro através de uma matriz unimodular U_α e uma matriz não singular R_α através de

$$F_2 = U_\alpha F_1, \quad C_2 = U_\alpha C_1 R_\alpha^{-1}, \quad G_2 = U_\alpha G_1, \quad V_2 = R_\alpha V_1 R_\alpha^T$$

A definição de uma forma canônica C_α^1 , pode ser vista como escolha de um único elemento dentro de cada classe P_α , que tenha uma característica especial T_β , geralmente a de ter uma estrutura mais simples. Os elementos para cada T_β constituirão o subconjunto canônico $C_\beta \subset P$. Obviamente não existirá apenas uma forma canônica dado que a sua definição depende do critério de simplicidade escolhido que será função da aplicação pretendida.

Exige-se, portanto, que qualquer mudança no valor dos parâmetros de uma forma canônica origine mudanças na densidade espectral ou matematicamente :

$$\text{Se } H_1 \text{ e } H_2 \text{ pertencem a } C_1 \text{ e } H_1 \textcircled{E} H_2 \text{ então } H_1 \equiv H_2$$

Também, normalmente, exige-se que dado qualquer elemento H de P , exista um único elemento H_1 em C_β tal que $H \textcircled{E} H_1$.

No problema de estimação, a primeira condição é indispensável, para que o algoritmo não tenha a possibilidade de oscilar entre duas soluções equivalentes. A segunda condição é necessária para garantir que a forma canônica utilizada seja geral, isto é, possa gerar qualquer tipo de densidade espectral

Obviamente, se for fornecida *ã priori* a estrutura do sistema que gera a densidade espectral, a estimação dos parâme

metros pode ser realizada utilizando-se esta informação e neste caso a segunda condição não precisa ser satisfeita. Neste caso, a estrutura define-se como uma forma pseudo-canônica [18].

Entre as formas canônicas existentes na literatura a apresentam se a seguir as formas canônicas I, II e III de Kashyap [36]. A última será utilizada no capítulo 3.

Forma Canônica I [37] : caracterizada por $F(Z^{-1})$ ser triangular, com $F(0) = I$ e $G(Z^{-1})$, $C(Z^{-1})$ e V arbitrários. Kashyap mostra que a vantagem desta forma canônica é de apresentar um número mínimo de parâmetros.

Forma Canônica II : caracterizada por $C(Z^{-1})$ ser diagonal, $F(0) = I$ e $F(Z^{-1})$, $G(Z^{-1})$, V arbitrários.

Forma Canônica III : caracterizada por $C(Z^{-1})$ e V serem diagonais e $F(Z^{-1})$ e $G(Z^{-1})$ arbitrários com exceção de $F(0)$ que é uma matriz triangular com elementos unitários na diagonal.

Com a utilização da forma canônica III, cada componente do ruído atua independentemente em cada saída do modelo, permitindo, como será mostrado no capítulo 3, a partição do problema de estimação de parâmetros num sistema MIMO, em n subproblemas distintos e paralelos. Também, tendo em vista a partição do estimador em n subproblemas de dimensão menor, utilizam-se no capítulo 3 as seguintes formas pseudo-canônicas :

Forma pseudo-canônica I : caracterizada por $C(Z^{-1})$, $D(Z^{-1})$ e V diagonais e $F(Z^{-1})$ e $G(Z^{-1})$ arbitrários, com $F(0) = G(0) = C(0) = D(0) = I$

Forma pseudo-canônica II: caracterizada por $F(Z^{-1})$, $G(Z^{-1})$ e V quaisquer com $F(0) = G(0) = I$ e :

$$C(Z^{-1}) = c(Z^{-1}) \cdot I$$

$$D(Z^{-1}) = d(Z^{-1}) \cdot I$$

com

$$c(Z^{-1}) = 1 + c_1 z^{-1} + \dots + c_{N_C} z^{-N_C} \quad e$$

⋮

$$d(z^{-1}) = 1 + d_1 z^{-1} + \dots + d_{N_D} z^{-N_D}$$

1.4. Conclusões

As formas canônicas e pseudo-canônicas apresentadas se
rão utilizadas no capítulo 3, tendo em vista a estimação parti
cionada dos parâmetros do modelo. Algoritmos para a obtenção das
referidas formas canônicas são descritos por Kashyap em [36].

CAPÍTULO 2

2. ALGORITMOS PARA ESTIMAÇÃO DE PARÂMETROS EM SISTEMAS MULTIVARIÁVEIS DESCRITOS POR EQUAÇÕES À DIFERENÇAS

2.1. Introdução

Neste capítulo apresenta-se um resumo dos algoritmos de identificação para sistemas multivariáveis existentes na literatura. Trata-se especificamente dos algoritmos dos mínimos quadrados, mínimos quadrados generalizado, matriz estendida e máxima verossimilhança.

Todos os algoritmos têm sua origem nos correspondentes algoritmos para sistemas SISO. A extensão ao caso multivariável é imediata em alguns casos (mínimos quadrados e matriz estendida) enquanto que em outros, apresenta dificuldades específicas que serão comentadas no decorrer da sua apresentação.

Os algoritmos estudados neste capítulo têm aparecido recentemente na literatura, à exceção do algoritmo dos mínimos quadrados. As principais referências estão contidas nos últimos congressos da IFAC em Identificação de Sistemas e Estimação de Parâmetros (Tbilisi (1976), Darmstadt (1979)) e nos últimos congressos do IEEE em Controle e Decisão.

A extensão direta ao caso multivariável original, geralmente, algoritmos com tempos de cálculo elevados que dificultam ou impossibilitam a sua utilização em controle "on-line". No capítulo 3 apresentam-se as modificações propostas para se diminuir o tempo de cálculo envolvido.

Encerra-se este capítulo com a formulação do problema de estimação através do modelo do erro previsto apresentado por Ljung [20]. Esta formulação permite unificar os diferentes modelos apresentados e será utilizada para se justificar as modificações propostas no capítulo 3.

2.2. Obtenção de um modelo de regressão linear

Os elementos das matrizes do modelo utilizado são os parâmetros desconhecidos que se desejam estimar. Para facilidade de obtenção do estimador de Markov, reescreve-se a equação 1.4 de modo que os elementos desconhecidos estejam contidos num vetor, isto é :

$$Y_k = -A_1 Y_{k-1} - \dots - A_{N_A} Y_{k-N_A} + B_1 u_{k-1} + \dots + B_{N_B} u_{k-N_B} + v_k$$

2.1

$$Y_k = \begin{bmatrix} -Y_{k-1}^T \dots -Y_{k-N_A}^T & u_{k-1}^T \dots u_{k-N_B}^T & \vdots & 0 & 0 & \dots & 0 & \vdots & \dots & \vdots \\ \vdots & \vdots \\ 0 & 0 & \dots & 0 & \vdots & -Y_{k-1}^T & \dots & u_{k-N_B}^T & \vdots & \dots \\ \vdots & \vdots \\ 0 & 0 & \dots & 0 & \vdots & 0 & \dots & 0 & \vdots & \dots \\ \vdots & \vdots \\ \vdots & 0 & \dots & 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots \\ \vdots & Y_{k-1}^T & \dots & u_{k-N_B}^T & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \cdot \begin{bmatrix} a^1 \\ \vdots \\ a^n \\ b^1 \\ \vdots \\ b^n \end{bmatrix} + v_k$$

$$Y_k = \psi_k \theta^{A,B} + v_k \quad Y_k, v_k \in R^n \quad 2.2$$

$\theta^{A,B} \in R^{np}$ onde p é dado por $p = (nN_A + mN_B)$ e ψ_k tem dimensão $n \times np$

$$\theta^{A,B} = \begin{bmatrix} a^1 T, \dots, a^n T & b^1 T & \dots & b^n T \end{bmatrix}$$

$a^i T$ - iésima linha de $A(z^{-1})$

$b^i T$ - iésima linha de $B(z^{-1})$

$$\psi_k = \text{Bloco diagonal} \begin{bmatrix} \phi_k^T & \dots & \phi_k^T \end{bmatrix}$$

$$\phi_k^T = \begin{bmatrix} -Y_{k-1}^T \dots u_{k-N_B}^T \end{bmatrix} \quad \phi_k \in R^p$$

Obtem-se outra forma de representação do modelo descrito pela equação 1.4, colocando-se os parâmetros desconhecidos numa matriz, isto é :

$$Y_k = -A_1 Y_{k-1} - \dots - A_{N_A} Y_{k-N_A} + B_1 u_{k-1} + \dots + B_{N_B} u_{k-N_B} + v_k \quad 2.3$$

$$Y_k = \theta^{T_{A,B}} \phi_k + v_k \quad 2.4$$

onde $\theta^{T_{A,B}}$ tem dimensão $n \times p$

$$\text{com } \theta^{T_{A,B}} = \begin{bmatrix} A_1, \dots, A_{N_A}, B_1, \dots, B_{N_B} \end{bmatrix}$$

Esta forma será utilizada, no item 2.4, na obtenção do estimador dos mínimos quadrados.

2.3. Estimador de Markov

Seja, inicialmente, um sistema de equações lineares relacionando o vetor de saídas Y_k com um vetor de parâmetros $\theta^{A,B}$ descrito pela equação 2.1, reescrita abaixo

$$Y_k = \psi_k \theta^{A,B} + v_k \quad 2.1$$

onde $\{ \psi_t, t = 1, 2, \dots, k \}$ é uma matriz conhecida de dimensão $n \times np$ e v_k é o vetor de ruído com uma matriz de covariância R_{ts}

onde

$$R_{ts} = \begin{bmatrix} E \{ v_t v_s^T \} \end{bmatrix} \quad 2.5$$

Após k medidas realizadas, pode-se escrever a equação 2.1 como :

$$\underline{Y} = \psi \underline{\theta}^{A,B} + \underline{V} \quad 2.6$$

Sendo $\underline{Y}^T = \left[\underline{Y}_1^T, \underline{Y}_2^T, \dots, \underline{Y}_k^T \right] \quad \underline{Y} \text{ (nk x 1)}$

$$\psi^T = \left[\psi_1^T : \psi_2^T : \dots : \psi_k^T \right] \quad \psi^T \text{ (np x nk)}$$

$$\underline{V}^T = \left[\underline{v}_1^T, \underline{v}_2^T, \dots, \underline{v}_k^T \right] \quad \underline{V} \text{ (nk x 1)}$$

O estimador de Markov é obtido minimizando-se

$$J = \left\| \underline{Y} - \psi \underline{\theta}^{A,B} \right\|_{R^{-1}}^2 \quad 2.7$$

onde $R = \text{cov}(V)$

A sua dedução é direta, pois o critério J é uma função quadrática nos parâmetros desconhecidos, resultando na equação

$$\hat{\underline{\theta}}^{A,B} = \left[\psi^T R^{-1} \psi \right]^{-1} \left[\psi^T R^{-1} \underline{Y} \right] \quad 2.8$$

conhecida como a solução de Markov.

São bem conhecidas as propriedades deste estimador como sendo não polarizado. Como a matriz ψ é de ordem $nk \times np$, em geral muito elevada, o tempo de cálculo envolvido no estimador de Markov pode ser excessivo. Aplicando-se os modelos apresentados no capítulo 1, estudam-se, nos itens 2.5 e 2.6, alternativas simplificadoras para o algoritmo de Markov. Nestes algoritmos, é mostrado que devido, à modelagem do ruído "colorido" a partir do ruído branco, não é necessário o conhecimento total da matriz de covariância R .

Supondo o ruído \underline{v}_k branco, a matriz de covariância R é descrita por

$$R = \text{bloco diagonal} \left[R_1 : R_2 : \dots : R_k \right] \quad 2.9$$

Com esta hipótese simplificadora resulta a seguinte equação para o estimador

$$\hat{\theta}_{A,B} = \left[\sum_{t=1}^k \psi_t^T R_t^{-1} \psi_t \right]^{-1} \left[\sum_{t=1}^k \psi_t^T R_t^{-1} y_t \right] \quad 2.10$$

onde, agora, pode-se observar que a dimensão das matrizes ψ_k é $n \times n_p$, o que reduz consideravelmente o volume computacional do algoritmo.

Supondo o ruído v_k branco, é possível obter-se o estimador de Markov através de um processo recursivo, onde a cada nova medida se atualiza o estimador calculado no instante anterior.

A transformação do problema de estimação num procedimento recursivo tem as seguintes vantagens :

- i) Não é necessária a inversão de uma matriz de ordem elevada;
- ii) O efeito de cada nova observação é facilmente incorporado no estimador, a cada passo.

Para se obter a versão recursiva do estimador de Markov define-se :

$$P_{k+1} = \left[\sum_{t=1}^k \psi_t^T R_t^{-1} \psi_t \right]^{-1} = \left[\sum_{t=1}^k \psi_t^T R_t^{-1} \psi_t + \psi_{k+1}^T R_{k+1}^{-1} \psi_{k+1} \right]^{-1} \quad 2.11$$

A partir da equação 2.11 e utilizando a identidade 2.12, descrita abaixo

$$\left[A + B^T C B \right]^{-1} = A^{-1} - A^{-1} B \left[C^{-1} + B A^{-1} B^T \right]^{-1} B A^{-1} \quad 2.12$$

obtem-se

$$P_{k+1} = P_k - P_k \psi_{k+1}^T \left[R_{k+1} + \psi_{k+1} P_k \psi_{k+1}^T \right]^{-1} \psi_{k+1} P_k \quad 2.13$$

Utilizando a equação 2.14 na equação do estimador, resulta :

$$\hat{\theta}_{k+1}^{A,B} = \left\{ P_k - P_k \psi_{k+1}^T \left[R_{k+1} + \psi_{k+1} P_k \psi_{k+1}^T \right]^{-1} \psi_{k+1} P_k \right\} \cdot \left[\sum_{t=1}^k \psi_t^T R_t^{-1} y_t + \psi_{k+1}^T y_{k+1} \right] \quad 2.14$$

Definindo $K_{k+1} = P_k \psi_{k+1}^T \left[R_{k+1} + \psi_{k+1} P_k \psi_{k+1}^T \right]^{-1}$ 2.15

e utilizando a expressão do estimador $\hat{\theta}_k^{A,B}$, resulta

$$\hat{\theta}_{k+1}^{A,B} = \hat{\theta}_k^{A,B} + K_{k+1} \left\{ y_{k+1} - \psi_{k+1} \hat{\theta}_k^{A,B} \right\} \quad 2.16$$

A equação 2.16 representa a versão recursiva do estimador de Markov.

Considerando-se o ruído v_k como sendo estacionário e branco a matriz de covariância R_k será constante. Este caso corresponde ao algoritmo dos mínimos quadrados que será tratado na seção seguinte.

2.4. Algoritmo dos mínimos quadrados [1,2,3,4,5,6,7]

O algoritmo é obtido na literatura utilizando o modelo descrito pela equação 2.3, repetida a seguir :

$$y_k = \theta^{A,B} \phi_k + v_k \quad 2.17$$

com $\theta^{A,B}$, ϕ_k como definidos no item 2.2.

É bem conhecido na literatura que o estimador dos mínimos quadrados é um estimador consistente dos parâmetros do processo, somente se a sequência $\{v_k\}$ for branca. [1,11,14,16,36].

Supondo que o ruído é branco, o estimador dos parâmetros $A_i B_i$ é obtido, na literatura, de modo a minimizar uma função de custo quadrática do tipo dado pela equação 2.7 com $R=I$.

$$\min_{\hat{\theta}^{A,B}} \left\{ \sum_{t=1}^k \| \underline{y}_t - \hat{\theta}^{A,B} \underline{\phi}_t \|^2 \right\} \quad 2.18$$

Como a equação 2.18 é uma função quadrática em $\hat{\theta}$, facilmente se obtém o estimador dos parâmetros desconhecidos após k medidas. Ele será descrito por:

$$\hat{\theta}_k^{A,B} = \left[\sum_{t=1}^k \underline{\phi}_t \underline{\phi}_t^T \right]^{-1} \left[\sum_{t=1}^k \underline{\phi}_t \underline{y}_t^T \right] \quad 2.19$$

Utilizando a identidade 2.12 pode-se, como no item anterior, colocar este resultado na forma recursiva resultando :

$$\hat{\theta}_{k+1}^{A,B} = \hat{\theta}_k^{A,B} + K_{k+1} \left\{ \begin{array}{c} \underline{y}_{k+1}^T - \underline{\phi}_{k+1}^T \hat{\theta}_k^{A,B} \end{array} \right\}$$

$$K_{k+1} = P_k \underline{\phi}_{k+1} / \left(1 + \underline{\phi}_{k+1}^T P_k \underline{\phi}_{k+1} \right) \quad 2.20$$

$$P_{k+1} = P_k - P_k \underline{\phi}_{k+1} \left(1 + \underline{\phi}_{k+1}^T P_k \underline{\phi}_{k+1} \right)^{-1} \underline{\phi}_{k+1}^T P_k$$

O estimador dos mínimos quadrados exige menos esforço computacional que o estimador de Markov, porque, no critério utilizado neste item a matriz de ponderação é igual à identidade, o que permite na equação 2.20 estimar cada linha do modelo separadamente (colunas da matriz $\hat{\theta}^{AB}$). No capítulo seguinte mostra-se que, mesmo com uma matriz de ponderação diferente da identidade, é possível estimar-se cada linha do modelo separadamente.

Entretanto muitas vezes este algoritmo não pode ser utilizado, porque os valores estimados são polarizados. Isto é devido ao fato do processo não poder ser modelado através de um ruído branco. Os algoritmos que se descrevem a seguir resolvem este problema através da modelagem do ruído e consequente estimação dos parâmetros. Os modelos utilizados são os descritos no capítulo 1.

2.5. Algoritmo dos mínimos quadrados generalizado

O algoritmo dos mínimos quadrados generalizado foi apresentado inicialmente por Clarke [41] para sistemas SISO. Uma versão recursiva aproximada foi posteriormente introduzida por Hasting-James [43]. A extensão ao caso MIMO foi concomitantemente realizada por Payne [17] e Keviczky [10].

Este método postula um modelo autoregressivo para o ruído \underline{v}_k , como indicado na equação 2.21

$$D(Z^{-1}) \underline{v}_k = \underline{e}_k \quad 2.21$$

onde $\{\underline{e}_k\}$ é um ruído branco com matriz de covariância V .

Com a utilização deste modelo não há necessidade do conhecimento total da matriz de covariância R do ruído \underline{v}_k , porque esta informação estará contida na matriz de transferência $D^{-1}(Z^{-1})$ e na matriz de covariância V do ruído branco. Este modelo permitirá obter um estimador com um tempo de cálculo menor do que o estimador de Markov descrito pela equação 2.8.

O modelo mostra, claramente, quais serão as limitações do algoritmo. Dado que a matriz de transferência que gera o ruído \underline{v}_k a partir do ruído branco \underline{e}_k conterà apenas polos, a ordem do polinômio $D(Z^{-1})$ poderá vir a ser infinita. Para fins práticos o polinômio será truncado e o seu grau será escolhido de forma a atingir a precisão desejada para o modelo.

Introduzindo o modelo do ruído (equação 2.21) no sistema descrito pela equação 2.1, obtém-se

$$D(Z^{-1}) A(Z^{-1}) \underline{y}_k = D(Z^{-1}) B(Z^{-1}) \underline{u}_k + \underline{e}_k \quad 2.22$$

No caso SISO o produto $D(Z^{-1}) A(Z^{-1})$ é comutativo e o método consiste essencialmente na estimação separada dos parâmetros em $A(Z^{-1})$ e $B(Z^{-1})$ e dos parâmetros em $D(Z^{-1})$. Para tanto utiliza-se um método de relaxação. Supõe-se disponível uma estimação $\hat{D}(Z^{-1})$ de $D(Z^{-1})$ e utiliza-se o algoritmo dos mínimos quadrados para se estimar $A(Z^{-1})$ e $B(Z^{-1})$ no modelo abaixo

$$A(Z^{-1}) \underline{y}_k^F = B(Z^{-1}) \underline{u}_k^F + \underline{e}_k^F \quad 2.23$$

onde $\underline{y}_k^F \triangleq \hat{D}(Z^{-1}) \underline{y}_k$ e $\underline{u}_k^F \triangleq \hat{D}(Z^{-1}) \underline{u}_k$

Com os valores estimados $\hat{A}(Z^{-1})$ e $\hat{B}(Z^{-1})$ obtem-se uma estimaco de \underline{v}_k , atravs de

$$\hat{\underline{v}}_k = \hat{A}(Z^{-1}) \underline{y}_k - \hat{B}(Z^{-1}) \underline{u}_k \quad 2.24$$

que possibilita a obteno de uma nova estimaco de $D(Z^{-1})$, atravs de um algoritmo dos mnimos quadrados, utilizando o modelo descrito pela equaco 2.21.

Portanto este algoritmo resulta num processo iterativo. Mostra-se [42] que tal processo converge para o valor verdadeiro dos parmetros desconhecidos se a relao sinal rudo  suficientemente grande.

A verso recursiva proposta por Hasting-James [43] simplesmente elimina este processo iterativo.

Dada a utilizao do algoritmo dos mnimos quadrados nota-se que o algoritmo obtido minimiza uma funo de custo do tipo indicado pela equaco 2.18 sem levar em conta a matriz de covarincia de \underline{e}_k .

No caso MIMO, o produto de matrizes no  comutativo e no se pode aplicar, imediatamente o mesmo procedimento do caso SIS0.

Keviczky [10] e Payne [17,34] contornaram esse problema utilizando as definies de produto de Kronecker [44,45] e vec de uma matriz [44,45].

Definio 1 : Dada uma matriz $A = [a_{ij}]$ (mxn) e uma matriz

B (sxt) ento o produto de Kronecker $A \times B$  definido como uma matriz (ms x nt), com

$$A \textcircled{x} B = [a_{ij}^B]$$

Definição 2 : Dada uma matriz $A = [a_{ij}]$ ($m \times n$), onde $a_{.j}$ denota a j-ésima coluna de A define-se o vetor coluna $\text{vec} A$ (mn) como

$$\text{vec } A = \begin{bmatrix} a_{.1} \\ \vdots \\ a_{.n} \end{bmatrix}$$

Com as definições acima pode-se reescrever a equação 2.22 e obtém-se

$$Y_k^F + \left[Y_k^T \textcircled{x} D(Z^{-1}) \right] \text{vec } A'(Z^{-1}) = \left[u_k^T \textcircled{x} D(Z^{-1}) \right] \text{vec } B(Z^{-1}) + e_{-k}$$

2.25

com $A'(Z^{-1}) = A(Z^{-1}) - I$, e $Y_k^F = D(Z^{-1}) Y_k$

Definindo $\underline{\theta}^{A,B} = \left[\text{vec}^T A'(Z^{-1}), \text{vec}^T B(Z^{-1}) \right]$ e

$$\psi_k^F = \left[Y_k^T \textcircled{x} D(Z^{-1}), u_k^T \textcircled{x} D(Z^{-1}) \right]$$

pode-se reescrever a equação 2.25 na seguinte forma

$$Y_k^F = \psi_k^F \underline{\theta}^{A,B} + e_{-k} \tag{2.26}$$

Esta equação é semelhante à sua equivalente, no caso SISO, (equação 2.19), existindo somente uma diferença básica na construção do filtro "embranquecedor".

Para se obter o estimador dos parâmetros do processo $(\underline{\theta}^{A,B})$, aplica-se o algoritmo dos mínimos quadrados no modelo descrito pela equação 2.26 resultando

$$\hat{\underline{\theta}}_{-k}^{A,B} = \left[\sum_{t=1}^k \psi_t^{FT} \psi_t^F \right]^{-1} \left[\sum_{t=1}^k \psi_t^{FT} Y_t^F \right] \tag{2.27}$$

Como no caso monovariável, os valores estimados de $A(Z^{-1})$ e $B(Z^{-1})$ permitem, agora, obter um estimador para a matriz $D(Z^{-1})$

Das equações 2.21 e 2.24 obtém-se :

$$D(Z^{-1}) \hat{v}_{-k} = e_{-k}$$

com

$$\hat{v}_{-k} = \hat{A}(Z^{-1}) y_k - \hat{B}(Z^{-1}) u_k \quad 2.28$$

Utilizando a definição da matriz polinomial $D(Z^{-1})$ pode-se desenvolver a equação 2.21 e obtém-se

$$\hat{v}_{-k} = -D_1 \hat{v}_{-k-1} - \dots - D_{N_D} \hat{v}_{-k} - N_D^+ e_{-k} \quad 2.29$$

$$e \quad \hat{v}_{-k} = \begin{bmatrix} D_1 & \dots & D_{N_D} \end{bmatrix} \begin{bmatrix} -\hat{v}_{-k-1} \\ \vdots \\ -\hat{v}_{-k-N_D} \end{bmatrix} + e_{-k} \quad 2.30$$

Definindo $\hat{\theta}^{D^T} = \begin{bmatrix} D_1 & \dots & D_{N_D} \end{bmatrix}$ e

$$m_{-k}^T = \begin{bmatrix} -\hat{v}_{-k-1}^T & \dots & -\hat{v}_{-k-N_D}^T \end{bmatrix}$$

a equação (2.30) resulta em

$$\hat{v}_{-k} = \hat{\theta}^{D^T} m_{-k}^T + e_{-k} \quad 2.31$$

Aplicando o estimador dos mínimos quadrados obtém-se o estimador dos parâmetros do ruído dado por :

$$\hat{\theta}_k^D = \left[\sum_{t=1}^k m_{-t} m_{-t}^T \right]^{-1} \left[\sum_{t=1}^k m_{-t} v_{-t}^T \right] \quad 2.32$$

Com o valor estimado $\hat{\theta}_k^D$, continua-se o processo iterativo calculando-se novos valores filtrados e um novo estimador dos parâmetros do processo.

Como mostrado por Ljung [12] no caso SISO, a extensão para o caso MIMO também pode apresentar múltiplos pontos de mínimos, se a relação sinal-ruído não for suficientemente elevada.

Obtem-se a versão recursiva do estimador dos mínimos quadrados generalizado no caso MIMO utilizando-se a idéia de Hasting-James, que consiste basicamente na aplicação da identidade de 2.12 nas expressões de $\hat{\theta}_k^{A,B}$ e $\hat{\theta}_k^D$ e a eliminação do processo iterativo. Esta idéia será analisada, mais detalhadamente, no capítulo seguinte quando serão estudadas simplificações para se obter algoritmos com um menor tempo de cálculo.

2.6. Algoritmo da Matriz Estendida

No ítem anterior, estudou-se um algoritmo que tinha como hipótese básica a de que o ruído podia ser modelado através de um processo autoregressivo.

A seguir considera-se um algoritmo que postula um modelo média-móvel para o ruído v_k , como indicado pela equação 2.33

$$v_k = C(z^{-1}) e_k \tag{2.33}$$

com $\{e_k\}$ um ruído branco.

O algoritmo que utiliza este modelo é normalmente conhecido como matriz estendida porque se obtém o novo vetor ϕ_k como uma extensão do vetor ϕ_k do estimador dos mínimos quadrados. Este algoritmo foi originalmente sugerido para sistemas SISO por Panuska [46] e Young [47]. Este método também foi analisado por Aström e Eykoff [11] e Soderstrom et al. [12]. A extensão ao caso MIMO foi realizada por Morf [4], estimando separadamente cada linha do modelo e não considerando a matriz de correlação. Para tanto ele reescreveu as equações do modelo, utilizando a hipótese do ruído ser modelado por um processo média-móvel.

$$A(z^{-1}) y_k = B(z^{-1}) u_k + C(z^{-1}) e_k \tag{2.34}$$

$$\underline{y}_k = \begin{bmatrix} A_1, A_2, \dots, A_{N_A}, B_1, \dots, B_{N_B}, C_1, \dots, C_{N_C} \end{bmatrix} \begin{bmatrix} - \underline{y}_{k-1} \\ \vdots \\ - \underline{y}_{k-N_A} \\ \underline{u}_{k-1} \\ \vdots \\ \underline{u}_{k-N_B} \\ \underline{e}_{k-1} \\ \vdots \\ \underline{e}_{k-N_C} \end{bmatrix} + \underline{e}_k \quad 2.35$$

Definindo

$$\theta^{A,B,C} = \begin{bmatrix} A_1 \dots A_{N_A}, B_1 \dots B_{N_B}, C_1 \dots C_{N_C} \end{bmatrix} \quad 2.36a$$

$$\underline{\phi}_k^T = \begin{bmatrix} - \underline{y}_{k-1}^T \dots - \underline{y}_{k-N_A}^T, \underline{u}_{k-1}^T \dots \underline{u}_{k-N_B}^T, \underline{e}_{k-1}^T \dots \underline{e}_{k-N_C}^T \end{bmatrix} \quad 2.36b$$

a equação 2.35 do sistema, toma a seguinte forma:

$$\underline{y}_k = \theta^{A,B,C} \underline{\phi}_k + \underline{e}_k \quad 2.37$$

Assumindo, inicialmente, que a sequência de ruído $\{\underline{e}_k\}$ é conhecida (mensurável), vê-se que a equação 2.37 é formalmente equivalente à equação 2.17, e os parâmetros desconhecidos $(\theta^{A,B,C})$ são obtidos resolvendo-se o problema do estimador dos mínimos quadrados com duas entradas, \underline{u}_k e \underline{e}_k . Logo o estimador matriz estendida resultante é equivalente ao descrito pela equação 2.19 com o novo $\underline{\phi}_k$ (estendido) definido acima, isto é :

$$\hat{\theta}_k^{A,B,C} = \left[\begin{array}{cc} k & \\ \sum_{t=1}^k \underline{\phi}_t \underline{\phi}_t^T & \end{array} \right]^{-1} \left[\begin{array}{c} k \\ \sum_{t=1}^k \underline{\phi}_t \underline{y}_t^T \end{array} \right] \quad 2.38$$

Como algumas componentes do ruído $\{\underline{e}_{k-1}, \dots, \underline{e}_{k-N_C}\}$ entram como elementos do vetor $\underline{\phi}_k$, (equação 2.36), e visto que esses valores são desconhecidos, torna-se necessário o seu cálculo aproximado. Para tal o algoritmo é iniciado com o cálculo de um valor de mínimos quadrados de $\hat{\theta}_{-M}^{A,B}$ considerando-se \underline{e}_k como ruído branco.

