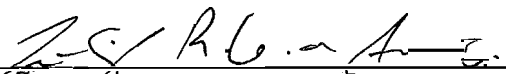


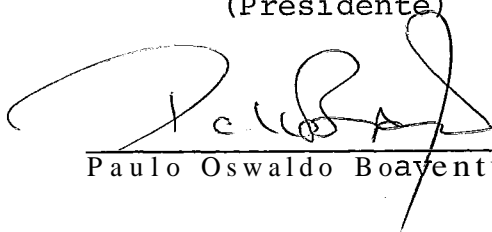
INVESTIGAÇÃO DE ESQUEMAS ADAPTATIVOS PARA OTIMIZAÇÃO DE HEURÍSTI
CAS

Renato Antonio Rabuske

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE
PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JA
NEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO
GRAU DE DOUTOR EM CIÊNCIAS (D.Sc.)

Aprovada por:


João Lázaro R. H. de Araujo
(Presidente)


Paulo Oswaldo Boaventura Netto


Nelson Maculan Filho


Rajamani Doraiswami


Demétrio Alonso Ribeiro

Rio de Janeiro, RJ - Brasil

Dezembro de 1980

RABUSKE, RENATO ANTONIO

INVESTIGAÇÃO DE ESQUEMAS ADAPTATIVOS PARA OTIMIZAÇÃO DE
HEURÍSTICAS (Rio de Janeiro, 1980)

, P. 29.7 cm (COPPE/UFRJ , D. Sc. , Engenharia de Siste
mas, 1980)

Tese - Universidade Federal do Rio de Janeiro. Faculda
de de Engenharia

ESQUEMAS ADAPTATIVOS - I. COPPE/UFRJ

II. Título (série)

À minha esposa Mária e aos
meus filhos Allan e Patrícia

AGRADECIMENTOS

- A DEUS, que com sua onisciência, assiste e lidera todo saber.
- Ao Professor João Lizardo Rodrigues Hermes de Araújo, pela eficiente orientação, pela solicitude no atendimento, pela disponibilidade e paciência, pela ajuda que deu.
- Aos professores da banca examinadora, por terem aceito participar da mesma.
- Aos professores Jaime Luiz Szwarcfiter e Paulo Oswaldo Boaventura Netto, pela participação na banca das provas de qualificação.
- A todos os colegas com que convivi durante o curso de Doutorado, pelo estímulo, humor e companhia que proporcionaram.
- Aos colegas, professores do Departamento de Ciências Estatísticas e da computação da Universidade Federal de Santa Catarina, por terem aceito suportar o Ônus de meu afastamento.
- À CAPES, pelo auxílio financeiro que proporcionou.
- À Universidade Federal de Santa Catarina por permitir meu afastamento.
- À Professora Vânia conceição Tavares, por ter defendido meus interesses temporais na função de Procuradora.
- A todos os que, direta ou indiretamente, contribuíram para a realização deste trabalho.

RESUMO

O presente trabalho tem como objetivo central montar um esquema adaptativo estocástico e testá-lo verificando suas potencialidades e limitações. O enfoque heurístico vale-se de resultados apresentados por ARAUJO³.

A evolução do trabalho processa-se levando avante concomitantemente aspectos teóricos e experimentais.

O Desenvolvimento da pesquisa abrange os seguintes aspectos:

- a) construção de um Esquema Adaptativo e otimização de seus parâmetros.
- b) Eliminação da tendenciosidade do esquema, seja através de fatores de correção, seja pelo uso de fórmulas onde se evita ao máximo o recurso a valores estimados.
- c) Otimização da Aplicação do Esquema e outros aspectos como escalarização de parâmetros, aceleração da busca, instabilidades ocasionais, sensibilidade do esquema e sua aplicação na determinação de caminhos mínimos ou quase-mínimos.
- d) comparação dos resultados obtidos com os esperados, permitindo concluir que o esquema apresenta um bom índice de desempenho.

A B S T R A C T

The present work has as its central objective to establish an adaptive stochastic scheme and test it checking its potentials and limitations. The heuristic approach uses results presented by ARAUJO³.

The evolution of the work proceeds taking into account simultaneously theoretical and experimental aspects.

The development of the research encompasses the following aspects:

- a) Construction of an Adaptive Scheme and optimization of its parameters.
- b) Elimination of the bias of the scheme, either by means of correction factors, or by use of formulas where the use of expected values is mostly avoided.
- c) Optimization of the Scheme Application and other aspects as parameter scaling, search acceleration, occasional instabilities, sensitivity of the scheme and its application in the determination of minimal or quasi-minimal paths.
- d) Comparison of the obtained results with the expected ones, allowing to conclude that the scheme presents a good performance index.

I N D I C E

RESUMO	iv
ABSTRACT	v
INDICE	vi
CAPÍTULO I	
1.1 - Introdução	1
CAPÍTULO II	
2 - CONCEITOS BÁSICOS DA PROGRAMAÇÃO HEURÍSTICA	4
2.1 - Grafo de um Problema	4
2.2 - Conceituação	5
2.3 - Busca Heurística	7
2.4 - Medidas de Desempenho	10
2.5 - O Algoritmo B*	12
CAPÍTULO III	
3 - HEURÍSTICA E BUSCA HEURÍSTICA	15
3.1 - Heurística em geral	15
3.2 - Heurística, Probabilidade e Processos Estocásticos ..	18
3.2.1 - Otimização de $\hat{h}(n)$ - Formulação do Problema .	21
3.2.2 - O caso linear - Forma geral de Solução	23
CAPÍTULO IV	
4 - CONSTRUÇÃO DE ESQUEMAS ADAPTATIVOS	27
4.1 - Introdução	27
4.2 - Definições - Formulação do Problema - Limitações ...!	28

4.3	- Esquema básico de Aproximação Estocástica	29
4.4	- Modelo para o grafo de Busca	30
4.5	- Esquema de A. E. para o caso $\hat{\delta}_m = \delta$	31
4.6	- Cálculo de \underline{A}_t	34
4.7	- Cálculo de \underline{S}_t	35
4.8	- Estimação dos P_i 's	35
4.9	- Cálculo dos L_i 's	38
4.10	- Análise experimental	39
4.11	- Resultados práticos do Esquema	41
4.12	- Novas fórmulas para $\tilde{\underline{A}}_+$ e $\bar{\underline{S}}_+$	41
4.13	- Outras formas para $\tilde{\delta}_+$	42
4.14	- Nova forma de cálculo para \bar{h}	43
4.15	- Resultados Experimentais com \tilde{h} ajustado	44
4.16	- Conclusões	45

CAPÍTULO V

5	- CORREÇÃO DA TENDENCIOSIDADE DE \underline{A}_t e \underline{S}_t	46
5.1	- Introdução	46
5.2	- Tentativa 1: Correção em \underline{A}_t e $\bar{\underline{S}}_t$	46
5.2.1	- Determinação da correção para $\bar{\underline{A}}_t$	47
5.2.2	- Cálculo de p	47
5.2.3	- Análise de \underline{S}_t	49
5.2.4	- Resultados Experimentais	52
5.3	- Tentativa 2: Correção em \underline{A}_t e $\tilde{\underline{S}}_t$	52
5.3.1	- Determinação da Correção para $\tilde{\underline{A}}_t$	52
5.3.2	- Determinação da Correção para $\tilde{\underline{S}}_t$	54
5.3.3	- Análise Experimental	56
5.4	- Tentativa 3: $\bar{\underline{A}}_t$ e $\bar{\underline{S}}_t$ sem C_i e δ	56
5.4.1	- Supressão de C_i e δ_t	56

5.4.2	- Mudança de $n\tilde{o}$ referencial	58
5.4.3	- Resultados Experimentais	58
5.5	- Tentativa 4: Uso de probabilidade	59
5.5.1	- Resultados Experimentais	59
5.6	- Conclusões	60

CAPÍTULO VI

6	- DISCUSSÃO DE ASPECTOS DOS ESQUEMAS PROPOSTOS E NOVAS PERSPECTIVAS	61
6.1	- Árvore de busca com expansão Dirigida	61
6.2	- Aproximação de \underline{S}_t por Matrizes Escalares	62
6.2.1	- Resultados Exp. com Expansão não dirigida e dirigida	63
6.3	- \underline{C}_t em separado	64
6.4	- \underline{C}_t fixado a priori	65
6.5	- Uso da diagonal de \underline{S}_t	65
6.6	- Instabilidade de \underline{S}^{-1}	66
6.6.1	- Soluções viáveis	66
6.6.2	- Enfoque em Aberto	67
6.7	- Manutenção da consistência entre \underline{A}_t e \underline{S}_t	68
6.7.1	- Uso da diferença máxima da Diagonal	68
6.7.2	- Resultados Experimentais	69
6.7.3	- Correção por Linha e coluna	69
6.8	- Estudo da Influência de δ na busca	75
6.9	- Variação da Dimensão de \underline{A} e \underline{S}	76
6.10	- Sobre o Menor Caminho	76
6.11	- Otimização da Aplicação do Esquema	82
6.12	- Sensibilidade e precisão do Esquema	85
6.13	- Conclusão	85

CAPÍTULO VII

7	- CONCLUSÕES	87
7.1	- Sobre a estimativa dos parâmetros	87
7.2	- Sobre a correção da tendenciosidade	88
7.3	- Sobre a aplicação do Esquema	89

APÊNDICES

1	- Resumo dos Experimentos Realizados	91
---	--	----

REFERÊNCIAS BIBLIOGRÁFICAS	147
----------------------------------	-----

ANEXOS

1	- Teoria subjacente à Tese	151
2	- Listagem da programação em Computador	161

C A P Í T U L O I

1.1 - INTRODUÇÃO

O presente trabalho propõe-se como objetivo, estudar um esquema adaptativo estocástico, observando suas deficiências, indicando alternativas, descobrindo suas limitações e estudando uma forma conveniente de aplicação.

Para conseguir o acima, foram seguidos os seguintes passos, traduzidos em diversos capítulos, cujos resumos seguem.

Capítulo II

É um apanhado das principais definições e conceitos necessários para um bom entendimento do trabalho. Os principais teoremas e corolários da programação heurística são aí enunciados, não havendo nenhuma preocupação em demonstrá-los, pois as provas se encontram nas referências bibliográficas. Acentuamos de modo especial o Algoritmo A*, base deste trabalho e acrescentamos o Algoritmo B*, que é uma das últimas novidades que surgiram na bibliografia especializada.

Capítulo III

Pretende ser uma revisão bibliográfica sobre heurística e busca heurística. Revela a escassez de autores que se aventuram em encarar aspectos probabilísticos dentro da heurística

tica. Na parte final do capítulo foi dado um enfoque mais detalhado sobre os antecedentes, ou seja, a origem deste trabalho.

Capítulo IV

Neste capítulo formulamos um problema e definimos seus parâmetros básicos. Em seguida sugerimos uma primeira proposta de solução e apresentamos as ferramentas de que o esquema se serve para atingir seu objetivo. A cada nova hipótese segue um experimento prático e uma análise dos resultados. Cada argumento, cada hipótese é questionada e tentativas de otimização são analisadas.

Capítulo V

Em vista de uma tendenciosidade latente no esquema, mostrada no capítulo IV, a análise foi dirigida para possíveis formas de eliminá-las. São efetuadas quatro tentativas, três delas apresentando visível sucesso.

Capítulo VI

Neste capítulo fazemos testes com hipóteses que no decorrer do trabalho foram ventiladas. Encaramos aspectos como \underline{S}_t matriz escalar, $\underline{\Sigma}_t$ conhecido ou estimado a priori, aspectos da instabilidade de \underline{S}_t^{-1} , influência de θ na busca, variação na dimensão de \underline{S} e \underline{A} , aspectos sobre o caminho mínimo e sua obtenção através da aplicação do esquema proposto, otimização da aplicação do esquema e sua sensibilidade, bem como uma nova forma de

expansão da árvore de busca.

Capítulo VII

Relata as conclusões e sugestões encarando em separado três aspectos do trabalho, quais sejam: Estimativa dos parâmetros (referente ao capítulo IV), Correção da tendenciosidade (cap. V) e Sobre a Aplicação do Esquema (cap. VI).

C A P Í T U L O I I

CONCEITOS BÁSICOS DA PROGRAMAÇÃO HEURÍSTICA

Os conceitos que seguem foram coletados em diversos livros e artigos especializados, destacando-se especialmente HART ET AL¹⁵, NILSSON²³, ARAUJO⁴ e GOMEZ¹², onde o leitor poderá aprofundar os tópicos enunciados neste capítulo.

O objetivo deste capítulo é reunir conceitos que posteriormente serão de alguma forma úteis, evitando-se definições extra-contexto.

2.1 - Grafo de um Problema

Suponhamos que um problema admita a seguinte formulação:

Sejam S o conjunto de estados iniciais (conhecidos), T o conjunto de estados desejados (terminais) e Γ um conjunto de operações sobre os estados, isto é, se e_i é um estado do problema e $\gamma \in \Gamma$ é uma operação aplicável a e_i , então $\gamma(e_i) = (e_i, C_i)$, onde e_i é um estado do problema e C_i é o custo associado a aplicação de γ a e_i .

Assim posto, o problema é isomorfo a um grafo ponderado e orientado a que chamaremos de grafo do problema e terá como notação $G_S(S, \Gamma)$, ou simplesmente G_S :

Definição 2.1

$G_S(S, \Gamma)$ é chamado Grafo de um problema, se e so-

mente se:

- (i) $S \subset V(G_s)$, o conjunto de vértices (ou nós) de G_s .
- (ii) $n_i \in V(G_s)$, $y \in \Gamma$ onde $y(n_i) = (n_j, C_{ij})$
 $n_j \in V(G_s)$ e C_{ij} = custo do arco (n_i, n_j) e n_j é de terminado sucessor de n_i .

2.2 - Conceituação

6 - Grafos

É a classe de grafos com custos de transição C_{ij} superiores a algum $\delta > 0$.

Custos Aditivos

Seja $L(n_o, n_r)$ um caminho de n_o a n_r . Os custos são aditivos se pudermos calcular o custo de L como

$$K_L(n_o, n_r) = \sum_{j=0}^{r-1} C_{j, j+1} \quad (\text{II-1})$$

Definição 2.2

seja $n \in V(G_s)$, então

$$g(n) = \begin{cases} \text{Min } K_L(s, n) & \text{se } L(s, n) \text{ existe} \\ s \in S & \\ + \infty & \text{caso contrário} \end{cases} \quad (\text{II-2})$$

$$h(n) = \text{Min}_{t \in T} K_L(n, t) \quad (\text{II-3})$$

Definição 2.3

O custo mínimo de um caminho de S a T , devendo passar por um nó $n \in V(G_S)$, é definido

$$f(n) = g(n) + h(n) \quad (\text{II-4})$$

Definição 2.4

$L(n_o, n_r)$ é dito solução de um problema se e somente se, $n_o \in S$ e $n_r \in T$.

Definição 2.5

$L(n_o, n_r)$ é uma solução Ótima se e somente se

- (i) $n_o \in S$
- (ii) $n_r \in T$
- (iii) $K_L(n_o, n_r) = f(S)$.

Definição 2.6

Um algoritmo de busca é admissível, se para todo δ -grafo tendo uma solução finita, o algoritmo garante encontrar uma solução Ótima em um número finito de passos. Note-se que resolver o problema equivale a explorar parcialmente G_S até achar uma solução.

Definição 2.7

Se a busca da solução é feita seletivamente, utilizando informações específicas ao domínio do problema, dizemos que a busca é heurística e o algoritmo usado de Algoritmo de busca

heurística.

2.3 - Busca Heurística

Suponhamos a possibilidade de calcular uma função de mérito $\hat{f}(n)$ para cada $n \in V(G_s)$. Podemos usar o algoritmo de busca ordenada de HART ET AL¹⁵, sendo necessário armazenarmos:

(i) Descrição do estado que o nó n representa.

(ii) Distinção entre nós abertos e fechados.

Seja $F(n) = 0 \rightarrow n$ aberto

$F(n) = 1 \rightarrow n$ fechado.

(iii) Valor de \hat{f} dos nós abertos.

(iv) Apontador P para seu melhor antecessor.

Sejam S : conjunto de vértices iniciais

T : conjunto de vértices terminais

Algoritmo de Busca Ordenada

P1 - Para todo $s \in S$, calcule $f(s)$, faça $F(s) = 0$, $P(s) = 1$ e guarde.

P2 - Escolha entre os nós abertos o nó n com melhor \hat{f} ; Se houver mais de um, desempate arbitrariamente, mas sempre a favor de nós terminais. Se não há nó aberto, pare com fracasso.

P3 - Se $n \in T$, faça $F(n) = 1$ e pare. Seguindo-se os apontadores P , obtém-se a solução.

P4 - Se não, faça $F(n) = 1$ e gere todos os sucessores de n ; Para cada sucessor m de n , calcule $\hat{f}(m)$. Se m é um novo nó, ou se o novo valor de

$\hat{f}(m)$ é melhor do que o anterior, faça $P(m)=n$,
 $F(m) = 0$ e guarde-o com o novo $\hat{f}(m)$.

P5 - Volte a P2.

Observações

Este algoritmo só para com fracasso se o problema não tem solução. O caso mais importante quanto à forma de \hat{f} é: $\hat{f}(n) = \hat{g}(n) + \hat{h}(n)$ (Algoritmo A^*), onde $\hat{h}(n)$ (função heurística), é uma estimativa de $h(n)$ e $\hat{g}(n) = k(s,n)$. Se $\hat{h}(n)$ depende do estado da busca, seu valor deve ser atualizado para todos os nós abertos ao fim do passo P4.

Teorema 2.1 (HART ET AL¹⁵)

Se $\hat{f}(n) = \hat{g}(n) + \hat{h}(n)$ e $\hat{h}(n) \leq h(n)$ para todos os nós de G_s , então A^* é admissível e além disso, somente nós tendo $\hat{f}(n) \leq f(s)$ são fechados. Este resultado é extensível a grafos finitos com custos de transformação não negativos.

Definição 2.8

Uma função heurística é dita consistente se e somente se, para todo m e $n \in V(G_s)$

(i) $\hat{h}(n) \leq h(n)$

(ii) $\hat{h}(m) - \hat{h}(n) \leq K_L(m,n) \rightarrow m < n$

Teorema 2.2 (HART ET AL¹⁵)

Se \hat{h} é consistente, quando A^* fecha um nó, já achou o caminho Ótimo até ele.

Proposição 2.1 (NILSSON²³)

Se \hat{h} é consistente, e a sequência de nós fechados por A^* é (n_1, n_2, \dots, n_r) , então, para todo par i, j ($i < j$) temos $\hat{f}(n_i) \leq \hat{f}(n_j)$.

Definição 2.9

Seja θ_n^A o conjunto de índices usado pelo algoritmo A no nó n . Então, se $\theta_n^{A^*} \subset \theta_n^A$ para todos os nós n em G_S , podemos dizer que o algoritmo A é não mais informado que o algoritmo A^* .

Seja \mathcal{Q}^* o conjunto de todos os algoritmos que agem como A^* quando não há empates, mas que resolvem os empates de forma diferente.

Teorema 2.4 (HART ET AL¹⁵)

Seja A qualquer algoritmo admissível, não mais informado que os algoritmos em \mathcal{Q}^* , e suponha que a condição de consistência seja satisfeita por $h(\cdot)$ usado nos algoritmos em \mathcal{Q}^* , então, para todo δ -grafo G_S , existe um $A \in \mathcal{Q}^*$ tal que todo nó não terminal expandido por A^* , também o será por A .

Corolário 2.1

Seja $N(A^*)$ e $N(A)$ o número de nós visitados pelos Algoritmos A^* e A respectivamente e as premissas do teorema 2.4 estejam satisfeitas, então, para qualquer δ -grafo G_S , $N(A^*) \leq N(A)$, ocorrendo a igualdade se e somente se A expande exatamente os mesmos nós não-terminais que A^* .

Corolário 2.2

Se $R(A^*)$ é o número de empates que ocorreu na aplicação de A^* e se as premissas do teorema 2.4 estão satisfeitas,

então, para qualquer 6-grafo G_s , $N(A^*) \leq N(A) + R(A^*)$.

Definição 2.11 (VANDERBRUG³⁰)

Um algoritmo A é dito Completo se, para todo 6-grafo G_s possuindo solução finita, ele encontra uma num número finito de passos.

Teorema 2.5 (ARAUJO⁴)

Se existe $a > 0$ e $z \in \mathbb{R}$ tais que $z \leq \hat{h}(n) \leq \alpha \hat{h}(n)$ para todo $n \in V(G_s)$, então A^* é completo.

Definição 2.12 (ARAUJO⁴)

Se \hat{h} é uma variável aleatória, então T , o número de nós fechados por A , também o é, e se o algoritmo satisfaz $E(T), E(T^2) < \infty$, então é chamado Completo em Média Quadrática.

2.4 - Medidas de Desempenho

Um dos primeiros métodos, proposto por DORAN⁸, foi a penetrância $\rho = k/T$, onde k é o custo da solução achada e T o custo de fechamento dos nós.

SLAGLE²⁸, introduziu a razão de profundidade $DR = \log T / \log N_o$, sendo N_o o custo total da busca (medido nas mesmas unidades que T), empregando $\hat{h} = 0$, ou seja, o algoritmo de Dijkstra.

NILSSON²³ propôs o fator de ramificação efetiva R de uma árvore de profundidade k , possuindo T nós, e R satisfaz

$$R \frac{R^k - 1}{R - 1} = T$$

Nilsson sugere que T seja o número total-de nós gerados, o que é um lapso, levando-se em conta a estrutura de A , pois, mesmo numa busca perfeita, o número de nós gerados é maior que k , pois todos os sucessores do nó fechado são gerados.

P e R dependem das dimensões dos problemas. Isto dificulta a avaliação de uma mesma função heurística empregada em problemas diferentes.

Embora DR não tenha este inconveniente, requer o conhecimento de N_0 .

ARAUJO⁴, contorna estes problemas propondo medir o conteúdo de informação de uma função heurística $\hat{h}(n)$, por um número, tal que a busca usando $\alpha h(n)$ fecha o mesmo número de nós (ou dá o mesmo custo de fechamento) que empregando $\hat{h}(n)$. α é aproximadamente independente das dimensões, sendo essencialmente determinado por $\hat{h}(n)$.

Assim, pressupondo que G_s seja uma árvore com fator de ramificação constante B (número de sucessores por nó), com arcos sendo percorridos em ambos os sentidos, tendo custo unitário e que exista uma solução Ótima de comprimento k , fornece como medida de desempenho a "informação heurística equivalente":

$$a = \frac{\log B - \log R}{\log B + \log R} ; R = B^{(1-\alpha)/(1+\alpha)}$$

onde R satisfaz a equação

$$\frac{R^k - 1}{R - 1} = T$$

$$e B = \frac{N? \text{ de nós gerados}}{N? \text{ de nós fechados}} - \frac{C}{T}$$

isto é

$$G = B.T = B \frac{R^k - 1}{R - 1}$$

*

2.5 - O Algoritmo B

* BERLINER⁷, apresenta um algoritmo, por ele denominado B, que prova que um arco na raiz de uma árvore de busca é melhor que qualquer outro dos concorrentes. Para cada nó provê duas avaliações: Uma otimista, outra pessimista. Estabelece com isto limites para avaliações na subárvore deste ramos. Estes limites são gradativamente reajustados (reduzindo o intervalo) durante a busca, até coincidirem.

*

O domínio de B inclui tanto problemas de busca 1-pessoa quanto de 2-pessoas (adversários). Nestes Últimos, uma pessoa procura maximizar uma função enquanto outra procura minimizá-la.

*

O que distingue B da busca "O melhor-primeiro":

(1) Uma busca o melhor-primeiro é entendida como "achar um objetivo a partir do nó que mais promete", sendo este fechado. B está interessado somente em achar "o melhor primeiro passo" em direção ao objetivo, enquanto a prova (*) não estiver completa. A sutileza está no fato de que é sem sentido estender uma ramificação, melhorando o seu valor, se isto não muda o "status" da prova.

(2) B* pode escolher uma estratégia, mesmo que esteja ainda na raiz da árvore.

O algoritmo usa as seguintes variáveis:

CURNODE : guarda a pista do nó corrente

DEPTH : recorda a distância do CURNODE à raiz
 OPTIM : limite a cada passo
 PESSIM : limite a cada passo
 PARENT : apontador para o pai do nó
 MAXOPTIM: guarda a pista do valor mais otimista de todos os sucessores de CURNODE;
 MAXPESS : guarda a pista do valor "pessimista" de todos os sucessores de CURNODE.

observação:

No passo 4 há dois testes que dependem do ponto de vista como é encarado o problema. O operador "<" indica que dependendo deste ponto de vista, > torna-se < e < torna-se >.

P1 - DEPTH \leftarrow 0; CURNODE \leftarrow 0; OPTIM (0) \leftarrow $-\infty$; PESSIM (0) \leftarrow $-\infty$.

P2 - Se CURNODE ainda não foi expandido, então gerar e avaliar os sucessores, dando a todos um apontador para CURNODE;

P3 - BESTNODE —nome do sucessor de CURNODE com melhor valor OPTIM;

ALTERN \leftarrow Nome do Sucessor com segundo melhor valor OPTIM;

MAXOPTIM \leftarrow OPTIM (BESTNODE);

MAXPESS \leftarrow Valor do melhor PESSIM de todos os sucessores;

P4 - Se $\left[\text{MAXOPTIM} > \text{PESSIM} (\text{CURNODE}) \right]$ E $\left[\text{MAXPESS} < \text{OPTIM} (\text{CURNODE}) \right]$, vai a P5

(*) Prova: refere-se à forma como trabalha B*, provando otimalidade para determinado sucessor em cada raiz da árvore de busca.

PESSIM (CURNODE) \leftarrow MAXOPTIM;

OPTIM (CURNODE) \leftarrow MAXPESS;

Se DEPTH > 0

então CURNODE \leftarrow PARENT (CURNODE);

DEPTH \leftarrow DEPTH - 1;

Vai a P3;

Se DEPTH = 0

então se PESSIM (BESTNODE) \geq OPTIM (ALTERN)

então sai com ANSWER = BESTNODE;

P5 - Se DEPTH = 0, então decide estratégia;

Se estratégia = DISPROVEREST, então CURNODE \leftarrow
ALTERN;

Se estratégia = PROVEREST, então CURNODE \leftarrow
BESTNODE;

P6 - Se (DEPTH \neq 0), então CURNODE \leftarrow BESTNODE;

P7 - DEPTH \leftarrow DEPTH + 1; Vai a P2.

O leitor poderá encontrar mais detalhes em [7], onde aparece também uma comparação com outros algoritmos de busca.

C A P Í T U L O I I I

HEURÍSTICA E BUSCA HEURÍSTICA

3.1 - Heurística em Geral

A busca heurística foi uma das idéias importantes no desenvolvimento da pesquisa em Inteligência artificial. É um conceito não totalmente definido e foi usado como um guarda-chuva para muitas técnicas computacionais que são difíceis de classificar ou analisar. Isto é bom porque leva a imaginação por caminhos isentos, a tentar técnicas que funcionam em problemas complexos.

Contudo, a não conceituação leva à repetição. Seguidamente análises analíticas levam a procedimentos mais eficientes.

É difícil atingir as origens do uso da heurística como processo facilitador de tarefas. Aparentemente sempre andou junto com a ciência como fonte inspiradora, mantendo-se transparente.

Contudo, em fins da década 50 e início da 60, apareceram diversos trabalhos que procuravam tornar concreta a intervenção da heurística na solução de diversos problemas.

Para rever o que ocorreu, é preciso ir a bibliografia da época, não tão específica quanto a atual, e proceder cuidadosa pesquisa.

Os programas de busca heurística de NEWELL, SIMON e SHAW^{21,22}, geraram grande entusiasmo. Aplicando as mesmas técnicas

nicas a diversos problemas de domínios diferentes, demonstraram empiricamente o caráter geral da busca heurística.

Já um pouco mais recente é a tentativa de DORAN e MICHIE^{8,18}, com seus Graph Traverser Programs (GTP), em que os mecanismos de busca são programas de busca em grafos, dirigidos por funções heurísticas que são estimadores de distâncias em grafos.

Os GTPs foram usados para comparar a eficiência de diferentes funções heurísticas na solução de quebra-cabeças.

HART, NILSSON e RAPHAEL¹⁵, baseados em experiências anteriores de Nilsson sobre determinação do menor caminho em grafos, descobriram como usar melhor a informação para melhorar a eficiência do cálculo do menor caminho. Eles generalizaram isto para o conceito de estimacão heurística e fizeram um algoritmo que usa esta informacão no cálculo do menor caminho. Também produziram os primeiros teoremas sobre a eficiência como funçãõ de exatidão da funçãõ heurística.

MARTELLI¹⁶, analisa a complexidade de algoritmos, em particular o A*, para o qual atribui $O(2^N)$ passos no pior caso.

Apresenta, além disso, um algoritmo B, que, conforme ARAUJO⁶, é dominado por uma subclasse de A*, com dadas regras de desempate.

GELPERIN¹¹, aponta algumas falhas em demonstracões sobre otimalidade de A*, feitas anteriormente por Nilsson, e apresenta provas novas. Além disso, mostra através de exemplos, que:

1) O valor de $\hat{h}(n)$ pode ser uma funçãõ do estado da busca assim como da informacão heurística disponível. Esta é uma

idéia já antiga, ventilada por Nilsson e repetida por Blaydon em [8].

2) Que existem algoritmos de busca admissíveis que não podem ser simulados por qualquer algoritmo A*.

POHL²⁶, apresenta um modelo particular de busca heurística para achar um caminho num grafo dirigido, o "Heuristic Path Algorithm" (HPA). O HPA tem seu protótipo em MOORE²⁰. O que o diferencia de A* é $f(n) = (1-w)g(n) + wh(n)$, sendo $0 \leq w \leq 1$. Se $w = 0.5$, HPA e A* serão idênticos.

A função avaliação pode ser estendida a uma forma linear geral,

$$f(n) = (1 - \sum_{i=1}^K w_i) g(n) + \sum_{i=1}^K w_i h_i(n)$$

$$\text{com } \sum_{i=1}^K w_i \geq 0 \quad ; \quad \sum_{i=1}^K w_i \leq 1$$

podendo os w_i serem funções dos nós. Isto permitiria um sistema com aprendizado, medir o desempenho de cada termo como contribuição para a estimação da distância.

HARRIS¹⁴, introduz mais uma restrição sobre a heurística, a "Bandwidth Condition" (BC), que permite à busca heurística lidar melhor com problemas práticos de tempo e espaço. O tempo para a busca pode ser reduzido se aproximarmos melhor os valores de $h(\cdot)$. Pode ser impraticável determinar uma heurística que nunca viole a condição de admissibilidade. Em muitos casos é fácil achar um novo valor para $h(\cdot)$, embora sendo este superestimado. BC requer que

$$h(n) - d \leq \hat{h}(n) \leq h(n) + e$$

o que é um afrouxamento da admissibilidade. Claro é que, se $e=0$ e $d > h(n)$ para todos os nós, BC se reduz a admissibilidade.

A busca BC minimiza a dependência crítica em relação à heurística e elimina a necessidade da fixação a priori de restrições ao processo de busca.

3.2 - Heurística, Probabilidade e Processos Estocásticos

No início da década 70 apareceram os primeiros trabalhos relacionando heurística à probabilidade e processos estocásticos. Está atrás desta tentativa de relacionamento, a intenção de mostrar que grande parte dos sucessos da primeira podem ter uma explicação segura nesta última, embora, até o momento, não se tenha a palavra final sobre o assunto. Creio mesmo que tão cedo não se terá, pois, a heurística é uma tentativa de usar a forma como o intelecto humano encara os problemas.

A maioria dos problemas admitem soluções "ad hoc", e então para haver maior síntese nos tipos de soluções, deve-se encarar a possibilidade de reuni-los em famílias de problemas, o que deve ser tentado com mais intensidade, considerando a importância do assunto.

Entre os autores que mais se destacaram no enfoque probabilístico em problemas heurísticos estão ARAUJO^{1,2,3,4 e 5} e VANDERBRUG³¹.

Abordaremos, dentro da ordem cronológica de publi

cação, os principais aspectos enfocados, bem como, os resultados obtidos.

VANDERBRUG³¹, determina o número total esperado de nós expandidos para diversas suposições probabilísticas em relação à profundidade do objetivo, do número de sucessores do nó, e do número de tais sucessores que são expandidos.

