

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA DE CAMPINAS

DEPARTAMENTO DE ENGENHARIA ELÉTRICA/ELETRÔNICA

IDENTIFICAÇÃO DE SISTEMAS DISCRETOS POR
MÉTODOS SEQUENCIAIS

WAGNER CARADORI DO AMARAL

Orientador: Prof. MANUEL DE JESUS MENDES

Tese de Mestrado apresentada à
Faculdade de Engenharia da Uni
versidade Estadual de Campinas

SETEMBRO - 1976

Am13i

1346/BC

UNICAMP
BIBLIOTECA CENTRAL

AGRADECIMENTOS

Desejo expressar meu agradecimento a todas as pessoas que direta ou indiretamente colaboraram na realização deste trabalho. Em particular sou grato a:

Dr. Manuel de Jesus Mendes, meu orientador, pela sua assistência, estímulo e inúmeras sugestões em todas as fases deste trabalho.

Srta. Elza Aoki, pela paciência e boa vontade no trabalho de datilografia.

Prof. Antonio Carlos O. Campos, responsável pela realização dos desenhos contidos neste trabalho.

RESUMO

Neste trabalho são analisados algoritmos de identificação para sistemas lineares, nos parâmetros.

É feita uma generalização do método dos mínimos quadrados, para os casos em que a perturbação é uma sequência de variáveis aleatórias correlatas, a fim de se obterem estimadores não polarizados.

São discutidos tres diferentes tipos de modelamento do ruído e uma versão recursiva dos algoritmos de identificação, para aplicações "on-line". Além disso é apresentado um método que torna possível o estudo analítico da convergência dos algoritmos estocásticos. Finalmente, os algoritmos deduzidos são utilizados para a estimação de parâmetros de diversos processos, simulados em computador.

ÍNDICE

CAPÍTULO I - INTRODUÇÃO	1
CAPÍTULO II - MÉTODOS DE ESTIMAÇÃO "OFF-LINE"	3
2.1 - Introdução	3
2.2 - Características dos estimadores	4
2.2.1 - Estimador não polarizado	4
2.2.2 - Estimador consistente	4
2.2.3 - Estimador eficaz	4
2.3 - Estimador de Markov	5
2.4 - Propriedades do estimador de Markov	8
2.4.1 - Polarização do estimador de Markov	8
2.4.2 - Eficiência do estimador de Markov	8
2.5 - Estimador dos mínimos quadrados	9
2.6 - Generalização do modelo	10
2.7 - Polarização do estimador	13
2.8 - Ruído com função de transferência conhecida	17
2.9 - Ruído com função de transferência desconhecida	18
2.10 - Resumo	20
CAPÍTULO III - MÉTODOS DE IDENTIFICAÇÃO RECURSIVA	21
3.1 - Introdução	21
3.2 - Estimador recursivo dos mínimos quadrados	22
3.3 - Estimador recursivo dos mínimos quadrados com o ruído modelado como auto-regressão	27

3.4	- Método da matriz estendida	30
3.5	- Ruído modelado como média-móvel	34
3.6	- Versão dos algoritmos em tempo real	36
3.7	- Aumento da taxa de convergência	39
3.8	- Inicialização dos algoritmos	40
3.9	- Determinação da ordem do modelo	40
3.10	- Resumo	42
CAPÍTULO IV - CONVERGÊNCIA DOS ALGORÍTMOS DE IDENTIFICAÇÃO		46
4.1	- Introdução	46
4.2	- Análise de convergência	46
4.3	- Obtenção das equações de Ljung	48
4.4	- Estudo da natureza dos pontos singu lares.	52
4.5	- Construção de um algoritmo média-mó vel que tenha solução estável para o valor verdadeiro	58
4.6	- Resumo	60
CAPÍTULO V - RESULTADO E COMPARAÇÕES DOS MÉTODOS APRESENTADOS		61
5.1	- Introdução.	61
5.2	- Sistemas simulados	61
5.3	- Escolha do fator de ponderação	64
5.4	- Resultados do método dos mínimos qua drados recursivo.	69
5.5	- Método dos mínimos quadrados genera lizado com o ruído modelado por au to-regressão.	76
5.6	- Simulação do algoritmo dos mínimos quadrados generalizado com o ruído modelado por auto-regressão e média- móvel.	76

5.7 - Simulação do algoritmo dos mínimos quadrados generalizado com o ruído modelado como média-móvel	85
5.8 - Simulação de um sistema com parâmetro variando no tempo	92
5.9 - Exemplo da determinação da ordem do modelo	94
5.10 - Comparação dos métodos de estimação	95
5.11 - Conclusões	96
CAPÍTULO VI - CONCLUSÕES	98
REFERÊNCIAS BIBLIOGRÁFICAS	100

CAPÍTULO 1

INTRODUÇÃO

Problemas de identificação tem recebido um grande interesse nos últimos anos por causa da aplicação da teoria de controle a problemas industriais.

Vários métodos em teoria de controle requerem uma descrição do sistema em termos de equações diferenciais ou às diferenças e a descrição das perturbações como processos estocásticos, caracterizados também por equações diferenciais ou às diferenças estocásticas ou por propriedades de segunda ordem, tais como, a função de covariância e a densidade espectral.

Em muitos problemas práticos, a descrição de um sistema e suas perturbações não é disponível, tendo a identificação o propósito de obter um modelo adequado.

São apresentados métodos de estimação recursiva, porque em muitas aplicações é altamente desejável obter-se resultados da identificação de um processo enquanto ele está se desenvolvendo. Por exemplo, isto ocorre no caso de aplicações industriais onde o sistema geralmente será de ordem elevada, mas como há interesse de analisar o processo em torno de um certo ponto de operação, pode-se utilizar modelos de menor ordem [2]. No entanto com esta aproximação os parâmetros do modelo podem variar ao longo do processo sendo necessário uma identificação on-line.

Também é interessante aplicar-se estimação recursiva em sistemas de controle adaptativo, porque muitos controladores adaptativos são construídos baseados na afirmação de que é possível separar a estimação em tempo real dos parâmetros, como mostra a figura 1.1.

Um controlador muito utilizado é o regulador "self-tuning" que tem a estrutura apresentada na figura 1.1, sendo a identificação baseada no método dos mínimos quadrados.

Entretanto controladores mais sofisticados podem ser

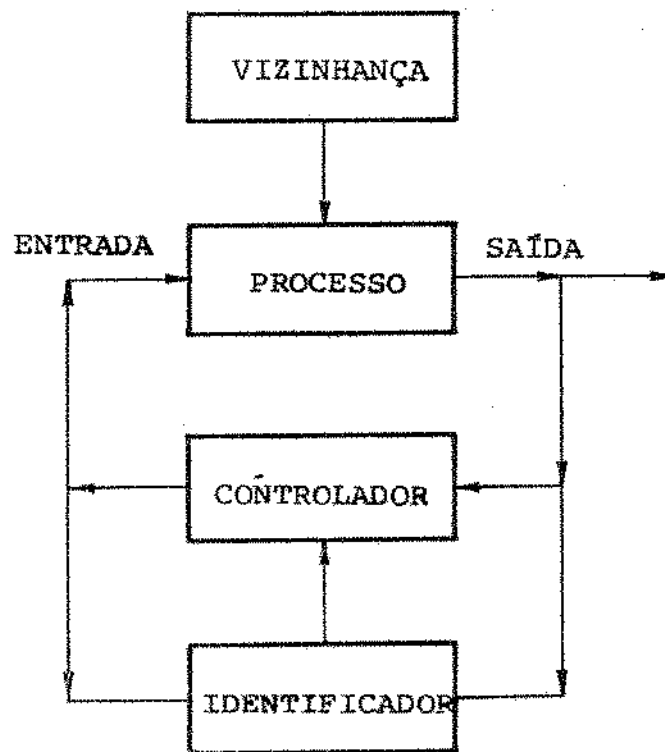


FIGURA 1.1

obtidos introduzindo-se outros métodos de identificação, assim este trabalho procura analisar alguns desses métodos.

A organização deste trabalho obedece ao seguinte esquema: no capítulo 2 formulam-se os problemas de identificação e estimação. Métodos de estimação recursiva são analisados no capítulo 3, bem como a determinação da ordem do modelo.

No capítulo seguinte são apresentados resultados sobre a convergência dos algoritmos. No capítulo 5 simulam-se vários exemplos ilustrativos e no último capítulo são apresentadas as conclusões.

CAPÍTULO 2

MÉTODOS DE ESTIMAÇÃO "OFF-LINE"

2.1 - INTRODUÇÃO

Para a aplicação da teoria de controle estocástico, o processo a ser controlado deve ser descrito em termos de uma equação diferencial ou às diferenças, linear nos parâmetros e as perturbações na forma de processos estocástico estacionários no sentido amplo, com um espectro racional.

Este trabalho procura uma técnica para determinar tais modelos a partir de medidas das entradas e saídas do processo.

Restringir-se-ã aos casos de equações às diferenças com uma única entrada e saída e modelos invariantes no tempo.

Com o objetivo de utilizar uma linguagem comum definem-se inicialmente os seguintes termos:

Identificação é a determinação de um modelo, baseado nos sinais de entrada e saída, de tal modo que seja equivalente ao processo sob teste, segundo algum critério definido [1].

Partindo-se do princípio que a estrutura do modelo é conhecida, estimação é o processo da determinação dos parâmetros do modelo [1].

Supõe-se conhecida a ordem do modelo e os cálculos para obtenção deste valor são mostrados num capítulo posterior.

Para a determinação dos parâmetros do modelo, distinguem-se os métodos em duas categorias básicas: [2]

1 - Métodos de estimação "off-line".

2 - Métodos de estimação "on-line".

Estes últimos são interessantes quando as características do processo ou do ruído variam gradualmente durante o controle do processo.

Este capítulo tratará da primeira classe de métodos, sendo a segunda analisada no capítulo seguinte.

2.2 - CARACTERÍSTICAS DOS ESTIMADORES

Seja o processo

$$\underline{y} = \underline{f}(\underline{\theta}) + \underline{n} \quad (2.1)$$

onde $\underline{\theta}$ é um vetor $n \times 1$ de parâmetros (determinísticos) desconhecidos, $\underline{y}$ é um vetor $m \times 1$ de medidas estocásticas (com ruído) e $\underline{n}$ é um vetor $m \times 1$ de ruído.

As propriedades desejadas para o estimador $\underline{\theta}$ dos parâmetros $\underline{\theta}$ fornecidas na literatura são: [2, 3].

2.2.1 Estimador não polarizado

Define-se polarização $\underline{p}$ como

$$\underline{p} \triangleq \underline{\theta} - E \hat{\underline{\theta}} \quad (2.2)$$

em que $\hat{\underline{\theta}}$ é o estimador de $\underline{\theta}$

Para um estimador sem polarização vale:

$$\underline{p} = \underline{0} \quad (2.3)$$

2.2.2 Estimador consistente

Um estimador é consistente se

$$\lim_{K \rightarrow \infty} P \left[|\hat{\underline{\theta}} - \underline{\theta}| > \epsilon \right] = 0 \quad (2.4)$$

onde K é o número de medidas realizadas e com ϵ arbitrariamente pequeno, isto é, $\hat{\underline{\theta}}$ converge em probabilidade para o valor real $\underline{\theta}$.

2.2.3 Estimador eficaz

Pela desigualdade de Cramér - Rao, a matriz de erro $\underline{p} = E(\underline{\theta} - \hat{\underline{\theta}})(\underline{\theta} - \hat{\underline{\theta}})^T$ é tal que

$$\underline{p} \geq \left[\underline{I} + \nabla_{\underline{\theta}} p(\underline{\theta}) \right] \left[E \left[\nabla_{\underline{\theta}} \ln p(\underline{y}; \underline{\theta}) \right] \left[\nabla_{\underline{\theta}} \ln p(\underline{y}; \underline{\theta}) \right]^T \right]^{-1} \left[\underline{I} + \nabla_{\underline{\theta}} p(\underline{\theta}) \right]^T \quad (2.5)$$

no sentido que a matriz diferença é semi-definida positiva.

No caso de estimadores não polarizados tem-se:

$$\underline{P} \geq \left[E \left[\underline{V}_{\underline{\theta}} \ln p(\underline{y} : \underline{\theta}) \right] \left[\underline{V}_{\underline{\theta}} \ln p(\underline{y} : \underline{\theta}) \right]^T \right]^{-1} \quad (2.5.a)$$

onde $p(\underline{y} : \underline{\theta})$ é a função densidade de probabilidade de $\underline{y}$ fixando-se $\underline{\theta}$ como parâmetro determinístico.

Um estimador é eficaz quando o limite mínimo da desigualdade de Cramér - Rao é satisfeito. A condição necessária e suficiente para se obter tal propriedade é:

$$\hat{\underline{\theta}}_{EF} = \underline{\theta} + \underline{F}(\underline{\theta}) \underline{V}_{\underline{\theta}} \ln p(\underline{y} : \underline{\theta}) \quad (2.6)$$

onde $\underline{F}(\underline{\theta})$ é uma função matricial de $\underline{\theta}$ [18]

2.3 - ESTIMADOR DE MARKOV

Seja um sistema de equações lineares relacionando um vetor de variáveis $\underline{y}$ com um vetor de parâmetros $\underline{\theta}$

$$\underline{y} = \underline{F} \underline{\theta} \quad (2.7)$$

onde $\underline{F}$ é uma matriz conhecida $m \times n$, $\underline{\theta}$ um vetor de dimensão $n \times 1$ desconhecido e $\underline{y}$ um vetor de observações de dimensão $m \times 1$.

Se a observação é exata e a matriz $\underline{F}$ tem "rank" n (isto implica $m \geq n$), facilmente se determina o valor de $\underline{\theta}$ que satisfaz a equação (2.7) :

$$\hat{\underline{\theta}} = (\underline{F}^T \underline{F})^{-1} \underline{F}^T \underline{y} \quad (2.8)$$

Se as observações $\underline{y}$ forem estocásticas, então pode-se decidir fazer um número de medidas $y_0, y_1, \dots, y_K$ e escolher $\underline{\theta}$ que satisfaz a equação (2.7) de modo a minimizar a soma dos quadrados dos erros $\underline{e} = \underline{y} - \underline{F} \underline{\theta}$

Com esta motivação, considera-se um sistema dinâmico cuja relação entrada/saída é descrito pela equação:

$$y(K) = b_0 u(K) + b_1 u(K-1) + \dots + b_m u(K-m) + n(K) \quad (2.9)$$

onde $\{n(K)\}$ é uma sequência de variáveis aleatórias normais, $y(K)$ a saída no instante K e $u(K)$ a entrada no instante K .

A equação anterior pode ser reescrita como:

$$\underline{y} = \underline{U} \underline{b} + \underline{n} \quad (2.10)$$

sendo

$$\underline{y} = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(K) \end{bmatrix} \quad (2.11.a)$$

$$\underline{U} = \begin{bmatrix} u(1) & \dots & u(1-m) \\ \vdots & & \vdots \\ u(K) & \dots & u(K-m) \end{bmatrix} \quad (2.11.b)$$

$$\underline{b} = \begin{bmatrix} b_0 \\ \vdots \\ b_m \end{bmatrix} \quad (2.11.c)$$

$$\underline{n} = \begin{bmatrix} n(1) \\ n(2) \\ \vdots \\ n(K) \end{bmatrix} \quad (2.11.d)$$

Tendo o ruído valor médio nulo, isto é; $E \underline{n} = 0$ e matriz de covariância dada por:

$$\underline{N} = E \underline{n} \underline{n}^T = \begin{bmatrix} E n(1) n(1) & \dots & E n(1) n(m) \\ \vdots & & \vdots \\ E n(m) n(1) & \dots & E n(m) n(m) \end{bmatrix} \quad (2.12)$$

Considera-se o modelo discreto invariante no tempo

$$y_M(K) = \sum_{i=0}^m \theta_i u(K-i) \quad (2.13)$$

em que se supõe conhecida sua ordem.

A equação (2.13) pode ser reescrita, quando todos os instantes $K = 1, \dots, N$ são considerados.

$$\underline{y}_M = \underline{U} \underline{\theta} \quad (2.14)$$

$$\underline{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_M \end{bmatrix} \quad (2.14.a)$$

onde $\underline{\theta}$ é o vetor de parâmetros desconhecidos.

$$\text{Minimizando-se } E = ||\underline{y} - \underline{U} \underline{\theta}||_{\underline{N}^{-1}}^2 \quad (2.15)$$

em que $\underline{N}^{-1}$ é uma matriz simétrica definida positiva de dimensão m (isto significa um conhecimento prévio da magnitude dos ruídos envolvidos), o valor ótimo $\hat{\underline{\theta}}$ é tal que

$$E(\hat{\underline{\theta}}) \leq E(\underline{\theta}) \quad (2.16)$$

para qualquer $\underline{\theta}$ no espaço Euclidiano m dimensional, de onde resulta a solução:

$$\hat{\underline{\theta}} = \left[\underline{U}^T \underline{N}^{-1} \underline{U} \right]^+ \underline{U}^T \underline{N}^{-1} \underline{y} \quad (2.17)$$

onde $\left[\underline{U}^T \underline{N}^{-1} \underline{U} \right]^+$ representa a matriz pseudo-inversa da matriz $\left[\underline{U}^T \underline{N}^{-1} \underline{U} \right]$. [19, 20]

2.4 - PROPRIEDADES DO ESTIMADOR DE MARKOV

2.4.1 Polarização do estimador de Markov

Teorema 2.4.1 - O estimador de Markov é não polarizado se o ruído tiver média nula e a entrada e o ruído forem independentes.

Prova:

$$\hat{\underline{\theta}} = \left[\underline{U}^T \underline{N}^{-1} \underline{U} \right]^+ \underline{U}^T \underline{N}^{-1} \underline{y}$$

$$\hat{\underline{\theta}} = \left[\underline{U}^T \underline{N}^{-1} \underline{U} \right]^+ \underline{U}^T \underline{N}^{-1} \underline{U} \underline{b} + \underline{n}$$

$$= \underline{b} + \left[\underline{U}^T \underline{N}^{-1} \underline{U} \right]^+ \underline{U}^T \underline{N}^{-1} \underline{n}$$

Se $\underline{U}$ e $\underline{n}$ são independentes

$$\underline{p} = \underline{b} - E \hat{\underline{\theta}} = E \left\{ \left[\underline{U}^T \underline{N}^{-1} \underline{U} \right]^+ \underline{U}^T \underline{N}^{-1} \right\} E \underline{n} = \underline{0}$$

$$\underline{p} = \underline{0}$$

2.4.2 - Eficiência do estimador de Markov

Para o estimador ser eficaz é necessário que a equação (2.6) seja satisfeita

A partir da equação (2.10)

$$\nabla_{\underline{b}} \ln p(\underline{y} : \underline{b}) = (\underline{U}^T \underline{N}^{-1} \underline{y} - \underline{U}^T \underline{N}^{-1} \underline{U} \underline{b}) \quad (2.18)$$

$$\nabla_{\underline{b}} \ln p(\underline{y} : \underline{b}) = \underline{U}^T \underline{N}^{-1} \underline{U} (\underline{U}^T \underline{N}^{-1} \underline{U})^+ \underline{U}^T \underline{N}^{-1} \underline{y} - \underline{b} \quad (2.19)$$

logo:

$$\hat{\underline{\theta}} = \underline{b} + [\underline{U}^T \underline{N}^{-1} \underline{U}]^+ \nabla_{\underline{b}} \ln p(\underline{y} : \underline{b}) \quad (2.20)$$

Portanto, o estimador de Markov satisfaz a equação (2.6) que é a condição para o estimador ser eficaz.

2.5 - ESTIMADOR DOS MÍNIMOS QUADRADOS

Assumindo-se o ruído $\underline{n}$ branco, isto é, com matriz de covariância $\underline{N} = \sigma^2 \underline{I}$ a equação (2.17) se reduz a:

$$\hat{\underline{\theta}} = [\underline{U}^T \underline{U}]^{-1} \underline{U}^T \underline{y} \quad (2.21)$$

Este estimador também é não polarizado, e possui matriz de covariância dada por:

$$\begin{aligned} \text{Cov } \hat{\underline{\theta}} &= E \left[[\underline{U}^T \underline{U}]^{-1} \underline{U}^T \underline{n} \underline{n}^T \underline{U} [\underline{U}^T \underline{U}]^{-1} \right] \\ &= [\underline{U}^T \underline{U}]^{-1} \underline{U}^T \sigma^2 \underline{I} \underline{U} [\underline{U}^T \underline{U}]^{-1} \\ \text{Cov } \hat{\underline{\theta}} &= \sigma^2 [\underline{U}^T \underline{U}]^{-1} \end{aligned} \quad (2.22)$$

Fixando-se um certo valor c^2 para o erro, é fácil mostrar que a forma quadrática pode ser colocada sob a forma:

$$\Omega_{\underline{\theta}} = \left\{ \underline{\theta} : (\underline{\theta} - \hat{\underline{\theta}})^T \underline{U}^T \underline{U} (\underline{\theta} - \hat{\underline{\theta}}) \leq c^2 \right\} \quad (2.23)$$

Onde $\Omega_{\underline{\theta}}$ é um elipsóide com centro em $\underline{\hat{\theta}}$ no espaço m-dimensional de vetores $\underline{\theta}$ [21].

O tipo e a localização do elipsóide são determinados por $\underline{\hat{\theta}}$ e $\underline{U}$. O centro do elipsóide é determinado por $\underline{\hat{\theta}}$, a orientação pelos autovetores de $\underline{U}^T \underline{U}$ e os comprimentos dos eixos são:

$$l_i = \frac{2c}{\lambda_i}$$

$\lambda_i = 1, 2, \dots, m$ são raízes quadradas dos autovalores da matriz $\underline{U}^T \underline{U}$.

2.6 - GENERALIZAÇÃO DO MODELO

Até este instante, foi analisado um modelo em que a função de transferência do sistema é descrita por um único polinômio de ordem finita. Nesta seção, é considerado um modelo generalizado, que estima os parâmetros, tanto do numerador quanto do denominador, da função de transferência do sistema.

Seja um sistema dinâmico discreto com uma entrada - uma saída, cuja relação entrada/saída pode ser descrita pela equação:

$$Y(K) + a_1 Y(K-1) + \dots + a_n Y(K-n) = b_1 u(K-1) + \dots + b_n u(K-n) + n(K) \quad (2.24)$$

Pela introdução do operador deslocamento Z , onde

$$\begin{aligned} Z Y(K) &= Y(K+1) \\ Z^{-1} Y(K) &= Y(K-1) \end{aligned} \quad (2.25)$$

e os polinômios:

$$\begin{aligned} A(Z^{-1}) &= 1 + a_1 Z^{-1} + \dots + a_n Z^{-n} \\ B(Z^{-1}) &= b_1 Z^{-1} + b_2 Z^{-2} + \dots + b_n Z^{-n} \end{aligned}$$

A equação (2.24) pode ser reescrita

$$A(Z^{-1}) Y(K) = B(Z^{-1}) u(K) \quad (2.26)$$

Assume-se que a descrição do sistema é tal que o par de polinômios $A(Z^{-1})$ e $B(Z^{-1})$ não tem fatores comuns, isto particularmente implica que todos os estados do sistema são controláveis.

Supõe-se também que o sistema é assintoticamente estável, significando que, para operações em malha aberta, o polinômio $A(Z^{-1})$ tem todos os zeros fora do círculo unitário.

Note-se também que os graus dos polinômios $A(Z^{-1})$ e $B(Z^{-1})$ são os mesmos; caso contrário alguns coeficientes de $B(Z^{-1})$ podem ser tomados iguais a zero.

Considere-se um modelo discreto linear, invariante no tempo, com uma entrada e uma saída.

Uma forma canônica para este modelo é:

$$Y_m(K) + \alpha_1 Y_m(K-1) + \dots + \alpha_n Y_m(K-n) = \beta_1 u(K-1) + \dots + \beta_n u(K-n)$$

onde u é a entrada do modelo e y_m é a saída do modelo.

Utilizando a equação (2.25) o modelo pode ser reescrito:

$$A_M(Z^{-1}) Y_M(K) = B_M(Z^{-1}) u(K) \quad (2.27)$$

onde

$$A_M(Z^{-1}) = 1 + \alpha_1 Z^{-1} + \dots + \alpha_n Z^{-n}$$

$$B_M(Z^{-1}) = \beta_1 Z^{-1} + \beta_2 Z^{-2} + \dots + \beta_n Z^{-n}$$

Agora será desenvolvido, para o caso do modelo generalizado, o estimador obtido anteriormente, em vista das interessantes propriedades apresentadas pelo mesmo.

Seguindo as mesmas linhas da seção 2.3 um estimador dos mínimos quadrados é obtido minimizando a função de custo descrita pela equação

$$E = \underline{e}^T \underline{e} \quad (2.28)$$

onde $\underline{e}$ é o erro generalizado definido por:

$$e(k) = A_M(z^{-1}) (y(k) - y_M(k)) \quad (2.29)$$

ou

$$e(k) = A_M(z^{-1}) y(k) - B_M(z^{-1}) u(k) \quad (2.30)$$

A figura 2.1 mostra como o erro generalizado pode ser obtido a partir das entradas e saídas.

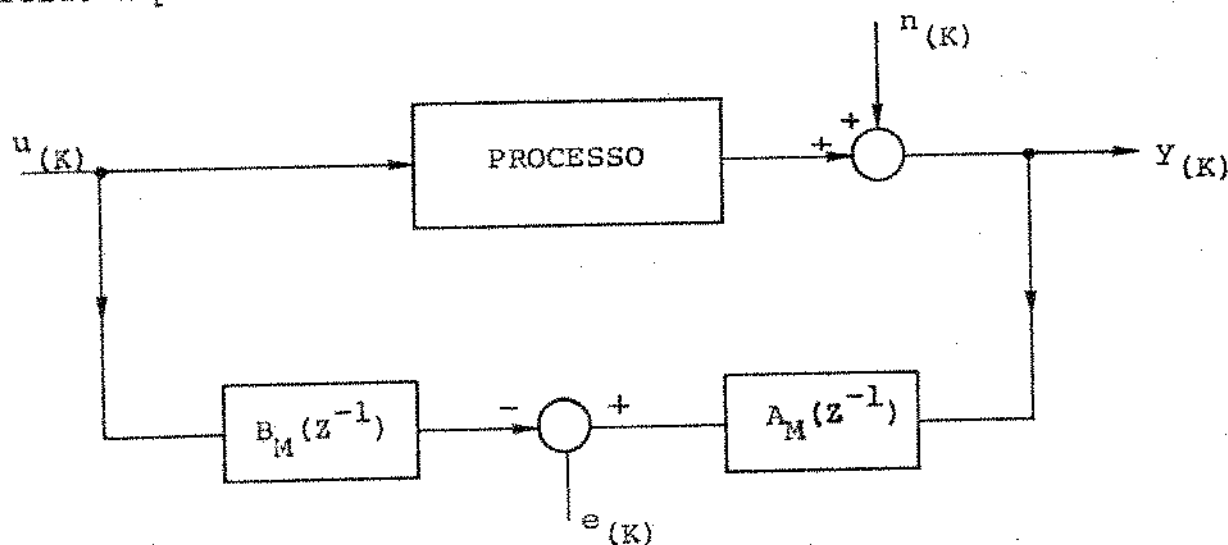


FIGURA 2.1

Da equação (2.30), para $k = 1, 2, 3, \dots$, tem-se:

$$\underline{e} = \underline{y} + \underline{y} \underline{\alpha} - \underline{u} \underline{\beta} = \underline{y} - \begin{bmatrix} \underline{u} \\ \underline{y} \end{bmatrix} \begin{bmatrix} \underline{\beta} \\ -\underline{\alpha} \end{bmatrix} \quad (2.31)$$

$$\text{onde } \underline{Y} = \begin{bmatrix} Y_{(0)} \cdots Y_{(1-m)} \\ \vdots \\ Y_{(K-1)} \cdots Y_{(K-m)} \end{bmatrix} \quad \text{e} \quad \underline{U} = \begin{bmatrix} u_{(0)} \cdots u_{(1-m)} \\ \vdots \\ u_{(K-1)} \cdots u_{(K-m)} \end{bmatrix}$$

ou

$$\underline{e} = \underline{Y} - \underline{\Omega} \underline{\Theta} \quad (2.32)$$

$$\text{onde } \underline{\Omega} = \begin{bmatrix} \underline{U} & \underline{Y} \end{bmatrix} \quad \text{e} \quad \underline{\Theta} = \begin{bmatrix} \beta \\ -\alpha \end{bmatrix}$$

A partir da minimização da expressão (2.28) os estimadores dos parâmetros a_i e b_i são obtidos como função das observações

$$\underline{\hat{\Theta}} = \left[\underline{\Omega}^T \underline{\Omega} \right]^{-1} \underline{\Omega}^T \underline{Y} \quad (2.33)$$

$$\text{onde } \underline{\Omega} = \begin{bmatrix} u_{(0)} \cdots u_{(1-m)} & Y_{(0)} \cdots Y_{(1-m)} \\ \vdots & \vdots \\ u_{(K-1)} \cdots u_{(K-m)} & Y_{(K-1)} \cdots Y_{(K-m)} \end{bmatrix} \quad (2.33.a)$$

2.7 - POLARIZAÇÃO DO ESTIMADOR

Este modelo tem a desvantagem do estimador ser polarizado, pelo fato da matriz de observação $\underline{\Omega}$ ser dependente do ruído.

Teorema 2.7.1 - O estimador dos mínimos quadrados generalizado é assintoticamente não polarizado se o ruído $\underline{n}$ é gerado a partir de uma sequência de variáveis aleatórias não correlatas.