Com esses valores estimados calcula-se uma primeira aproximação de \underline{e}_k dado por

$$\hat{\underline{e}}_t = \underline{y}_t - \hat{\theta}_{MQ}^{T A, B} \underline{\phi}_t^{MQ} \quad t = 1, 2, \dots, k \quad 2.39a$$

com

$$\underline{\phi}_k^{MQ} = \left[-\underline{y}_{k-1}, \dots, -\underline{y}_{k-N_A}, \underline{u}_{k-1}, \dots, \underline{u}_{k-N_B} \right]$$

Os valores de $\hat{\underline{e}}_k$ são introduzidos na equação 2.36b de $\underline{\phi}_k$. Deste modo obtem-se o primeiro valor estimado de $\hat{\theta}^{A, B, C}$ através da equação 2.38.

Com os valores estimados da matriz $\hat{\theta}^{A, B, C}$ calcula-se uma nova aproximação de \underline{e}_k com

$$\hat{\underline{e}}_t = \underline{y}_t - \hat{\theta}^{T A, B, C} \underline{\phi}_t \quad t = 1, 2, \dots, k \quad 2.39b$$

O procedimento iterativo continua, levando-se os novos valores estimados de $\hat{\underline{e}}_k$ no vetor $\underline{\phi}_k$ e obtendo-se um novo valor estimado de $\theta^{A, B, C}$ através da equação 2.38. O processo iterativo cessa quando os valores absolutos das diferenças dos estimadores de $\theta^{A, B, C}$, em sucessivas iterações, são menores do que um certo valor especificado.

Morf [4] também apresentou uma versão recursiva deste estimador, descrita pela equação 2.40 que é obtida aplicando-se a identidade 2.12 na equação 2.38.

$$\begin{aligned} \hat{\theta}_{k+1}^{A, B, C} &= \hat{\theta}_k^{A, B, C} + K_{k+1} \left\{ \underline{y}_{k+1}^T - \underline{\phi}_{k+1}^T \hat{\theta}_k^{A, B, C} \right\} \\ K_{k+1} &= P_k \underline{\phi}_{k+1} \left\{ 1 + \underline{\phi}_{k+1}^T P_k \underline{\phi}_{k+1} \right\}^{-1} \\ P_{k+1} &= P_k - P_k \underline{\phi}_{k+1} \left[1 + \underline{\phi}_{k+1}^T P_k \underline{\phi}_{k+1} \right]^{-1} \underline{\phi}_{k+1}^T P_k \end{aligned} \quad 2.40$$

em que agora o estimador do ruído é dado por

$$\hat{\underline{e}}_k = \underline{y}_k - \hat{\theta}_{k+1}^{T A, B, C} \underline{\phi}_k$$

Para se obter uma maior precisão nos valores estimados \hat{e}_k , normalmente, itera-se uma vez a cada nova medida realizada.

2.7. Estimador de máxima verossimilhança

O algoritmo de máxima verossimilhança normalmente utiliza para o ruído um modelo média-móvel, como descrito no item anterior, Este estimador foi apresentado inicialmente por Astrom-Bohlin [48] para sistemas SISO.

Supondo que $\{e_k\}$ é uma sequência de variáveis aleatórias gaussianas estacionária e com matriz de covariância V então o logaritmo da função verossimilhança é dada por :

$$L = -\frac{mn}{2} \ln 2\pi - \frac{K}{2} \ln \det V - \frac{1}{2} \sum_{t=1}^k e_t^T V^{-1} e_t \quad 2.41$$

Obtem-se o estimador de máxima verossimilhança maximizando-se a função L . Supondo-se que a matriz de covariância é conhecida, a maximização do logaritmo da função de verossimilhança é equivalente à minimização de $J_1(\underline{\theta}^{A,B,C})$, onde

$$\begin{aligned} J_1(\underline{\theta}^{A,B,C}) &= \frac{1}{2} \sum_{t=1}^k e_t^T V^{-1} e_t \\ &= \frac{1}{2} \text{traço } V^{-1} \sum_{t=1}^k e_t e_t^T \end{aligned} \quad 2.42$$

e $\underline{\theta}^{A,B,C}$ é um vetor contendo todos os parâmetros desconhecidos. Os valores de e_t são substituídos por valores estimados recursivamente a partir das medidas realizadas e do vetor de parâmetros $\underline{\theta}^{A,B,C}$ estimado.

$$\hat{e}_t = \hat{A}(z^{-1}) y_t - \hat{B}(z^{-1}) u_t - \hat{C}(z^{-1}) \hat{e}_t \quad 2.43$$

$$t = 1, 2, \dots, k$$

Como o critério $J_1(\underline{\theta}^{A,B,C})$ é uma função não quadrática nos parâmetros desconhecidos, a minimização não pode ser realizada de uma forma fechada. Realiza-se a minimização através de algum procedimento iterativo de procura do mínimo.

O algoritmo mais citado [10,31,34] utiliza o cálculo :
 i) do gradiente da função de custo $J_1(\underline{\theta})$, e
 ii) da matriz hessiana das derivadas de segunda ordem do critério $J_1(\underline{\theta})$.

Neste caso o gradiente da função de custo é dado por,

$$J'(\underline{\theta}^{A,B,C}) = \frac{\partial J(\underline{\theta}^{A,B,C})}{\partial \underline{\theta}} = \sum_{t=1}^k \frac{\partial e_t^T}{\partial \underline{\theta}} V^{-1} e_t \quad 2.44$$

Os elementos da matriz derivada de segunda ordem são:

$$J''(\underline{\theta}^{A,B,C}) = \frac{\partial^2 J(\underline{\theta})}{\partial \theta_i \partial \theta_j} = \sum_{t=1}^k \frac{\partial e_t^T}{\partial \theta_i} V^{-1} \frac{\partial e_t}{\partial \theta_j} + \sum_{t=1}^k e_t^T V^{-1} \frac{\partial^2 e_t}{\partial \theta_i \partial \theta_j} \quad 2.45$$

Através destas relações Keviczky [10], Payne [34], Astrom [48,50], Saridis [49] construíram um procedimento iterativo utilizando um algoritmo do tipo de Newton-Raphson.

$$\hat{\theta}_{i+1}^{A,B,C} = \hat{\theta}_i^{A,B,C} + \left[\frac{\partial^2 J(\hat{\theta}_i)}{\partial \hat{\theta}_i \partial \hat{\theta}_i^T} \right]^{-1} \cdot \frac{\partial J(\hat{\theta}_i)}{\partial \hat{\theta}_i} \quad 2.46$$

Como em geral o cálculo da derivada segunda através da equação 2.45 exige muito esforço computacional Astrom [48], Keviczky [10], Payne [34] e Kashyap [31,36], aproximaram o cálculo da matriz das derivadas segundas pelo primeiro termo da equação 2.45, devido à pouca influência do outro termo na convergência [10,34]. Esta aproximação também não causará dificuldades na implementação do estimador pois é bem conhecido que o algoritmo de Newton-Raphson fornece resultados satisfatórios, mesmo quando somente um valor aproximado da derivada de segunda ordem é disponível.

O algoritmo consiste dos seguintes passos :

- i) Inicializa-se a estimação com um valor $\hat{\theta}_i^{A,B,C}$ que pode ser escolhido arbitrariamente ou como resultado de outro estimador;

- ii) Utilizando, o valor de $\hat{\theta}_{-i}^{A,B,C}$ calculam-se os valores de $J'(\hat{\theta}_{-i}^{A,B,C})$ e $J''(\hat{\theta}_{-i}^{A,B,C})$;
- iii) Obtem-se $\hat{\theta}_{-i+1}^{A,B,C}$ utilizando a expressão 2.46;
- iv) A iteração termina quando os valores absolutos das diferenças dos estimadores de $\hat{\theta}_{-i}^{A,B,C}$, em sucessivas iterações, são menores que um valor pequeno especificado ou se o valor absoluto da soma das componentes do vetor derivada é menor que um valor especificado [31].

Como o procedimento computacional para o cálculo do estimador de máxima verossimilhança utiliza técnicas relacionadas com métodos de gradiente, o algoritmo convergirá para o valor verdadeiro somente quando os valores iniciais dos estimadores não estiverem "muito longe" do valor correto dos parâmetros [31].

Quando não se conhece a matriz de covariância, Payne [34] mostrou que recai-se na minimização de $J_2(\theta)$, onde

$$J_2(\theta) = \frac{K}{2} \ln \det \frac{1}{k} \sum_{t=1}^k e_t e_t^T \quad 2.47$$

A utilização deste algoritmo em sistemas MIMO está obviamente limitada pelo tempo de cálculo envolvido.

2.8. Formulação do problema de estimação através do modelo erro previsto

São chamados métodos de previsão do erro os métodos que minimizam algum critério função do erro previsto. Estes métodos estão intimamente ligados com o estimador de máxima verossimilhança pois é possível calcular-se a função de verossimilhança como um produto das densidades condicionais da previsão da saída.

Para se obter o estimador através da formulação do erro previsto o primeiro passo será a construção de um modelo que relacione o erro previsto com as medidas e os parâmetros desconhecidos.

2.8.1. Modelo do erro previsto

Um sistema determinístico discreto pode ser modelado através de uma função que calcula as várias saídas futuras do modelo a partir das saídas e entradas anteriores como descrito pela equação 2.48.

$$Y_{k+1} = f(Y_k, U_k, \theta) \quad 2.48$$

onde as matrizes $Y_k = \{y_1, \dots, y_{k-1}\}$ e $U_k = \{u_1, \dots, u_k\}$ contêm as medidas realizadas e Y_{k+1} é o vetor de medidas no instante $k+1$ pertencente a R^n . O vetor θ contém os parâmetros desconhecidos. Sabe-se que a saída de um processo real não pode ser, em geral, exatamente determinada dadas as suas características estocásticas. Utilizam-se então modelos estocásticos como descrito pela equação 2.49.

$$Y_{k+1} = f(Y_k, U_k, \theta) + e_{k+1} \quad 2.49$$

onde $\{e_k\}$ é a sequência aleatória caracterizada por:

$$E \{e_{k+1} / Y_k, U_k\} = 0$$

O processo $\{e_k\}$ é chamado a nova informação ou o processo de inovação de $\{Y_k\}$ [38], [51].

A seguir mostra-se como se chega ao modelo do erro previsto a partir de modelos entrada/saída.

Já foi visto no capítulo anterior, que um sistema linear com perturbações estacionárias pode ser modelado por

$$Y_k = H_1(Z) u_k + H_2(Z) e_k \quad 2.50$$

onde $H_1(Z)$ e $H_2(Z)$ e $V = \text{cov} \{e_k\}$ satisfazem as condições dadas pelo teorema da representação.

A equação anterior pode ser escrita como :

$$H_2^{-1}(z) \underline{y}_k = H_2^{-1}(z) H_1(z) \underline{u}_k + \underline{e}_k \quad 2.51a$$

$$\underline{y}_k = z \left[I - H_2^{-1}(z) \right] \underline{y}_{k-1} + H_2^{-1}(z) H_1(z) \underline{u}_k + \underline{e}_k \quad 2.51b$$

Comparando as equações 2.49 e 2.51b vê-se que esta última corresponde ao modelo do erro previsto, Payne [34] e Ljung [38] também mostraram como se obtém o modelo do erro previsto a partir de modelos de estado.

O modelo descrito pela equação 2.49 é conhecido como um modelo de previsão do erro e para um valor particular estimado de $\underline{\theta} = \hat{\underline{\theta}}$ o erro de previsão do vetor de saída \underline{y}_k , no instante K, é dado por :

$$\underline{w}_k(\hat{\underline{\theta}}) \triangleq \underline{y}_k - \underline{f}(\underline{y}_{k-1}, \underline{u}_{k-1}, \hat{\underline{\theta}}) = \underline{y}_k - \hat{\underline{y}}_k \quad 2.52$$

onde $\hat{\underline{y}}_k \triangleq \underline{f}(\underline{y}_{k-1}, \underline{u}_{k-1}, \hat{\underline{\theta}})$ é o valor previsto da saída no instante K.

Por outro lado, este modelo explicita o cálculo da saída prevista, o que é essencial na maioria dos algoritmos de controle estocástico.

Os parâmetros desconhecidos serão obtidos minimizando algum tipo de critério função do erro previsto.

2.8.2. Definição do critério. Hipóteses simplificadoras

Em Ljung [38] mostram-se as propriedades que o critério escolhido deve satisfazer para que a minimização seja possível. Na literatura é normalmente utilizado um critério do tipo quadrático, como descrito pela equação 2.53, devido a facilidade de manuseio, tanto teórica como computacional.

$$J(\underline{\theta}) = \frac{1}{K} \sum_{t=1}^k \underline{w}_t^T(\underline{\theta}) Q_t^{-1} \underline{w}_t \quad 2.53$$

Q_t é uma sequência de matrizes positiva definida; a sua escolha está relacionada com a importância dada às componentes do erro previsto.

Outro critério utilizado é o descrito pela equação 2.54, devido à sua relação com o estimador de máxima verossimilhança, [10,34], como será mostrado no ítem 2.8.3.

$$J(\underline{\theta}) = \frac{1}{2} \ln \det \left[\frac{1}{K} \sum_{t=1}^k \underline{w}_t(\underline{\theta}) \underline{w}_t^T(\underline{\theta}) \right] \quad 2.54$$

No caso geral, o erro de previsão é uma função não linear dos parâmetros desconhecidos, e para minimizar-se o critério, são utilizadas técnicas de otimização não linear, isto é, resolve-se o problema de otimização através de métodos numéricos. Na escolha do método numérico leva-se em consideração problemas de tempo de cálculo devido à dimensão das matrizes envolvidas em sistemas MIMO e presença de vários pontos locais de mínimo.

Para se obterem algoritmos recursivos com tempo de cálculo compatível com a sua utilização em controle de processos, realizam-se, neste trabalho, as seguintes hipóteses simplificadoras.

Supõe-se que:

- i) o modelo utilizado é linear;
- ii) a perturbação atuando no processo é uma sequência de variáveis aleatórias gaussianas;
- iii) o processo estocástico é estacionário e gaussiano.

Utiliza-se a hipótese de modelo linear porque, como será visto no capítulo 3, tal hipótese permite obter estimadores recursivos; e um estimador recursivo tem a possibilidade de adaptar-se continuamente a variações no ponto de funcionamento que eventualmente envolvam modificações nos parâmetros do modelo linearizado. A segunda hipótese permitirá relacionar os diferentes algoritmos de estimação com o método de máxima verossimilhança.

Supõe-se que o ruído é um processo estocástico estacionário para não se modelar a evolução da covariância do processo,

e não se aumentar a complexidade do problema de estimação tornando-se impossível obter-se estimadores recursivos com tempo de cálculo compatível para aplicações "on-line".

2.8.3. Relação entre a metodologia de previsão do erro e o estimador de máxima verossimilhança

Por definição sabe-se que a função de verossimilhança de um conjunto de K medidas é dado por :

$$L(Y_k : \underline{\theta}) = p(Y_k = Y_{obs} : \underline{\theta}) \quad 2.55$$

Utilizando a regra de Bayes, é possível desenvolver a expressão da densidade de probabilidade resultando:

$$L(Y_k : \underline{\theta}) = p(y_k / y_{k-1}, \underline{\theta}) p(y_{k-1} / y_{k-2}, \underline{\theta}) \dots \quad 2.56$$

e

$$L(Y_k : \underline{\theta}) = \prod_{t=1}^k p(y_t / y_{t-1} : \underline{\theta}) \quad 2.57$$

Utilizando o modelo de previsão do erro, tem-se que :

$$P(y_k / y_{k-1} : \underline{\theta}) = p_{e_k} (w_k(\underline{\theta}) / \underline{\theta}) \quad 2.58$$

logo

$$L(Y_k : \underline{\theta}) = \prod_{t=1}^k p_{e_k} (w_t(\underline{\theta}) / \underline{\theta}) \quad 2.59$$

Com a hipótese que as inovações são normalmente distribuídas, com média zero e matriz de covariância V_t a função de verossimilhança toma a seguinte forma:

$$L(Y_k : \underline{\theta}) = \prod_{t=1}^k \left\{ \left[(2\pi)^n \det V_t \right]^{-\frac{1}{2}} \exp \left[-\frac{1}{2} w_t^T V_t^{-1} w_t \right] \right\} \quad 2.60$$

Aplicando logaritmo nos dois lados da expressão 2.60 obtém-se :

$$\ln L(Y_k; \underline{\theta}) = -\frac{1}{2} n K \ln 2\pi - \frac{1}{2} \sum_{t=1}^k \ln \det V_t - \frac{1}{2} \sum_{t=1}^k \underline{w}_t^T(\underline{\theta}) V_t^{-1} \underline{w}_t(\underline{\theta}) \quad 2.61$$

Conhecendo-se as matrizes de covariância V_t , o estimador de máxima verossimilhança é obtido como o vetor que minimiza a função de custo $J_1(\underline{\theta})$, descrita pela equação 2.62

$$J_1(\underline{\theta}) = \frac{1}{2} \sum_{t=1}^k \underline{w}_t^T(\underline{\theta}) V_t^{-1} \underline{w}_t(\underline{\theta}) \quad 2.62$$

Portanto, quando a série de inovações é gaussiana e com matriz de covariância V_t conhecida, o estimador de máxima verossimilhança é um caso particular do método de previsão do erro minimizando o critério $J_1(\underline{\theta})$.

Quando a matriz de covariância é desconhecida, deve-se realizar a maximização da função de verossimilhança, em relação à V_t e ao vetor de parâmetros desconhecidos. Utilizando a terceira hipótese simplificadora, obtem-se uma expressão analítica para o cálculo da matriz de covariância.

Neste caso o logaritmo da função de verossimilhança é descrito pela equação 2.63

$$L_1 = \ln L(\underline{\theta}, V) = -\frac{1}{2} n K \ln 2\pi - \frac{1}{2} K \ln \det V - \frac{1}{2} \sum_{t=1}^k \underline{w}_t^T V^{-1} \underline{w}_t \quad 2.63$$

Maximizando em relação à V resulta :

$$\frac{\partial L_1}{\partial V} = \frac{K}{2} V^{-1} - \frac{1}{2} V^{-1} \left[\sum_{t=1}^k \underline{w}_t \underline{w}_t^T \right] V^{-1} = 0 \quad 2.64$$

ou

$$V = \frac{1}{K} \sum_{t=1}^k \underline{w}_t(\underline{\theta}) \underline{w}_t^T(\underline{\theta}) = G(\underline{\theta}) \quad 2.65$$

Portanto, para um dado valor de $\underline{\theta}$, $G(\underline{\theta})$ maximiza a função de verossimilhança em relação à V .

Substituindo a última expressão na equação 2.63, vê-se que L_1 depende diretamente de $\underline{\theta}$, pois :

$$L_1 = -\frac{1}{2} n k \ln 2\pi - \frac{1}{2} k \ln \det G(\underline{\theta}) - \frac{1}{2} \sum_{t=1}^k \underline{w}_t^T G^{-1}(\underline{\theta}) \underline{w}_t$$

$$= -\frac{1}{2} n k \ln 2\pi - \frac{kn}{2} - \frac{k}{2} \ln \det G(\underline{\theta}) \quad 2.66$$

Então, quando a matriz de covariância é desconhecida, mas constante no tempo, o estimador de máxima verossimilhança é equivalente ao método de previsão do erro que minimiza o critério $J_2(\underline{\theta})$, descrito pela equação 2.67.

$$J_2(\underline{\theta}) = \frac{1}{2} \ln \det \left\{ \frac{1}{k} \sum_{t=1}^k \underline{w}_t(\underline{\theta}) \underline{w}_t^T(\underline{\theta}) \right\} \quad 2.67$$

Keviczky [10] obteve a mesma equação do estimador de máxima verossimilhança descrito no item 2.7 minimizando o critério $J_2(\underline{\theta})$.

2.8.4. Modelo linear para o erro de previsão

Para que o critério, descrito pela equação 2.49 seja quadrático nos parâmetros desconhecidos, e se obtenha uma equação explícita para o estimador, procurar-se-á obter o erro de previsão como uma função linear dos parâmetros.

$$J(\underline{\theta}) = \frac{1}{2} \sum_{t=1}^k \underline{w}_t^T(\underline{\theta}) V^{-1} \underline{w}_t(\underline{\theta}) \quad 2.68$$

Se o critério descrito pela equação 2.68 for quadrático nos parâmetros desconhecidos é possível obter uma equação explícita para o estimador. Para tanto é necessário que o erro de previsão possa ser escrito como uma função linear de $\underline{\theta}$.

Supondo o erro de previsão como uma função linear dos parâmetros desconhecidos, pode-se escrever que :

$$\underline{w}_k^{(0)} = \underline{y}_k - \underline{\psi}_k \underline{\theta} \quad 2.69$$

onde $\underline{\psi}_k$ é uma matriz contendo informação do processo em instantes anteriores a K.

Supondo que a matriz V é conhecida, obtem-se o estimador de máxima verossimilhança minimizando J($\underline{\theta}$)

$$\begin{aligned} \min_{\underline{\theta}} J = \min_{\underline{\theta}} \frac{1}{2} \sum_{t=1}^k \underline{w}_t^T(\underline{\theta}) V^{-1} \underline{w}_t(\underline{\theta}) &= \min_{\underline{\theta}} \frac{1}{2} \sum_{t=1}^k (\underline{y}_t - \underline{\psi}_t \underline{\theta})^T \\ &\cdot V^{-1} (\underline{y}_t - \underline{\psi}_t \underline{\theta}) \end{aligned} \quad 2.70$$

Da expressão 2.70, obtem-se o estimador descrito por:

$$\hat{\underline{\theta}} = \left[\sum_{t=1}^k \underline{\psi}_t^T V^{-1} \underline{\psi}_t \right]^{-1} \left[\sum_{t=1}^k \underline{\psi}_t^T V^{-1} \underline{y}_t \right] \quad 2.71$$

Este resultado é o mesmo que foi obtido para o estimador de Markov na seção 2.3 (equação 2.10 com $R_t=V$). No capítulo seguinte utiliza-se a formulação do erro de previsão, porque ela simplificará a análise das condições a serem impostas nas matrizes ψ e V, para se obter algoritmos particionados.

Basicamente tais algoritmos resultam de diferentes formas de se obter um modelo linear nos parâmetros.

2.9. Conclusões

Neste capítulo foram apresentados os principais resultados existentes na literatura concernentes à estimação de parâmetros em modelos entrada-saída para sistemas MIMO. Os diferentes métodos de estimação são extensões dos algoritmos existentes para o caso SISO e supõem geralmente um grande aumento em tempo de cálculo.

Viu-se que a extensão ao caso MIMO dos diferentes algoritmos não considera em geral, a influência da matriz de covariância do vetor de ruído.

A formulação do erro previsto permite unificar os diferentes algoritmos. Mostrou-se a importância de um modelo linear para o erro previsto, dado que implica numa forma explícita para o estimador.

CAPÍTULO 3

3. ESTIMAÇÃO RECURSIVA E PARTICIONADA DE SISTEMAS POR EQUAÇÕES À DIFERENÇAS

3.1. Introdução

Neste capítulo apresentam-se extensões e modificações dos algoritmos analisados no capítulo 2. O objetivo é obter algoritmos com tempo de cálculo reduzido, através de uma partição do problema de estimação. A estimação particionada do vetor de parâmetros desconhecidos permite a decomposição do problema de estimação num conjunto de subproblemas de dimensão menor. Por outro lado, esta partição permite uma implementação do algoritmo de estimação com um certo grau de paralelismo o que novamente reduz o tempo de cálculo envolvido.

A partição estudada consiste, à exceção de um caso, na estimação separada dos parâmetros de cada linha do modelo.

As condições necessárias para que a partição seja possível são analisadas formulando o problema de estimação como um problema de minimização do erro previsto. Para tanto estudam-se as modificações necessárias para se conseguir um erro previsto linear nos parâmetros desconhecidos e mostram-se os casos em que a matriz de correlação do ruído branco não influencia a estimação, sendo este um dos fatores que permitem a estimação separada de cada linha do modelo, mesmo na presença de perturbações correlatas.

Os algoritmos estudados são os seguintes:

- algoritmo dos mínimos quadrados;
- algoritmo dos mínimos quadrados linearizado, que, no caso particular de modelo média móvel para o ruído, corresponde a uma generalização do método de máxima verossimilhança aproximado apresentado por Soderstrom (18) para sistemas SISO;
- algoritmo da matriz estendida;
- algoritmo particionado que, no caso SISO e modelo autoregressivo para o ruído coincide com o algoritmo desenvolvido por Hsia (54);
- algoritmo dos mínimos quadrados generalizado.

Em todos os casos, à exceção dos mínimos quadrados, considera-se um modelo ARMA para o ruído.

\underline{b}_i^T : vetor constituído pelos parâmetros da iésima linha de $B(Z^{-1})$

Definindo $\underline{\phi}_K^T \triangleq \left[-\underline{y}_{K-1}^T \cdots -\underline{y}_{K-N_A}^T \quad \underline{u}_{K-1}^T, \dots, \underline{u}_{K-N_B}^T \right]$ e

$\underline{\psi}_K \triangleq$ bloco diagonal $\begin{bmatrix} \underline{\phi}_K^T & \underline{\phi}_K^T & \dots & \underline{\phi}_K^T \end{bmatrix}$

tem-se:

$$\underline{w}_K(\underline{\theta}^{A,B}) = \underline{y}_K - \underline{\psi}_K \underline{\theta}^{A,B} \tag{3.3}$$

Sendo $p \triangleq nN_A + mN_B$, então np é o número de parâmetros desconhecidos, $\underline{\phi}_K \in R^p$ e $\underline{\psi}_K$ é uma matriz de dimensão $nxnp$.