Apresenta um modelo probabilístico para estratégias de busca heurística. O modelo considera como medida de desempenho o valor esperado do número de nós a serem fechados, até que se ache a solução. O modelo pode ser aplicado a uma classe de problemas caracterizada por restrições em suas estruturas e na localização dos nós objetivos.

A fim de exemplificar, Vanderbrug desenvolve um modelo nas seguintes condições: O espaço de busca deve ser uma árvore sendo Poisson tanto o número de sucessores por nó como a profundidade da melhor solução.

As principais limitações do modelo são:

- i) O espaço de busca é árvore.
- ii) O pior caso é considerado como se produzisse resultados semelhantes a busca breadth-first.
- iii) Considera como medida de eficiência o número de nós expandidos (não considera custos computacionais, etc ...).

ARAUJO⁴, aborda em primeiro lugar a eficiência da busca heurística, já anteriormente enfocada. Em seguida aborda a busca heurística sob o ponto de vista randômico. Como resultado preliminar estima o número de nós fechados $E(T)$ e o número de nós gerados $E(G)$, respeitadas as condições do problema.

Considera dois casos:

- i) A distribuição de $\hat{\alpha}$ (vide 2.4), tem átomos.
- ii) A distribuição de $\hat{\alpha}$ é absolutamente contínua.

Na parte final é apresentado um algoritmo para avaliar $E(T)$ e $E(G)$. Embora um pouco trabalhosa, a implementação deste algoritmo é sem dúvida viável em relação ao tempo e espaço. O importante é que ele proporciona condições para investigar que valor ou valores dos coeficientes de $\hat{h}(\cdot)$ minimizarão os custos da busca.

Em [1], Araujo aprofunda alguns dos resultados obtidos em [4] e discute rapidamente os resultados de VANDERBRUG³¹, fazendo as seguintes considerações:

- a) O modelo não tem relação com a estrutura do algoritmo de busca.
- b) Para que os resultados seja corretos, requer-se:
 - b1) Todos os nós à profundidade $K+1$ sejam terminais, (K é o comprimento de caminho ótimo).
 - b2) A expressão que calcula o número esperado de nós fechados, não inclua o último.
 - b3) Ao iniciar-se a busca, o algoritmo já saiba que a melhor solução tem comprimento K .

ARAUJO², analisa o problema probabilístico do caminho mínimo, visto através da programação heurística.

Sendo o problema do caminho mínimo definido como: Achar uma tripla (s^*, t^*, γ^*) , com $s^* \in S$, $t^* \in T$, $\gamma^* \in \Gamma(s^*, t^*)$ tal que $K_{\gamma^*}(s^*, t^*) = \min\{K_{\gamma}(s, t) / s \in S, t \in T, \gamma \in \Gamma(s, t)\}$, e sendo as letras com chapéu consideradas variáveis aleatórias, en

foca dois casos:

i) Achar (s^*, t^*, γ^*) tal que

$$E[\hat{K}_{\gamma^*}(s^*, t^*)] = \min_{s, t, \gamma} \{E[\hat{K}_{\gamma}(s, t)]\}$$

(Problema do caminho mínimo esperado).

ii) Dado $a \in (0, 1)$, achar $(s^*, t^*, \gamma^*, d^*)$, tal que

$$\Pr[\hat{K}_{\gamma^*}(s^*, t^*) \leq d^*] \geq a \quad \text{e} \quad d^* = \min\{d / \Pr[\hat{K}_{\gamma}(s, t) \leq d] \geq a\}$$

A esta formulação denomina "Problema Probabilístico do Caminho Mínimo". Os custos dos arcos são aproximados por normais independentes (custos não aditivos).

Para resolver (ii), apresenta um algoritmo, uma extensão de A^* que denomina \tilde{A} .

ARAUJO³, considera buscas heurísticas como processos estocásticos, e procura formular o problema da otimização das heurísticas empregadas.

Como este trabalho inspirou em parte a presente pesquisa e dele adotamos parte da terminologia, vamos estender-nos um pouco mais sobre o mesmo.

3.2.1 - Otimização de $\hat{h}(n)$ - Formulação do Problema

Suponhamos que $\hat{h}(n)$ é da forma $\phi(\hat{H}(n), w)$, onde $\hat{H}(n)$ é um vetor de medições sobre n e $w \in \Omega$ é um vetor de parâmetros a determinar de modo a otimizar o desempenho de A^* .

Seja: $\hat{Z}(W)$, o custo total da geração de nós durante a busca;
 G_s , um grafo probabilístico, portanto $\hat{Z}(W)$ será uma va
riável aleatória;
 $P(w, \epsilon)$, a probabilidade de pararmos numa solução de custo
 igual ou superior a $f(S) + \epsilon$; ($f(S)$ também é variá
vel aleatória).

Para chegar à formulação final do Problema, Araujo descarta diversas alternativas, que, sob aspectos ponderáveis, tornam-se inviáveis e chega a seguinte:

Minimizar em Ω

$$E\left[(\Phi(\hat{H}(h), W) - \theta h)^2\right] \quad (\text{III-1})$$

ou

$$E\left[\left(\frac{\Phi(\hat{H}(h), W)}{h} - \theta\right)^2\right] \quad (\text{III-1}')$$

onde θ é um escalar determinado em separado.

Seja $W(\theta)$ a solução de (III-1) ou (III-1') que dá a menor dispersão das estimativas em torno de $\theta h(n)$. Se a parametrização de W por θ for tratável, então pode-se escolher θ de modo a minimizar.

$$\{E[\hat{Z}(W(\theta))] / \theta \in R^+, P(W(\theta)), \epsilon\} \leq \alpha \quad (\text{III-1}'')$$

reduzindo a otimização a um problema unidimensional. Para obter $E[\hat{Z}(W(\theta))]$, recorreremos aos modelos de ARAUJO^{1,4}. Como, devido a dificuldades no cálculo e desconhecimento das distribuições en

volvidas, a formulação (III-1") pode tornar-se inconveniente, e por outro lado, se temos uma idéia da ordem de grandeza de $f(S)$ e não conhecemos a distribuição de $h(n)$ (o que é frequente), a seguinte formulação é viável:

i) $\forall \theta \in \mathbb{R}^+$, $W(\theta)$ minimiza

$$E\left[\left(\Phi(\hat{A}(h), W) - \theta h\right)^2 / h = h_0\right]$$

onde h_0 é especificado.

ii) $\theta = \theta_0$, onde θ_0 satisfaz $P(\theta_0, \epsilon) \leq a$.

Para valores pequenos de ϵ e a , Araujo sugere para θ_0 , valor próximo a

$$\theta(\alpha, \epsilon) = \sup\{\theta \in \mathbb{R}^+ / P(\theta, \epsilon) \leq \alpha\}$$

e apresenta para $P(\theta, \epsilon)$ longa e complexa dedução.

3.2.2 - O caso Linear - Forma Geral de solução

$$\text{Se } \mu(h, W) = E\left[\frac{\Phi(\hat{H}(h), W)}{h} / h\right] \text{ e}$$

$$\sigma^2(h, W) = \text{Var}\left[\frac{\Phi(\hat{H}(h), W)}{h} / h\right]$$

então (III-1) e (III-1') 'podem ser assim reescritos:

Minimizar em $W \in \Omega$

$$E \left[h^2 \{ \sigma^2(h, W) + (\mu(h, W) - \theta)^2 \} \right] \quad (\text{III-2})$$

ou

$$E \left[\sigma^2(h, W) + (\mu(h, W) - \theta)^2 \right] \quad (\text{III-2}')$$

Suponhamos agora que $\hat{H}(n) \in W \in \mathbb{R}^p$, e

$$\Phi(\hat{H}(n), W) = W^T \hat{H}(n) \quad (\text{III-3})$$

Chamado

$$\bar{H}(h) = E[\hat{H}(n) / h(n) = h]$$

$$M(h) = \bar{H}(h) / h$$

$$\Sigma(h) = \text{Cov}[\hat{H}(n) / h(n) = h]$$

$$S(h) = C(h) / h^2$$

teremos

$$\mu(h, W) = W^T M(h) \quad (\text{III-4})$$

$$\sigma^2(h, W) = W^T S(h) W \quad (\text{III-5})$$

que podem ser substituídos em (III-2) e (III-2') .

Sendo $\bar{S} = E[S(h)]$

$$\bar{M} = E[M(h)]$$

$$R = E[M(h)M^T(h)]$$

Chamado

$$\bar{H}(h) = E[\hat{H}(n) / h(n) = h]$$

$$M(h) = \bar{H}(h) / h$$

$$\Sigma(h) = \text{Cov}[\hat{H}(n) / h(n) = h]$$

$$S(h) = C(h) / h^2$$

teremos

$$\mu(h, W) = W^T M(h) \quad (\text{III-4})$$

$$\sigma^2(h, W) = W^T S(h) W \quad (\text{III-5})$$

que podem ser substituídos em (III-2) e (III-2').

sendo $\bar{S} = E[S(h)]$

$$\bar{M} = E[M(h)]$$

$$\bar{R} = E[M(h)M^T(h)]$$

minimizar (III-2'), equivale então a minimizar a forma quadrática

$$W^T (\bar{S} + \bar{R}) W - 2\theta W^T \bar{M} + \theta^2 \quad (\text{III-6})$$

o que ocorre para

$$W(\theta) = \theta (\bar{S} + \bar{R})^{-1} \bar{M} \quad (\text{III-7})$$

Se o critério for o de minimizar (III-2), a forma da solução será a mesma de (III-7), mas

$$\bar{R} = E[\bar{H}(h)\bar{H}^T(h)]$$

$$\bar{S} = E[\Sigma(h)]$$

$$\bar{M} = E[h\bar{H}(h)] .$$

A solução \hat{e} , assim, parametrizada por θ . A direção de $W(\theta)$ é totalmente determinada pela lei de probabilidade de $(h, \hat{H}(h))$, e θ influencia apenas a norma.

$W(\theta)$ descreve pois uma raia ótima em \mathbb{R}^P .

C A P Í T U L O I V

CONSTRUÇÃO DE ESQUEMAS ADAPTATIVOS

4.1 - Introdução

Continuaremos empregando o termo "busca Heurística" no sentido de ARAUJO³ e HART ET AL¹⁵.

Suponhamos um problema, passível de representação via um grafo de estados, e sua solução equivalha à determinação de um caminho mínimo entre dois conjuntos de seus vértices, S a origem e T o destino (*).

Seja $G_s(S, \Gamma)$ este grafo, orientado, onde Γ é operador sucessor. A cada arco (n, m) de G_s está associado um custo $C_{n, m}$. Adotando o algoritmo A*, podemos determinar o menor caminho de S a T. Este problema é enfocado na forma convencional, caso os valores dos custos for a única informação disponível.

Suponhamos contudo que temos informações adicionais de caráter estocástico, que permitem formular uma melhor heurística. Podemos então recorrer a métodos de Estimação e aproximação.

(*) - A partir do Capítulo IV, vale a seguinte notação:

'^' sobre uma letra simboliza variável aleatória.

Letras maiúsculas, sublinhadas com uma barra, simbolizam vetores.

Letras maiúsculas, sublinhadas com duas barras, simbolizam matrizes.

Letras encimadas com til '~', indicam estimativa.

Barra '-' encimando letras, indica média (valor esperado).

'*' junto a letras indica ótimo (valor ótimo).

mação Estocástica para otimizar a heurística. Prosseguiremos dentro do Último aspecto.

4.2 - Definições - Formulação do Problema - Limitações

A formulação do problema segue a de ARAUJO³ e alguns dos tópicos tem sua base em ARAUJO⁵.

Seja \underline{W} um vetor de pesos,

$\hat{\underline{H}}(n)$ um vetor de medições sobre o nó n ,

$h(n)$ da forma $\Phi(\underline{W}, \hat{\underline{H}}(n))$, de modo a otimizar A^* .

Especificamente, para D sucessores de n ,

$$\hat{h}(n) = \sum_{i=1}^D w_i \hat{h}_i(n) = \underline{W}^T \hat{\underline{H}}(n) \quad (\text{IV-1})$$

Assumimos também que $\hat{\underline{H}}(n)$ é proporcional a $h(n)$, ou seja, sendo $\hat{\underline{A}}(n)$ um vetor aleatório com

$$\begin{aligned} E \left[\hat{\underline{A}}(n) \right] &= \underline{A} \quad e \\ \text{Cov} \left[\hat{\underline{A}}(n) \right] &= \underline{\Sigma}, \quad \text{então} \end{aligned} \quad (\text{IV-2})$$

$$\hat{\underline{H}}(n) = h(n) \cdot \hat{\underline{A}}(n)$$

e a matriz de autocorrelação

$$\underline{\underline{S}} = E \left[\hat{\underline{A}}(n) \cdot \hat{\underline{A}}(n)^T \right] = \underline{\Sigma} + \underline{A} \cdot \underline{A}^T \quad (\text{IV-3})$$

Neste caso, o problema

$$\text{Min } E \left[(\hat{h}(n) - \theta h(n))^2 \right] \quad (\text{IV-4})$$

ou

$$\text{Min } E \left[\left(\frac{1}{h(n)} \underline{W}^T \hat{\underline{H}}(n) - \theta \right)^2 \right] \quad (\text{IV-4'})$$

tem uma solução dada por

$$\underline{W}_* = \theta \underline{S}^{-1} \bar{\underline{A}} \quad (\text{IV-5})$$

sendo o parâmetro θ decidido por fora. O problema é que não conhecemos $\bar{\underline{A}}$ nem \underline{S} , nem tampouco partimos de \underline{W}_* .

Para contornar o impasse, vamos montar um esquema adaptativo que convirja, em sentido estocástico, para \underline{W}_* .

Em outras palavras, queremos determinar uma sequência \underline{W}_i ($i = 0, 1, \dots, t$), tal que

$$\text{a) } \forall t > 0, \underline{W}_t = f_t(\underline{W}_{t-1}, \dots, \underline{W}_0)$$

$$\text{b) } \{ \underline{W}_t \} \text{ converge (em algum sentido) para } \underline{W}_*.$$

Para satisfazer estes dois quesitos, precisamos obter \underline{A} e \underline{S} que convirjam (em algum sentido) para os verdadeiros valores.

Para resolver este problema, recorreremos a aproximação estocástica por ser esta um instrumento simples e de fácil implementação.

4.3 - Esquema Básico de Aproximação Estocástica

Suponhamos que α_t e β_t ($t = 0, \dots, \infty$) sejam duas sequências de números reais positivos, tais que:

$$i) \lim_{n \rightarrow \infty} \sum_{t=0}^n \alpha_t = \lim_{n \rightarrow \infty} \sum_{t=0}^n \beta_t = \infty \quad (\text{IV-6})$$

$$ii) \lim_{n \rightarrow \infty} \sum_{t=0}^n \alpha_t^2 < \infty; \lim_{n \rightarrow \infty} \sum_{t=0}^n \beta_t^2 < \infty \quad (\text{IV-6'})$$

$$iii) \lim_{n \rightarrow \infty} \alpha_n = 0 \quad ; \quad \lim_{n \rightarrow \infty} \beta_n = 0 \quad (\text{IV-6''})$$

Suponhamos também que, a cada iteração t , dispomos de estimativas $\tilde{\underline{A}}_t$ e $\tilde{\underline{S}}_t$, satisfazendo

$$E \left[\tilde{\underline{A}}_t \right] = \underline{A} \quad e \quad E \left[\tilde{\underline{S}}_t \right] = \underline{S}.$$

Então, se \underline{A}_0 e \underline{S}_0 são estimativas arbitrárias de \underline{A} e \underline{S} , as sequências \underline{A}_t e \underline{S}_t , definidas como:

$$\underline{A}_{t+1} = \underline{A}_t + \alpha_t (\tilde{\underline{A}}_{t+1} - \underline{A}_t) \quad , \quad t \geq 0 \quad (\text{IV-7})$$

$$\underline{S}_{t+1} = \underline{S}_t + \beta_t (\tilde{\underline{S}}_{t+1} - \underline{S}_t) \quad , \quad t \geq 0 \quad (\text{IV-7'})$$

convergem com probabilidade um para \underline{A} e \underline{S} , pois, basta observar que (IV-7) e (IV-7') satisfazem as condições do teorema de Dvoretzky (vide anexo 1).

Resta obter uma forma para $\tilde{\underline{A}}_t$ e $\tilde{\underline{S}}_t$, o que impõe maiores hipóteses sobre o grafo de busca, ou seja, um modelo para o grafo.

4.4 - Modelo para o Grafo de Busca

No que se segue, restringir-nos-emos a arborescências, ou seja, faremos tal aproximação para G_S .

A figura IV-1, mostra a expansão do nó n na iteração t , com B_t sucessores, sendo C_i ($i=1, \dots, B_t$) os custos atribuídos a seus arcos.

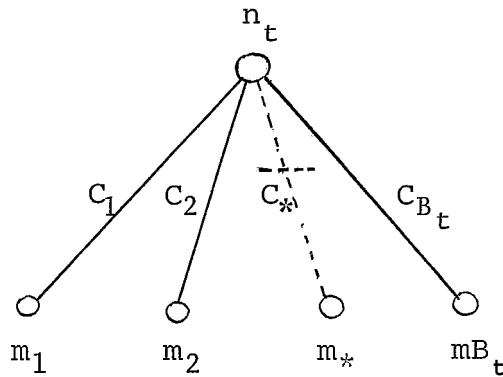


Fig : IV-1

Seja m_* o sucessor ótimo de n_t , isto é $f(m_*) = f(n_t)$. Para todo $m \in \{n_t - \{m_*\}\}$, definimos $\hat{\delta}_m$ como uma variável aleatória não negativa, ou seja, por hipótese

$$\hat{\delta}_m = f(m) - f(n_t) \geq 0$$

Um caso à parte é $\hat{\delta}_m = \delta > 0$, com distribuição constante ou determinística.

Observando que não conhecemos o valor de δ a priori, precisamos ainda estimá-lo.

4.5 - Esquema de Aproximação Estocástica para o caso

$$\hat{\delta}_m = \delta$$

Além das sequências $\{A_t\}$, $\{S_t\}$, $\{\alpha_t\}$ e $\{\beta_t\}$, pre

consideraremos de $\{\delta_t\}$ e $\{\gamma_t\}$ para estimas 6. A seqüência de números reais γ_t obedece a (IV-6), (IV-6') e (IV-6'').

Se conhecêssemos m_* , o problema estaria resolvido. Como não o conhecemos, podemos recorrer à Probabilidade total e escrever

$$\tilde{\delta}_t = \sum_{i=1}^{B_t} \tilde{\delta}_{t/i} \cdot P_i \quad (\text{IV-8})$$

onde: P_i = probabilidade de que $m_i = m_*$ (m_* é o nó Ótimo)

$\tilde{\delta}_{t/i}$ = estimador de δ supondo que $m_i = m_*$.

Análogamente

$$\tilde{A}_t = \sum_{i=1}^{B_t} \tilde{A}_{t/i} \cdot P_i \quad (\text{IV-9})$$

e

$$\tilde{S}_t = \sum_{i=1}^{B_t} \tilde{S}_{t/i} \cdot P_i \quad (\text{IV-10})$$

O problema recai, então, em calcular

i) $\tilde{\delta}_{t/i}$, $\tilde{A}_{t/i}$, $\tilde{S}_{t/i}$

ii) P_i

Para isso, temos de fazer hipóteses sobre a distribuição de A , (donde sobre $\hat{h}(n)$).

Temos dois casos a considerar

a) $m_i = m_*$, o que implica em $f(m_i) = f(n_t)$

b) $m_i \neq m_*$, o que implica em $\delta_i = \delta = f(m_i) - f(n_t) > 0$.

Para o caso (a), lembramos que não temos os valores de f . Para (b), podemos fazer

$$\tilde{\delta}_{t/i} = \frac{1}{B_t - 1} \sum_{j \neq i_*} (\tilde{f}_t(m_j) - f_t(n_t)). \quad (\text{IV-11})$$

Se \tilde{f}_t estima f , podemos escrever, $\tilde{f}_t(n) = \hat{g}(n) + \tilde{h}_t(n)$,

Como trabalhamos com arborescências, $\hat{g}(n) = g(n)$. A estimativa $\tilde{h}_t(n)$, levando em conta (IV-1) e (IV-2), pode ser obtida da seguinte observação:

$$E \left[\underline{W}^T \hat{H}(n) \right] = \theta \underline{A}^T \underline{S}^{-1} E \left[\hat{H}(n) \right] = h(n) \theta \underline{A}^T \underline{S}^{-1} \underline{A}$$

Donde, se $\delta = (\underline{A}^T \underline{S}^{-1} \underline{A})^{-1}$, temos $E \left[\hat{h}(n) \right] = h(n)$, isto é, usamos

$$\tilde{h}_t(n) = \frac{1}{\underline{A}_{t-1}^T \underline{S}_{t-1}^{-1} \underline{A}_{t-1}} \underline{z}_{t-1} \underline{S}_{t-1}^{-1} \hat{H}(n)$$

ou ainda

$$\tilde{h}_t(n) = \tilde{\underline{W}}_t^T \hat{H}(n) \quad \text{com} \quad \tilde{\underline{W}}_t^T = \frac{\underline{H}_{t-1}^T \underline{S}_{t-1}^{-1}}{\underline{A}_{t-1}^T \underline{S}_{t-1}^{-1} \underline{A}_{t-1}}$$

Substituindo os \tilde{f}_t em (IV-11) e simplificando, temos

$$\begin{aligned} \tilde{\delta}_{t/i} &= \frac{1}{B_t - 1} \sum_{j \neq i} (\hat{g}(n_t) + C_j + \tilde{h}_t(m_j) - g(n_t) - \tilde{h}(n_t)) = \\ &= \frac{1}{B_t - 1} \sum_{j \neq i} (C_j + \tilde{h}(m_j) - h(n_t)) \end{aligned}$$

$$\tilde{\delta}_{t/i} = \frac{1}{B_t - 1} \sum_{j \neq i} C_j + \tilde{W}_t^T \cdot \frac{1}{B_t - 1} \left[\sum_{j \neq i} \hat{H}(m_j) - \hat{H}(n_t) \right] \quad (\text{IV-12})$$

Substituindo $\tilde{\delta}_{t/i}$ em (IV-8), temos

$$\begin{aligned} \tilde{\delta}_t &= \sum_{i=1}^{B_t} P_i \left[\frac{1}{B_t - 1} \sum_{j \neq i} C_j + \tilde{W}_t^T \left\{ \frac{1}{B_t - 1} \sum_{j \neq i} \hat{H}(m_j) - \hat{H}(n_t) \right\} \right] = \\ &= \frac{1}{B_t - 1} \sum_{i=1}^{B_t} \left\{ C_i (1 - P_i) + \tilde{W}_t^T (1 - P_i) \hat{H}(m_i) \right\} - \tilde{W}_t^T \hat{H}(n_t) = \\ &= \tilde{W}_t^T \left\{ \frac{1}{B_t - 1} \sum_{i=1}^{B_t} (1 - P_i) \hat{H}(m_i) \right\} + \frac{1}{B_t - 1} \sum_{i=1}^{B_t} C_i (1 - P_i) - \tilde{W}_t^T \hat{H}(n_t) \end{aligned}$$

ou seja

$$\tilde{\delta}_t = \tilde{W}_t^T \left\{ \frac{1}{B_t - 1} \sum_{i=1}^{B_t} (1 - P_i) \hat{H}(m_i) - \hat{H}(n_t) \right\} + \frac{1}{B_t - 1} \sum_{i=1}^{B_t} C_i (1 - P_i) \quad (\text{IV-13})$$

Donde, podemos calcular

$$\delta_t = \delta_{t-1} + \gamma_t (\tilde{\delta}_t - \delta_{t-1}) \quad (\text{IV-14})$$

onde, γ_t é um ganho com as mesmas propriedades de α_t e β_t .

4.6 - Cálculo de \underline{A}_t

Mantendo para $\tilde{\underline{A}}_{t/i}$ a abordagem adotada para $\tilde{\delta}_{t/i}$, podemos escrever

$$\tilde{\underline{A}}_{t/i} = \frac{1}{B_t} \left\{ \sum_{j \neq i} \frac{\hat{H}(m_j)}{h(n_t) - C_j + \delta} + \frac{\hat{H}(m_i)}{h(n_t) - C_i} \right\} \quad (\text{IV-15})$$

Como não temos $h(n_t)$ e δ , podemos substituí-los por δ_t e por $\tilde{h}(n_t) = \tilde{W}_t^T \cdot \hat{H}(n_t)$ respectivamente. Donde, substituindo $\underline{A}_{t/i}$ em $\tilde{\underline{A}}_t$, temos:

$$\tilde{\underline{A}}_t = \sum_{i=1}^{B_t} P_i \tilde{\underline{A}}_{t/i}$$

$$\tilde{\underline{A}}_t = \frac{1}{B_t} \left\{ \sum_{i=1}^{B_t} \left(\frac{P_i}{\tilde{h}_t(n_t) - C_i} + \frac{1 - P_i}{\tilde{h}_t(n_t) - C_i + \delta_t} \right) \hat{H}(m_i) \right\} \quad (\text{IV-16})$$

e

$$\underline{A}_t = \underline{A}_{t-1} + \alpha_t (\tilde{\underline{A}}_t - \underline{A}_{t-1}) \quad (\text{IV-17})$$

4.7 - Cálculo de \underline{S}_t

De modo análogo a (IV-15), podemos escrever

$$\tilde{\underline{S}}_{t/i} = \frac{1}{B_t} \left\{ \sum_{j \neq i} \frac{\hat{H}(m_j) \hat{H}^T(m_j)}{[h(n_t) - C_j + \delta]^2} + \frac{\hat{H}(m_i) \hat{H}^T(m_i)}{[h(n_t) - C_i]^2} \right\} \quad (\text{IV-18})$$

e

$$\tilde{\underline{S}}_t = \frac{1}{B_t} \sum_{i=1}^{B_t} \left[\frac{P_i}{[\tilde{h}(n_t) - C_i]^2} + \frac{1 - P_i}{[\tilde{h}(n_t) - C_i + \delta_t]^2} \right] \hat{H}(m_i) \hat{H}^T(m_i) \quad (\text{IV-19})$$

e

$$\underline{S}_t = \underline{S}_{t-1} + \beta_t (\tilde{\underline{S}}_t - \underline{S}_{t-1}) \quad (\text{IV-20})$$

4.8 - Estimação dos P_i

Seja P_i = probabilidade de que $m_i = m_*$. Vamos estimá-la por máxima verossimilhança.

O teorema da probabilidade total dá, para a ocorrência dos B h_i ,

$$\Pr(\hat{h}_1, \dots, \hat{h}_B) = \sum_{i=1}^B P_i \cdot \Pr(\hat{h}_1, \dots, \hat{h}_B / m_i = m_*).$$

Como por hipótese

$$h(n) = h(n)W^T \hat{A}(n) = h(n)\hat{\alpha}(n),$$

onde os $\hat{\alpha}$ são independentes entre si,

$$\begin{aligned} \Pr(\hat{h}_1, \dots, \hat{h}_B) &= \sum_{i=1}^B P_i \cdot f(\hat{h}_i / m_i = m_*) \cdot \prod_{j \neq i} f(\hat{h}_j / m_j \neq m_*) = \\ &= \left[\prod_{j=1}^B f(\hat{h}_j / m_j \neq m_*) \right] \cdot \left[\sum_{i=1}^B P_i \frac{f(\hat{h}_i / m_i = m_*)}{f(\hat{h}_i / m_i \neq m_*)} \right] \end{aligned} \quad (\text{IV-21})$$

sendo $f(\hat{h}_j / (\dots))$ densidades condicionais de probabilidade de \hat{h}_j dado (\dots) . A razão

$$L_i = \frac{f(\hat{h}_i / m_i = m_*)}{f(\hat{h}_i / m_i \neq m_*)}$$

é chamada "razão de verossimilhança".

Teorema 4.1 : $\Pr(\hat{h}_1, \dots, \hat{h}_B)$ é máxima de $\sum_{i=1}^B P_i L_i$ o for.

Prova: Imediata, levando-se em conta que o primeiro fator de

(IV-21) é não negativo.

Teorema 4.2 : Se $L_i = \max\{L_j, (j=1, \dots, B)\}$ é único, o conjunto $\{P_i, \dots, P_B\}$, que torna $\sum_{i=1}^B P_i L_i$ máximo, é:

$$\begin{cases} P_i = 1 & \text{para o 'i' cujo } L_i \text{ é máximo} \\ P_j = 0 & \text{para } j \neq i \end{cases}$$

Prova: (Ver anexo 1)

Teorema 4.3: Se $\{i_1, \dots, i_k\}$ maximizam L_i então, $P_{i_1} = \dots = P_{i_k} = \frac{1}{k}$ e $P_j = 0$ para $j \notin \{i_1, \dots, i_k\}$.

Prova: Semelhante a do teorema 4.2.

Observemos que, se o máximo é Único, as expressões de $\tilde{\delta}_t$ (IV-13), \tilde{A}_t (IV-16), \tilde{S}_t (IV-19) ficam simplificadas. Assim, seja i_* o índice que maximiza L_i . Então, $\tilde{\delta}_t$ passa a

$$\tilde{\delta}_t = \frac{W_t^T}{B_t - 1} \left\{ \frac{1}{\sum_{i \neq i_*} \hat{H}(m_i) - \hat{H}(n_t)} \right\} + \frac{1}{B_t - 1} \sum_{i \neq i_*} C_i \quad (\text{IV-22})$$

\tilde{A}_t passa a

$$\tilde{A}_t = \frac{1}{B_t} \left\{ \frac{\hat{H}(m_{i_*})}{\tilde{h}_t(n_t) - C_{i_*}} + \sum_{j \neq i_*} \frac{\hat{H}(m_j)}{\tilde{h}_t(n_t) - C_j + \delta_t} \right\} \quad (\text{IV-23})$$

\tilde{S}_t passa a

$$\tilde{S}_t = \frac{1}{B_t} \left\{ \frac{\hat{H}(m_{i_*}) \hat{H}^T(m_{i_*})}{[\tilde{h}_t(n_t) - C_{i_*}]^2} + \sum_{j \neq i_*} \frac{\hat{H}(m_j) \hat{H}^T(m_j)}{[\tilde{h}_t(n_t) - C_j + \delta_t]^2} \right\} \quad (\text{IV-24})$$

4.9 - Cálculo dos L_i

Se $\hat{H}(n) = N(h(n)A, h^2(n) \Sigma)$, então

$$h(n) = \underline{W}^T \hat{H}(n) = N(h(n) \underline{W}_t^T A, h^2(n) \underline{W}_t^T \Sigma \underline{W}_t) = N(h(n) \mu_t, h^2(n) \sigma_t^2).$$

Fazendo $\hat{f}(h(n)) = \frac{1}{\sqrt{2\pi} h(n) \sigma} \exp \left[-\frac{1}{2} \frac{(\hat{h}(n) - h(n) \mu_t)^2}{h^2(n) \sigma_t^2} \right]$

e chamando $k_i = h(n_t) - C_i$, temos

$$\hat{f}(h_i / m_i = m_*) = \frac{1}{\sqrt{2\pi} \sigma k_i} \exp \left[-\frac{1}{2} \frac{(\hat{h}_i - k_i \mu_t)^2}{k_i^2 \sigma_t^2} \right] \quad (\text{IV-25})$$

$$\hat{f}(h_i / m_i \neq m_*) = \frac{1}{\sqrt{2\pi} \sigma (k_i + \delta)} \exp \left[-\frac{1}{2} \frac{(\hat{h}_i - (k_i + \delta) \mu_t)^2}{(k_i + \delta)^2 \sigma_t^2} \right] \quad (\text{IV-26})$$

De (IV-25) e (IV-26), temos

$$L_i = \left(1 + \frac{\delta}{k_i}\right) \exp \left[-\frac{1}{2\sigma_t^2} \left\{ \frac{(\hat{h}_i - k_i \mu_t)^2}{k_i^2} - \frac{(\hat{h}_i - (k_i + \delta) \mu_t)^2}{(k_i + \delta)^2} \right\} \right]$$

Logaritmando, obtemos

$$l_i = \ln(L_i) = \ln \left(1 + \frac{\delta}{k_i}\right) - \frac{1}{2\sigma_t^2} \left\{ \frac{(\hat{h}_i - k_i \mu_t)^2}{k_i^2} - \frac{(\hat{h}_i - (k_i + \delta) \mu_t)^2}{(k_i + \delta)^2} \right\} \quad (\text{IV-27})$$

Maximizar l_i corresponde a minimizar $-l_i$

$$-l_i = \frac{1}{2\sigma_t^2} \left\{ \frac{(\hat{h}_i - k_i \mu_t)^2}{k_i^2} - \frac{(\hat{h}_i - (k_i + \delta) \mu_t)^2}{(k_i + \delta)^2} - \ln \left(1 + \frac{\delta}{k_i}\right) \right\} \quad (\text{IV-28})$$

Como não conhecemos os verdadeiros valores de δ , μ_t , σ_t , e k_i , podemos substituí-los por estimativas como:

$$\delta = \delta_t$$

$$\mu_t = \underline{W}_t^T \cdot \underline{A}_t$$

$$\sigma_t^2 = \underline{W}_t^T \cdot \underline{\Sigma} \cdot \underline{W}_t$$

$$k_i = \tilde{h}(n_t) - C_i$$

4.10 - Análise Experimental

Os testes com o "Experimento Base" da figura IV-2, compreendem duas partes: A Estabilidade e a Convergência.