Prova:

Seja o processo discreto linear e invariante no tempo.

$$x_{(K)} + a_1 x_{(K-1)} + \dots + a_m x_{(K-m)} = b_1 u_{(K-1)} + \dots + b_m u_{(K-m)} \quad (2.34.a)$$

com medidas

$$Y_{(K)} = x_{(K)} + n_{(K)} \quad (2.34.b)$$

Pela equação (2.33), o estimador é dado por:

$$\underline{\theta} = \left[\underline{\Omega}^T \underline{\Omega} \right]^{-1} \underline{\Omega}^T \underline{Y} \quad (2.35)$$

onde

$$\underline{\Omega} = \begin{bmatrix} \underline{U} & \vdots & \underline{Y} \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \underline{U} & \vdots & \underline{X} \\ \vdots & \vdots & \vdots \end{bmatrix} + \begin{bmatrix} \vdots & n_{(0)} & \dots & n_{(1-m)} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & n_{(K-1)} & \dots & n_{(K-m)} \end{bmatrix} = \underline{\Omega}_{U,X} + \underline{\Omega}_{0,n} \quad (2.36)$$

Portanto

$$\underline{Y} = \underline{\Omega}_{U,X} \underline{\theta} + \underline{n} \quad (2.37)$$

onde

$$\underline{\theta} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \\ \hline -a_1 \\ \vdots \\ -a_m \end{bmatrix} \quad (2.37.a)$$

logo

$$E \hat{\underline{\theta}} = \hat{\underline{\theta}} - E \left[\left(\underline{\Omega}^T \underline{\Omega} \right)^{-1} \left[\underline{\Omega}^T \underline{\Omega}_{0,n} \underline{\theta} - \underline{\Omega}^T \underline{n} \right] \right] \quad (2.38)$$

Estudam-se a seguir, as propriedades assintóticas do estimador (isto é, para um grande número de observações).

O segundo termo do lado direito da equação (2.38) indica a polarização do estimador

$$\underline{p} = E \left[(\underline{\Omega}^T \underline{\Omega})^{-1} \left[\underline{\Omega}^T \underline{\Omega}_{0,n} \underline{\theta} - \underline{\Omega}^T \underline{n} \right] \right] \quad (2.39)$$

Utilizando a noção de limite em probabilidade, tem-se:

$$\lim_{K \rightarrow \infty} E \left[\hat{\underline{\theta}} \right] = P(\lim_{K \rightarrow \infty} \hat{\underline{\theta}}) \quad (2.40.a)$$

Também sabe-se que:

$$P(\lim_{K \rightarrow \infty} A^{-1}(K) B(K)) = \left[P(\lim_{K \rightarrow \infty} A(K)) \right]^{-1} P(\lim_{K \rightarrow \infty} B(K)) \quad (2.40.b)$$

se A(K) e B(K) convergem em probabilidade (Teorema de Slutsky)

Portanto utilizando as equações acima, obtém-se

$$P(\lim_{K \rightarrow \infty} \underline{p}) = \left[P \lim_{K \rightarrow \infty} \frac{\underline{\Omega}^T \underline{\Omega}}{K} \right]^{-1} \left[P \lim_{K \rightarrow \infty} \frac{\underline{\Omega}^T \underline{\Omega}_{0,n}}{K} \underline{\theta} - P \lim_{K \rightarrow \infty} \frac{\underline{\Omega}^T \underline{n}}{K} \right] \quad (2.41)$$

onde

$$P \lim_{K \rightarrow \infty} \frac{\underline{\Omega}^T \underline{\Omega}_{0,n}}{K} = \left[\begin{array}{c|c} \underline{0} & \underline{0} \\ \hline \underline{0} & \begin{array}{c} \psi_{nn}(0) \dots \dots \psi_{nn}(1-n) \\ \vdots \\ \psi_{nn}(1-n) \dots \dots \psi_{nn}(0) \end{array} \end{array} \right] = \left[\begin{array}{c|c} \underline{0} & \underline{0} \\ \hline \underline{0} & \underline{N} \end{array} \right] \quad (2.42.a)$$

$$P \lim_{K \rightarrow \infty} \frac{\Omega^T \Omega}{K} = \underline{C}$$

e

$$P \lim_{K \rightarrow \infty} \frac{\Omega^T \underline{n}}{K} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \psi_{nn}(1) \\ \vdots \\ -\psi_{nn}(n) \end{bmatrix} \quad (2.42.b)$$

Assumindo que a matriz de covariância $\underline{C}$ é definida positiva, obtém-se:

$$\lim_{K \rightarrow \infty} E[\hat{\underline{\theta}}] = P(\lim_{K \rightarrow \infty} \hat{\underline{\theta}}) = \underline{\theta} - \underline{C}^{-1} \begin{bmatrix} \frac{0}{\vdots} \frac{0}{\vdots} \frac{0}{N} \end{bmatrix} \underline{\theta} - \begin{bmatrix} \frac{0}{\vdots} \psi_{nn}(1) \vdots \psi_{nn}(n) \end{bmatrix} \quad (2.43)$$

Consequentemente não há polarização assintótica se:

$$\underline{N} \underline{a} = - \begin{bmatrix} \psi_{nn}(1) \\ \vdots \\ \psi_{nn}(n) \end{bmatrix} \quad (2.44)$$

isto é, se o ruído $n_{(K)}$ é gerado a partir de uma sequência não correlata $\{\zeta_{(K)}\}$

$$n_{(K)} + a_1 n_{(K-1)} + \dots + a_n n_{(K-n)} = \zeta_{(K)} \quad (2.45)$$

Corolário 2.7.1 - A condição de estimador não polarizado corresponde à condição de erro generalizado "branco".

Prova:

Da equação (2.30) tem-se:

$$e_{(K)} = A_M(Z^{-1}) y_{(K)} - B_M(Z^{-1}) u_{(K)}$$

Se os valores estimados convergem para os valores corretos, resulta:

$$e_{(K)} = A(Z^{-1}) \cdot x_{(K)} + A(Z^{-1}) n_{(K)} - B(Z^{-1}) u_{(K)} \quad (2.46)$$

portanto:

$$e_{(K)} = A(Z^{-1}) n_{(K)} = \zeta_{(K)} \quad (2.47)$$

2.8 - RUÍDO COM FUNÇÃO DE TRANSFERÊNCIA CONHECIDA

No caso geral, o erro nem sempre é branco, portanto, é necessário alterar o estimador determinado pela equação (2.33) para se obter um valor não polarizado.

Seja o processo descrito pelas equações:

$$A(Z^{-1}) x_{(K)} = B(Z^{-1}) u_{(K)} \quad (2.48)$$

com medidas

$$y_{(K)} = x_{(K)} + n_{(K)} \quad (2.49)$$

Supondo que a correlação de $n_{(K)}$ é conhecida, isto é:

$$n_{(K)} = \frac{C(Z^{-1})}{D(Z^{-1})} \zeta_{(K)} \quad (2.50)$$

onde $\zeta_{(K)}$ é uma sequência de variáveis aleatórias não correlatas; a equação que descreve o processo pode ser reescrita:

$$A(Z^{-1}) y_{(K)} = B(Z^{-1}) u_{(K)} + \frac{E(Z^{-1})}{D(Z^{-1})} \zeta_{(K)} \quad (2.51.a)$$

ou

$$A(Z^{-1}) \tilde{Y}_{(K)} = B(Z^{-1}) \tilde{u}_{(K)} + \zeta_{(K)} \quad (2.51.b)$$

onde

$$E(Z^{-1}) = A(Z^{-1}) C(Z^{-1}) \quad e \quad (2.52)$$

$$\tilde{Y}_{(K)} = \frac{D(Z^{-1})}{E(Z^{-1})} Y_{(K)} \quad (2.53.a)$$

$$\tilde{u}_{(K)} = \frac{D(Z^{-1})}{E(Z^{-1})} u_{(K)} \quad (2.53.b)$$

são as medidas e as entradas filtradas do sistema.

Usando o modelo generalizado e operando com as medidas filtradas, os resíduos serão não correlatos e o estimador será assintoticamente não polarizado.

2.9 - RUÍDO COM FUNÇÃO DE TRANSFERÊNCIA DESCONHECIDA

Entretanto, a suposição de $\frac{E(Z^{-1})}{D(Z^{-1})}$ ser conhecido, pode não ser realista; será então apresentado um novo método [5] que procurará estimar $\frac{E(Z^{-1})}{D(Z^{-1})}$, bem como $A(Z^{-1})$ e $B(Z^{-1})$

Seja o processo descrito por:

$$A(Z^{-1}) Y_{(K)} = B(Z^{-1}) u_{(K)} + r_{(K)} \quad (2.54)$$

onde $r_{(K)}$ é uma sequência de variáveis aleatórias correlatas.

Conforme apresentado anteriormente, assume-se o modelo

$A_M(z^{-1}) y_{(K)} = B_M(z^{-1}) u_{(K)}$ e, como já visto na equação (2.33)

$$\hat{\underline{\theta}}_{(i)} = \left[\underline{\Omega}^T \underline{\Omega} \right]^{-1} \underline{\Omega}^T \underline{y} \quad (2.55)$$

e o erro generalizado $\underline{e}_{(i)} = \underline{y} - \underline{\Omega} \hat{\underline{\theta}}_{(i)}$ (2.56)

Desde que no limite $\hat{\underline{\theta}}$ tenda para $\underline{\theta}$, o erro generalizado aproxima-se do ruído e, portanto, uma boa aproximação para este último será:

$$\hat{\underline{r}}_{(i)} \approx \underline{e}_{(i)} \quad (2.57)$$

Para que o estimador seja não polarizado é necessário introduzir uma transformação apropriada nas medidas u e y . Esta transformação é a filtragem das medidas pelo inverso da função de transferência pulsada que caracteriza a vizinhança.

O ruído $\underline{r}$, presente na medida $\underline{y}$, será filtrado supondo-se um modelo auto-regressivo, isto é, o ruído no instante i dependerá da entrada do filtro no instante i e dos instantes anteriores do ruído. Outros tipos de modelamento do ruído serão apresentados no próximo capítulo.

O modelo utilizado é o seguinte:

$$G(z^{-1}) \hat{\underline{r}}_{(i)} = \zeta_{(i)} \quad (2.58)$$

e

$$G(z^{-1}) = 1 + g_1 z^{-1} + \dots + g_s z^{-s} \quad (2.59)$$

Da equação (2.58) para $i = 1, 2, 3, \dots$, tem-se:

$$\hat{\underline{r}} = -\underline{\hat{R}} \underline{g}_{(i)} + \underline{\zeta} \quad (2.60)$$

onde

$$\hat{\underline{R}} = \begin{bmatrix} \hat{r}_{(0)} & \dots & \hat{r}_{(i-s)} \\ \vdots & & \vdots \\ \hat{r}_{(i-1)} & \dots & \hat{r}_{(i-s)} \end{bmatrix} \quad (2.61)$$

O estimador de $\underline{q}(i)$ é dado por:

$$\hat{\underline{q}}(i) = - \left[\hat{\underline{R}}^T \hat{\underline{R}} \right]^{-1} \hat{\underline{R}}^T \hat{\underline{r}} \quad (2.62)$$

Este estimador é não polarizado porque os elementos de $\hat{\underline{R}}$ são independentes de $\underline{r}$.

A partir de $\hat{\underline{q}}(i)$ determina-se o polinômio $G(Z^{-1})$, e em seguida um novo estimador $\hat{\underline{\theta}}$ é obtido utilizando as medidas filtradas $\tilde{u}_{(i)} = G(Z^{-1}) u_{(i)}$ e $\tilde{y}_{(i)} = G(Z^{-1}) y_{(i)}$. Com este estimador calcula-se um outro erro, estima-se nova sequência de parâmetros da auto-regressão, resultando um novo estimador $\hat{\underline{\theta}}$.

Esta recorrência cessará quando a função de custo atingir um mínimo, obtendo-se para $\hat{\underline{\theta}}$, um estimador não polarizado.

2.10 - RESUMO

O método dos mínimos quadrados foi apresentado, sendo descrito pela equação (2.33). Quando o ruído é colorido sugeriu-se uma estimação dos parâmetros do ruído de modo a obter um estimador não polarizado. Este algoritmo foi descrito pelas equações (2.55) a (2.62).

CAPÍTULO 3

MÉTODOS DE IDENTIFICAÇÃO RECURSIVA

3.1 - INTRODUÇÃO

Em muitas aplicações, é altamente desejável que se obtenham os resultados da identificação de um processo enquanto ele está decorrendo, principalmente quando as características do processo ou do ruído são variáveis com o tempo.

A identificação do processo descrito acima pode ser obtida por estimação recursiva, isto é, a cada nova medida, modifica-se o valor estimado que foi calculado na medida anterior.

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

1) Não é necessária a inversão de uma matriz de ordem elevada. A partir da equação (2.21) observa-se que, se a ordem do modelo for elevada, será necessário inverter uma matriz de alta ordem; isto é indesejável pelo tempo consumido e a quantidade de memória rápida que deve ser usada no computador;

2) O efeito de cada nova observação pode ser facilmente incorporada no estimador.

A seguir, modificar-se-ão os algoritmos deduzidos no capítulo 2, de modo a que os parâmetros do processo possam ser simultaneamente estimados com a sua evolução.

A grosso modo, a idéia será a de atualizar o estimador a cada nova medida, tal como mostra a figura 3.1

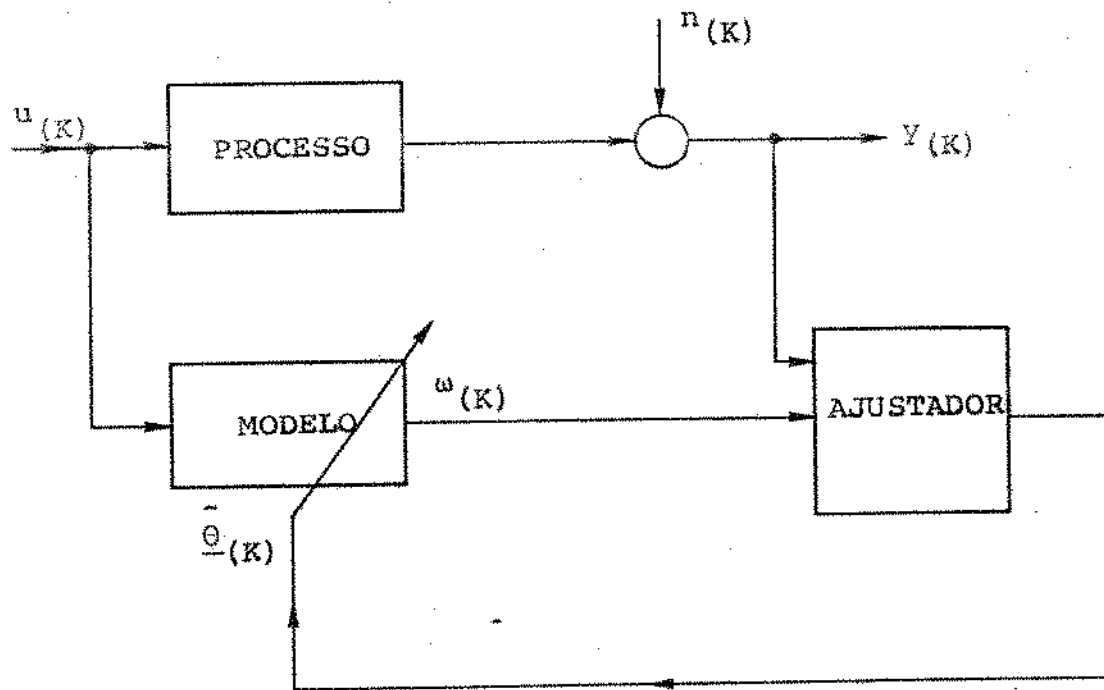


FIGURA 3.1

onde

$$\hat{\theta}(k+1) = \hat{\theta}(k) + A(k) e(k)$$

$A(k)$ - ganho no instante k

$e(k)$ - erro entre a saída do processo e do modelo, no instante k

3.2 - ESTIMADOR RECURSIVO DOS MÍNIMOS QUADRADOS

Neste item, considerando as propriedades apresentadas no item anterior, o estimador dos mínimos quadrados será calculado sequencialmente [2].

Seja o conjunto de observações estocásticas mostradas pela equação (3.1) que é uma generalização da equação (2.10):

$$y_{(i)}^* = H_i \theta + n_i \quad i = 0, 1, 2, 3, \dots \quad (3.1)$$

onde

$y_{(i)}^*$ é um vetor de medidas $n \times 1$ no instante i ,

$$\underline{H}_i = \begin{bmatrix} \underline{u}_{1,i}^T \\ \underline{u}_{2,i}^T \\ \vdots \\ \underline{u}_{n,i}^T \end{bmatrix} \quad \text{e} \quad \underline{u}_{J,i}^T = \begin{bmatrix} u_J(i) & \dots & u_J(i-m) \end{bmatrix}$$

e

$\underline{n}_i$ é um vetor de ruído $n \times 1$

O valor ótimo de $\underline{\theta}$, obtido a partir das medidas estocásticas $\underline{y}_i$ é determinado minimizando-se:

$$E_K(\underline{\theta}) = \sum_{i=0}^K || \underline{H}_i \underline{\theta} - \underline{y}_{(i)}^* ||^2 \quad (3.2)$$

Da equação (3.2) resulta que:

$$E_{K+1}(\underline{\theta}) = E_K(\underline{\theta}) + || \underline{H}_{K+1} \underline{\theta} - \underline{y}_{(K+1)}^* ||^2 \quad (3.3)$$

O valor ótimo $\hat{\underline{\theta}}_K$ é tal que

$$E_K(\hat{\underline{\theta}}_K) \leq E_K(\underline{\theta}) \quad (3.4)$$

para qualquer $\underline{\theta}$ no espaço Euclidiano m dimensional.

Das equações (3.2) e (3.4) e considerando-se que $\underline{\theta} = \hat{\underline{\theta}}_K + \Delta \underline{\theta}$, obtém-se:

$$\Delta \underline{\theta}^T \sum_{i=0}^K \underline{H}_i^T (\underline{H}_i \hat{\underline{\theta}}_K - \underline{y}_{(i)}^*) = 0 \quad (3.5)$$

para qualquer $\Delta \underline{\theta}$ no espaço Euclidiano m dimensional.

Considerando a relação

$$\hat{\theta}_{(K+1)} = \hat{\theta}_{(K)} + \Delta \theta_{(K+1)} \quad (3.6)$$

onde o estimador no instante $K+1$ é igual ao estimador no instante K mais uma parcela que conterá a nova informação disponível.

Reescrevendo a equação (3.3), obtem-se:

$$E_{K+1}(\hat{\theta}_{(K+1)}) = E_K(\hat{\theta}_{(K)}) + \sum_{i=0}^K (\underline{H}_i \Delta \theta_{(K+1)})^T (\underline{H}_i \Delta \theta_{(K+1)}) + \\ + ||\underline{H}_{K+1} \hat{\theta}_{(K)} - \underline{Y}_{(K+1)}^* + \underline{H}_{K+1} \Delta \theta_{(K+1)}||^2$$

$$E_{K+1}(\hat{\theta}_{(K+1)}) = E_K(\hat{\theta}_{(K)}) + ||\underline{H}_{K+1} \hat{\theta}_{(K)} - \underline{Y}_{(K+1)}^*||^2 + \\ + 2 \Delta \theta_{(K+1)}^T \underline{H}_{K+1}^T (\underline{H}_{K+1} \hat{\theta}_{(K)} - \underline{Y}_{(K+1)}^*) + \\ + \Delta \theta_{(K+1)}^T \left(\sum_{i=0}^{K+1} \underline{H}_i^T \underline{H}_i \right) \Delta \theta_{(K+1)} \quad (3.7)$$

O valor ótimo $\Delta \hat{\theta}_{(K+1)}$ que minimiza a equação (3.7) com respeito a $\Delta \theta_{(K+1)}$ é:

$$\Delta \hat{\theta}_{(K+1)} = - \left(\sum_{i=0}^{K+1} \underline{H}_i^T \underline{H}_i \right)^+ \underline{H}_{K+1}^T (\underline{H}_{K+1} \hat{\theta}_{(K)} - \underline{Y}_{(K+1)}^*) \quad (3.8)$$

Se a matriz inversa da matriz $\left(\sum_{i=0}^{K+1} \underline{H}_i^T \underline{H}_i \right)$ existe e

$$\text{definindo } \underline{P}_{(K+1)} = \left(\sum_{i=0}^{K+1} \underline{H}_i^T \underline{H}_i \right)^- \quad (3.9)$$

obtem-se:

$$\hat{\theta}_{(K+1)} = \hat{\theta}_{(K)} + \underline{P}_{(K+1)} \underline{H}_{K+1}^T (\underline{y}_{(K+1)}^* - \underline{H}_{K+1} \hat{\theta}_{(K)}) \quad (3.10)$$

Para o caso de uma saída, isto é, $n = 1$ resulta:

$$\hat{\theta}_{(K+1)} = \hat{\theta}_{(K)} + \underline{P}_{(K+1)} \underline{u}_{(K+1)} (y_{(K+1)} - \underline{u}_{(K+1)}^T \hat{\theta}_{(K)}) \quad (3.11)$$

e

$$\underline{P}_{(K+1)} = \left(\sum_{i=0}^{K+1} \underline{u}_i \underline{u}_i^T \right)^{-1} = \left[\underline{U}^T(K+1) \underline{U}(K+1) \right]^{-1} \quad (3.11.)$$

com a matriz $\underline{U}(K)$ definida pela equação (2.11.b)

A partir da equação (3.11.b) e utilizando a identidade (3.12)

$$\left(\underline{A}^T \underline{A} + \underline{a} \underline{a}^T \right)^{-1} = \left(\underline{A}^T \underline{A} \right)^{-1} - \frac{\left(\underline{A}^T \underline{A} \right)^{-1} \underline{a} \underline{a}^T \left(\underline{A}^T \underline{A} \right)^{-1}}{1 + \underline{a}^T \left(\underline{A}^T \underline{A} \right)^{-1} \underline{a}} \quad (3.12)$$

onde $\underline{A}$ é uma matriz e $\underline{a}$ um vetor, resulta que:

$$\underline{P}_{(K+1)} = \underline{P}_{(K)} - \frac{\underline{P}_{(K)} \underline{u}_{(K+1)} \underline{u}_{(K+1)}^T \underline{P}_{(K)}}{\gamma(K)} \quad (3.13)$$

$$\text{com } \gamma(K) = 1 + \underline{u}_{(K+1)}^T \underline{P}_{(K)} \underline{u}_{(K+1)} \quad (3.14)$$

Utilizando as equações (3.13) e (3.14) a equação para o estimador pode ser reescrita como segue:

$$\hat{\theta}_{(K+1)} = \hat{\theta}_{(K)} + \frac{\underline{P}_{(K)} \underline{u}_{(K+1)}}{\gamma(K)} \left(y_{(K)} - \underline{u}_{(K+1)}^T \hat{\theta}_{(K)} \right) \quad (3.15)$$

A equação (3.15) também pode ser vista como:

$$\hat{\underline{\theta}}_{(K+1)} = \hat{\underline{\theta}}_{(K)} + \frac{\underline{P}_{(K)} \underline{u}_{(K+1)}}{Y_{(K)}} \varepsilon_{(K+1)} \quad (3.16)$$

onde

$$\varepsilon_{(K+1)} = Y_{(K+1)} - \underline{u}_{(K+1)}^T \hat{\underline{\theta}}_{(K)} \quad (3.17)$$

e

$\hat{\underline{\theta}}_{(K)}$ denota o estimador de $\underline{\theta}$, baseado em K medidas.

O resíduo $\varepsilon_{(K+1)}$ é o erro de predição, representado pela diferença entre o valor medido $Y_{(K+1)}$ e o valor predito $\underline{u}_{(K+1)}^T \hat{\underline{\theta}}_{(K)}$.

O par $(\hat{\underline{\theta}}, \underline{P})$ representa o menor número de variáveis, caracterizando os dados de entrada-saída, que devem ser atualizados ao longo da estimação. Esquematicamente este fato é representado pela figura 3.2.

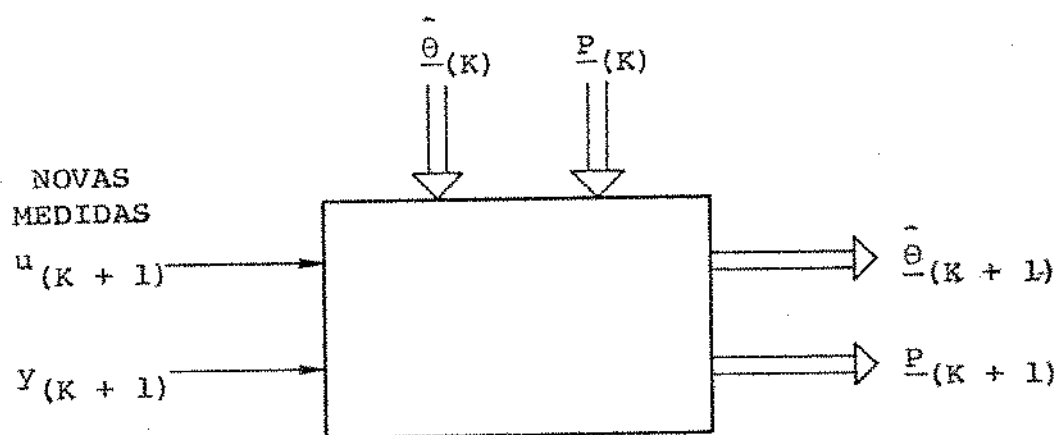


FIGURA 3.2

A partir da equação (3.9), observa-se que a matriz $\underline{P}_{(K)}$ é simétrica, o que portanto só é necessário atualizar os elementos da matriz de um dos lados da diagonal principal.

Neste ponto deve-se ressaltar que este estimador é idêntico ao estimador de mínimos quadrados obtido no capítulo anterior, podendo-se portanto afirmar que:

$$E \left[\hat{\underline{\theta}}_{(K)} \right] = \underline{\theta} \quad (3.18)$$

e

$$\text{Cov} \left[\hat{\underline{\theta}}_{(K)} \right] = \underline{P}_{(K)} \underline{U}_{(K)}^T \underline{N} \underline{U}_{(K)} \underline{P}_{(K)} \quad (3.19)$$

Se $\underline{n}$ é uma sequência de variáveis aleatórias não correlatas, com matriz de covariância $\underline{N} = \sigma_n^2 \underline{I}$, obtém-se:

$$\text{Cov} \left[\hat{\underline{\theta}}_{(K)} \right] = \sigma_n^2 \underline{P}_{(K)} \quad (3.20)$$

e conseqüentemente $\underline{P}_{(K)}$ é proporcional à matriz de covariância do estimador.

Ainda sobre o estimador sequencial observa-se, das equações (3.13) e (3.15), que este método é simples e fácil de ser aplicado, devido à facilidade de tratamento de suas equações por computador.

Permanece a restrição, do ruído ter de ser branco para que o estimador seja não polarizado.

3.3 - ESTIMADOR RECURSIVO DOS MÍNIMOS QUADRADOS COM O RUÍDO MODELADO COMO AUTO-REGRESSÃO

No capítulo anterior verificou-se que o estimador dos mínimos quadrados resulta em valores estimados polarizados, se a sequência de ruído é correlata. A técnica apresentada, no fim do capítulo 2, para resolução do problema acima mencionado, consiste

em filtrar os valores observados (u, y) de modo a se obter resíduos não correlatos e consequentemente estimadores não polarizados.

A seguir, modificar-se-á o procedimento apresentado no item 2.9 de modo a transformá-lo num procedimento recursivo [6], [2], [10].

Parte-se do item 2.9 no qual, além das fórmulas do estimador do processo, foram deduzidos as fórmulas do estimador dos parâmetros do ruído, isto é, dos parâmetros do erro generalizado.

Como os valores observados, naquele item, eram filtrados, a equação (2.55) pode ser reescrita como segue:

$$\hat{\underline{\theta}}_{(K)} = \left[\begin{array}{cc} \underline{\Omega}_{(K)}^{FT} & \underline{\Omega}_{(K)}^F \end{array} \right]^{-1} \underline{\Omega}_{(K)}^{FT} \underline{Y}_{(K)}^F \quad (3.21)$$

onde $\underline{\Omega}_{(K)}^F$ representa a matriz de regressão das medidas filtradas e $\underline{Y}_{(K)}^F$ o vetor de saída filtrado.

Da equação (3.21) obtém-se:

$$\begin{aligned} \hat{\underline{\theta}}_{(K+1)} &= \left[\begin{array}{cc} \underline{\Omega}_{(K+1)}^{FT} & \underline{\Omega}_{(K+1)}^F \end{array} \right]^{-1} \underline{\Omega}_{(K+1)}^{FT} \underline{Y}_{(K+1)}^F = \\ &= \left(\left[\begin{array}{c} \underline{\Omega}_{(K)}^F \\ \vdots \\ \underline{\Omega}_{(K+1)}^{FT} \end{array} \right]^T \left[\begin{array}{c} \underline{\Omega}_{(K)}^F \\ \vdots \\ \underline{\Omega}_{(K+1)}^F \end{array} \right] \right)^{-1} \left[\begin{array}{c} \underline{\Omega}_{(K)}^F \\ \vdots \\ \underline{\Omega}_{(K+1)}^{FT} \end{array} \right]^T \left[\begin{array}{c} \underline{Y}_{(K)}^F \\ \vdots \\ \underline{Y}_{(K+1)}^F \end{array} \right] \end{aligned} \quad (3.22)$$

$$\text{onde } \underline{\Omega}_{(K+1)}^{FT} = \left[\begin{array}{c} u_{(K)}^F \cdots u_{(K+1-m)}^F \\ Y_{(K)}^F \cdots Y_{(K+1-m)}^F \end{array} \right]$$

Desenvolvendo a equação (3.22) de modo análoga ao realizado no item 3.2 resulta:

$$\hat{\underline{\theta}}_{(K+1)} = \hat{\underline{\theta}}_{(K)} + \frac{\underline{P}_{(K)} \underline{\omega}_{(K+1)}^F}{\gamma_{(K)}} \left(y_{(K+1)}^F - \underline{\omega}_{(K+1)}^{FT} \hat{\underline{\theta}}_{(K)} \right) \quad (3.23)$$

De

$$\underline{P}_{(K)} = \left[\begin{array}{cc} \underline{\Omega}_{(K)}^{FT} & \underline{\Omega}_{(K)}^F \end{array} \right]^{-1} \quad (3.24)$$

e

$$\gamma_{(K)} = 1 + \underline{\omega}_{(K+1)}^{FT} \underline{P}_{(K)} \underline{\omega}_{(K+1)}^F$$

resulta

$$\underline{P}_{(K+1)} = \underline{P}_{(K)} - \frac{\underline{P}_{(K)} \underline{\omega}_{(K+1)}^F \underline{\omega}_{(K+1)}^{FT} \underline{P}_{(K)}}{\gamma_{(K)}} \quad (3.25)$$

Em seguida apresenta-se a equação recursiva para a determinação da função de transferência do ruído, levando-se em conta que o ruído está sendo aproximado pelo erro entre a saída verdadeira e a saída do modelo.