O estimador dos mínimos quadrados é obtido através da minimização de:

$$J = \frac{1}{2} \sum_{t=1}^K \underline{w}_t^T(\underline{\theta}) V^{-1} \underline{w}_t(\underline{\theta}) \tag{3.4}$$

onde V é a matriz de covariância do ruído \underline{e}_K suposto estacionário, resultando na expressão indicada na equação 3.5.

$$\underline{\hat{\theta}}^{A,B} = \left[\sum_{t=1}^K \underline{\psi}_t^T V^{-1} \underline{\psi}_t \right]^{-1} \left[\sum_{t=1}^K \underline{\psi}_t^T V^{-1} \underline{y}_t \right] \tag{3.5}$$

A matriz de covariância deste estimador é descrita por:

$$C_{\underline{\hat{\theta}}^{A,B}} = E \left\{ (\underline{\hat{\theta}}^{A,B} - \underline{\theta}^{A,B}) (\underline{\hat{\theta}}^{A,B} - \underline{\theta}^{A,B})^T \right\} \tag{3.6}$$

$$C_{\underline{\hat{\theta}}^{A,B}} = \left[\sum_{t=1}^K \underline{\psi}_t^T \underline{\psi}_t \right]^{-1} \left\{ \left[\sum_{t=1}^K \underline{\psi}_t^T V \underline{\psi}_t \right] \right\} \left[\sum_{t=1}^K \underline{\psi}_t^T \underline{\psi}_t \right]^{-1} \tag{3.7}$$

Da equação 3.5 nota-se que a matriz de covariância entra na equação do estimador. O objetivo agora é obter uma partição no vetor de parâmetros e analisar a influência da matriz V . Para tanto, utilizam-se as seguintes proposições:

Proposição 1: Dada uma matriz $R = \text{bloco diagonal } [\underline{r}, \underline{r}, \dots, \underline{r}]$ de dimensão $\ell \times p$ onde \underline{r} é um vetor de dimensão ℓ , e uma matriz simétrica S de dimensão p então:

$$R.S = (S \otimes I_{\ell}) . R$$

onde I_{ℓ} é uma matriz identidade $\ell \times \ell$

Prova: Seja $S = \{S_{ij}\}$ e $S_{ij} = S_{ji}$ para $i \neq j$ e $i, j = 1, \dots, p$ então:

$$R.S = \text{bloco diagonal } [\underline{r}, \underline{r}, \dots, \underline{r}] . S$$

$$R.S = \begin{bmatrix} \underline{r} S_{11} & \underline{r} S_{12} & \dots & \underline{r} S_{1p} \\ \underline{r} S_{12} & \underline{r} S_{22} & \dots & \underline{r} S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{r} S_{1p} & \underline{r} S_{2p} & \dots & \underline{r} S_{pp} \end{bmatrix}$$

$$R.S = \begin{bmatrix} S_{11} I_{\ell} & S_{12} I_{\ell} & \dots & S_{1p} I_{\ell} \\ S_{12} I_{\ell} & S_{22} I_{\ell} & \dots & S_{2p} I_{\ell} \\ \vdots & \vdots & \ddots & \vdots \\ S_{1p} I_{\ell} & S_{2p} I_{\ell} & \dots & S_{pp} I_{\ell} \end{bmatrix} . \text{ bloco diag. } [\underline{r}, \dots, \underline{r}]$$

logo $R.S = (S \otimes I_{\ell}) . R$

Δ

Proposição 1b: Dada uma matriz $R = \text{bloco diagonal } [\underline{r}_1, \underline{r}_2, \dots, \underline{r}_p]$ de dimensão $\ell \times p$ onde \underline{r}_i é um vetor de dimensão ℓ e uma matriz diagonal S de dimensão $p \times p$ então:

$$R.S = (S \otimes I_{\ell}) . R$$

I_ℓ é uma matriz identidade de dimensão $\ell \times \ell$

Prova: Seja $S = \{S_{ij}\}$ e $S_{ij} = 0$ para $i \neq j$ $i, j = 1, \dots, p$

$$R \cdot S = \text{bloco diagonal } [\underline{r}_1, \underline{r}_2, \dots, \underline{r}_p] S$$

$$R \cdot S = \text{bloco diagonal } [\underline{r}_1 S_{11}, \underline{r}_2 S_{22}, \dots, \underline{r}_p S_{pp}]$$

$$= \left[\begin{array}{c|c|c|c} S_{11} I_\ell & & & \\ \hline & S_{22} I_\ell & & \\ & & \ddots & \\ & & & S_{pp} I_\ell \end{array} \right] \text{bloco diagonal } [\underline{r}_1, \dots, \underline{r}_p]$$

$$= (S \otimes I_\ell) R$$

Δ

Proposição 2: Dada uma matriz $R = \text{bloco diagonal } [\underline{r}_1, \underline{r}_2, \dots, \underline{r}_p]$ de dimensão $\ell \times p$, onde \underline{r}_i é um vetor de dimensão ℓ , e um vetor \underline{S} de dimensão p , tal que $\underline{S}^T = [s^1, \dots, s^p]$ então:

$$\left[R \ R^T \right]^{-1} \left[R \ \underline{S} \right] = \underline{m} \quad \text{onde}$$

$$\underline{m}^T = \left[\underline{m}_1^T, \dots, \underline{m}_p^T \right] \quad \text{e} \quad \underline{m}_i = (\underline{r}_i \ \underline{r}_i^T)^{-1} \underline{r}_i s^i$$

Prova:

$$\begin{aligned}
 [R R^T]^{-1} [R \underline{S}] &= \begin{bmatrix} \underline{r}_1 & & & \\ & \underline{r}_2 & & \\ & & \ddots & \\ & & & \underline{r}_p \end{bmatrix} \begin{bmatrix} \underline{r}_1^T & & & \\ & \underline{r}_2^T & & \\ & & \ddots & \\ & & & \underline{r}_p^T \end{bmatrix} R \underline{S} \\
 &= \begin{bmatrix} \underline{r}_1 & \underline{r}_1^T \\ & \underline{r}_2 & \underline{r}_2^T \\ & & \ddots & \\ & & & \underline{r}_p & \underline{r}_p^T \end{bmatrix}^{-1} \begin{bmatrix} \underline{r}_1 & & & \\ & \underline{r}_2 & & \\ & & \ddots & \\ & & & \underline{r}_p \end{bmatrix} \underline{S} \\
 &= \begin{bmatrix} (\underline{r}_1 & \underline{r}_1^T)^{-1} & & & \\ & (\underline{r}_2 & \underline{r}_2^T)^{-1} & & \\ & & \ddots & & \\ & & & (\underline{r}_p & \underline{r}_p^T)^{-1} \end{bmatrix} \begin{bmatrix} \underline{r}_1 & & & \\ & \underline{r}_2 & & \\ & & \ddots & \\ & & & \underline{r}_p \end{bmatrix} \begin{bmatrix} s^1 \\ s^2 \\ \vdots \\ s^p \end{bmatrix} \\
 &= \begin{bmatrix} (\underline{r}_1 & \underline{r}_1^T)^{-1} & \underline{r}_1 & s^1 \\ \vdots & & & \\ (\underline{r}_p & \underline{r}_p^T)^{-1} & \underline{r}_p & s^p \end{bmatrix} = \underline{m}
 \end{aligned}$$

Δ

Utilizando a proposição 1 pode-se reescrever a equação 3.5 e tem-se:

$$\hat{\underline{\theta}}^{A,B} = \left[\sum_{t=1}^K V^{-1} \otimes I_p \psi_t^T \psi_t \right]^{-1} \left[\sum_{t=1}^K V^{-1} \otimes I_p \psi_t^T \underline{y}_t \right] \quad 3.8$$

Supondo que o ruído branco é estacionário resulta:

$$\hat{\theta}^{A,B} = \left[\sum_{t=1}^K \begin{bmatrix} \psi_t^T & \psi_t \end{bmatrix} \right]^{-1} \left[v^{-1} \otimes I_p \right]^{-1} \left[v^{-1} \otimes I_p \right] \sum_{t=1}^K \begin{bmatrix} \psi_t^T & \underline{y}_t \end{bmatrix} \quad 3.9$$

e

$$\hat{\theta}^{A,B} = \left[\sum_{t=1}^K \begin{bmatrix} \psi_t^T & \psi_t \end{bmatrix} \right]^{-1} \left[\sum_{t=1}^K \begin{bmatrix} \psi_t^T & \underline{y}_t \end{bmatrix} \right] \quad 3.10$$

Utilizando a proposição 2, dado que ψ_t tem estrutura bloco diagonal, a equação 3.10 pode ser separada em n equações, uma para cada linha do modelo:

$$\hat{\theta}_i^{A,B} = \left[\sum_{t=1}^K \begin{bmatrix} \phi_t & \phi_t^T \end{bmatrix} \right]^{-1} \left[\sum_{t=1}^K \begin{bmatrix} \phi_t & y_t^i \end{bmatrix} \right] \quad i=1,2,\dots,n \quad 3.11$$

onde y_t^i é a medida da i-ésima componente do vetor de saída. Nota-se que, além da possibilidade de calcular em paralelo os vetores $\hat{\theta}_i^{A,B}$, tem-se uma grande redução em tempo de cálculo devido a dimensão da matriz a ser invertida (pxp em lugar de npxnp no caso global).

Utilizando a identidade 2.12 pode-se obter a versão recursiva do algoritmo dos mínimos quadrados, descrita pela equação 3.11.b

$$\hat{\theta}_{i,K+1}^{A,B} = \hat{\theta}_{i,K}^{A,B} + \frac{i}{K+1} \left\{ y_{K+1}^i - \phi_{K+1}^T \hat{\theta}_{i,K}^{A,B} \right\} \quad i=1,2,\dots,n$$

$$\frac{i}{K+1} = P_K^i \phi_{K+1} / (1 + \phi_{K+1}^T P_K^i \phi_{K+1}) \quad 3.11b$$

$$P_{K+1}^i = P_K^i - \frac{K^i}{K+1} \phi_{K+1}^T P_K^i$$

Quando não existe uma diferença significativa na incerteza associada aos valores iniciais dos parâmetros desconhecidos escolhem-se os mesmos valores iniciais para todas as matrizes P_K^i . Neste caso, os ganhos K_K^i são iguais e obtém-se uma significativa redução no tempo de computação do algoritmo porque se atualiza uma única matriz quadrada de dimensão p em lugar de n matrizes quadradas de dimensão p .

Também pode-se obter a equação recursiva do estimador global (equação 3.10) aplicando-se a identidade 2.8 na equação 3.9 e definindo-se:

$$\underline{P}_K = \left[\begin{array}{c} K \\ \sum_{t=1}^K \psi_t^T \psi_t \end{array} \right]^{-1} \quad 3.12$$

O estimador recursivo é então descrito pela equação 3.13

$$\hat{\theta}_{K+1}^{A,B} = \hat{\theta}_K^{A,B} + K_{K+1} \left\{ \underline{y}_{K+1} - \psi_{K+1} \hat{\theta}_K^{A,B} \right\}$$

$$K_{K+1} = \underline{P}_{K+1} \psi_{K+1}^T \quad 3.13$$

$$\underline{P}_{K+1} = \underline{P}_K - \underline{P}_K \psi_{K+1}^T \left[I + \psi_{K+1} \underline{P}_K \psi_{K+1}^T \right]^{-1} \psi_{K+1} \underline{P}_K$$

Se o valor inicial da matriz P_K é constituído pelos valores iniciais das matrizes P_K^i , isto é,

$$P_0 = \text{bloco diagonal} \left[P_0^1, \dots, P_0^n \right], \quad P_0 \text{ (npxnp)} \quad 3.14$$

então o ganho K_K é dado por:

$$K_K = \text{bloco diagonal} \left[\begin{array}{c|c} P_K^1 & \phi_K \\ \hline \dots & \\ \hline P_K^n & \phi_K \end{array} \right] \quad 3.15a$$

ou

$$K_K = \text{bloco diagonal} \begin{bmatrix} \underline{K}_K^1 & | & \dots & | & \underline{K}_K^n \\ \hline & & & & \end{bmatrix}, K_K \text{ (npxn)} \quad 3.15b$$

Esta última equação mostra como que o ganho do estimador global recursivo é obtido, a partir dos ganhos dos estimadores recursivos de cada linha do modelo.

A seguir obtêm-se a matriz de covariância do estimador global (equação 3.6) a partir do estimador particionado.

Para tal reescreve-se a equação 3.7 como segue:

$$C_{\hat{\theta}^{A,B}} = \begin{bmatrix} K \\ \sum_{t=1} \psi_t^T \psi_t \end{bmatrix}^{-1} V \otimes I_P \begin{bmatrix} K \\ \sum_{t=1} \psi_t^T \psi_t \end{bmatrix} \begin{bmatrix} K \\ \sum_{t=1} \psi_t^T \psi_t \end{bmatrix}^{-1} \quad 3.16$$

$$C_{\hat{\theta}^{A,B}} = V \otimes I_P \begin{bmatrix} K \\ \sum_{t=1} \psi_t^T \psi_t \end{bmatrix}^{-1} \quad 3.17$$

$$C_{\hat{\theta}^{A,B}} = V \otimes I_P \left[\text{bloco diagonal} \begin{bmatrix} \underline{K}_K^1, \dots, \underline{K}_K^n \end{bmatrix} \right]^{-1} \quad 3.18$$

Portanto, com os ganhos \underline{K}_K^i dos estimadores de cada linha do modelo, obtêm-se a matriz de covariância do estimador global utilizando a equação 3.18.

O estimador dos mínimos quadrados representado pela equação 3.11.b é bem conhecido na literatura [1, 7, 10, 11] mas é obtido em todos os casos a partir de um critério que não leva em consideração a matriz de covariância do ruído.

Através das proposições 1 e 2, mostrou-se que, no caso multivariável, se a perturbação é um processo estocástico estacionário branco, então o estimador dos mínimos quadrados ponderados independe da matriz de covariância do ruído e, deste modo pode-se realizar a estimação linha por linha do modelo, acarretando uma sensível redução no tempo de cálculo.

Entretanto, como já foi dito no capítulo anterior, sabe-se que o estimador dos mínimos quadrados é consistente somente se o ruído for uma sequência de variáveis aleatórias brancas.

A seguir desenvolvem-se técnicas de estimação particionada para algoritmos que eliminem este problema através da modelagem do ruído e consequente estimação dos parâmetros do modelo. Nestes casos a equação do erro previsto é não linear nos parâmetros a serem estimados e torna-se necessária algum tipo de aproximação para a obtenção de uma forma explícita para o estimador.

3.3. Estimador dos mínimos quadrados linearizado

A aproximação utilizada para a obtenção deste algoritmo consiste na linearização do erro previsto em torno do último valor estimado do vetor de parâmetros desconhecidos. Esta linearização foi utilizada por Payne (34) no desenvolvimento do algoritmo de máxima verossimilhança aproximado, para sistemas SISO, utilizando um modelo média móvel para o ruído. O algoritmo foi inicialmente obtido por Soderstrom (18) utilizando uma aproximação de segunda ordem para a função de custo. Apresenta-se a seguir a generalização desse algoritmo para Sistemas MIMO com modelo ARMA do ruído.

Esta seção é constituída pela dedução do algoritmo linearizado, cálculo da derivada do erro previsto e estimação dos ruídos.

Seja o modelo descrito pela seguinte equação:

$$A(z^{-1})\underline{y}_K = B(z^{-1})\underline{u}_K + \underline{v}_K$$

$$D(z^{-1})\underline{v}_K = C(z^{-1})\underline{e}_K \tag{3.19}$$

Neste caso a equação do erro previsto é descrita pela equação 3.20.

$$\underline{w}_K(\underline{\theta}^{A,B,C,D}) = \left[C(Z^{-1}) \right]^{-1} D(Z^{-1}) \left\{ A(Z^{-1}) \underline{y}_K - B(Z^{-1}) \underline{u}_K \right\} \quad 3.20$$

onde $\underline{\theta}^{A,B,C,D}$ é um vetor contendo os parâmetros desconhecidos das matrizes polinomiais $A(Z^{-1})$, $B(Z^{-1})$, $C(Z^{-1})$ e $D(Z^{-1})$.

Na equação 3.20 observa-se que \underline{w}_K é uma função não linear dos parâmetros desconhecidos.

Para se obter uma equação explícita para o estimador realiza-se, nesta seção, uma linearização do erro previsto em torno do último valor estimado ($\hat{\underline{\theta}}_K^{A,B,C,D}$), isto é:

$$\underline{w}_K(\underline{\theta}^{A,B,C,D}) \cong \underline{w}_K(\hat{\underline{\theta}}_{K-1}^{A,B,C,D}) + \frac{\partial \underline{w}_K}{\partial \underline{\theta}} \bigg|_{\underline{\theta}=\hat{\underline{\theta}}_{K-1}^{A,B,C,D}} \begin{pmatrix} \underline{\theta}^{A,B,C,D} - \hat{\underline{\theta}}_{K-1}^{A,B,C,D} \end{pmatrix} \quad 3.21$$

onde $\hat{\underline{\theta}}_K^{A,B,C,D}$ é o estimador dos parâmetros desconhecidos no instante K .

O estimador é obtido minimizando-se

$$J_K(\underline{\theta}^{A,B,C,D}) = \frac{1}{2} \sum_{t=1}^K \underline{w}_t^T(\underline{\theta}^{A,B,C,D}) V^{-1} \underline{w}_t(\underline{\theta}^{A,B,C,D}) \quad 3.22$$

Para simplicidade da notação seja

$$\hat{\underline{\theta}}_K = \hat{\underline{\theta}}_K^{A,B,C,D} \quad \text{e} \quad \underline{\theta} = \underline{\theta}^{A,B,C,D}$$

A função de custo no instante $K+1$ é descrita por:

$$J_{K+1}(\underline{\theta}) = \frac{1}{2} \sum_{t=1}^{K+1} \underline{w}_t^T(\cdot) V^{-1} \underline{w}_t(\cdot) \quad 3.23$$

com $\underline{w}_t(\cdot) = \underline{w}_t(\underline{\theta}) = \underline{w}_t(\underline{\theta}^{A,B,C,D})$

Obtêm-se diretamente o estimador recursivo considerando-se que $\hat{\underline{\theta}}_K$ é um ponto de mínimo de $J_K(\underline{\theta})$ e utilizando uma aproximação quadrática para $J_{K+1}(\underline{\theta})$ em torno de $\hat{\underline{\theta}}_K$

$$J_{K+1}(\underline{\theta}) = J_{K+1}(\hat{\underline{\theta}}_K) + \underline{J}_{K+1}'^T(\hat{\underline{\theta}}_K) (\underline{\theta} - \hat{\underline{\theta}}_K) + \frac{1}{2} (\underline{\theta} - \hat{\underline{\theta}}_K)^T \underline{J}_{K+1}''^T(\hat{\underline{\theta}}_K) (\underline{\theta} - \hat{\underline{\theta}}_K) \quad 3.24$$

Minimizando o critério $J_{K+1}(\underline{\theta}) = J_{K+1}(\cdot)$ resulta o estimador no instante $K+1$ descrito por:

$$\hat{\underline{\theta}}_{K+1} = \hat{\underline{\theta}}_K - \underline{J}_{K+1}''^T(\hat{\underline{\theta}}_K)^{-1} \underline{J}_{K+1}'^T(\hat{\underline{\theta}}_K) \quad 3.25$$

Esta última equação corresponde à primeira iteração do algoritmo de máxima verossimilhança apresentado no capítulo 2.

Da equação 3.25 observa-se que, para se obter um estimador recursivo, deve-se calcular o gradiente e a matriz Hessiana da função de custo no instante $K+1$ como uma função do instante anterior e do novo instante.

Da equação 3.23 resulta:

$$J_{K+1}(\underline{\theta}) = J_K(\underline{\theta}) + \frac{1}{2} \underline{w}_{K+1}^T(\cdot) \underline{V}^{-1} \underline{w}_{K+1}(\cdot) \quad 3.26$$

Derivando ambos os lados da equação 3.26 em relação a $\underline{\theta}$ obtêm-se:

$$\underline{J}_{K+1}'^T(\underline{\theta}) = \underline{J}_K'(\underline{\theta}) + \underline{w}_{K+1}'(\cdot) \underline{V}^{-1} \underline{w}_{K+1}(\cdot) \quad 3.27$$

Com a hipótese que $\underline{\theta} = \hat{\underline{\theta}}_K$ é um ponto de mínimo do critério 3.22 resulta:

$$J'_{K+1}(\hat{\theta}_{-K}) = w'_{K+1}(\hat{\theta}_{-K}) V^{-1} w'_{K+1}(\hat{\theta}_{-K}) \quad 3.28$$

Com a linearização do erro previsto em torno do último valor estimado (equação 3.21) obtém-se, a partir da equação 3.27 a seguinte expressão para a matriz Hessiana.

$$J''_{K+1}(\hat{\theta}_{-K}) = J''_K(\hat{\theta}_{-K}) + w'_{K+1}(\hat{\theta}_{-K}) V^{-1} w'_{K+1}(\hat{\theta}_{-K}) \quad 3.29$$

Supondo que os valores estimados, em instantes sucessivos, estão suficientemente próximos pode-se assumir que:

$$J''_K(\hat{\theta}_{-K}) = J''_K(\hat{\theta}_{-K-1}) \quad 3.30$$

Levando a equação 3.30 na equação 3.29 obtém-se:

$$J''_{K+1}(\hat{\theta}_{-K}) = J''_K(\hat{\theta}_{-K-1}) + w'^T_{K+1}(\hat{\theta}_{-K}) V^{-1} w'_{K+1}(\hat{\theta}_{-K}) \quad 3.31$$

Definindo: i) $\psi_{K+1} = w'_{K+1}(\hat{\theta}_{-K}) \quad 3.32$

ii) $P_{K+1} = \left[J''_{K+1}(\hat{\theta}_{-K}) \right]^{-1} \quad 3.33a$

A partir da equação 3.31 tem-se que:

$$P_{K+1} = \left[J''_K(\hat{\theta}_{-K-1}) + \psi'^T_{K+1} V^{-1} \psi_{K+1} \right]^{-1} \quad 3.33b$$

Utilizando a equação da inversa recursiva (equação 2.12) resulta:

$$P_{K+1} = P_K - P_K \psi'^T_{K+1} \left[V + \psi_{K+1} P_K \psi'^T_{K+1} \right]^{-1} \psi_{K+1} P_K \quad 3.33c$$

Levando as equações 3.28, 3.31 e 3.33a na equação do estimador, obtêm-se

$$\hat{\theta}_{K+1}^{A,B,C,D} = \hat{\theta}_K^{A,B,C,D} + K_{K+1} w_{K+1} (\hat{\theta}_K^{A,B,C,D})$$

$$K_{K+1} = P_{K+1} \psi_{K+1}^T V^{-1} \quad 3.34$$

$$P_{K+1} = P_K - P_K \psi_{K+1}^T \left[V + \psi_{K+1} P_K \psi_{K+1}^T \right]^{-1} \psi_{K+1} P_K$$

A equação 3.34 representa o estimador dos mínimos quadrados linearizado. Nesta equação observa-se que, para se obter o estimador, é necessário ainda o cálculo da derivada do erro previsto (equação 3.32) em relação aos parâmetros desconhecidos.

- Cálculo da derivada do erro previsto

Sabe-se que a equação do erro previsto é descrita por:

$$w_K(\cdot) = C^{-1}(z^{-1}) D(z^{-1}) \{A(z^{-1}) \underline{y}_K - B(z^{-1}) \underline{u}_K\}$$

ou de acordo com a Eq. (3.35):

$$w_K(\cdot) = C^{-1}(z^{-1}) D(z^{-1}) \{ \underline{y}_K - \psi_K^P \underline{\theta}^{A,B} \} \quad 3.35$$

onde:

$$\psi_K^P = \text{bloco diagonal} \{ \phi_{K'}^{PT}, \phi_K^{PT}, \dots, \phi_K^{PT} \}$$

$$\phi_K^{PT} = \begin{bmatrix} -\underline{y}_{K-1}^T & -\underline{y}_{K-2}^T & \underline{y}_{K-NA}^T & \underline{u}_{K-1}^T & \dots & \underline{u}_{K-NB}^T \end{bmatrix}$$

e $\underline{\theta}^{A,B}$ como definido na seção 3.2.

Reescrevendo a equação 3.35 resulta:

$$C(Z^{-1}) \underline{w}_K(\cdot) = D(Z^{-1}) \{ \underline{y}_K - \psi_K^P \underline{\theta}^{A,B} \} \quad 3.36$$

Definindo $\underline{\theta}^{TC} = \begin{bmatrix} \underline{c}_1^T \\ \underline{c}_2^T \\ \dots \\ \underline{c}_n^T \end{bmatrix}$ e $\underline{\theta}^{TD} = \begin{bmatrix} \underline{d}_1^T \\ \underline{d}_2^T \\ \dots \\ \underline{d}_n^T \end{bmatrix}$ com

\underline{c}_1^T : vetor constituído pelos parâmetros da iésima linha de $C(Z^{-1})$

\underline{d}_1^T : vetor constituído pelos parâmetros da iésima linha de $D(Z^{-1})$

Derivando a equação do erro previsto em relação a $\underline{\theta}^{A,B}$, obtém-se:

$$\frac{\partial}{\partial \underline{\theta}^{A,B}} C(Z^{-1}) \underline{w}_K(\cdot) = C(Z^{-1}) \frac{\partial}{\partial \underline{\theta}^{A,B}} \underline{w}_K(\cdot) = -D(Z^{-1}) \psi_K^P \quad 3.37$$

Derivando a equação 3.36 em relação a θ^C resulta:

$$\frac{\partial}{\partial \theta^C} C(Z^{-1}) \underline{w}_K(\cdot) = \left[\frac{\partial C(Z^{-1})}{\partial \theta^C} \Big| \underline{w}_K(\cdot) \right] + C(Z^{-1}) \frac{\partial \underline{w}_K(\cdot)}{\partial \theta^C} = 0 \quad 3.38$$

A partir de definição de $\underline{\theta}^C$ obtém-se

$$\begin{aligned}
\frac{\partial C(z^{-1})}{\partial \underline{\theta}^c} \underline{w}_K(\cdot) &= \begin{bmatrix} z^{-1} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{bmatrix} \underline{w}_K(\cdot) \left| \begin{bmatrix} 0 & z^{-1} & \dots & 0 \\ 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{bmatrix} \underline{w}_K(\cdot) \right| \dots \\
\dots & \begin{bmatrix} 0 & 0 & \dots & z^{-1} \\ 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{bmatrix} \underline{w}_K(\cdot) \left| \begin{bmatrix} z^{-2} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{bmatrix} \underline{w}_K(\cdot) \right| \dots \left| \begin{bmatrix} 0 & 0 & \dots & z^{-2} \\ 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{bmatrix} \underline{w}_K(\cdot) \right| \\
\dots & \left| \begin{bmatrix} z^{-N_C} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{bmatrix} \underline{w}_K(\cdot) \right| \dots \left| \begin{bmatrix} 0 & 0 & \dots & z^{-N_C} \\ 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{bmatrix} \underline{w}_K(\cdot) \right| \dots \\
\dots & \left| \begin{bmatrix} 0 & 0 & \dots & 0 \\ z^{-1} & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{bmatrix} \underline{w}_K(\cdot) \right| \dots \left| \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & z^{-N_C} \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{bmatrix} \underline{w}_K(\cdot) \right| \dots \left| \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & z^{-N_C} \end{bmatrix} \underline{w}_K(\cdot) \right| \\
&= \underline{w}_{K-1}^T(\cdot) \quad \underline{w}_{K-2}^T(\cdot) \quad \dots \quad \underline{w}_{K-N_C}^T(\cdot) \quad \left| \begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \end{array} \right| \\
&\quad \quad \left| \begin{array}{c} \underline{w}_{K-1}^T(\cdot) \quad \underline{w}_{K-2}^T(\cdot) \quad \dots \quad \underline{w}_{K-N_C}^T(\cdot) \end{array} \right| \\
\dots & \left| \begin{array}{c} 0 \\ \vdots \\ \underline{w}_{K-1}^T(\cdot) \quad \dots \quad \underline{w}_{K-N_C}^T(\cdot) \end{array} \right|
\end{aligned}$$

$$\left[\frac{\partial C(z^{-1})}{\partial \underline{\theta}^c} \underline{w}_K(\cdot) \right] = + \psi_K^w = + \text{bloco diagonal} \begin{bmatrix} \underline{m}_K^T & & & \\ & \underline{m}_K^T & & \\ & & \dots & \\ & & & \underline{m}_K^T \end{bmatrix}$$

1 2 ... n

$$\text{com } \underline{m}_K^T = \left[\underline{w}_{K-1}^T(\cdot), \underline{w}_{K-2}^T(\cdot), \dots, \underline{w}_{K-N_C}^T(\cdot) \right]$$

logo:

$$C(z^{-1}) \frac{\partial \underline{w}_K(\cdot)}{\partial \underline{\theta}^c} = - \psi_K^w$$

Para o cálculo da derivada do erro previsto em relação a $\underline{\theta}^D$ tem-se, a partir da equação do erro previsto, que:

$$\frac{\partial}{\partial \underline{\theta}^D} C(Z^{-1}) \underline{w}_K(\cdot) = C(Z^{-1}) \frac{\partial}{\partial \underline{\theta}^D} \underline{w}_K(\cdot) = \left[\begin{array}{c} \frac{\partial D(Z^{-1})}{\partial \underline{\theta}^D} \\ \vdots \\ \vdots \end{array} \middle| \underline{v}_K(\cdot) \right]$$

com $\underline{v}_K(\cdot) = A(Z^{-1}) \underline{y}_K - B(Z^{-1}) \underline{u}_K$

Analogamente a derivada do erro previsto em relação aos parâmetros $\underline{\theta}^C$, obtém-se que

$$C(Z^{-1}) \frac{\partial}{\partial \underline{\theta}^D} \underline{w}_K(\cdot) = - \text{bloco diagonal} \left[\underline{m}_K^{1T}, \underline{m}_K^{1T}, \dots, \underline{m}_K^{1T} \right] \quad 3.40$$

$$= - \psi_K^V$$

e $\underline{m}_K^{1T} = \left[-\underline{v}_{K-1}^T(\cdot), -\underline{v}_{K-2}^T(\cdot), \dots, -\underline{v}_{K-N_D}^T(\cdot) \right]$

Portanto, $\psi_K = -\frac{\partial \underline{w}_K(\cdot)}{\partial \underline{\theta}^{A,B,C,D}}$ é obtido a partir das equações 3.37, 3.39 e 3.40.

- Estimação do ruído

Para a utilização posterior das equações da derivada do erro previsto na estimação é necessário conhecer-se os ruídos \underline{v}_t para $t=K-1, \dots, K-N_D$ e \underline{w}_t para $t=K-1, \dots, K-N_C$.

Se os parâmetros a estimar fossem conhecidos os valores dos ruídos poderiam ser calculados pelas definições de \underline{y}_K e \underline{w}_K , isto é:

$$\underline{v}_K(\underline{\theta}) = \underline{y}_K - \psi_K^P \underline{\theta}^{A,B} \quad 3.41$$

$$\underline{w}_K(\underline{\theta}) = \underline{v}_K - \left[+ \psi_K^W - \psi_K^V \right] \underline{\theta}^{C,D} \quad 3.42$$

Quando o vetor $\underline{\theta}^{A,B,C,D}$ é desconhecido, após a K-ésima medida e com o estimador do instante anterior $(\hat{\underline{\theta}}_{K-1}^{A,B,C,D})$, calculam-se os valores estimados de $\{\hat{\underline{v}}_t(\hat{\underline{\theta}}_{K-1})$ e $\hat{\underline{w}}_t(\hat{\underline{\theta}}_{K-1})$, $t=0,1,\dots,K\}$ através das equações 3.41 e 3.42 respectivamente, onde $\underline{\theta}^{A,B}$ e $\underline{\theta}^{C,D}$ são substituídos pelos valores estimados no instante K-1.