A Estabilidade é um teste efetuado, supondo-se que o experimento base tivesse atingido um estado de "regime". Pressupõe que os valores δ_t , \underline{A}_t e $\underline{\Sigma}_t$ sejam Ótimos no início do experimento e os fatores de ganho α_t , β_t e γ_t sejam razoavelmente pequenos, situação em que ocorre a estabilidade, caso o esquema seja realmente estável. Havendo estabilidade, o teste mostrará que os valores observados comportam-se aproximadamente como os esperados.

O teste de convergência é efetuado, começando-se com valores aleatórios para δ_t , \underline{A} , e $\underline{\Sigma}_t$, como também $\alpha_t = \beta_t = \gamma_t = 1/t$ ($t=1,2,\dots$).

Para efeitos de praticidade, vamos separar Convergência Estocástica de Potência Heurística. Limitar-nos-emos por enquanto à primeira, deixando a última junto a análise global.

Quanto ao teste de convergência, este é aplicado apenas quando a estabilidade produz resultados promissores.

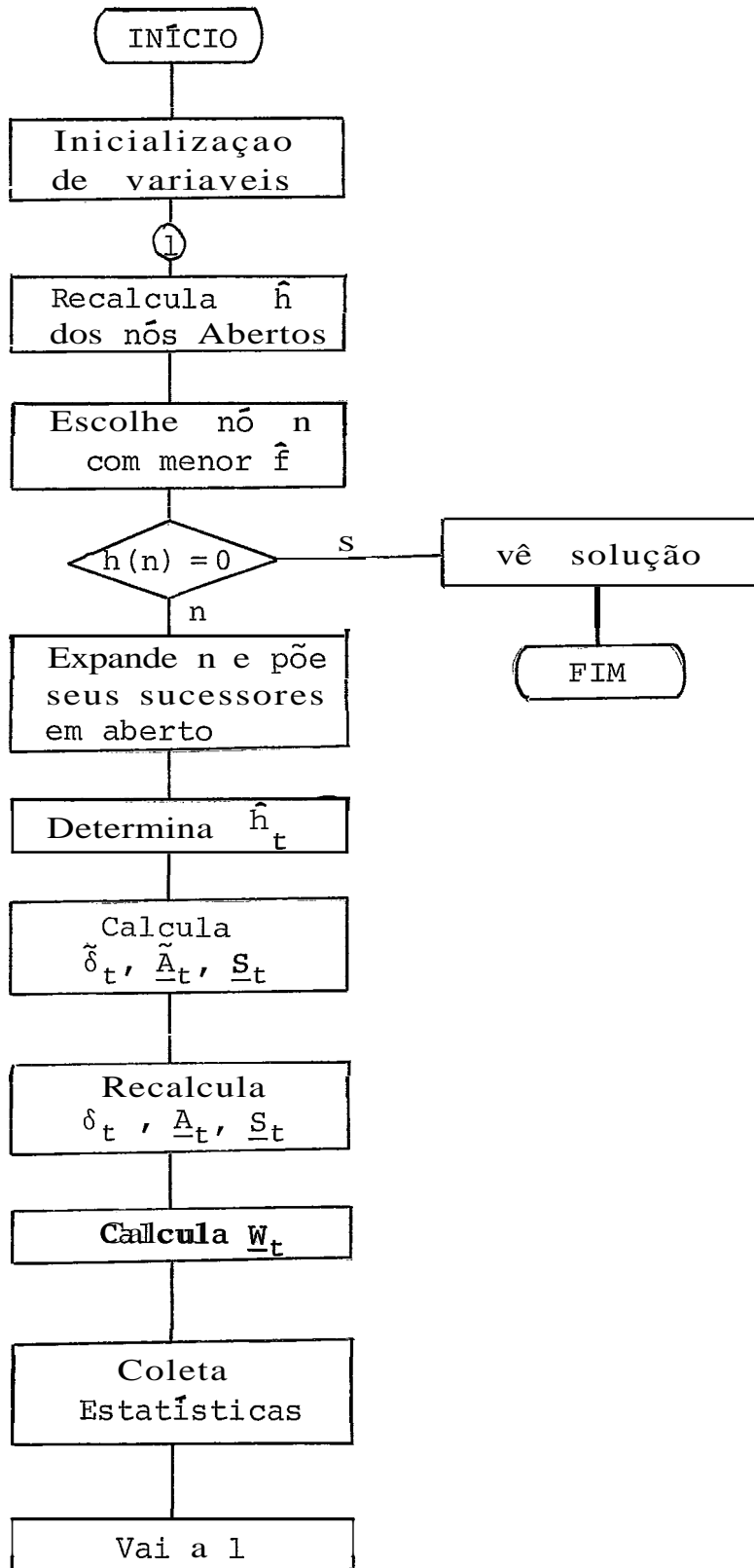


Fig. IV-2

4.11 - Resultados práticos do Esquema com (IV-22), (IV-23) e (IV-24) (*)

Os resultados práticos não tem sido promissores. Os valores de \underline{A}_t e \underline{S}_t aparecem sistematicamente com tendenciosidade acentuada e sempre para valores menores que os esperados. O teste 1 e 0 gráfico 1 (apêndice 1), dão uma idéia melhor do ocorrido.

Como os resultados tendenciosos se apresentavam como norma, novas investigações a respeito do modelo foram necessárias e uma análise mais profunda, sugeriu uma série de tentativas de melhoramento, a seguir analisados.

4.12 - Novas fórmulas para \tilde{A}_t e \tilde{S}_t

A primeira tentativa foi uma nova formulação para \tilde{A}_t e \tilde{S}_t .

$$\tilde{A}_t = \frac{1}{B_t} \left\{ \frac{\hat{H}(n_t) - \hat{H}(m_*)}{C_{i_*}} + \sum_{i \neq i_*} \frac{\hat{H}(n_t) - \hat{H}(m_i)}{C_i - \delta_t} \right\} \quad (\text{IV-29})$$

$$\tilde{S}_t = \frac{1}{B_t} \left\{ \frac{\left[\hat{H}(n_t) - \hat{H}(m_*) \right] \left[\hat{H}(n_t) - \hat{H}(m_*) \right]^T}{C_{i_*}^2} + \sum_{j \neq i_*} \frac{\left[\hat{H}(n_t) - \hat{H}(m_j) \right] \left[\hat{H}(n_t) - \hat{H}(m_j) \right]^T}{(C_i - \delta_t)^2} \right\} \quad (\text{IV-30})$$

(*) O esquema (IV-23), (IV-24) e (IV-22) ou (IV-32) é referido no apêndice 1 como "tentativa 0".

A vantagem destas fórmulas é recorrerem o mínimo a estimadores (apenas δ_t). Amparada nas vantagens, vem porém uma grande desvantagem, a instabilidade que o denominador ($C_{i^*} - \delta_t$) pode gerar quando C e δ_t forem da mesma ordem de Grandeza.

Como no experimento descrito na seção 4.11, os valores de δ_t também não foram satisfatórios, vamos tentar melhorá-lo experimentando novas formas para $\tilde{\delta}_t$.

4.13 - Outras formas para $\tilde{\delta}_t$

Consideremos novamente a figura IV-1. Devido à formulação do problema, para todo sucessor m_j de n_t , sendo m_{i^*} o sucessor considerado Ótimo, $m_{i^*} \neq m_j$, e supondo \hat{H} unidimensional (um só componente), sem dispersão, podemos escrever:

$$\frac{\hat{H}(n_t) - \hat{H}(m_{i^*})}{C_{i^*}} = \frac{\hat{H}(n_t) - \hat{H}(m_j)}{C_j - \delta_t} \quad (\text{IV-31})$$

e generalizando para \hat{H} com D componentes, temos

$$\tilde{\delta}_{t,k} = \frac{1}{B_t - 1} \sum_{j \neq i^*} \left[C_j - C_{i^*} \cdot \frac{\hat{h}_k(n_t) - \hat{h}_k(m_j)}{\hat{h}_k(n_t) - \hat{h}_k(m_{i^*})} \right]$$

onde $\hat{h}_k(n)$ é a k-ésima componente de $\hat{H}(n)$, e

$$\tilde{\delta}_t = \frac{1}{D} \sum_{k=1}^D \tilde{\delta}_{t,k} \quad (\text{IV-32})$$

Esta formulação parece bastante aceitável, pois não envolve diretamente nenhum estimador.

Submetendo o esquema (IV-32), (IV-23), (IV-24) ao computador, os resultados foram melhores que os das tentativas anteriores, já justificando seu uso.

No entanto, há um viés na divisão de duas variáveis aleatórias pois $E \left[\frac{\hat{x}}{\hat{v}} = \frac{\hat{y}}{\hat{z}} \right] \neq \frac{E(\hat{x} - \hat{y})}{E(\hat{x} - \hat{z})}$. Pela observação dos resultados experimentais, este viés não perturbou muito. (como a presente observação foi feita com o trabalho, já concluído, fica a segunda forma como sugestão para experimentos futuros).

Uma outra experiência que poderia ser feita é tentar obter $\tilde{\delta}_+$ em função de h . Contudo, a expressão (IV-33) nos daria um $\tilde{\delta}_t$

$$\tilde{\delta}_t = \frac{1}{B_t - 1} \sum_{j \neq i_*} (\tilde{h}_j - \tilde{h}_{i_*}) \quad (\text{IV-33})$$

coerente com \tilde{h} , ou seja, $\tilde{\delta}_+$ seria super ou subestimado, segundo h o fosse. Isto não tornaria propício seu uso em (IV-31).

A diferença entre os resultados apresentados por (IV-22) e (IV-32) diminui bastante quando foi imposta a restrição $\delta_+ \leq C_i$ médio dos sucessores do nó em expansão, sem alterar significativamente o índice de acertos, revelando assim uma importância secundária para 8.

Cabe então investigar melhor h .

4.14 - Nova forma de cálculo para $\tilde{h}(n)$

Seja n_t o nó em expansão e y o conjunto de nós pertencentes ao melhor caminho de n_t até a origem. Vamos ajustar uma primeira reta aos pontos y , considerando como coordenadas os va-

lores $\left[\hat{H}(1, y), g(y) \right]$; Uma segunda reta com coordenadas $\left[\hat{H}(2, y), g(y) \right]$; e assim até a última posição (m) dos vetores $\hat{H}(y)$. Se os valores do esquema forem bons, as retas ajustadas terão inclinação em torno de 135° e podem ser usadas para calcular $\tilde{h}(n)$.

Supondo as retas da forma $r(g(n)) = a_0 + a_1 g(n)$, cada uma das m retas fornecerá um valor $\tilde{h}(n) = r(g(n)) / |a_1|$. Como estimativa a ser efetivamente usada, podemos tomar a média dos m valores. Denominaremos esta forma de cálculo de \tilde{h} de "h ajustado".

Embora o processo pareça dispendioso em relação ao volume de cálculos, não o será todavia, se forem armazenados alguns valores do cálculo em cada ajustamento efetuado, pois os dados do ajustamento são de caráter cumulativo.

As retas ajustadas fornecem elementos adicionais (por exemplo: inclinação), que permitem ao esquema munir-se de defesas contra estimativas \tilde{h} excessivamente distantes de h.

A principal vantagem de usar \tilde{h} ajustado é que o seu processo de cálculo não envolve diretamente os valores do esquema.

A principal desvantagem é que pode viciar-se no início com valores disparatados de \hat{H} .

Em vista disto, pode-se encarar a possibilidade de usar um \tilde{h} que mantenha certa coerência entre o \tilde{h} ajustado e um outro, obtido por exemplo através da fórmula $\tilde{h} = W^T \hat{H}$.

4.15 - Resultados Experimentais com \tilde{h} ajustado

Usando \tilde{h} ajustado, foi repetido o experimento padrão com (IV-32), (IV-23) e (IV-24). A estabilidade (teste 2), foi

bastante superior a do teste 1 (apêndice 1). A convergência (testes 3 e 4), pode ser considerado boa, uma vez que, já na iteração 30 (teste 3), \underline{A}_t tinha uma boa aproximação para \bar{A} , e \underline{S}_t foi melhorando gradativamente.

4.16 - Conclusões

Em vista da tendenciosidade que ainda aparece nos resultados (tentativa 0), vamos investigar esquemas alternativos que a minimizem, ou mesmo eliminem. Estes esquemas são denominados no capítulo V de "tentativas".

C A P Í T U L O V

CORREÇÃO DA TENDENCIOSIDADE DE \underline{A}_t E \underline{S}_t 5.1 - Introdução

Os resultados práticos do capítulo IV não negam um bom desempenho de (IV-23) e (IV-24). Porém, algumas experiências mostraram que os valores de \tilde{h} , apesar dos possíveis bons métodos em obtê-lo, podem levar estas duas expressões a uma acentuada tendenciosidade. Os valores de δ_t , situam-se em segundo plano, levando-se em conta seu menor peso em relação a h .

Apresentamos neste capítulo, algumas tentativas para diminuir, ou se possível, eliminar a tendenciosidade de \underline{A}_t e \underline{S}_t .

5.2 - Tentativa 1: Correção em $\tilde{\underline{A}}_t$ e $\tilde{\underline{S}}_t$

Como ponto de partida podemos retomar a expressão (IV-29), abandonando a parte que a inviabiliza e efetuando uma operação corretiva sobre a outra, obtendo:

$$\tilde{\underline{A}}_t = \frac{\hat{H}(n_t) - \hat{H}(m)}{C_{i*}} \cdot \text{CORREÇÃO} \quad (\text{V-1})$$

Com esta expressão, sem dúvida há perda de informação, pois os elementos considerados não ótimos são simplesmente

ignorados, o que pode pesar na velocidade de convergência . Em contrapartida esperamos uma diminuição do viés em \underline{A}_t .

5.2.1 - Cálculo da Correção para $\underline{\tilde{A}}_t$

Para calcular o fator de correção, podemos considerar

$$\underline{\tilde{A}}_t = \underline{\hat{A}}_* \cdot p + \hat{A}_0 \frac{C_* - \delta}{C_*} (1-p) \quad (V-2)$$

onde \hat{A}_0 indica que \mathbf{i} , está sendo considerado como não ótimo, e p = probabilidade de \mathbf{i} , ser Ótimo, ou seja, que a escolha foi acertada.

Então $\underline{\tilde{A}}_t = \underline{\hat{A}} \{ p + (1 - \delta / C_*) (1 - p) \}$ e

$$\underline{\tilde{A}}_t = \underline{\hat{A}} \{ 1 - (\delta / C_*) (1 - p) \}$$

Logo

$$\text{CORREÇÃO} = 1 / [1 - \delta / C_* (1 - p)] \quad (V-3)$$

5.2.2 - Cálculo de p

p é a probabilidade de \mathbf{i} , ser Ótimo. Então

$$p = \text{pr} [C_* + \underline{w}_t^T \underline{H}(m_*) \leq C_i + \underline{w}_t^T \underline{\hat{H}}(m_i), \forall \mathbf{i} \neq *] .$$

Considerando os C_i 's iguais, podemos eliminá-los.

Fazendo $\hat{\underline{H}}_* = h_* \hat{\underline{A}}_*$

$$\hat{\underline{H}}_i = h_i \hat{\underline{A}}_i = (h_* + \delta) \hat{\underline{A}}_i$$

temos
$$p = \text{pr} \left[h_* \underline{W}_t^T \hat{\underline{A}}_* \leq (h_* + \delta) \underline{W}_t^T \hat{\underline{A}}_i \right]$$

Dividindo por h_* , temos

$$p = \text{pr} \left[\underline{W}_t^T \hat{\underline{A}}_* \leq (1 + \delta/h_*) \underline{W}_t^T \hat{\underline{A}}_i \right]$$

fazendo $\mu = \underline{W}_t^T \bar{\underline{A}}$

$$\sigma^2 = \underline{W}_t^T \underline{\Sigma} \underline{W}_t$$

$$\alpha = 1 + \delta/h_*$$

$$y = \underline{W}_t^T \underline{A}_*$$

$$x = \underline{W}_t^T \underline{A}_i$$

temos
$$p = \text{pr}(y \leq \alpha x)$$

Supondo $y, x \in N(\mu, \sigma^2)$

$$p = \text{pr}(y \leq \alpha x) = E_x[\text{pr}\{y \leq \alpha x / x = X\}] =$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left(\int_{-\infty}^{\alpha x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} d_y d_x \right)$$

fazendo $w = (y - \mu)/\sigma$, então, $y = w\sigma + \mu \leq ax$, e

$$p = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \int_{-\infty}^{(ax-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}w^2} dw dx$$

fazendo $z = (x - \mu)/\sigma$, dá $x = z\sigma + \mu$ e

$$p = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \int_{-\infty}^{[\alpha(z\sigma + \mu)]/\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}w^2} dw dz$$

ou

$$p = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \int_{-\infty}^{[\alpha z + \mu/\sigma(\alpha-1)]} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}w^2} dw dz$$

(V-4)

(Nesta forma p é facilmente determinado com o auxílio da subrotina NDTR do anexo 1).

Então, p é a probabilidade de h_{i^*} ser menor que um h_i não Ótimo. Considerando independência entre sucessores, temos $p = p^{(B_t-1)}$

5.2.3 - Análise de \tilde{S}_t

Seguiremos para \tilde{S}_t a mesma linha de raciocínio que para \tilde{A}_t . Seja m o nó em expansão e seja n seu sucessor ótimo.

Então para o $n \neq m$ podemos escrever:

$$\begin{aligned}\Phi &= \frac{1}{C_*^2} \left[\underline{\hat{H}}(m) - \underline{\hat{H}}(n) \right] \cdot \left[\underline{\hat{H}}(m) - \underline{\hat{H}}(n) \right]^T \\ &= \frac{1}{C_*^2} \left[\underline{\hat{H}}(m) \underline{\hat{H}}^T(m) + \underline{\hat{H}}(n) \underline{\hat{H}}^T(n) - \underline{\hat{H}}(m) \underline{\hat{H}}(n) - \underline{\hat{H}}(n) \underline{\hat{H}}^T(m) \right] \\ \Phi &= \frac{1}{C_*^2} \left[h^2(m) \underline{\hat{A}}_{m-m} \underline{\hat{A}}_{m-m}^T + (h(m) - C_*) \underline{\hat{A}}_{n-n} \underline{\hat{A}}_{n-n}^T - h(m) (h(m) - C_*) \underline{\hat{A}}_{m-n} \underline{\hat{A}}_{m-n}^T - \right. \\ &\quad \left. - h(m) (h(m) - C_*) \underline{\hat{A}}_{-n} \underline{\hat{A}}_{-n}^T \right]\end{aligned}$$

$$\begin{aligned}E(\Phi) &= \frac{1}{C_*^2} \left\{ h^2(m) E \left[\underline{\hat{A}}_{m-m} \underline{\hat{A}}_{m-m}^T \right] + (h(m) - C_*)^2 E \left[\underline{\hat{A}}_{-n} \underline{\hat{A}}_{-n}^T \right] - \right. \\ &\quad \left. - h(m) (h(m) - C_*) E \left[\underline{\hat{A}}_{m-n} \underline{\hat{A}}_{m-n}^T \right] - h(m) (h(m) - C_*) E \left[\underline{\hat{A}}_{-n} \underline{\hat{A}}_{-m}^T \right] \right\}\end{aligned}$$

Lembrando a independência de $\underline{\hat{A}}_{-n}$ e $\underline{\hat{A}}_{-m}$, temos

$$E(\underline{\hat{A}}_{-n} \underline{\hat{A}}_{-m}^T) = E(\underline{\hat{A}}_{-n}) \cdot E(\underline{\hat{A}}_{-m}^T) = \underline{\bar{A}}_{-n} \cdot \underline{\bar{A}}_{-m}^T$$

portanto

$$\begin{aligned}E(\Phi) &= \frac{1}{C_*^2} \left[h^2(m) \underline{\underline{S}} + (h(m) - C_*)^2 \underline{\underline{S}} - 2h(m) (h(m) - C_*) \underline{\bar{A}}_{-n} \cdot \underline{\bar{A}}_{-m}^T \right] \\ &= \frac{1}{C_*^2} \left[h^2(m) \underline{\underline{S}} + h^2(m) \underline{\underline{S}} + C_*^2 \underline{\underline{S}} - 2C_* h(m) \underline{\underline{S}} - \right.\end{aligned}$$

$$\begin{aligned}
 & - 2h^2(m) \bar{\underline{A}} \bar{\underline{A}}^T + 2C_* h(m) \bar{\underline{A}} \bar{\underline{A}}^T \Big] \\
 & = \frac{1}{C_*^2} \left[2h^2(m) (\underline{\underline{S}} - \bar{\underline{A}} \bar{\underline{A}}^T) - 2h(m) C_* (\underline{\underline{S}} - \bar{\underline{A}} \bar{\underline{A}}^T) + C_*^2 \underline{\underline{S}} \right].
 \end{aligned}$$

Como

$$\underline{\underline{\Sigma}} = \underline{\underline{S}} - \bar{\underline{A}} \bar{\underline{A}}^T, \quad \text{vem}$$

$$E(\Phi) = \underline{\underline{S}} + \frac{2h(m) (h(m) - C_*)}{C_*^2} \underline{\underline{\Sigma}}$$

$$\hat{\underline{\underline{S}}} = \Phi - \frac{2h(m) (h(m) - C_*)}{C_*^2} \underline{\underline{\Sigma}} \quad (V-5)$$

$$\tilde{\underline{\underline{S}}} = \hat{\underline{\underline{S}}}_* \cdot p + \hat{\underline{\underline{S}}}_0 \frac{C_* - \delta}{C_*} \quad (1-p) \quad (V-6)$$

Em (V-6), $\hat{\underline{\underline{S}}}_0$ implica na suposição que n não é ótimo. Para obter $\hat{\underline{\underline{S}}}_0$, trocamos em (V-5), $(h(m) - C_*)$ por $(h(m) - C_* - \delta)$ e teremos:

$$\hat{\underline{\underline{S}}}_0 = \Phi - \frac{2h(m) (h(m) - C_* + \delta)}{C_*^2} \underline{\underline{\Sigma}} \quad (V-7)$$

Substituindo (V-5) e (V-7) em (V-6), temos

$$\underline{\tilde{S}} = \left[\Phi - \frac{2h(m)(h(m)-C_*)}{C_*^2} \underline{\Sigma} \right] p + \left[\Phi - \frac{2h(m)(h(m)-C_*+\delta)}{C_*^2} \underline{\Sigma} \right] (1-p)$$

$$= \Phi - \frac{2h(m)(h(m)-C_*)}{C_*^2} \underline{\Sigma} - \frac{2h(m)\delta}{C_*^2} (1-p) \underline{\Sigma} \quad \text{ou}$$

$$\underline{\tilde{S}} = \Phi + \frac{2h(m)}{C_*^2} \left[\delta p - h(m) + C_* - \delta \right] \underline{\Sigma} \quad (V-8)$$

5.2.4 - Resultados Experimentais

Repetindo o "experimento base" com (V-1) e (V-8), teste de Estabilidade, nota-se oscilações muito bruscas no início do teste, levando a valores completamente sem sentido. Isto apesar de α , β e γ iniciais serem bem pequenos.

5.3 - Tentativa 2: Correção em \underline{A}_t e \underline{S}_t

Em vista do fracasso das fórmulas (V-1) e (V-8) em proporcionar uma estabilidade ao esquema adaptativo, vamos tentar uma nova composição. Consideremos as fórmulas (IV-23) e (IV-24).

5.3.1 - Cálculo da Correção para $\underline{\tilde{A}}_t$

Se à semelhança da tentativa 1, tomarmos apenas a primeira parte, compensando a segunda por um fator de correção, podemos então escrever:

$$\tilde{\hat{A}}_t = \frac{\hat{H}(m_*)}{\tilde{h}(n_t) - C(m_*)} \cdot \text{CORREÇÃO} \quad (\text{V-9})$$

Para calcular a CORREÇÃO, temos a definir, à semelhança da tentativa 1, as seguintes variáveis:

\hat{A}_* : \hat{A} , considerando m , como Ótimo.

\hat{A}_0 : \hat{A} , considerando m , como não Ótimo.

p = propabilidade de m , ser ótimo, ou seja, de que a escolha foi acertada; p é calculado com (V-4).

Mantendo a relação

$$\tilde{\hat{A}} = \hat{A}_* \cdot p + \hat{A}_0 (1-p) \quad (\text{V-10})$$

vamos fazer

$$\theta = \frac{\hat{H}(m_*)}{h(n_t) - C_{m_*}}$$

então

$$E(\theta) = \frac{1}{h(n_t) - C_*} \quad E(h(m_*) \hat{A}_*) = \frac{h(m_*)}{h(n_t) - C_*} \hat{A}$$

Se m , é Ótimo, então $E(\theta) = \hat{A}$.

Se m , é considerado como não ótimo, então

$$E(\theta) = \frac{h(n_t) - C_* + \delta}{h(n_t) - C_*} \hat{A}$$

Englobando os dois casos, podemos escrever:

$$E(\theta) = \hat{A}p + \frac{h(n_t) - C_* + \delta}{h(n_t) - C_*} \hat{A}(1-p) \quad (V-11)$$

ou

$$E(\theta) = \hat{A} \cdot \left[p + \frac{h(n_t) - C_* + \delta}{h(n_t) - C_*} (1-p) \right]$$

donde

$$\tilde{A}_t = \theta / \left[p + \frac{h(n_t) - C_* + \delta}{h(n_t) - C_*} (1-p) \right] \quad (V-12)$$

ou seja

$$\text{CORREÇÃO} = \frac{1}{p + \frac{h(n_t) - C_* + \delta}{h(n_t) - C_*} (1-p)} \quad (V-13)$$

Para $h(n_t)$ e δ continuam as estimativas $\tilde{h}(n_t)$ e $\tilde{\delta}_t$.

5.3.2 - Cálculo da Correção para \tilde{S}_t

Repetindo para \tilde{S}_t um raciocínio semelhante ao se guido na seção (5.2.3), temos:

$$\text{seja} \quad \tilde{S}_t = \frac{\hat{H}(m_*) \hat{H}^T(m_*)}{(h(n_t) - C_*)^2} \cdot \text{CORREÇÃO} \quad (V-14)$$

fazendo

$$\psi = \frac{\hat{H}(m_*) \hat{H}^T(m_*)}{(h(n_t) - C_*)^2}, \text{ então}$$

$$E(\psi) = \frac{1}{(h(n_t) - C_*)^2} \cdot E\left[h^2(m_*) \hat{A}_* \hat{A}_*^T\right] = \frac{h^2(m_*)}{(h(n_t) - C_*)^2} \underline{\underline{S}}$$

Caso a: quando m_* é ótimo, então

$$h(m_*) = h(n_t) - C_*$$

$$\text{logo } E(\psi) = \underline{\underline{S}}.$$

Caso b: quando m_* é considerado como não ótimo, então

$$h(m_*) = h(n_t) - C_* + \delta$$

$$\text{logo } E(\psi) = \frac{(h(n_t) - C_* + \delta)^2}{(h(n_t) - C_*)^2} \underline{\underline{S}}$$

Levando em conta os casos "a" e "b", à semelhança da expressão (V-11), temos

$$E(\psi) = p \cdot \hat{\underline{\underline{S}}} + (1-p) \frac{(h(n_t) - C_* + \delta)^2}{(h(n_t) - C_*)^2} \underline{\underline{S}} \quad (\text{V-15a})$$

ou seja

$$\tilde{\underline{\underline{S}}}_t = \psi / (p + (1-p) \left(1 + \frac{\delta}{h(n_t) - C_*}\right)^2)$$

e o fator de correção em ψ , será

$$\text{CORREÇÃO} = \frac{1}{p + (1-p) \left(1 + \frac{\delta}{h(n_t) - C_*}\right)^2} \quad (\text{V-15b})$$

5.3.3 - Comportamento Experimental

Observando os resultados dos testes de Estabilidade (testes 5, 6, 7), podemos classificá-los como bastante bons.

$\underline{\underline{A}}_t$ e $\underline{\underline{S}}_t$ mantiveram uma proporcionalidade em suas variações e $\underline{\underline{W}}_t$ manteve-se razoavelmente perto de $\underline{\underline{W}}_*$.

Por outra, os testes de convergência (testes 8, 9, 10, 11), feitos com variação na forma de δ_t e nos valores de $\bar{\underline{\underline{A}}}$, deram resultados satisfatórios.

5.4 - Tentativa 3: $\tilde{\underline{\underline{A}}}_t$ e $\tilde{\underline{\underline{S}}}_t$ sem C_i e δ

5.4.1 - Seria possível suprimir C_i e δ ?

Observamos que, se substituirmos nas expressões de $\tilde{\underline{\underline{A}}}_t$ e $\tilde{\underline{\underline{S}}}_t$ em (IV-23) e (IV-24), os termos $[h, (n_t) - C_{i*}]$ e $[\tilde{h}(n_t) - C_j + S_t]$ por $\tilde{h}(m_{i*})$ e $\tilde{h}(m_j)$ respectivamente, as expressões se

simplificam para

$$\hat{\underline{A}}_t = \frac{1}{B_t} \sum_{j=1}^{B_t} \frac{\hat{H}(m_j)}{\tilde{h}(m_j)} \quad (V-16)$$

$$\tilde{\underline{S}}_t = \frac{1}{B_t} \sum_{j=1}^{B_t} \frac{\hat{H}(m_j) \hat{H}^T(m_j)}{\tilde{h}^2(m_j)} \quad (V-17)$$

e não há necessidade de calcular δ_t . Seria possível suprimir C_i e δ ?

Verificando novamente o caso unidimensional, sem dispersão, temos

$$\tilde{a}_t = \frac{1}{B_t} \sum_{j=1}^{B_t} \frac{ah(m_j)}{\frac{a}{a_t} h(m_j)} = a_t$$

o que mostra que não há ganho de informação, ou seja, o esquema seria inócuo. É portanto necessário utilizar a informação de C_i e δ_t , ou diversificar as fontes de informação, o que se consegue com o uso de \tilde{h}_t ajustado.

5.4.2 - Mudança de nó Referencial

Como (V-16) e (V-17) são um pouco inconsistentes com a formulação do problema, vamos substituir o nó em referência m_j por n_t , conforme a figura IV-1, e temos:

$$\tilde{\underline{A}}_t = \frac{\hat{\underline{H}}(n_t)}{h(n_t)} \quad (V-18)$$

$$\tilde{\underline{S}}_t = \frac{\hat{\underline{H}}(n_t) \hat{\underline{H}}^T(n_t)}{\tilde{h}^2(n_t)} \quad (V-19)$$

As principais vantagens desta nova formulação são:

- a) Usa apenas um valor estimado ($\tilde{h}(n_t)$).
- b) Requer menos cálculos que os anteriores.
- c) Não usa C_i e δ_t .

A principal desvantagem é de que coleta toda sua informação nos nós fechados, os quais, de modo geral, representam uma fração razoavelmente pequena em relação aos abertos.

5.4.3 - Resultados Experimentais

Olhando os resultados dos testes de Estabilidade no anexo 1, (testes 12, 13, 14), com variações em δ_t e $\tilde{\underline{A}}$, não há

razão para não considerá-los como bons.

Quanto aos testes de convergência (testes 15, 16, 17 e 18), feitos novamente com variações em \hat{c}_t e \hat{A} , agradam tanto na aceleração da convergência como nos valores obtidos.