Portanto o ruído que atua na (K + 1)-ésima observação da saída é estimado utilizando-se a equação (3.26)

$$\hat{r}_{(K+1)} = y_{(K+1)}^F - \underline{\omega}_{(K+1)}^{FT} \hat{\underline{\theta}}_{(K+1)} \quad (3.26)$$

o que permite que a equação (2.62) seja reescrita como:

$$\hat{\underline{g}}_{(K+1)} = - \left(\begin{array}{c} \left| \begin{array}{c} \hat{\underline{R}}_{(K)} \\ \vdots \\ \hat{\underline{r}}_{(K+1)} \end{array} \right|^T \left| \begin{array}{c} \hat{\underline{R}}_{(K)} \\ \vdots \\ \hat{\underline{r}}_{(K+1)} \end{array} \right| \end{array} \right)^{-1} \begin{array}{c} \left| \begin{array}{c} \hat{\underline{R}}_{(K)} \\ \vdots \\ \hat{\underline{r}}_{(K+1)} \end{array} \right|^T \left| \begin{array}{c} \hat{\underline{r}}_{(K)} \\ \vdots \\ \hat{\underline{r}}_{(K+1)} \end{array} \right| \end{array} \quad (3.27)$$

onde $\underline{r}_{(K+1)}^T = \left[\hat{r}_{(K+1-1)} \cdots \hat{r}_{(K+1-p)} \right]$ e

$$\underline{r}_{(K)}^T = \left[\hat{r}_{(1)} \cdots \hat{r}_{(K)} \right]$$

Tomando as equações (3.12) e (3.27) e desenvolvendo de modo análogo ao caso anterior, obtém-se o estimador dos parâmetros do ruído, que é colocado na equação (3.28)

$$\hat{q}_{(K+1)} = \hat{q}_{(K)} - \frac{\underline{Q}_{(K)} \underline{r}_{(K+1)}^T \left(\hat{r}_{(K+1)} + \underline{r}_{(K+1)}^T \hat{q}_{(K)} \right)}{d_{(K)}} \quad (3.28)$$

$$d_{(K)} = 1 + \underline{r}_{(K+1)}^T \underline{Q}_{(K)} \underline{r}_{(K+1)} \quad (3.29)$$

$$\underline{Q}_{(K)} = \left[\begin{array}{cc} \hat{R}_{(K)}^T & \hat{R}_{(K)} \end{array} \right]^{-1}$$

e

$$\underline{Q}_{(K+1)} = \underline{Q}_{(K)} - \frac{\underline{Q}_{(K)} \underline{r}_{(K+1)}^T \underline{r}_{(K+1)}^T \underline{Q}_{(K)}}{d_{(K)}} \quad (3.30)$$

Portanto com a utilização deste método obtém-se estimadores recursivos para os parâmetros do sistema e do ruído.

3.4 - MÉTODO DA MATRIZ EXTENDIDA

O método de estimação dos parâmetros, apresentado no item anterior, é baseado na afirmação de que o ruído pode ser representado por uma auto-regressão, isto é,

$$r_{(K)} = \frac{1}{G(Z^{-1})} \zeta_{(K)} \quad (3.31)$$

onde $G(z^{-1}) = 1 + g_1 z^{-1} + \dots + g_p z^{-p}$ e $\{\zeta_{(K)}\}$ é uma sequência de variáveis aleatórias não correlatas. Contudo, este método de modelamento do ruído pode se tornar inconveniente quanto ao número de parâmetros envolvidos.

Torna-se, portanto, necessário introduzir uma descrição completa do tipo média-móvel-auto-regressão, para a função de transferência do ruído, a fim de diminuir o número de parâmetros a serem estimados [7].

Esta representação é mais adequada, porque a dinâmica do ruído é tratada da mesma maneira que a do processo, isto é, por estimação dos termos do numerador e denominador da função de transferência e do sistema linear invariante no tempo, que caracteriza o ruído.

O método apresentado no item 3.3 estima separadamente os parâmetros do processo e os parâmetros do ruído, enquanto que o método que será deduzido a seguir, média-móvel-auto-regressão, estima simultaneamente os parâmetros do processo e do ruído.

Seja o processo dado pelas equações:

$$A(z^{-1}) Y_{(K)} = B(z^{-1}) u_{(K)} + n_{(K)} \quad (3.32)$$

Modelado por:

$$A_M(z^{-1}) Y_{(K)} = B_M(z^{-1}) u_{(K)} + e_{(K)} \quad (3.33)$$

Como uma boa aproximação para o ruído é o erro generalizado, supor-se-á que o ruído possa ser descrito como a saída de um filtro com entrada do tipo ruído branco, como descrito pela equação (3.34)

$$e_{(K)} = n_{(K)} = \frac{1 + C_M(z^{-1})}{D_M(z^{-1})} \zeta_{(K)} \quad (3.34)$$

Utilizando a equação (3.34) é possível reescrever a equação (3.33), resultando:

$$A_M(z^{-1}) Y_{(K)} = B_M(z^{-1}) u_{(K)} - D_M(z^{-1}) n_{(K)} + C_M(z^{-1}) \zeta_{(K)} + \zeta_{(K)} \quad (3.35)$$

onde $A_M(z^{-1}) = 1 + \alpha_1 z^{-1} + \dots + \alpha_n z^{-n}$

$$B_M(z^{-1}) = \beta_1 z^{-1} + \dots + \beta_n z^{-n}$$

$$D_M(z^{-1}) = 1 + \delta_1 z^{-1} + \dots + \delta_r z^{-r}$$

$$C_M(z^{-1}) = \gamma_1 z^{-1} + \dots + \gamma_s z^{-s}$$

Utilizando-se a notação matricial e supondo-se $r=s$, obtém-se:

$$\underline{Y} = \underline{\Omega} \underline{\Theta} + \underline{\zeta} \quad (3.36)$$

onde $\underline{Y} = \begin{bmatrix} Y_{(1)} \\ \vdots \\ Y_{(K)} \end{bmatrix}$ e

$$\underline{\Omega} = \begin{bmatrix} u_{(0)} \dots u_{(1-n)} & | & Y_{(0)} \dots Y_{(1-n)} & | & \zeta_{(0)} \dots \zeta_{(1-s)} & | & n_{(0)} \dots n_{(1-s)} \\ \vdots & & \vdots & & \vdots & & \vdots \\ u_{(K-1)} \dots u_{(K-n)} & | & Y_{(K-1)} \dots Y_{(K-n)} & | & \zeta_{(K-1)} \dots \zeta_{(K-s)} & | & n_{(K-1)} \dots n_{(K-s)} \end{bmatrix}$$

$$\underline{\Omega} = \begin{bmatrix} \underline{u} & | & \underline{Y} & | & \underline{\zeta} & | & \underline{N} \end{bmatrix}$$

$$\underline{\theta} = \begin{bmatrix} \underline{\beta} \\ -\underline{\alpha} \\ \underline{\gamma} \\ -\underline{\delta} \end{bmatrix} \quad \underline{\zeta} = \begin{bmatrix} \zeta(1) \\ \zeta(2) \\ \vdots \\ \zeta(K) \end{bmatrix}$$

Deve ser notado que a ordem do processo e do ruído são assumidas conhecidas.

Analogamente ao deduzido para os mínimos quadrados recursivos, no item 3.2, obtém-se:

$$\underline{\theta}_{(K+1)} = \underline{\theta}_{(K)} + \frac{\underline{P}_{(K)} \underline{\omega}_{(K+1)}}{\gamma_{(K)}} \left\{ Y_{(K+1)} - \underline{\omega}_{(K+1)}^T \underline{\theta}_{(K)} \right\} \quad (3.37)$$

$$\underline{P}_{(K+1)} = \underline{P}_{(K)} - \frac{\underline{P}_{(K)} \underline{\omega}_{(K+1)} \underline{\omega}_{(K+1)}^T \underline{P}_{(K)}}{\gamma_{(K)}} \quad (3.38)$$

$$\gamma_{(K)} = 1 + \underline{\omega}_{(K+1)}^T \underline{P}_{(K)} \underline{\omega}_{(K+1)} \quad (3.39)$$

$$\underline{\omega}_{(K+1)}^T = \begin{bmatrix} u_{(K)}, \dots, u_{(K+1-n)}, Y_{(K)}, \dots, Y_{(K+1-n)}, \\ \hat{\zeta}_{(K)}, \dots, \hat{\zeta}_{(K+1-s)}, \hat{n}_{(K)}, \dots, \hat{n}_{(K+1-s)} \end{bmatrix} \quad (3.40)$$

onde

$\hat{\zeta}$ e $\hat{n}$ são valores estimados dos elementos da matriz $\underline{\Psi}$ e $\underline{N}$ respectivamente, obtidos a partir das seguintes equações:

$$\hat{n}_{(K)} = Y_{(K)} - \underline{\omega}_{(K)}^T \begin{bmatrix} \underline{\beta}_{(K)} \\ -\underline{\alpha}_{(K)} \end{bmatrix} \quad (3.41)$$

$$\hat{\underline{z}}_{(K)} = \hat{\underline{n}}_{(K)} - \underline{\omega}_{(K)}^{2T} \begin{bmatrix} \underline{y}_{(K)} \\ -\hat{\underline{\delta}}_{(K)} \end{bmatrix} \quad (3.42)$$

onde

$$\underline{\omega}_{(K)}^{1T} = \begin{bmatrix} u_{(K-1)} \cdots u_{(K-n)} & y_{(K-1)} \cdots y_{(K-n)} \end{bmatrix} \quad e$$

$$\underline{\omega}_{(K)}^{2T} = \begin{bmatrix} \hat{\zeta}_{(K-1)} \cdots \hat{\zeta}_{(K-s)} & \hat{n}_{(K-1)} \cdots \hat{n}_{(K-s)} \end{bmatrix}$$

Portanto as equações (3.37), (3.38) e (3.39) representam o estimador para um modelo média-móvel-auto-regressão.

3.5 - RUÍDO MODELADO COMO MÉDIA-MÓVEL

No item anterior considerou-se um processo, com relação entrada-saída, descrito por:

$$A(z^{-1}) y_{(K)} = B(z^{-1}) u_{(K)} + n_{(K)} \quad (3.43)$$

e modelado por:

$$A_M(z^{-1}) y_{(K)} = B_M(z^{-1}) u_{(K)} + e_{(K)} \quad (3.44)$$

$$\text{onde } e_{(K)} \triangleq n_{(K)} = \frac{1 + C_M(z^{-1})}{D_M(z^{-1})} \zeta_{(K)}$$

Um caso particular deste algoritmo resulta quando se supõe que $D_M(z^{-1}) = 1$, isto é, quando o ruído é modelado como um processo média-móvel [8].

Este modelo também pode ser usado para modelar séries temporais, isto é, sem os parâmetros de B_M .

Analogamente, ao item anterior, obtém-se:

$$Y(K) = \underline{\Omega} \underline{\Theta} + \underline{\zeta}(K) \quad (3.45)$$

onde agora

$$\underline{\Omega} = \begin{bmatrix} u(0) \cdots \cdots u(1-n) & | & -Y(0) \cdots \cdots -Y(1-n) & | & \zeta(0) \cdots \cdots \zeta(1-s) \\ \vdots & & \vdots & & \vdots \\ u(K-1) \cdots \cdots u(K-n) & | & -Y(K-1) \cdots \cdots -Y(K-n) & | & \zeta(K-1) \cdots \cdots \zeta(K-s) \end{bmatrix} \quad (3.46)$$

e

$$\underline{\Theta} = \begin{bmatrix} \underline{\beta} \\ \underline{\alpha} \\ \underline{\gamma} \end{bmatrix}$$

Destas equações resulta o algoritmo média-móvel descrito pelas equações abaixo

$$\underline{\Theta}(K+1) = \hat{\underline{\Theta}}(K) + \frac{P(K) \underline{\omega}(K+1)}{\gamma(K)} \varepsilon(K+1) \quad (3.47)$$

$$P(K+1) = P(K) - \frac{P(K) \underline{\omega}(K+1) \underline{\omega}^T(K+1) P(K)}{\gamma(K)} \quad (3.48)$$

$$\varepsilon(K+1) = Y(K+1) - \underline{\omega}^T(K+1) \hat{\underline{\Theta}}(K) \quad (3.49)$$

$$\gamma(K) = 1 + \underline{\omega}^T(K+1) P(K) \underline{\omega}(K+1) \quad (3.50)$$

$$\underline{w}_{(K+1)}^T = \begin{bmatrix} u_{(K)} \cdots u_{(K+1-n)} & -Y_{(K)} \cdots -Y_{(K+1-n)} & \hat{\zeta}_{(K)} \cdots \hat{\zeta}_{(K+1-s)} \end{bmatrix} \quad (3.51)$$

Entretanto, como não é possível medir $\zeta_{(K)}$, faz-se uma aproximação deste valor por $\hat{\zeta}_{(K)}$, onde:

$$\hat{\zeta}_{(K)} = Y_{(K)} - \underline{w}_{(K)}^T \hat{\theta}_{(K-1)} \quad (3.52)$$

3.6 - VERSÃO DOS ALGORÍTMOS EM TEMPO REAL

Para a estimação de parâmetros variáveis no tempo, os algoritmos deduzidos neste capítulo deverão sofrer algumas alterações. Neste caso, devido à evolução dos parâmetros, deve-se dar um maior peso às últimas medidas feitas, em relação às anteriores. Uma maneira de atingir este objetivo, consiste em modificar os algoritmos deduzidos, introduzindo um fator de ponderação que atue mais sobre as últimas medidas. Por exemplo, no algoritmo dos mínimos quadrados, este fato pode ser representado modificando-se a equação (3.11.b)

$$\underline{w}_{(K+1)}^T \underline{u}_{(K+1)} = \left[C \underline{u}_{(K)} \right]^T \left[C \underline{u}_{(K)} \right] + \underline{u}_{(K+1)}^T \underline{u}_{(K+1)} \quad (3.53)$$

com $0 < C \leq 1$

Esquemáticamente, esta idéia é representada na figura

3.3

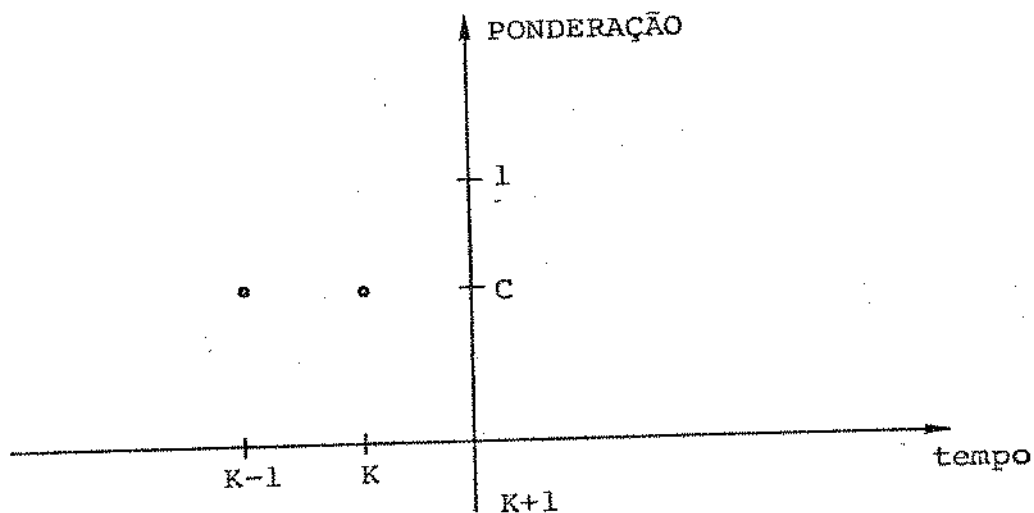


FIGURA 3.3

Desenvolvendo (3.53), resulta:

$$C^{-2} \underline{P}_{(K+1)}^{-1} = \underline{P}_{(K)}^{-1} + \underline{u}_{(K+1)} C^{-2} \underline{u}_{(K+1)}^T \quad (3.54)$$

Sendo $\rho = C^2$, obtem-se: (3.55)

$$\underline{P}_{(K+1)}^{-1} = \rho \underline{P}_{(K)}^{-1} + \underline{u}_{(K+1)} \underline{u}_{(K+1)}^T \quad (3.56)$$

$$\begin{aligned} \underline{P}_{(K+1)} &= \underline{u}_{(K+1)}^T \underline{u}_{(K+1)}^{-1} = \left(\begin{bmatrix} C \underline{u}_{(K)} \\ \underline{u}_{(K+1)}^T \end{bmatrix}^T \begin{bmatrix} C \underline{u}_{(K)} \\ \underline{u}_{(K+1)}^T \end{bmatrix} \right)^{-1} = \\ &= \left((C \underline{u}_{(K)})^T (C \underline{u}_{(K)}) + \underline{u}_{(K+1)} \underline{u}_{(K+1)}^T \right)^{-1} = \\ &= \rho^{-1} (\underline{u}_{(K)}^T \underline{u}_{(K)})^{-1} = \\ &= \frac{\left((C \underline{u}_{(K)})^T C \underline{u}_{(K)} \right)^{-1} \underline{u}_{(K+1)} \underline{u}_{(K+1)}^T \left((C \underline{u}_{(K)})^T C \underline{u}_{(K)} \right)^{-1}}{1 + \rho^{-1} \underline{u}_{(K+1)}^T \underline{P}_{(K)} \underline{u}_{(K)}} \\ &= \rho^{-1} \underline{P}_{(K)} - \frac{\underline{P}_{(K)} \underline{u}_{(K+1)} \underline{u}_{(K+1)}^T \rho^{-1} \underline{P}_{(K)}}{\rho + \underline{u}_{(K+1)}^T \underline{P}_{(K)} \underline{u}_{(K+1)}} \\ \underline{P}_{(K+1)} &= \rho^{-1} \left[\underline{P}_{(K)} - \frac{\underline{P}_{(K)} \underline{u}_{(K+1)} \underline{u}_{(K+1)}^T \underline{P}_{(K)}}{\gamma(K)} \right] \quad (3.57) \end{aligned}$$

$$Y_{(K)} = \rho + \underline{u}_{(K+1)}^T \underline{P}_{(K)} \underline{u}_{(K+1)} \quad (3.58)$$

Já foi mostrado que:

$$\hat{\underline{\theta}}_{(K+1)} = \underline{P}_{(K+1)} \underline{u}_{(K+1)}^T Y_{(K+1)} \quad (3.59)$$

Portanto

$$\hat{\underline{\theta}}_{(K+1)} = \left[\begin{pmatrix} C \underline{u}_{(K)} \\ \underline{u}_{(K+1)}^T \end{pmatrix}^T \quad \begin{pmatrix} C \underline{u}_{(K)} \\ \underline{u}_{(K+1)}^T \end{pmatrix} \right]^{-1} \begin{bmatrix} C \underline{u}_{(K)} \\ \underline{u}_{(K+1)}^T \end{bmatrix}^T \begin{bmatrix} C Y_{(K)} \\ Y_{(K+1)} \end{bmatrix} \quad (3.60)$$

$$= \begin{bmatrix} C^2 \underline{u}_{(K)}^T \underline{u}_{(K)} + \underline{u}_{(K+1)} \underline{u}_{(K+1)}^T \end{bmatrix}^{-1} \cdot \begin{bmatrix} C^2 \underline{u}_{(K)}^T Y_{(K)} + \underline{u}_{(K+1)} Y_{(K+1)} \end{bmatrix} \quad (3.61)$$

$$= \rho^{-1} \left[\underline{P}_{(K)} - \frac{\underline{P}_{(K)} \underline{u}_{(K+1)} \underline{u}_{(K+1)}^T \underline{P}_{(K)}}{Y_{(K)}} \right] \cdot \rho \left[\underline{u}_{(K)}^T Y_{(K)} + \frac{\underline{u}_{(K+1)} Y_{(K+1)}}{\rho} \right] \quad (3.62)$$

$$\hat{\underline{\theta}}_{(K+1)} = \hat{\underline{\theta}}_{(K)} + \underline{K}_{(K)} \left\{ Y_{(K)} - \underline{u}_{(K+1)}^T \hat{\underline{\theta}}_{(K)} \right\} \quad (3.63)$$

$$\underline{K}_{(K)} = \frac{\underline{P}_{(K)} \underline{u}_{(K+1)}}{Y_{(K)}} \quad (3.64)$$

onde ρ é o fator de ponderação.

A influência do fator de ponderação (ρ) no comportamento transitório e estacionário do estimador será analisada no capítulo 5.

Quando são estimados os parâmetros do sistema e também os do ruído, há necessidade de mudanças nas equações obtidas no item 3.3, para que as últimas medidas possam influir mais que as anteriores. Isto deve ocorrer porque o estimador do sistema opera com dados filtrados que variam continuamente com o desenvolvimento do processo.

Nos outros dois algoritmos apresentados, também é necessário introduzir este fator de ponderação, porque, no início, os valores estimados são pouco confiáveis e conseqüentemente obtêm-se maus valores estimados para o ruído.

A introdução do fator de ponderação e as modificações efetuadas nas equações dos algoritmos são análogas às modificações apresentadas anteriormente para o algoritmo dos mínimos quadrados.

3.7 - AUMENTO DA TAXA DE CONVERGÊNCIA

Os estimadores podem, em algumas condições, convergir muito lentamente. A magnitude das variações dos estimadores, em cada passo, depende do ganho $K_{(K)}$, que geralmente decresce com o tempo. Pode ocorrer que o ganho se torne muito pequeno, de modo que os estimadores demorem a atingir os valores corretos. Uma maneira óbvia de evitar este fato é aumentar o ganho $K_{(K)}$, o que pode ser realizado, como no parágrafo anterior, introduzindo um fator de ponderação. Entretanto, pode-se verificar (será mostrado no capítulo 5) que os estimadores finalizam com os elementos da matriz de covariância maiores do que os obtidos sem o fator de ponderação. Para resolver este problema introduz-se um fator de ponderação variável no tempo [7], do seguinte tipo:

$$\rho_{(K+1)} = \rho_0 \rho_{(K)} + (1 - \rho_0) \quad (3.65)$$

onde ρ_0 é escolhido tendendo para 1, (por exemplo 0.99).

A equação (3.65) mostra que $\rho(k)$ tende exponencialmente para a unidade quando o número de medidas cresce.

Esta mudança pode ser encarada como uma generalização do algoritmo em tempo real, generalização esta que permite uma maior ponderação nos primeiros valores estimados, além de que, como o fator é construído de modo a tender para a unidade, a precisão dos parâmetros estimados será melhor que aquela obtida no item anterior.

Outra maneira de solucionar este problema é introduzir o fator de ponderação constante, um pouco menor do que a unidade, para um determinado número de medidas, (por exemplo 300 a 400), tomando-se a partir desta o fator de ponderação igual à unidade, o que obriga a variância do estimador não ser muito maior que a obtida sem o fator de ponderação.

3.8 - INICIALIZAÇÃO DOS ALGORÍTMOS

Valores iniciais dos parâmetros $\hat{\theta}(k)$ e da matriz $P(k)$ devem ser fornecidos quando se inicializam quaisquer dos algoritmos anteriores. A matriz $P_{(0)}$ deve refletir a confiança nos valores iniciais de $\hat{\theta}$.

Se os parâmetros são completamente desconhecidos, uma escolha comum é colocá-los iguais a zero e, neste caso, os elementos da matriz de covariância devem ter valores elevados (por exemplo $P_{(0)} = (10 \div 100) \underline{I}$, onde $\underline{I}$ é a matriz identidade).

Se existe uma aproximação inicial para os parâmetros do sistema, esses valores podem ser usados como condição inicial de $\hat{\theta}$. Neste caso os elementos da matriz $P_{(0)}$ podem assumir valores "pequenos", (por exemplo $P_{(0)} = (10 \div 10^{-4}) \underline{I}$), traduzindo a confiança inicial nos valores fornecidos para $\hat{\theta}_{(0)}$.

3.9 - DETERMINAÇÃO DA ORDEN DO MODELO

Na prática a ordem do sistema deve ser conhecida, caso contrário, sérios erros ocorrem por modelamento errado.

Assim é importante dispor de alguns métodos para se determinar esta ordem.

Descreve-se nesta secção um método para avaliação do número de parâmetros de um modelo. A fim de determinar a ordem do sistema fazem-se estimações para modelos de diferentes ordens e analisam-se os resíduos, ou seja, a função de custo [2], [17].

Considerando o modelo generalizado, obtém-se:

$$\underline{e} = \underline{y} - \underline{\Omega} \hat{\underline{\theta}} = \underline{\Omega} \underline{\theta} + \underline{n} - \underline{\Omega} \hat{\underline{\theta}} \quad (3.66)$$

$$\underline{e} = \underline{\Omega} (\underline{\theta} - \hat{\underline{\theta}}) + \underline{n} \quad (3.67)$$

e $E = \underline{e}^T \underline{e}$, onde $\underline{e}$ é o erro generalizado e E a função de custo.

Quando não se tem nenhum conhecimento a priori, inicia-se o algoritmo com modelo de ordem zero e calcula-se esta função. Em seguida, a ordem do modelo é aumentada, obtendo-se nova função de custo. Repetindo-se este procedimento para ordens crescentes, várias funções de custo são determinadas. A ordem correta do modelo é encontrada analisando estas funções.

Seja, por exemplo, um modelo de ordem q e outro de ordem m , com $m > q$. Para se comparar um modelo com outro e concluir que é necessário um de maior ordem é necessário definir-se um critério.

Seja

$$\underline{e}_q = \underline{y} - \underline{\Omega}_q \hat{\underline{\theta}}_q \quad \text{logo} \quad E_q = \underline{e}_q^T \underline{e}_q \quad (3.68)$$

e

$$\underline{e}_m = \underline{y} - \underline{\Omega}_m \hat{\underline{\theta}}_m \quad \text{logo} \quad E_m = \underline{e}_m^T \underline{e}_m \quad (3.69)$$

Definindo-se a variável

$$\chi = \frac{(\bar{E}_q - \bar{E}_m) / (m - q)}{\bar{E}_m / (K - m)} \quad (3.70)$$

onde K é o número de medidas realizadas e χ é uma variável aleatória distribuída segundo F_{K-m}^m , que para $K \gg m$ é uma qui-quadrada $\left(\frac{\lambda_m^2}{m - q} \right)$, se o ruído for gaussiano [1], [2], [3].

Definindo um erro admissível de 5%, da tabela da qui-quadrada, sabe-se que:

$$P \left[\lambda^2 (1) < 3 \right] = .95 \quad (3.71)$$

Desta maneira, obtido o valor de χ pela expressão (3.70), se $\chi > \frac{\lambda_m^2}{m - q}$, a ordem m deverá ser utilizada como q para um novo procedimento, até que $\chi < \frac{\lambda_m^2}{m - q}$, adotando-se então q como a ordem deste modelo.

3.10 - RESUMO

Neste capítulo apresentou-se a versão recursiva para os algoritmos de estimação do capítulo 2.

Algoritmo dos mínimos quadrados recursivo.

$$\hat{\theta}_{(K+1)} = \hat{\theta}_{(K)} + \underline{P}_{(K)} \underline{u}_{(K+1)} \left[y_{(K)} - \underline{u}_{(K+1)}^T \hat{\theta}_{(K)} \right] \quad (3.72)$$

$$\underline{P}_{(K+1)} = \underline{P}_{(K)} - \frac{\underline{P}_{(K)} \underline{u}_{(K+1)} \underline{u}_{(K+1)}^T \underline{P}_{(K)}}{\gamma_{(K)}} \quad (3.73)$$

$$\gamma_{(K)} = 1 + \underline{u}_{(K+1)}^T \underline{P}_{(K)} \underline{u}_{(K+1)} \quad (3.74)$$

Algoritmo auto-regressivo.

$$\hat{\underline{\theta}}_{(K+1)} = \hat{\underline{\theta}}_{(K)} + \frac{\underline{P}_{(K)} \underline{\omega}_{(K+1)}^F}{\gamma_{(K)}} \left[Y_{(K+1)}^F - \underline{\omega}_{(K+1)}^{FT} \hat{\underline{\theta}}_{(K)} \right] \quad (3.75)$$

$$\underline{P}_{(K+1)} = \underline{P}_{(K)} - \frac{\underline{P}_{(K)} \underline{\omega}_{(K+1)}^F \underline{\omega}_{(K+1)}^{FT} \underline{P}_{(K)}}{\gamma_{(K)}} \quad (3.76)$$

$$\hat{\underline{q}}_{(K+1)} = \hat{\underline{q}}_{(K)} - \frac{\underline{Q}_{(K)} \underline{r}_{(K+1)}^1}{d_{(K)}} \left(r_{(K+1)} + \underline{r}_{(K+1)}^{1T} \hat{\underline{q}}_{(K)} \right) \quad (3.77)$$

$$\underline{Q}_{(K+1)} = \underline{Q}_{(K)} - \frac{\underline{Q}_{(K)} \underline{r}_{(K+1)}^1 \underline{r}_{(K+1)}^{1T} \underline{Q}_{(K)}}{d_{(K)}} \quad (3.78)$$

com

$$\underline{\omega}_{(K+1)}^{FT} = \left[\underline{u}_{(K)}^F \cdots \underline{u}_{(K+1-m)}^F \right] \begin{vmatrix} Y_{(K)}^F & \cdots & Y_{(K+1-m)}^F \end{vmatrix} \quad (3.79)$$

$\underline{u}_{(K)}^F$ e $Y_{(K)}^F$ são as medidas filtradas no instante K .

$$\gamma_{(K)} = 1 + \underline{\omega}_{(K+1)}^{FT} \underline{P}_{(K)} \underline{\omega}_{(K+1)}^F \quad (3.80)$$

$$\underline{r}_{(K+1)}^{1T} = \left[\hat{r}_{(K)} \cdots \hat{r}_{(K+1-p)} \right] \quad (3.81)$$

$$\hat{r}_{(K)} = Y_{(K)} - \underline{\omega}_{(K)}^{FT} \hat{\underline{\theta}}_{(K)} \quad (3.82)$$

$$d_{(K)} = 1 + \underline{r}_{(K+1)}^T \underline{Q}_{(K)} \underline{r}_{(K+1)}^1 \quad (3.83)$$

Algoritmo média-móvel auto-regressivo.

$$\hat{\underline{\theta}}_{(K+1)} = \hat{\underline{\theta}}_{(K)} + \frac{\underline{P}_{(K)} \underline{\omega}_{(K+1)}}{\gamma_{(K)}} \left[y_{(K+1)} - \underline{\omega}_{(K+1)}^T \hat{\underline{\theta}}_{(K)} \right] \quad (3.84)$$

$$\underline{P}_{(K+1)} = \underline{P}_{(K)} - \frac{\underline{P}_{(K)} \underline{\omega}_{(K+1)} \underline{\omega}_{(K+1)}^T \underline{P}_{(K)}}{\gamma_{(K)}} \quad (3.85)$$

$$\gamma_{(K)} = 1 + \underline{\omega}_{(K+1)}^T \underline{P}_{(K)} \underline{\omega}_{(K+1)} \quad (3.86)$$

$$\underline{\omega}_{(K+1)}^T = \left[u_{(K)} \cdots u_{(K+1-n)} y_{(K)} \cdots y_{(K+1-n)} \hat{\zeta}_{(K)} \cdots \hat{\zeta}_{(K+1-s)} \hat{n}_{(K)} \cdots \hat{n}_{(K+1-s)} \right] \quad (3.87)$$

$$\hat{n}_{(K)} = y_{(K)} - \underline{\omega}_{(K)}^T \begin{bmatrix} \underline{\beta}_{(K)} \\ -\underline{\alpha}_{(K)} \end{bmatrix} \quad (3.88)$$

$$\hat{\zeta}_{(K)} = \hat{n}_{(K)} - \underline{\omega}_{(K)}^2 \begin{bmatrix} \underline{\gamma}_{(K)} \\ -\underline{\delta}_{(K)} \end{bmatrix} \quad (3.89)$$

$$\underline{\omega}_{(K)}^1 = \begin{bmatrix} u_{(K-1)} \cdots u_{(K-n)} y_{(K-1)} \cdots y_{(K-n)} \end{bmatrix} \quad (3.90)$$

$$\underline{\omega}_{(K)}^2 = \begin{bmatrix} \hat{\zeta}_{(K-1)} \cdots \hat{\zeta}_{(K-s)} \hat{n}_{(K-1)} \cdots \hat{n}_{(K-s)} \end{bmatrix} \quad (3.91)$$

Algoritmo média-móvel

$$\hat{\theta}_{(K+1)} = \hat{\theta}_{(K)} + \frac{P_{(K)} \omega_{(K+1)}}{\gamma_{(K)}} \quad (3.92)$$

$$P_{(K+1)} = P_{(K)} - \frac{P_{(K)} \omega_{(K+1)} \omega_{(K+1)}^T P_{(K)}}{\gamma_{(K)}} \quad (3.93)$$

$$\varepsilon_{(K+1)} = Y_{(K+1)} - \omega_{(K+1)}^T \hat{\theta}_{(K)} \quad (3.94)$$

$$\gamma_{(K)} = 1 + \omega_{(K+1)}^T P_{(K)} \omega_{(K+1)} \quad (3.95)$$

$$\omega_{(K+1)}^T = \left[u_{(K)} \dots u_{(K+1-n)} \quad -Y_{(K)} \dots -Y_{(K+1-n)} \quad \hat{\zeta}_{(K)} \dots \hat{\zeta}_{(K+1-s)} \right] \quad (3.96)$$

$$\hat{\zeta}_{(K)} = Y_{(K)} - \omega_{(K)}^T \hat{\theta}_{(K-1)} \quad (3.97)$$

Neste capítulo também se apresentou um método para determinação da ordem do modelo.