Neste caso o algoritmo é descrito pelas seguintes equações:

$$\hat{\underline{\theta}}_{K+1}^{A,B,C,D} = \hat{\underline{\theta}}_{K+1}^{A,B,C,D} + K_{K+1} \underline{w}_{K+1} (\hat{\underline{\theta}}_{K+1}^{A,B,C,D})$$

$$K_{K+1} = P_{K+1} \psi_{K+1}^T V^{-1} \quad 3.43$$

$$P_{K+1} = P_K - P_K \psi_{K+1}^T \left[V + \psi_{K+1} P_K \psi_{K+1}^T \right]^{-1} \psi_{K+1} P_K$$

onde

$$\psi_{K+1} = - \frac{\partial \underline{w}_{K+1}}{\partial \underline{\theta}^{A,B,C,D}} (\cdot) \quad 3.44.a$$

e

$$C_{K-1}(Z^{-1}) \frac{\partial \underline{w}_K}{\partial \underline{\theta}^{A,B}} (\cdot) = -D_{K-1}(Z^{-1}) \psi_K^P$$

$$C_{K-1}(Z^{-1}) \frac{\partial \underline{w}_K}{\partial \underline{\theta}^C} (\cdot) = - \psi_K^W \quad 3.44.b$$

$$C_{K-1}(Z^{-1}) \frac{\partial \underline{w}_K}{\partial \underline{\theta}^D} (\cdot) = - \psi_K^V$$

e os ruídos são estimados através da equação 3.45 para valores de $t=0,1,\dots,K$.

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

$$\hat{w}_K(\hat{\theta}_{K-1}) = \hat{v}_K(\hat{\theta}_{K-1}) - \begin{bmatrix} \psi_K^W & -\psi_K^V \end{bmatrix} \hat{\theta}_{K-1}^{C,D} \quad 3.45$$

- Simplificação na estimação dos ruídos

Entretanto, como este procedimento requer muito esforço computacional, para valores elevados de K, utiliza-se, neste algoritmo, a idéia formulada por Young [47] que atualiza somente o valor estimado do ruído no instante da nova medida. Isto é, a cada medida realizada (\underline{y}_K), estimam-se os ruídos através de:

$$\hat{v}_K(\hat{\theta}) = \underline{y}_K - \psi_K^P \hat{\theta}_{K-1}^{A,B} \quad 3.46$$

$$\hat{w}_K(\hat{\theta}) = \hat{v}_K(\hat{\theta}) - \begin{bmatrix} \psi_K^W & -\psi_K^V \end{bmatrix} \hat{\theta}_{K-1}^{C,D} = \underline{y}_K - \begin{bmatrix} \psi_K^P & +\psi_K^W & -\psi_K^V \end{bmatrix} \hat{\theta}_{K-1}^{A,B,C,D} \quad 3.47$$

Observa-se que, neste caso, os erros cometidos na estimação dos valores de \hat{v} e \hat{w} no passado são maiores, em virtude da falta de informação inicial sobre os parâmetros a estimar. Por isso pode ser conveniente a introdução de um fator de ponderação que de maior peso às novas medidas, como proposto por Talman em [15].

Além disso, nota-se que as hipóteses simplificadoras do algoritmo somente são válidas para valores elevados de K. Por isso normalmente o algoritmo apresentado deve ser inicializado com os resultados de algum outro estimador.

- Conclusões

Nesta secção deduziu-se um novo algoritmo de estimação de parâmetros para sistemas MIMO.

O algoritmo linearizado foi obtido através de uma linearização do erro previsto.

Na dedução deste algoritmo utilizou-se um modelo geral (ARMA) para o ruído, isto é:

$$D(Z^{-1}) \underline{v}_K = C(Z^{-1}) \underline{e}_K$$

onde \underline{e}_K é um ruído branco com matriz de covariância V .

Soderstrom [18] obteve a versão monovariável deste algoritmo, considerando somente um modelo média móvel para o ruído, isto é, utilizou $D(Z^{-1})=I$ no modelo da perturbação descrito acima. Neste caso, ele é conhecido na literatura como estimador de máxima verossimilhança aproximado Bányas e Gertler [52] estudaram a versão monovariável do algoritmo linearizado quando se utiliza somente um modelo autoregressivo para o ruído.

Da equação 3.44 nota-se que, se $C(Z^{-1})$ e $D(Z^{-1})$ são matrizes polinomiais quaisquer, ψ_K não terá estrutura bloco diagonal. Assim as proposições 1 e 2 da secção 3.2 não são aplicáveis e não se pode particionar o vetor de parâmetros desconhecidos.

A seguir estudam-se três situações em que é possível a partição do estimador linearizado.

3.3.1. Algoritmo linearizado particionado utilizando a forma pseudo-canônica I

Mostra-se, nesta secção, que a utilização da forma pseudo-canônica I permite a partição do problema de estimação em n subproblemas de estimação que podem ser resolvidos paralelamente.

Seja inicialmente, um sistema modelado através da forma pseudo-canônica I, isto é, pressupõe-se que o ruído pode ser modelado como um processo ARMA com $C(Z^{-1})$ e $D(Z^{-1})$ diagonais.

Utilizando a definição de P_{K+1} (equação 3.33a e a equação 3.34) vê-se que o estimador linearizado é descrito por:

$$\hat{\underline{\theta}}_{-K+1}^{A,B,C,D} = \hat{\underline{\theta}}_K^{A,B,C,D} + \left[\sum_{t=1}^{K+1} \begin{bmatrix} \psi_t^T & V^{-1} & \psi_t \end{bmatrix} \right]^{-1} \psi_{K+1}^T V^{-1} \underline{w}_{K+1}(\cdot) \quad 3.48$$

Como neste caso as matrizes $C(Z^{-1})$ e $D(Z^{-1})$ são diagonais, resulta da equação 3.44 que $\{\psi_t, t=1, \dots, K+1\}$ tem uma estrutura bloco diagonal pois ψ_K^P , ψ_K^W e ψ_K^V tem uma estrutura bloco diagonal e $C_{K-1}(Z^{-1})$ e $D_{K-1}(Z^{-1})$ são diagonais.

Reagrupando os elementos da matriz ψ_K resulta

$$\psi_K = \text{bloco diagonal}\{\xi_{1,K}, \xi_{2,K}, \dots, \xi_{n,K}\} \quad 3.49$$

com $c_{i,K-1}(Z^{-1})\xi_{i,K} = \underline{m}_{i,K}^3$

$$\underline{m}_{i,K}^3 = \begin{bmatrix} -\underline{m}_{i,K'}^{2T} & -\underline{m}_{i,K}^{4T} & \underline{m}_{i,K}^{5T} \end{bmatrix}$$

$$\underline{m}_{i,K}^2 = d_{i,K-1}(Z^{-1}) \underline{\phi}_K \quad 3.50$$

$c_{i,\cdot}(Z^{-1})$ e $d_{i,\cdot}(Z^{-1})$ são respectivamente, elementos da i ésima linha de $C(Z^{-1})$ e $D(Z^{-1})$, e:

$$\underline{m}_{i,K}^{4T} = \begin{bmatrix} w_{i,K-1} & w_{i,K-2} & \dots & w_{i,K-N_C} \end{bmatrix}$$

$$\underline{m}_{i,K}^{5T} = \begin{bmatrix} v_{i,K-1} & v_{i,K-2} & \dots & v_{i,K-N_D} \end{bmatrix}$$

Sabendo que ψ_K é bloco diagonal e a matriz de covariância V é diagonal e utilizando a proposição 1.b resulta o seguinte estimador:

$$\hat{\theta}_{K+1}^{A,B,C,D} = \hat{\theta}_K^{A,B,C,D} + \left[\sum_{t=1}^{K+1} \psi_t^T \psi_t \right]^{-1} \psi_{K+1} \underline{w}_{K+1}(\cdot) \quad 3.51$$

Como ψ_t tem estrutura bloco diagonal pode-se aplicar a

proposição 2 e resulta o estimador de cada linha do modelo separadamente, descrito por:

$$\hat{\theta}_{i,K+1} = \hat{\theta}_{i,K} + \frac{K_{K+1}^i}{K_{K+1}} w_{K+1}^i (\cdot)$$

$$K_{K+1}^i = P_{K+1}^i \xi_{K+1}^i$$

3.52

$$P_{K+1}^i = (P_K^i - P_K^i \xi_{K+1}^i (1 + \xi_{K+1}^{iT} P_K^i \xi_{K+1}^i)^{-1} \xi_{K+1}^{iT} P_K^i)$$

Portanto, neste caso, como no algoritmo dos mínimos quadrados (secção 3.2) em vez de se trabalhar com uma matriz quadrada de dimensão np , onde $p=nN_A+nN_C+nN_D+mN_B$, atualizam-se n matrizes quadradas de dimensão P que podem ser calculadas em paralelo, reduzindo aproximadamente n^2 vezes o tempo de cálculo do algoritmo da seção anterior.

3.3.2. Algoritmo linearizado particionado utilizando a forma canônica III

A forma canônica III apresentada por Kashyap [10] e descrita na seção 1.3 permite reescrever de uma forma conveniente o modelo dado pela equação 3.19 repetido a seguir:

$$A(z^{-1})\underline{y}_K = B(z^{-1})\underline{u}_K + \underline{v}_K$$

$$D(z^{-1})\underline{v}_K = C(z^{-1})\underline{e}_K$$

Introduzindo o modelo do ruído na primeira equação tem-se:

$$D(z^{-1})A(z^{-1})\underline{y}_K = D(z^{-1})B(z^{-1})\underline{u}_K + C(z^{-1})\underline{e}_K \quad \text{cov}(\underline{e}_K) = V$$

Utilizando a forma canônica III de Kashyap obtêm-se:

$$A^*(z^{-1})\underline{y}_K = B^*(z^{-1})\underline{u}_K + C^*(z^{-1})\underline{e}_K^* \quad \text{cov}(\underline{e}_K^*) = V^*$$

onde $A^*(z^{-1})$, $B^*(z^{-1})$, $C^*(z^{-1})$ e V^* têm as características definidas na seção 1.3.

Interessa aqui especificamente o fato de $C^*(z^{-1})$ e V^* serem diagonais. Neste caso o algoritmo é descrito pela equação 3.52 com

$$\underline{m}_{i,K}^2 = \underline{\phi}_K \quad \text{para qualquer } i \quad 3.53$$

e

$$\underline{m}_{i,K}^{3T} = \begin{bmatrix} -\underline{m}_{i,K}^{2T} & \underline{m}_{i,K}^{4T} \end{bmatrix} \quad 3.54$$

sendo possível a estimação separada de cada linha do processo.

Convém ressaltar que, neste caso, não se está fazendo nenhuma hipótese restritiva sobre o modelo do ruído porque se utiliza de uma forma canônica, ao contrário do caso da seção anterior em que a parametrização das matrizes $C(z^{-1})$ e $D(z^{-1})$ nem sempre representará uma aproximação válida.

A utilização da forma canônica III envolve um aumento na ordem das matrizes polinomiais $A^*(z^{-1})$ e $B^*(z^{-1})$ em relação a $A(z^{-1})$ e $B(z^{-1})$ [31,53]. Da mesma forma que no algoritmo particionado descrito na seção anterior este estimador independe da matriz de covariância do ruído (V^*) e o ganho \underline{K}^i é diferente para cada linha. Na próxima seção utiliza-se a forma pseudo-canônica II que leva a um ganho único para todas as linhas.

3.3.3. Algoritmo linearizado particionado utilizando a forma pseudo-canônica II

Para se obter um algoritmo com um ganho único para todas as linhas do modelo a matriz ψ_K deve ter uma estrutura com blocos iguais na diagonal.

Da equação 3.36 vê-se que ψ_K^P , ψ_K^W , ψ_K^V têm uma estrutura bloco diagonal. Deste modo pode-se obter ψ_K com a estrutura desejada utilizando a forma pseudo-canônica II, onde o modelo dinâmico de cada uma das componentes do ruído é o mesmo, permitindo apenas diferenças na matriz de covariância.

Com as hipóteses das matrizes $C(Z^{-1})$ e $D(Z^{-1})$ diagonais pode-se, como na seção 3.3.1 rearranjar-se a matriz das derivadas e obter:

$$\psi_{K+1} = \text{bloco diagonal} \begin{bmatrix} \xi_{K'} & \xi_{K'} & \dots & \xi_{K'} \\ 1 & 2 & \dots & n \end{bmatrix} \quad 3.55$$

com

$$C(Z^{-1}) \xi_{-K} = \underline{m}_K^3$$

e

$$\underline{m}_K^{3T} = \begin{bmatrix} -\underline{m}_K^{2T} & -\underline{m}_K^{4T} & \underline{m}_K^{5T} \end{bmatrix} \quad 3.56$$

e

$$\underline{m}_K^2 = d(Z^{-1}) \phi_K$$

$$\underline{m}_K^{4T} = \begin{bmatrix} w_{i,K-1} & w_{i,K-2} & \dots & w_{i,K-NC} \end{bmatrix}$$

$$\underline{m}_K^{5T} = \begin{bmatrix} v_{i,K-1} & v_{i,K-2} & \dots & v_{i,K-NC} \end{bmatrix}$$

Utilizando a proposição 1 e 2 obtêm-se

$$\hat{\theta}_{-i,K+1} = \hat{\theta}_{-i,K} + \frac{K_{-K+1}^i}{-K+1} w_{K+1}^i(\cdot) \quad i=1,2,\dots,n$$

$$\frac{K_{-K+1}^i}{-K+1} = P_K^i \xi_{-K+1} / (1 + \xi_{-K+1}^T P_K^i \xi_{-K+1}) \quad 3.57$$

$$P_{K+1}^i = P_K^i - \frac{K_{-K+1}^i}{-K+1} \xi_{-K+1}^T P_K^i$$

Como no caso do estimador dos mínimos quadrados, quando é possível considerar-se \underline{p}_0^i igual para qualquer i , obtêm-se uma sensível redução no tempo de cálculo.

Convém ressaltar que esta parametrização das matrizes $C(Z^{-1})$ e $D(Z^{-1})$ não é uma forma canônica e portanto nem sempre qualquer densidade espectral poderá ser representada através dela.

A estimação dos parâmetros de $C(Z^{-1})$ e $D(Z^{-1})$ neste modo pode ser realizada considerando apenas uma componente de $\underline{w}_K(\cdot)$ e $\underline{v}_K(\cdot)$, respectivamente sendo, neste caso, a escolha meramente subjetiva. Quando a matriz de covariância for diagonal, escolhe-se a componente que corresponde à menor variância do ruído.

3.4. Algoritmo de Estimação da Matriz Estendida

Seja um sistema descrito por:

$$A(Z^{-1}) \underline{y}_K = B(Z^{-1}) \underline{u}_K + \underline{v}_K \quad 3.58$$

$$D(Z^{-1}) \underline{v}_K = C(Z^{-1}) \underline{e}_K$$

Já foi visto na seção anterior que a equação do erro previsto é descrita por:

$$\underline{w}_K(\theta) = C^{-1}(Z^{-1}) D(Z^{-1}) \{A(Z^{-1}) \underline{y}_K - B(Z^{-1}) \underline{u}_K\} \quad 3.59$$

Como o erro previsto é uma função não linear dos parâmetros desconhecidos, deve-se realizar alguma hipótese simplificadora, para se obter uma equação explícita para o estimador.

Da equação 3.59, resulta:

$$C(Z^{-1}) \underline{w}_K(\cdot) = D(Z^{-1}) \{A(Z^{-1}) \underline{y}_K - B(Z^{-1}) \underline{u}_K\} \quad 3.60$$

Definindo $C^*(Z^{-1}) \triangleq C(Z^{-1}) - I$ e $D^*(Z^{-1}) = D(Z^{-1}) - I$ reescreve-se a equação 3.60 e obtêm-se:

$$\underline{w}_K(\cdot) + C^*(Z^{-1}) \underline{w}_K(\cdot) = A(Z^{-1}) \underline{y}_K - B(Z^{-1}) \underline{u}_K + D^*(Z^{-1}) \{A(Z^{-1}) \underline{y}_K - B(Z^{-1}) \underline{u}_K\}$$

3.61

e

$$\underline{w}_K(\cdot) = A(Z^{-1}) \underline{y}_K - B(Z^{-1}) \underline{u}_K + D^*(Z^{-1}) \underline{v}_K(\theta_{A,B}) - C(Z^{-1}) \underline{w}_K(\cdot)$$

3.62

Supondo que $\{\underline{w}_K(\cdot)\}$ e $\{\underline{v}_K(\cdot)\}$ são mensuráveis vê-se na equação 3.62 que neste caso $\underline{w}_K(\underline{\theta})$ é uma função linear dos parâmetros desconhecidos e pode-se obter um estimador aplicando o mesmo procedimento do algoritmo dos mínimos quadrados apresentado no item 3.2.

Da equação 3.62 obtêm-se, como no caso do estimador dos mínimos quadrados, que:

$$\underline{w}_K(\cdot) = \underline{y}_K - \psi_K \underline{\theta}^{A,B,C,D}$$

3.63

com

$$\underline{\theta}^{A,B,C,D} = \left[\begin{array}{c} \underline{\theta}^{A,B,C,D} \\ 1 \end{array}, \dots, \left[\begin{array}{c} \underline{\theta}^{A,B,C,D} \\ n \end{array} \right] \right]$$

$$\underline{\theta}_i^{A,B,C,D} = \left[\begin{array}{c} \underline{a}_i^T, \underline{b}_i^T, \underline{c}_i^T, \underline{d}_i^T \end{array} \right]$$

\underline{a}_i^T - elementos da iésima linha de $A(Z^{-1})$

\underline{b}_i^T - elementos da iésima linha de $B(Z^{-1})$

\underline{c}_i^T - elementos da iésima linha de $C(Z^{-1})$

\underline{d}_i^T - elementos da iésima linha de $D(Z^{-1})$

$$\psi_K = \text{bloco diagonal} \left[\begin{array}{cccc} \underline{\phi}_K^T & \underline{\phi}_K^T & \dots & \underline{\phi}_K^T \\ 1 & 2 & \dots & n \end{array} \right]$$

$$\underline{\phi}_K^T = \left[\underline{y}_{K-1}^T, \dots, \underline{y}_{K-N_A}^T, \underline{u}_{K-1}^T, \dots, \underline{u}_{K-N_B}^T, \underline{w}_{K-1}^T(\cdot), \dots, \underline{w}_{K-N_C}^T(\cdot), \underline{v}_{K-1}^T(\cdot), \dots, \underline{v}_{K-N_D}^T(\cdot) \right]$$

Definindo $p \triangleq nN_A + nN_B + nN_C + nN_D$, então $\underline{\phi}_K$ é um vetor de dimensão p , $\underline{\theta}$ é um vetor de dimensão np e a matriz ψ_K tem dimensão $n \times np$.

Como a estrutura de ψ_K é bloco diagonal podem-se aplicar as proposições 1 e 2 como no item 3.2. O algoritmo recursivo é obtido minimizando-se o critério quadrático J descrito pela equação 3.46 e reescrito a seguir:

$$J = \frac{1}{2} \sum_{t=1}^K \underline{w}_t^T(\underline{\theta}^{A,B,C}) V^{-1} \underline{w}_t(\underline{\theta}^{A,B,C})$$

O estimador recursivo é então descrito pela equação 3.64

$$\hat{\underline{\theta}}_{i,K+1}^{A,B,C,D} = \hat{\underline{\theta}}_{i,K}^{A,B,C,D} + \frac{K_{K+1}}{K+1} \{ \underline{y}_{K+1}^i - \underline{\phi}_{K+1}^T \hat{\underline{\theta}}_{i,K}^{A,B,C,D} \} \quad 3.64$$

$$\frac{K_{K+1}}{K+1} = P_K \underline{\phi}_{K+1} / (1 + \underline{\phi}_{K+1}^T P_K \underline{\phi}_{K+1}) \quad i=1, 2, \dots, n$$

$$P_{K+1} = P_K - \frac{K_{K+1}}{K+1} \underline{\phi}_{K+1}^T P_K$$

Como as perturbações $\underline{w}_K(\cdot)$ e $\underline{v}_K(\cdot)$ não são mensuráveis é necessária a sua estimação. O algoritmo da matriz estendida aplica o mesmo procedimento da seção 3.3, obtendo-se as seguintes equações para os estimadores dos ruídos:

$$\hat{\underline{v}}_K(\hat{\underline{\theta}}) = \underline{y}_K - \psi_K^{A,B} \hat{\underline{\theta}}_{K-1}^{A,B} \quad 3.65$$

e

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

com

$$\psi_K^{A,B} = \text{bloco diagonal} \left[\begin{array}{c} \phi_K^{T_{A,B}} \\ \phi_K^{T_{A,B}} \\ \dots \\ \phi_K^{T_{A,B}} \end{array} \right] \quad 3.66$$

e

$$\phi_K^{T_{A,B}} = \left[\begin{array}{c} -Y_{K-1}^T, \dots, -Y_{K-NA}^T, \\ U_{K-1}^T, \dots, U_{K-NB}^T \end{array} \right]$$

O estimador da matriz estendida representado pela equação 3.64 é bem conhecido na literatura [4, 7, 11] mas é obtido, em todos os casos, a partir da minimização de um critério que não considera a matriz de covariância do ruído. A partir das proposições 1 e 2 mostrou-se que, no caso multivariável, o estimador da matriz estendida independe da matriz de covariância e, deste modo, pode-se realizar uma partição no algoritmo e estimar-se, separadamente, cada linha do modelo, ocasionando uma sensível redução no tempo de cálculo do algoritmo.

Em Morf [4] é obtido o algoritmo da matriz estendida minimizando um critério que não considera a matriz de covariância do ruído e utilizando somente um modelo média-móvel para a perturbação. A versão monovariável deste algoritmo foi apresentada por Talmon [15], considerando um modelo autoregressivo e média móvel para o ruído.

Em relação ao algoritmo dos mínimos quadrados linearizado descrito na seção anterior este algoritmo utiliza uma aproximação do erro de previsão menos refinada. Pode-se esperar um desempenho mais baixo, mas a vantagem de um tempo de cálculo claramente inferior devido ao processo de aproximação utilizado. Convém também ressaltar que se obteve uma partição no estimador sem a utilização de nenhuma forma pseudo-canônica específica.

3.5. Algoritmo dos mínimos quadrados particionado

Este algoritmo permitirá estimar separadamente os parâmetros do processo ($\theta^{A,B}$) e os parâmetros do ruído ($\theta^{C,D}$) através de uma partição na matriz ψ_K . O algoritmo pode ser visto como uma generalização do método de mínimos quadrados com correção da polarização introduzido por Hsia [54] para sistemas SISO, que considerou somente um modelo autoregressivo para o ruído. Sua generalização para sistemas MIMO foi apresentada em Amaral [55].

Nesta seção obtêm-se o algoritmo particionado para um modelo descrito pela equação 3.58. Mostra-se que, além da partição acima mencionada pode-se aplicar as proposições 1 e 2 e consequentemente calcular-se cada linha do modelo separadamente.

O algoritmo é obtido a partir do estimador da Matriz Estendida realizando-se uma partição no estimador, dos parâmetros do processo e do ruído, apresentado na seção anterior. Devido a esta partição o algoritmo resultante apresenta uma redução no tempo de cálculo em relação ao estimador da Matriz Estendida. A grosso modo, a diminuição no tempo de cálculo corresponde à redução da dimensão da matriz P_K , pois no algoritmo da matriz estendida atualiza-se uma matriz quadrada de dimensão $p \times p$ onde $p = nN_A + mN_B + n(N_C + N_D)$, enquanto que no algoritmo particionado atualizam-se duas matrizes quadradas de dimensões $p_1 \times p_1$ e $p_2 \times p_2$ onde $p_1 = nN_A + mN_B$ e $p_2 = n(N_C + N_D)$.

Seja um sistema descrito pela equação 3.58 reescrita a seguir:

$$A(z^{-1})\underline{y}_K = B(z^{-1})\underline{u}_K + \underline{v}_K$$

$$D(z^{-1})\underline{v}_K = C(z^{-1})\underline{e}_K$$

Mostrou-se, na seção anterior, que com este modelo minimizando o critério 3.4 obtêm-se, após a aplicação das proposições 1 e 2, um estimador descrito pela seguinte equação:

$$\hat{\theta}_i^{A,B,C,D} = \left[\sum_{t=1}^K \begin{bmatrix} \phi_t^T \\ \phi_t \end{bmatrix} \right]^{-1} \left[\sum_{t=1}^K \begin{bmatrix} \phi_t^T y_t^i \\ \phi_t \end{bmatrix} \right] \quad i=1,2,\dots,n \quad 3.67$$

com

$$\underline{\phi}_K^T = \left[\begin{array}{c} -\underline{y}_{K-1}^T, \dots, -\underline{y}_{K-N_A}^T, \underline{u}_{K-1}^T, \dots, \underline{u}_{K-N_B}^T, \underline{w}_{K-1}^T(\cdot), \dots, \underline{w}_{K-N_C}^T(\cdot), \\ -\underline{v}_{K-1}^T(\cdot), \dots, -\underline{v}_{K-N_D}^T(\cdot) \end{array} \right] \quad 3.68$$

Definindo

$$\underline{\phi}_K^P = \left[-\underline{Y}_{K-1}^T, \dots, -\underline{Y}_{K-N_A}^T, \underline{u}_{K-1}^T, \dots, \underline{u}_{K-N_B}^T \right]$$

e

$$\underline{\phi}_K^R = \left[\underline{w}_{K-1}^T(\cdot), \dots, \underline{w}_{K-N_C}^T(\cdot), -\underline{v}_{K-1}^T(\cdot), \dots, -\underline{v}_{K-N_D}^T(\cdot) \right]$$

então

$$\underline{\phi}_K^T = \left[\underline{\phi}_K^P, \underline{\phi}_K^R \right]$$

Utilizando a partição da matriz $\underline{\phi}_K$ reescreve-se a equação 3.67, resultando:

$$\begin{bmatrix} \hat{\theta}_i^P \\ \hat{\theta}_i^R \end{bmatrix} = \begin{bmatrix} \hat{\theta}_i^{A,B} \\ \hat{\theta}_i^{C,D} \end{bmatrix} = \left[\begin{array}{c|c} \sum_{t=1}^K \underline{\phi}_t^P \underline{\phi}_t^{P^T} & \sum_{t=1}^K \underline{\phi}_t^P \underline{\phi}_t^{R^T} \\ \hline \sum_{t=1}^K \underline{\phi}_t^R \underline{\phi}_t^{R^T} & \sum_{t=1}^K \underline{\phi}_t^R \underline{\phi}_t^{P^T} \end{array} \right]^{-1} \begin{bmatrix} \sum_{t=1}^K \underline{\phi}_t^P Y_t^i \\ \sum_{t=1}^K \underline{\phi}_t^R Y_t^i \end{bmatrix} \quad 3.69$$

Para se obter a matriz inversa, da equação anterior, utiliza-se o lema de inversa particionada,

Lema da inversa particionada:

Dado $F = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]$, com A,B,C,D matrizes quadradas

então a matriz inversa F^{-1} é dada por

$$F^{-1} = \left[\begin{array}{c|c} A^{-1}(I+BE^{-1}CA^{-1}) & -A^{-1}BE^{-1} \\ \hline -E^{-1}CA^{-1} & E^{-1} \end{array} \right] \quad 3.70$$

com $E = D - CA^{-1}B$

Δ

Levando a expressão da matriz inversa particionada na equação do estimador, resulta:

$$\hat{\theta}_i^P = A^{-1} \sum_{t=1}^K \phi_t^P Y_t^i + A^{-1} B E^{-1} C A^{-1} \sum_{t=1}^K \phi_t^P Y_t^i - A^{-1} B E^{-1} \sum_{t=1}^K \phi_t^R Y_t^i \quad 3.71$$

e

$$\hat{\theta}_i^R = -E^{-1} C A^{-1} \sum_{t=1}^K \phi_t^P Y_t^i + E^{-1} \sum_{t=1}^K \phi_t^R Y_t^i \quad 3.72$$

sendo $A = \sum_{t=1}^K \phi_t^P \phi_t^{P^T}$; $B = \sum_{t=1}^K \phi_t^P \phi_t^{R^T}$; $C = \sum_{t=1}^K \phi_t^R \phi_t^{R^T}$ e $D = \sum_{t=1}^K \phi_t^R \phi_t^{R^T}$

Reescrevendo as equações 3.71 e 3.72, obtêm-se:

$$\hat{\theta}_i^P = A^{-1} \sum_{t=1}^K \phi_t^P Y_t^i - A^{-1} B E^{-1} \left\{ -C A^{-1} \sum_{t=1}^K \phi_t^P Y_t^i + \sum_{t=1}^K \phi_t^P Y_t^i \right\} \quad 3.73$$

e

$$\hat{\theta}_i^R = E^{-1} \left\{ -C A^{-1} \sum_{t=1}^K \phi_t^P Y_t^i + \sum_{t=1}^K \phi_t^R Y_t^i \right\} \quad 3.74$$

Substituindo os valores das matrizes A, B e do vetor de ruído $\hat{\theta}_i^R$, suposto conhecido, na equação 3.73 resulta:

$$\hat{\theta}_i^P = \left[\sum_{t=1}^K \phi_t^P \phi_t^{P^T} \right]^{-1} \left[\sum_{t=1}^K \phi_t^P Y_t^i \right] - \left[\sum_{t=1}^K \phi_t^P \phi_t^{P^T} \right]^{-1} \left[\sum_{t=1}^K \phi_t^P \phi_t^{R^T} \hat{\theta}_i^R \right] \quad 3.75$$

Nesta equação nota-se que o estimador do processo é obtido a partir do cálculo de duas parcelas. A primeira parcela da equação 3.75, corresponde ao estimador dos mínimos quadrados, obtido na seção 3.2, e a segunda parcela é um termo corretivo, correspondente à polarização do estimador dos mínimos quadrados [4, 14, 15].