5.5 - Tentativa 4: Uso de probabilidade

Voltando a aproveitar a probabilidade p , calculada na seção 5.2.2, podemos sugerir mais uma forma de calcular \hat{A}_t e \hat{S}_t , semelhante a (IV-16) e (IV-19).

$$\hat{A}_t = \frac{\hat{H}(m_*)}{\hat{h}(n_t) - C_{m_*}} \cdot p + \frac{\hat{H}(m_*)}{\hat{h}(n_t) - C_{m_*} + \delta_t} \cdot (1-p) \quad (V-20)$$

$$\hat{S}_t = \frac{\hat{H}(m_*) \hat{H}^T(m_*)}{(\hat{h}(n_t) - C_{m_*})^2} \cdot p + \frac{\hat{H}(m_*) \hat{H}^T(m_*)}{(\hat{h}(n_t) - C_{m_*} + \delta_t)^2} \cdot (1-p) \quad (V-21)$$

A massa de informação arregimentada nesta tentativa é superior à da tentativa 3, situando-se no mesmo nível da tentativa 2, da qual parece ser um complemento ou variação.

5.5.1 - Resultados Experimentais

Efetuada os testes de Estabilidade (testes 19 e 20), vê-se que os resultados não negam um bom desempenho ao esquema.

Os resultados dos testes de convergência (testes 21 e 22), são satisfatórios. Nos resultados do teste 21, os va

lores de \underline{A}_t e \underline{S}_t aparecem bastante deturpados em relação às respectivas médias, devido ao fato de a busca ter-se precipitado em profundidade. Cabe para este caso outra consideração, feita no capítulo seguinte,

O teste 22 é um paradigma de outros testes efetuados, retratando os níveis de precisão atingidos.

5.6 - Conclusão

Excluindo a tentativa 1, as outras, 0, 2, 3 e 4 apresentaram resultados em geral bons e até muito bons, podendo, portanto, serem testadas em aplicações práticas.

C A P Í T U L O VI

DISCUSSÃO DE ASPECTOS DOS ESQUEMAS PROPOSTOS
E NOVAS PERSPECTIVAS

Veremos no decorrer deste capítulo, tentativas de solução para alguns problemas que o esquema apresenta, comentários sobre experimentos e tópicos: \underline{S}_t matriz escalar, \underline{C}_t fixado a priori, \underline{S}_t através de \underline{A}_t e \underline{S}_t , influência de θ , maior dimensão para \underline{A}_t e \underline{S}_t , o menor caminho, Otimização da aplicação do esquema, Sensibilidade do esquema e Aspectos da árvore de busca.

6.1 - Árvore de busca com Expansão Dirigida

Anteriormente, salvo menção explícita em contrário, fizemos os experimentos usando A com função de mérito $\hat{f} = g + \hat{h}$. Nesta forma, a árvore de busca é expandida livremente, podendo ser selecionado qualquer nó para ser fechado, independente da profundidade em que este se situa.

Esta forma de expansão dificulta comparar quantitativamente entre si dois esquemas ou duas heurísticas, devido ao fato de em muitos casos não se chegar à solução.

*

Para contornar este problema, propomos A com esquema de poda, chamado "Busca Dirigida" e que se comporta a meio termo entre uma busca horizontal e uma busca "vertical em faixa". A idéia é, poder atingir soluções próximas ao Ótimo com maior frequência e num tempo razoável.

Nos experimentos práticos foi adotado o seguinte procedimento:

Para $(g=1, \dots, k)$, o número de nós expandidos na profundidade g é menor ou igual a g , sendo expandidos o maior número de nós possíveis dentro deste Limite. Para $g > k$, o número de nós expandidos nas diversas profundidades é k . Os nós em Aberto são ordenados segundo \hat{h} .

Nos experimentos efetuados usamos $K = 10$.

6.2 - Aproximação de \underline{S}_t por Matrizes Escalares

Suponhamos que \underline{S} é matriz escalar, i.e., $\underline{S} = \epsilon \underline{I}$, onde ϵ é o valor dos autovalores de \underline{S} .

Observamos que, se os autovalores de \underline{S} (matriz positiva-definida) forem iguais a λ , então os de \underline{S}^{-1} serão λ^{-1} e, então,

$$\underline{S} \underline{A} = (\underline{A} \underline{A}^T + \epsilon \underline{I}) \underline{A} = \underline{A} \|\underline{A}\|^2 + \epsilon \underline{A} = (\|\underline{A}\|^2 + \epsilon) \underline{A}$$

o que permite escrever

$$\underline{W} = \underline{S}^{-1} \underline{A} = \frac{1}{\|\underline{A}\|^2 + \epsilon}$$

Embora este caso não seja muito freqüente, podemos usar esta aproximação quando a dispersão é pequena.

Quando $\underline{S} \neq \epsilon \underline{I}$, então, definindo

$$\bar{\epsilon} = \frac{1}{n} \sum_{i=1}^n \underline{c}_i^2,$$

se

$$\bar{\epsilon} \ll \|\underline{A}\|^2, \text{ então}$$

$$\underline{W} \approx \frac{1}{\|\underline{A}\|^2 + \epsilon} \underline{A} \quad (\text{VI-L})$$

O principal interesse desta aproximação está em que, para pequenas dispersões (i.e., $\underline{\Sigma} + \underline{A} \underline{A}^T \approx \underline{A} \underline{A}^T$), $\underline{S} = \underline{\Sigma} + \underline{A} \underline{A}^T$ é quase singular e conseqüentemente \underline{S}^{-1} é instável. A aproximação por escalar permite contornar este problema.

6.2.1 - Resultados Experimentais

Foram efetuados dois tipos de testes: no primeiro a árvore de busca foi expandida segundo o algoritmo A (Expansão não Dirigida), e no segundo foi usada a Expansão Dirigida.

6.2.1.1 - Expansão não Dirigida

$$\text{Fazendo } \tilde{s}_t = \epsilon + \|\tilde{\underline{A}}_t\|^2,$$

$$s_t = s_{t-1} + \beta_t (\tilde{s}_t - s_{t-1}) \quad \text{e} \quad (\text{VI-2})$$

$$\underline{W}_t = \frac{1}{s_t} \underline{A}_t$$

foram feitos diversos experimentos e os seus resultados foram semelhantes. Os índices de acertos giraram em torno de 37%, muito baixos em relação aos outros testes do esquema. \underline{A}_t também não mostrou estabilidade na maioria dos casos.

6.2.1.2 - Expansão Dirigida

$$\text{Fazendo } \underline{W}_t = \frac{1}{\|\underline{A}_t\| + \epsilon} \underline{A}_t$$

e usando Expansão Dirigida, foram repetidos os experimentos. A grande surpresa do teste desta vez apareceu no índice de acertos 45,5%, semelhante ao caso da matriz S não escalar (43 a 46%), e no caminho subótimo obtido 44, contra 45 a 46 da matriz \underline{S} não escalar. Os valores apresentados são médias obtidas de 10 experimentos, melhor descritos no teste 23 e gráfico 23. \underline{A}_t em geral aparece superestimado.

O gráfico 23 foi feito a partir do experimento 1 (E=1) da tabela do teste 23.

6.3 - $\underline{\Sigma}$ em Separado

Calculando $\underline{\Sigma}_t = \underline{\tilde{S}}_t - \underline{A}_t \underline{A}_t^T$ e fazendo

$$\underline{\Sigma}_t = \underline{\Sigma}_{t-1} + \gamma_t (\underline{\tilde{S}}_t - \underline{\Sigma}_{t-1}),$$

podemos obter

$$\underline{S}_t = \underline{A}_t \underline{A}_t^T + \underline{\Sigma}_t \quad (\text{VI-3})$$

6.3.1 - Resultados Experimentais

Muito ruins. Como o volume de Cálculos não oferecia vantagem sobre o dos experimentos anteriores, o Experimento foi abandonado.

6.4 - Σ fixado a priori

Partindo-se de um conhecimento prévio dos possíveis valores de $\underline{\Sigma}$, poderíamos calcular \underline{S}_t a partir de \underline{A}_t e $\underline{\Sigma}$, como

$$\underline{S}_t = \underline{A}_t \underline{A}_t^T + \underline{\Sigma} \quad (\text{VI-4})$$

6.4.1 - Resultados Experimentais

Os resultados experimentais foram melhores que os da seção anterior, porém, aquém dos resultados do capítulo V. Combinar esta forma com outras sugeridas mais adiante, deve possivelmente proporcionar bons resultados, especialmente no tangente ao volume de Cálculos.

6.5 - Uso da diagonal de \underline{S}_t

Foram feitos alguns testes de convergência com

$$\underline{W}_t(i) = \underline{A}_t(i) / \underline{S}_t(i, i) \quad (\text{VI-5})$$

O índice de acertos nas decisões não foi bom e \tilde{h} apresentou valores bastante instáveis.

6.6 - Instabilidade de \underline{S}_t^{-1}

No decorrer dos testes realizados no capítulo V, a pareceram alguns casos em que \tilde{h} oscilava muito e por conseqüên
tia, embora com menos intensidade, \tilde{h} ,

Observando atentamente estes casos, notou-se que a origem desta instabilidade residia em geral na relação entre \underline{A}_t e \underline{S}_t , ou melhor precisando, a instabilidade ocorria quando $(\underline{A}_t^i)^2 > \underline{S}_t^{ii}$ para algum ou todos os i . Ocorria em particular com \underline{C} pequeno.

6.6.1 - Soluções viáveis

Como a instabilidade de \tilde{h} gera maus resultados na busca, é preciso contorná-la, e para isto, duas idéias foram pos
tas em prática.

A primeira, que serve para a maioria dos casos, é efetuar testes sobre \underline{S}_t^{ii} e $\underline{A}_t^i \underline{A}_t^{iT}$, mantendo \underline{S}_t^{ii} maior que $(\underline{A}_t^i)^2$.

O que seria mais correto, era testar se \underline{S}_t é posi
tivo definida. Caso não o seja, então, tomemos $h_n < 0$, o menor autovalor de \underline{S}_t . Fazemos então a correção

$$\underline{S}_t \leftarrow \underline{S}_t + (|\lambda_n| + \epsilon) \underline{I} ,$$

onde $\epsilon > 0$ qualquer.

i.e.,

$$\underline{\underline{S}}_t^{ii} \leftarrow \underline{\underline{S}}_t^{ii} + \varepsilon + |\lambda_n|, \quad \forall i$$

$$\underline{\underline{S}}_t^{ij} \leftarrow \underline{\underline{S}}_t^{ij}, \quad \forall j \neq i.$$

A nova matriz $\underline{\underline{S}}_t$ é positivo-definida. Não deve ser esquecido o E, caso contrário, $\underline{\underline{S}}_t$ pode ficar singular.

Como fica meio lento chamar uma subrotina de autovalores todas as vezes, parece prático, usar como alternativa (segunda idéia), recursos de programação "sense switch" ou algo parecido, afim de que, quando uma inversão é inviável, ao invés de abortar o programa, soma-se um ε à diagonal principal de $\underline{\underline{S}}_t$ e refaz-se o cálculo de $\underline{\underline{S}}_t^{-1}$. Isto parece razoável e foi implantado na subrotina de Inversão de Matrizes "INVERT".

6.6.2 - Enfoque em Aberto

Aceitando a hipótese que a instabilidade do esquema possa originar-se de um vetor \underline{N} normal a \underline{A} , então, em vez de trabalhar apenas com $\underline{\underline{A}}_t$, poder-se-ia tentar uma forma de decomposição

$$\underline{\underline{A}}_t = \rho \underline{A} + \underline{N} \tag{VI-7}$$

onde ρ é um escalar qualquer e \underline{N} um vetor Normal a \underline{A} .

Embora (VI-7) seja sugestiva, não conduziu a uma forma concreta de realização, sendo abandonada.

6.7 - Manutenção da consistência entre \underline{A}_t e \underline{S}_t

Os resultados experimentais mostraram, o que é esperado, que a busca produz melhores resultados quando $\underline{\Sigma}_t = \underline{S}_t - \underline{A}_t \underline{A}_t^T \approx \underline{\Sigma}$. Para que $\underline{\Sigma}_t \approx \underline{\Sigma}$, é preciso que \underline{A}_t e \underline{S}_t mantenham uma certa consistência entre si. Os piores resultados apareceram quando $\underline{\Sigma}_t$ não é positivo definida, e é para este caso que sugerimos a correção que segue:

6.7.1 - Uso da Diferença Máxima da Diagonal

Suponhamos que $\underline{S}_t - \underline{A}_t \underline{A}_t^T$ não seja positiva definida; supondo também que $\underline{S}_t - \underline{A}_t \underline{A}_t^T$ é diagonal, e chamando

$$m = \max\{\delta_i = (\underline{A}_t^i)^2 - s_t^{ii} / \delta_i > 0\}.$$

Se fizermos

$$\underline{S}'_t = \underline{A}_t \underline{A}_t^T + A, \quad \text{com}$$

$$\Delta_t = \underline{S}_t - \underline{A}_t \underline{A}_t^T + m \underline{I},$$

Δ_t será não-negativa definida. (Para garantir positiva definida, poderíamos usar $m' = m + \epsilon > 0$).

Para evitar vieses cumulativos podemos fazer

$$\bar{\delta} = \frac{1}{n} \sum_{i=1}^n (s_t^{ii} - (\underline{A}_t^i)^2) \quad \text{e} \quad \Delta_t \rightarrow \frac{1}{m + |\bar{\delta}|} \Delta_t.$$

6.7.2 - Resultados Experimentais

O teste experimental revelou que este método só pode ser aplicado após certo número de nós expandidos, caso contrário pode levar a deformação em \underline{S}_t , as quais o esquema não consegue recuperar posteriormente. Experimentos mostraram também que, nas condições dos testes efetuados, 20 nós expandidos antes de aplicar a correção, em geral produzem resultados satisfatórios.

Repetindo, com a correção ora sugerida, experimentos em que ocorria \underline{C}_t não positiva definida, houve sensível melhora no índice de acertos na escolha do sucessor ótimo, e, em consequência, da solução obtida. Os resultados nivelaram-se aos casos em que \underline{C}_t era positiva definida sem precisar correção.

6.7.3 - correção por linha e coluna

Uma segunda forma é efetuar uma correção multiplicativa apenas sobre a linha e coluna i de \underline{S}_t quando $(\underline{A}_t^i)^2 > \underline{S}_t^{ii}$. Neste caso, o fator de correção $FC_i = (\underline{A}_t^i)^2 / \underline{S}_t^{ii}$ pode ser usado.

Os resultados experimentais aparecem no teste 24, notando-se que o esquema é bem estável. O gráfico 24, referente ao mesmo teste, mostra o comportamento aproximado de \tilde{h} a medida que $\underline{\Sigma}_t$ crescia.

Mostra por outra, o viés de \underline{S}_t em relação ao seu valor esperado, calculado através da expressão:

$$S_k = \sum_{i,j=1}^n | \underline{S}_{ij}^T - \bar{S}_{ij} | .$$

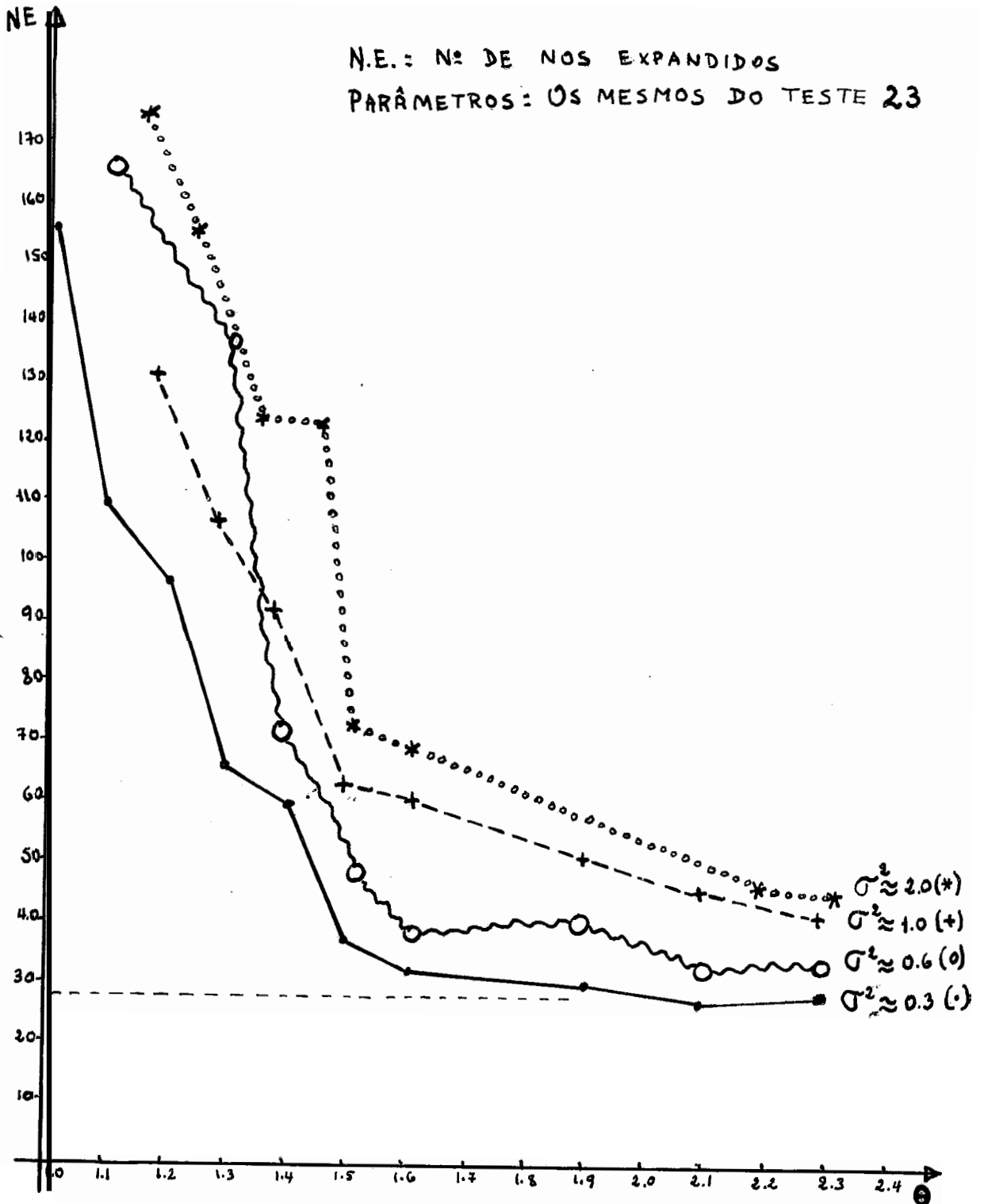


fig.VI-1

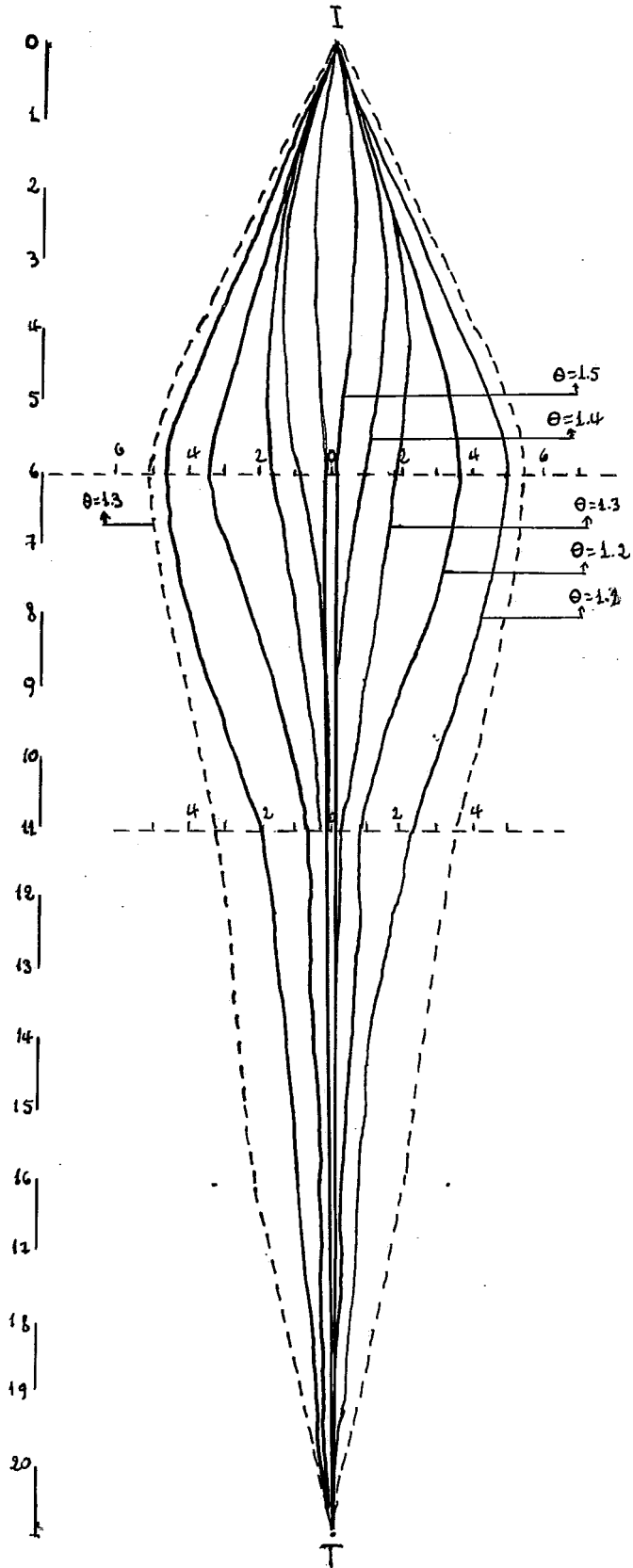


fig. VI-2

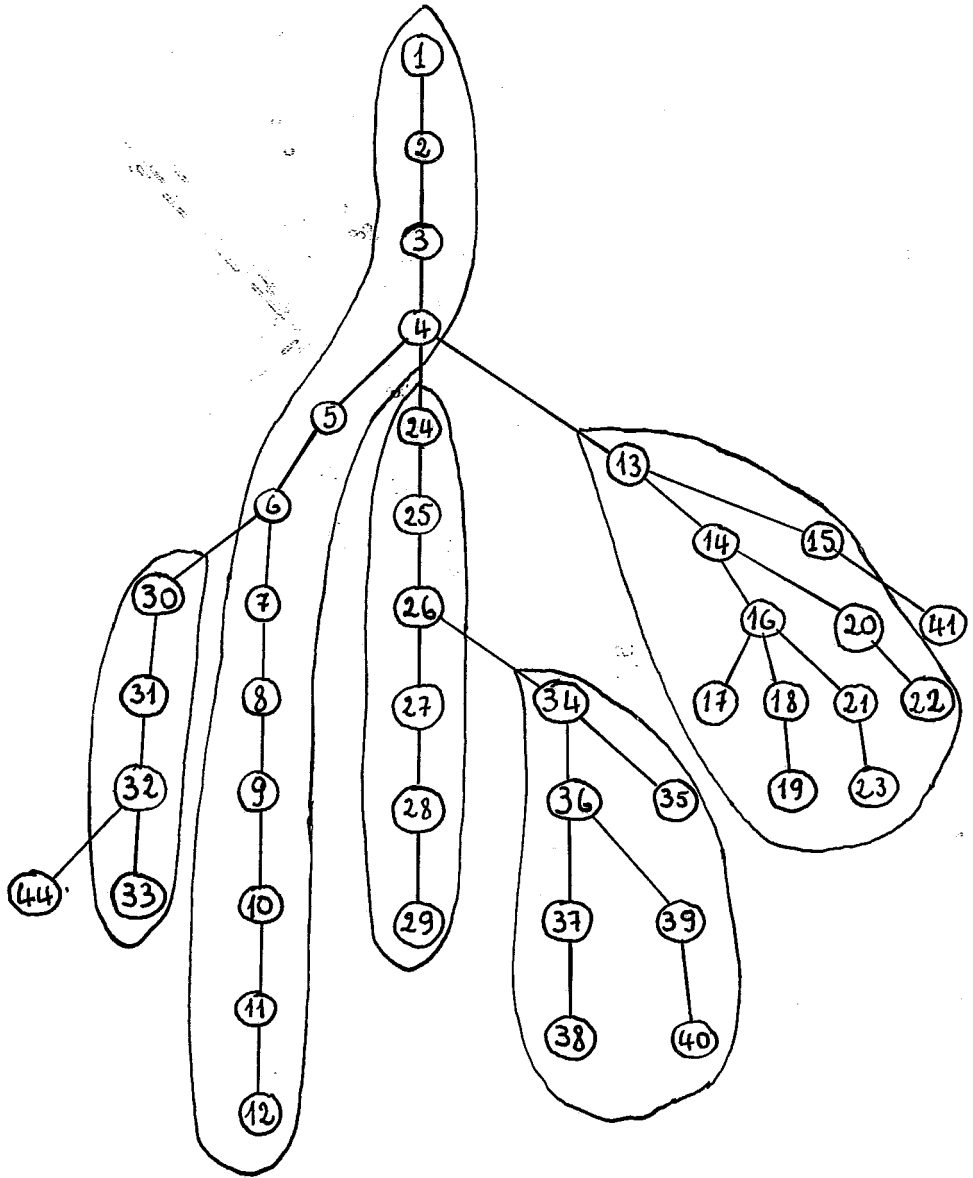


fig. VI-3

M.C. = MENOR CAMINHO
 N.E. = Nº DE NÓS EXPANDIDOS

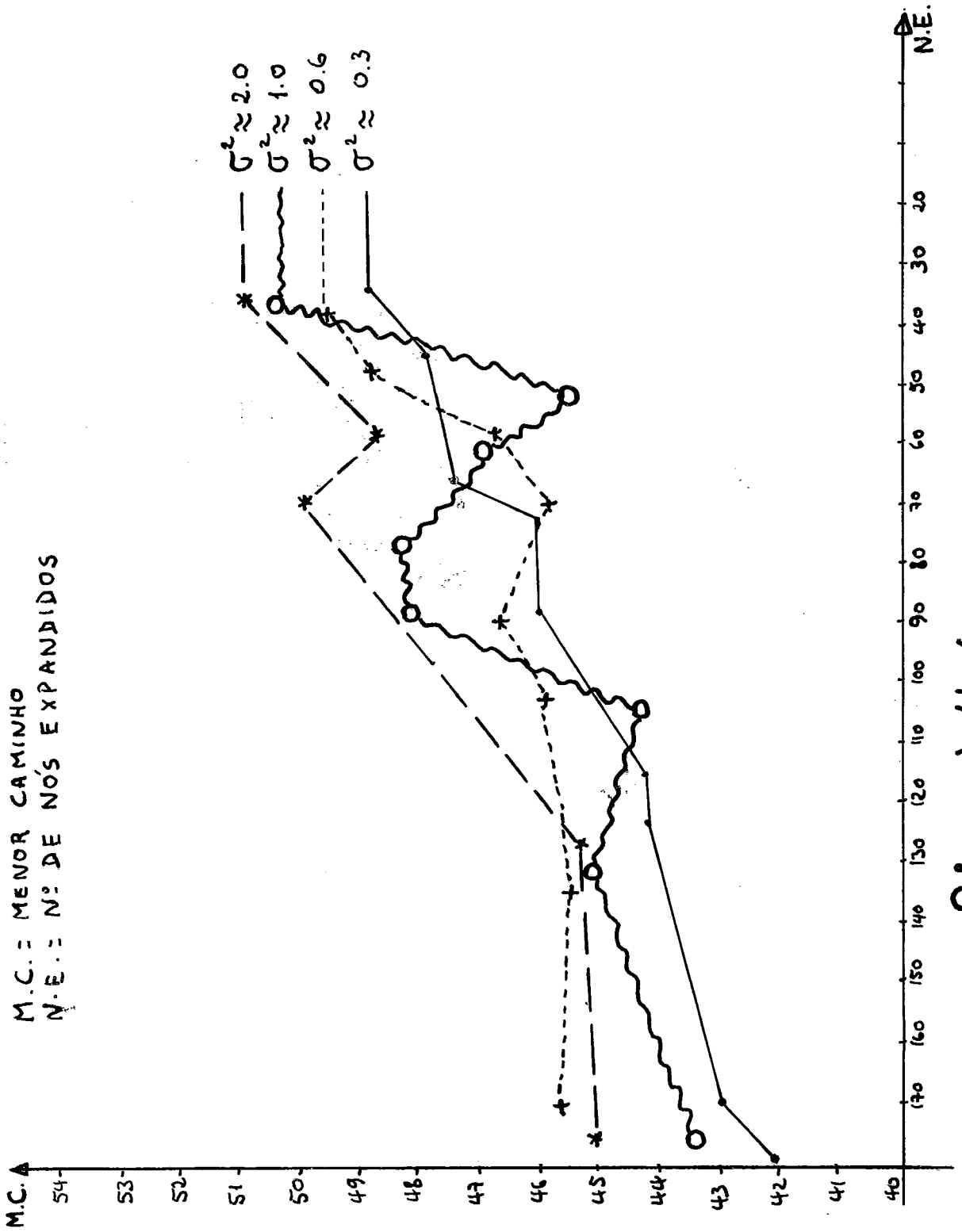


fig. VI-4

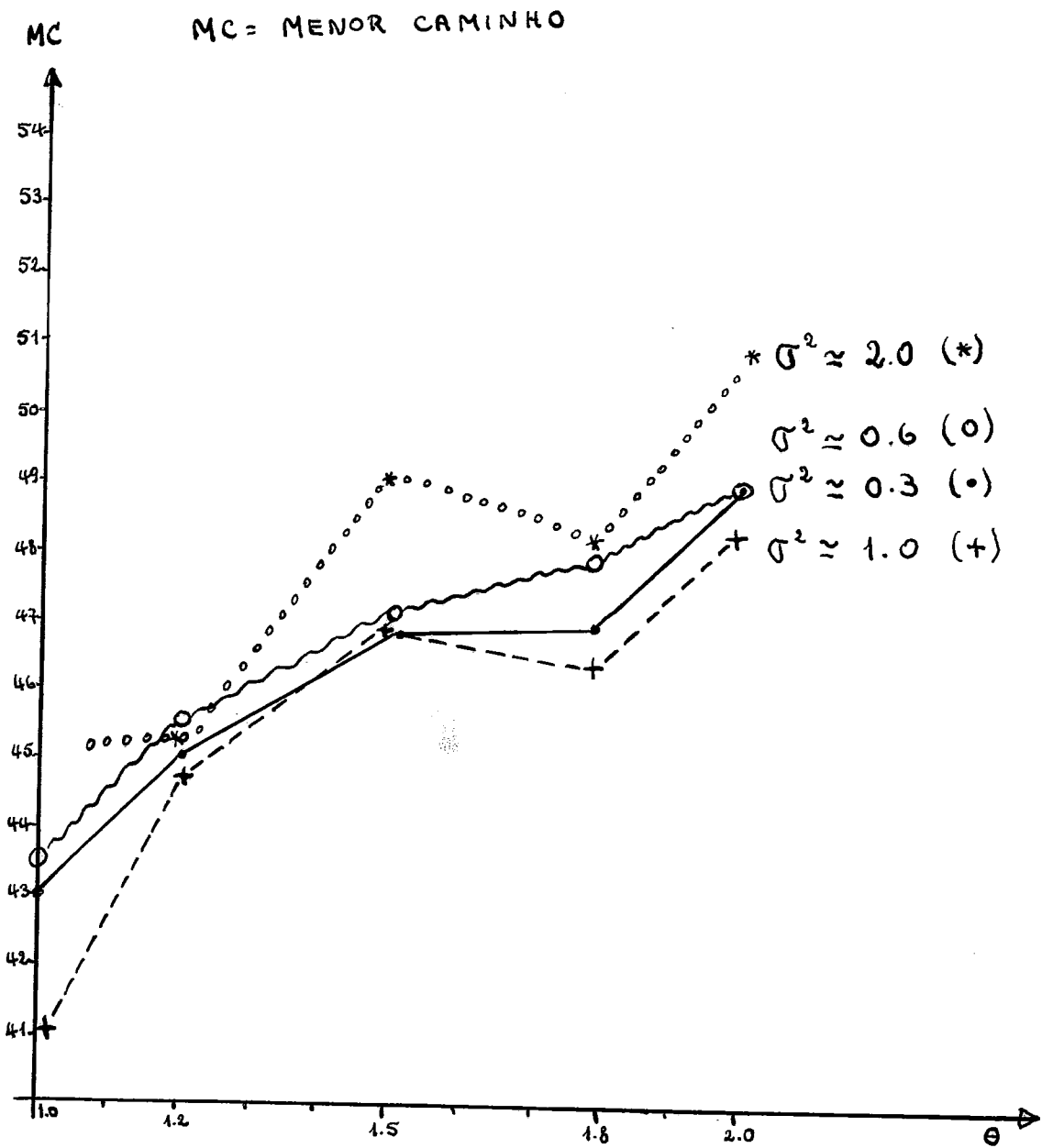


fig. VI-5

6.8 - Estudo da Influência de δ na busca

Quando efetuamos a operação $\hat{h} \leftarrow \theta \hat{h}$, introduzimos uma maior ou menor aceleração da busca no sentido de profundidade. Se $\delta > 1$, aumenta a aceleração, mas se $\delta < 1$, diminui, processando-se uma busca mais horizontal.