CAPÍTULO 4

CONVERGÊNCIA DOS ALGORÍTMOS DE IDENTIFICAÇÃO

4.1 - INTRODUÇÃO

Os algoritmos deduzidos são equações às diferenças estocásticas não lineares, variáveis no tempo, com propriedades de convergência difíceis de serem analisadas. Outros pontos importantes são a consistência do estimador e a unicidade da solução.

Uma das possibilidades de se estudar os algoritmos consiste na simulação dos mesmos através do Método de Monte-Carlo, para evitar possíveis erros devido às particulares sequências de ruído. Entretanto tal tipo de simulação aumenta sensivelmente o tempo de computação e a memória requerida, podendo ser substituído por métodos analíticos conforme descrito neste capítulo.

Ljung [8], [9] fez uma primeira tentativa para análise dos algoritmos deduzidos no capítulo anterior e mostra no seu trabalho como um sistema de equações diferenciais ordinárias determinísticas pode ser associada com o algoritmo estocástico.

Neste capítulo as principais idéias de Ljung são apresentadas, dando-se uma introdução heurística da solução encontrada para a associação dos algoritmos estocásticos com uma equação diferencial ordinária determinística e os principais resultados obtidos.

4.2 - ANÁLISE DE CONVERGÊNCIA

Para que se possa exemplificar, de maneira mais compreensível, o proposto no item 4.1, é feito um estudo heurístico do algoritmo estocástico, descrito pela equação (4.1), que é uma simplificação do algoritmo dos mínimos quadrados apresentado no capítulo 3 [8], [9], [11].

$$\text{Seja} \quad \hat{\theta}_{(K)} = \hat{\theta}_{(K-1)} + \frac{1}{K} \omega_{(K)}^T \varepsilon_{(K)} \quad (4.1)$$

$$\text{com } \varepsilon_{(K)} = Y_{(K)} - \underline{\omega}_{(K)} \hat{\underline{\theta}}_{(K-1)} \quad (4.2.a)$$

$$\text{e } \underline{\omega}_{(K)} = \begin{bmatrix} -Y_{(K-1)}, \dots, u_{(K-1)}, \dots, \varepsilon_{(K-1)} \dots \end{bmatrix} \quad (4.2.b)$$

onde $\underline{\omega}_{(K)}$ é o vetor que contém as informações das medidas anteriores ao instante K e $\varepsilon_{(K)}$ é o erro entre o valor medido no instante K e o valor predito da saída neste instante.

$$\text{Portanto } \varepsilon_{(K)} = \varepsilon(K; \hat{\underline{\theta}}_{(K-1)}, \hat{\underline{\theta}}_{(K-2)}, \dots, \hat{\underline{\theta}}_{(0)}) \quad \text{e}$$

$$\underline{\omega}_{(K)} = \underline{\omega}(K; \hat{\underline{\theta}}_{(K-1)}, \hat{\underline{\theta}}_{(K-2)}, \dots, \hat{\underline{\theta}}_{(0)})$$

A interrelação entre $\hat{\underline{\theta}}_{(K)}$, $\varepsilon_{(K)}$ e $\underline{\omega}_{(K)}$ dificulta sensivelmente a análise do algoritmo apresentado.

Entretanto, pode ser feita uma simplificação supondo-se que o estimador converge para um certo valor $\underline{\theta}_c$; neste caso as funções $\varepsilon_{(K)}$ e $\underline{\omega}_{(K)}$ tendem para as funções $\varepsilon_{(K)} = \varepsilon^*(K, \underline{\theta}_c)$ e $\underline{\omega}_{(K)} = \underline{\omega}^*(K, \underline{\theta}_c)$ respectivamente, obtidas das equações (4.2.a) e (4.2.b), substituindo-se $\hat{\underline{\theta}}_{(.)}$ por $\underline{\theta}_c$.

Tendo-se como base esta afirmação, pode-se concluir que as funções $\varepsilon_{(N)}$ e $\underline{\omega}_{(N)}$, tendem para $\varepsilon^*(N, \hat{\underline{\theta}}_{(N-1)})$ e $\underline{\omega}^*(N, \hat{\underline{\theta}}_{(N-1)})$ respectivamente, quando $N \rightarrow \infty$, visto que da equação (4.1) conclui-se que a diferença $\hat{\underline{\theta}}_{(K)} - \hat{\underline{\theta}}_{(K-1)}$ decresce à medida que o valor de K aumenta e que de modo geral o estimador $\hat{\underline{\theta}}_{(K)}$ varia lentamente quando o valor de K cresce.

Portanto da equação (4.1) resulta:

$$\hat{\underline{\theta}}_{(K+T)} = \hat{\underline{\theta}}_{(K)} + \sum_{i=K+1}^{K+T} \frac{1}{i} \underline{\omega}_{(i)}^T \varepsilon_{(i)} \quad (4.3)$$

$$\hat{\underline{\theta}}_{(K+T)} = \hat{\underline{\theta}}_{(K)} + \sum_{i=K+1}^{K+T} \frac{1}{i} \left[\underline{\omega}_{(i)}^{*T}, \hat{\underline{\theta}}_{(.)} \varepsilon_{(i)}^*, \hat{\underline{\theta}}_{(.)} \right] \quad (4.4)$$

Da equação (4.1) vê-se que a diferença $\hat{\theta}_{(K)} - \hat{\theta}_{(K-1)}$ varia lentamente quando K tende para infinito, sendo portanto necessário um grande número de amostras, por exemplo T , para que o estimador sofra uma pequena variação.

Esta variação é causada pelas variáveis aleatórias $\omega_{(i)}^{*T}, \hat{\theta}_{(i)}^*$, as quais tem aproximadamente o mesmo valor médio $E \left[\omega_{(K)}^{*T}, \hat{\theta}_{(K)}^* \right]$. Logo é razoável assumir que a variação do estimador é proporcional a este valor, isto é:

$$\hat{\theta}_{(K+T)} = \hat{\theta}_{(K)} + \Delta \tau E \left[\omega_{(K)}^{*T}, \hat{\theta}_{(K)}^* \right] \quad (4.5)$$

$$\text{onde} \quad \Delta \tau = \sum_{i=K+1}^{K+T} \frac{1}{i} \quad e \quad (4.5.a)$$

A equação (4.5) pode ser vista como uma aproximação da equação diferencial ordinária descrita por:

$$\frac{d}{d\tau} \hat{\theta}_{(\tau)} = f \left[\hat{\theta}_{(\tau)} \right] \quad (4.6)$$

$$\text{onde} \quad f[\theta] = E \left[\omega_{(K)}^{*T}, \theta^* \right]$$

4.3 - OBTENÇÃO DAS EQUAÇÕES DE LJUNG

Relembrando os algoritmos deduzidos no capítulo anterior é fácil mostrar que eles podem ser unificados como segue:

$$\hat{\theta}_{(K+1)} = \hat{\theta}_{(K)} + \frac{K}{(K+1)} \epsilon_{(K+1)} \quad (4.7.a)$$

$$\underline{K}_{(K+1)} = \frac{\underline{P}_{(K)} \underline{\omega}_{(K+1)}^T}{1 + \underline{\omega}_{(K+1)}^T \underline{P}_{(K)} \underline{\omega}_{(K+1)}} \quad (4.7.b)$$

$$\underline{P}_{(K+1)} = \underline{P}_{(K)} - \frac{\underline{P}_{(K)} \underline{\omega}_{(K+1)}^T \underline{\omega}_{(K+1)} \underline{P}_{(K)}}{1 + \underline{\omega}_{(K+1)}^T \underline{P}_{(K)} \underline{\omega}_{(K+1)}} \quad (4.7.c)$$

onde $\underline{\omega}_{(K+1)}$ e $\epsilon_{(K+1)}$ são calculados para cada algoritmo de acordo com o apresentado no capítulo anterior.

Utilizando o mesmo argumento apresentado no item anterior é possível mostrar [8] que uma equação diferencial ordinária (Equação de Ljung), também pode ser associada com o algoritmo descrito pelas equações (4.7). Obtendo-se como resultado:

$$\dot{\underline{\theta}}(t) = \underline{R}^{-1} \underline{f}(\underline{\theta}) \quad (4.8)$$

$$\underline{R} = \underline{G}(\underline{\theta}) - \underline{R}$$

$$\text{onde } \underline{f}(\underline{\theta}) = E \left\{ \underline{\omega}_{(K, \underline{\theta})}^{*T} \epsilon_{(K, \underline{\theta})}^* \right\} \text{ e}$$

$$\underline{G}(\underline{\theta}) = E \left[\underline{\omega}_{(K, \underline{\theta})}^{*T} \underline{\omega}_{(K, \underline{\theta})}^* \right]$$

A obtenção das equações de Ljung é trabalhosa e demorada, sendo apresentada em [8]. Em particular demonstra-se que, sendo $\underline{\theta}^*$ um ponto singular do sistema de equações diferenciais (4.8), tal que $\underline{\theta}_{(K)} = \underline{\theta}^*$ é uma solução assintoticamente estável, então $\underline{\theta}_{(K)}$ gerado pelas equações (4.7) tende para $\underline{\theta}^*$ com probabilidade 1 quando K tende para infinito [8]. As trajetórias das equações (4.8) são as trajetórias assintóticas das equações (4.7) que diferem de um dado valor δ da trajetória correspondente, $\underline{\theta}(t)$. Este desvio torna-se arbitrariamente pequeno quando K tende para infinito.

A partir do sistema de equações (4.8) encontram-se os seus pontos singulares, bem como a natureza de cada um deles.

Os possíveis pontos de convergência da equação (4.8) pertencem ao conjunto D que satisfaz $D = \left\{ \underline{\theta}^* / \underline{f}(\underline{\theta}^*) = \underline{0} \right\}$.

Verifica-se facilmente que o valor verdadeiro $\underline{\theta}_v$ sempre pertence ao conjunto D . Por exemplo para o algoritmo média-móvel encontra-se:

$$\underline{f}(\underline{\theta}_v) = E \left[\begin{matrix} \omega^{*T} \\ \omega(K, \underline{\theta}_v) \end{matrix} \quad \begin{matrix} \epsilon^* \\ \epsilon(K, \underline{\theta}_v) \end{matrix} \right] \quad (4.9)$$

$$\underline{f}(\underline{\theta}_v) = E \left\{ \begin{bmatrix} -Y(K-1) \\ \vdots \\ -Y(K-n_a) \\ u(K-1) \\ \vdots \\ u(K-n_b) \\ \zeta(K-1, \underline{\theta}_v) \\ \vdots \\ \zeta(K-n_c, \underline{\theta}_v) \end{bmatrix} \zeta(K, \underline{\theta}_v) \right\} = \underline{0} \quad (4.10)$$

Entretanto o conjunto D poderá conter mais pontos singulares, além de $\underline{\theta}_v$, todos estes valores devem ser analisados.

Por se ter associado ao algoritmo (4.7) a equação diferencial ordinária (4.8), é natural assumir que as condições de estabilidade do sistema de equações diferenciais ordinárias (4.8), linearizados em torno de $\underline{\theta} = \underline{\theta}^*$ e $\underline{R} = \underline{G}(\underline{\theta}^*)$, são as condições necessárias para a convergência do algoritmo (4.7) para $\underline{\theta}^* \in D$ [8].

Linearizando-se a equação (4.8) em torno de $\underline{\theta} = \underline{\theta}^*$, obtém-se:

$$\Delta \underline{\underline{\theta}} = \underline{\underline{G}}_{(\underline{\underline{\theta}}^*)}^{-1} \left[\underline{\underline{f}}_{(\underline{\underline{\theta}}^*)} + \frac{d}{d_{\underline{\underline{\theta}}}} \underline{\underline{f}}_{(\underline{\underline{\theta}})} \Big|_{\underline{\underline{\theta}} = \underline{\underline{\theta}}^*} \right] \quad (4.11)$$

$$\Delta \underline{\underline{\theta}} = \underline{\underline{G}}_{(\underline{\underline{\theta}}^*)}^{-1} \underline{\underline{H}}_{(\underline{\underline{\theta}}^*)} \Delta \underline{\underline{\theta}} \quad (4.12)$$

e

$$\underline{\underline{H}}_{(\underline{\underline{\theta}}^*)} = \frac{d}{d_{\underline{\underline{\theta}}}} \underline{\underline{f}}_{(\underline{\underline{\theta}})} \Big|_{\underline{\underline{\theta}} = \underline{\underline{\theta}}^*} \quad (4.13)$$

Se a matriz $\underline{\underline{G}}_{(\underline{\underline{\theta}}^*)}$ for inversível e todos autovalores da matriz $\underline{\underline{G}}_{(\underline{\underline{\theta}}^*)}^{-1} \underline{\underline{H}}_{(\underline{\underline{\theta}}^*)}$ tiverem parte real negativa, conclui-se que $\underline{\underline{\theta}} = \underline{\underline{\theta}}^*$ é um ponto singular estável, portanto o algoritmo converge para o ponto $\underline{\underline{\theta}}^*$.

Em seguida, para cada algoritmo, determinar-se-ão os possíveis pontos de convergência, isto é todos os elementos do conjunto D.

O primeiro caso a ser tratado, na determinação dos pontos do conjunto D, será o método dos mínimos quadrados recursivo, que é um método equivalente ao método dos mínimos quadrados.

Para este caso é fácil mostrar que a solução $\underline{\underline{\theta}}_v$ é única.

Da equação (4.8) tem-se:

$$\underline{\underline{f}}_{(\underline{\underline{\theta}}^*)} = E \begin{bmatrix} \omega_{(K, \underline{\underline{\theta}}^*)}^{*T} & \epsilon_{(K, \underline{\underline{\theta}}^*)}^* \end{bmatrix} \quad (4.14)$$

Para os mínimos quadrados recursivo resulta:

$$\underline{\underline{f}}_{(\underline{\underline{\theta}}^*)} = E \begin{bmatrix} \omega_{(K, \underline{\underline{\theta}}^*)}^{*T} & \left\{ y_{(K)} - \omega_{(K)} \underline{\underline{\theta}}^* \right\} \end{bmatrix} \quad (4.15)$$

$$\underline{f}(\underline{\theta}^*) = E \left[\underline{\omega}_{(K, \underline{\theta}^*)}^{*T} \left\{ \underline{\omega}_{(K)}^* \underline{\theta}_v + \zeta(K) - \underline{\omega}_{(K)}^* \underline{\theta}^* \right\} \right] \quad (4.16)$$

Para $\underline{f}(\underline{\theta}^*) = \underline{0}$ obtém-se da equação (4.16) que:

$$E \left[\underline{\omega}_{(K, \underline{\theta}^*)}^{*T} \underline{\omega}_{(K, \underline{\theta}^*)}^* \right] \left[\underline{\theta}_v - \underline{\theta}^* \right] = \underline{0} \quad (4.17)$$

Como a matriz $E \left[\underline{\omega}_{(K, \underline{\theta}^*)}^{*T} \underline{\omega}_{(K, \underline{\theta}^*)}^* \right]$ é definida positiva [15], $\underline{\theta}^* = \underline{\theta}_v$ é a única solução da equação (4.17), e portanto o único elemento do conjunto D.

Para o algoritmo dos mínimos quadrados generalizado recursivo, onde o ruído é estimado supondo um modelo auto-regressivo, Söderström [13], [14] mostra que, para relação sinal-ruído** muito maior do que a unidade, a solução $\underline{\theta}^* = \underline{\theta}_v$ é única; caso contrário múltiplas soluções podem ser obtidas para $\underline{f}(\underline{\theta}^*) = \underline{0}$.

A associação da equação diferencial ordinária ao algoritmo estocástico, não é ainda suficiente para determinar se os outros dois algoritmos recursivos apresentam outros pontos singulares, além de $\underline{\theta}_v$.

4.4 - ESTUDO DA NATUREZA DOS PONTOS SINGULARES

Neste item serão feitas considerações sobre a natureza dos pontos singulares do sistema de equações diferenciais, com o intuito de se determinarem os pontos estáveis.

Para o método dos mínimos quadrados recursivo, a solução $\underline{\theta} = \underline{\theta}_v$ além de ser única, é sempre estável.

Como foi visto no item anterior, para que ocorra a estabilidade, é necessário que os autovalores da matriz $\underline{G}_{(\underline{\theta}^*)}^{-1} \underline{H}_{(\underline{\theta}^*)}$ tenham parte real negativa.

** Definido no capítulo 5

Portanto a partir da equação (4.8) obtem-se:

$$\underline{G}_{(\underline{\theta}_v)}^{-1} = \left[E \left\{ \underline{\omega}_{(K, \underline{\theta}_v)}^{*T} \underline{\omega}_{(K, \underline{\theta}_v)}^* \right\} \right]^{-1} \quad (4.18)$$

e

$$\underline{H}_{(\underline{\theta}_v)} = \frac{d}{d\underline{\theta}} E \left. \underline{\omega}_{(K, \underline{\theta})}^{*T} \epsilon_{(K, \underline{\theta})}^* \right|_{\underline{\theta} = \underline{\theta}_v} \quad (4.19)$$

logo

$$\underline{H}_{(\underline{\theta}_v)} = E \left\{ \left. \frac{d}{d\underline{\theta}} \underline{\omega}_{(K, \underline{\theta})}^{*T} \right|_{\underline{\theta} = \underline{\theta}_v} \epsilon_{(K, \underline{\theta}_v)}^* + \underline{\omega}_{(K, \underline{\theta}_v)}^{*T} \frac{d}{d\underline{\theta}} \epsilon_{(K, \underline{\theta})}^* \right|_{\underline{\theta} = \underline{\theta}_v} \right\} \quad (4.20)$$

$$\text{Como } \epsilon_{(K, \underline{\theta})}^* = Y_{(K)} - \underline{\theta} \underline{\omega}_{(K, \underline{\theta})}^T \quad (4.21)$$

$$\text{e } Y_{(K)} = \underline{\theta}_v \underline{\omega}_{(K)}^T + \zeta_{(K)} \quad (4.22)$$

$$\text{obtem-se } \epsilon_{(K, \underline{\theta}_v)}^* = \zeta_{(K)} \quad (4.23)$$

logo

$$E \left\{ \left. \frac{d}{d\underline{\theta}} \underline{\omega}_{(K, \underline{\theta})}^{*T} \right|_{\underline{\theta} = \underline{\theta}_v} \epsilon_{(K, \underline{\theta}_v)}^* \right\} = E \left\{ \zeta_{(K)} \right\} E \left\{ \left. \frac{d}{d\underline{\theta}} \underline{\omega}_{(K, \underline{\theta})}^{*T} \right|_{\underline{\theta} = \underline{\theta}_v} \right\} = 0 \quad (4.24)$$

pois $\{ \zeta_{(K)} \}$ é uma sequência de variáveis aleatórias não correlatas com média-zero.

Portanto a equação (4.20) pode ser reescrita como segue:

$$\underline{H}(\underline{\theta}_v) = \epsilon \left[\frac{\omega^{*T}(K, \underline{\theta}_v)}{d_{\underline{\theta}}} \frac{d}{d_{\underline{\theta}}} \epsilon^*(K, \underline{\theta}) \right]_{\underline{\theta} = \underline{\theta}_v} \quad (4.25)$$

$$\text{mas} \quad \frac{d}{d_{\underline{\theta}}} \epsilon^*(K, \underline{\theta}) \Big|_{\underline{\theta} = \underline{\theta}_v} = -\omega^*(K, \underline{\theta}_v) \quad (4.26)$$

portanto

$$\underline{G}_{(\underline{\theta}_v)}^{-1} \underline{H}(\underline{\theta}_v) = -\underline{I} \quad (4.27)$$

Então todos os autovalores da matriz $\underline{G}_{(\underline{\theta}_v)}^{-1} \underline{H}(\underline{\theta}_v)$ são iguais a -1 , e conseqüentemente $\underline{\theta}_v$ é uma solução estável da equação diferencial ordinária e o algoritmo dos mínimos quadrados recursivo converge para $\underline{\theta}_v$.

Entretanto, o segundo algoritmo, mínimos quadrados generalizado recursivo com ruído auto-regressivo, pode, sob determinadas condições, (relação sinal-ruído muito menor do que a unidade), apresentar outros pontos no conjunto D que resultam em soluções estáveis [13], [14], ou seja, o algoritmo pode convergir para valores diferentes, dependendo da inicialização.

Para os outros dois algoritmos pode se mostrar [8], como segue abaixo, que o valor $\underline{\theta}^* = \underline{\theta}_v$ pode produzir soluções instáveis.

Da equação (4.12) observa-se que, para que $\underline{\theta}^* = \underline{\theta}_v$ apresente soluções estáveis, é necessário que a matriz $\underline{G}_{(\underline{\theta}_v)}$ seja invertível e que todos os autovalores da matriz $\underline{G}_{(\underline{\theta}_v)}^{-1} \underline{H}(\underline{\theta}_v)$ tenham parte real negativa.

Para o algoritmo com modelamento do ruído por média-móvel, mostra-se a seguir que pelo menos uma das condições acima não é satisfeita.

Seja o processo:

$$A(Z^{-1}) Y_{(K)} = B(Z^{-1}) u_{(K)} + C(Z^{-1}) \zeta_{(K)}$$

onde A, B e C são polinômios de ordem n_a , n_b e n_c respectivamente e um modelo com a mesma estrutura do processo.

Das equações (4.9) e (4.13) obtém-se:

$$\underline{H}(\underline{\theta}_v) = \frac{d}{d\underline{\theta}} E \left\{ \underline{\omega}_{(K, \underline{\theta})}^{*T} \underline{\varepsilon}_{(K, \underline{\theta})}^* \right\} \Big|_{\underline{\theta} = \underline{\theta}_v} \quad (4.28)$$

$$\underline{H}(\underline{\theta}_v) = E \left\{ \frac{d}{d\underline{\theta}} \underline{\omega}_{(K, \underline{\theta})}^{*T} \Big|_{\underline{\theta} = \underline{\theta}_v} \underline{\varepsilon}_{(K, \underline{\theta}_v)}^* + \underline{\omega}_{(K, \underline{\theta}_v)}^{*T} \frac{d}{d\underline{\theta}} \underline{\varepsilon}_{(K, \underline{\theta})}^* \Big|_{\underline{\theta} = \underline{\theta}_v} \right\} \quad (4.29)$$

Para o algoritmo média-móvel, tem-se:

$$\underline{\varepsilon}_{(K, \underline{\theta})}^* = Y_{(K)} - \underline{\theta} \underline{\omega}_{(K, \underline{\theta})}^T \quad (4.30)$$

Portanto:

$$\underline{\varepsilon}_{(K, \underline{\theta}_v)}^* = \zeta_{(K)} \quad (4.31)$$

logo

$$E \left\{ \frac{d}{d\underline{\theta}} \underline{\omega}_{(K, \underline{\theta})}^{*T} \Big|_{\underline{\theta} = \underline{\theta}_v} \underline{\varepsilon}_{(K, \underline{\theta}_v)}^* \right\} = E \zeta_{(K)} E \frac{d}{d\underline{\theta}} \underline{\omega}_{(K, \underline{\theta})}^{*T} \Big|_{\underline{\theta} = \underline{\theta}_v} = 0 \quad (4.32)$$

pois $\{\zeta_{(K)}\}$ é uma sequência de variáveis aleatórias com média zero.

Portanto a equação (4.29) pode ser reescrita como segue:

$$\underline{H}(\underline{\theta}_V) = E \left[\underline{\omega}^{*T}(K, \underline{\theta}_V) \frac{d}{d\underline{\theta}} \varepsilon^*(K, \underline{\theta}) \Big|_{\underline{\theta} = \underline{\theta}_V} \right] \quad (4.33)$$

Para este algoritmo tem-se que:

$$\varepsilon(K) = \frac{1}{C(Z^{-1})} \left\{ A(Z^{-1}) Y(K) - B(Z^{-1}) u(K) \right\} \quad (4.34)$$

Portanto

$$\frac{d}{d\underline{\theta}} \varepsilon^*(K, \underline{\theta}) \Big|_{\underline{\theta} = \underline{\theta}_V} = \underline{\tilde{\omega}}^T(K) = - \frac{1}{C(Z^{-1})} \underline{\omega}^T(K) \quad (4.35)$$

Substituindo a equação (4.35) na equação (4.33), obtem-se:

$$\underline{H}(\underline{\theta}_V) = E \underline{\omega}^T(K) \underline{\tilde{\omega}}(K) \quad \text{onde} \quad \underline{\omega}^T(K) = \underline{\omega}^{*T}(K, \underline{\theta}_V) \quad (4.36)$$

Portanto o traço da matriz $\underline{H}(\underline{\theta}_V)$ pode ser escrito como:

$$\text{tr } \underline{H}(\underline{\theta}_V) = n_a \left[-E Y(K) \tilde{Y}(K) \right] + n_b \left[-E u(K) \tilde{u}(K) \right] + n_c \left[-E \zeta(K) \tilde{\zeta}(K) \right] \quad (4.37)$$

onde

$$C(Z^{-1}) \tilde{Y}(K) = Y(K), \quad C(Z^{-1}) \tilde{u}(K) = u(K) \quad \text{e} \quad C(Z^{-1}) \tilde{\zeta}(K) = \zeta(K)$$

Assumindo que a sequência de entrada $\left\{ u(K) \right\}$ tem ca-

racterísticas de ruído branco com variância unitária e que $E\{\zeta_{(K)}^2\} = 1$, então a covariância da saída é dada por: [15]

$$E\{Y_{(K)} \bar{Y}_{(K)}\} = \frac{1}{2\pi i} \int_{|Z|=1} \frac{1}{Z} \frac{1}{C(Z)} \left[\frac{B(Z) B(Z^{-1})}{A(Z) A(Z^{-1})} + \frac{C(Z) C(Z^{-1})}{A(Z) A(Z^{-1})} \right] dZ \quad (4.38)$$

$$\text{portanto } \text{tr } \underline{H}_{(\underline{\theta}_v)} = -(n_a + n_c) - n_a E Y_{(K)} \bar{Y}_{(K)} \quad (4.39)$$

Como $E Y_{(K)} \bar{Y}_{(K)}$ pode ser feita negativa, o traço da matriz $\underline{H}_{(\underline{\theta}_v)}$ poderá ser positivo e a matriz $\underline{H}_{(\underline{\theta}_v)}$ terá pelo menos um autovalor positivo ou com parte real positiva, portanto pode ser encontrado um exemplo onde a matriz $\underline{G}_{(\underline{\theta}_v)}^{-1} \underline{H}_{(\underline{\theta}_v)}$ terá autovalores instáveis.

Os autovalores da matriz $\underline{G}_{(\underline{\theta}_v)}^{-1} \underline{H}_{(\underline{\theta}_v)}$ podem então ser calculados através de processos numéricos.

Exemplo 4.1 : Seja um sistema descrito por:

$$A(Z^{-1}) Y_{(K)} = B(Z^{-1}) u_{(K)} + C(Z^{-1}) \zeta_{(K)} \quad (4.40)$$

onde

$$A(Z^{-1}) = 1 + 0,9Z^{-1} + 0,95Z^{-2}$$

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

$$C(Z^{-1}) = 1 + 1,5Z^{-1} + 0,75Z^{-2}$$

Utilizando-se métodos numéricos para calcular os autovalores da matriz $\underline{G}_{(\underline{\theta}_v)}^{-1} \underline{H}_{(\underline{\theta}_v)}$, obtem-se um autovalor complexo, com parte real positiva, cujo valor numérico é $(0,164 \pm j 1,379)$.