Definindo o termo corretivo através da equação 3.76 como:

$$\hat{\theta}_i^C = \left[\sum_{t=1}^K \phi_t^P \phi_t^{P^T} \right]^{-1} \left[\sum_{t=1}^K \phi_t^P \phi_t^{R^T} \right] \hat{\theta}_t^R \quad 3.76$$

resulta:

$$\hat{\theta}_i^P = \hat{\theta}_i^{MQ} - \hat{\theta}_i^C \quad 3.77$$

onde $\hat{\theta}_i^{MQ}$ é o estimador dos mínimos quadrados dos parâmetros de $A(Z^{-1})$ e $B(Z^{-1})$.

Portanto para se obter o estimador dos parâmetros do processo, deve-se estimar pelo método dos mínimos quadrados os parâmetros do processo, como no item 3.2, e a seguir estimar-se os parâmetros dos ruídos para se obter o termo corretivo. Para tal reescreve-se a equação do erro previsto como segue:

$$C(Z^{-1}) \underline{w}_K(\underline{\theta}) = D(Z^{-1}) \underline{v}_K(\cdot) \quad 3.78$$

com $\underline{v}_K(\cdot)$ definido como na seção anterior.

Reescrevendo a equação 3.78 obtêm-se:

$$\underline{w}_K(\cdot) = \underline{v}_K(\cdot) - \psi_K^R \underline{\theta}^R \quad 3.79$$

$$\psi_K^R = \text{bloco diagonal} \left[\begin{array}{c} \phi_K^R \phi_K^R, \dots, \phi_K^R \end{array} \right]$$

$$\phi_K^R = \left[\begin{array}{c} w_{K-1}^T(\cdot), \dots, w_{K-N_C}^T(\cdot), -v_{K-1}^T, \dots, -v_{K-N_D}^T \end{array} \right]$$

$$\underline{\theta}^R = \left[\begin{array}{c} \theta_1^R, \theta_2^R, \dots, \theta_n^R \end{array} \right] \quad \text{e} \quad \underline{\theta}_i^R = \left[\begin{array}{c} c_i^T, d_i^T \end{array} \right]$$

c_i^T : vetor contendo os parâmetros da iésima linha de $C(Z^{-1})$

d_i^T : vetor contendo os parâmetros da iésima linha de $D(Z^{-1})$

Supondo $\underline{v}_K(\cdot)$ e $\underline{w}_K(\cdot)$ mensuráveis e minimizando o critério 3.4 obtêm-se o estimador dos parâmetros do ruído descrito pela seguinte equação:

$$\hat{\theta}_{-K}^R = \left[\sum_{t=1}^K \begin{matrix} \psi_t^R & \psi_t^{R^T} \\ \psi_t^R & \psi_t^{R^T} \end{matrix} \right]^{-1} \left[\sum_{t=1}^K \begin{matrix} \psi_t^R & \psi_t^{R^T} \\ \psi_t^R & \psi_t^{R^T} \end{matrix} \begin{matrix} v_t \\ v_t \end{matrix} \right] \quad 3.80$$

Devido a estrutura bloco diagonal das matrizes ψ_t^R , pode-se estimar cada linha do modelo separadamente e obtêm-se:

$$\hat{\theta}_{-i,K}^R = \left[\sum_{t=1}^K \begin{matrix} \phi_t^R & \phi_t^{R^T} \\ \phi_t^R & \phi_t^{R^T} \end{matrix} \right]^{-1} \left[\sum_{t=1}^K \begin{matrix} \phi_t^R & \phi_t^{R^T} \\ \phi_t^R & \phi_t^{R^T} \end{matrix} v_t^i \right] \quad i=1,2,\dots,n \quad 3.81$$

Obtêm-se a versão recursiva do algoritmo particionado aplicando-se a equação 2.12 sucessivamente nas equações do estimador dos mínimos quadrados, do estimador dos parâmetros do ruído e do termo corretivo.

$$\hat{\theta}_{-i,K+1}^{MQ} = \hat{\theta}_{-i,K}^{MQ} + \frac{1}{K+1} \{ y_{K+1}^i - \phi_{K+1}^{P^T} \hat{\theta}_{-i,K}^{MQ} \}$$

$$\frac{1}{K+1} = P_K \phi_{K+1}^P / (1 + \phi_{K+1}^{P^T} P_K \phi_{K+1}^P)$$

$$P_{K+1} = P_K - \frac{1}{K} \phi_{K+1}^{P^T} P_K$$

$$\hat{\theta}_{-i,K+1}^R = \hat{\theta}_{-i,K}^R + \frac{2}{K+1} \{ \tilde{v}_{K+1}^i - \phi_{K+1}^{R^T} \hat{\theta}_{-i,K}^R \} \quad 3.82$$

$$\frac{2}{K+1} = Q_K \phi_{K+1}^R (1 + \phi_{K+1}^{R^T} Q_K \phi_{K+1}^R)^{-1}$$

$$Q_{K+1} = Q_K - \frac{2}{K} \phi_{K+1}^{R^T} Q_K$$

$$\hat{\theta}_{-i,K+1}^C = \hat{\theta}_{-i,K}^C + \frac{1}{K+1} \{ \phi_{K+1}^{R^T} \hat{\theta}_{-i,K+1}^R - \phi_{K+1}^{P^T} \hat{\theta}_{-i,K}^C \}$$

$$\hat{\theta}_{-i,K+1}^P = \hat{\theta}_{-i,K+1}^{MQ} - \hat{\theta}_{-i,K+1}^C \quad i=1,2,\dots,n$$

As perturbações $\underline{v}_K(\cdot)$ e $\underline{w}_K(\cdot)$ são estimadas como na seção anterior, isto é, através das seguintes equações:

$$\hat{\underline{v}}_K(\hat{\underline{\theta}}) = \underline{y}_K - \psi_K^P \hat{\underline{\theta}}_{K-1}^{A,B} \quad 3.83$$

$$\hat{\underline{w}}_K(\hat{\underline{\theta}}) = \hat{\underline{v}}_K(\hat{\underline{\theta}}) - \psi_K^R \hat{\underline{\theta}}_{K-1}^{C,D} \quad 3.84$$

Nesta seção foi desenvolvido um algoritmo que particiona o algoritmo da seção anterior estimando separadamente os parâmetros do processo e do ruído acarretando num algoritmo com tempo de cálculo reduzido. Foi apresentada a versão recursiva do algoritmo e mostrou-se que as equações do estimador independem da matriz de covariância do ruído.

3.6. Estimador dos mínimos quadrados generalizado

A seguir estuda-se o algoritmo dos mínimos quadrados generalizado utilizando o modelo descrito pela equação 3.58, em vez do modelo autoregressivo considerado por Payne [17] e Keviczky [10]. Analisam-se a seguir as simplificações necessárias no modelo, para que seja possível a partição do estimador dos mínimos quadrados generalizado.

Seja a equação do erro previsto descrita por:

$$C(Z^{-1}) \underline{w}_K(\cdot) = D(Z^{-1}) \{A(Z^{-1}) \underline{y}_K - B(Z^{-1}) \underline{u}_K\}$$

Utilizando a definição de ψ_K e $\underline{\theta}^{A,B}$, do item 3.3 reescreve-se a equação do erro previsto como:

$$C(Z^{-1}) \underline{w}_K(\cdot) = D(Z^{-1}) \{\underline{y}_K - \psi_K \underline{\theta}^{A,B}\} \quad 3.85$$

Para se obter os parâmetros desconhecidos $(\underline{\theta}^{A,B,C,D})$ minimiza-se o critério descrito pela equação 3.86:

$$\min_{\underline{\theta}^{A,B}, \underline{\theta}^{C,D}} J = \min_{\underline{\theta}^{A,B}, \underline{\theta}^{C,D}} \frac{1}{2} \sum_{t=1}^K \underline{w}_K^T(\underline{\theta}^{A,B}, \underline{\theta}^{C,D}) V^{-1} \underline{w}_K(\underline{\theta}^{A,B}, \underline{\theta}^{C,D}) \quad 3.86$$

onde $\underline{\theta}^{C,D} = \underline{\theta}^R$, como na seção anterior, contém os parâmetros desconhecidos das matrizes $C(Z^{-1})$ e $D(Z^{-1})$ e $\underline{\theta}^{A,B} = \underline{\theta}^P$ os parâmetros conhecidos das matrizes $A(Z^{-1})$ e $B(Z^{-1})$.

Na equação 3.86 o critério não é uma função quadrática dos elementos $\underline{\theta}^{A,B,C,D}$ a estimar.

Entretanto, supondo que as matrizes $C(Z^{-1})$ e $D(Z^{-1})$ sejam conhecidas, isto é, relaxando a condição sobre os parâmetros do ruído, então a função de custo J será quadrática no vetor dos parâmetros do processo.

Nesse caso, da equação 3.85, obtêm-se

$$C(Z^{-1}) \underline{w}_K(\cdot) = \underline{y}_K^{F1} - \psi_K^{F1} \underline{\theta}^{A,B} \quad 3.87$$

onde

$$\underline{y}_K^{F1} = D(Z^{-1}) \underline{y}_K \quad \text{e} \quad \psi_K^{F1} = D(Z^{-1}) \psi_K$$

Reescrevendo a equação 3.87 resulta:

$$\underline{w}_K(\cdot) = \underline{y}_K^F - \psi_K^F \underline{\theta}^{A,B} \quad 3.88$$

$$\text{com} \quad C(Z^{-1}) \underline{y}_K^F = \underline{y}_K^{F1} \quad \text{e} \quad C(Z^{-1}) \psi_K^F = \psi_K^{F1}$$

e \underline{y}_K^F correspondendo ao valor filtrado da saída no instante K , ψ_K^F aos valores filtrados das medidas realizadas em instantes anteriores à K .

Substituindo o valor do erro previsto na equação 3.86 e minimizando o critério J em relação aos parâmetros $\underline{\theta}^{A,B}$, analogamente ao realizado para o estimador dos mínimos quadrados, obtêm-se:

$$\underline{\hat{\theta}}^P = \underline{\hat{\theta}}^{A,B} = \left[\sum_{t=1}^K \psi_t^F V^{-1} \psi_t^F \right]^{-1} \left[\sum_{t=1}^K \psi_t^F V^{-1} \underline{y}_t^F \right] \quad 3.89$$

Esta equação representa o estimador dos parâmetros do processo, $A(Z^{-1})$ e $B(Z^{-1})$.

Como as matrizes $C(Z^{-1})$ e $D(Z^{-1})$ não são normalmente diagonais a matriz ψ_K^F não será bloco diagonal, como a matriz ψ_K . Deste modo não se pode realizar uma partição no estimador dos parâmetros do processo, porque não se pode aplicar as proposições 1 e 2.

Para se estimar os parâmetros da perturbação, $\underline{\theta}^R = \underline{\theta}^{C,D}$ supõem-se que as matrizes $A(Z^{-1})$ e $B(Z^{-1})$ são conhecidas, isto é, relaxa-se as condições sobre os parâmetros do processo.

Definindo:

$$\underline{v}_K(\cdot) = A(Z^{-1}) \underline{y}_K - B(Z^{-1}) \underline{u}_K \quad 3.90$$

e levando a equação 3.90 na equação do erro previsto, obtêm-se:

$$C(Z^{-1}) \underline{w}_K(\cdot) = D(Z^{-1}) \underline{v}_K(\cdot) \quad 3.91$$

que pode ser reescrita como na seção anterior resultando:

$$\underline{w}_K(\cdot) = \underline{v}_K(\cdot) - \psi_K^R \underline{\theta}^R \quad 3.92$$

com $\psi_K^R = \text{bloco diagonal} \left[\begin{matrix} \phi_K^{R^T} & & \\ & \dots & \\ & & \phi_K^{R^T} \end{matrix} \right]$

e

$$\underline{\phi}_K^{RT} = \left[+ \underline{w}_{K-1}^T(\cdot), \dots, + \underline{w}_{K-N_C}^T(\cdot), - \underline{v}_{K-1}^T, \dots, - \underline{v}_{K-N_D}^T \right]$$

Deste modo o critério J é quadrático nos parâmetros desconhecidos $\underline{\theta}^{C,D}$, e obtêm-se o estimador minimizando J em relação a $\underline{\theta}^{C,D}$

$$\underline{\hat{\theta}}^R = \underline{\hat{\theta}}^{C,D} = \left[\sum_{t=1}^K \psi_t^R V^{-1} \psi_t^{RT} \right]^{-1} \left[\sum_{t=1}^K \psi_t^R V^{-1} \underline{v}_t(\cdot) \right] \quad 3.93$$

Devido a estrutura bloco diagonal de ψ_t^R pode-se aplicar as suposições 1 e 2 e obter-se uma partição para o estimador dos parâmetros do ruído.

$$\underline{\hat{\theta}}_i^R = \underline{\hat{\theta}}^{C,D} = \left[\sum_{t=1}^K \phi_t^R \phi_t^{RT} \right]^{-1} \left[\sum_{t=1}^K \phi_t^R v_{i,t}(\cdot) \right] \quad 3.94$$

$i=1,2,\dots,n$

onde $v_{i,K}$ é a i-ésima componente do vetor \underline{v}_K .

A versão recursiva deste algoritmo é obtida utilizando-se a equação 2.8, da inversa recursiva, nas equações dos estimadores $\underline{\hat{\theta}}^P$ e $\underline{\hat{\theta}}_i^R$ (equações 3.79 e 3.84 respectivamente).

$$\underline{\hat{\theta}}_{K+1}^P = \underline{\hat{\theta}}_K^P + K_{K+1} \{ \underline{y}_{K+1}^F - \psi_{K+1}^F \underline{\hat{\theta}}_K^P \}$$

$$K_{K+1} = P_K \psi_{K+1}^{FT} (V + \psi_{K+1}^F P_K \psi_{K+1}^{FT})^{-1} \quad 3.95$$

$$P_{K+1} = P_K - K_{K+1} \psi_{K+1}^F P_K$$

$$\underline{\hat{\theta}}_{i,K+1}^R = \underline{\hat{\theta}}_{i,K}^R + K_{K+1}^1 \{ v_{i,K+1} - \phi_{K+1}^R \underline{\hat{\theta}}_{i,K}^R \}$$

$$K_{K+1}^1 = Q_K \phi_{K+1}^R / (1 + \phi_{K+1}^{RT} Q_K \phi_{K+1}^R) \quad i=1,2,\dots,n \quad 3.96$$

$$Q_{K+1} = Q_K - K_{K+1}^1 \psi_{K+1}^{RT} Q_K$$

Os valores de $\underline{v}_K(\cdot)$ e $\underline{w}_K(\cdot)$ são estimados como no algoritmo dos mínimos quadrados particionados, isto é, através das equações 3.83 e 3.84.

Nesta seção foi apresentada a extensão multivariável do algoritmo dos mínimos quadrados generalizado. Foi utilizado um modelo geral autoregressivo e média-móvel para o ruído. Mostrou-se que devido a ψ_K^F não ter estrutura bloco diagonal não se pode particionar o vetor de parâmetros do processo, acarretando um maior esforço computacional no cálculo do estimador.

A versão existente na literatura [10, 17] utiliza, como é tradicional para este algoritmo, um modelo autoregressivo para o ruído. A partir da equação 3.89 nota-se que tal hipótese não possibilita a estimação separada dos parâmetros do processo.

A seguir estudam-se situações em que é possível a partição do estimador dos parâmetros do processo.

3.6.1. Algoritmo dos mínimos quadrados generalizado particionado utilizando a forma pseudo-canônica I

Neste item utiliza-se a forma pseudo-canônica I para, como na seção 3.3.1, obter-se um estimador que calcule separadamente e paralelamente as n linhas do modelo utilizado.

Considera-se, inicialmente, um sistema modelado através da forma pseudo-canônica I, isto é, pressupõem-se que o ruído pode ser modelado como um processo ARMA com $C(Z^{-1})$ e $D(Z^{-1})$ diagonais.

Já foi visto que a equação do erro previsto é dada por:

$$C(Z^{-1}) \underline{w}_K(\cdot) = \underline{y}_K^F - \psi_K^F \underline{\theta}^{A,B} \quad 3.97$$

com $\underline{y}_K^F = D(Z^{-1}) \underline{y}_K$; $\psi_K^F = D(Z^{-1}) \psi_K$ e $D(Z^{-1}) = \text{diagonal}(d^i(Z^{-1}))$

Devido a matriz $D(Z^{-1})$ ser diagonal, ψ_K^F é agora uma matriz bloco diagonal.

Desenvolvendo a equação 3.97, obtêm-se:

$$\underline{w}_K(\cdot) = \underline{Y}_K^F - \psi_K^F \underline{\theta}^{A,B} \quad 3.98$$

com $C(Z^{-1}) \underline{Y}_K^F = \underline{Y}_K^{F1}$ e $C(Z^{-1}) \psi_K^F = \psi_K^{F1}$

e $C(Z^{-1}) = \text{diagonal } (c^i(Z^{-1}))$

Devido a característica das matrizes $C(Z^{-1})$ e ψ_K^{F1} , tem-se a matriz ψ_K^F com estrutura bloco diagonal do seguinte tipo:

$$\psi_K^F = \text{bloco diagonal } \begin{bmatrix} \psi_{1,K}^{FT} & & & \\ & \psi_{2,K}^{FT} & & \\ & & \dots & \\ & & & \psi_{n,K}^{FT} \end{bmatrix} \quad 3.99$$

com

$$c^i(Z^{-1}) \psi_{i,K}^F = d^i(Z^{-1}) \psi_{i,K}^F$$

$$c^i(Z^{-1}) = 1 + c_1^i Z^{-1} + \dots + c_{N_C}^i Z^{-N_C}$$

$$d^i(Z^{-1}) = 1 + d_1^i Z^{-1} + \dots + d_{N_D}^i Z^{-N_D}$$

Supondo as matrizes $C(Z^{-1})$ e $D(Z^{-1})$ conhecidas e minimizando o critério J obtêm-se

$$\underline{\hat{\theta}}^{A,B} = \left[\sum_{t=1}^K \psi_t^{FT} V^{-1} \psi_t^F \right]^{-1} \left[\sum_{t=1}^K \psi_t^{FT} V^{-1} \underline{Y}_t^F \right] \quad 3.100$$

Aplicando a proposição 1b, e a proposição 2, com V uma matriz diagonal e ψ_K^F blocos diagonais diferentes resulta o estimador descrito pela seguinte equação:

$$\underline{\hat{\theta}}_i^P = \underline{\hat{\theta}}_i^{A,B} = \left[\sum_{t=1}^K \phi_{i,t}^F \phi_{i,t}^{FT} \right]^{-1} \sum_{t=1}^K \phi_{i,t}^F Y_t^{F^i} \quad i=1,2,\dots,n \quad 3.101$$

Colocando este estimador na forma recursiva, obtêm-se:

$$\hat{\underline{p}}_{i,K+1}^P = \hat{\underline{p}}_{i,K}^P + \underline{K}_{K+1}^i \{ \underline{Y}_{K+1}^{Fi} - \underline{\phi}_{i,K}^{FT} \hat{\underline{p}}_{i,K} \}$$

$$\underline{K}_{K+1}^i = \underline{P}_K^i \underline{\phi}_{i,K}^F / (1 + \underline{\phi}_{i,K}^{FT} \underline{P}_K \underline{\phi}_{i,K}^F) \quad 3.102$$

$$\underline{P}_{K+1}^i = \underline{P}_K^i - \underline{K}_{K+1}^i \underline{\phi}_{i,K}^{FT} \underline{P}_K \quad i=1,2,\dots,n$$

Portanto utilizando a forma pseudo-canônica I obteve-se um algoritmo generalizado que particiona o estimador dos parâmetros do processo. Como neste caso o modelo do ruído contém somente parâmetros na diagonal das matrizes polinomiais $C(Z^{-1})$ e $D(Z^{-1})$ também se pode simplificar o estimador dos parâmetros do ruído apresentado na seção anterior.

Os parâmetros do ruído são estimados supondo-se, como no item anterior, que as matrizes $A(Z^{-1})$ e $B(Z^{-1})$ são conhecidas e definindo-se:

$$\underline{v}_K(\cdot) = A(Z^{-1}) \underline{y}_K - B(Z^{-1}) \underline{u}_K \quad 3.103$$

Depois de obter-se $\underline{v}_K(\cdot)$ resulta, a partir da equação do erro previsto, que:

$$C(Z^{-1}) \underline{w}_K(\cdot) = D(Z^{-1}) \underline{v}_K(\cdot) \quad 3.104$$

Neste caso, as matrizes $C(Z^{-1})$ e $D(Z^{-1})$ são diagonais e pode-se reescrever a i-ésima componente da equação 3.104.

$$\underline{w}_{i,K}(\cdot) = \underline{v}_{i,K}(\cdot) \left[+ w(\cdot)_{i,K-1} \dots + w(\cdot)_{i,K-N_C} - v_{i,K-1} \dots - v(\cdot)_{i,K-N_D} \right] \begin{bmatrix} c_1^i \\ \vdots \\ c_{N_C}^i \\ d_1^i \\ \vdots \\ d_{N_D}^i \end{bmatrix}$$

Definindo

$$\underline{\phi}_{i,K}^{RT} \left[\begin{array}{c} -w_{i,K-1} \dots -w_{i,K-N_C} \quad v_{i,K-1} \dots v_{i,K-N_D} \end{array} \right]$$

$$\underline{\theta}_i^{RT} = \left[\begin{array}{c} c_1^i \dots c_{N_C}^i \quad d_1^i \dots d_{N_D}^i \end{array} \right] \quad \text{resulta}$$

$$w_{i,K}(\cdot) = v_{i,K}(\cdot) - \underline{\phi}_{i,K}^{RT} \cdot \underline{\theta}_i^R \quad 3.105$$

Com o valor do erro previsto, dado pela equação 3.105 e minimizando o critério J, equação 3.86, resulta o estimador da i-ésima linha do modelo da perturbação:

$$\hat{\underline{\theta}}_i^R = \left[\sum_{t=1}^K \underline{\phi}_{i,t}^R \underline{\phi}_{i,t}^{RT} \right]^{-1} \left[\sum_{t=1}^K \underline{\phi}_{i,t}^R v_{i,t}(\cdot) \right] \quad i=1,2,\dots,n$$

Como nos estimadores anteriores pode-se obter versão recursiva do estimador dos parâmetros do ruído, descrito por:

$$\hat{\underline{\theta}}_{i,K+1}^R = \hat{\underline{\theta}}_{i,K}^R + \underline{K}_{K+1}^{1,i} (v_{i,K+1} - \underline{\phi}_{i,K+1}^{RT} \hat{\underline{\theta}}_{i,K}^R) \quad 3.106$$

$$\underline{K}_{K+1}^{1,i} = Q_K^i \underline{\phi}_{i,K+1}^R / (1 + \underline{\phi}_{i,K+1}^{RT} Q_K^i \underline{\phi}_{i,K+1}^R) \quad i=1,2,\dots,n$$

$$Q_{K+1}^i = Q_K^i - \underline{K}_{K+1}^{1,i} \underline{\phi}_{i,K+1}^{RT} Q_K^i$$

3.6.2. Algoritmo dos mínimos quadrados generalizado particionado utilizando a forma canônica III

Como já foi mostrado na seção 3.3.2, a forma canônica III de Kashyap permite reescrever de uma forma conveniente o modelo geral descrito pela seguinte equação:

$$A(z^{-1}) \underline{y}_K = B(z^{-1}) \underline{u}_K + \underline{v}_K$$

$$D(z^{-1}) \underline{v}_K = C(z^{-1}) \underline{e}_K$$

Introduzindo o modelo do ruído na primeira equação ob-
têm-se:

$$D(z^{-1}) A(z^{-1}) \underline{y}_K = D(z^{-1}) B(z^{-1}) \underline{u}_K + C(z^{-1}) \underline{e}_K \quad \text{cov } \underline{e}_K = V$$

Utilizando a forma canônica III de Kashyap resulta:

$$A^*(z^{-1}) \underline{y}_K = B^*(z^{-1}) \underline{u}_K + C^*(z^{-1}) \underline{e}_K^* \quad \text{cov } (e_K^*) = V^*$$

onde $A^*(z^{-1})$, $B^*(z^{-1})$, $C^*(z^{-1})$ e V^* tem as características defini-
das na seção 1.3.

Com $C^*(z^{-1})$ e V^* diagonais o algoritmo dos mínimos qua-
drados generalizado particionado é descrito pela equação 3.102
com $D(z^{-1})=I$ nas equações de filtragem das medidas (equações 3.97,
3.98 e 3.99).

Nesta seção não se fez nenhuma hipótese restritiva so-
bre o modelo do ruído porque se utiliza uma forma canônica, ao
contrário da seção anterior em que a parametrização das matrizes
 $C(z^{-1})$ e $D(z^{-1})$ nem sempre representa uma aproximação válida.

No algoritmo dos mínimos quadrados generalizado parti-
cionado das seções 3.6.1 e 3.6.2 observa-se que o ganho \underline{K}^i é di-
ferente para cada linha do modelo do processo. Na seção seguinte
utiliza-se a forma pseudo-canônica II que resulta num ganho único
para todas as linhas.

3.6.3. Algoritmo dos mínimos quadrados generalizado particio- nado utilizando a forma pseudo-canônica II

Para se obter um algoritmo que permita a estimação li-

na por linha dos parâmetros do processo, no estimador generalizado, a matriz ψ_K^F deve ser bloco diagonal igual. Como ψ_K é bloco diagonal igual, vê-se das equações 3.87 e 3.88 que se obtém a estrutura desejada para ψ_K^F utilizando-se a forma pseudo-canônica II, onde as matrizes $C(Z^{-1})$ e $D(Z^{-1})$ são diagonais iguais.

Com estas hipóteses pode-se rearranjar a equação do erro previsto resultando:

$$C(Z^{-1}) \underline{w}_K(\cdot) = D(Z^{-1}) \{ \underline{y}_K - \psi_K \underline{\theta}^{A,B} \} \quad 3.107$$

com as simplificações das matrizes $C(Z^{-1})$ e $D(Z^{-1})$ obtém-se

$$\underline{w}_K(\cdot) = \underline{y}_K^F - \psi_K^F \underline{\theta}^{A,B} \quad 3.108$$

com $\psi_K^F =$ bloco diagonal $\left[\begin{array}{cccc} \phi_{K1}^{F^T} & \phi_{K2}^{F^T} & \cdots & \phi_{Kp}^{F^T} \end{array} \right]$

e $\phi_{Kj}^{F^T} = \left[\begin{array}{cccc} -\underline{y}_{K-1}^{F^T} & \cdots & -\underline{y}_{K-N_A}^{F^T} & \underline{u}_{K-1}^{F^T} \cdots \underline{u}_{K-N_B}^{F^T} \end{array} \right]$

$$c(Z^{-1}) \underline{y}_K^F = d(Z^{-1}) \underline{y}_K$$

e

$$c(Z^{-1}) \underline{u}_K^F = d(Z^{-1}) \underline{u}_K$$

Minimizando a função de custo J (equação 3.86), obtém-se o estimador simplificado descrito pela seguinte equação:

$$\hat{\underline{\theta}}^{A,B} = \left[\sum_{t=1}^K \psi_t^{F^T} V^{-1} \psi_t^F \right]^{-1} \left[\sum_{t=1}^K \psi_t^{F^T} V^{-1} \underline{y}_t^F \right] \quad 3.109$$

Entretanto, agora ψ_t^F tem uma estrutura de blocos diagonais iguais e aplicando as proposições 1 e 2 obtém-se:

$$\hat{\theta}_i^{A,B} = \left[\sum_{t=1}^K \phi_t^F \phi_t^{F^T} \right]^{-1} \left[\sum_{t=1}^K \phi_t^F y_t^{F^i} \right] \quad i=1,2,\dots,n \quad 3.110$$

A versão recursiva deste algoritmo é descrita pela equação 3.111.

$$\hat{\theta}_{i,K+1}^{A,B} = \hat{\theta}_{i,K}^{A,B} + K_{K+1} \left\{ y_{K+1}^{F^i} - \phi_{K+1}^T \hat{\theta}_{i,K}^{A,B} \right\}$$

$$K_{K+1} = P_K \phi_{K+1}^F / (1 + \phi_{K+1}^{F^T} P_K \phi_{K+1}^F) \quad 3.111$$

$$P_{K+1} = P_K - K_{K+1} \phi_{K+1}^T P_K \quad i=1,2,\dots,n$$

Estas simplificações foram introduzidas por Morf [4] no caso particular de $C(Z^{-1})=I$. Com a hipótese simplificadora da matriz $D(Z^{-1})$ Morf obteve o algoritmo descrito pela equação 3.111, com $C(Z^{-1})=I$, minimizando um critério não ponderado, como descrito pela equação 2.18, do capítulo anterior.