Para melhor aquilatar a utilidade de θ , apresentamos uma breve análise, baseada em experimentos práticos. Os resultados devem aproximar-se razoavelmente dos valores esperados, uma vez que utilizamos nos testes $\tilde{h}=h(\text{real})$. Os gráficos da figura VI.1, dão uma idéia da relação entre θ e o número de nós expandidos (NE). Estes gráficos foram traçados ligando os pontos experimentais. Os pontos representam a média de três ou mais experimentos $\bar{\sigma}^2$ é a média das variâncias de Δ .

A figura VI.2, nos dá uma idéia da forma que assume a árvore de expansão como um todo. Enquanto a ordenada indica profundidade, a abscissa indica o número de nós que tem esta profundidade. Note-se que a altura onde a largura da árvore é máxima, depende de δ .

Para o tracejado que envolve todas as outras, foi usado $\bar{\sigma}^2 \approx 2$.

A figura **VI-3**, dá uma idéia mais detalhada da expansão da árvore de busca. A ordem da expansão (números cerca-dos), mostra um cacheamento, que revela o peso dos últimos nós expandidos.

Os gráficos da figura VI.5 relacionam δ com o menor caminho. Os pontos representam médias tiradas dos resultados experimentais do anexo 2.

6.9 - variação da Dimensão de Δ e \underline{S}

A medida que m (número de componentes de \underline{A}_t) cresce, o esquema, quando converge, aproxima mais rapidamente \hat{h} de h , o que se explica pelo maior volume de dados que entram no ajustamento. Um melhor \hat{h} leva a melhores valores de \underline{A}_t e \underline{S}_t . Os testes experimentais com dimensões 2 e 4, apresentaram resultados semelhantes quando a média das variâncias em $\underline{\Sigma}$ eram aproximadamente iguais.

Podemos ter uma idéia do comportamento do esquema através dos testes 25 e 26 e do gráfico 25. Observemos que na iteração 30, \underline{W}_t tinha praticamente a mesma precisão que na 300. Acelerando a busca através de θ , a média de 9 experimentos deu um índice de 45% de acertos, para um h mínimo de 40.

6.10 - Sobre o Menor Caminho

A maioria dos experimentos foi efetuada com $h(1) = 40$, ou seja, menor caminho igual a 40. A solução esperada, caso não fosse usado \hat{h} e em seu lugar fosse feito um sorteio para determinar o n^o a ser expandido, numa busca em profundidade, seria ≈ 52.66 .

É evidente que os resultados dependem de $\underline{\Sigma}$, como podemos ver nos gráficos da figura VI-4. Cada um dos pontos que aparecem nesta figura, representa a média aritmética de pelo menos três resultados do experimento, (σ^2 é a média aritmética da diagonal de $\underline{\Sigma}$).

Para eliminar efeitos devidos a h , foi usado $h = h$ real.

A obtenção dos pontos para $\sigma^2 \approx 0.5$, $\sigma^2 \approx 1.0$ e $\sigma^2 \approx 2.0$, não foi tão tranqüila quanto para $\sigma^2 \approx 0.3$, sendo em alguns casos necessário aumentar o número de experimentos.

Outros resultados sobre menor caminho aparecem nas tabelas VI.1, VI.2 e VI.3.

A tabela VI.1 refere-se ao índice de acertos do esquema por profundidade, parando o experimento com a primeira solução que for detectada. Os recursos das tentativas 2, 3 e 4 do Capítulo V foram usados, fazendo variar os valores de Σ e sendo $h(1)=40$.

Sobre a tabela VI.1 podemos fazer algumas observações:

a) Considerando o número de nós fechados, o caminho ótimo médio para $\delta = 1.0$ e $\theta = 1.1$ é aproximadamente igual.

b) A percentagem de acertos por intervalo só fica significativa a partir da profundidade $(T-20)$ aproximadamente.

c) O percentual de acertos para $\theta = 1.5$ é elevado devido ao pequeno número de nós expandidos. Neste caso os nós expandidos perto de T, onde o índice de acertos é maior, tem maior peso.

A tabela VI.2 é um resumo de experimentos usando a tentativa 3 do Capítulo V, com $h(1)=40$. Até a expansão de 25 nós é usado \underline{W}_t calculado pelo esquema a cada passo. Depois de 25, é usado \underline{W}_{25} .

Sobre a tabela VI.2 podemos fazer as seguintes observações:

a) A tabela representa 10 experimentos, um para cada Σ .

b) O menor caminho médio g^* é a média dos 10 experimentos.

c) Comparando $\bar{g}^* = 44.5$ da tabela VI.2 com $\bar{g}^* = 44.9$ da tabela VI.3, vê-se, com surpresa, que o valor obtido com \underline{W} equi vale ao obtido com \underline{W}_t não Ótimo. (Embora não conclusivo, o dado é sugestivo).

A tabela VI.3, é uma réplica dos experimentos da tabela VI.2, fazendo $\underline{W}_t = \underline{W}_*$.

Um outro dado indicador de desempenho do esquema, é o número de nós que são fechados antes de encontrar o caminho ótimo. Como não é fácil determinar tal número, pois este varia segundo $\underline{\Sigma}$, contentamo-nos com alguns resultados, obtidos com \underline{W}_* , que junto com as tabelas VI.2 e VI.3 podem informar-nos sobre isto. Assim, com $\sigma^2 \approx 0.3$, $h(1)=40$, o melhor caminho em dois experimentos com \underline{W} de dimensão 4, foi detectado depois de 345 e 446 nós expandidos. Com $\sigma^2 \approx 0.6$, depois de 584 nós expandidos, surgiu uma solução subótima $g=42$. Com $\sigma^2 \approx 1.0$, para encontrar $g=41$, 533 nós foram fechados.

É digno de nota também, o baixo desempenho do esquema em melhorar uma solução subótima quando \hat{h} é superestimado. A explicação é simples: Neste caso, o esquema seleciona os nós com menor \hat{h} , pois g terá um peso pequeno e a escolha se concentrará na área em que o esquema já atuara com eficiência quando obteve a primeira solução (observemos o maior índice de acertos perto de T na Tabela VI.1).

Como exemplo apresentamos o caso em que, com $h(1) = 40$, foi encontrada a primeira solução ($g=47$) com o fechamento de 35 nós, e até 660 nós fechados, a melhor solução que apareceu não passou de $g=46$.

ÍNDICE DE ACERTOS

Expansão não dirigida

T : Conjunto de g's obtidos como solução

\tilde{A} e $\underline{\Sigma}$: Os mesmos do teste 23

Int = Intervalo (g de 5 em 5)

Acum = Acumulado

	$\theta = 1.0$		$\theta = 1.1$		$\theta = 1.5$	
	$\underline{W}_t \neq \underline{W}_*$		$\underline{W}_t = \underline{W}_*$		$\underline{W}_t = \underline{W}_*$	
	% de acertos		% de acertos		% de acertos	
Profundidade	Int	Acum	Int	Acum	Int	Acum
(T-50) - (T-45)	38	45	35	44	33	55
(T-45) - (T-40)	40	46	36	44	44	57
(T-40) - (T-35)	40	46	38	45	48	58
(T-35) - (T-30)	40	46	40	46	48	60
(T-30) - (T-25)	41	47	42	51	50	63
(T-25) - (T-20)	42	55	44	54	56	67
(T-20) - (T-15)	52	60	52	62	58	70
(T-15) - (T-10)	54	63	59	68	67	74
(T-10) - (T-5)	66	72	77	79	80	82
(T-5) - (T)	78	78	90	90	84	84
Caminho Médio	45.0		46.2		47.4	
Nós Fechados (média)	101		68		37	
Nº de Experi- mentos	8		12		8	

Tabela VI.1

MENOR CAMINHO

Expansão Dirigida

 \hat{A} e $\underline{\Sigma}$: Os mesmos do teste 23Dim = N? de componentes de \underline{W}_t \bar{o}^2 = Média da diagonal principal de $\underline{\Sigma}$ g_* = Melhor caminho obtido

NNE = N? de nós expandidos (fechados)

Dim	\bar{o}^2	g_*	NNE	\underline{W}_t			
4	0.3	43	192	0.14	0.10	0.05	-0.01
	0.6	45	198	0.05	-0.07	0.06	0.26
	1.0	49	219	0.12	0.00	0.04	0.11
	Esperado			0.02	0.10	0.06	0.07
3	0.3	44	190	-0.13	0.18	0.22	
	0.6	44	191	-0.15	0.21	0.08	
	1.0	45	198	0.00	0.22	0.07	
	Esperado			-0.09	0.19	0.13	
2	0.3	42	182	0.12	0.15		
	0.7	42	184	0.06	0.20		
	1.0	47	209	-0.05	0.23		
	2.0	44	189	0.07	0.17		
	Esperado			0.08	0.17		
Média geral	0.78	44.5	195				

Tabela VI. 2

EXPERIMENTOS COM \underline{W}_*

Expansão Dirigida

 \bar{A} e $\underline{\Sigma}$: Os mesmos do teste 23Dim: Dimensão de \underline{W}_* $\bar{\sigma}^2$: Média da diagonal principal de $\underline{\Sigma}$ g_* : Melhor caminho obtido

NNE: Número de nós expandidos (fechados)

Dim	$\bar{\sigma}^2$	g_*	NNE	\underline{w}_*
4	0.3	42	180	0.02 0.10 0.06 0.08
	0.6	45	198	
	1.0	47	208	
3	0.3	46	181	- 0.09 0.19 0.13
	0.6	43	191	
	1.0	44	191	
2	0.3	44	189	0.08 0.17
	0.6	45	199	
	1.0	48	208	
	2.0	45	199	
Média geral	0.78	44.9	194.4	

Tabela VI.3

6.11 - Otimização da Aplicação do Esquema

Embora o esquema, nas suas rotinas básicas, apresente complexidade polinomial $O(n^2)$, sua aplicação a cada expansão de n pode tornar-se por demais onerosa quando n for grande (o problema como um todo é de natureza exponencial).

No sentido de aliviar o volume de trabalho, a aplicação do esquema foi feita de diversas maneiras. Evidentemente haverá prejuízo em seu desempenho, mas o que importa é: Até onde podemos diminuir o volume de cálculos, mantendo o esquema como instrumento suficientemente útil?

Para verificar este ponto, foram checadas diversas idéias que apareceram.

a) Calcular \underline{W}_t com um intervalo maior que um.

Para não perdermos muita informação, os valores de $\tilde{\underline{S}}_t$ calculados em cada intervalo, seriam armazenados e no fim do intervalo, calculado um $\underline{\underline{S}}_t$ como a média aritmética dos valores acumulados.

Submetido à máquina, tal procedimento em princípio revelou-se em fracasso, conforme se pode ver no teste e gráfico 27.

Para não abandonar a idéia, foi aplicado um fator corretivo FC em $\underline{\underline{S}}_t$, multiplicando todos os elementos de $\underline{\underline{S}}_t$ por FC para que $\underline{\underline{S}}_t^{ii} > (\underline{A}_t^i)^2$. Esta correção produziu um bom efeito, elevando o índice médio de acertos para 41%. Contudo, esta correção só pode ser feita depois que \underline{A}_t e $\underline{\underline{S}}_t$ "tomaram forma" (quando forem mais ou menos iguais a \underline{A} e $\underline{\underline{S}}$), caso contrário, pode produzir distorções catastróficas. Nos experimentos realizados, a

correção foi iniciada após 20 expansões, o que permitiu obter o índice de 41% acima anunciado.

Se a correção for feita apenas sobre a diagonal, os resultados não são tão bons. Para dimensão 4 as estimativas h foram muito boas.

Determinação FC

Seja D o maior quociente $(\underline{A}_t^i)^2 / \underline{S}_t^{ii}$, \forall_i .

O fator Correção (FC) foi assim determinado:

$$\left\{ \begin{array}{l} FC = 1, \text{ se } \underline{S}_t^{ii} > (\underline{A}_t^i)^2 \forall_i \\ (FC = D + 0.0001, \text{ caso contrário.} \end{array} \right.$$

Modo de efetuar a Correção

a.1) Correção sobre todo \underline{S}_t

$$\underline{S}_t^{ij} = \underline{S}_t^{ij} \times FC \quad \forall_i, \forall_j$$

a.2) Correção apenas sobre a diagonal de \underline{S}_t

$$\underline{S}_t^{ii} = \underline{S}_t^{ii} \times FC \quad \forall_i$$

b) Abandonando a idéia de média, uma segunda forma a ser encarada é calcular a cada iteração \underline{A}_t , δ_+ e \underline{S}_t , calculando \underline{p}_t e \underline{W}_t com intervalo maior que um.

Os resultados experimentais com um intervalo de cinco, revelaram um índice de acertos semelhantes ao caso(a) acima, ou seja 41%.

c) Estendendo esta idéia, podemos calcular \underline{A}_t , δ_t , \underline{S}_t e \underline{W}_t em intervalos maiores. Assim, com um intervalo de cinco entre um cálculo e outro, os resultados práticos surpreendentemente mostraram uma média de acertos semelhantes aos de (b), 43%. Para obter esta média foram efetuados 7 experimentos, podendo-se creditar a pequena diferença à casualidade.

d) Prosseguindo com a idéia (c), truncamos completamente o cálculo de \underline{A}_t , δ_t , \underline{S}_t e \underline{W}_t após 25 nós expandidos. Ou seja, a partir desta expansão, sempre foi usado o mesmo \underline{W}_t até o final do teste. O índice médio de acertos foi de 43%, igual ao de (d). Notou-se porém que, quando a dimensão de \underline{S} era 4 (a maior testada), o índice de acertos foi de 47%, o que mostra uma maior rapidez na convergência para dimensões maiores, devido à força de obter h.

e) Sem nenhuma interferência (experimentos realizados com as tentativas 2, 3 e 4 do capítulo V), o índice de acertos foi de 46%.

Nas tentativas a, b, c, d acima, \underline{A}_t e \underline{S}_t foram calculados segundo a tentativa 3 do Capítulo V.

As percentagens aqui apresentadas são globais, servindo apenas para fins de seleção para a forma em que o esquema' será aplicado.

6.12 - Sensibilidade e Precisão do Esquema

Podemos determiná-lo tanto considerando a qualidade das estimativas quanto a quantidade de nós expandidos ou índices de acerto.

No aspecto qualitativo, o esquema produziu bons resultados no referente às estimativas \tilde{h} e \underline{A}_t , e razoáveis para h , \underline{s}_t e δ_t .

No aspecto quantitativo os resultados devem ser julgados comparativamente, tomando-se como paradigma resultados obtidos com \underline{W}_* . Neste sentido podemos consultar a tabela VI-1.

As experiências com \underline{W}_* também mostraram, o que era esperado, que h sempre subestima levemente θh (para compensar a dispersão), exigindo um $\theta - 1.1$ para igualá-los. Este detalhe também foi notado em aproximadamente 80% dos experimentos realizados com \underline{W}_t determinado pelo esquema.

6.13 - Conclusão

Comparando os resultados do esquema com aqueles obtidos com \underline{W} , podemos afirmar que o esquema aprova e que algumas das sugestões apresentadas nas seções 6.2 e 6.11 são perfeitamente viáveis.

Cumpramos salientar também que não são as rotinas do esquema as únicas a serem questionadas em relação ao tempo e volume de trabalho, mas, também as rotinas de atualização e reordenação dos nós abertos devem ter programação e tratamento adequado.

A seção 6.2 sugere que o teste com "Expansão não Dirigida" apresenta dificuldades na avaliação da força da heurística

tica para problemas de maior porte, sendo que indicamos como complemento, a Expansão Dirigida.

C A P Í T U L O VII

CONCLUSÕES

Para, numa apreciação sucinta, apresentar as principais conclusões relativas ao presente trabalho, baixamos novamente aos seus tres aspectos, a saber: Estimativa dos parâmetros, Correção da tendenciosidade e Aspectos da aplicação do esquema.

7.1 - Sobre a Estimativa dos Parâmetros

Sem recorrer aos recursos empregados na correção da tendenciosidade, os melhores valores obtidos para \underline{A}_t , \underline{S}_t e δ_t foram conseguidos usando as fórmulas (IV-23), (IV-24) e (IV-32) respectivamente. Em quase todos os experimentos realizados, \underline{A}_t e \underline{S}_t mantiveram entre si uma quase-proporcionalidade, que permitiu que os valores de h calculados numa mesma iteração, fossem boas heurísticas.

Como a precisão das estimativas dos parâmetros depende do estado da busca, houve necessidade de atualização de \hat{h} segundo o avanço da busca. A medida que os nós abertos cresciam em número, crescia também o tempo de reordenação após cada atualização nos \hat{h} . Para equacionar o problema, as atualizações de h foram efetuadas em iterações (t) espaçadas por intervalos iguais a m , ($m = \text{integer}(t/20)+1$).

Os fatores de convergência (ganhos por iteração) α , β , γ usados, foram $\alpha = \beta = \gamma = 1/t$, os quais cumpriram seu papel satisfatoriamente.

\underline{W}_t , que inicialmente tinha sido calculado através da expressão $\underline{W}_t = \underline{W}_{t-1} + \gamma_t (\underline{W}_{t-1} - \tilde{W}_t)$, teve seu cálculo mudado para $\underline{W}_t = \underline{S}_t^{-1} \underline{A}_t$, devido a excessiva fixidez da primeira forma. Realmente os resultados com a primeira forma não eram bons e os valores iniciais pesavam demasiadamente.

$\tilde{h}(n_t)$, obtido via ajustamento, também deve ser considerado satisfatório e foi, sem dúvida, a partida da estabilização do esquema. $\tilde{h} = \tilde{W}_t \hat{H}$ não satisfazia de forma nenhuma.

P_i , a probabilidade atribuída aos sucessores dos nós expandidos, deve ser avaliada em função dos resultados que o esquema produziu. Para verificar estes resultados, consultemos a tabela VI-1, que mostra a pequena diferença entre testes com \underline{W}_t e \underline{W}_* .

Os valores obtidos para \hat{h} foram bastante bons, basta para isto consultar m e v nos diversos gráficos apresentados. Consultando estes mesmos gráficos, vemos também através do símbolo \hat{h} , que há em geral uma coerência nas variações de \hat{h} e \tilde{h} , conseqüência direta do peso do Último na determinação de \tilde{A}_t e \tilde{S}_t .

7.2 - Sobre a Correção da Tendenciosidade

As tentativas 2 e 3 apresentaram resultados ligeiramente superiores aos do experimento sem a correção de tendenciosidade. A estimativa de $\underline{\Sigma}$ é claramente melhor e isto é revelado através de \hat{h} e \mathbf{h} dos gráficos. A tentativa 3, apesar de abranger menor volume de informação do que a 2, não forneceu nos experimentos elementos que demonstrassem inferioridade em relação a ela. Além do mais, exige menor volume de cálculos.

A tentativa 4, embora \underline{A}_t e \underline{S}_t de modo geral fossem

bons, mostrou uma grande instabilidade em \hat{h} , revelada nos gráficos através de h , v e m . Esta instabilidade contra-indica o esquema sob esta forma nos casos em que ele é usado sem atualizar a cada passo os h em "Aberto". Este fato, além do volume de cálculos, põe a tentativa 4 em desvantagem em relação às tentativas 2 e 3.

Em suma, como a aplicação de uma das tentativas (2 ou 3) não aumenta significativamente o custo em cálculos, não há porque não optar por uma delas.

7.3 - Sobre a Aplicação do Esquema

Para manter a estabilidade de h é imperioso evitar a inatibilidade se \underline{s}^{-1} . A singularidade de \underline{s} pode ser contornada através de uma programação tipo "sense switch" que adotamos na subrotina de Inversão de Matrizes. A instabilidade de h pode ser detectada através de testes entre \underline{S}_t e \underline{A}_t como sugerido e experimentado no capítulo VI, e corrigida através de correções sobre \underline{S}_t . A estabilidade de \hat{h} é importante para uma aplicação eficiente do esquema (como abaixo sugerido), não convindo, portanto, omitir este teste que permite repor o equilíbrio devido entre \underline{A}_t e \underline{S}_t .

A dimensão da matriz \underline{s} não preocupa no que tange à precisão dos resultados, pois, as experiências mostraram que as estimativas dos parâmetros até melhoram, devido em parte à melhor estimativa de \hat{h} . A convergência processa-se mais rapidamente com dimensões maiores, o que permite abandonar antes o cálculo de \underline{W}_t , \underline{A}_t , \underline{S}_t e δ_t como sugerido no capítulo VI. O Único inconveniente de uma matriz \underline{s} com dimensões razoavelmente grandes, é o volume

de cálculos.

Considerando a teoria e os experimentos acima enunciados, sugerimos uma forma de aplicar eficientemente o esquema:

Expandir uns 30 nós, acelerando na profundidade, pegando para \underline{W}_t a média dos 5 últimos obtidos. A partir de então, manter o mesmo \underline{W}_t até obter a solução, aplicando uma expansão dirigida. É evidente que esta forma encerra riscos quando há possibilidade de as soluções serem grandes, mas garante em geral uma boa solução (veja tabela VI.3).

Merecem aqui citação especial, os experimentos realizados aproximando \underline{S} por matriz escalar, descritos na seção 6.2. A principal vantagem do esquema adaptativo aplicado desta forma, é sem dúvida, o pequeno volume de cálculos por iteração. Contudo, a expansão dirigida tem seus inconvenientes tais como a rigidez no número de nós expandidos, dificuldades em prosseguirabuscacaso a solução ótima seja a meta e esta ainda não tenha sido determinada, restringir-se a \hat{h} não levando em conta g .

Assim, dependendo do objetivo da busca, a aplicação de \underline{S} matriz escalar com Expansão dirigida pode ser uma boa alternativa, especialmente se tiver muitos componentes em \hat{H} e a dispersão for pequena.

Em suma, embora não tivéssemos feito comparações de eficiência com outros métodos, em parte por falta de tempo, os resultados mostraram-se satisfatórios ao serem comparados com os ótimos previstos, usando os parâmetros do problema.

Enfim, pode o esquema de aproximação estocástico não ser o mais eficiente de todos, mas é seguramente simples, de fácil implementação e bom desempenho.

A P Ê N D I C E 1

Este apêndice contém um resumo dos diversos experimentos realizados, numerados seqüencialmente.

Para cada teste realizado, são apresentados os principais parâmetros e fórmulas usadas, valores iniciais, uma tabela resumo e, quando proveitoso, um gráfico.

A tabela e o gráfico representam apenas uma amostra dos resultados experimentais obtidos com o esquema na forma apresentado no teste.

Não apresentamos todos os resultados, porque isso não é importante e também, porque é praticamente inviável apresentar tamanho volume de dados.

SIMBOLOGIA E ABREVIACÕES ADOTADAS NOS TESTES

C_i : Custo da aresta i da árvore expandida (foi mantida constante e igual a 2).

$h(1)$: Custo da solução Ótima (única)

B_t : Nº de sucessores do nó expandido (foi mantido constante e igual a 3).

PVO : Abreviação de Parâmetros e Valores Ótimos.

\underline{C} : Matriz de Covariância de \hat{H} .

Estabilidade: Indica que o teste proposto serve para verificar a estabilidade do esquema.

convergência: Indica que o teste serve para verificar a convergência do esquema.

Tentativa "n":Relaciona o teste com a análise teórica correspondente.

\tilde{W}_t : Obtido de observações locais, considerados apenas os sucessores do nó em expansão.

Expansão : Indica se foi usada expansão dirigida ou não dirigida.

θ : Fator de \hat{h} para aceleração da busca.

$\bar{\sigma}^2$: Média da diagonal principal de $\underline{\Sigma}$.

ELEMENTOS USADOS NOS GRÁFICOS

Símbolo	Fórmula	Escala	Significado
\hat{h}	$(h-h)$	1.0	Desvio de \hat{h}
v	$(\underline{W}_t - \underline{W}^*)^T \underline{\Sigma} (\underline{W}_t - \underline{W}^*)$	10.0	Variância (Desvio do \hat{h} ótimo)
m	$(\underline{W}_t - \underline{W}^*)^T \underline{\bar{A}}$	10.0	Média (Desvio de \hat{h} ótimo)
\underline{A}_t	$\sum_{i=1}^D (\underline{A}_{t,i} - \underline{\bar{A}}_i) $	1.0	Desvio de \underline{A}_t
\underline{S}_t	$\sum_{i=1}^D (\underline{S}_{t,ii} - \underline{S}_{ii}) $	1.0	Desvio de \underline{S}_t
δ_t	δ_t	10.0	Delta
\tilde{h}	$(h-h)$	1.0	Desvio de h
ε	$\frac{\varepsilon}{\ \underline{A}_t\ ^2}$	1.0	Relação $\varepsilon / \ \underline{A}\ ^2$

TESTE 1

Tentativa: 0

Tipo de teste: Estabilidade

Experimento com IV.22, IV.23 e IV.24

$$h(1) = 30$$

$$\tilde{h}(n_t) = \tilde{W}_t \hat{H}(n)$$

Nº de iterações: 50

$$C_i = 2.0$$

$$B_t = 3$$

$$\underline{\underline{\Sigma}} = \begin{pmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{pmatrix}$$

Resultados Extraídos do Experimento

t	1 (PVO)	16	32	50
\underline{A}_t	2.00	1.30	1.28	0.98
	5.00	3.76	2.85	2.20
\underline{W}_t	0.08	0.25	0.43	1.18
	0.17	0.09	0.02	-0.30
δ_t	1.00	1.86	2.00	2.00
$\underline{\underline{S}}_t$	4.5 9.7	3.5 7.7	2.8 6.2	2.3 5.1
	9.7 25.0	7.7 19.7	6.2 15.8	5.1 13.0

Resumo: a) Tendenciosidade acentuada em \underline{A}_t , \underline{W}_t , δ_t e $\underline{\underline{S}}_t$

b) Índice de acertos: 40%

c) Caminho subótimo: Não determinado no teste

d) \underline{W}_t bastante instável (ver gráfico 1)

(*) PVO: Parâmetro ou valor Ótimo

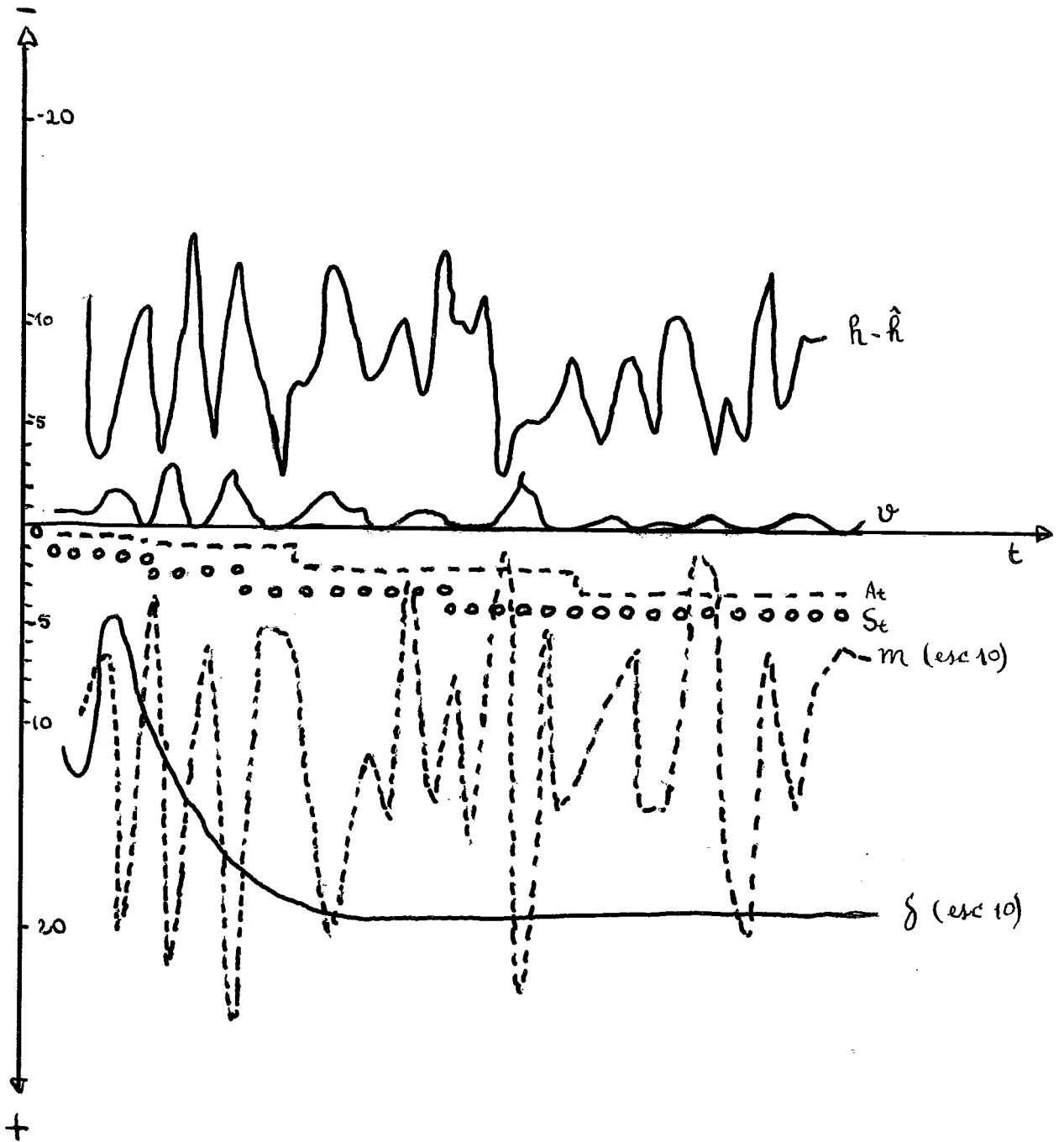


gráfico:1

T E S T E 2

Tentativa: 0

Experimento com IV-32, IV-23 e IV-24

Tipo de teste: Estabilidade

$$h(1) = 40$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

Nº de iterações: 200

$$C_i = 2.0$$

$$B_t = 3$$

$$\underline{\underline{\Sigma}} = \begin{pmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{pmatrix}$$

Resultados obtidos em 200 iterações

t	1 (PVO)	30	115	200
\underline{A}_t	2.00	2.07	2.23	2.26
	5.00	5.10	5.54	5.69
\underline{W}_t	0.08	0.07	0.06	0.06
	0.17	0.17	0.15	0.15
δ_t	1.0	1.45	1.92	1.89
	4.6 10.0	4.9 10.5	5.6 12.4	5.8 12.8
	10.0 25.4	10.5 26.5	12.4 31.3	12.8 33.1

Resumo: a) Estabilidade boa para \underline{A}_t , \underline{W}_t e δ_t

b) Índice de acertos: 42%

c) Caminho subótimo: Não determinado no teste

d) $\underline{\underline{S}}_t$ com tendenciosidade mais acentuada (gráfico 2)

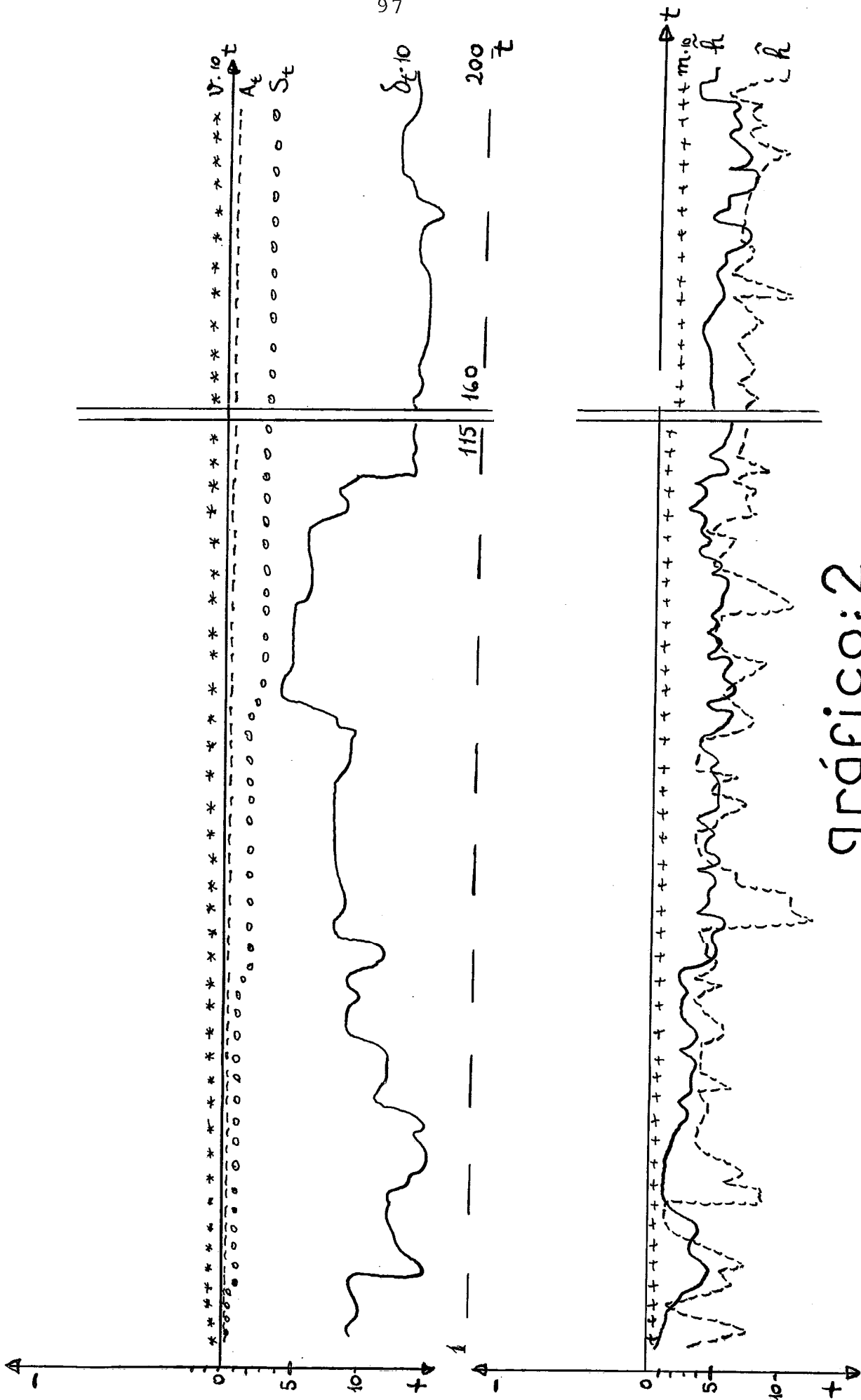


Gráfico:2

T E S T E 3

Tentativa: 0

Experimento com IV.32, IV.23 e IV.24.