Portanto o valor verdadeiro $\underline{\theta}_v$ é uma solução instável.

A simulação deste sistema é apresentada em detalhe no capítulo 5.

4.5 - CONSTRUÇÃO DE UM ALGORÍTMO MÉDIA-MÓVEL QUE TENHA SOLUÇÃO ESTÁVEL PARA O VALOR VERDADEIRO

Neste item procurar-se-á mudar o algoritmo anterior de modo que o valor verdadeiro $\underline{\theta}_v$ resulte sempre em soluções estáveis [8] .

Seja o algoritmo

$$\hat{\underline{\theta}}_{(K)} = \hat{\underline{\theta}}_{(K-1)} + \underline{K}_{(K)} \varepsilon_{(K)} \quad (4.41.a)$$

$$\underline{K}_{(K+1)} = \frac{\underline{P}_{(K)} \underline{Z}_{(K+1)}}{1 + \underline{Z}_{(K+1)}^T \underline{P}_{(K)} \underline{Z}_{(K+1)}} \quad (4.41.b)$$

$$\underline{P}_{(K+1)} = \underline{P}_{(K)} - \frac{\underline{P}_{(K)} \underline{Z}_{(K+1)} \underline{Z}_{(K+1)}^T \underline{P}_{(K)}}{1 + \underline{Z}_{(K+1)}^T \underline{P}_{(K)} \underline{Z}_{(K+1)}} \quad (4.41.c)$$

$$\varepsilon_{(K)} = Y_{(K)} - \hat{\underline{\theta}}_{(K-1)}^T \underline{\omega}_{(K}, \hat{\underline{\theta}}_{(K-1)}) \quad (4.41.d)$$

onde $\underline{Z}_{(K)}$ é definido por:

$$\underline{C}_K - 1 \left(\underline{Z}^{-1} \right) \underline{Z}_{(K)} = \underline{\omega}_{(K)}$$

A equação diferencial ordinária associada às equações (4.40) será análoga à equação (4.8) [8] , onde

$$\underline{\dot{\theta}}(\underline{\theta}) = \underline{E} \underline{Z}_{(K}, \underline{\theta})^* \varepsilon_{(K}, \underline{\theta})^* \quad (4.42)$$

$$\underline{G}(\underline{\theta}) = \epsilon \underline{Z}_{(K, \underline{\theta})}^{*T} \underline{Z}_{(K, \underline{\theta})}^*$$

e $\underline{Z}_{(K, \underline{\theta})}^*$ denota, analogamente a $\underline{\omega}_{(K, \underline{\theta})}^*$ a função obtida se $\underline{Z}_{(., .)}$ for calculado utilizando um parâmetro fixo $\underline{\theta}_{(K)}$.

Da equação (4.41) resulta que:

$$\underline{f}(\underline{\theta}_V) = \epsilon \underline{Z}_{(K, \underline{\theta}_V)}^{*T} \epsilon_{(K, \underline{\theta}_V)}^* = \epsilon \underline{Z}_{(K, \underline{\theta}_V)}^{*T} \zeta_{(K)} \quad (4.43)$$

$$\underline{f}(\underline{\theta}_V) = \epsilon \underline{Z}_{(K, \underline{\theta}_V)}^{*T} \epsilon \zeta_{(K)} = \underline{0} \quad (4.44)$$

logo, como na seção anterior, obtém-se:

$$\underline{H}(\underline{\theta}_V) = \frac{d}{d\underline{\theta}} \left[\epsilon \underline{Z}_{(K, \underline{\theta})}^{*T} \epsilon_{(K, \underline{\theta})}^* \right] \bigg|_{\underline{\theta} = \underline{\theta}_V} \quad (4.45)$$

$$\underline{H}(\underline{\theta}_V) = \epsilon \underline{Z}_{(K, \underline{\theta}_V)}^{*T} \underline{\tilde{\omega}}_{(K)} \quad (4.46)$$

onde a partir da equação (4.41.) tem-se que:

$$\underline{Z}_{(K, \underline{\theta}_V)}^* = -\underline{\tilde{\omega}}_{(K)} \quad (4.47)$$

Portanto

$$\underline{G}_{(\underline{\theta}_V)}^{-1} \underline{H}(\underline{\theta}_V) = \left[\epsilon \underline{Z}_{(K, \underline{\theta}_V)}^{*T} \underline{Z}_{(K, \underline{\theta}_V)}^{*T} \right]^{-1} \cdot \left[-\epsilon \underline{Z}_{(K, \underline{\theta}_V)}^{*T} \underline{Z}_{(K, \underline{\theta}_V)}^* \right] \quad (4.48)$$

$$\underline{G}_{(\underline{\theta}_V)}^{-1} \underline{H}(\underline{\theta}_V) = -\underline{I}$$

Logo o valor verdadeiro θ_v é sempre um ponto de convergência.

4.6 - RESUMO

Neste capítulo foi apresentado um método que torna possível o estudo de convergência dos algoritmos estocásticos. Este método associa a cada algoritmo um sistema de equações diferenciais ordinárias e, a partir delas, estuda a sua convergência.

Utilizando este procedimento, mostrou-se que o algoritmo dos mínimos quadrados recursivo tem sempre uma única solução estável, enquanto que o dos mínimos quadrados generalizado recursivo com o ruído modelado por um processo auto-regressivo pode apresentar mais de uma solução quando a relação sinal-ruído não é muito maior do que a unidade.

Além disso, foi introduzido neste capítulo uma alteração do algoritmo média-móvel, para que o mesmo convirja para o valor verdadeiro.

CAPÍTULO 5

RESULTADOS E COMPARAÇÕES DOS MÉTODOS APRESENTADOS

5.1 - INTRODUÇÃO

Os métodos apresentados anteriormente foram simulados exaustivamente e neste capítulo os principais resultados são apresentados comparativamente.

Estuda-se a escolha do fator de ponderação, bem como é feita a análise de um exemplo em que um parâmetro de processo é variável no tempo.

Nos outros itens são apresentadas simulações com os tres tipos de modelamento do ruído, deduzidos no capítulo 3.

Por último é exemplificada a determinação da ordem do modelo.

5.2 - SISTEMAS SIMULADOS

Os sistemas simulados são descritos pela equação:

$$A(Z^{-1}) Y_{(K)} = B(Z^{-1}) u_{(K)} + C(Z^{-1}) \zeta_{(K)} \quad (5.1)$$

onde

$$A(Z^{-1}) = 1 + a_1 Z^{-1} + \dots + a_n Z^{-n}$$

$$B(Z^{-1}) = b_1 Z^{-1} + \dots + b_n Z^{-n}$$

$$C(Z^{-1}) = 1 + c_1 Z^{-1} + \dots + c_n Z^{-n}$$

Estes processos são modelados por:

$$A_M(Z^{-1}) Y_{(K)} = B_M(Z^{-1}) u_{(K)} + e_{(K)} \quad (5.2)$$

$$e_{(K)} \equiv H(Z^{-1}) \zeta_{(K)}$$

$$e \quad H(Z^{-1}) = \begin{cases} C_M(Z^{-1}) & \text{algoritmo média-móvel} \\ 1/C_M(Z^{-1}) & \text{algoritmo auto-regressivo} \\ C_M(Z^{-1})/D_M(Z^{-1}) & \text{algoritmo média-móvel e auto-regressivo} \end{cases} \quad (5.3)$$

Neste capítulo são apresentados resultados para os seguintes processos:

1) P₁ - Sistema de primeira ordem

$$\begin{aligned} A(Z^{-1}) &= 1 - 0,8Z^{-1} \\ B(Z^{-1}) &= Z^{-1} \\ C(Z^{-1}) &= 1 + 0,7Z^{-1} \end{aligned} \quad (5.4)$$

2) P₂ - Sistema oscilante de segunda ordem

$$\begin{aligned} A(Z^{-1}) &= 1 - 1,5Z^{-1} + 0,7Z^{-2} \\ B(Z^{-1}) &= Z^{-1} + 0,5Z^{-2} \\ C(Z^{-1}) &= 1 - Z^{-1} + 0,2Z^{-2} \end{aligned} \quad (5.5)$$

3) P₃ - Sistema de segunda ordem de fase não mínima

$$\begin{aligned} A(Z^{-1}) &= 1 - 1,425Z^{-1} + 0,496Z^{-2} \\ B(Z^{-1}) &= -0,102Z^{-1} + 0,173Z^{-2} \\ C(Z^{-1}) &= 1 - Z^{-1} + 0,2Z^{-2} \end{aligned} \quad (5.6)$$

4) P₄ - Sistema de segunda ordem

$$\begin{aligned} A(Z^{-1}) &= 1 + 0,9Z^{-1} + 0,95Z^{-2} \\ B(Z^{-1}) &= Z^{-1} \\ C(Z^{-1}) &= 1 + 1,5Z^{-1} + 0,75Z^{-2} \end{aligned} \quad (5.7)$$

O sinal de entrada foi escolhido como sendo uma sequência binária pseudo-aleatória, porque este sinal, excita todas as frequências do sistema [2], [13].

O ruído $\zeta(K)$ foi obtido a partir de um gerador de números aleatórios, produzindo uma sequência uniformemente distribuída entre -1 e $+1$.

Nos sistemas simulados, a amplitude do sinal de entrada foi escolhida conforme a relação sinal-ruído desejada.

A relação sinal-ruído $\frac{S}{N}$ é definida por:

$$\frac{S}{N} = \frac{\sigma_x^2}{\sigma_n^2} \quad \text{sendo} \quad (5.8)$$

$$\sigma_x^2 = \frac{\sigma_u^2}{2\pi i} \int_{|Z|=1} H_1(Z) H_1(Z^{-1}) \frac{dZ}{Z} \quad (5.9)$$

onde σ_u^2 é a potência do sinal de entrada e

$$H_1(Z) = \frac{B(Z)}{A(Z)} \quad \text{com} \quad \begin{aligned} A(Z) &= a_1 Z^n + a_2 Z^{n-1} + \dots + a_{n+1} \\ B(Z) &= b_0 Z^n + b_1 Z^{n-1} + \dots + b_n \end{aligned} \quad (5.10)$$

$$e \quad \sigma_n^2 = \frac{\sigma_\zeta^2}{2\pi i} \int_{|Z|=1} H_2(Z) H_2(Z^{-1}) \frac{dZ}{Z} \quad (5.11)$$

onde σ_ζ^2 é a potência do ruído e

$$H_2(Z) = \frac{C(Z)}{A(Z)} \quad \text{com} \quad C(Z) = c_0 Z^n + \dots + c_n \quad (5.12)$$

As integrais descritas pelas equações (5.9) e (5.11) são facilmente calculadas numericamente [16]. Portanto, conhecendo os valores das integrais e a potência do sinal de entrada e do ruído, pode-se calcular a relação sinal-ruído definida pela equação (5.8).

5.3 - ESCOLHA DO FATOR DE PONDERAÇÃO

A influência do fator de ponderação no comportamento do estimador é analisada no exemplo a seguir

Exemplo 5.1 : Seja o processo descrito por:

$$y(k) = -0,5u(k-1) + \zeta(k)$$

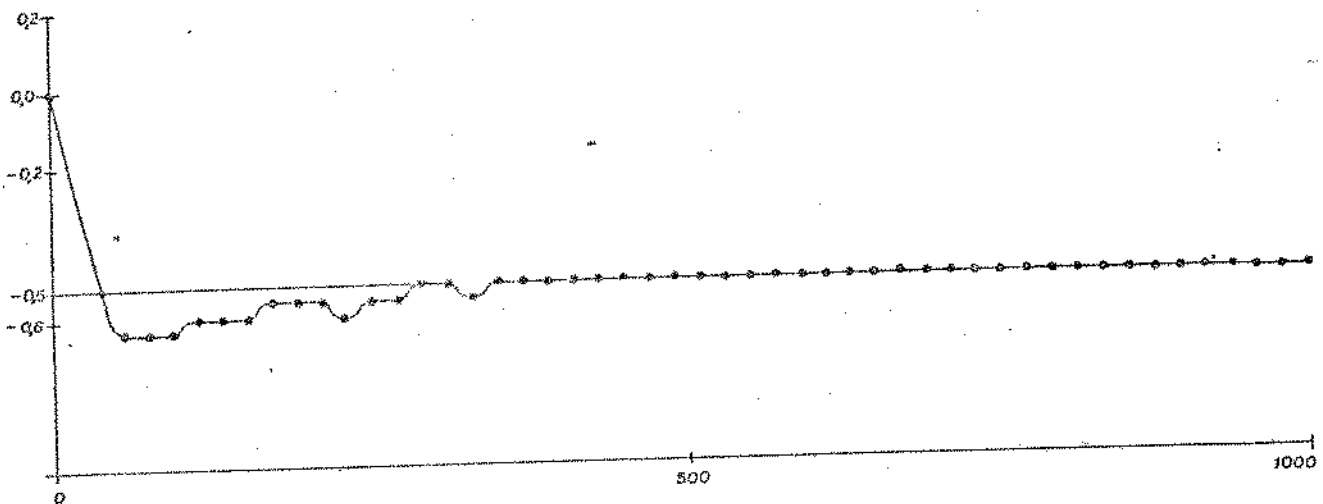


FIGURA 5.1.a

$$\rho = 1 \quad \hat{b}_{(0)} = 0 \quad P_{(0)} = 10$$

$$\text{VAR}(\hat{b})_{1000 \text{ MEDIDAS}} = 6,50 \times 10^{-4}$$

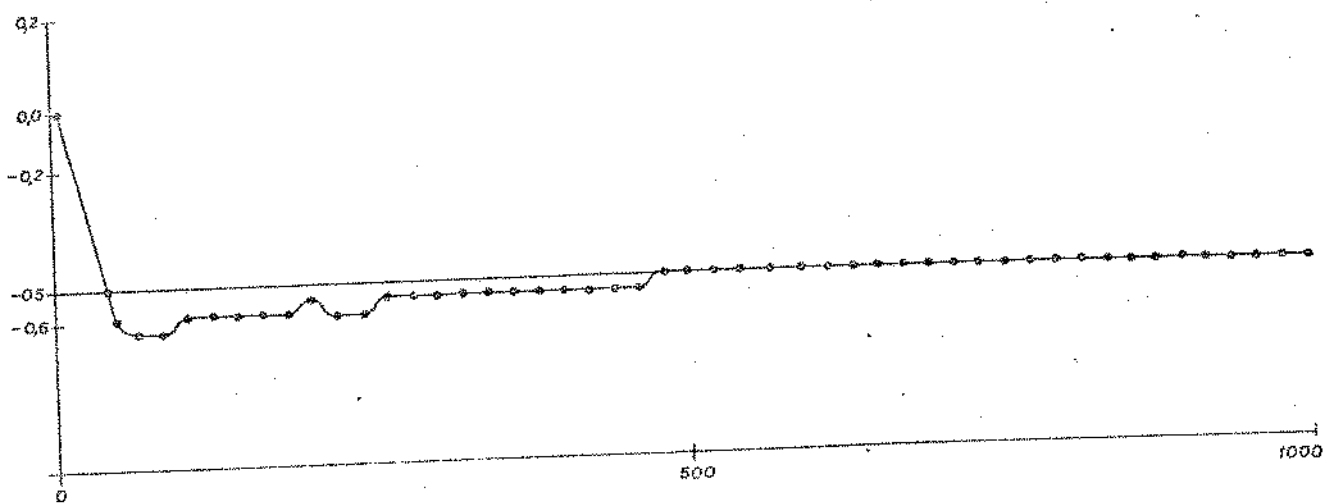


FIGURA 5.1.b

$$\rho = 0,99 \quad \hat{b}_{(0)} = 0 \quad P_{(0)} = 10$$

$$\text{VAR}(\hat{b})_{1000 \text{ MEDIDAS}} = 6,10 \times 10^{-3}$$

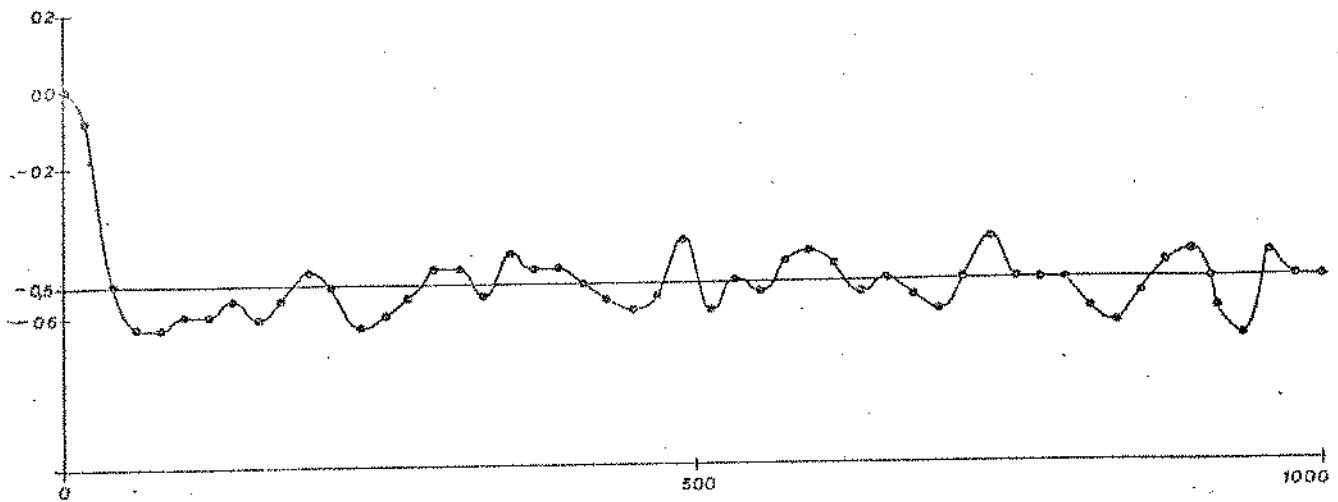


FIGURA 5.1.c

$$\rho = 0,95 \quad \hat{b}_{(0)} = 0 \quad P_{(0)} = 10$$

$$\text{VAR}(\hat{b}) = 3,42 \times 10^{-2}$$

1000 MEDIDAS

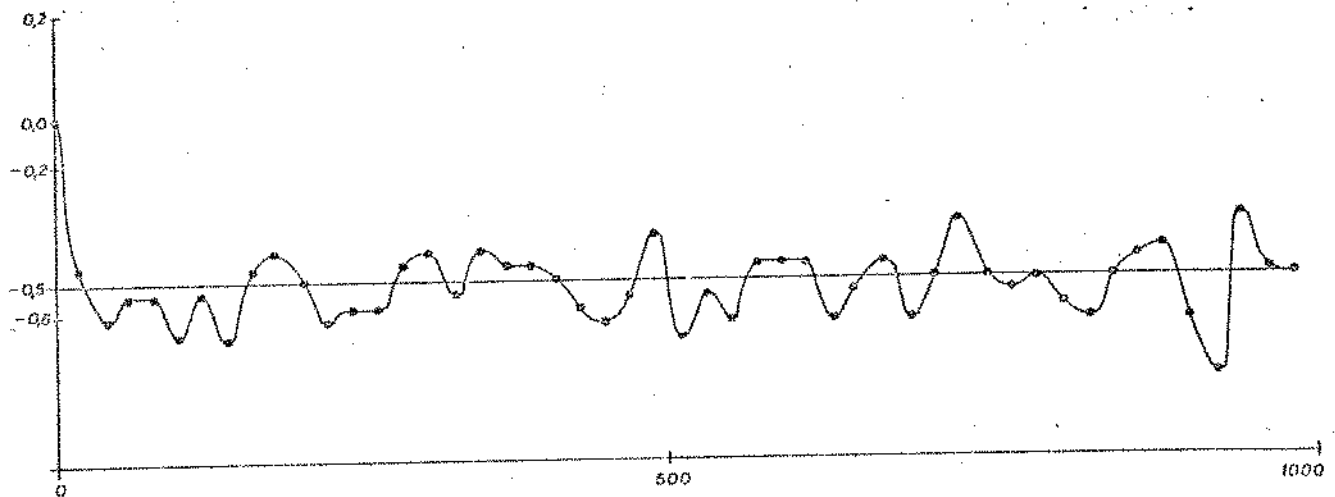


FIGURA 5.1.d

$$\rho = 0,9 \quad \hat{b}_{(0)} = 0 \quad P_{(0)} = 10$$

$$\text{VAR}(\hat{b}) = 3,50 \times 10^{-2}$$

1000 MEDIDAS

Nas figuras 5.1.a a 5.1.d observa-se que o estimador converge mais rapidamente, à medida que o fator de ponderação é diminuído. Concomitantemente a variância do estimador é aumentada como visto pela equação (3.49).

Para melhorar a precisão do estimador e não diminuir a taxa de convergência, sugere-se a utilização de um fator de ponderação variável no tempo, isto é;

$$\rho(K) = \rho_0 \rho(K-1) + (1 - \rho_0)$$

onde $\rho(K)$ é o fator de ponderação na K-ésima medida e $\rho_0 = .9$ e $\rho(0) = .9$

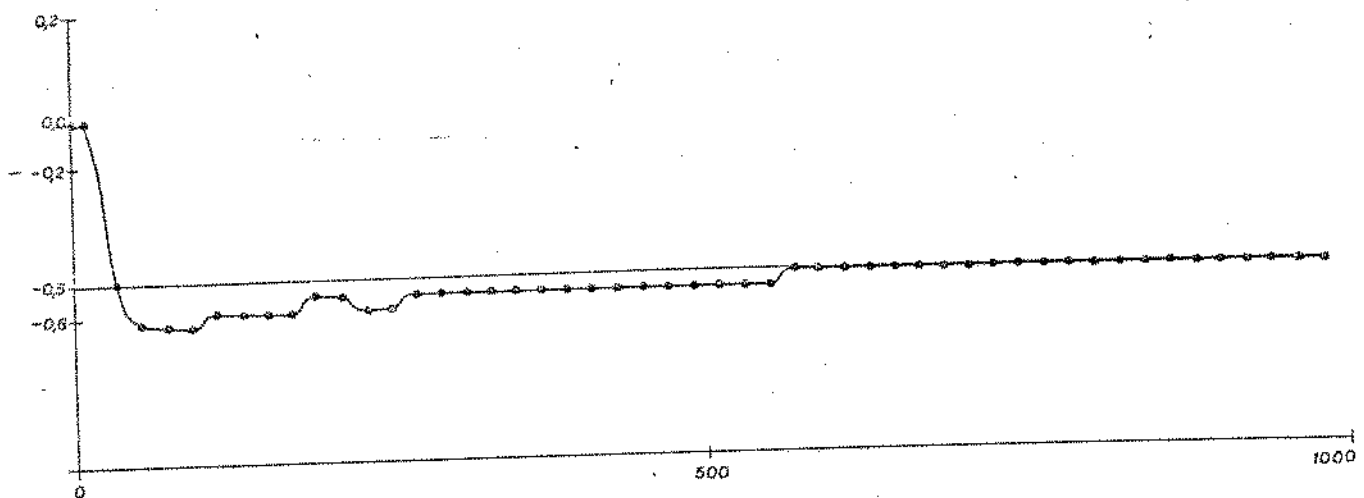


FIGURA 5.2

$$\begin{aligned} \hat{b}(0) &= 0 & P(0) &= 10 \\ \text{VAR}(\hat{b}) \\ 1000 \text{ MEDIDAS} &= 6,03 \times 10^{-4} \end{aligned}$$

Os resultados obtidos utilizando este fator de ponderação são apresentados na figura 5.2, verifica-se que a variância do estimador é diminuída, isto é, a precisão do valor estimado aumenta.

Outro método que permite aumentar a precisão do estima-

dor consiste em utilizar o fator de ponderação menor do que a unidade somente nas primeiras medidas (por exemplo 200-400). O resultado obtido com esta técnica é apresentado na figura 5.3.a-5.3.c

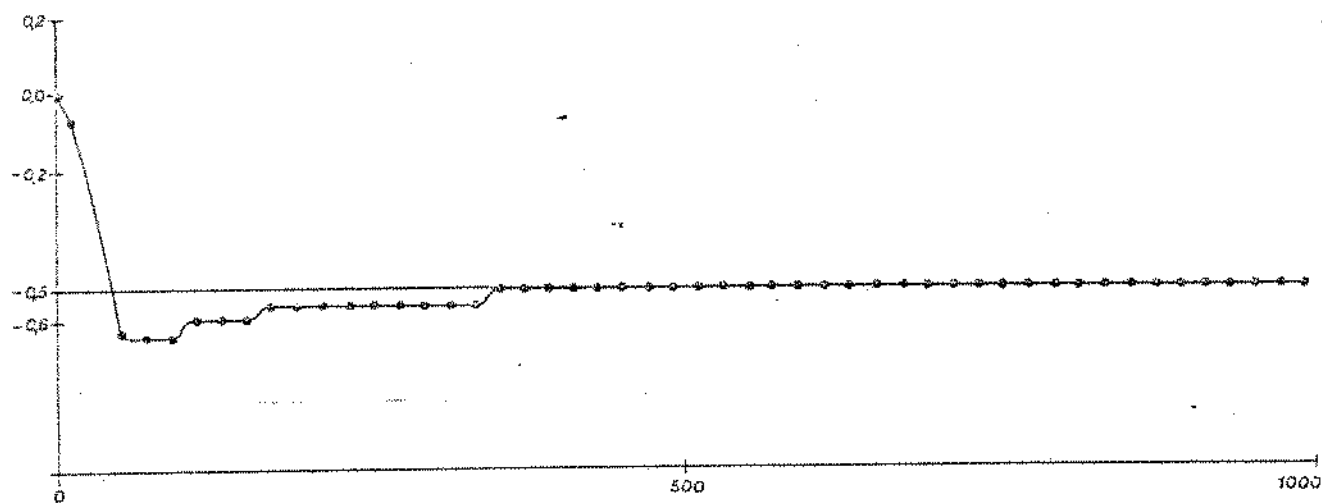


FIGURA 5.3.a

$$\hat{b}_{(0)} = 0 \quad P_{(0)} = 10$$

$$\rho = 0,99 \quad \text{atê} \quad 200 \text{ MEDIDAS}$$

$$\text{VAR } \hat{b}_{(b)} = 6,23 \times 10^{-4}$$

1000 MEDIDAS

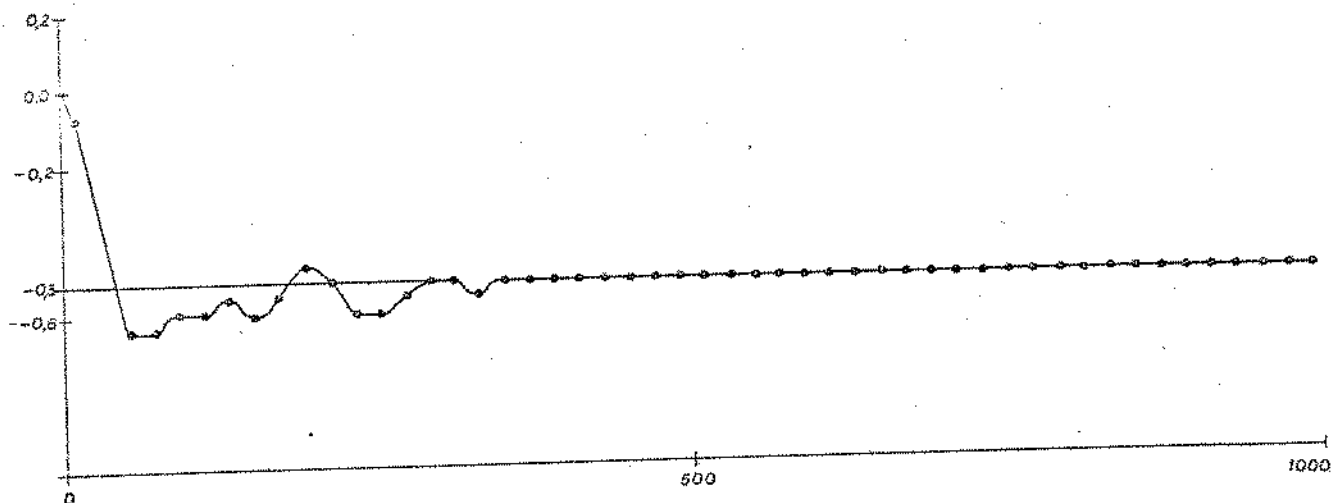


FIGURA 5.3.b

$$\begin{aligned} \hat{b}_{(0)} &= 0 & P_{(0)} &= 10 \\ \rho &= 0,95 \text{ até } 200 \text{ MEDIDAS} \\ \text{VAR}(\hat{b}) &= 6,70 \times 10^{-4} \\ 1000 \text{ MEDIDAS} \end{aligned}$$

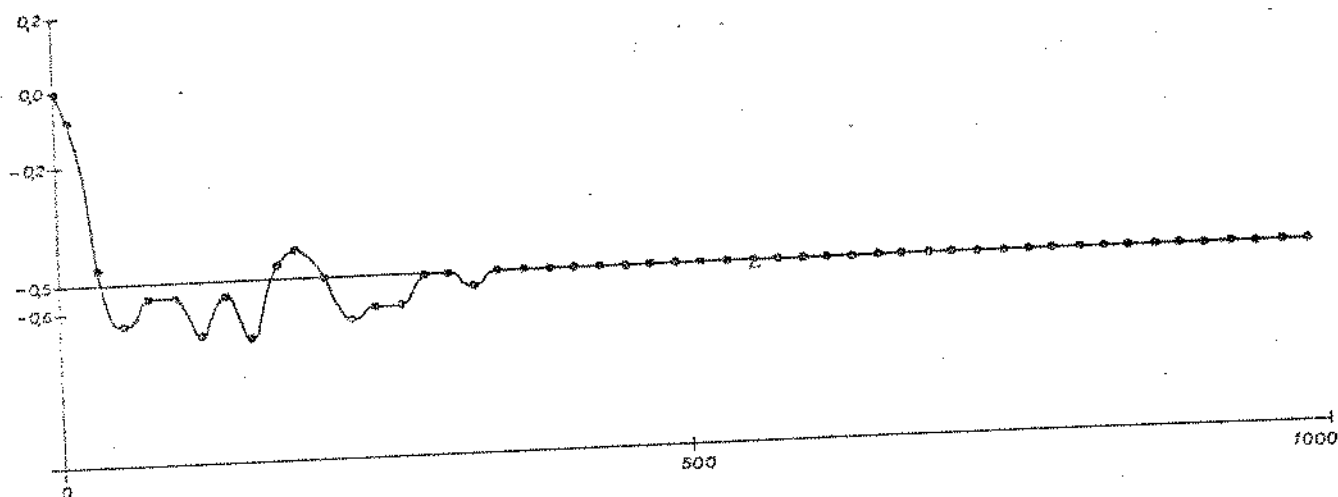


FIGURA 5.3.c

$$\begin{aligned} \hat{b}_{(0)} &= 0 & P_{(0)} &= 10 \\ \rho &= 0,90 \text{ até } 200 \text{ MEDIDAS} \\ \text{VAR}(\hat{b}) &= 6,90 \times 10^{-4} \\ 1000 \text{ MEDIDAS} \end{aligned}$$

Conclusões

- As figuras 5.1 a 5.3 mostram que a escolha do fator de ponderação deve ser um compromisso entre a taxa de convergência e a "precisão" desejada.