Convém ressaltar que esta parametrização das matrizes $C(Z^{-1})$ e $D(Z^{-1})$ não é uma forma canônica e portanto nem sempre representará uma aproximação válida.

Realiza-se a estimação dos parâmetros de $C(Z^{-1})$ e $D(Z^{-1})$ considerando-se apenas uma das saídas, sendo que em Morf [4] não é proposta nenhuma escolha particular.

Neste caso a equação do erro previsto é dada pela seguinte equação:

$$C(Z^{-1}) \underline{w}_K(\cdot) = D(Z^{-1}) \underline{v}_K(\cdot) \quad 3.112$$

com $C(Z^{-1}) = c(Z^{-1}) I$

$$D(Z^{-1}) = d(Z^{-1}) I$$

No caso da matriz de covariância diagonal escolhe-se a componente com uma menor variância, isto é, supondo que a i ésima componente da medida é a de menor variância reescreve-se a i ésima linha da equação 3.112 como:

$$w_{i,K}(\cdot) = v_{i,K} + d_1 v_{i,K-1} + \dots + d_{N_D} v_{i,K-N_D} - c_1 w_{i,K-1} - \dots - c_{N_C} w_{i,K-N_C} \quad 3.113$$

$$w_{i,K}(\cdot) = v_{i,K} - \underline{\phi}_K^{R^T} \underline{\theta}^R \quad 3.114$$

com $\underline{\phi}_K^{R^T} = \left[w_{i,K-1}, w_{i,K-2}, \dots, w_{i,K-N_C}, -v_{i,K-1}, \dots, -v_{i,K-N_D} \right]$

$$\underline{\theta}^{R^T} = \left[c_1, \dots, c_{N_C}, d_1, \dots, d_{N_D} \right]$$

Obtém-se o estimador para os parâmetros do ruído minimizando o critério J_1 , descrito pela equação 3.115:

$$J_1 = \sum_{t=1}^K ||w_{i,t}||^2 \quad 3.115$$

A partir da minimização do critério J resulta:

$$\hat{\underline{\theta}}^R = \left[\sum_{t=1}^K \underline{\phi}_t^R \underline{\phi}_t^{R^T} \right]^{-1} \left[\sum_{t=1}^K \underline{\phi}_t^R v_{i,t} \right] \quad 3.116$$

A versão recursiva do estimador dos parâmetros do ruído é descrito pela seguinte equação:

$$\hat{\underline{\theta}}_{K+1}^R = \hat{\underline{\theta}}_K^R + \underline{K}_{K+1}^1 \{ v_{i,K} - \underline{\phi}_{K+1}^{R^T} \hat{\underline{\theta}}_K^R \}$$

$$\underline{K}_{K+1}^1 = Q_K \underline{\phi}_{K+1}^R / (1 + \underline{\phi}_{K+1}^{R^T} Q_K \underline{\phi}_{K+1}^R) \quad 3.117$$

$$Q_{K+1} = Q_K - \underline{K}_{K+1}^1 \underline{\phi}_{K+1}^T Q_K \quad i=1, 2, \dots, n$$

Quando a matriz de covariância não é diagonal não se sabe, a priori, qual informação deve ser escolhida. Para resolver este problema, propõem-se estimar o ruído, através de um novo método, que leve em consideração todas as medidas realizadas.

Para tanto, deve-se reescrever a equação do erro previsto como segue adiante.

$$\underline{w}_K(\cdot) = \underline{v}_K(\cdot) - \psi_K^R \underline{\theta}^R \quad 3.118$$

com

$$\psi_K^R = \left[-w_{K-1}, \dots, -w_{K-N_C}, v_{K-1}, v_{K-2}, \dots, v_{K-N_D} \right]$$

$$\underline{\theta}^{R^T} = \left[c_1, \dots, c_{N_C}, d_1, \dots, d_{N_D} \right]$$

Analogamente aos casos estudados anteriormente minimizando o critério J descrito pela equação 3.86, e supondo as matrizes $A(z^{-1})$ e $B(z^{-1})$ conhecidas, obtém-se o estimador para o vetor de parâmetros desconhecidos $\underline{\theta}^R$:

$$\hat{\underline{\theta}}^R = \left[\sum_{t=1}^K \psi_t^T V^{-1} \psi_t \right]^{-1} \left[\sum_{t=1}^K \psi_t^T V^{-1} \underline{v}_t \right] \quad 3.119$$

Entretanto, neste caso, devido a estrutura de ψ_K não ser bloco diagonal, não se aplicam as proposições 1 e 2 e consequentemente não se reduz o tempo de cálculo deste algoritmo.

A versão recursiva deste algoritmo é descrita pela equação 3.120 onde se nota a necessidade da inversão de uma matriz de ordem n , o que aumenta o esforço computacional deste estimador.

$$\hat{\theta}_{K+1}^R = \hat{\theta}_K^R + K_{K+1}^R \{ v_{K+1} - \psi_{K+1} \hat{\theta}_K^R \}$$

$$K_{K+1}^R = Q_K \psi_{K+1}^T R_{K+1}^{-1}$$

$$R_{K+1} = V + \psi_{K+1} P_K \psi_{K+1}^T$$

3.120

$$Q_{K+1} = Q_K - Q_K \psi_{K+1}^T R_{K+1}^{-1} \psi_{K+1} Q_K$$

Observa-se que, neste caso, não se pode particionar o estimador dos parâmetros do ruído. Deste modo este algoritmo somente será interessante quando a ordem do modelo do ruído não for elevada.

3.7. Conclusões

Neste capítulo foi analisado o problema da partição dos algoritmos de estimação com o objetivo de diminuir o tempo de cálculo envolvido. Mostrou-se que a partição depende essencialmente do modelo utilizado, da matriz de covariância e do algoritmo de estimação. Os diferentes algoritmos foram obtidos através da minimização de uma função de custo quadrática do erro de previsão incluindo a matriz de covariância.

Quando o ruído pode ser modelado como um ruído branco, mostrou-se que o estimador dos mínimos quadrados leva diretamente a uma estrutura particionada independente da matriz de covariância do ruído.

Quando o ruído não é branco a obtenção de uma forma explícita para o estimador supõe a construção de uma aproximação linear para o erro de previsão.

Os métodos analisados para obter tal aproximação levaram aos seguintes algoritmos: i) mínimos quadrados linearizado, ii) matriz estendida, iii) mínimos quadrados particionado e iv) mínimos quadrados generalizado.

O algoritmo dos *mínimos quadrados linearizado* foi desenvolvido utilizando um modelo ARMA para o ruído. Mostrou-se que a partição é possível utilizando a forma canônica III de Kashyap. Conseguiu-se uma maior simplificação no cálculo do ganho do estimador em sistemas com uma estrutura particular, utilizando a forma pseudo-canônica II.

Generalizou-se o algoritmo da *matriz estendida* para o caso multivariável com modelo ARMA para o ruído.

Neste caso obteve-se diretamente um algoritmo particionado no qual não influi a matriz de covariância do ruído.

O algoritmo dos *mínimos quadrados generalizado* foi estendido para modelos ARMA no ruído, mostrando-se que a partição a penas é possível quando se utiliza a forma canônica III ou a forma pseudo-canônica I.

Maiores simplificações foram obtidas, para sistemas com estrutura particular, utilizando a forma pseudo-canônica II.

Por último foi desenvolvido o algoritmo dos *mínimos quadrados particionado*, partindo do algoritmo da matriz estendida e realizando uma partição suplementar para estimar separadamente os parâmetros do processo e os parâmetros do ruído. A nova partição acarretou uma sensível redução no tempo de cálculo envolvido.

No Capítulo 4 será apresentado um estudo comparativo dos diversos algoritmos utilizando exemplos de simulação.

CAPÍTULO 4

4. ANÁLISE DOS ALGORITMOS ATRAVÉS DE EXEMPLOS DE SIMULAÇÃO

4.1. Introdução

Neste capítulo apresentam-se alguns dos resultados obtidos através de exemplos de simulação que ilustram o comportamento dos algoritmos desenvolvidos no capítulo anterior.

A seção 4.2 contém uma apresentação dos critérios utilizados para a comparação dos diferentes algoritmos. Na seção 4.3 apresentam-se os métodos utilizados para a estimação do tempo de cálculo e memória requerida em cada um dos estimadores.

A seção 4.5 contém resultados que ilustram o comportamento das extensões para modelos ARMA apresentadas no capítulo 3.

Inicialmente analisa-se a situação em que o modelo ARMA para o ruído tem uma representação exata nas formas autoregressiva (AR) e média-móvel (MA). Os resultados mostram o comportamento esperado para as extensões propostas, equivalentes para cada algoritmo, independentemente do modelo utilizado para o ruído.

A seguir utiliza-se um modelo ARMA para o ruído para o qual não existe equivalência exata nas formas AR e MA. Este exemplo permite ilustrar o interesse na utilização de um modelo ARMA, quando a estrutura do modelo não é conhecida, dada a diminuição no número de parâmetros a estimar.

A seção 4.6 contém os resultados obtidos com o novo algoritmo linearizado apresentado no capítulo 3. Utiliza-se a forma canônica III de Kashyap para permitir uma estimação particionada dos parâmetros. Os resultados desta simulação são comparados com os resultados obtidos com outros métodos que também permitem a estimação particionada (Matriz Estendida, algoritmo dos mínimos quadrados particionado) mostrando-se o seu melhor desempenho. O esforço computacional envolvido é contudo muito maior que o dos outros algoritmos, devido à ordem elevada do modelo.

Finalmente, faz-se a análise dos resultados de estimação de um modelo experimental do nível de aço líquido, no molde de uma máquina de lingotamento contínuo de uma usina siderúrgica. Este exemplo evidencia, numa situação real, o interesse da utili

zação do estimador dos mínimos quadrados linearizado, visto que, em geral, a ordem dos modelos utilizados é baixa.

4.2. Critérios para comparação de desempenho dos algoritmos de estimação

A comparação do desempenho experimental dos diferentes algoritmos tem como base os resultados obtidos a partir de simulações de Monte Carlo com médias sobre dez realizações. Para fins de comparação utiliza-se o cálculo da função de custo acumulado descrito pela equação 4.1.

$$J_{K,i} = \sum_{t=1}^K \underline{w}_t^T(\hat{\theta}) V^{-1} \underline{w}_t(\hat{\theta}) \quad 4.1$$

e

$$\bar{J}_K = \frac{1}{MC} \sum_{i=1}^{MC} J_{K,i}$$

onde MC é o número de realizações de Monte Carlo, $\underline{w}_t(\hat{\theta})$ é o erro de previsão no instante t, descrito por

$$\underline{w}_t(\hat{\theta}) \triangleq y_t - \hat{y}_t(\hat{\theta}_{t-1})$$

e V a matriz de covariância do ruído \underline{e}_K , suposto estacionário.

A função de custo acumulado será comparada com a função de custo acumulado teórica descrita pela equação 4.2:

$$J_K^T = \sum_{t=1}^K E \{ \underline{e}_t^T V^{-1} \underline{e}_t \} \quad 4.2$$

Sendo $\{\underline{e}_t\}$ uma sequência branca de variáveis aleatórias gaussianas, com matriz de covariância V, a equação 4.2 pode ser reescrita na forma

$$J_K^T = \sum_{t=1}^K \text{traço } V^{-1} E \{ \underline{e}_t^T \underline{e}_t \} = n.K \quad 4.3$$

A equação 4.3 mostra que a função de custo acumulado teórico é uma reta com inclinação proporcional ao número de saídas do sistema (n).

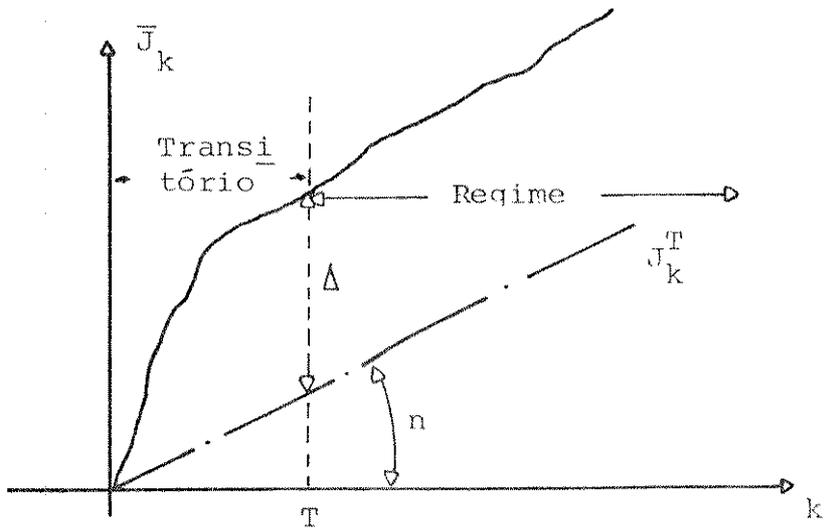


Figura 4.1 - Função de custo

A função de custo acumulado para os diferentes exemplos de simulação será representada como indicado na figura 4.1. O comportamento em regime, de cada algoritmo, poderá ser avaliado, comparando-se a inclinação do custo acumulado teórico para valores de K suficientemente elevados. Quando as curvas se tornam paralelas, o comportamento em regime do estimador, atinge o desempenho ótimo teórico.

O valor absoluto do custo acumulado depende do comportamento transitório do algoritmo. A diferença Δ entre as curvas, calculada num certo instante adequado T (Figura 4.1), é uma medida do comportamento transitório do algoritmo, sendo que, quanto menor for esta distância melhor será o desempenho do algoritmo no período inicial de estimação.

Os algoritmos também são analisados através do cálculo da distância paramétrica normalizada do processo [56] definida pela seguinte equação:

$$\delta_{K,i}^2 = \sum_{i=1}^r \left[\frac{\Delta \theta_K^i}{\theta_K^i} \right]^2$$

4.4

$$\delta_K^{-2} = \frac{1}{MC} \sum_{i=1}^{MC} \delta_{K,i}^2$$

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 r o número de parâmetros estimados do processo (vetor $\hat{\theta}^{AB}$ formado pelos parâmetros das matrizes polinomiais $A(z^{-1})$ e $B(z^{-1})$ do processo).

A variável δ_K^{-2} é, em particular, interessante na análise da influência do modelo do ruído na estimação do vetor de parâmetros $\theta^{A,B}$

4.3. Comparação da eficiência de cálculo dos estimadores

4.3.1. Tempo de cálculo dos estimadores

Como já foi descrito no capítulo inicial, estudam-se algoritmos de estimação com o objetivo de aplicá-los em esquemas de controle adaptativo como mostrado na figura 4.2.

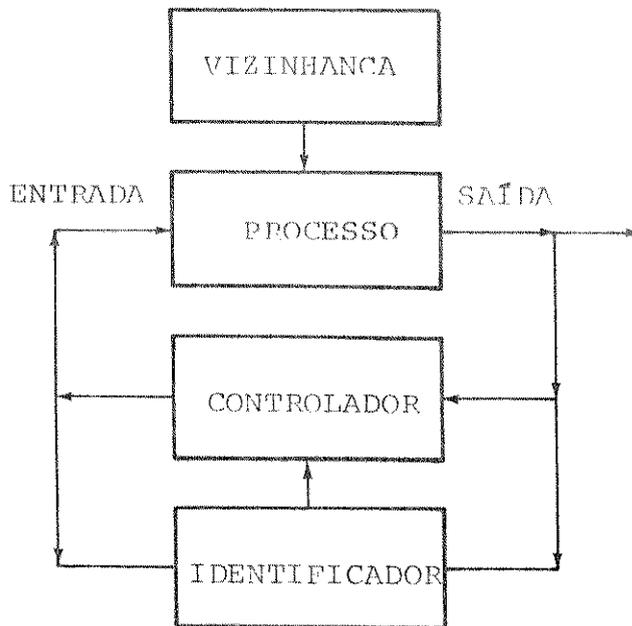


Figura 4.2 - Esquema de Controle Adaptativo.

Portanto, para o estabelecimento dos tempos de amostragem do sistema, bem como do tipo de computador necessário para a execução do estimador é importante conhecer o tempo de cálculo de um ciclo completo de cada algoritmo.

Como se viu no capítulo 3 os algoritmos deduzidos são obtidos pela realização das seguintes operações matriciais: adição, subtração, transposição, inversão, multiplicação e multiplicação por escalar.

As subrotinas para execução de cada uma destas operações, podem ser realizadas a partir de um conjunto de instruções básicas, cujos tempos de execução são conhecidos e expressos numa unidade básica de tempo do computador.

TABELA 4.1

INSTRUÇÕES E TEMPO DE EXECUÇÃO	
Instrução	tempo de execução
Add to A	2
Sub from A	2
Load A or B	2
Store A or B	2
Mul A with memory	MUL
Div A by memory	DIV
Mark place and transfer	2
Return branch	2
Transfer on zero or minus	1
Transfer unconditional	1
Transfer on plus	1
Increment A	1
Increment index register R_i	1
Decrement index register R_i	1
Add to index register R_i	1
Sub to index register R_i	1
Increment B register	1
Skip on index register zero	1

J.M. Mendel [58] descreve subrotinas em linguagem de

máquina para a execução das operações citadas. Estas subrotinas utilizam o conjunto de instruções da tabela 4.1 que contém os tempos correspondentes de execução. A partir dessa tabela o tempo de computação pode ser calculado em função das dimensões das matrizes envolvidas.

Os tempos de execução das instruções de multiplicação e divisão dependem fortemente do modo como tais instruções são implementadas e do tipo de computador utilizado. Para o computador PDP-10 esse tempo é dado por MUL=4.4µS, DIV=7.4µS e a unidade de tempo para o mesmo é de .11µS [59].

Os algoritmos deduzidos no capítulo anterior, podem ser descritos, de uma maneira geral, pelas equações:

$$\begin{aligned} \hat{\underline{\theta}}_{K+1} &= \hat{\underline{\theta}}_K + K_{K+1} \{ y_{K+1} - \phi_{K+1}^T \hat{\underline{\theta}}_K \} \\ K_{K+1} &= P_K \phi_{K+1} \cdot \left[\Lambda + \phi_{K+1}^T P_K \phi_{K+1} \right]^{-1} \\ P_{K+1} &= P_K - K_{K+1} \phi_{K+1}^T P_K \end{aligned} \quad 4.5$$

onde ϕ_{K+1} e Λ dependem do algoritmo utilizado. Convém ressaltar ainda que certos algoritmos (como o dos mínimos quadrados generalizado) são constituídos por dois conjuntos de equações do tipo da equação 4.5.

Partindo das equações do estimador (equação 4.5), das dimensões das variáveis envolvidas e utilizando-se a tabela 4.2 obtém-se o tempo necessário para o cálculo de um ciclo completo do algoritmo. Obviamente este tempo depende do número de entradas e saídas do sistema e das ordens dos polinômios $A(Z^{-1})$, $B(Z^{-1})$, $C(Z^{-1})$ e $D(Z^{-1})$.

Para ilustrar como evolui o tempo de cálculo por ciclo de cada algoritmo, supõe-se, sem perda de generalidade, que o número de entradas é igual ao número de saídas do sistema ($n=m$).

TABELA 4.2

OPERAÇÕES COM MATRIZES E TEMPO TOTAL DE COMPUTAÇÃO		
Operação		tempo total de computação (em unidades de tempo)
Soma de matrizes	$C_{MN} = A_{MN} + B_{MN}$	MUL + 27 + 7MN
Subtr. de matrizes	$C_{MN} = A_{MN} - B_{MN}$	MUL + 27 + 7MN
Produto de matrizes	$C_{ML} = A_{MN} B_{NL}$	10 + 8MNL + 19ML + 16M + MUL(MNL)
Produto pela matriz transposta	$C_{MN} = A_{ML} (B_{NL}^T)$	10 + 8MNL + 19NL + 16M + MUL(MNL)
Inversão de matriz	$A_{NN} \rightarrow A_{NN}^{-1}$	10 + 7,5N ⁴ + 43N ³ + 139,5N ² + 92N + + DIV(2N ² +N) + MUL(N ³ +0,5N ² +2,5N)
Produto de matriz por escalar	$C_{N1} = \mu A_{N1}$	8N + MUL(N)

A tabela 4.3 apresenta os valores do tempo de processamento, por ciclo, para um computador de médio porte (PDP-10) e para o micro computador INTEL 8086 munido do processador de ponto flutuante INTEL 8232. Neste caso, tem-se MUL=24µS e DIV=28µS.

Observa-se a partir da tabela 4.3 que os algoritmos que estimam separadamente cada linha do modelo (M.E. e M.Q.P.) apresentam um tempo de cálculo reduzido quando comparados com os algoritmos M.Q.L. e M.Q.G.

A partir dos tempos de cálculo destes algoritmos também se verifica que estes dois últimos estimadores são mais sensíveis à variação no número de saídas do que à variação na ordem das matrizes polinomiais. Isto ocorre porque a dimensão da matriz P_K nestes algoritmos é proporcional ao quadrado do número de saídas. Este problema é sensivelmente minimizado quando se utilizam as formas pseudo-canônicas como mostra a tabela 4.3, para o algoritmo M.Q.L. O mesmo comportamento é obtido com a forma canônica III de Kashyap, que não restringe a estrutura do modelo utilizado.

ALGORITMO		Tempo de cálculo em ms														
ORDEM	N _D	n	ME		MQL		MQL FPCI		MQL FPCII		MQP		MQG		MQL FCIII	
			PDP10	INTEL 8086	PDP 10	INTEL 8086	PDP10	INTEL 8086	PDP10	INTEL 8086	PDP10	INTEL 8086	PDP10	INTEL 8086	PDP10	INTEL 8086
1	0	2	1,0	4,0	7,5	32,0	1,3	5,2	0,75	2,9	0,87	3,0	3,5	14,0	1,8	7,1
1	1	2	1,6	6,5	13,0	55,0	1,8	7,1	1,0	3,9	1,2	4,5	3,8	15,0	-	-
2	2	2	5,6	23,0	48,0	213,0	6,5	25,0	3,3	13,0	3,4	14,0	12,0	51,0	5,3	21
2	2	3	12,0	50,0	351	1,6(s)	17,0	65,0	5,8	23,0	7,3	30,0	75,0	330	14	58
3	3	3	26,0	108,0	783	3,5(s)	35,0	142,0	12,0	50,0	15,0	61,0	163	725	30	120

Tabela 4.3 - Tempos de cálculo para os algoritmos

- Matriz Estendida - ME
- Mínimos Quadrados Linearizados - MQL
- Mínimos Quadrados Linearizado com forma pseudo-canônica I - MQL FPCI
- Mínimos Quadrados Linearizado com forma pseudo-canônica II - MQL FPCII
- Mínimos Quadrados Particionado - MQP
- Mínimos Quadrados Generalizado - MQG
- Mínimos Quadrados Linearizado com forma canônica III - MQL FCIII

Da tabela anterior também se observa que o algoritmo MQL é o estimador que apresenta maior tempo de processamento. Entretanto, quando aplicado a sistemas de ordem baixa, este valor não é excessivamente elevado, sendo viável a sua utilização em algoritmos de controle "on-line".

Uma maior redução, nos tempos de cálculo dos estimadores, pode ser obtida considerando a simetria da matriz P_K .

4.3.2. Memória

A memória requerida para os algoritmos é, analogamente ao tempo de cálculo, função das dimensões das matrizes θ, P, ϕ, K . Além dessas variáveis devem ser armazenadas as subrotinas para operações com matrizes e previsto espaço para a memória de trabalho.

A tabela 4.4 indica a quantidade de memória, requerida para cada algoritmo, em função do número de saídas, e da ordem do sistema.

Algoritmo	Memória (palavra)
ME	$120 + n^2 \left[2N_G^2 + N_G + 2N_G/n \right]$
M.Q.L.	$348 + n^4 \left[2N_G^2 + \frac{2N_G}{n} + \frac{N_G}{n^3} \right]$
M.Q.G.	$348 + n^2 \left[2N_P^2 + \frac{2N_P^2}{n} + \frac{N_P}{n^2} + \frac{2N_R^2}{n^2} + \frac{N_R}{n^2} + \frac{2N_R}{n^3} \right]$
M.Q.P.	$120 + n^2 \left[2(N_P^2 + N_R^2) + N_P + N_R + \frac{2}{n} \cdot (N_P + N_R) \right]$
M.Q.L. FPC1	$120 + n \left[nN_P + N_R \right]^2 + nN_P^2 + N_R^2 + 2nN_P + 2N_R + n^2 N_P + nN_R$
M.Q.L. FPC2	$120 + 2 \left[nN_P + N_R \right]^2 + n^2 N_P + nN_R + 2nN_P + 2N_R$

Tabela 4.4 - Memória

n = número de saídas

$$N_G = N_A + N_B + N_C + N_D$$

$$N_P = N_A + N_B$$

$$N_R = N_C + N_D$$

Na tabela 4.5 indica-se a memória requerida em dois casos particulares.

ALGORITMO ORDEM	Memória em K palavras						
	$N_A=N_B=N_C=N_D=n$	ME	MQL	MQL FPCI	MQL FPCII	MQL FCIII	MQP
2	0,7	2,5	0,5	0,4	0,5	0,4	1,3
3	3,0	24,0	2,1	1,4	2,1	1,6	9,0

Tabela 4.5 - Memória

Destas tabelas verifica-se que os algoritmos M.Q.L. e M.Q.G. são mais sensíveis à variação no número de saídas do que os algoritmos particionados, devido à quantidade de memória requerida nestes dois algoritmos ser uma função da quarta potência de n . Este problema, entretanto, é sensivelmente minimizado quando se utilizam as formas pseudo-canônicas, ou a forma canônica III de Kashyap, que permitem a estimação particionada dos parâmetros desconhecidos. Convém ressaltar que neste último caso não se está fazendo nenhuma hipótese restritiva sobre o modelo do ruído.

4.4. Geração das medidas e inicialização

4.4.1. Geração das medidas

Obtém-se as medidas \underline{y}_K por meio de um modelo multivariável do tipo ARMA para o ruído, em que a entrada \underline{u}_K e a perturbação \underline{e}_K são geradas independentemente. O ruído \underline{e}_K é uma sequência branca de variáveis aleatórias gaussianas com média zero e matriz de covariância V . Para a entrada \underline{u}_K utiliza-se um ruído branco de média zero, onde a variância de cada componente da entrada é escolhida de acordo com o valor desejado para a relação sinal-ruído das saídas.

A relação sinal-ruído da i ésima componente da saída é a relação entre a variância da i ésima entrada observada ($u_{i,K}^*$) e a variância da correspondente componente do ruído ($v_{i,K}$) [57], onde

$$u_{i,K}^* = b_i(z^{-1}) \underline{u}_K$$

e $b_i(z^{-1})$ é a i ésima linha da matriz polinomial $B(z^{-1})$, definida no capítulo 3.

Nas simulações deste capítulo utiliza-se sempre uma relação sinal-ruído igual à unidade.

4.4.2. Inicialização

Para todos os algoritmos devem ser fornecidas inicializações dos valores estimados $\hat{\theta}(0)$ e da matriz $P(0)$. A matriz $P(0)$ deve refletir o grau de conhecimento "à priori" dos valores iniciais $\hat{\theta}(0)$. Nos exemplos deste capítulo escolhem-se os valores iniciais dos estimadores iguais à zero e a matriz $P(0) = 100.I$, onde I é a matriz de identidade.

Para que sejam satisfeitas as hipóteses simplificadores utilizados na seção 3.3. na dedução do algoritmo M.Q.L. utiliza-se este estimador sem filtragem durante as primeiras cinquenta medidas.

Nos algoritmos M.Q.P. e M.Q.G. calculam-se parâmetros do ruído somente após as primeiras cinquenta medidas, isto é, durante as primeiras cinquenta medidas os resultados destes algoritmos são idênticos ao do estimador M.Q.

4.5. Estimadores com modelos ARMA para o ruído

Normalmente, nos trabalhos publicados na área, cada algoritmo de identificação é utilizado somente com um certo tipo de modelo (autoregressivo ou média-móvel). As extensões propostas no capítulo anterior para utilização de modelos ARMA, em geral, diminuem o tempo de cálculo devido ao menor número de parâ-

metros a serem estimados.

Nesta seção utiliza-se inicialmente um exemplo em que a representação ARMA do ruído tem equivalente autoregressivo e média-móvel exatos. A fim de se obter tais representações para o ruído, utiliza-se um modelo ARMA que envolve somente matrizes unimodulares. Os resultados obtidos mostram comportamentos equivalentes para cada algoritmo independentemente do modelo utilizado para o ruído.

A seguir escolhe-se um exemplo em que o modelo ARMA do ruído não tem equivalentes exatos para as estruturas média-móvel e autoregressiva. Ilustra-se, assim, a superioridade do modelo ARMA porque, em geral, a utilização de outra estrutura dá origem a um acréscimo no número de parâmetros a serem estimados e conseqüentemente um aumento no tempo de cálculo do algoritmo considerado.