Tipo de teste: convergência

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \begin{Bmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{Bmatrix}$$

Resultados obtidos em 300 iterações

t	1	30	165	300	(PVO)
	1.0	2.42	2.67	2.53	2.0
	1.0	6.37	6.70	6.47	5.0
	0.33	0.03	0.04	0.04	0.08
	0.33	0.11	0.13	0.13	0.17
	0.50	1.96	1.57	0.92	1.0
	2.0 1.0	13.5 19.3	9.4 18.5	8.0 16.7	4.6 9.7
	1.0 2.0	19.3 54.6	18.5 48.0	16.7 44.0	9.7 25.0

Resumo: a) convergência razoável para \underline{A}_t , δ_t , \underline{W}_t .

b) Índice de acertos: 39%

c) Caminho subótimo: não determinado no teste

d) Note-se que a certa proporcionalidade de \underline{A}_t e \underline{S}_t , proporcionou um bom \underline{W}_t .

e) Gráfico associado: 3

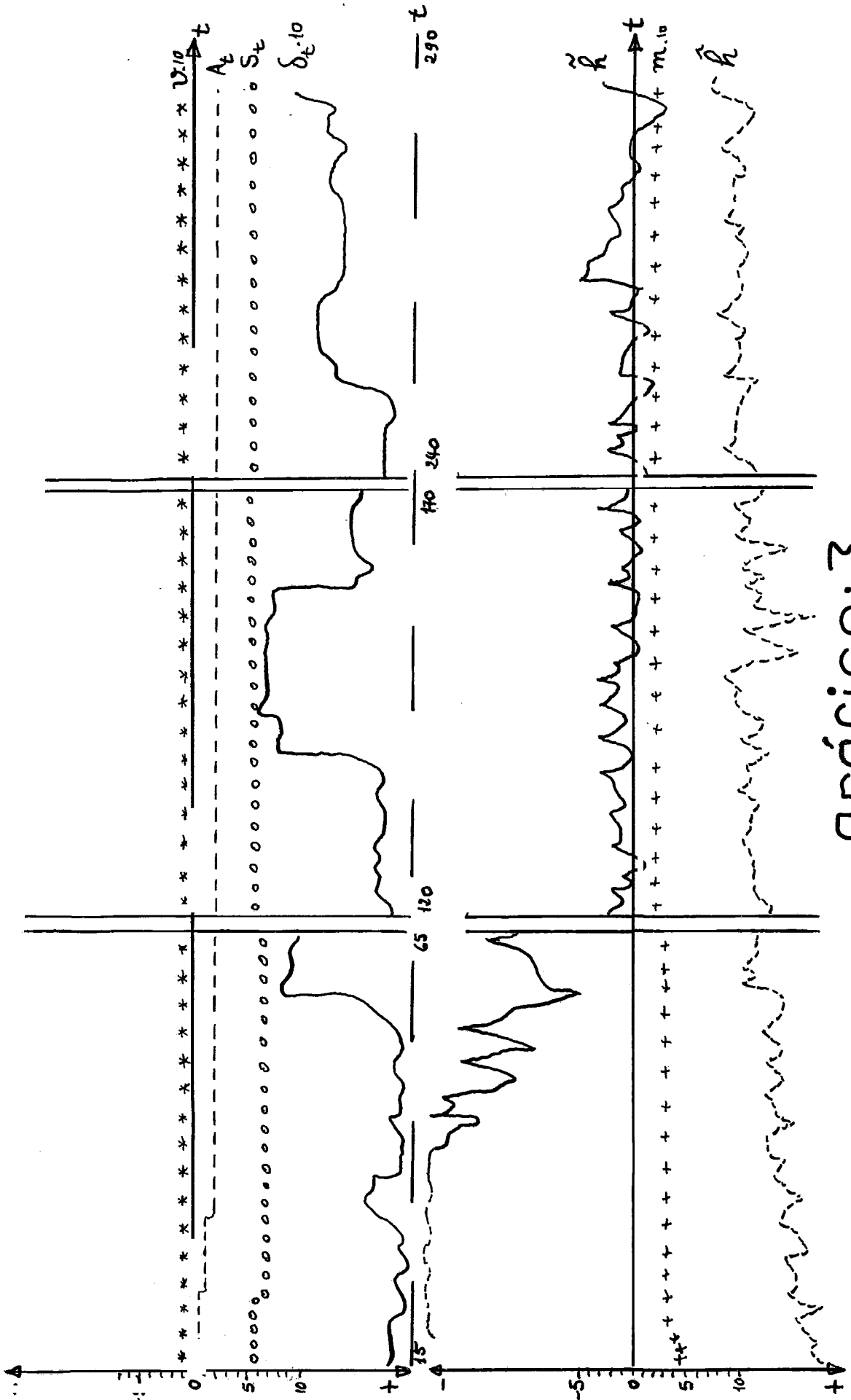


Gráfico:3

TESTE 4

Tentativa: 0

Experimento com IV.32, IV.23 e IV.24

Tipo de teste: convergência

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\theta = 1.5$$

$$\tilde{h}(n_t) = h \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \begin{pmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{pmatrix}$$

Resultados obtidos em 28 iterações

t	1	10	20	25	(PVO)
\underline{A}_t	1.0	1.0.4	1.85	2.11	2.0
	1.0	2.69	4.46	4.99	5.0
\underline{W}_t	0.33	0.02	0.02	0.02	0.08
	0.33	0.14	0.15	0.15	0.17
δ_t	0.50	0.79	1.53	1.57	1.0
\underline{S}_t	2.0 1.0	13.3 5.4	10.6 11.0	10.8 13.5	4.6 9.7
	1.0 2.0	5.4 18.1	11.0 28.8	13.5 33.6	9.7 25.5

Resumo: a) convergência boa para \underline{A}_t , \underline{W}_t e δ_t

b) Caminho Subótimo :48

c) Gráfico associado: 4

d) Índice de acertos: 60%

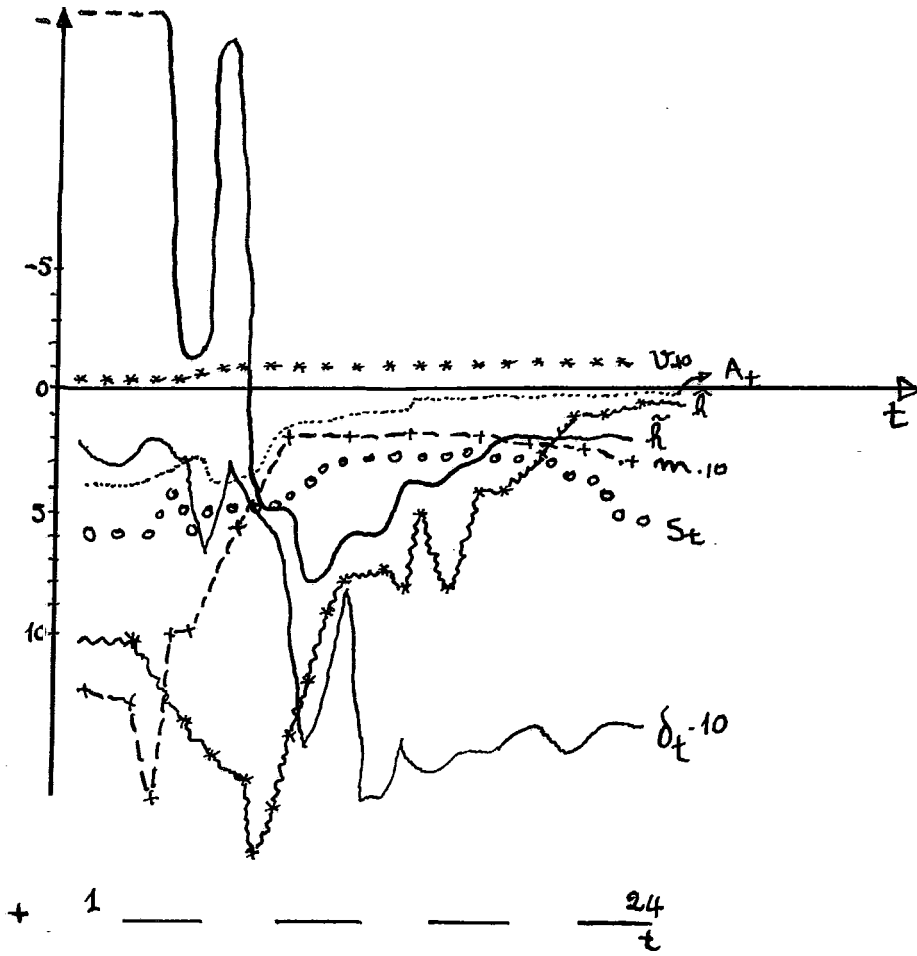


gráfico:4

TESTE 5

Tentativa: 2

Experimento com IV.32, V.9 e V.14

Tipo de Teste: Estabilidade

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = h \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \left\{ \begin{array}{cc} 0.550 & -0.087 \\ -0.087 & 0.453 \end{array} \right\}$$

Resultados obtidos em 200 iterações

t	1 (PVO)	30	115	200
\underline{A}_t	2.00	1.96	2.04	2.10
	5.00	4.90	5.33	5.40
\underline{W}_t	0.08	0.07	0.07	0.06
	0.17	0.17	0.15	0.15
δ_t	1.00	1.73	1.03	0.82
$\underline{\underline{S}}_t$	4.6 9.7	4.6 9.7	5.0 11.2	5.4 12.0
	9.7 25.4	9.7 25.0	11.2 30.0	12.0 31.6

Resumo: a) Estabilidade boa para \underline{A}_t , \underline{W}_t , δ_t e $\underline{\underline{S}}_t$

b) Caminho subótimo: não determinado no teste

c) Índice de acertos: 41%

d) Gráfico associado: 5

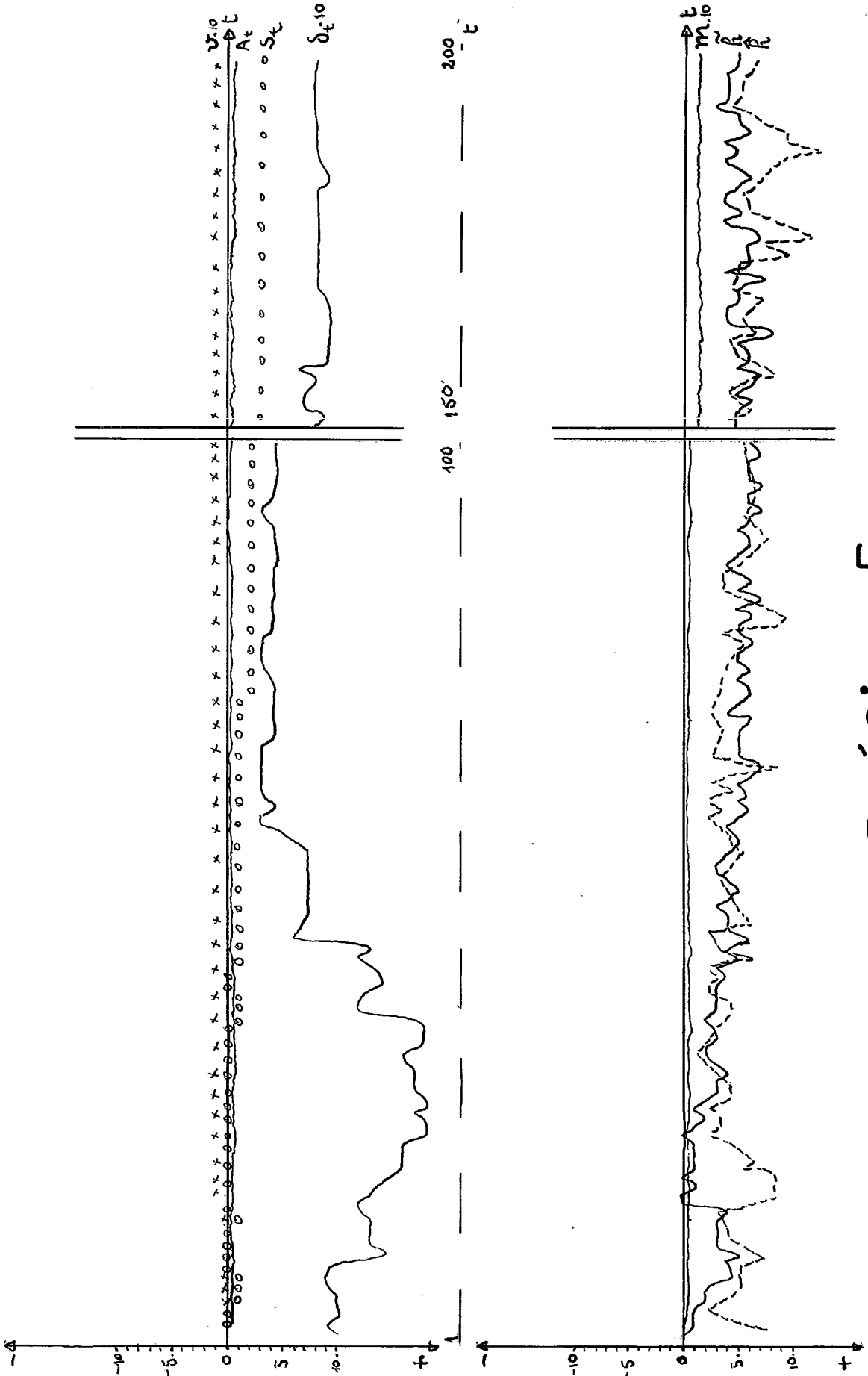


Gráfico:5

T E S T E 6

Tentativa: 2

Experimento com IV.22, V.9 e V.14

Tipo de teste: Estabilidade

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{\Sigma} = \left\{ \begin{array}{cc} 0.550 & -0.087 \\ -0.087 & 0.453 \end{array} \right\}$$

Resultados obtidos em 300 iterações

t	1 (PVO)	30	165	300
\underline{A}_t	2.00	1.95	2.06	2.07
	5.00	4.89	5.28	5.26
\underline{W}_t	0.08	0.07	0.06	0.07
	0.17	0.17	0.15	0.14
δ_t	1.00	2.00	2.00	2.00
\underline{S}_t	4.6 9.7	4.6 9.7	5.3 11.9	5.4 11.9
	9.7 25.4	9.7 25.0	11.9 31.1	11.9 31.1

- Resumo: a) Estabilidade boa para \underline{A}_t , \underline{W}_t , δ_t e \underline{S}_t
 b) Caminho subótimo não determinado no teste
 c) Índice de acertos: 42%
 d) Gráfico associado: 6

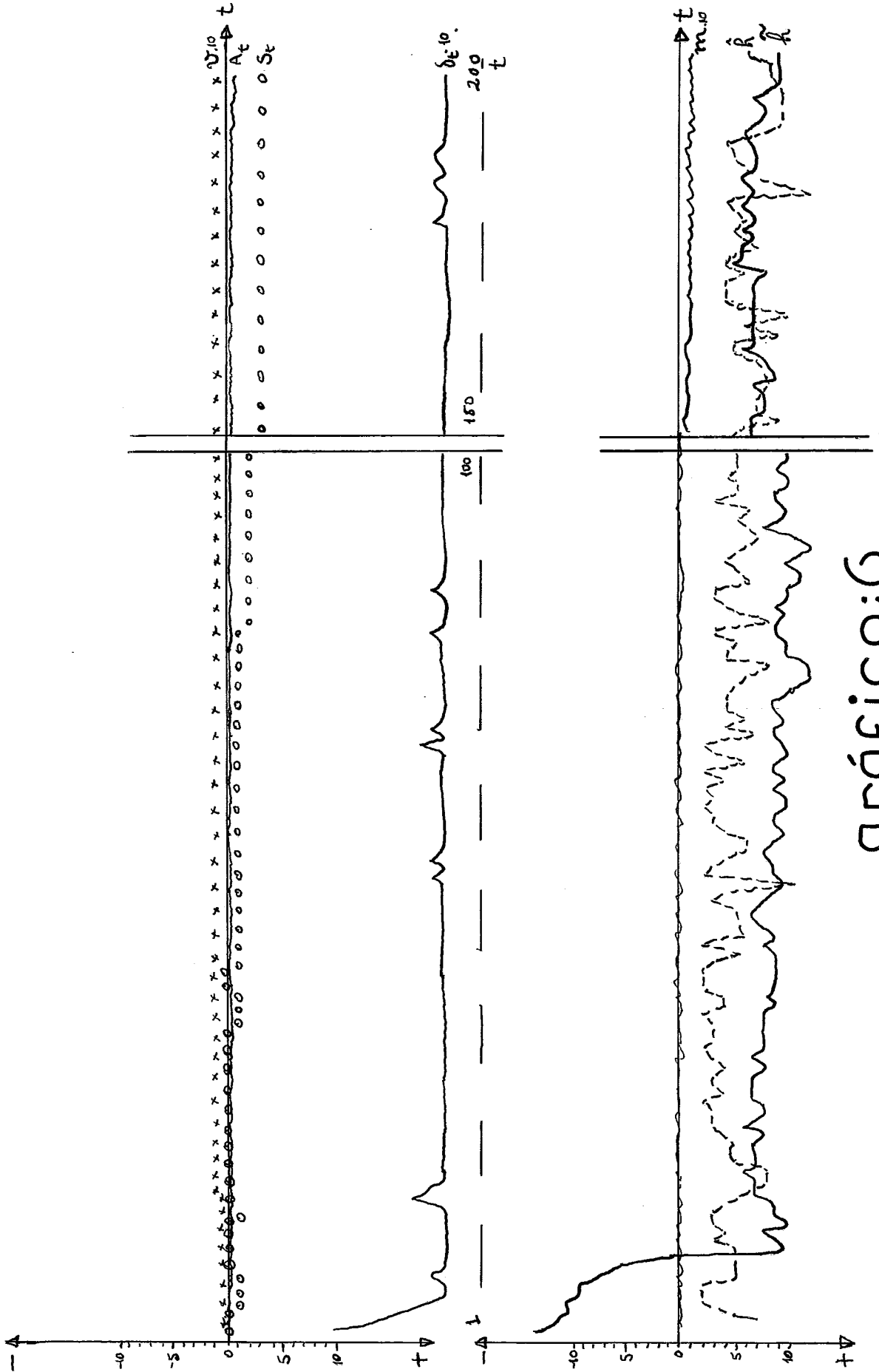


Gráfico:6

T E S T E 7

Tentativa: 2

Experimento com IV.32, V.9 e V.14

Tipo de Experimento: Estabilidade

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = h \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \begin{Bmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{Bmatrix}$$

Resultados obtidos em 120 iterações

t	1 (PVO)	30	75	120
\underline{A}_t	4.00	3.98	4.04	4.15
	3000	3.01	3.09	3.13
\underline{W}_t	0.14	0.14	0.14	0.13
	0.14	0.14	0.13	0.13
δ_t	1.00	1.15	0.99	0.40
$\underline{\underline{S}}_t$	16.6 11.9	16.7 12.1	17.4 12.8	18.5 13.5
	11.9 9.4	12.1 9.7	12.8 10.4	13.5 10.7

Resumo: a) Estabilidade boa para \underline{A}_t , \underline{W}_t e $\underline{\underline{S}}_t$

b) Caminho subótimo: não determinado no teste

c) Gráfico associado: 7

d) Índice de acertos: 43%

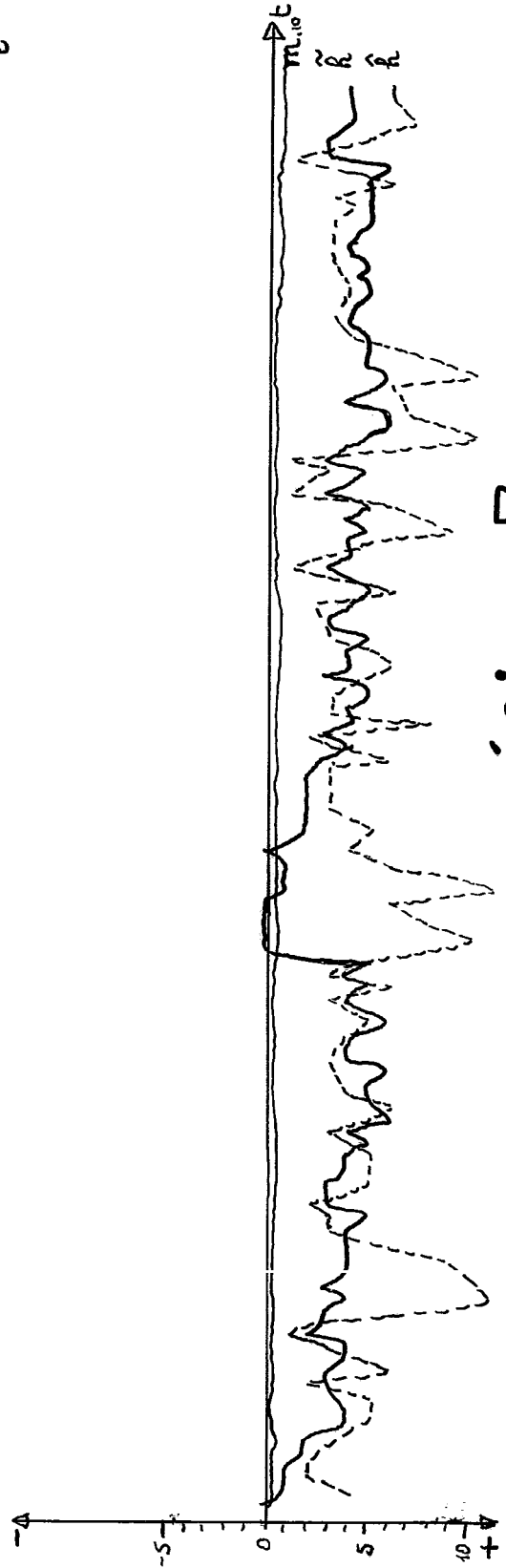
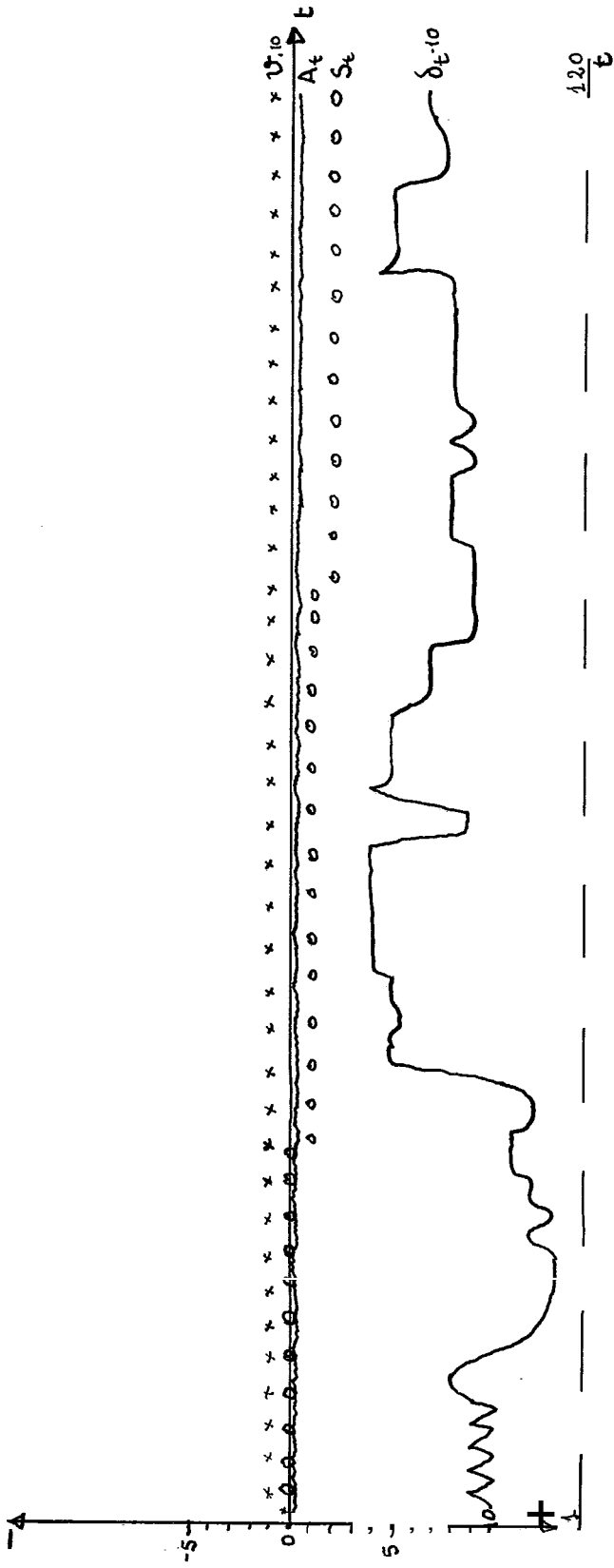


gráfico:7

TESTE 8

Tentativa: 2

Experimento com IV.22, V.9 e V.14

Tipo de teste: Convergência

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$h = h \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \begin{pmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{pmatrix}$$

Resultados obtidos em 200 iterações

t	1	30	115	200	(PVO)
\underline{A}_t	1.0	1.15	1.58	1.70	2.00
	1.0	2.65	4.12	4.44	5.00
\underline{W}_t	0.33	0.04	0.08	0.08	0.08
	0.33	0.25	0.17	0.16	0.17
δ_t	0.50	2.00	2.00	2.00	1.0
\underline{S}_t	2.0 1.0	2.2 4.3	3.4 7.7	3.9 8.8	4.6 9.7
	1.0 2.0	4.3 10.1	7.7 20.62	8.8 23.6	9.7 25.4

Resumo: a) convergência boa para \underline{A}_t , \underline{W}_t e \underline{S}_t

b) Caminho Subótimo: não determinado no teste

c) Gráfico associado: 8

d) Índice de acertos: 42%

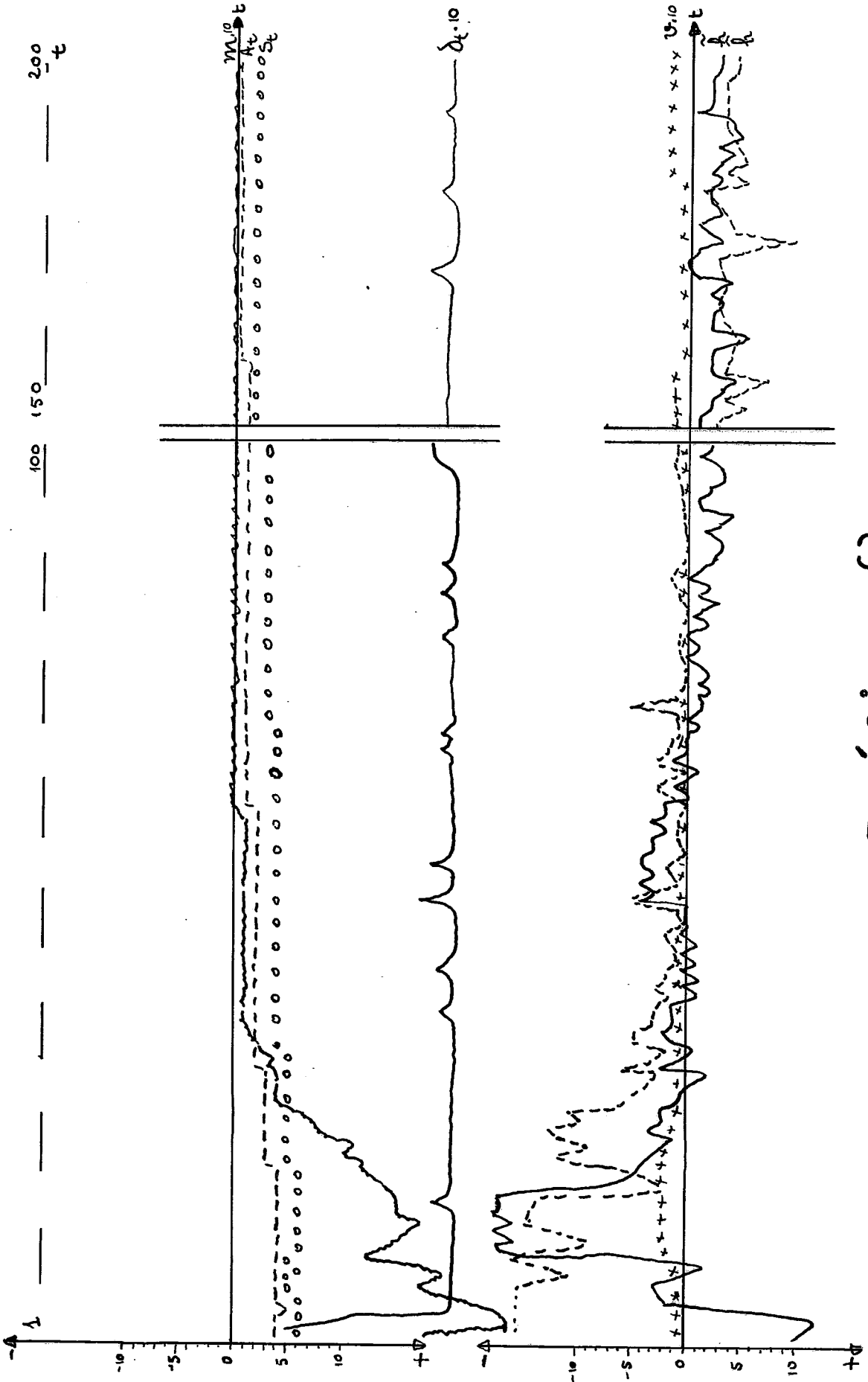


Gráfico: {}

T E S T E 9

Tentativa: 2

Experimento com IV.22, V.9 e V.14

Tipo de teste: convergência

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h} = \tilde{h} \text{ ajustado}$$

$$\underline{\Sigma} = \begin{pmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{pmatrix}$$