- Na prática também deve ser levado em conta que as características do processo ou das perturbações podem variar com o tempo. Utilizando o fator de ponderação menor do que a unidade, é possível que os algoritmos acompanhem esta variação nos parâmetros.

- A experiência com muitos exemplos mostrou que o valor do fator de ponderação entre .90 e .95 até aproximadamente 300 medidas e depois igual a 1 oferece os melhores desempenhos de transitório e precisão final dos valores estimados.

5.4 - RESULTADOS DO MÉTODO DOS MÍNIMOS QUADRADOS RECURSIVO

Utilizou-se o algoritmo dos mínimos quadrados recursivo na estimação dos processos P_1 , P_2 e P_3 inicialmente com ruído branco, isto é, $C(Z^{-1}) = 1$ e a seguir com ruído "colorido", ou seja, com polinômio $C(Z^{-1})$ consoante indicado em 5.4, 5.5, 5.6.

Na figura 5.4.a é apresentada a simulação do processo P_1 com ruído branco e na figura 5.4.b com ruído "colorido".

Nas figuras 5.5.a e 5.5.b são apresentados ambos os resultados para o processo P_2 , enquanto nas figuras 5.6.a e 5.6.b se considera o processo P_3 .

Conclusões

Baseado nas figuras 5.4 a 5.6 pode-se verificar que:

1) Para o caso de ruído "branco" os estimadores serão não polarizados, como deduzido no capítulo 2;

2) Os estimadores serão polarizados quando o ruído é "colorido", porque neste caso o modelo e o sistema tem estruturas diferentes;

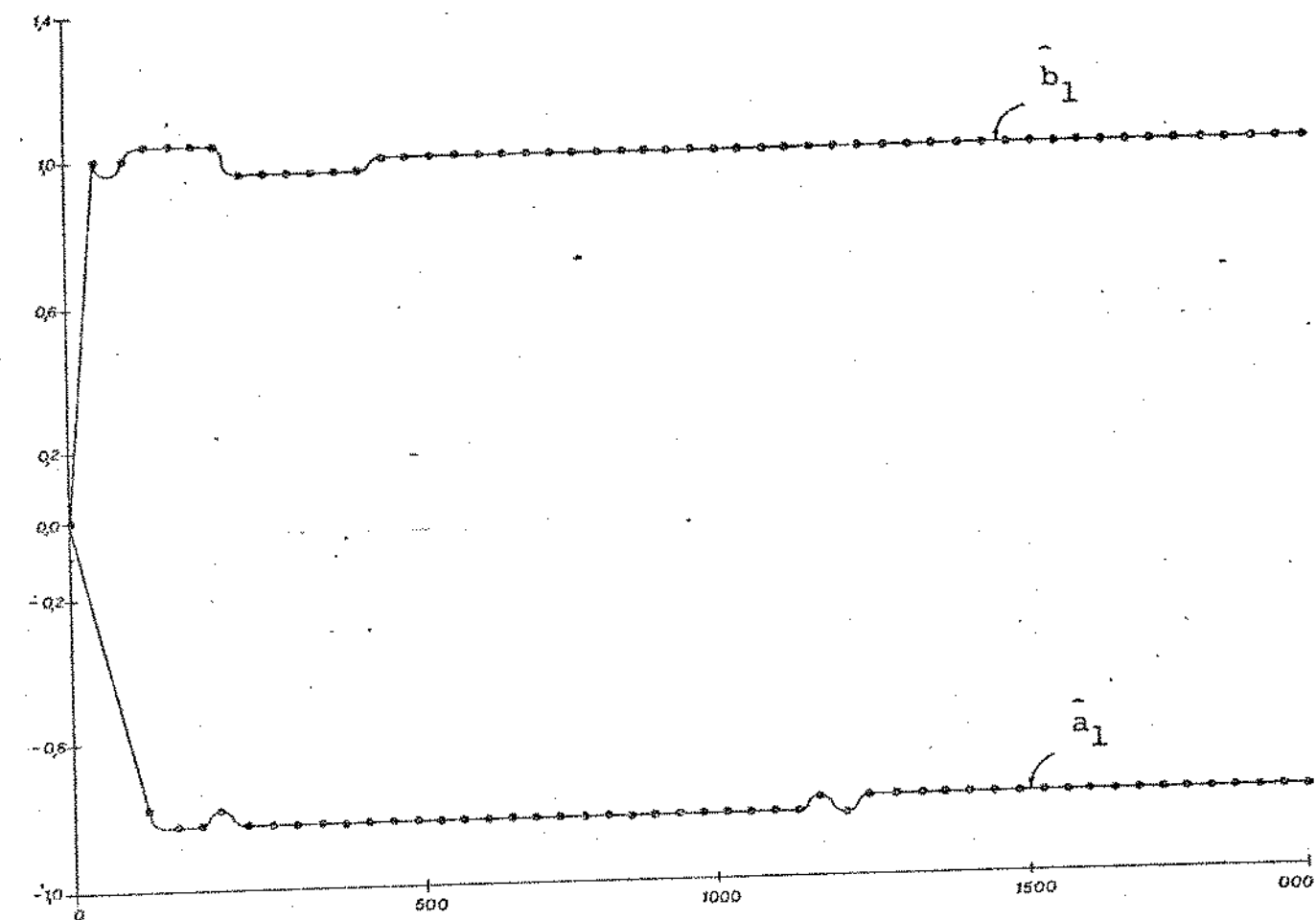


FIGURA 5.4.a

$$\bar{\theta}(0) = 0 \quad \underline{P}(0) = 10 \underline{I}$$

$$\frac{\underline{S}}{N} = 1$$

PROCESSO P_1

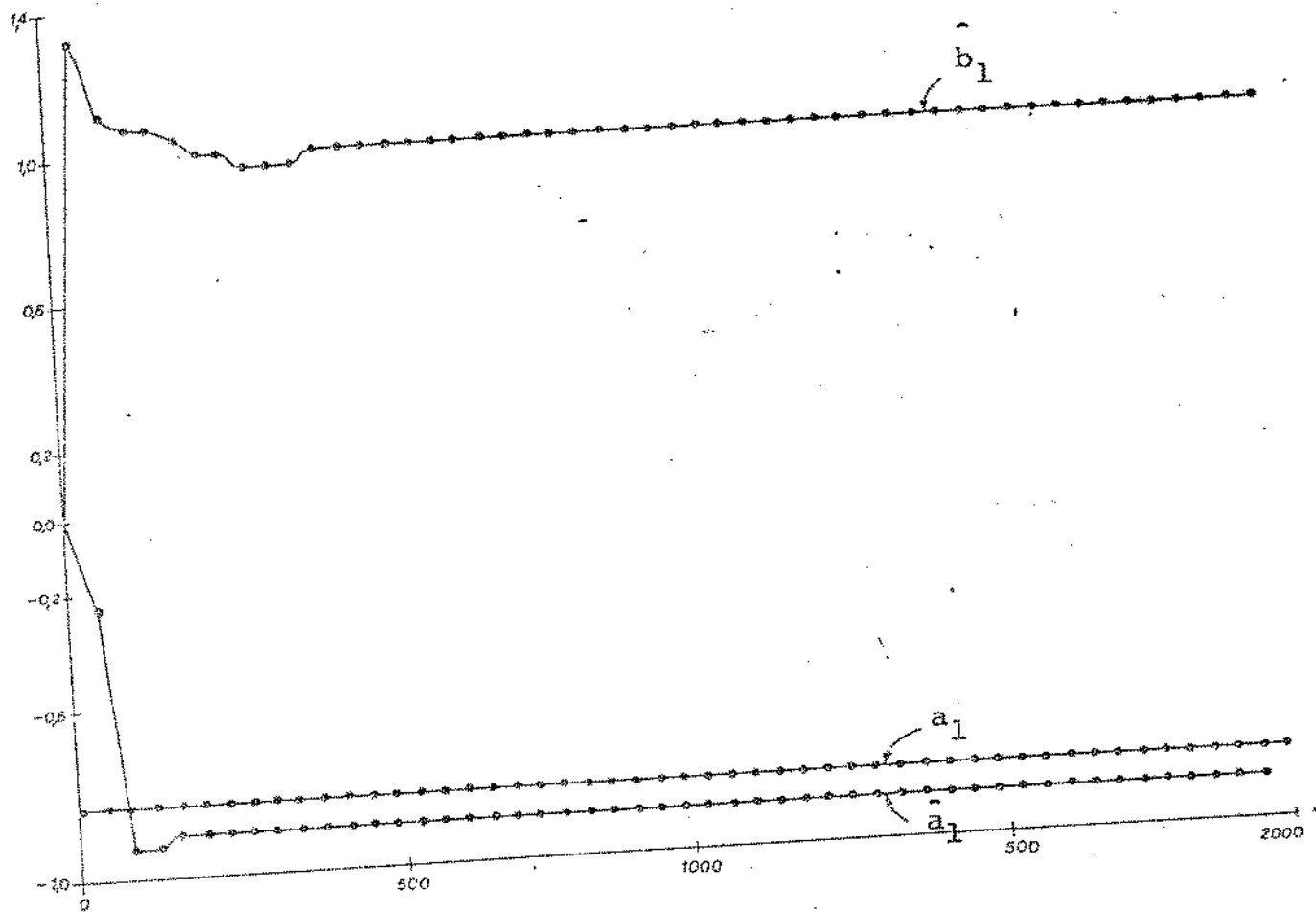


FIGURA 5.4.b

$$\hat{\theta}(0) = 0 \quad \underline{P}(0) = 10 \underline{I}$$

$$\frac{S}{N} = 1$$

PROCESSO P_1

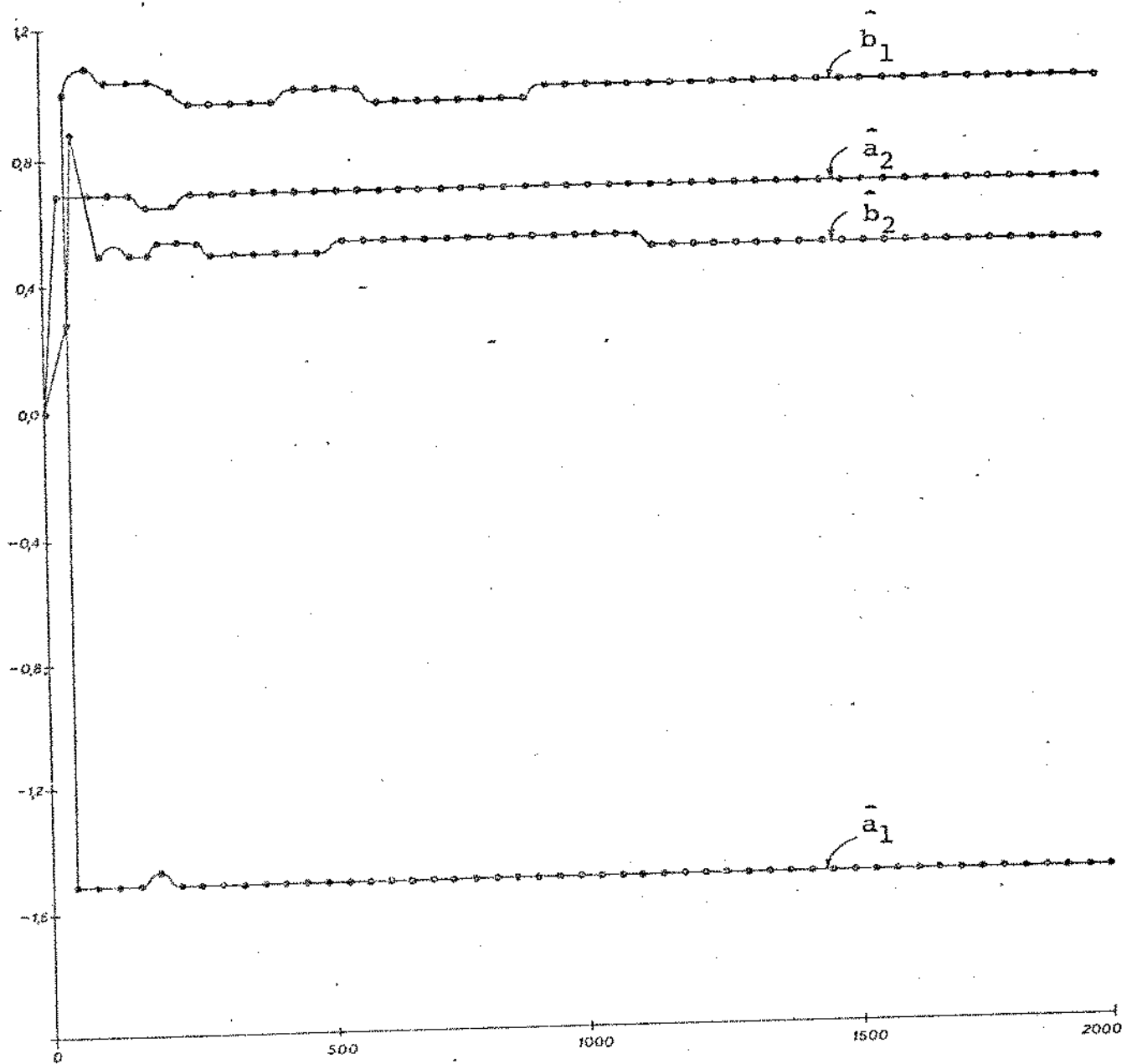


FIGURA 5.5.a

$$\hat{\theta}_{(0)} = 0 \quad \underline{p}_{(0)} = 10 \underline{I}$$

$$\frac{S}{N} = 1$$

PROCESSO P_2

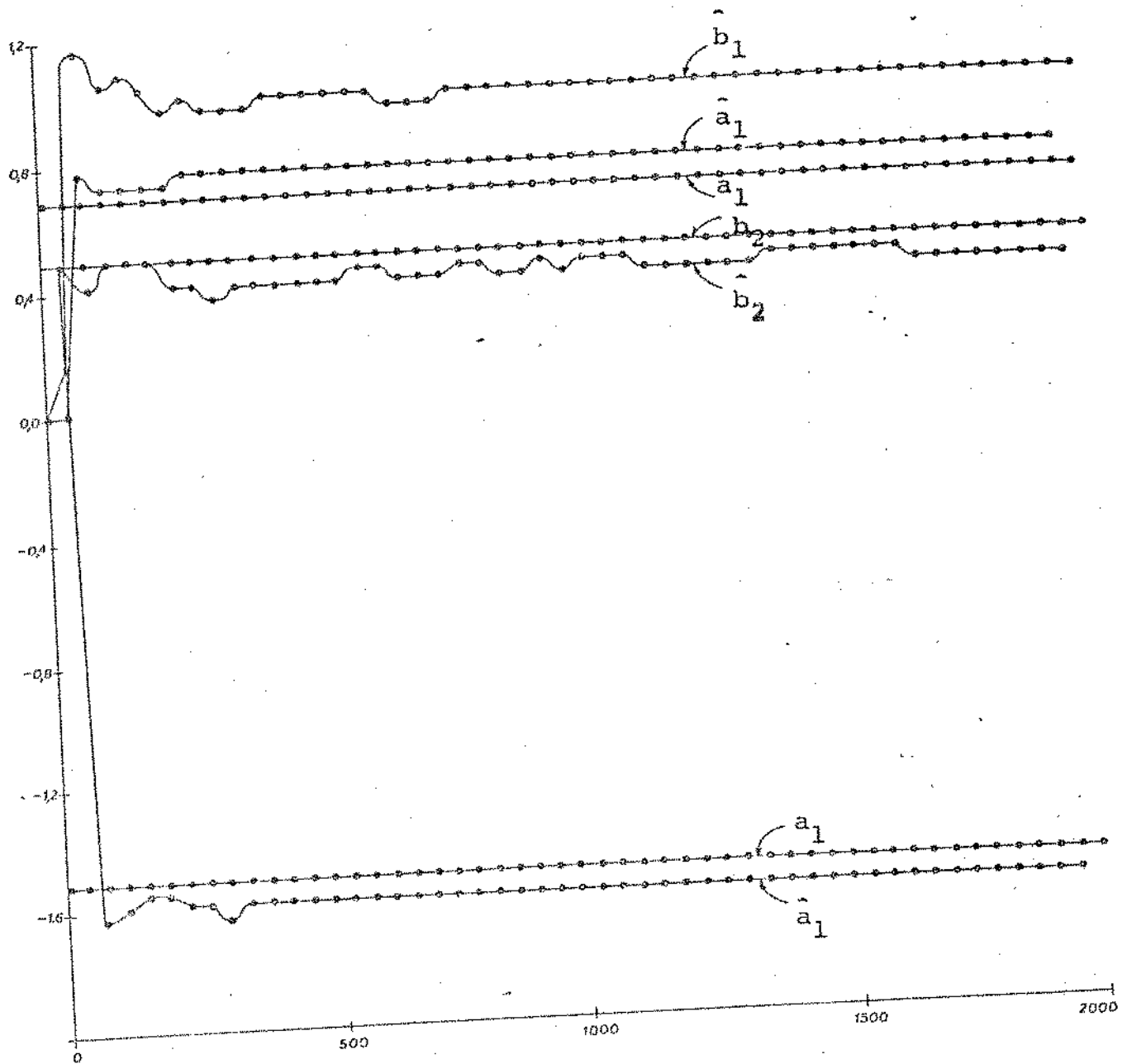


FIGURA 5.5.b

$$\hat{\theta}_{(0)} = \underline{0} \quad \underline{P}_{(0)} = 10 \underline{I}$$

$$\underline{S} = 1$$

PROCESSO $\underline{P}_2$

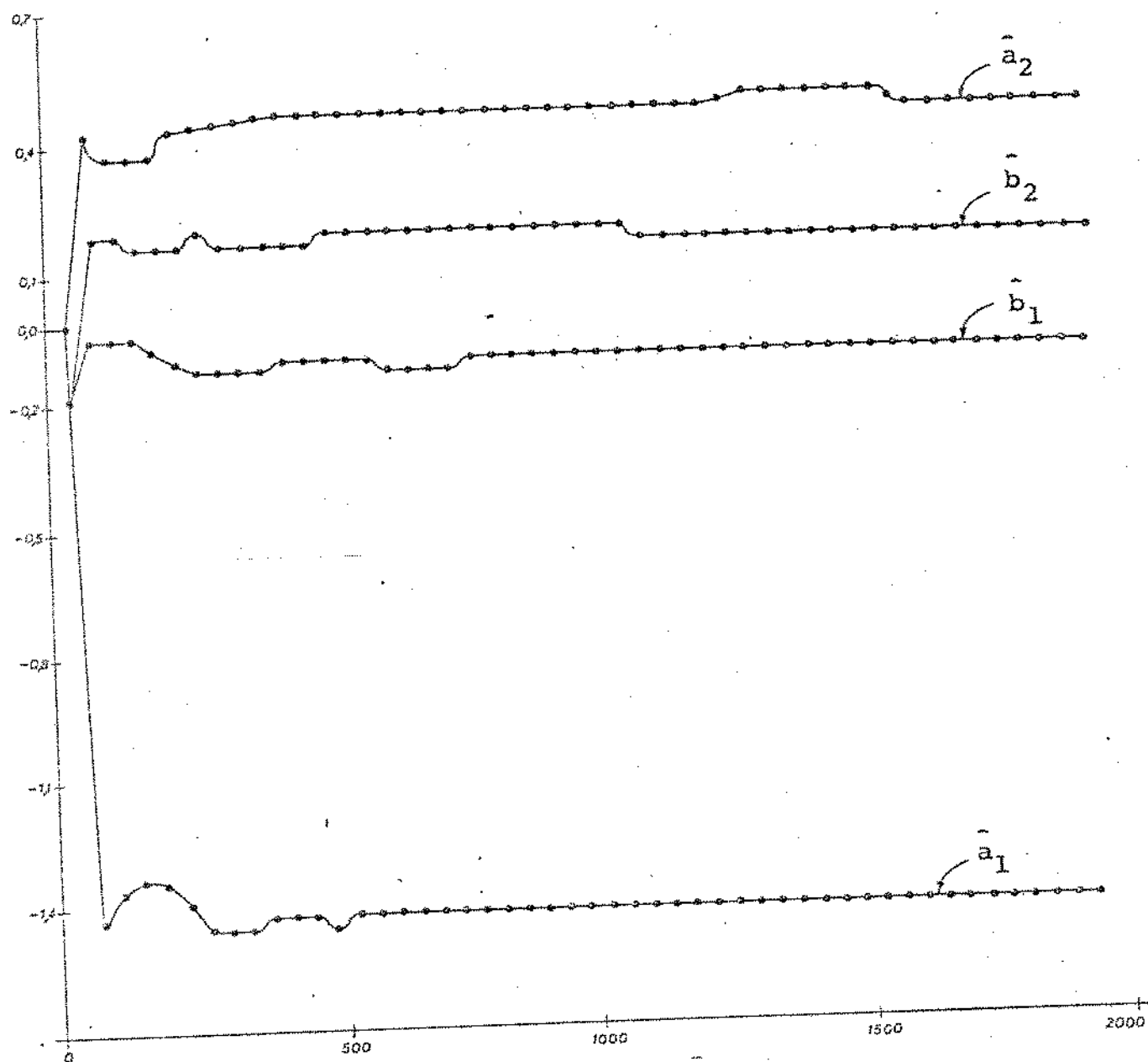


FIGURA 5.6.a

$$\hat{\theta}_{(0)} = \underline{0}$$

$$\underline{P}_{(0)} = 10 \underline{I}$$

$$\frac{\underline{S}}{N} = 1$$

PROCESSO $\underline{P}_3$

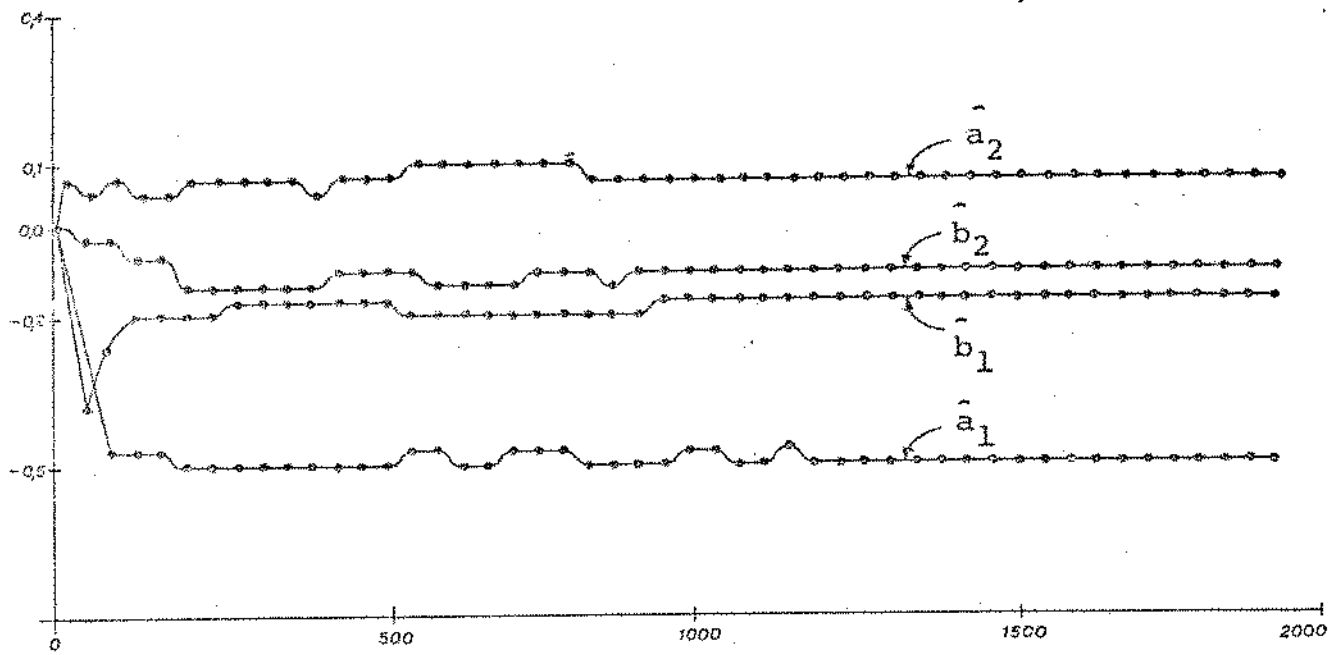


FIGURA 5.6.b

$$\hat{\theta}_{(0)} = \underline{0} \quad \underline{P}_{(0)} = 10 \underline{I}$$

$$\frac{S}{N} = 1$$

PROCESSO $\underline{P}_3$

Através de simulações também verificou-se que aumentando a relação sinal-ruído pode-se reduzir o desvio entre o valor verdadeiro e o valor estimado, mas mesmo assim o estimador será polarizado.

5.5 - MÉTODO DOS MÍNIMOS QUADRADOS GENERALIZADO COM O RUÍDO MODELADO POR AUTO-REGRESSÃO

Este método foi simulado para os processos P_1 , P_2 , P_3 . Os resultados são apresentados nas figuras 5.7 a 5.11.

Principais Conclusões

- Para o sistema de primeira ordem, figura 5.7, observa-se que os estimadores tendem para os valores corretos após poucas medidas (aproximadamente $80 \div 100$), enquanto que para os sistemas de segunda ordem, figuras 5.8 e 5.9, a convergência é mais lenta;

- É possível aumentar a taxa de convergência deste algoritmo, inicializando os parâmetros do ruído corretamente, quando os mesmos são conhecidos. A simulação desta situação é apresentada na figura 5.10;

- De acordo com o capítulo 4 este método apresenta uma única solução quando a relação sinal-ruído é muito maior do que a unidade, caso contrário existirão múltiplas soluções. Nas figuras 5.11.a e 5.11.b apresenta-se um exemplo de primeira ordem (processo P_1) que ilustra o exposto acima. Verificou-se que, dependendo da inicialização, obtém-se soluções diferentes, ou seja, o valor verdadeiro não é a única solução possível.

5.6 - SIMULAÇÃO DO ALGORÍTMO DOS MÍNIMOS QUADRADOS GENERALIZADO COM O RUÍDO MODELADO POR AUTO-REGRESSÃO E MÉDIA-MÓVEL

Seja o processo P_2 onde $C(z^{-1}) = 1 + 0,3z^{-1}$ e $D(z^{-1}) = 1 - 0,5z^{-1}$.