4.5.1. Equivalentes exatos para o modelo ARMA

Seja um sistema descrito pela equação 4.6

$$A(Z^{-1})y_K = B(Z^{-1})u_K + v_K \quad 4.6$$

$$D(Z^{-1})v_K = C(Z^{-1})e_K$$

com

$$A(Z^{-1}) = I + \begin{bmatrix} .8 & -2.9 \\ .4 & -1.4 \end{bmatrix} Z^{-1}$$

$$B(Z^{-1}) = IZ^{-1}$$

$$C(Z^{-1}) = I + \begin{bmatrix} .2 & -.4 \\ .1 & -.2 \end{bmatrix} Z^{-1} \quad v = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$D(Z^{-1}) = I + \begin{bmatrix} -.4 & .8 \\ -.2 & .4 \end{bmatrix} Z^{-1}$$

As matrizes $C(Z^{-1})$ e $D(Z^{-1})$ foram escolhidas como matrizes unimodulares para que possam existir equivalentes exatos nas formas autoregressiva e média-móvel (equações 4.7 e 4.9).

$$\text{MA} \quad C'(Z^{-1}) = I + \begin{bmatrix} .6 & -1.2 \\ -3 & -.6 \end{bmatrix} Z^{-1} \quad 4.7$$

$$\text{AR} \quad D'(Z^{-1}) = I + \begin{bmatrix} -.6 & 1.2 \\ -.3 & 6 \end{bmatrix} Z^{-1} \quad 4.8$$

Nesta seção e na seguinte são apresentados os resultados das simulações dos algoritmos dos mínimos quadrados (M.Q.), dos mínimos quadrados linearizado (M.Q.L.), da matriz estendida (ME), dos mínimos quadrados particionado (M.Q.P.) e dos mínimos quadrados generalizado (M.Q.G.) descritos nas seções 3.2, 3.3, 3.4, 3.5 e 3.6 respectivamente.

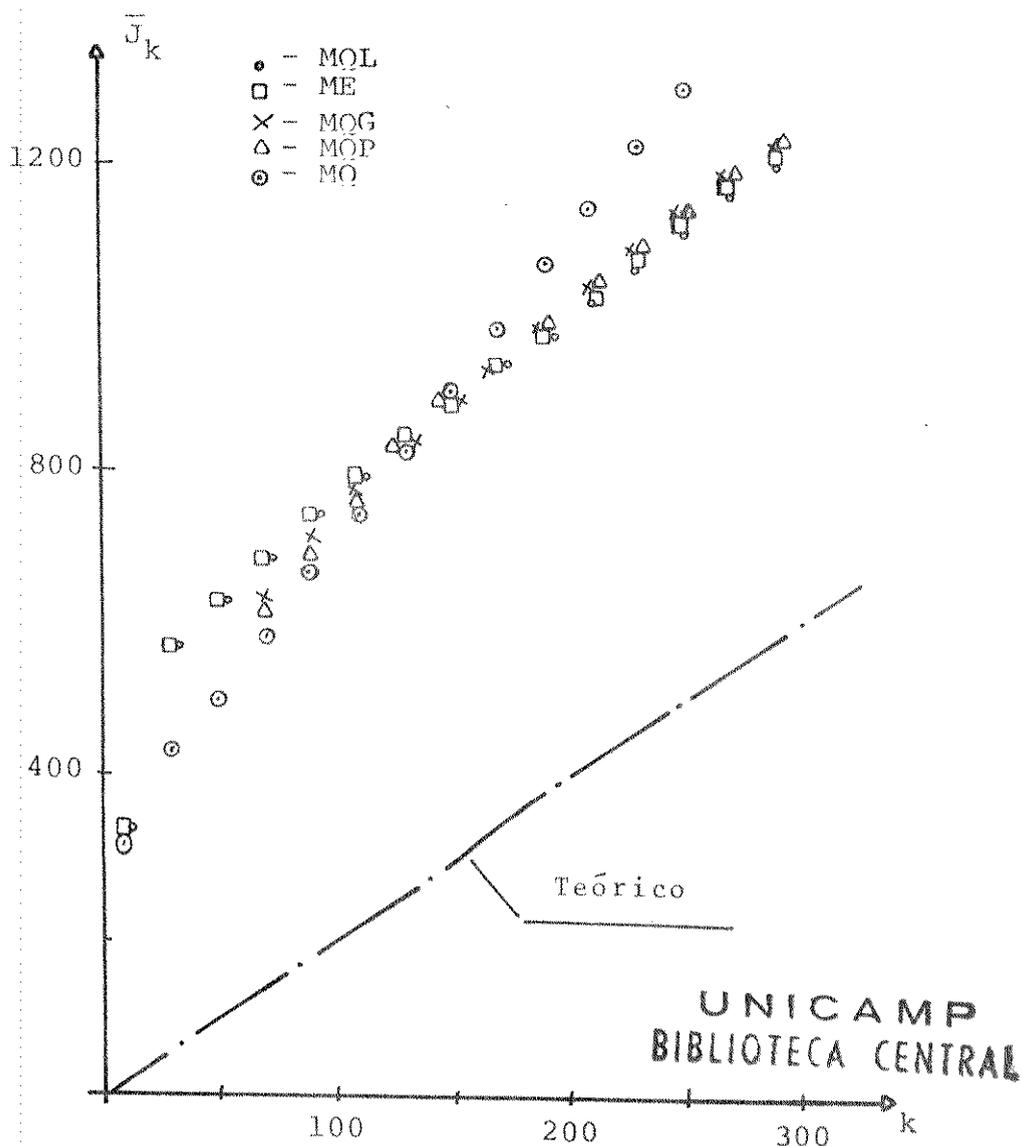


Figura 4.3 - Evolução da função de custo utilizando um modelo ARMA para o ruído.

A figura 4.3 mostra a evolução do critério $\bar{J}_K(\underline{\theta})$ para os diferentes algoritmos desenvolvidos no capítulo 3, quando se utiliza um modelo ARMA para o ruído. A reta tracejada representa a curva do custo teórico definido pela equação 4.3.

A tabela 4.6 apresenta o valor do custo acumulado após 600 medidas quando se utiliza um modelo ARMA para o ruído. Nesta tabela $\sigma_{\bar{J}_{600}}$ representa o desvio padrão da média do custo acumulado após 600 medidas e J_{600}^{\min} e J_{600}^{\max} representam respectivamente os valores mínimo e máximo do custo final obtido nas dez realizações de Monte Carlo.

Algoritmo	\bar{J}_{600}	$\sigma_{\bar{J}_{600}}$	J_{600}^{\min}	J_{600}^{\max}
M.Q.	2668	218	2373	3115
M.E.	1891	230	1586	2235
M.Q.P.	1931	200	1673	2366
M.Q.G.	1939	233	1653	2330
M.Q.L.	1886	230	1568	2227

Tabela 4.6 - Custo acumulado após 600 medidas

Os resultados apresentados na figura 4.3 e na tabela 4.6 mostram um comportamento em regime, em relação ao erro previsto, semelhante para os algoritmos que estimam os parâmetros do ruído. Da figura 4.3 verifica-se, como era esperado, que estes algoritmos apresentam um melhor comportamento do que o estimador dos mínimos quadrados. Em particular, após o transitório, todas as curvas de custo acumulado apresentam praticamente a mesma inclinação, sensivelmente idêntica à inclinação da curva de custo teórico. Estes resultados ilustram o bom comportamento dos algoritmos desenvolvidos para modelos ARMA.

Na tabela 4.7 apresentam-se as mesmas variáveis da tabela 4.6 quando se varia a estrutura do modelo utilizado. Nesta tabela δ representa a distância paramétrica definida na seção 4.2.1.

Algoritmo	\bar{J}_{600}	$\sigma_{\bar{J}_{600}}$	J_{\max}^{\min}	J_{600}^{\max}	δ
Mínimos Quadrados	2668	218	2373	3115	0,343
ME ARMA	1891	230	1586	2235	0,091
ME MA	1824	192	1604	2232	0,094
ME AR	1829	232	1596	2338	0,092
MQP ARMA	1931	200	1673	2366	0,150
MQP MA	1901	198	1653	2321	0,142
MQP AR	1931	202	1673	2366	0,150
MQG ARMA	1939	233	1653	2330	0,260
MQG MA	1914	210	1615	2213	0,260
MQG AR	1852	177	1610	2165	0,220
MQL ARMA	1886	230	1568	2227	0,115
MQL MA	1805	193	1583	2229	0,120
MQL AR	1820	200	1539	2328	0,120

Tabela 4.7 - Valores do custo acumulado para os diferentes algoritmos quando se varia a estrutura do modelo.

A tabela 4.7 mostra que a função de custo em $K=600$ atinge, para cada algoritmo, valores semelhantes independentemente do tipo de modelagem do ruído. Este fato reflete que cada algoritmo atinge o mesmo desempenho quando a ordem do modelo é correta. Desta forma se ilustra que todas as extensões realizadas nos algoritmos deduzidos no capítulo anterior são válidas.

Verifica-se ainda que não é necessário utilizar-se um algoritmo de identificação para cada tipo de modelo, como ocorre usualmente na literatura.

Os resultados mostram que o algoritmo da matriz estendida com o modelo AR, ou com um modelo MA utilizado na literatura, apresenta o mesmo comportamento. Este fato é especialmente importante pois, como visto na tabela 4.3, a utilização exclusiva de modelos AR no algoritmo M.Q.G. acarreta tempos de cálculo elevados quando comparados com os do algoritmo ME.

O maior valor para a distância paramétrica no estimador M.Q.G. é um resultado normalmente encontrado na literatura quando a relação sinal-ruído é baixa [42].

4.5.2. Modelos ARMA sem equivalentes exatos

Nesta seção utiliza-se um exemplo em que o modelo ARMA para o ruído não tem equivalentes exatos para as estruturas média-móvel e autoregressiva. Isto é, os correspondentes modelos AR e MA para o ruído têm uma ordem infinita. Como não se deseja utilizar um modelo de ordem muito elevada, devido o aumento no tempo de cálculo é necessário realizar um truncamento, o que originará, em geral, uma degradação do desempenho do estimador.

O exemplo utilizado nesta seção tem a mesma estrutura do exemplo da seção 4.5.1. mas com as matrizes polinomiais $A(Z^{-1})$, $B(Z^{-1})$, $C(Z^{-1})$ e $D(Z^{-1})$ descritas por:

$$A(Z^{-1}) = \begin{bmatrix} 1 + 0,4Z^{-1} & 0 \\ 0 & 1 + 0,6Z^{-1} \end{bmatrix} \quad B(Z^{-1}) = IZ^{-1}$$

$$C(Z^{-1}) = \begin{bmatrix} 1 - 0,2Z^{-1} & 0 \\ 0 & 1 - 0,1Z^{-1} \end{bmatrix} \quad 4.9$$

$$D(Z^{-1}) = \begin{bmatrix} 1 + 0,9Z^{-1} & 0 \\ 0 & 1 + 0,8Z^{-1} \end{bmatrix}$$

$$V = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

Com as matrizes $C(Z^{-1})$ e $D(Z^{-1})$ descritas na equação 4.9 obtêm-se os seguintes modelos do ruído na forma autoregressiva e média-móvel.

Modelo média-móvel

$$C'(Z^{-1})=D^{-1}(Z^{-1})C(Z^{-1})=\begin{bmatrix} 1-0,82Z^{-1}+0,656Z^{-2}-0,525Z^{-3}+\dots & 0 \\ 0 & 1-0,91Z^{-1}+0,819Z^{-2}-0,737Z^{-3}+\dots \end{bmatrix} \quad 4.10$$

Modelo autoregressivo

$$D'(Z^{-1})=C^{-1}(Z^{-1})D(Z^{-1})=\begin{bmatrix} 1+0,91Z^{-1}+0,0091Z^{-2}+\dots & 0 \\ 0 & 1+0,82Z^{-1}+0,0164Z^{-2}+\dots \end{bmatrix} \quad 4.11$$

Das equações 4.10 e 4.11 verifica-se que o modelo ARMA do ruído pode ser, neste exemplo, aproximado por um modelo autoregressivo de menor ordem do que o modelo média-móvel. Neste caso o modelo média-móvel necessita de uma ordem elevada para aproximar o sistema descrito pela equação 4.9.

As figuras 4.4 a 4.7 mostram a evolução do critério $\bar{J}_K(\hat{\theta})$ para os diferentes algoritmos analisados.

A tabela 4.8 apresenta os valores do custo acumulado após 600 medidas. Também são apresentados os valores da distância paramétrica do processo (δ) para análise da influência do modelo do ruído na estimação dos parâmetros $\theta^{A,B}$.

A partir dos resultados numéricos apresentados na tabela 4.8 e da evolução do custo acumulado para os diferentes algoritmos (Figuras 4.4 a 4.7) podem ser feitas as seguintes observações:

1) Verifica-se que em todos os estimadores obtêm-se vantagens tanto no tempo de cálculo como no desempenho do algoritmo, quando se utiliza uma estrutura ARMA para o ruído em relação à estrutura média-móvel. Explica-se esse comportamento, considerando-se que a ordem do modelo média-móvel deve ser muito elevada

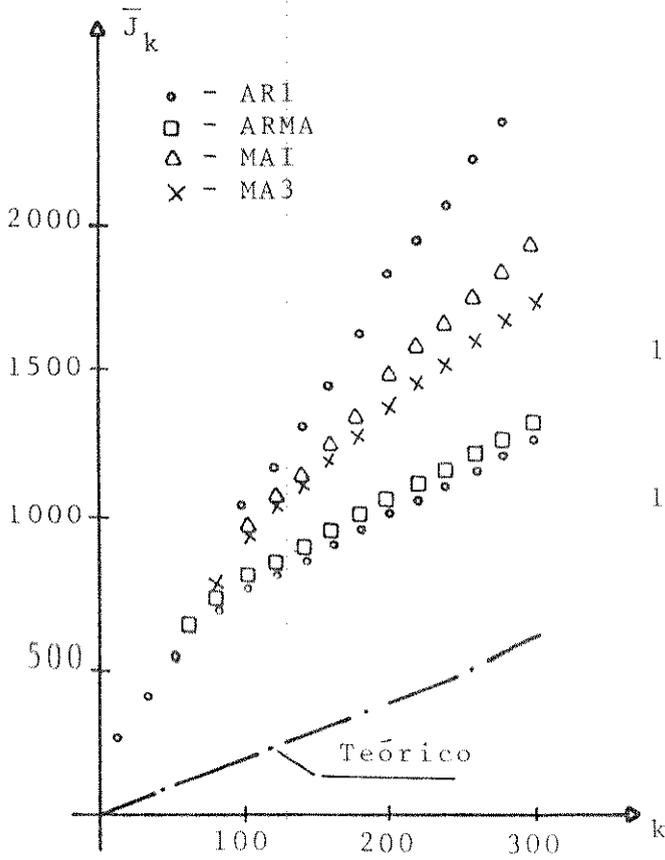


Figura 4.4 - Custo acumulado para o estimador M.Q.P.

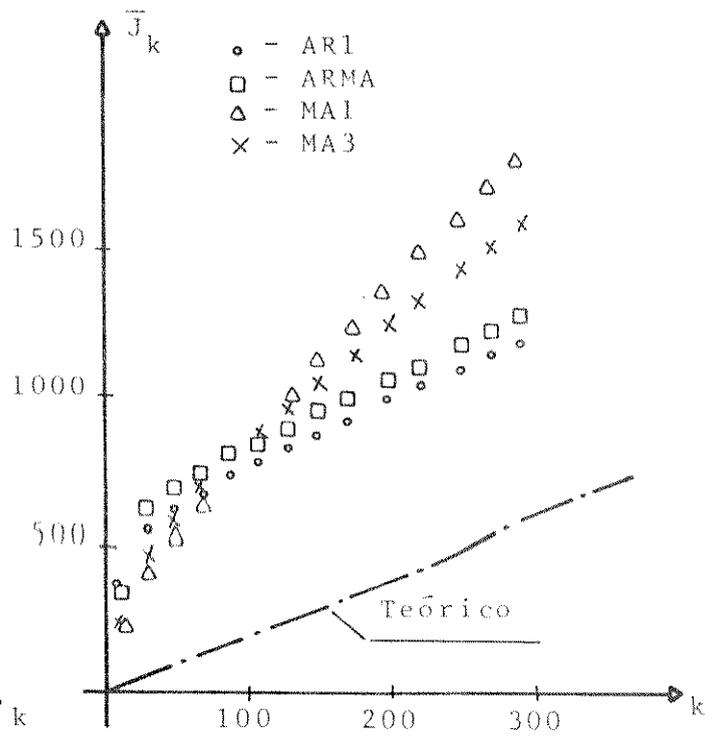


Figura 4.5 - Custo acumulado para o estimador M.E.

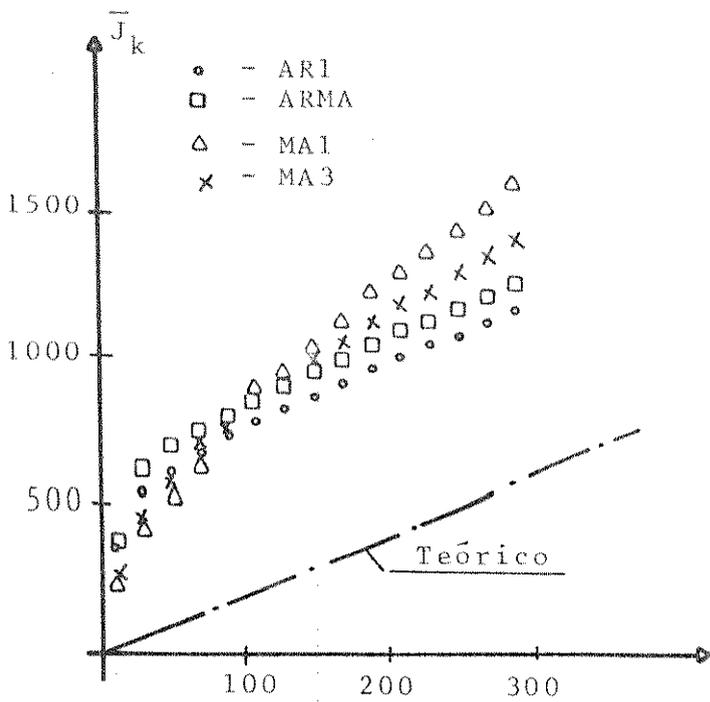


Figura 4.6 - Custo acumulado para o estimador M.Q.L.

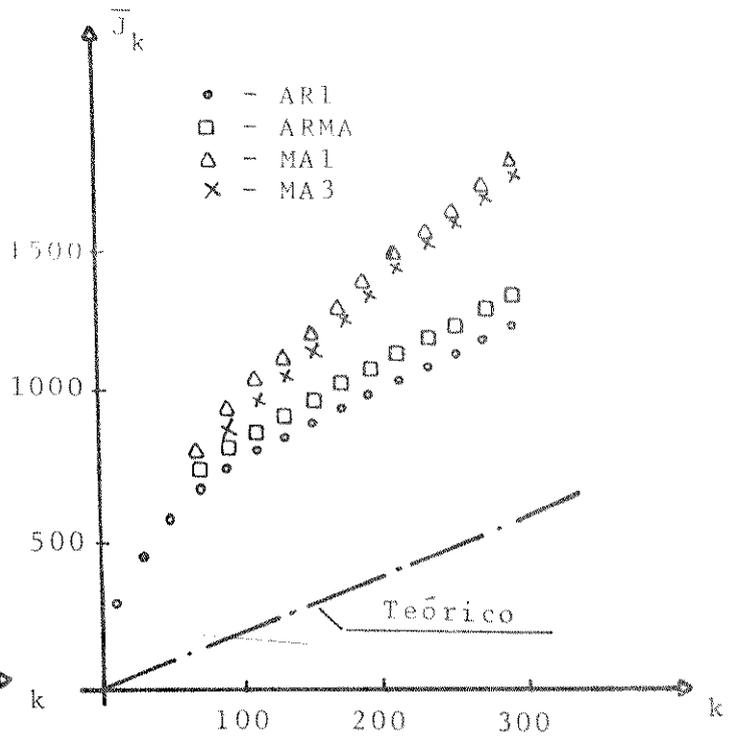


Figura 4.7 - Custo acumulado para o estimador M.Q.G.

Modelo	\bar{J}_{600}	$\sigma_{\bar{J}_{600}}$	J_{600}^{\min}	J_{600}^{\max}	δ	T (ms)
Mínimos Quadrados	4910	810	4031	6244	0,652	0,52
ME ARMA	1978	337	1635	2653	0,208	1,62
ME MA1	3456	468	2914	4255	0,424	1,00
ME MA3	2735	385	2304	3364	0,299	2,4
ME AR1	1916	365	1574	2807	0,204	1,00
MQP ARMA	2106	307	1714	2857	0,315	1,1
MQP MA1	3530	530	2904	4391	0,438	0,87
MQP MA3	2870	450	2394	3667	0,373	1,70
MQP AR1	2060	292	1668	2734	0,312	0,87
MQG ARMA	2095	313	1654	2878	0,297	3,8
MQG MA1	3353	510	2846	4285	0,391	3,5
MQG MA3	3060	359	2576	3630	0,430	4,3
MQG AR1	1977	333	1605	2875	0,264	3,8
SQL ARMA	1915	264	1611	2405	0,223	12,0
SQL MA1	2961	319	2521	3599	0,400	7,5
SQL MA3	2349	223	2060	2691	0,286	19,0
SQL AR1	1828	289	1567	2562	0,205	7,5

Tabela 4.8 - Valores do custo acumulado e do tempo de cálculo para os diferentes algoritmos quando se varia a estrutura do modelo.

T - tempo de cálculo por ciclo.

para que a aproximação do sistema seja razoável. Desde modo a utilização de um modelo MA de ordem três, que necessita de um tempo de cálculo aproximadamente 1,5 vezes maior do que um modelo ARMA, apresenta sempre um maior custo acumulado e uma maior distância paramétrica nos parâmetros do processo.

2) Os algoritmos utilizando um modelo ARMA apresentam comportamentos semelhantes aos estimadores com modelo ARL, devido a este modelo aproximar satisfatoriamente o modelo ARMA do ruído. O modelo AR apresenta um melhor desempenho transitório devido ao menor número de parâmetros a serem estimados. Além disso necessita somente da estimação da perturbação \underline{v}_K , cálculo que é realizado com os valores estimados de $A(Z^{-1})$ e $B(Z^{-1})$ (equação 3.46 do capítulo 3). É bem conhecido na literatura [12, 22] que estes valores convergem mais rapidamente do que os parâmetros do ruído. A tabela 4.7 também mostra que, em virtude do modelo ARL aproximar satisfatoriamente o modelo ARMA, a distância paramétrica nos parâmetros do processo é praticamente a mesma quando se utilizam estas duas estruturas para o ruído.

3) Os algoritmos ME e M.Q.P. apresentam um desempenho satisfatório quando utilizados com modelo AR (Figuras 4.5 e 4.6 respectivamente). Verifica-se, mais uma vez, a possibilidade de utilizarmos estes algoritmos, extremamente rápidos em modelos AR.

4) O algoritmo M.Q.G. (figura 4.7) não apresenta uma melhora significativa no custo acumulado quando se varia a ordem do modelo com estrutura média-móvel de um para três. Este comportamento pode ser explicado a partir dos resultados apresentados na seção 3.6 onde foi mostrado que a utilização de um modelo MA acarreta, para o estimador M.Q.G., na filtragem de novas variáveis (\underline{y}^F e \underline{u}^F) que são calculadas recursivamente utilizando os valores estimados dos parâmetros de $C(Z^{-1})$. Neste caso, esses valores são de baixa precisão devido ao truncamento efetuado.

Nas figuras 4.8 e 4.9 compara-se o custo acumulado, para todos os algoritmos em estudo, utilizando respectivamente modelos ARMA e ARL para o ruído.

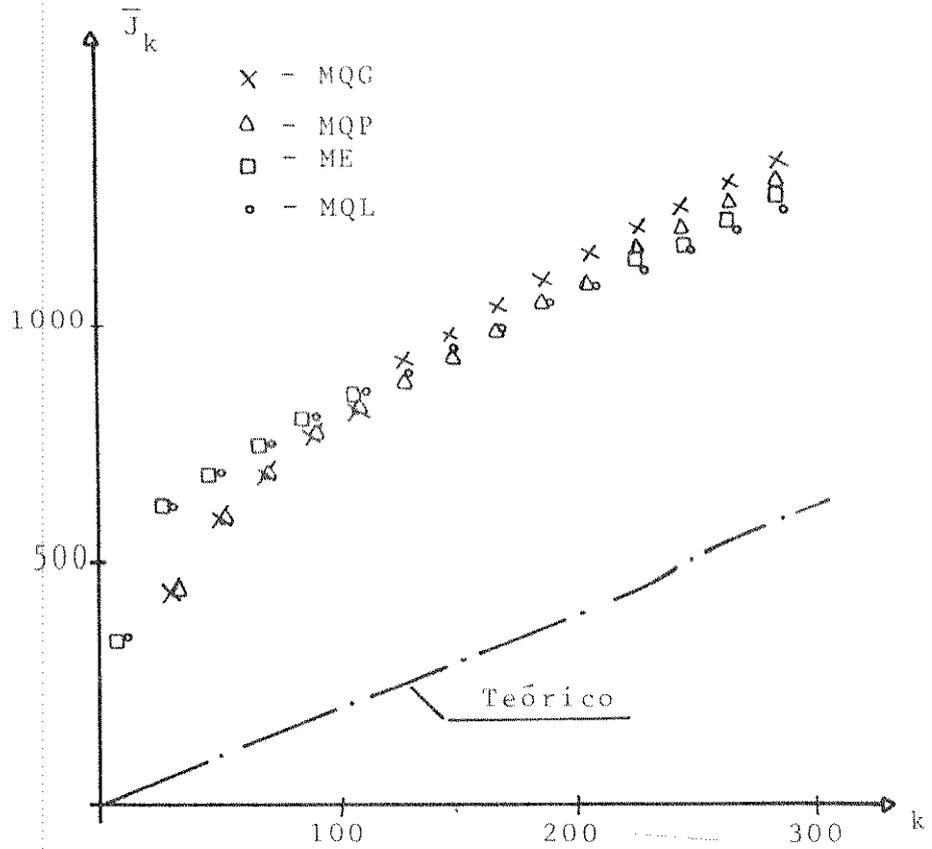


Figura 4.8 - Custo acumulado utilizando estrutura ARMA

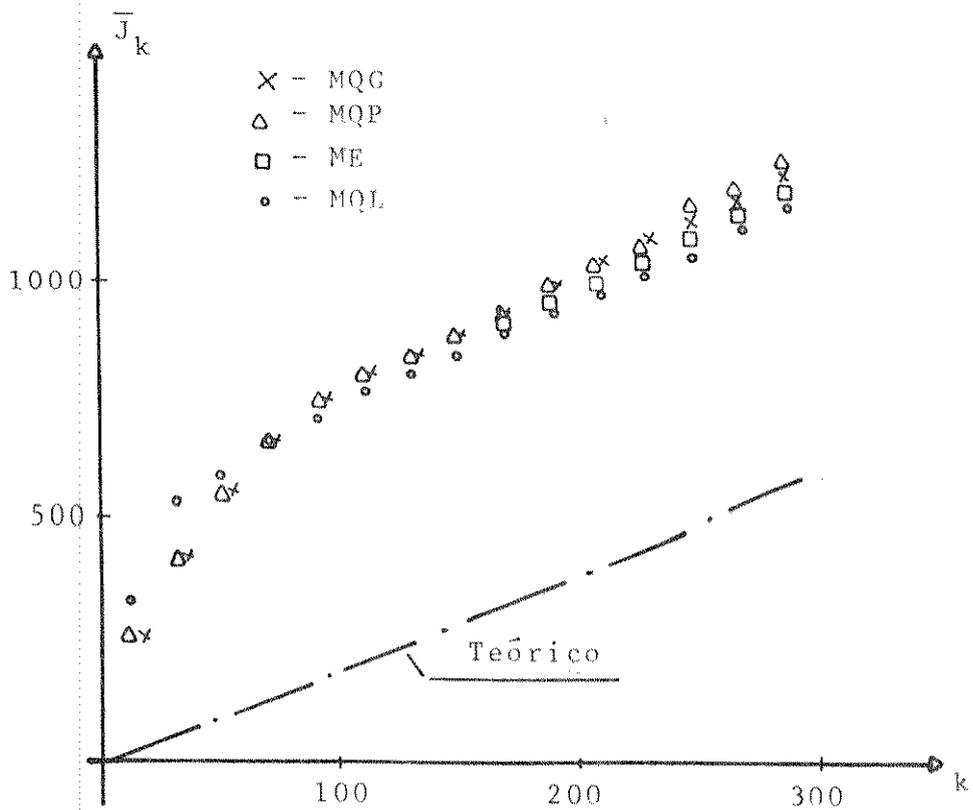


Figura 4.9 - Custo acumulado utilizando modelo ARI

Nestas figuras verifica-se que, quando a estrutura do sistema for conhecida, todos os estimadores apresentam praticamente o mesmo comportamento em regime. A escolha de um desses algoritmos poderá portanto ser feita levando em consideração principalmente a eficiência computacional (tempo de cálculo e memória ocupada).