Resultados obtidos em 200 iterações

t	1	30	115	200	(PVO)
\underline{A}_t	1.0	2.86	3.54	3.65	4.0
	1.0	1.91	2.49	2.59	3.0
\underline{W}_t	0.33	0.08	0.11	0.13	0.14
	0.33	0.29	0.18	0.14	0.14
δ_t	0.50	2.00	2.00	2.00	1.0
	2.0 1.0	11.3 6.9	15.1 10.1	15.9 10.9	16.5 11.9
	1.0 2.0	6.9 4.8	10.1 7.4	10.9 8.1	11.9 9.4

Resumo: a) convergência boa para \underline{A}_t , \underline{W}_t e $\underline{\Sigma}_t$

b) Caminho subótimo: não determinado no teste

c) Gráfico associado: 9

d) Índice de acertos: 42%

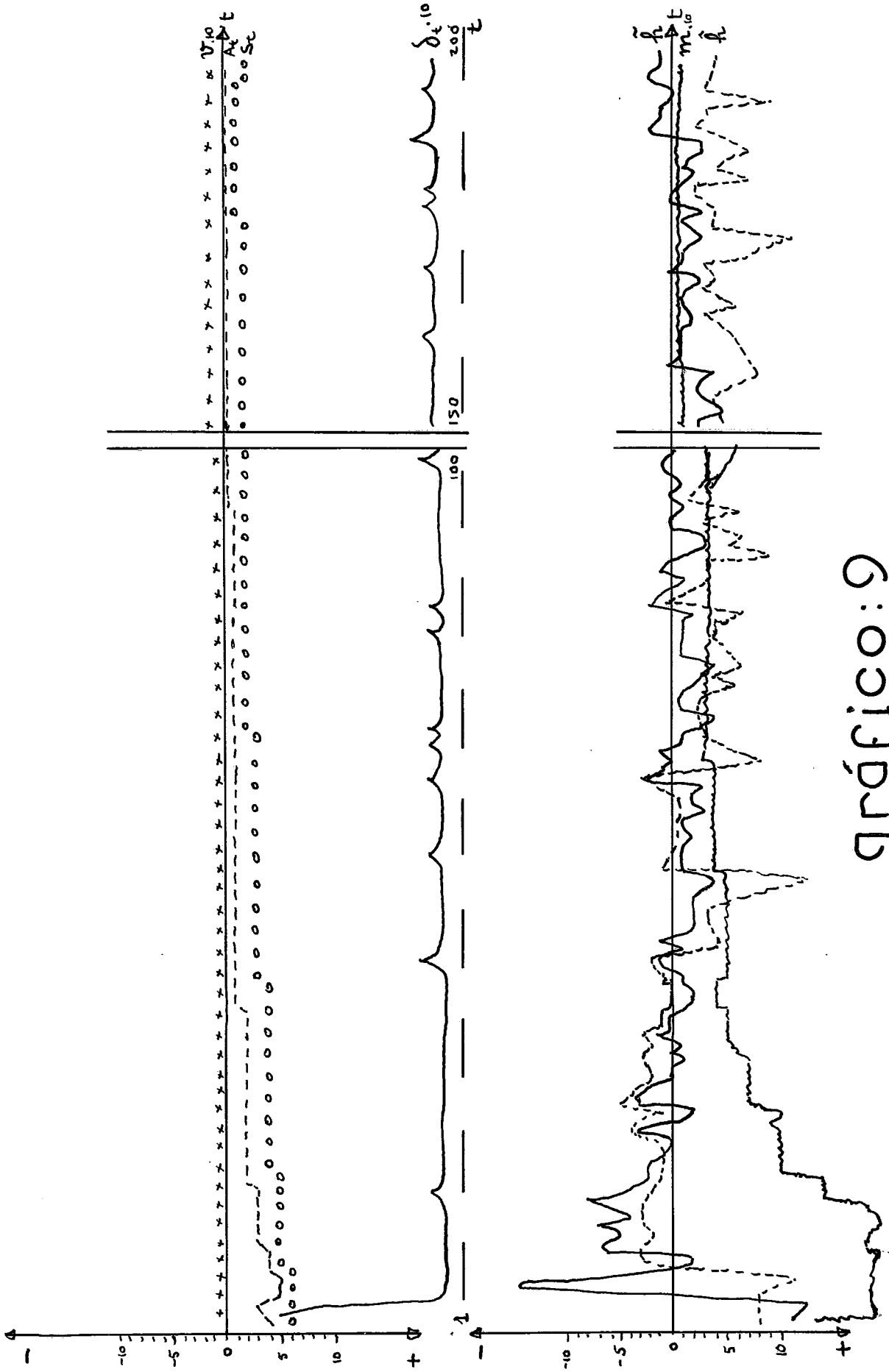


Gráfico: 9

T E S T E 10

Tentativa: 2

Experimento com IV.32, V.9 e V.14

Tipo de teste: Convergência

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$h = \tilde{h} \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \left\{ \begin{array}{cc} 0.550 & -0.087 \\ -0.087 & 0.453 \end{array} \right\}$$

Resultados obtidos em 200 iterações

t	1	30	115	200	(PVO)
\underline{A}_t	1.0	2.75	2.80	2.60	2.00
	1.0	7.30	7.30	6.66	5.00
\underline{W}_t	0.33	0.02	0.03	0.04	0.08
	0.33	0.09	0.10	0.11	0.17
δ_t	0.50	0.95	1.88	0.74	1.0
$\underline{\underline{S}}_t$	2.0 1.0	18.2 26.1	12.2 23.8	9.9 19.8	4.5 9.9
	1.0 2.0	26.1 73.7	23.8 64.4	19.8 52.1	9.9 25.4

Resumo: a) convergência boa para \underline{W}_t e δ_t

b) Caminho subótimo: não determinado no teste

c) Gráfico associado: 10

d) índice de caetos: 37%

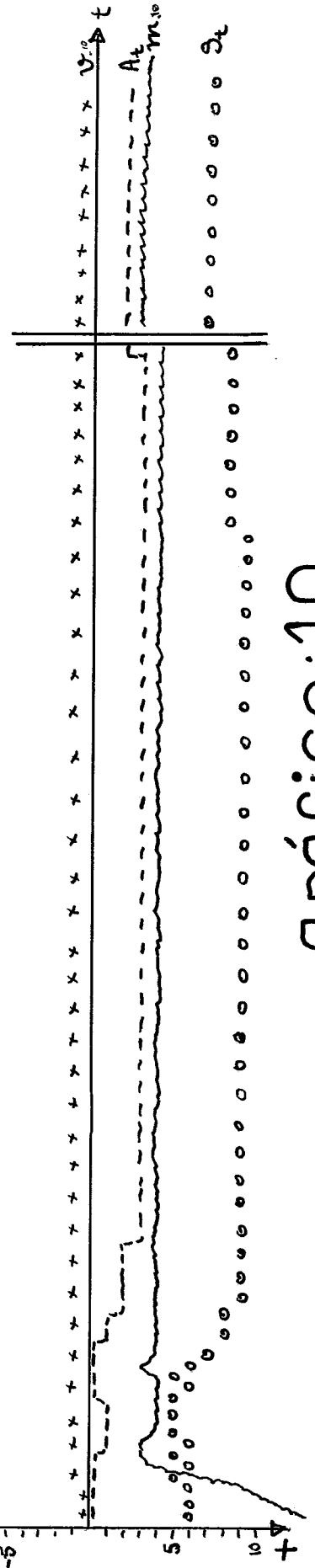
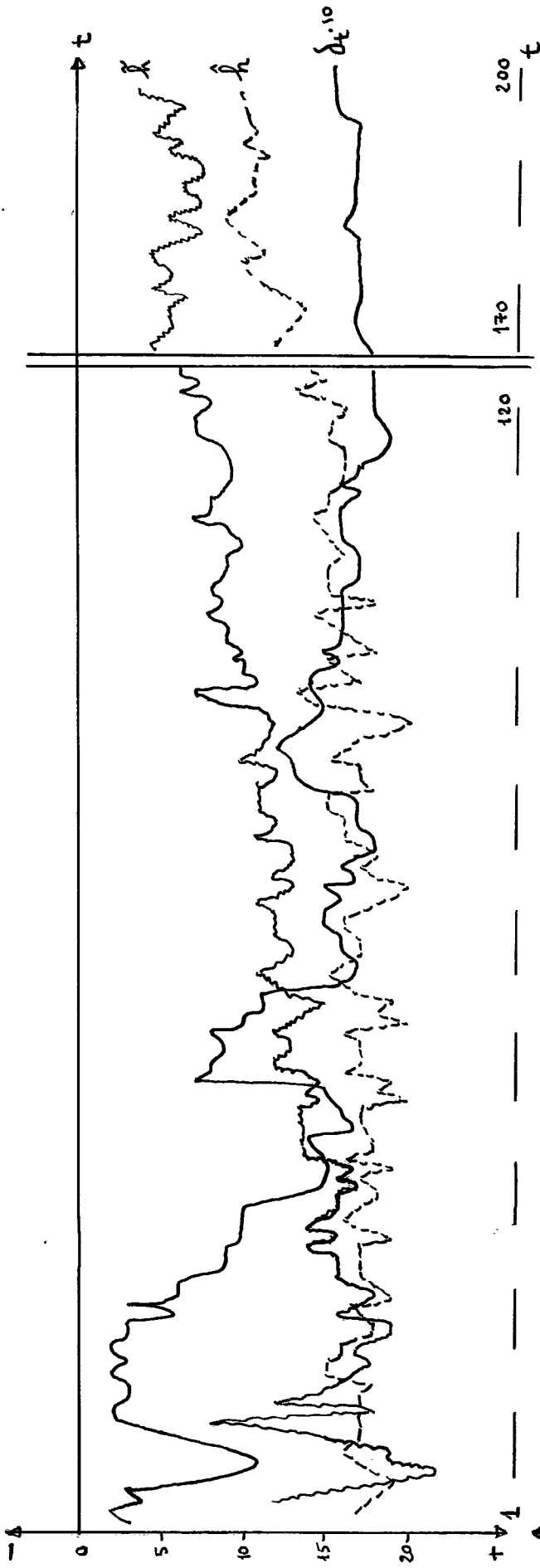


Gráfico:10

TESTE 11

Tentativa: 2

Experimento com IV.32, V.9 e V.14

Tipo de teste: convergência

$$h(1) = 40$$

$$C_i = 2.00$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \begin{Bmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{Bmatrix}$$

Resultados obtidos em 120 iterações

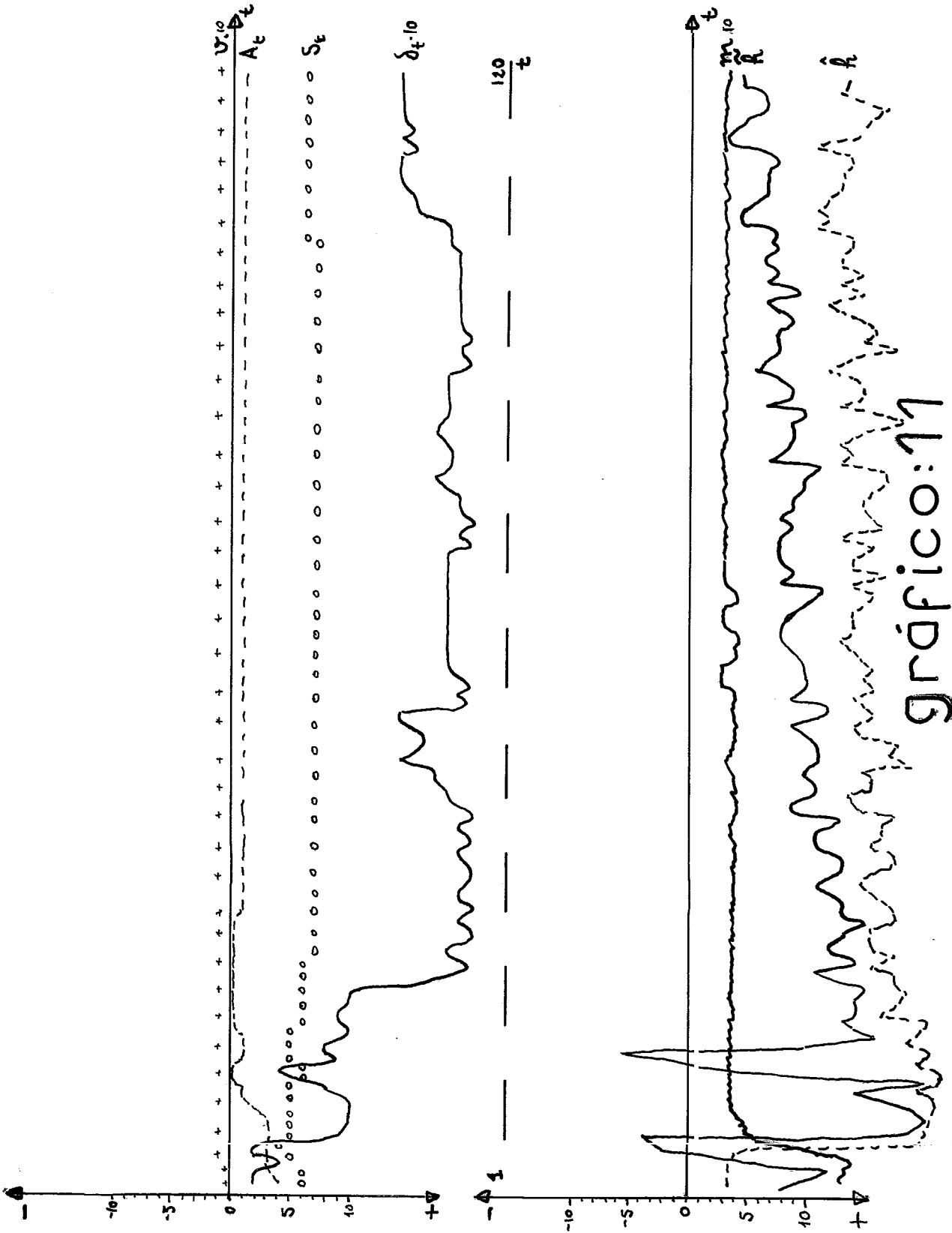
t	1	30	75	120	(PVO)
\underline{A}_t	1.00	4.59	5.16	5.07	4.00
	1.00	3.41	3.70	3.56	3.00
\underline{W}_t	0.33	0.03	0.05	0.06	0.14
	0.33	0.15	0.14	0.14	0.14
δ_t	0.50	2.00	1.97	1.45	1.0
$\underline{\underline{S}}_t$	2.0 1.0	38.7 23.3	37.7 23.4	34.1 21.5	16.6 11.9
	1.0 2.0	23.3 19.6	23.4 18.0	21.5 16.2	11.9 9.5

Resumo: a) convergência apenas razoável para \underline{A}_t , \underline{W}_t e $\underline{\underline{S}}_t$

b) Caminho subótimo: não determinado no teste

c) Gráfico associado: 11

d) Índice de acertos: 39%



TESTE 12

Tentativa: 3

Experimento com IV.22, V.18 e V.19

Tipo de teste: Estabilidade

$$h(1) = 40$$

$$C_i = 2.00$$

$$B_t = 3$$

$$\tilde{h}(n_t) = h \text{ ajustado}$$

$$\underline{\Sigma} = \begin{pmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{pmatrix}$$

Resultados obtidos em 300 iterações

t	1 (PVO)	30	165	300
\underline{A}_t	2.00	1.97	2.09	2.06
	5.00	4.98	5.19	5.16
\underline{W}_t	0.08	0.08	0.06	0.06
	0.17	0.17	0.17	0.17
δ_t	1.00	2.00	2.00	2.00
\underline{S}_t	4.5 9.9	4.4 9.7	5.0 10.7	4.9 10.5
	9.9 25.4	9.7 25.2	10.7 27.3	10.5 26.9

- Resumo: a) Estabilidade muito boa para \underline{A}_t , \underline{W}_t e \underline{S}_t
 b) Caminho subótimo: não determinado no teste
 c) Gráfico associado: 12
 d) Índice de acertos: 43%

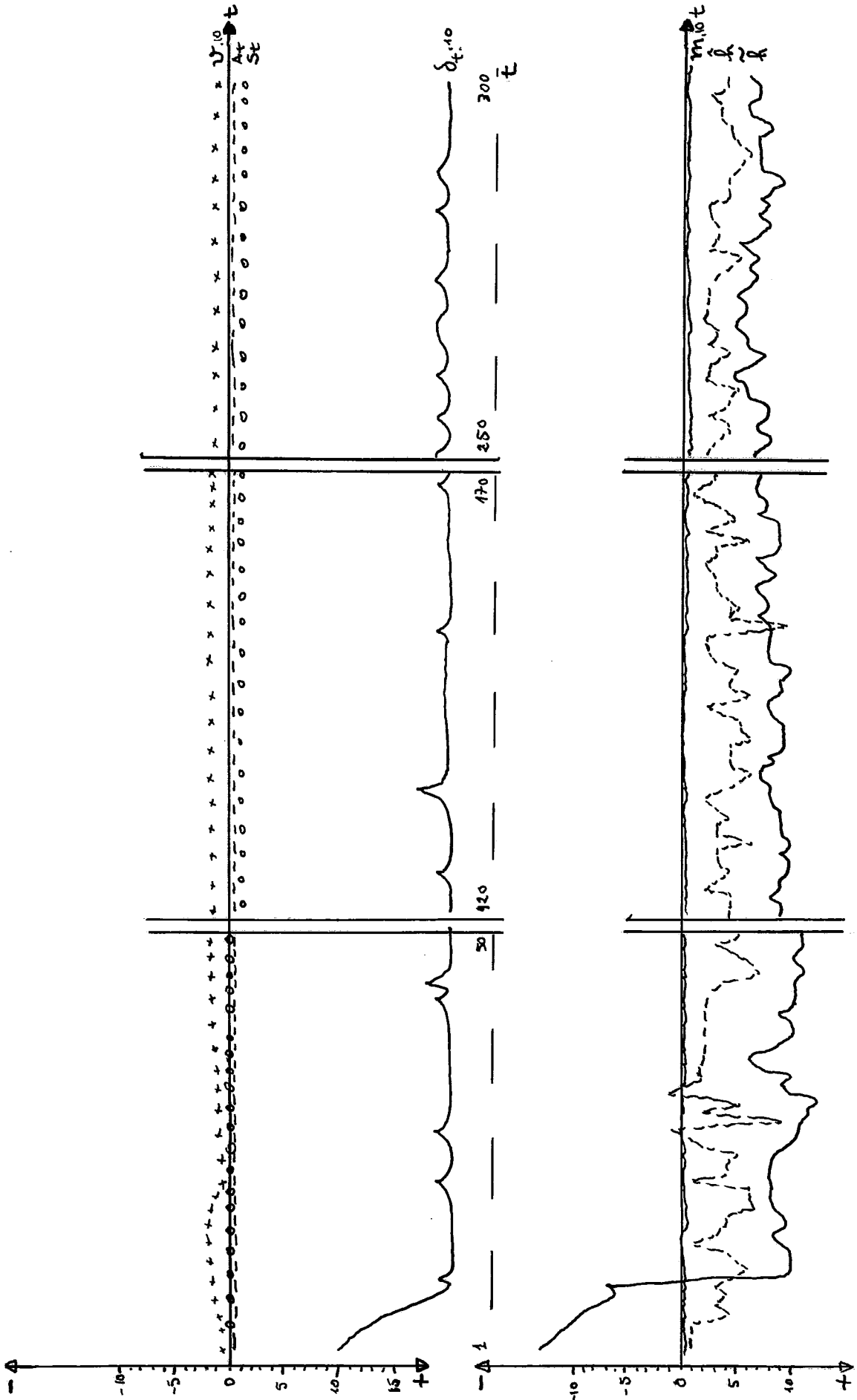


Gráfico:12

T E S T E 13

Tentativa: 3

Experimento com IV.32, V.18 e V.19

Tipo de teste: Estabilidade

$$h(1) = 40$$

$$C_i = 2.00$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{\Sigma} = \begin{Bmatrix} 0.550 & -0.087 \\ -0.087 & 0.550 \end{Bmatrix}$$

Resultados obtidos em 133 iterações, tendo encontrado uma solução.

t	1 (PVO)	30	75	130
\underline{A}_t	2.0	1.94	1.87	1.79
	5.0	4.88	4.77	4.65
\underline{W}_t	0.08	0.08	0.09	0.09
	0.17	0.17	0.17	0.18
δ_t	1.00	0.87	1.23	1.33
\underline{S}_t	4.5 9.7	4.3 9.4	4.1 8.8	3.8 8.2
	9.7 25.4	9.4 29.3	8.8 23.1	8.2 22.0

Resumo: a) Estabilidade boa para \underline{A}_t , \underline{W}_t , δ_t e \underline{S}_t

b) Caminho subótimo: 45

c) Gráfico associado: 13

d) Índice de acertos: 47%

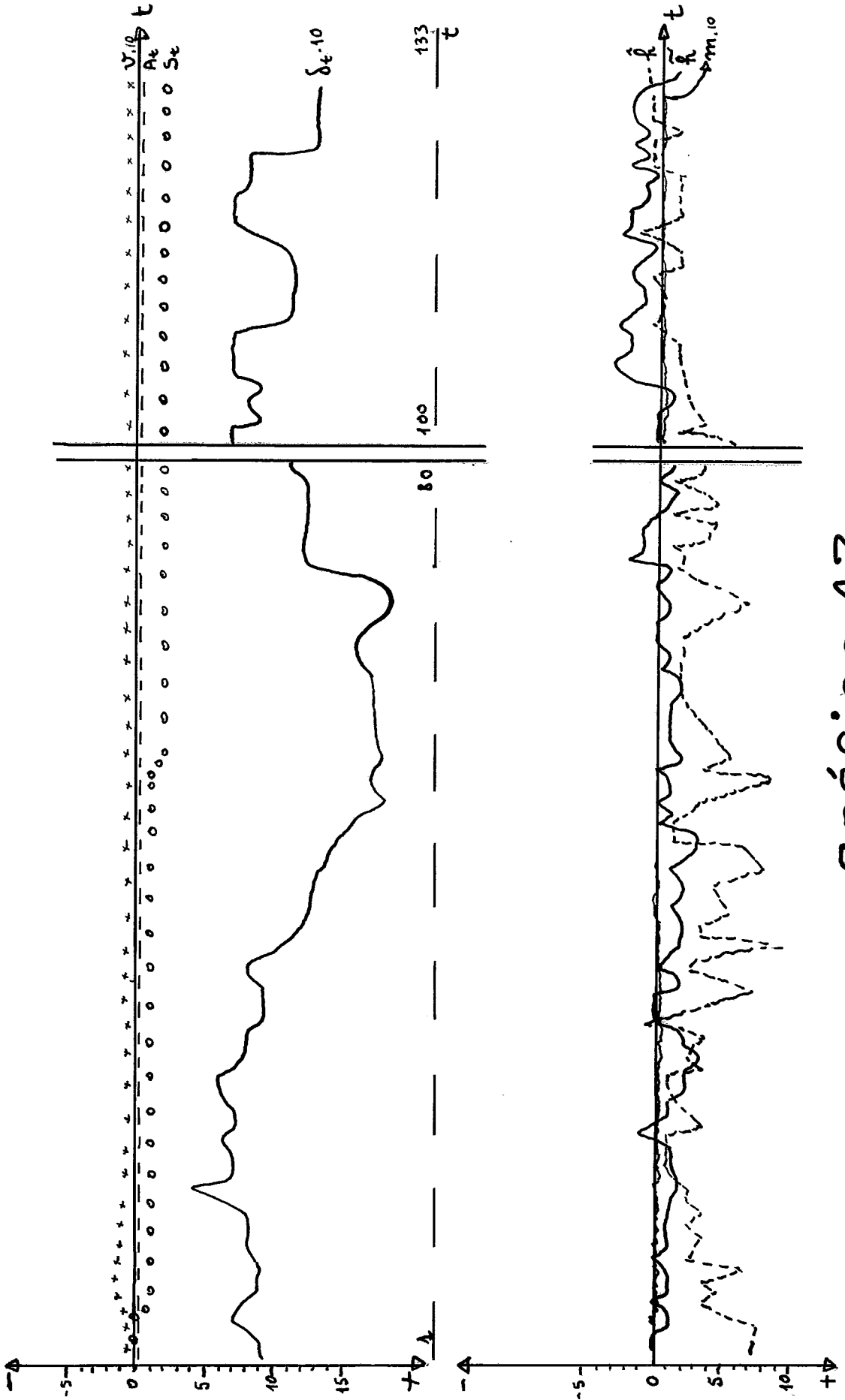


Gráfico:13

T E S T E 14

Tentativa: 3

Experimento com IV.32, V.18 e V.19

Tipo de teste: Estabilidade

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \left\{ \begin{array}{cc} 0.550 & -0.087 \\ -0.087 & 0.453 \end{array} \right\}$$

Resultados obtidos em 120 iterações

t	1 (PVO)		30		75		120	
\underline{A}_t	4.0		3.90		3.85		3.80	
	3.0		2.95		2.90		2.84	
\underline{W}_t	0.15		0.15		0.16		0.16	
	0.13		0.14		0.13		0.14	
δ_t	1.0		0.84		1.86		1.80	
$\underline{\underline{S}}_t$	16.4	11.6	15.8	11.3	15.3	11.0	14.9	10.6
	11.6	9.1	11.3	9.2	11.0	8.9	10.6	8.5

- Resumo: a) Estabilidade bastante boa para \underline{A}_t , \underline{W}_t , δ_t e $\underline{\underline{S}}_t$
 b) Caminho subótimo: não determinado
 c) Gráfico associado: 14
 d) índice de acertos: 43%

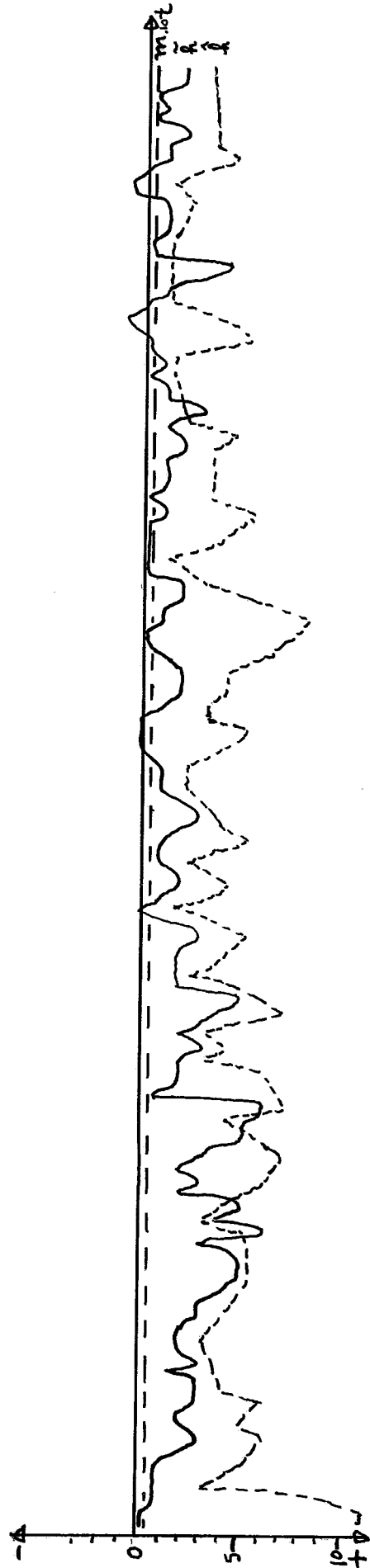
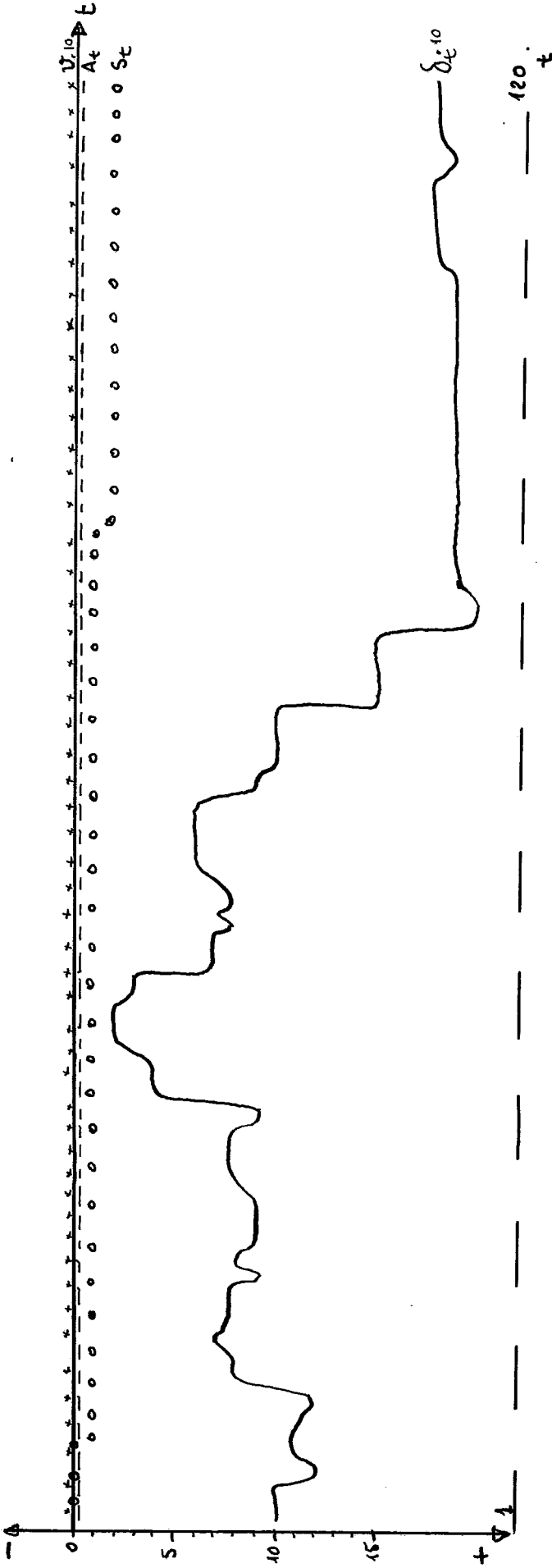


Gráfico 14

T E S T E 15

Tentativa: 3

Experimento com IV.32, V.18 e V.19

Tipo de teste: convergência

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{\Sigma} = \begin{pmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{pmatrix}$$

Resultados obtidos em 200 iterações

t	1	30	115	200	(PVO)
\underline{A}_t	1.0	1.96	1.76	1.66	2.00
	1.0	6.39	5.35	4.97	5.00
\underline{W}_t	0.33	0.06	0.09	0.11	0.08
	0.33	0.09	0.13	0.15	0.17
δ_t	0.55	0.52	0.95	0.20	1.0
\underline{S}_t	2.0 1.0	11.7 13.4	5.7 9.6	4.3 8.3	4.6 9.9
	1.0 2.0	13.4 59.7	9.6 34.7	8.3 28.2	9.9 25.4