Caso 1 :-

Na figura 5.12 apresenta-se o resultado da simulação do processo descrito acima, com o ruído modelado por auto-regressão e média-móvel, isto é:

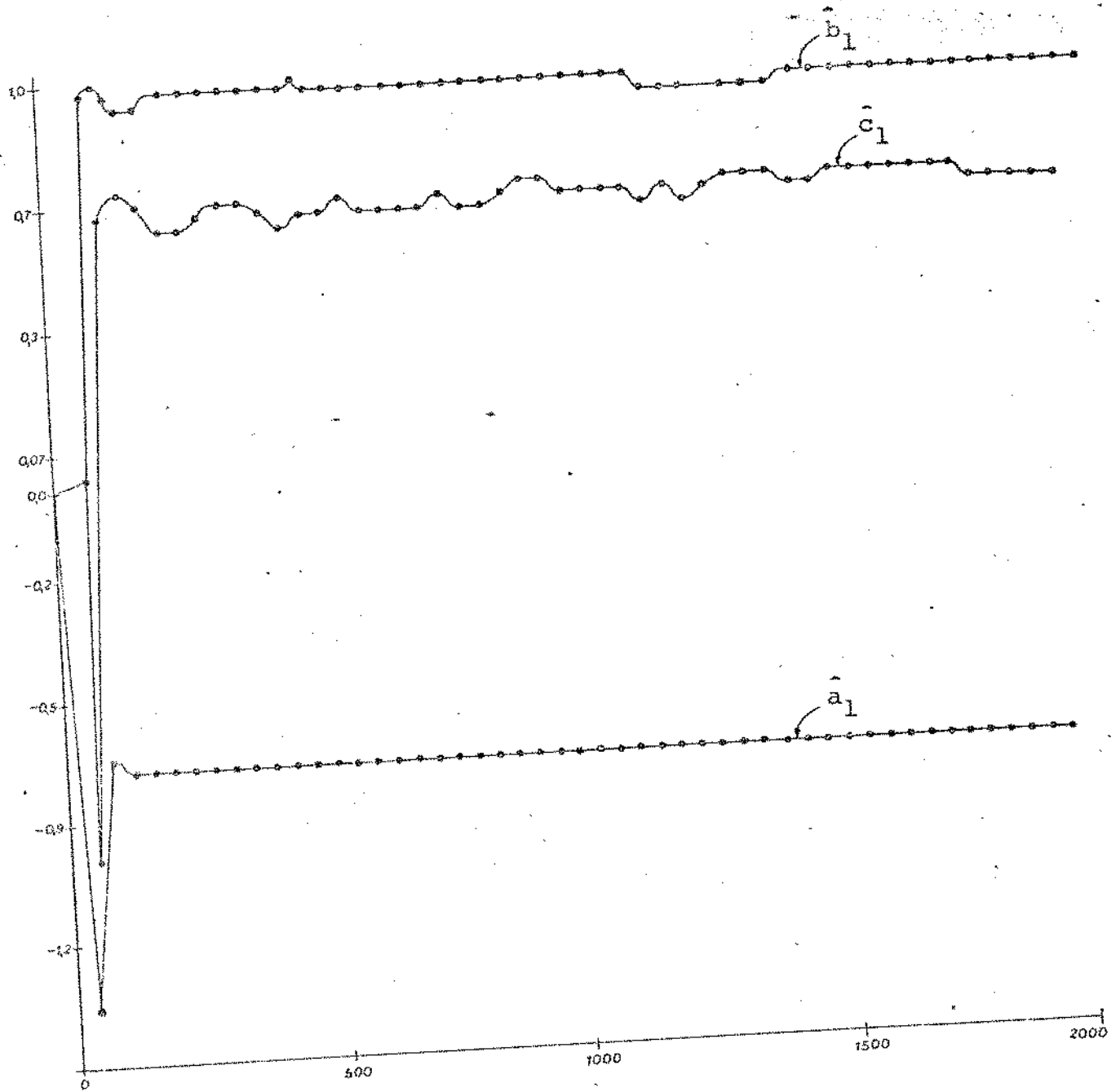


FIGURA 5.7

$$\underline{\theta}(0) = \underline{0}$$

$$\underline{P}(0) = 10 \underline{I}$$

$$\frac{\underline{S}}{N} = 1$$

$$\underline{Q}(0) = 10 \underline{I}$$

PROCESSO $\underline{P}_1$

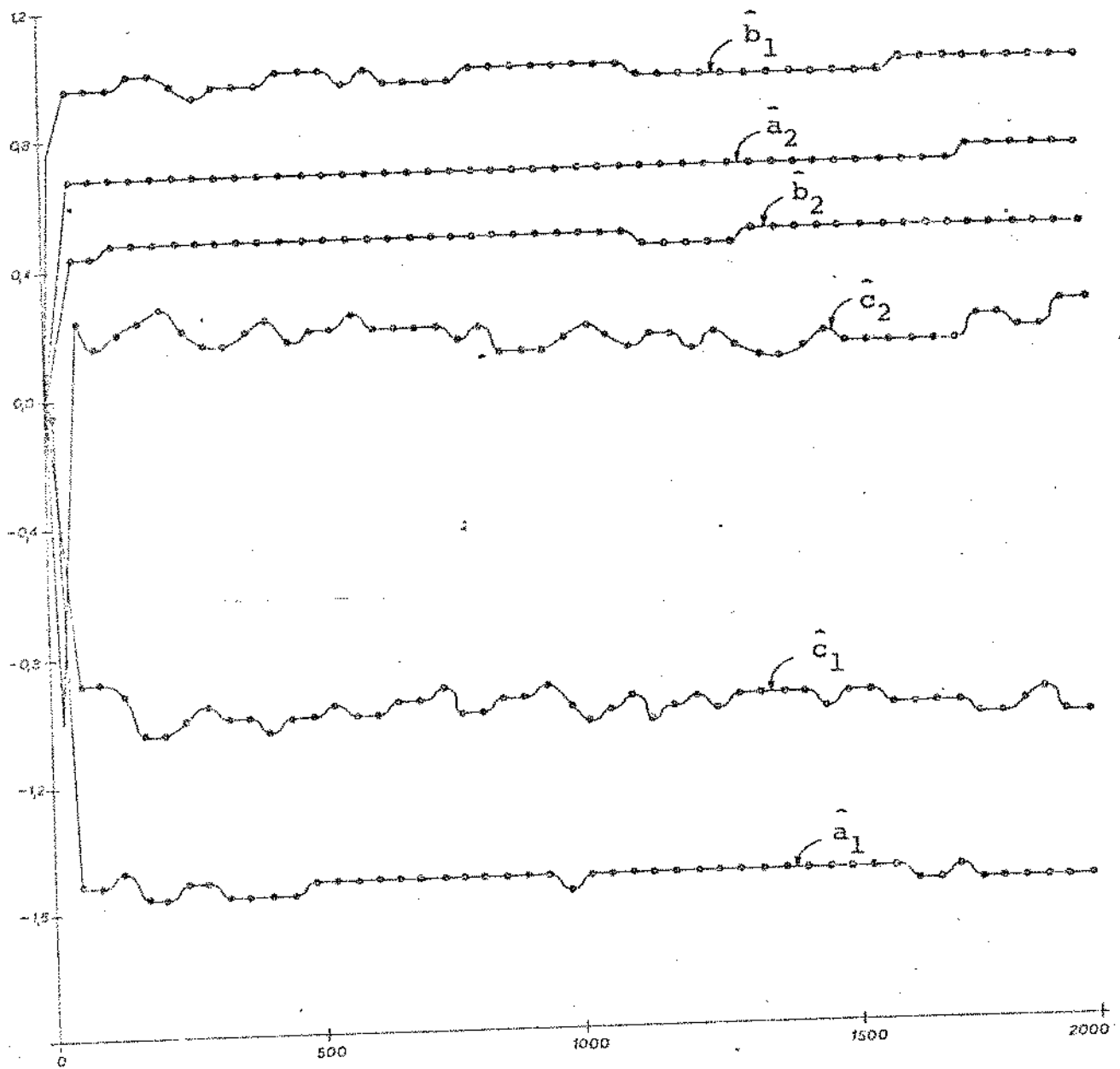


FIGURA 5.8

$$\hat{\theta}(0) = \underline{0} \quad \underline{P}(0) = 10 \underline{I}$$

$$\frac{\underline{S}}{N} = 1 \quad \underline{Q}(0) = 10 \underline{I}$$

PROCESSO P_2

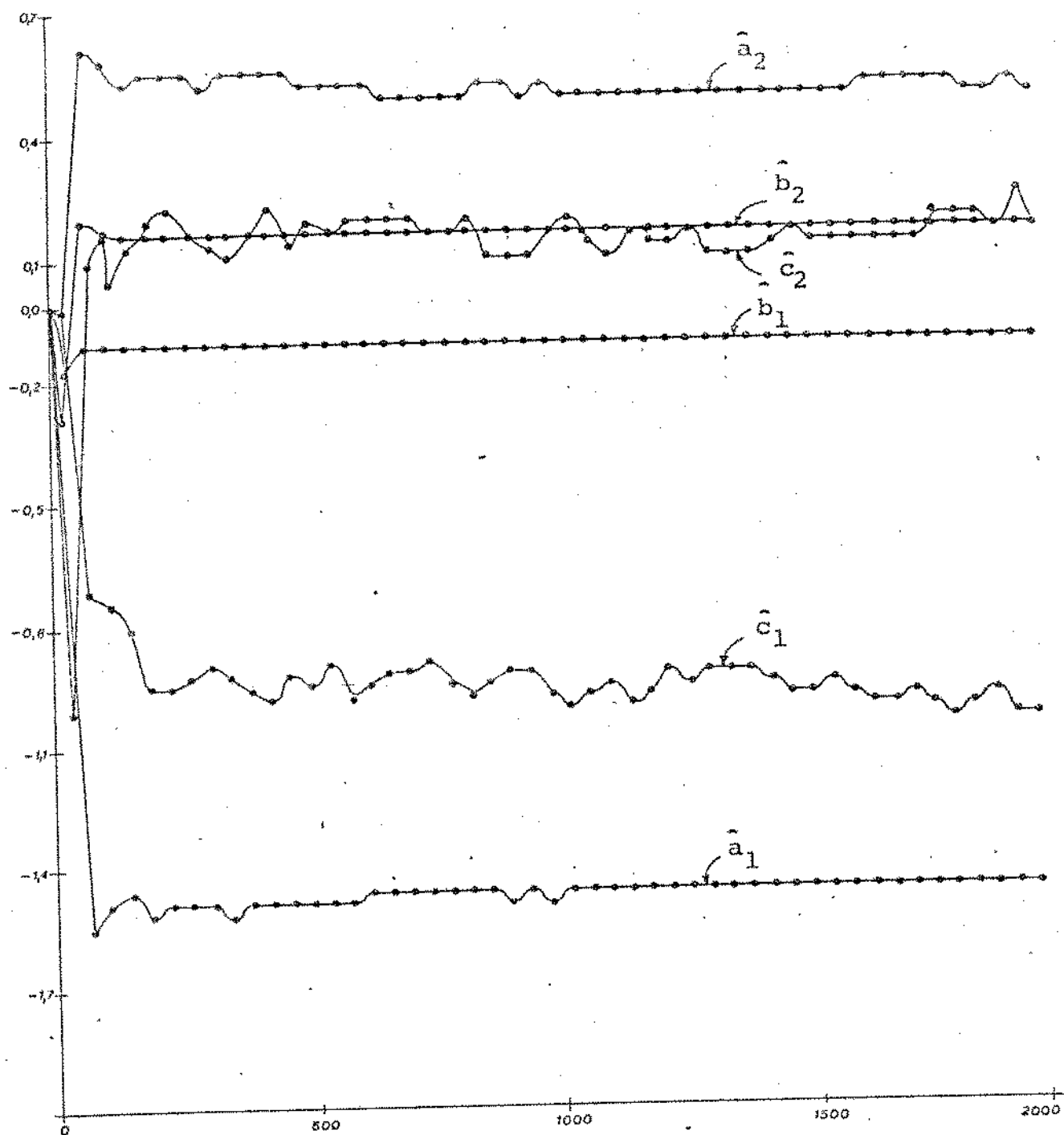


FIGURA 5.9

$$\hat{\theta}(0) = 0$$

$$\underline{P}(0) = 10 \underline{I}$$

$$\underline{Q}(0) = 10 \underline{I}$$

$$\underline{S} = 1$$

N

PROCESSO P_3

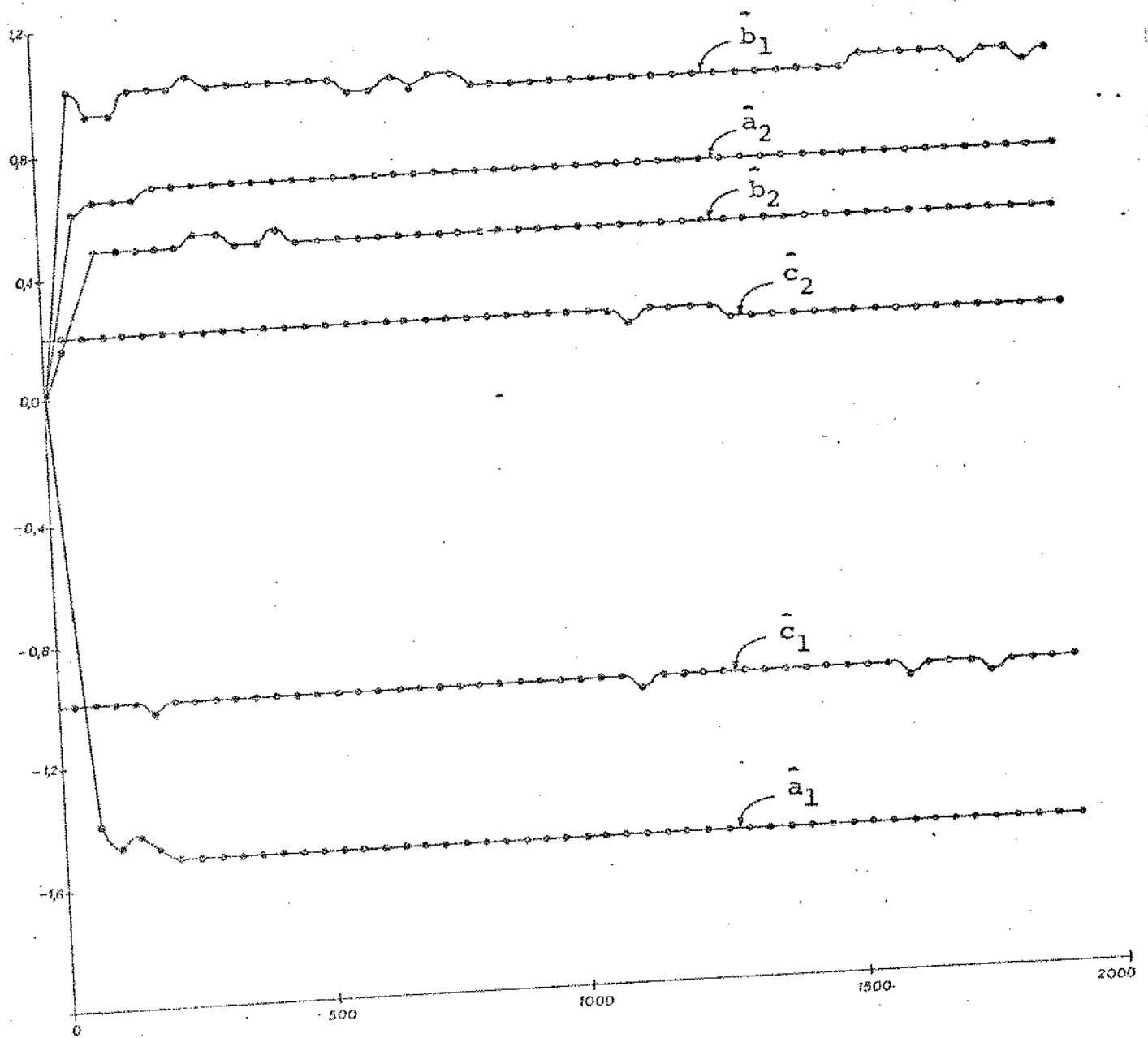


FIGURA 5.10

$$\hat{C}_1(0) = 0,2$$

$$\underline{P}(0) = 10 \underline{I}$$

$$\hat{C}_2(0) = -1,0$$

$$\underline{Q}(0) = 10^{-3} \underline{I}$$

$$\frac{\underline{S}}{N} = 1$$

PROCESSO $\underline{P}_2$

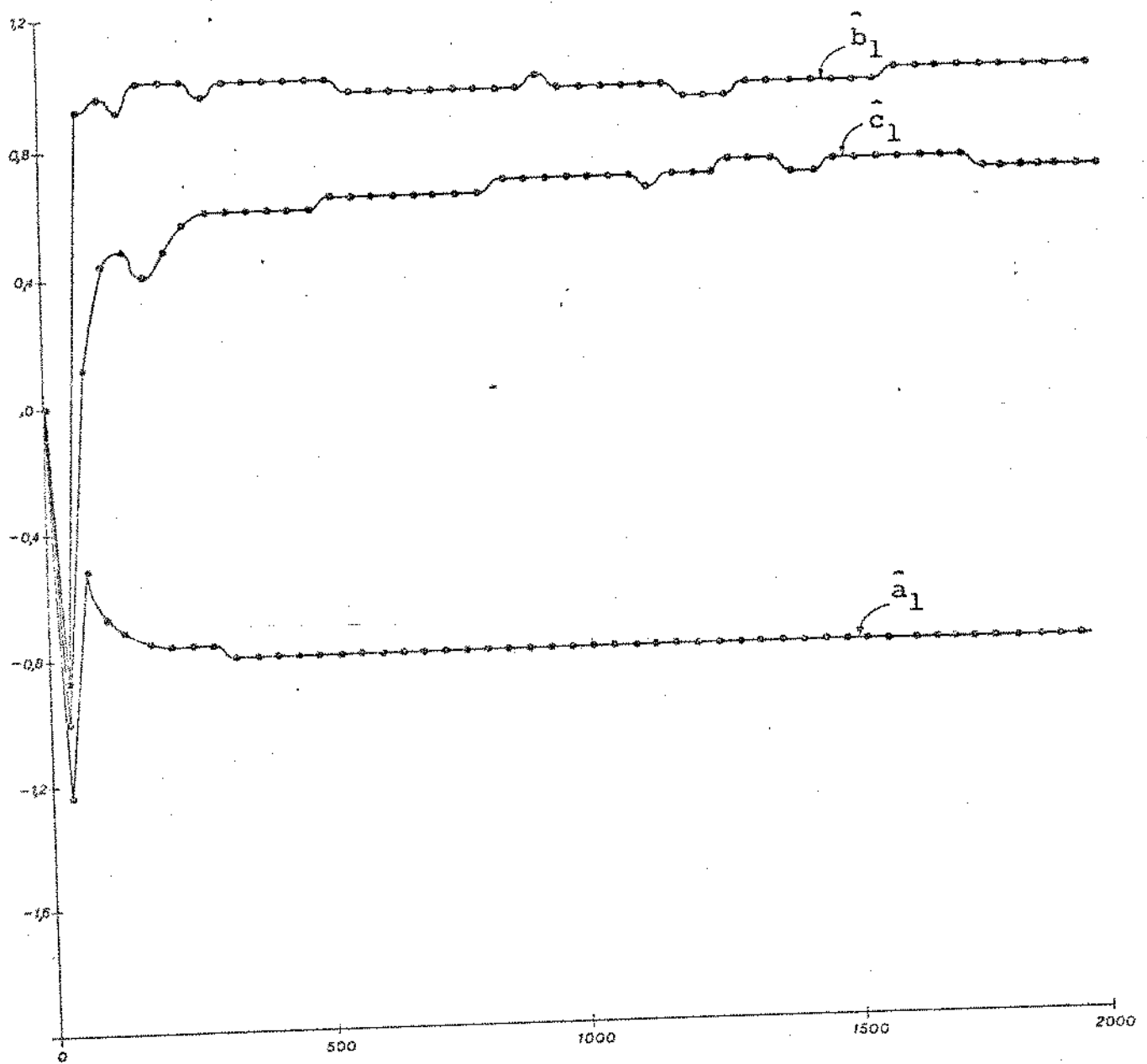


FIGURA 5.11.a

$$\hat{a}_1(0) = 0 \quad \hat{b}_1(0) = 0 \quad \hat{c}_1(0) = 0$$

$$\underline{P}(0) = 1 \underline{I} \quad \underline{Q}(0) = 1 \underline{I}$$

$$\frac{S}{N} = 1$$

PROCESSO P_1

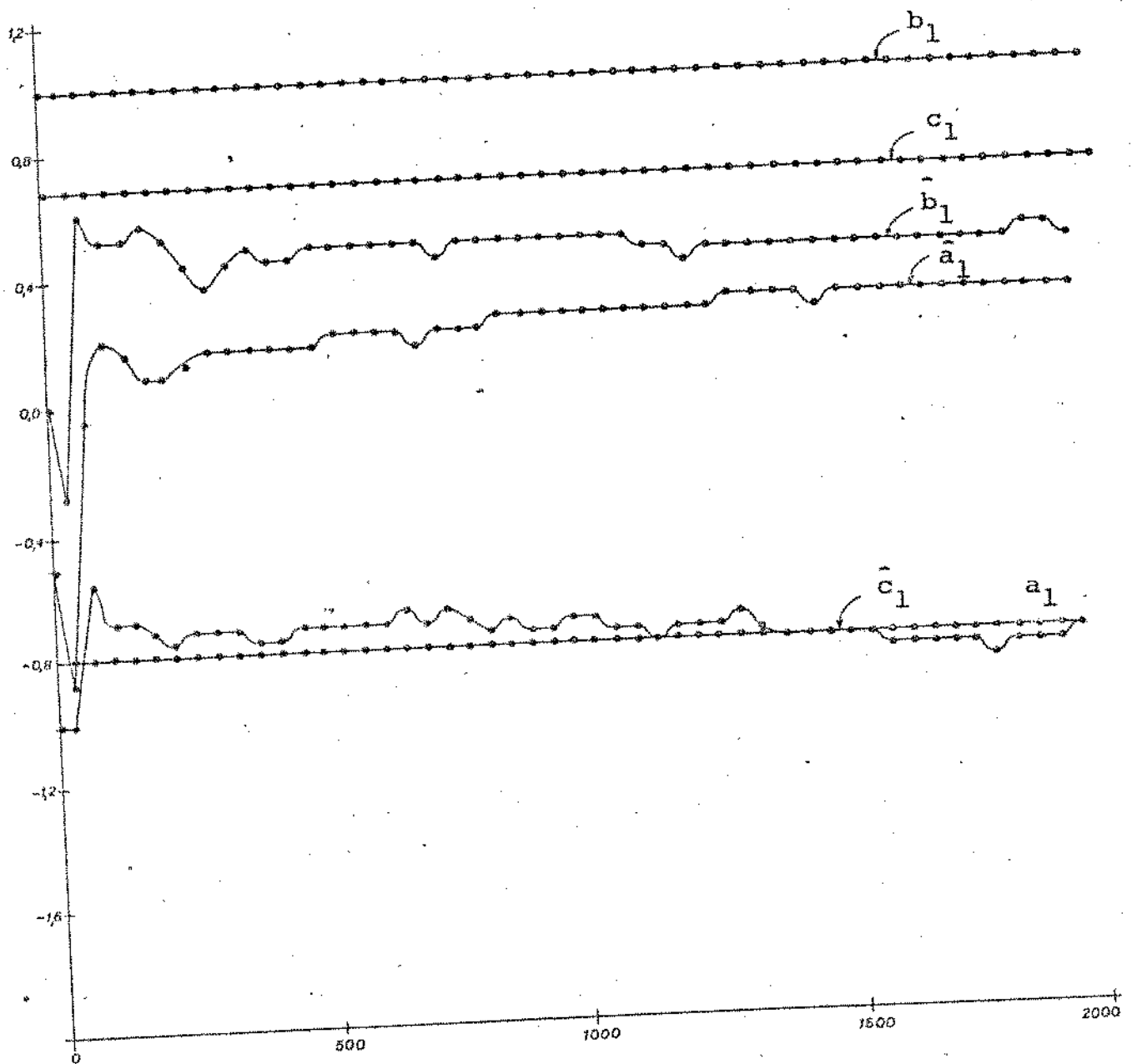


FIGURA 5.11.b

$$\hat{a}_1(0) = -0,5 \quad \hat{b}_1(0) = 0 \quad \hat{c}_1(0) = -1,0$$

$$\underline{P}(0) = 1 \underline{I} \quad \underline{Q}(0) = 1 \underline{I}$$

$$\frac{S}{N} = 1$$

PROCESSO P_1

$$e(k) = \frac{C_M(z^{-1})}{D_M(z^{-1})} \zeta(k) = \frac{1 + \hat{c}_1 z^{-1}}{1 + \hat{d}_1 z^{-1}} \zeta(k)$$

Portanto o algoritmo estimará tanto os parâmetros do numerador como do denominador da função de transferência do ruído.

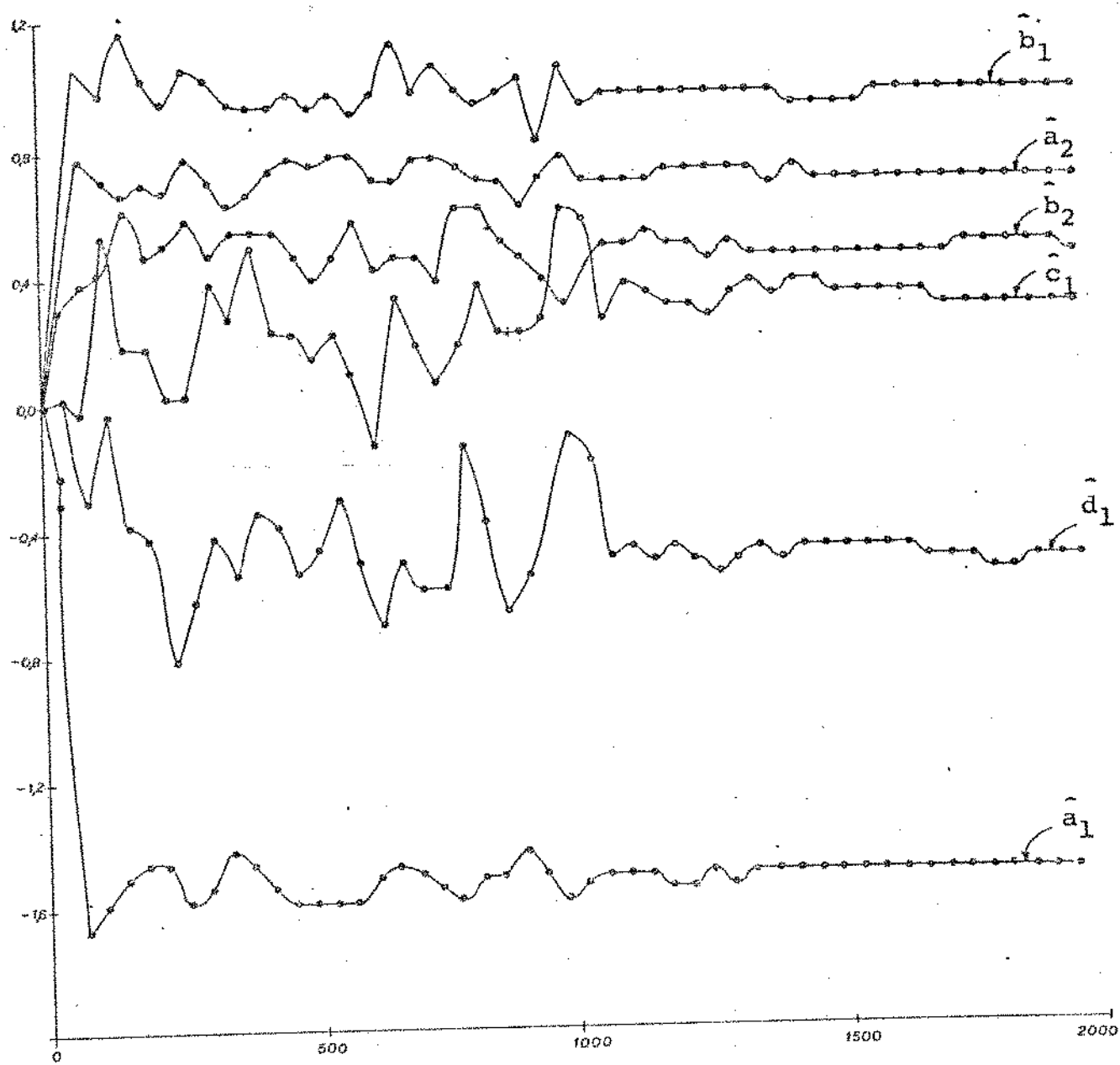


FIGURA 5.12

$$\hat{\theta}(0) = \underline{0} \quad \underline{P}(0) = 10 \underline{I}$$

$$\frac{S}{N} = 1$$

PROCESSO P_2

Caso 2 :-

PARÂMETRO	VALOR VERDADEIRO	VALOR ESTIMADO 2000 MEDIDAS
a_1	-1.5	-1.510 $\pm$.038
a_2	.7	.715 $\pm$.039
b_1	1.0	1.003 $\pm$.047
b_2	.5	.492 $\pm$.048
d_1	-1.0	..778 $\pm$.103
d_2	.2	.151 $\pm$.102

TABELA 5.1

A tabela 5.1 apresenta os resultados para o caso onde o modelo para o ruído é uma auto-regressão.

Neste caso o modelo de ruído é descrito por:

$$e(k) = \frac{1}{D_M(z^{-1})} \zeta(k)$$

Portanto o algoritmo estimará o quociente dos polinômios $D(z^{-1})/C(z^{-1})$, o que equivale a estimar os coeficientes da equação (5.14)

$$n(k) = \frac{1}{1 - \underset{d_1}{0,8A}z^{-1} + \underset{d_2}{0,24}z^{-2} + \underset{d_3}{0,072}z^{-3} + \dots} \zeta(k)$$

Neste exemplo verifica-se que um modelo de segunda ordem é suficiente porque d_3 é muito pequeno.

Caso 3 :-

As tabelas 5.2 e 5.3 apresentam os resultados obtidos para os casos onde o modelo do ruído é interpretado como uma média-móvel, isto é:

$$e_{(K)} = C_M'(z^{-1}) \zeta_{(K)}$$

Portanto este algoritmo estimará o quociente dos polinômios $C(z^{-1})/D(z^{-1})$, o que equivale a estimar os coeficientes da equação (5.15)

$$n_{(K)} = (1 + 0,8q^{-1} + 0,4q^{-2} + 0,2q^{-3} + 0,1q^{-4} + 0,05q^{-5} + \dots) \zeta_{(K)} \quad (5.15)$$

Observa-se das tabelas 5.2 e 5.3 que um aumento no número de parâmetros do modelo do ruído melhora os valores estimados; este fato ocorre dado que só o parâmetro c_5 é desprezível.

Portanto, verifica-se que, estimando por média-móvel, o modelo deverá ser de maior ordem do que o modelo para estimação auto-regressivo. Isto ocorre em particular quando, os parâmetros da auto-regressão decrescem mais rapidamente que os de média-móvel.

Principal Conclusão

- Este algoritmo permite modelar o ruído como média-móvel ou auto-regressão, possibilitando-se, como mostrado acima, a utilização de modelos de menor ordem.

5.7 - SIMULAÇÃO DO ALGORITMO DE MÍNIMOS QUADRADOS GENERALIZADO COM O RUÍDO MODELADO COMO MÉDIA-MÓVEL

Este método foi usado para estimação de modelos de primeira e segunda ordem.

Nas figuras 5.13 e 5.14 são apresentadas as evoluções dos parâmetros, para realizações do processo P_1 e P_2 respectivamente.