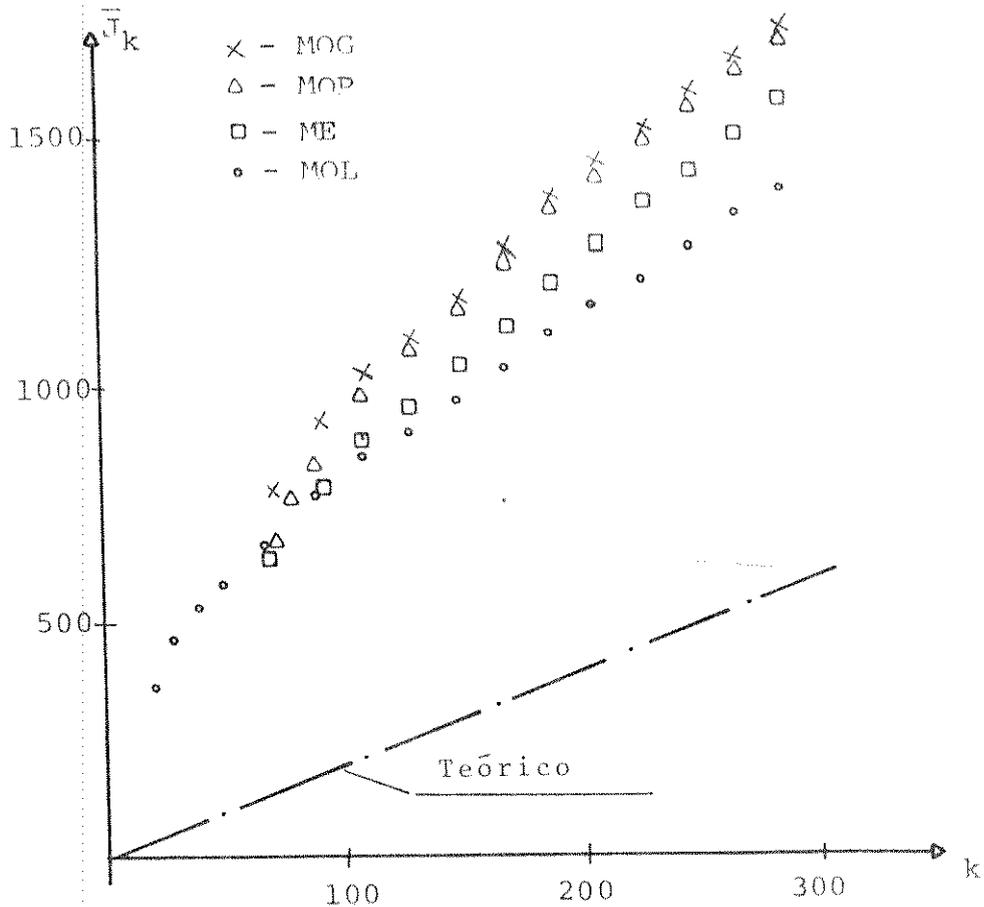


Figura 4.10 - Custo acumulado para o modelo MA3

Na figura 4.10 comparam-se os custos acumulados quando se utiliza um modelo MA3.

Destes resultados numéricos conclui-se que, numa situação de modelo com truncamento, o algoritmo M.Q.L. apresenta um menor custo acumulado do que os outros estimadores. Este resultado já era esperado porque, como foi mostrado na seção 3.3, este estimador é obtido através de uma aproximação mais refinada na equação do erro previsto, que é linearizado em torno do último valor estimado, enquanto que nos algoritmos ME e M.Q.P. o

erro previsto é diretamente escolhido como linear nos parâmetros desconhecidos.

4.6. Algoritmo M.Q.L. com forma canônica III de Kashyap

Nesta seção apresentam-se os resultados obtidos com o estimador M.Q.L. desenvolvido no capítulo 3, utilizando a forma canônica III de Kashyap, que permite particionar o problema de estimação e conseqüentemente aumentar a eficiência computacional. Os resultados são comparados com aqueles obtidos na simulação dos algoritmos ME e M.Q.P., estimadores esses que também permitem o cálculo particionado dos parâmetros desconhecidos.

No item seguinte apresentam-se alguns dos resultados obtidos na estimação de parâmetros de um modelo experimental do nível de aço líquido, no molde de uma máquina de lingotamento con tínuo. Esta aplicação demonstra a utilidade do estimador M.Q.L., devido a melhora sensível na previsão da altura de aço no molde, com um tempo de cálculo só levemente superior ao do estimador da ME, em virtude da ordem baixa dos modelos utilizados.

4.6.1. Comparação do algoritmo M.Q.L. com os algoritmos ME e M.Q.P.

Neste item ilustra-se a utilização do algoritmo M.Q.L. com a forma canônica III de Kashyap, que permite a estima ção particionada dos parâmetros desconhecidos.

Seja o sistema descrito pela equação 4.12

$$A(Z^{-1})Y_K = B(Z^{-1})u_{-K} + C(Z^{-1})e_{-K} \quad 4.12$$

com

$$A(Z^{-1}) = I + \begin{bmatrix} 0,8 & 0,0 \\ 0,2 & 0,4 \end{bmatrix} Z^{-1}; \quad B(Z^{-1}) = IZ^{-1}$$

$$C(Z^{-1}) = I + \begin{bmatrix} 0 & 0,5 \\ -0,5 & -1,0 \end{bmatrix} Z^{-1}; \quad v = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

Utilizando o algoritmo apresentado em Kashyap [31] para a construção da forma canônica III, obtêm-se a seguinte apresentação:

$$A'(Z^{-1})Y_K = B'(Z^{-1})u_K + C'(Z^{-1})e_K \quad 4.13$$

$$A'(Z^{-1}) = \begin{bmatrix} 1,0 & 0,0 \\ -1,0 & 1,0 \end{bmatrix} + \begin{bmatrix} -2,0 & -0,5 \\ 0,9 & 0,9 \end{bmatrix} Z^{-1} + \begin{bmatrix} -0,9 & -0,2 \\ 1,3 & 0,2 \end{bmatrix} Z^{-2}$$

$$B'(Z^{-1}) = \begin{bmatrix} 1,0 & 0,0 \\ -1,0 & 1,0 \end{bmatrix} Z^{-1} + \begin{bmatrix} -1,0 & -0,5 \\ 0,5 & 0,0 \end{bmatrix} Z^{-2}$$

$$C'(Z^{-1}) = I + \begin{bmatrix} -1,0 & 0,0 \\ 0,0 & -1,0 \end{bmatrix} Z^{-1} + \begin{bmatrix} 0,25 & 0 \\ 0 & 0,25 \end{bmatrix} Z^{-2}$$

$$V' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Este exemplo ilustra o caso geral em que a utilização da forma canônica III de Kashyap aumenta o número de parâmetros a serem estimados.

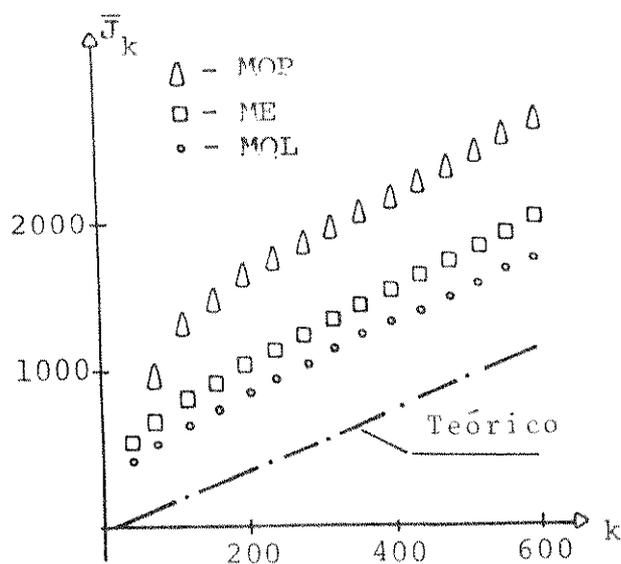


Figura 4.11 - Custo acumulado
algoritmos M.Q.L., M.E. e M.Q.P.

Na figura 4.11 comparam-se os custos acumulados quando se utilizam os algoritmos M.E., M.Q.P. e o estimador M.Q.L. com a forma canônica III de Kashyap.

Estes resultados evidenciam um melhor comportamento transitório do estimador M.Q.L. e desempenho em regime análogo para os três algoritmos. O estimador M.Q.L. apresenta um melhor comportamento transitório devido à sua aproximação mais refinada para a equação do erro previsto. Neste exemplo, o tempo de cálculo para o algoritmo M.Q.L. é aproximadamente oito vezes maior do que para o algoritmo M.E., independentemente da utilização ou não da forma canônica III. Este fato é devido ao aumento no número de parâmetros quando se utiliza a forma canônica, o que elimina a redução de tempo de cálculo originada pela estimação particionada.

Como já foi visto na seção anterior (figura 4.10) o algoritmo M.Q.L. pode fornecer bons resultados com modelos de ordem mais baixa devido a aproximação realizada. Este fato será ilustrado na próxima seção.

4.6.2. Identificação de um sistema de lingotamento contínuo

Nesta seção apresentam-se alguns dos resultados obtidos na estimação de parâmetros de um modelo experimental do nível de aço líquido, no molde de uma máquina de lingotamento contínuo, mostrada na figura 4.12.

Os algoritmos ME e M.Q.L. são comparados a partir da evolução no tempo da função de custo. As medidas utilizadas foram geradas a partir do modelo linearizado de uma máquina de lingotamento contínuo descrito em Amaral [61]. O modelo utilizado para a estimação de parâmetros é do tipo indicado na equação 4.14.

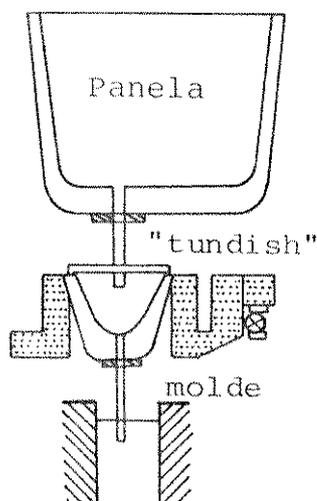


Figura 4.12 - Elementos de uma máquina de lingotamento contínuo.

$$A(Z^{-1})Y_K = B(Z^{-1})u_K + C(Z^{-1})e_{-K} \quad 4.14$$

com

$$Y_K^T = \begin{bmatrix} HT_K & HM_K \end{bmatrix} \quad e \quad u_K = UT_K$$

onde HT e HM são os níveis de aço no "tundish" e no molde e UT é a abertura da válvula de alimentação do molde.

Diferentes valores de N_A , N_B e N_C foram analisados sendo que os melhores resultados, utilizando o teste - F [1], correspondem a

$$\text{M.E.} \quad N_A = 3, N_B = 1, N_C = 3$$

$$\text{M.Q.L.} \quad N_A = 3, N_B = 3, N_C = 3$$

A figura 4.13 representa a evolução do custo acumulado, em uma realização, nas duas situações descritas anteriormente.

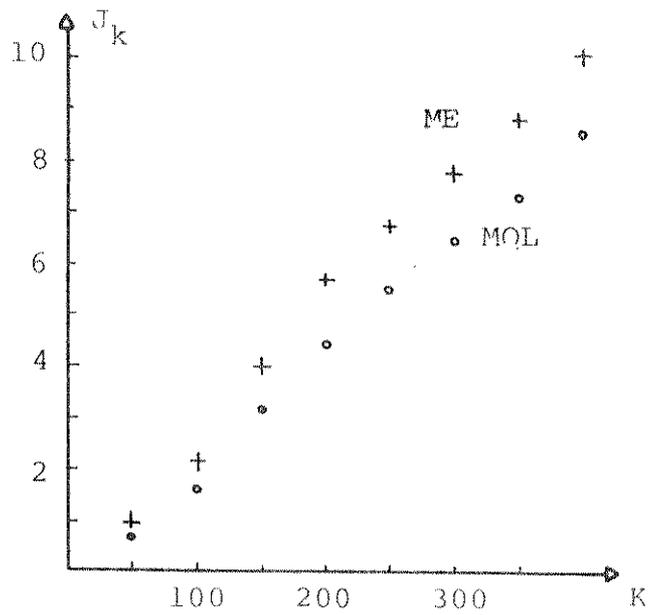


Figura 4.13 - Custo acumulado

Na figura 4.14 comparam-se os erros de previsão no nível de aço no molde quando se utilizam os dois métodos de estimação. A melhora ocasionada pelo algoritmo M.Q.L. é bem visível.

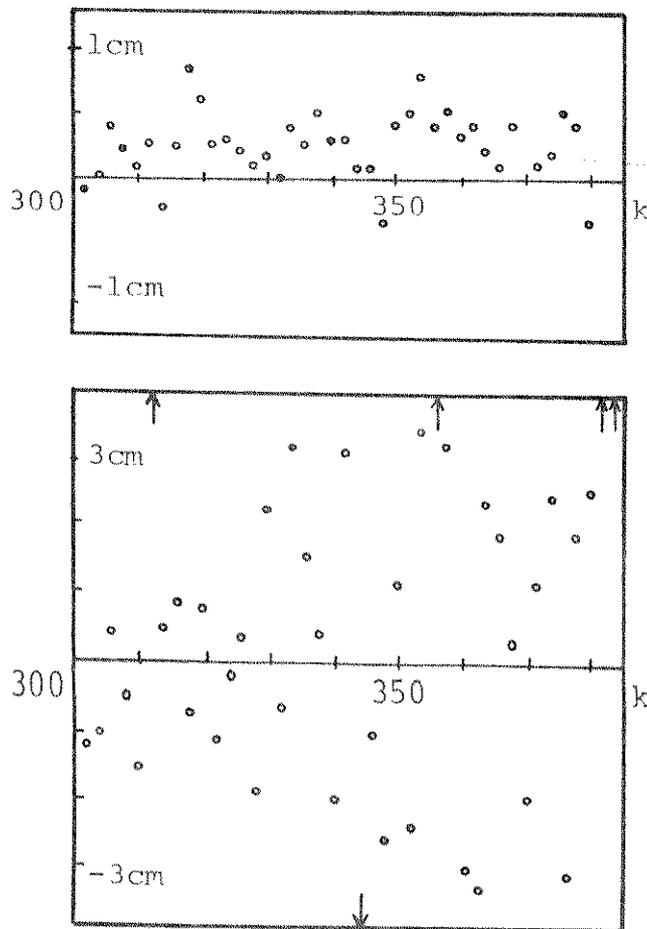


Figura 4.12 - Erro previsto no nível do molde.

O tempo de cálculo para o algoritmo dos M.Q.L. foi duas vezes maior do que o algoritmo da M.E. Dispondo-se de uma estrutura paralela para o cálculo de cada vetor de ganho \underline{K}^i o tempo de cálculo do estimador dos M.Q.L. poderia ser reduzido pela metade.

Comparando-se a ordem dos polinômios utilizada em cada um dos algoritmos nota-se que não houve aumento em N_C . Certamente a transformação do modelo utilizado pelo algoritmo ME na forma canônica III originará um valor de N_C maior do que três. Os resultados obtidos com o estimador M.Q.L. com as ordens indicadas mostram, como já foi descrito na seção 4.5.2, que com este algoritmo podem ser obtidos erros de previsão pequenos mesmo com ordens reduzidas nas matrizes polinomiais. Obtém-se então um compromisso entre a qualidade da estimação e o tempo de cálculo no algoritmo M.Q.L. utilizando-se a forma canônica III de Kashyap.

CAPÍTULO 5

5. CONCLUSÕES

As conclusões que resultam deste trabalho são de dois tipos. Por um lado foram apresentadas as generalizações, para o caso MI MO com modelos ARMA para o ruído, dos algoritmos de identificação recursiva mais utilizados na literatura. Todos os algoritmos foram obtidos minimizando uma função de custo na forma de um somatório dos erros de previsão ponderados. Mostrou-se que a obtenção de equações explícitas para os estimadores requer uma aproximação linear para o erro de previsão, sendo que os diferentes algoritmos caracterizam-se por aproximações mais ou menos refinadas. Neste sentido foram desenvolvidos os seguintes algoritmos: mínimos quadrados linearizado, mínimos quadrados generalizado, matriz estendida e mínimos quadrados particionado.

Além da obtenção unificada de todos estes algoritmos para o caso MIMO/ARMA, a principal contribuição deste trabalho consiste no desenvolvimento destes algoritmos para uma estrutura geral no modelo do ruído, eliminando o vínculo subjacente nos trabalhos publicados nesta área onde, via de regra, cada algoritmo é utilizado com um modelo particular para o ruído. O comportamento destes algoritmos no caso geral ARMA foi ilustrado através de exemplos de simulação.

Os algoritmos de identificação recursiva foram analisados, por outro lado, do ponto de vista da sua implementação para controle em tempo real. Para tanto estudou-se, com especial ênfase, a possibilidade de redução do tempo de cálculo envolvido através da partição do problema global em subproblemas passíveis de serem implementados de uma forma paralela. A partição utilizada foi a de estimar separadamente cada uma das linhas do modelo, decompondo o problema em n subproblemas, onde n é o número de saídas. Neste sentido mostrou-se que os algoritmos dos mínimos quadrados e da matriz estendida podem ser sempre particionados, independentemente da forma da matriz de correlação do ruído. Uma partição suplementar permitiu a obtenção do algoritmo da matriz estendida, significando uma sensível redução no tempo de cálculo. Os métodos utilizados para a linearização do erro previsto nos algoritmos linearizado e generalizado não permitem uma partição direta do problema de estimação. Nestes casos a partição só é possível

utilizando a forma canônica III apresentada por Kashyap ou em casos particulares através da utilização de formas pseudo-canônicas.

Como futuras linhas de trabalho sugere-se, entre outras, as seguintes:

1. Análise de formas canônicas com um número mínimo de parâmetros nos modelos ou que permitam a partição do estimador: como ficou demonstrado no trabalho, as formas AR, MA ou ARMA do modelo do ruído não são determinantes para os algoritmos, mas sim o número de parâmetros do modelo.
2. Implementação dos algoritmos estudados em sistemas multiprocessadores, com processamento paralelo; a rápida evolução da microeletrônica (já foram anunciados pela INTEL os micro-computadores de 32 bits) permite a realização, a baixo custo, de estruturas de cálculo sofisticadas e a consequente implementação "on-line" e em tempo real de algoritmos complexos de controle.
3. Estudo dos problemas de identificação e controle descentralizados: a evolução do processamento distribuído é hoje um elemento chave para o incremento da produtividade em controle de processos industriais. Na literatura têm sido considerados, principalmente, estudos de métodos de controle descentralizado, a partir da representação de estados do sistema. Parece-nos, contudo, mais promissora, a consideração de modelos do tipo ARMA e a generalização dos resultados de controle adaptativo clássico (centralizado) para essa classe de sistemas. Em particular, será necessário o desenvolvimento de métodos de identificação descentralizados, que normalmente utilizem os sub-conjuntos de medidas dos sub-sistemas locais. A troca de informação entre os sub-sistemas será feita de forma esporádica, de modo a minimizar os custos de transmissão: essa informação deverá ser essencial para a correção dos modelos locais, em particular, a polarização de estimação.

Referências Bibliográficas

- [1] Eykoff, P. - System identification: parameter and state estimation. London, Wiley, 1974.
- [2] Fasol, K.H. e Jorgl, H.P. - Modelling and Identification Tutorials on System Identification. 5th IFAC - Symposium on Identification and System Parameter Estimation, Darmstadt, Setembro 1979.
- [3] Strejc, V. - "On Least Squares and Regression Methods in Identification : A survey of Progresses and Trends". Proc. of the 7th IFAC World Congress, Helsinki 1978.
- [4] Morf, M.; Kailath, T.; Ljung, L. - "Fast algorithms for Recursive Identification". Proc. 1976 IEEE Conf. on Decision and Control. Dezembro 1976.
- [5] Ljung, L.; Morf, M. e Falconer, D. - "Fast calculation of gain matrices for recursive estimation schemes". Int. J. Control 27, 1 - 19.
- [6] Robins, A.J. e Wellstead, P.E. - " An appraisal of Fast Algorithms". Proc. 5 th IFAC Symposium on Identification and System Parameter Estimation, Darmstadt, Setembro 1979.
- [7] Borisson, U. - Self-tuning regulators - Industrial Application and Multivariable Theory. Lund Institute of Technology, Dep. Automatic Control, 1975 (Report 7513, Outubro 1975).
- [8] Borisson, U. e Syding, R. - "Self-Tuning control of an ore crusher". IFAC Symposium on Stochastic Control. Budapest, 1974.
- [9] Sinha, N.K. e Sen, A. - "Critical evaluation of on-line identification methods". Proc. IEE 122, 1153-1158 (1975).

- [10] Keviczky, L.K. e Bányász, Cs.M. - "Some new results on multiple input multiple output identification methods". Proc. 4th IFAC Symposium Identification and System Parameter Estimation, Tbilisi (1976)
- [11] Astrom, K.J. e Eykoff - "System identification, a survey" : Identification and Process Parameter Estimation, Czechoslovakia, Academia Prague, 1970. pt.1, p.01 (IFAC Symposium, 2, Prague, 1970).
- [12] Soderstrom, T.; Ljung, L.; Gustavsson, I. - A comparative study of recursive identification methods. Lund Institute of Technology, Dep. Automatic Control, 1974 (Report 7427, Dec. 1974).
- [13] Young, P.C., "The use of linear regression and related procedures for the identification of dynamic processes". Proc. 7th IEEE Symposium Adaptive Proc. (1968)
- [14] Strejec, V. - "Least Squares Parameter Estimation" - Tutorials on System Identification . 5th IFAC Symposium on Identification and System Parameter Estimation. Darmstadt, Setembro 1979.
- [15] Talman, J.L.; Van Den Boom, A.J.W. - "On the estimation of process and noise dynamic using a single-stage estimation". In: Eykoff, P. - Identification and system parameters estimation. Amsterdam, North-Holland Pub. Co, 1973. pt. 2, p. 711-720 (IFAC Symposium, 3, Netherlands, 1973).
- [16] Soderstrom, T.; Ljung, L. e Gustavsson, I. - "A theoretical Analysis of Recursive Identification Methods". Automatica Oxford, 14 : 231-244, 1978.
- [17] Goodwin, G.C., Payne, P.L., Kabaila, P. "On Canonical forms and algorithms for Multivariable Time Series analysis". Proc. 4th IFAC Symposium Identification and System Parameter Est., Tbilisi (1976).

- [18] Soderstrom, T.-An on-line Algorithm for Approximate Maximum Likelihood Identification of Linear Dynamic Systems. Lund Institute of Technology, Dep. Automatic Control, 1973 (Report 7308, 1973).
- [19] Mendel, J. - Discrete techniques of parameter estimation - Marcel Dekker. New York, 1973.
- [20] Foigel, J.K. e Richalet, J. - "Self-adpting IDCOM". Proc. 5th IFAC Symposium on Identification and System Parameter Estimation, Darmstadt, Setembro 1979.
- [21] Akaike, H. - "Canonical Correlation Analysis of Time Series and the use of an informations criterion" in System Identification Advances and Case Studies. Academic Press, 1977.
- [22] Amaral, W.C. - "Identificação de Sistemas Discretos por Mé todos Sequenciais". Tese de Mestrado. FEC, 1976.
- [23] Van Den Boom, A.J.W.; Van Den Enden, A.W.M. - "The determi nation of the order of processes and noise dynamics". In: Eykhooft, P. ed. - Identification and system parameter esti mation. Amsterdam, North-Holland Pub. Co., 1973, pt.2,p.929-938.
- [24] Isermann, R. - "Practical aspects of Process Identification". Tutorials on System Identification . 5th IFAC- Symposium on Identification and System Parameter Estimation, Darmstadt, Setembro 1979.
- [25] Niederliński, A. e Hajdasiński, A . "Multivariable System Identification - A survey". Proc. 5 th IFAC Symposium on Identification and System Parameter Estimation, Darmstadt , Setembro 1979.
- [26] Tse, E., Anton, J.J.- On the identifiability of parameters. IEEE Trans. Automatic Control, AC-17, 637-646, 1972.

- [27] Gustavsson, I.; Ljung, L.; Soderstrom, T. - Identification of processes in closed-loop-identifiability and accuracy aspects. Automatica, 13, 59-75, 1977
- [28] Popov, V.M. - Some properties of the control Systems with irreducible matrix - transfer functions. In lecture Notes in Math., 144, Springer Verlag, New York.
- [29] Wolovich, W.A. - Linear multivariable systems. Springer Verlag, New York, 1974.
- [30] Rosenbrock, H.H. - State-space and multivariable theory. Thomas Nelson, London, 1970.
- [31] Kashyap, R.L. e Rao, A. Ramachandra - Dynamic Stochastic Models from Empirical Data. Academic Press, 1976.
- [32] Ljung, L e Rissanen, J. - "On Canonical Forms, Parameter Identifiability and the concept of Complexity". Proc. 4 th IFAC Symp. Ident. and System Parameter Estimation, Tbilisi (1976).
- [33] Dickinson, B.W., Kailath, T., Morf, M. - Canonical matrix fraction and state space description for deterministic and stochastic linear systems. IEEE Trans. Automat. Control, AC-19, 656-666, 1974.
- [34] Payne, R.L. e Goodwin, G.C. - Dynamic System Identification : Experiment Design and Data Analysis. Academic Press, 1977.
- [35] Astrom, K.J. Introduction to stochastic control theory. Academic Press, New York, 1970.
- [36] Kashyap, R.L., Nasburg, R.E. - Parameter estimation in multivariable stochastic difference equations. IEEE Trans. Automatic Control, AC-19, 784-797, 1974
- [37] Hannan, E.J. - The identification of vector mixed autoregressive moving average systems. Biometrika 58 - 223-225.

- [38] Mehra, R.K. e Lainiotis, D.G. - System Identification : Advances and Case Studies. Academic Press, 1977.
- [39] Tzafestas, S.G. - Some Computer - aided estimators in stochastic control systems identification. Int. J. Control, vol. 12, 385 - 399, 1970.
- [40] Astrom, K.J., Witternmark, B., On Self-Tuning Regulators, Automatic, 9. 195-199, 1973.
- [41] Clarke, D.W. - "Generalized least-squares estimation of the parameters of a dynamical model". Identification in Automatic Control System, Prague, 1967. IFAC Symposium.
- [42] Soderstrom, T. - On the convergence Properties of the Generalized Least Squares Identification Mehtod. Lund Institute of Technology, Dep. Automatic Control, 1972 (Report 7228, 1972).
- [43] Hastings-James, R. e Sage, M.W. - "Recursive generalizes least square procedure for on-line identification of process parameters". Proc. IEE, London, 116 (12) : 2057-2062, Dec., 1969.
- [44] Neudecker H.- "Some Theorems on Matrix Differentiation with Special Reference to Kronecker Matrix Products". American Stat. Ass. Journal, 953-963 - 1969.
- [45] Tracy, D.S.; Dwyer, P.S. - "Multivariable Maxima and Minima with Matrix Derivates" . American Stat. Ass. Journal, 1576-1594, 1969.
- [46] Pannuska, V. - An adaptive recursive least squares identification algorithm. Proc. JACC, 1968.
- [47] Young, P.C. "Comments on On-line identification with applications to Kalamn filtering". IEEE Trans. Autom. Control, April - 1972.

- [48] Astrom, K.J. e Bohlin, T. - "Numerical identification of linear dynamic systems from normal operating records". Proc. 2nd IFAC Symp. on the theory of Self Adaptive Control Systems. Teddington, 1965.
- [49] Saridis, G.N. - Self organization control of stochastic systems. New York, Marcel Dekker, 1977.
- [50] Astrom, K.J. - Maximum likelihood and Prediction error methods. Tutorials on System Identification .5th IFAC Symp. on Identification and System Parameter Estimation, Darmstadt, Setembro - 1979.
- [51] Kailath, T.-The innovations approach to detection and estimation theory. Proc. IEEE, 58, 680-695, 1970.
- [52] Gertler, J. e Bányász, Cs. - A recursive (on-line) maximum likelihood Identification method. IEEE Trans. AC-19, 816-820, 1974.
- [53] Amaral, W.C. e Latre, L.G. - Estimacão Recursiva de Parâmetros em Sistemas Multivariáveis. Proc. do 3º Congresso Brasileiro de Automática, Rio de Janeiro, Setembro 1980.
- [54] Hsia, T.C. -"On least Squares for system Parameter Identification". IEEE Trans. AC., February, 1976.
- [55] Amaral W.C. , Gimeno L., "Recursive Multivariable Least Squares Identification with bias estimation". in Advances in Control , Lainiotis D.G., and Tzannes N.S., Reidel, Holland, 1979.
- [56] Isermann, R.; Baur, V.; Bamberger, W.; Kneppo, P.; Siebert, H.- Comparision of six on-line identification and parameter estimation methods. Automatica, 10, 81-103, 1974.
- [57] Brotherton, T.W. e Caines, P.E. - "ARMA System Identification via the Cholesky Least Squares Method". IEEE Trans. Automatic Control, 23(4), 698 - 701, Aug., 1978.

- [58] Mendel , J.M. - "Computational requerements for a discrete Kalman filter". IEEE Trans. Automatic Control, 16(p): 748-759, DEc., 1974.

- [59] PDP-10. System Reference. Digital Equipment Corporation. 1975.

- [60] Amaral, W.C. et all. Mould Level Control in Continuous Casting Machines". Proc. 6th IFAC/IFIP conf., Dusseldorf, October 1980.