PARÂMETRO	VALOR VERDADEIRO	VALOR ESTIMADO 2000 MEDIDAS
a_1	-1.5	$-1.512 \pm .044$
a_2	.7	$.712 \pm .043$
b_1	1.0	$1.010 \pm .050$
b_2	.5	$.491 \pm .050$
c_1	.8	$.745 \pm .098$
c_2	.4	$.359 \pm .120$
c_3	.2	$.140 \pm .094$

TABELA 5.2

PARÂMETRO	VALOR VERDADEIRO	VALOR ESTIMADO 200 MEDIDAS
a_1	-1.5	-1.500 $\pm$.046
a_2	.7	.707 $\pm$.043
b_1	1.0	1.005 $\pm$.050
b_2	.5	.493 $\pm$.050
c_1	.8	.755 $\pm$.099
c_2	.4	.395 $\pm$.123
c_3	.2	.228 $\pm$.122
c_4	.1	.107 $\pm$.095

TABELA 5.3

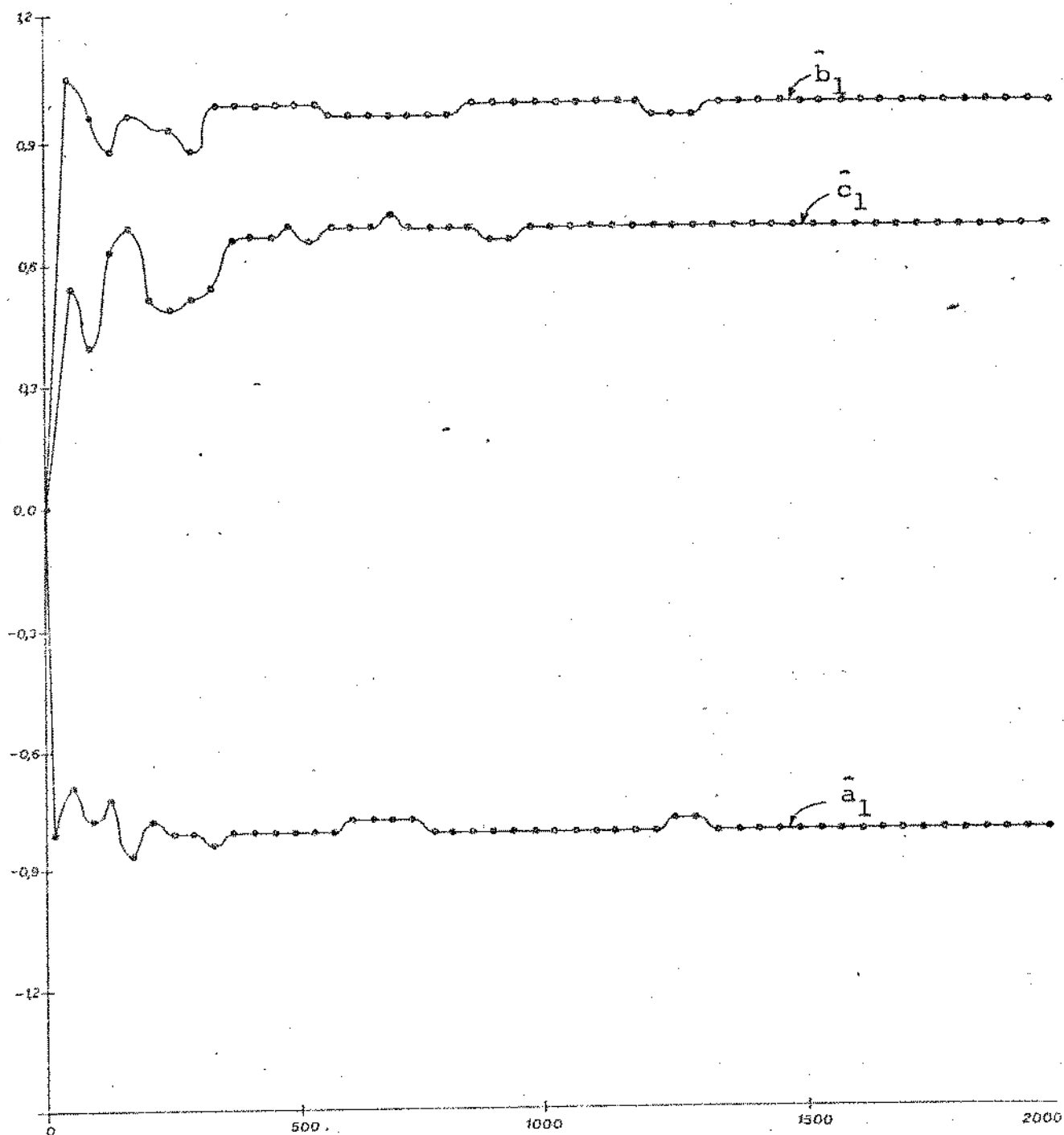


FIGURA 5.13

$$\hat{\theta}(0) = \underline{0} \quad \underline{P}(0) = 10 \underline{I}$$

$$\frac{\underline{S}}{N} = 1$$

PROCESSO P_1

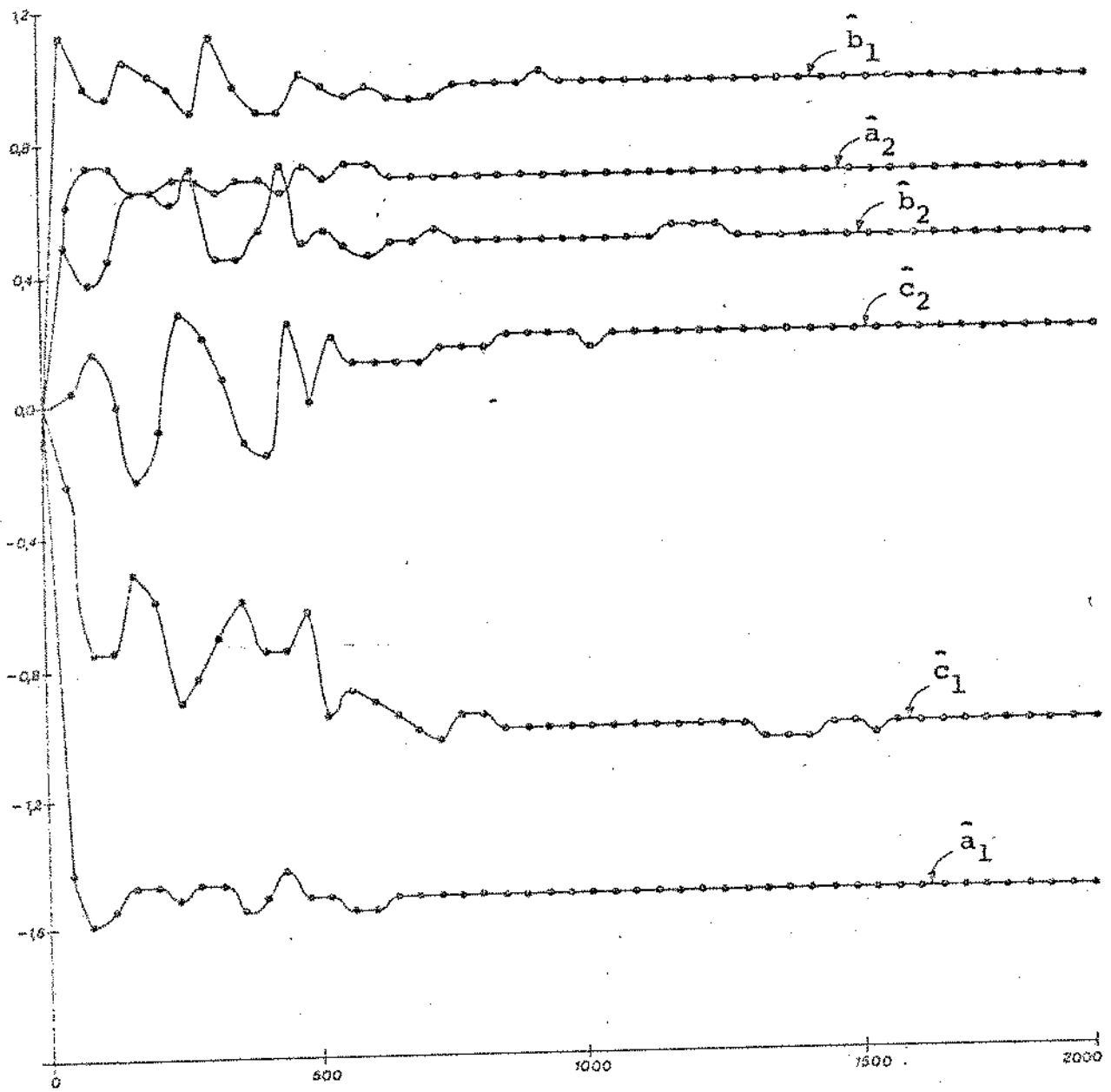


FIGURA 5.14

$$\hat{\theta}_{(0)} = \underline{0} \quad \underline{P}_{(0)} = 10 \underline{I}$$

$$\frac{S}{N} = 1$$

PROCESSO $\underline{P}_2$

Simula-se a seguir o exemplo apresentado no capítulo 4, sem a modificação sugerida, onde os valores estimados não convergem para os valores verdadeiros.

Como visto no capítulo anterior, o processo P_4 apresenta autovalores na matriz $G_{(\theta_v)}^{-1} H_{(\theta_v)}$ com parte real positiva, e portanto este algoritmo não deve convergir para o valor verdadeiro.

Na tabela 5.4, onde são mostrados os parâmetros estimados, vê-se que, mesmo quando o algoritmo é inicializado corretamente, ocorrem valores estimados diferentes dos verdadeiros

PARÂMETRO	VALOR VERDADEIRO	VALOR ESTIMADO $P_{(0)} = 10^{-4} I$ 2000 MEDIDAS
a_1	.90	.902 $\pm$.012
a_2	.95	.930 $\pm$.012
b_1	1	1.012 $\pm$.047
c_1	1.5	1.238 $\pm$.071
c_2	.75	.381 $\pm$.073

TABELA 5.4

Utilizando a modificação sugerida no capítulo anterior, é simulado o mesmo processo com condições iniciais nulas e $P_{(0)} = 10 I$. Observa-se na figura 5.15 que o algoritmo converge para o valor verdadeiro, como era esperado teoricamente.

Principais Conclusões

- Fator de Ponderação: a escolha do fator de ponderação igual a .96 até 500 medidas foi a que forneceu melhor precisão.

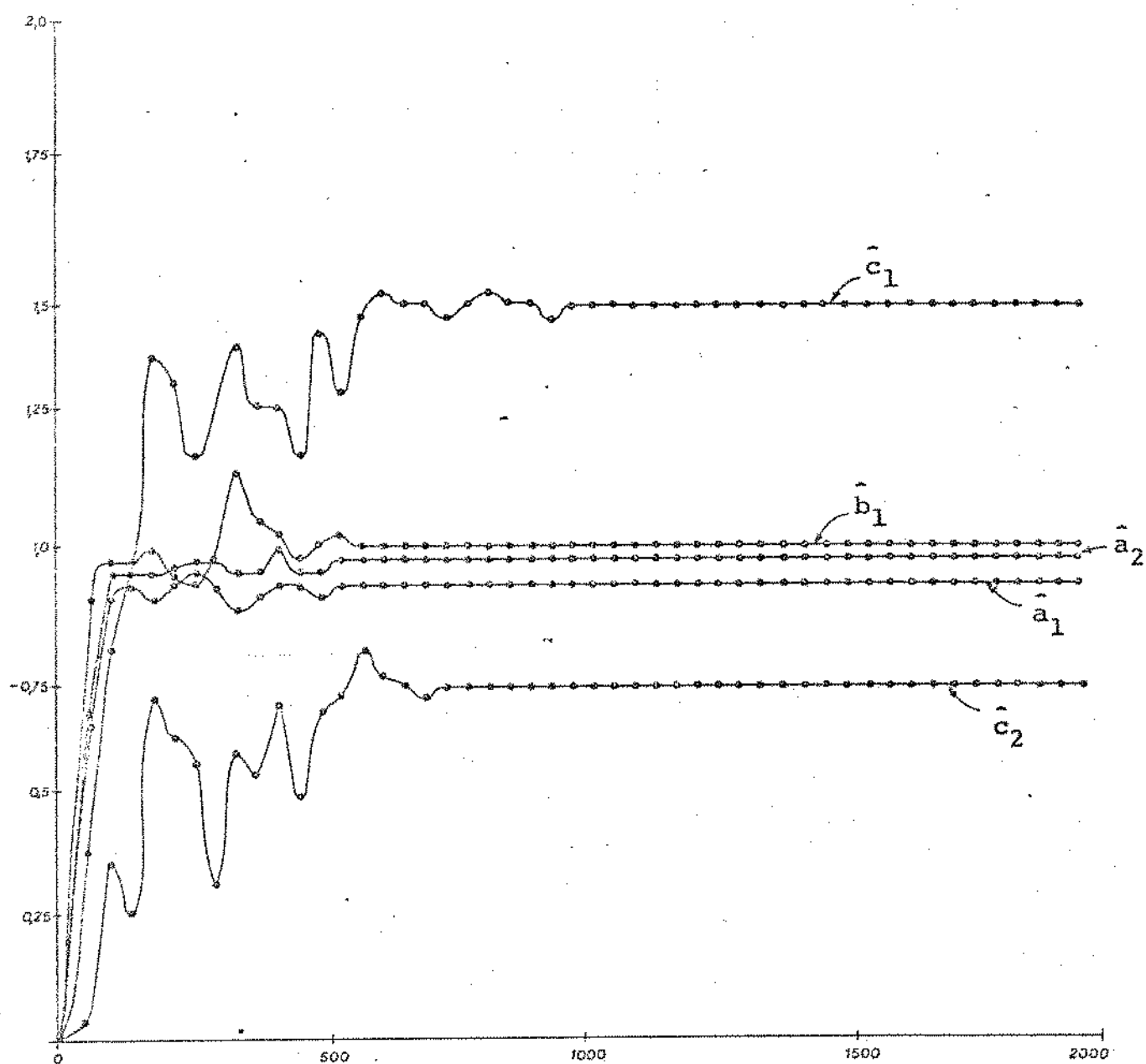


FIGURA 5.15

$$\hat{\theta}(0) = 0 \quad \underline{P}(0) = 10 \underline{I}$$

$$\frac{S}{N} = 1$$

Com o fator de ponderação igual à unidade o algoritmo tem uma taxa de convergência muito lenta para a estimação dos parâmetros do ruído;

- Convergência: analogamente aos casos anteriores, observa-se que a estimação para o sistema de primeira ordem converge mais rapidamente que para os sistemas de segunda ordem.

5.8 - SIMULAÇÃO DE UM SISTEMA COM PARÂMETRO VARIANDO NO TEMPO

Para demonstrar o desempenho dos algoritmos quando os parâmetros variam no tempo, discutem-se os resultados da simulação para um exemplo simples, (sistema de ordem zero), com o parâmetro b variando deterministicamente como mostrado na figura 5.16.

Seja o processo:

$$Y(K) = b(K) u(K) + n(K)$$

com $n(K) = (1 + 0,7q^{-1}) \xi(K)$ e

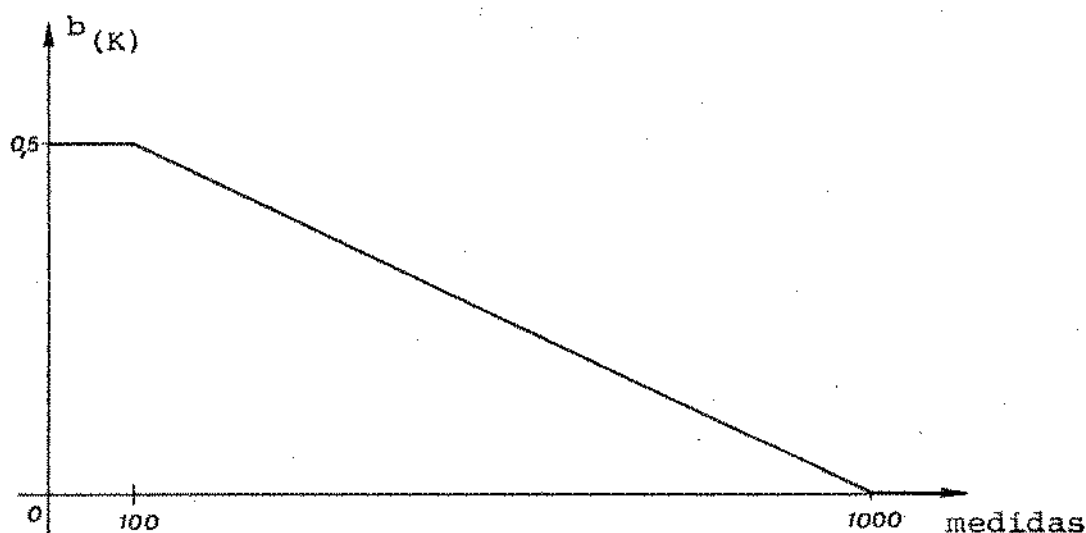


FIGURA 5.16

Os resultados obtidos são apresentados nas figuras 5.17 e 5.18.

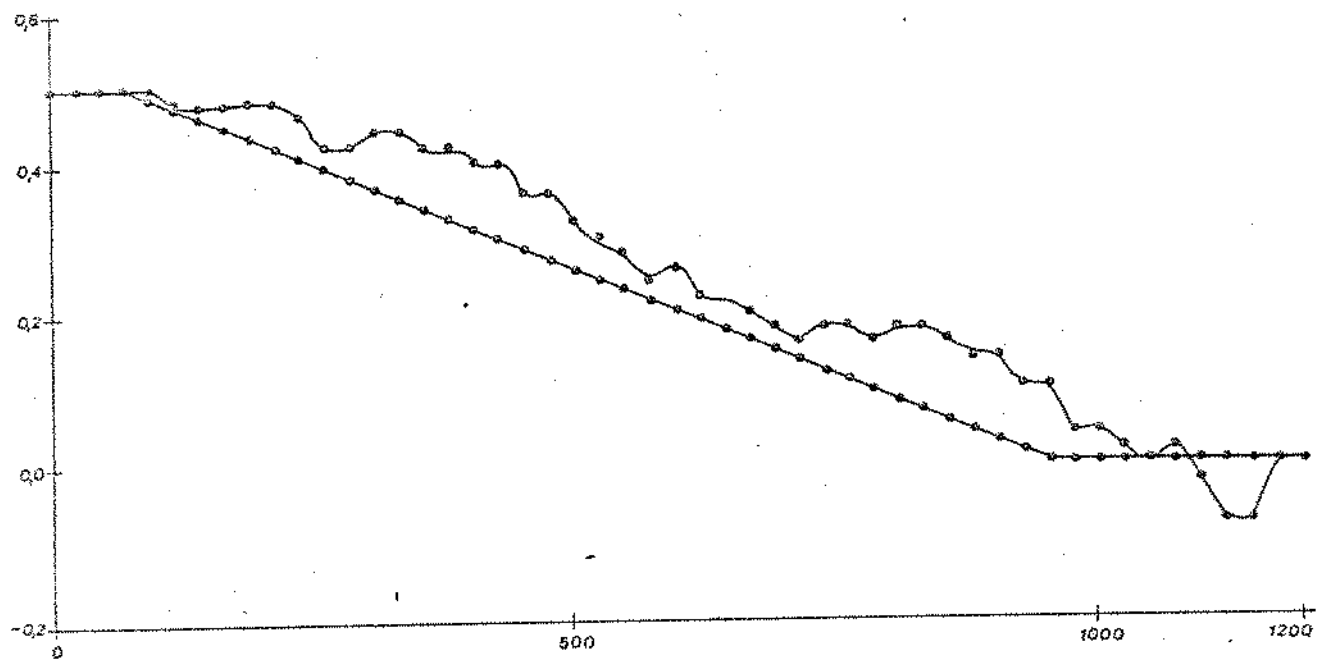


FIGURA 5.17

PARÂMETRO VARIANDO

$$\rho = 0,99$$

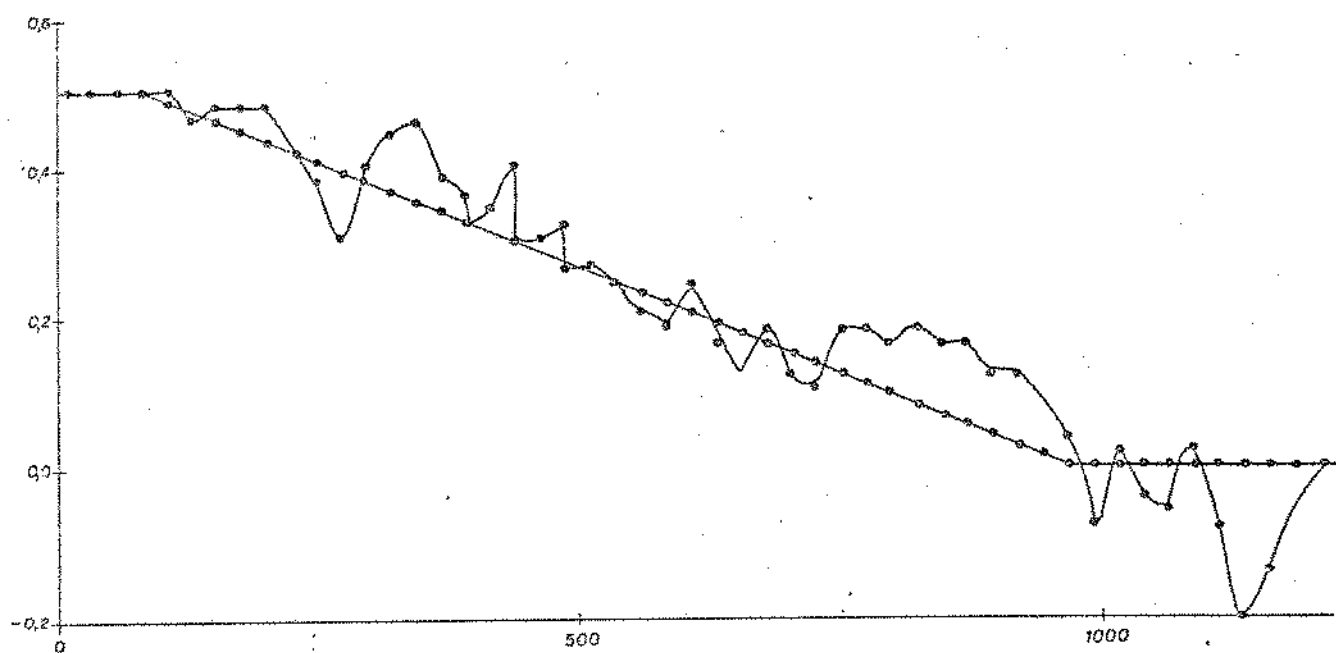


FIGURA 5.18

PARÂMETRO VARIANDO

$$\rho = 0,97$$

Observa-se que com o fator de ponderação menor que a unidade o algoritmo acompanha melhor a variação do parâmetro b , mas a variância assume valores maiores, como previsto teoricamente.

5.9 - EXEMPLO DE DETERMINAÇÃO DA ORDEM DO MODELO

Para exemplificar a idéia apresentada no terceiro capítulo, para a determinação da ordem do modelo, foram feitas simulações para o seguinte processo de segunda ordem:

$$Y(K) = 1,5Y(K-1) - 0,7Y(K-2) + u(K-1) + 0,5u(K-2) + \zeta(K) + 1,5\zeta(K-1) + 0,1\zeta(K-2)$$

Na tabela 5.5 são fornecido os valores da função de custo para várias ordens do modelo e com a ordem do ruído assumindo os valores 1 e 2.

ORDEM MODELO	ORDEM RUÍDO	FUNÇÃO DE CUSTO
1	1	4,24
2	1	1,79 (*)
3	1	1,80
4	1	1,81
1	2	2,98
2	2	1,80 (*)
3	2	1,91
4	2	1,81

(*) ordem escolhida

TABELA 5.5

Neste exemplo, tanto para o modelo de ruído de primeira ordem como de segunda, obteve-se o mesmo modelo do processo, pois

o coeficiente do termo de segunda ordem do ruído é muito menor que os de ordem inferior.

5.10 - COMPARAÇÃO DOS MÉTODOS DE ESTIMAÇÃO

Neste item apresenta-se nas tabelas 5.6 e 5.7 os resultados da simulação de um sistema de primeira e um de segunda ordem, para o método dos mínimos quadrados, média-móvel e auto-regressivo.

Das tabelas 5.6 e 5.7 verifica-se que modelando corretamente o ruído, e a relação sinal-ruído não sendo muito baixa, tanto o algoritmo auto-regressivo como a média-móvel produzem valores estimados corretos.

EXEMPLO 5.10.1

Sistema de primeira ordem (Processo P_1) com $\frac{S}{N} = 10$ e 2000 medidas

$$\underline{\theta}_{(0)} = \underline{0}$$

$$\underline{P}_{(0)} = 10 \underline{I}$$

PARÂMETRO	VALOR VERDADEIRO	MÍNIMOS QUADRADOS	MÉDIA-MÓVEL	AUTO-REGRESSIVO
a_1	-.8	$-.846 \pm .020$	$-.799 \pm .002$	$-.799 \pm .004$
b_1	1.	$1.007 \pm .045$	$.999 \pm .002$	$.999 \pm .010$
c_1	.7	—	$.679 \pm .052$	$.740 \pm .150$

TABELA 5.6

EXEMPLO 5.10.2

Sistema de segunda ordem (Processo P_2) com $\frac{S}{N} = 10$ e 2000 medidas

$$\hat{\theta}(0) = 0$$

$$\underline{P}(0) = 10 \underline{I}$$

PARÂMETRO	VALOR VERDADEIRO	MÍNIMOS QUADRADOS	MÉDIA-MÓVEL	AUTO-REGRESSIVO
a_1	-1.5	-1.389 $\pm$.025	-1.492 $\pm$.040	-1.499 $\pm$.002
a_2	.7	.596 $\pm$.024	.692 $\pm$.035	.701 $\pm$.002
b_1	1.0	1.017 $\pm$.045	1.002 $\pm$.120	.999 $\pm$.002
b_2	.5	.596 $\pm$.051	.502 $\pm$.160	.499 $\pm$.002
c_1	-1.0	—	-.983 $\pm$.095	-.987 $\pm$.098
c_2	.2	—	.207 $\pm$.091	.177 $\pm$.097

TABELA 5.7

5.11 - CONCLUSÕES

- A partir das simulações conclui-se, como esperado teoricamente, que o algoritmo dos mínimos quadrados recursivo produz valores estimados não polarizados somente quando o ruído é "branco" e portanto quando isto ocorre, é interessante utilizar este

método de estimação, porque o mesmo requer um menor número de cálculos e conseqüentemente um menor tempo de computação;

- Os outros métodos considerados, desenvolvem modelos para os parâmetros do processo do ruído, necessitando portanto de maior tempo de computação.

- Para o auto-regressivo os parâmetros estimados convergirão para os valores verdadeiros se a relação sinal-ruído for elevada. Por outro lado, se esta relação é baixa o algoritmo obterá valores estimados polarizados; neste caso a convergência para os valores corretos dependerá da inicialização do algoritmo.

- Para o algoritmo média-móvel foi mostrado que o algoritmo inicial poderia não convergir para os valores verdadeiros, conforme mostrado no exemplo (4.1). Com a introdução de modificações, mostrou-se que o mesmo apresenta boas propriedades de convergência.

Para estudos futuros sugere-se a análise de métodos para a determinação da ordem do modelo que não levem em conta a necessidade do conhecimento *ã priori* da ordem do ruído, assim como a aplicação dos algoritmos estudados em controladores adaptativos, e a determinação de algoritmos de identificação para sistemas multivariáveis.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ASTROM, K.J. & EYKHOOF, P. - "System identification, a survey". In: Identification and Process Parameter Estimation, Czechoslovakia, Academia Prague, 1970. pt.1, p.01. (IFAC Symposium, 2, Prague, 1970).
2. EYKHOOF, P. - System identification: parameter and state estimation. London, Wiley, 1974.
3. TZAFESTAS, S.G. - "Some computer - aided estimators in stochastic control systems identification". Int. Journal Control, 12(3): 385-399, 1970.
4. ASTROM, K.J. & BOHLIN, T. - "Numerical identification of linear dynamic systems from normal operating records". In: IFAC Symposium of Self - Adaptive Control Systems Teddington, Engl., 1966, p.96-110.
5. CLARKE, D.W. - "Generalized-least-squares estimation of the parameters of a dynamic model". In: Identification in Autom. Control System, Prague, 1967. p.3.17. (IFAC Symposium, Prague, 1967).
6. HASTINGS - JAMES, R. & SAGE, M.W. - "Recursive generalized least square procedure for on-line identification of process parameters". Proc.IEE, London, 116(12): 2057 - 2062, Dec., 1969.
7. TALMAN, J.L. & VAN DEN BOOM, A.J.W. - "On the estimation of the transfer function parameters of process and noise dynamic using a single-stage estimation". In: EYKHOOF, P. - Identification and system parameters estimation. Amsterdam, North-Holland Pub. Co, 1973. pt.2, p.711-720. (IFAC Symposium, 3 Hague/Deeft, Netherlands, 1973).
8. LJUNG, L. ; SODERSTROM, T. ; GUSTAUSSON, I. - "Counterexamples to general convergence of commonly used recursive identification method". IEEE Trans. Aut. Control, 20(5): 643-652, Oct., 1975.

9. LJUNG, L. & WITTENMARK, B. - "Analysis of a class of adaptive regulators". In: Stochastic Control Symposium. Budapest, IFAC, 1974, p.431-437.
10. ISERMANN, R. et all. - "Comparison and evaluation of six on-line identification and parameter estimation methods with three simulated process". In: EYKHOOF, P. ed. - Identification and system parameter estimation. Amsterdam, North-Holland Pub. Co, 1973 pt.2, p.1061-1082. (IFAC Symposium, 3. Hargue/Deeft, Netherlands, 1973)
11. AMARAL, W.C. ; MAGALHÃES, L.P. e MENDES, M.J. - Identificação de processos, através de diferentes modelos de ruído (a ser publicado).
12. AMARAL, W.C. ; MAGALHÃES, L.P. ; MENDES, M.J. - Determinação da ordem do Modelo. Trabalho apresentado a Sociedade Brasileira para Progresso da Ciência. Brasília, 7-14 de julho, 1976.
13. SÖDERSTRÖM, T. ; LJUNG, L. ; GUSTAUSSON, I. - A comparative study of recursive identification methods. Lund Institute of Technology, Dep. Automatic Control, 1974. (Report 7427, Dec. 1974).
14. SÖDERSTRÖM, T. - "Convergence properties of the generalised least square identification method". Automatica, Oxford, 10(6): 617-626, Nov., 1974.
15. SÖDERSTRÖM, T ; LJUNG, L. ; GUSTAUSSON, I. - A comparative study of recursive identification methods. Lund Institute of Technology, Dep. Automatic Control, 1974. 138p. (Report 7427, Dec. 1974)
16. ÅSTRÖM, K.J. - Introduction to stochastic control theory. New York, Academic Press, 1972. 299p. (Mathematics in science and engineering, v.70).
17. VAN DEN BOOM, A.J.W. & VAN DEN ENDEN, A.W.M. - "The determination of the order of processes and noise dynamics". In: EYKHOOF, P. ed. - Identification and system parameter estimation. Amsterdam, North-Holland Pub. Co., 1973. pt.2, p. 929-938.

18. BALAKRISHMAN, A.V. - Communication theory. New York, McGraw-Hill, 1968. (Inter-University Electronic series, v.6).
19. AOKI, M. - Optimization of stochastic systems: topics in discrete-time systems. London, Academic Press, 1967. 354p. (Mathematics in science and engineering, v.32).
20. ALBERT, A. - Regression and the Moore-Penrose pseudoinverse. New York, Academic Press, 1972. 380p. (Mathematics in science and engineering, v.94)
21. SCHWEPPE, F.C. - Uncertain dynamic system. Englewood Cliffs, Prentice-Hall, 1973. 563p. (Prentice-Hall electrical engineering series).