Resumo: a) convergência boa para \underline{A}_t , \underline{W}_t e \underline{S}_t

b) Caminho subótimo: não determinado no teste

c) Gráfico associado: 15

d) índice de acertos: 42%

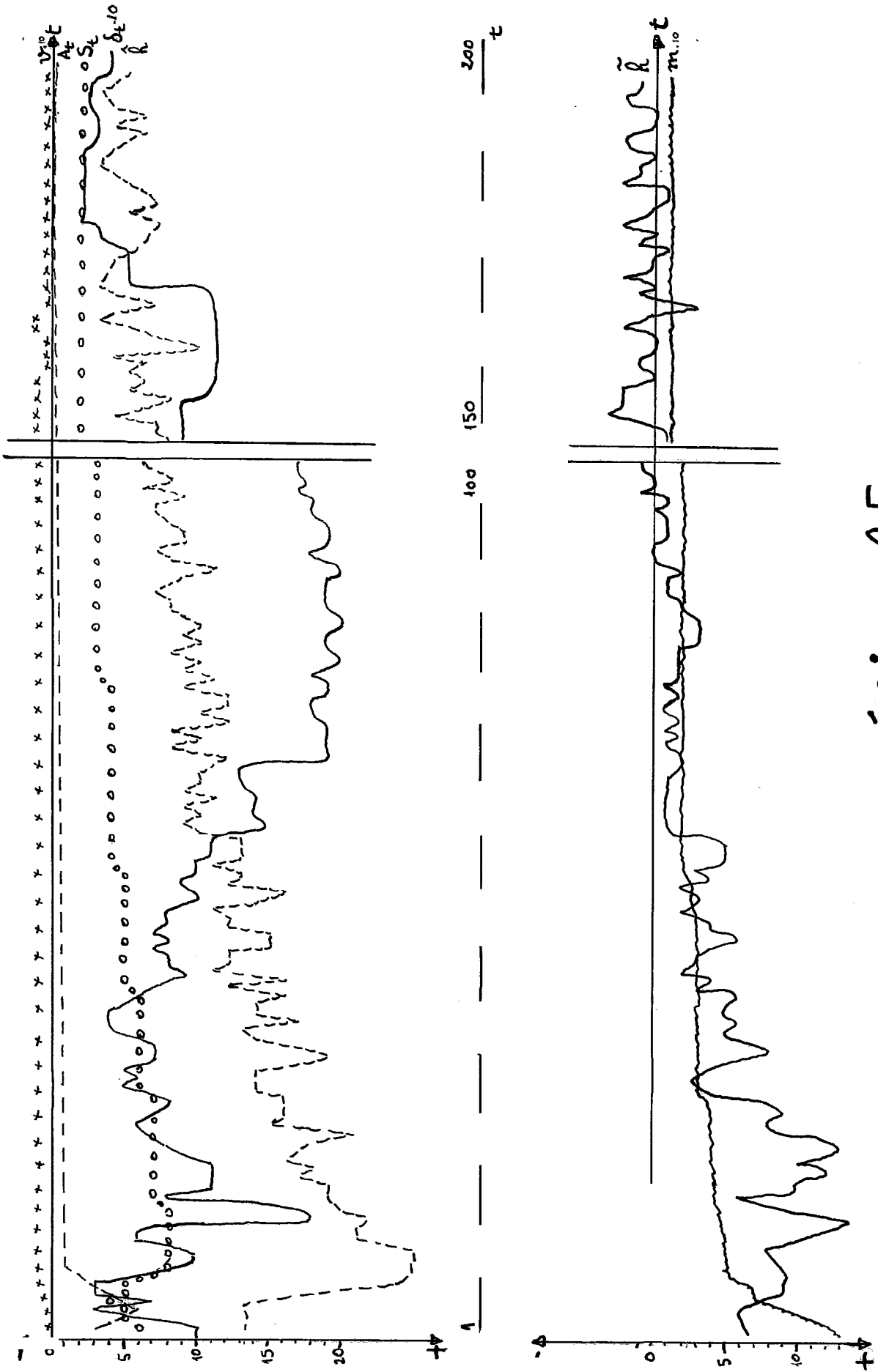


Gráfico:15

T E S T E 16

Tentativa: 3

experimento com IV. 32, V. 18 e V. 19

Tipo de teste: convergência

$$h(1) = 40$$

$$C_i = 2.00$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{\Sigma} = \begin{pmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{pmatrix}$$

Resultados obtidos em 120 iterações

t	1	30	75	120	(PVO)
\underline{A}_t	1.0	5.63	5.50	5.01	4.0
	1.0	4.10	4.01	3.65	3.0
\underline{W}_t	0.33	0.10	0.11	0.12	0.14
	0.33	0.08	0.09	0.09	0.14
δ_t	0.50	1.82	2.00	1.80	1.00
\underline{S}_t	2.0 1.0	38.5 23.9	33.5 22.2	27.7 18.5	16.5 11.9
	1.0 2.0	23.9 22.5	22.2 18.9	18.5 15.6	11.9 9.4

- Resumo: a) Convergência apenas regular para $\underline{A}_t, \underline{W}_t, \delta_t$ e \underline{S}_t
 b) Caminho subótimo: não determinado no teste
 c) Gráfico associado- 16
 d) Índice de caetos: 37%

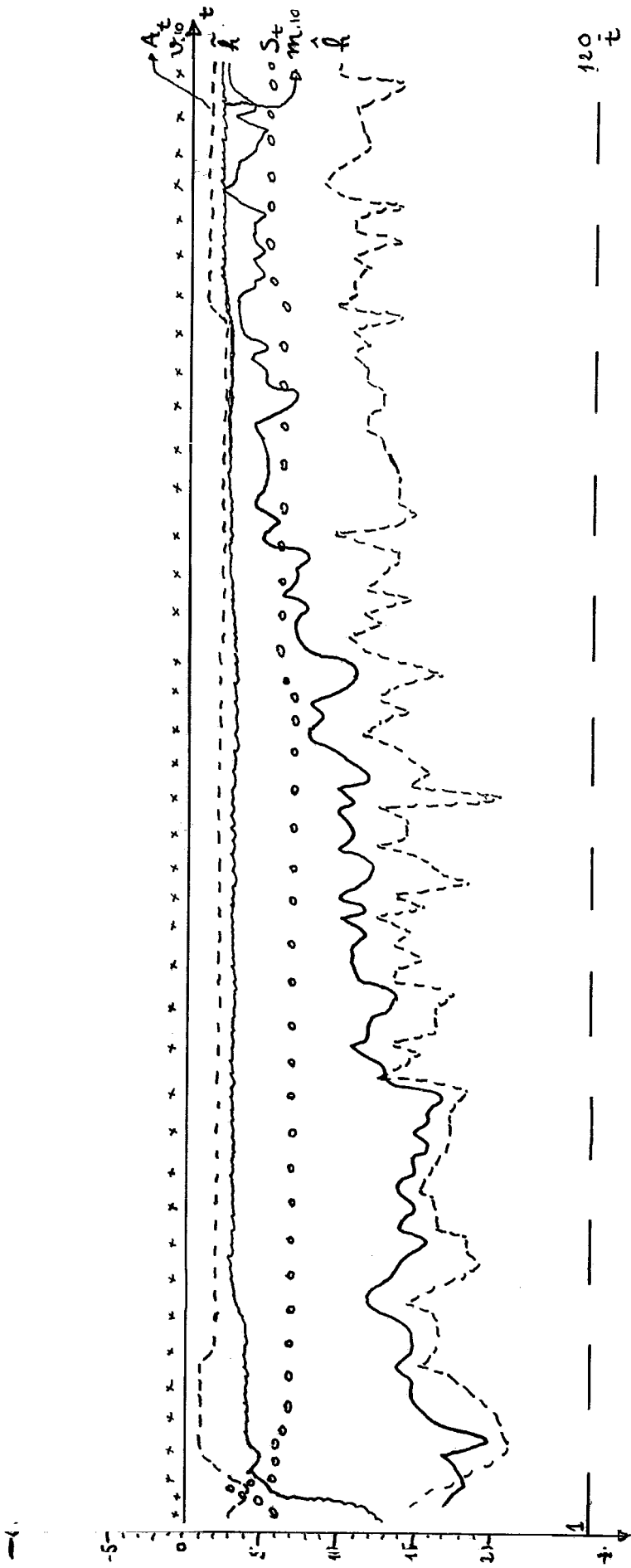


gráfico:16

T E S T E 17

Tentativa: 3

Experimento com IV.22, V.18 e V.19

Tipo de teste: convergência

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \begin{pmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{pmatrix}$$

Resultados obtidos em 200 iterações

t	1	30	115	200	(PVO)
\underline{A}_t	1.0	1.46	1.67	1.75	2.0
	1.0	4.49	5.42	5.50	5.0
\underline{W}_t	0.33	0.14	0.12	0.12	0.08
	0.33	0.15	0.14	0.14	0.17
δ_t	0.50	2.00	2.00	2.00	1.00
$\underline{\underline{S}}_t$	2.0 1.0	2.9 7.2	3.4 9.0	3.6 9.5	4.6 9.9
	1.0 2.0	7.2 24.2	9.0 31.0	9.5 31.4	9.9 25.4

Resumo: a) convergência regular em \underline{A}_t , \underline{W}_t , e boa em $\underline{\underline{S}}_t$

b) Caminho subótimo: não determinado no teste

c) Gráfico associado: 17

d) Índice de acertos: 45%

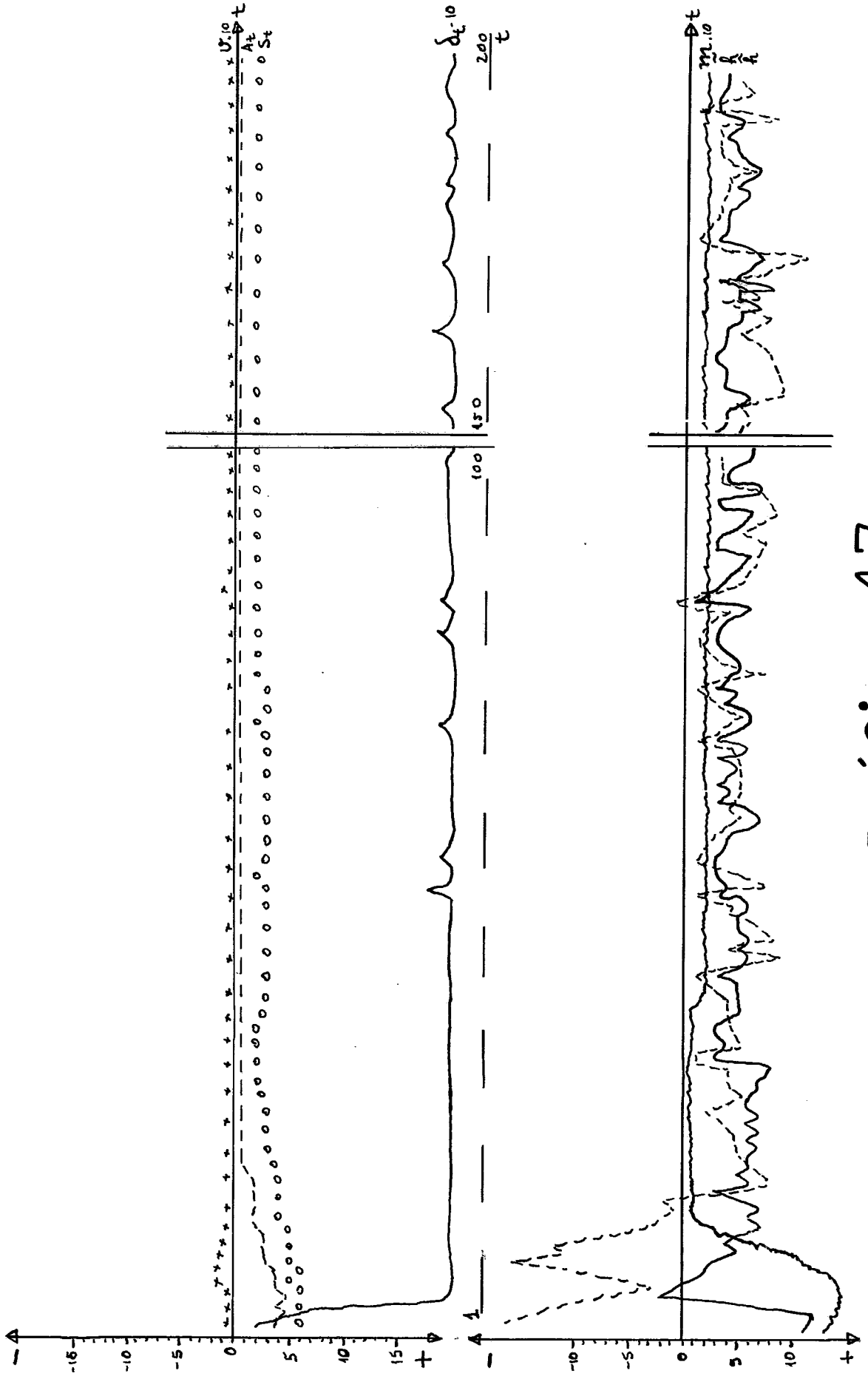


gráfico:17

T E S T E 18

Tentativa: 3

Experimento com IV.22, V.18 e V.19

Tipo de teste: convergência

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = h \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \begin{pmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{pmatrix}$$

Resultados obtidos em 200 iterações, tendo encontrado uma solução.

t	1	30	115	200	(PVO)
\underline{A}_t	1.0	4.24	4.56	4.56	4.00
	1.0	2.62	2.88	2.95	3.00
\underline{W}_t	0.33	0.01	0.09	0.09	0.14
	0.33	0.27	0.18	0.18	0.14
δ_t	0.50	2.00	2.00	2.00	1.00
$\underline{\underline{S}}_t$	2.0 1.0	26.0 14.9	23.5 13.9	22.7 13.7	16.6 11.9
	1.0 2.0	14.9 9.3	13.9 9.2	13.7 9.4	11.9 9.5

Resumo: a) convergência boa para \underline{A}_t e razoável para \underline{W}_t e $\underline{\underline{S}}_t$

b) Caminho subótimo: 44

c) Gráfico associado: 18

d) Índice de acertos: 39%

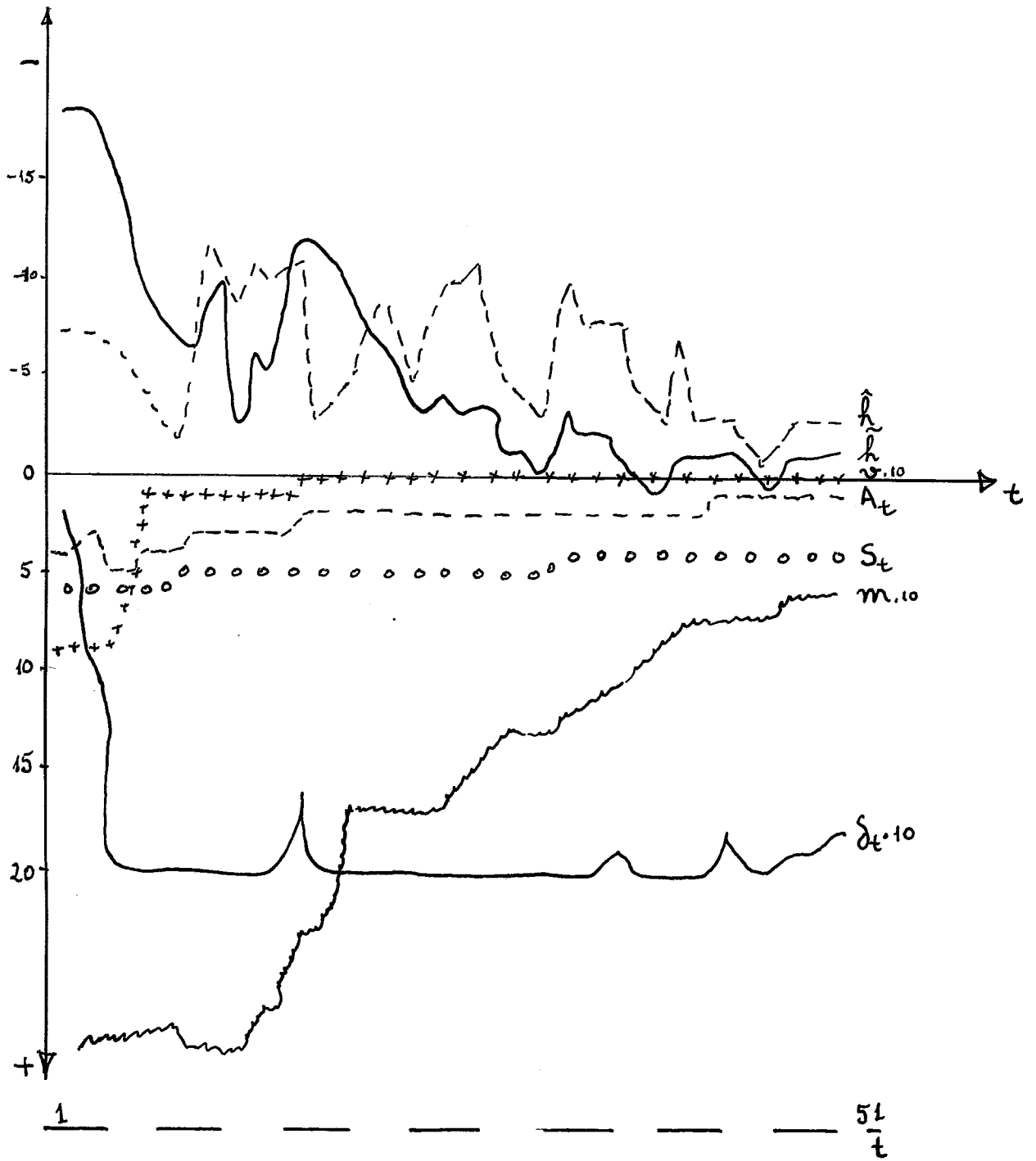


gráfico:18}

T E S T E 19

Tentativa: 4

Experimento com IV.32, V.20 e V.21

Tipo de teste: Estabilidade

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$h(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{\Sigma} = \begin{Bmatrix} 0.550 & -0.087 \\ -0.087 & 0.453 \end{Bmatrix}$$

Resultados obtidos em 141 iterações, tendo encontrado uma solução.

t	1 (PVO)	30	115	140
\underline{A}_t	2.00	2.11	2.25	2.16
	5.00	4.96	4.66	4.47
\underline{W}_t	0.08	-0.18	-0.40	-0.56
	0.17	0.30	0.44	0.52
δ_t	1.0	0.69	1.29	0.78
\underline{S}_t	4.6 9.9	5.3 10.2	5.9 10.5	6.0 10.5
	9.9 25.5	10.2 22.7	10.5 20.1	10.5 19.8

Resumo: a) Estabilidade muito boa para \underline{A}_t , δ_t e \underline{S}_t , mas \underline{W}_t instável e longe de \underline{W}_* .

b) Caminho subótimo: 47

c) Gráfico associado: 19

d) Índice de acertos: 43%

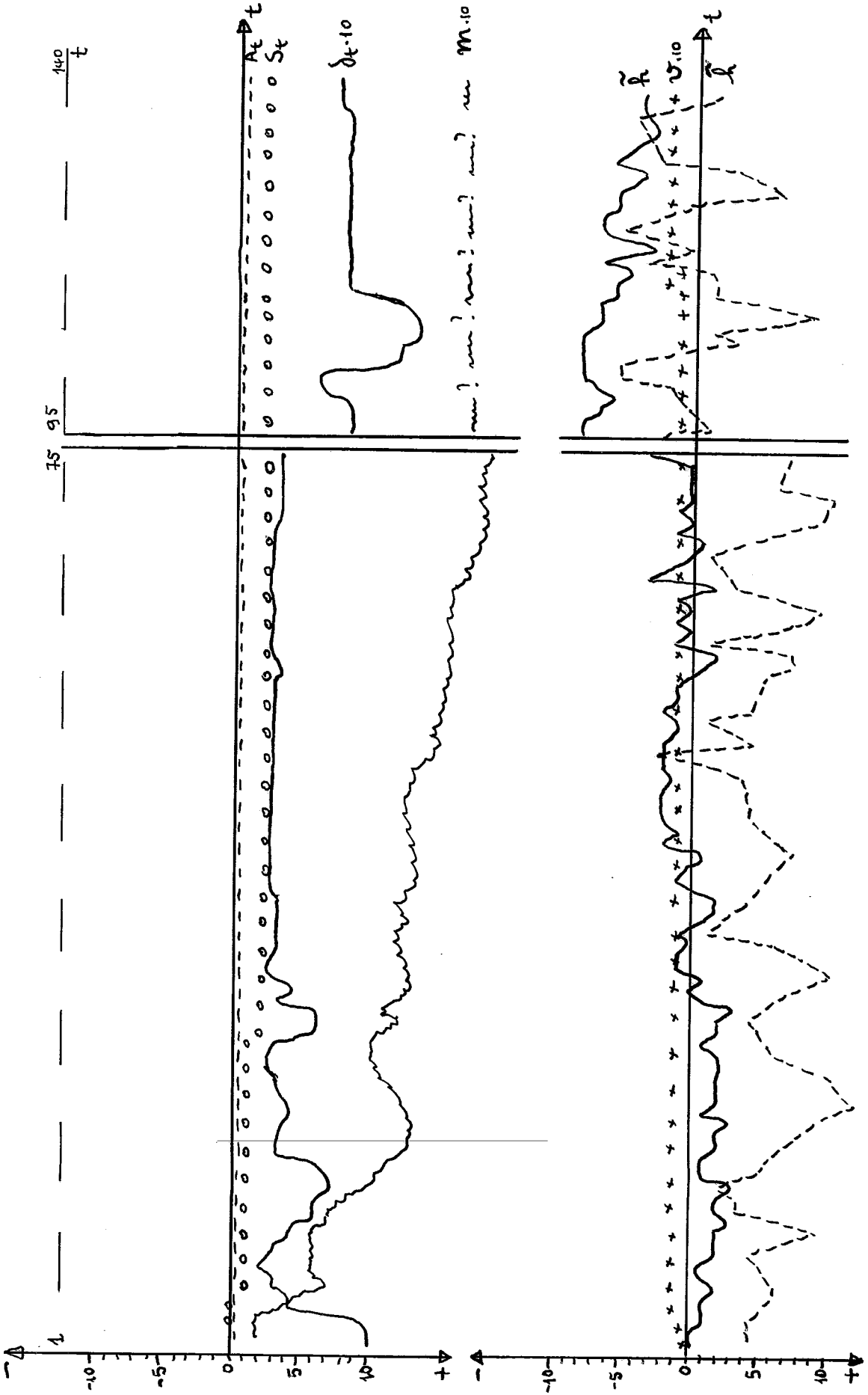


Gráfico:19

TESTE 20

Tentativa: 4

Experimento com IV.32, V.20 e V.21

Tipo de teste: Estabilidade

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{\Sigma} = \left\{ \begin{array}{cc} 0.550 & -0.087 \\ -0.087 & 0.453 \end{array} \right\}$$

Resultados obtidos em 120 iterações

t	1 (PVO)		30		75		120	
\underline{A}_t	4.0		3.91		4.01		4.20	
	3.0		2.91		3.04		3.22	
\underline{W}_t	0.14		0.22		0.19		0.20	
	0.14		0.04		0.11		0.12	
δ_t	1.0		0.56		0.72		1.20	
\underline{S}_t	16.6	11.9	15.5	11.4	14.9	11.1	14.6	11.0
	11.9	9.5	11.4	9.0	11.1	8.7	11.0	8.6

- Resumo: a) convergência muito boa para \underline{A}_t , \underline{W}_t , δ_t e \underline{S}_t
 b) Caminho subótimo: não determinado no teste
 c) Gráfico associado: 20
 d) índice de acertos: 41%

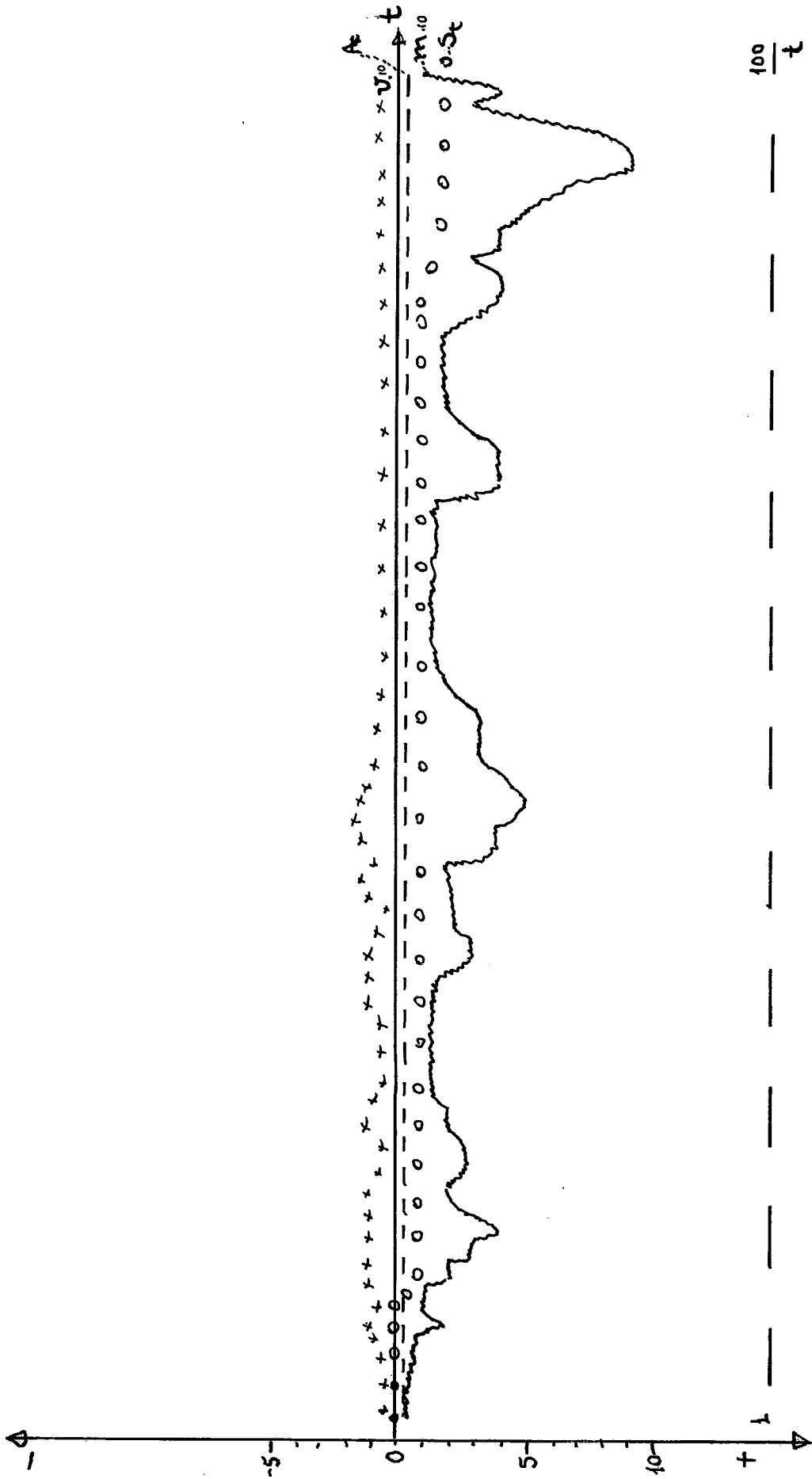


Gráfico:20

T E S T E 21

Tentativa: 4

Experimento com IV. 32, V. 20 e V. 21

Tipo de teste: convergência

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \left\{ \begin{array}{cc} 0.550 & -0.087 \\ -0.087 & 0.453 \end{array} \right\}$$

Resultados obtidos em 34 iterações, tendo encontrado uma solução.

t	1	10	20	30	(PVO)
\underline{A}_t	1.0	1.42	2.05	2.78	2.00
	1.0	4.75	6.45	8.74	5.00
\underline{W}_t	0.33	0.09	0.33	0.67	0.08
	0.33	0.18	0.35	0.62	0.17
δ_t	0.50	1.30	1.06	0.80	1.00
	2.0 1.0	45.0 -18.6	38.4 -31.3	36.2 -35.5	4.6 9.9
	1.0 2.0	-18.6 34.8	-31.3 47.8	-35.5 52.2	9.9 25.5

Resumo: a) Convergência boa **apenas** para δ_t .

b) Caminho subótimo: prejudicado

c) Gráfico associado: 21

d) Índice de acertos: 56%

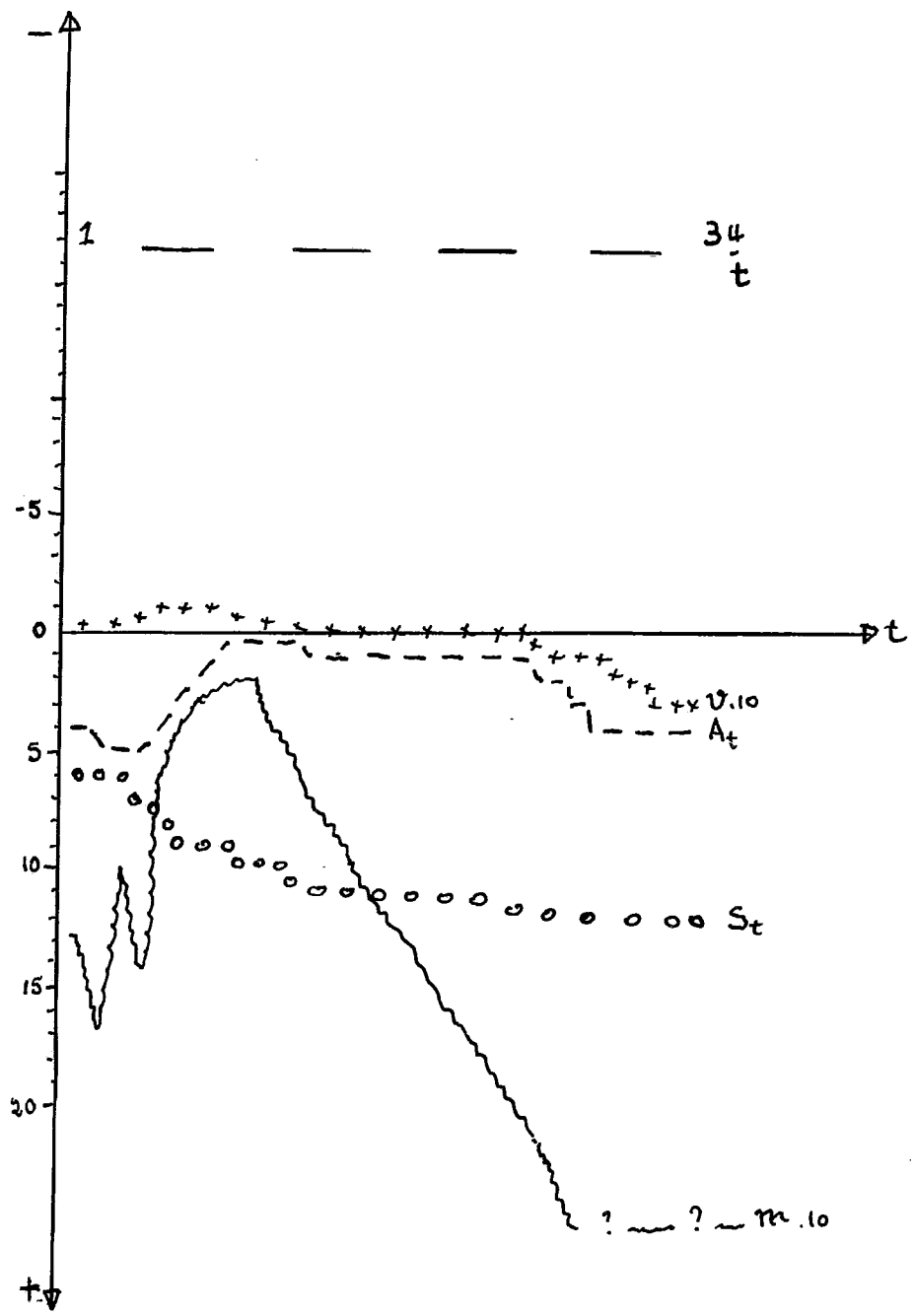


gráfico:21

[Faint handwritten text]

T E S T E 22

Tentativa: 4

Experimento com IV.22, V.20, e V.21

Tipo de teste: convergência

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = h \text{ ajustado}$$

$$\underline{\underline{\Sigma}} = \left\{ \begin{array}{cc} 0.550 & -0.087 \\ -0.087 & 0.453 \end{array} \right\}$$

Resultados obtidos em 140 iterações tendo encontrado uma solução.

t	1	30	100	140	(PVO)
\underline{A}_t	1.0	2.86	3.60	3.75	4.0
	1.0	1.90	2.50	2.64	3.0
\underline{W}_t	0.33	0.07	0.13	0.14	0.14
	0.33	0.34	0.19	0.17	0.14
δ_t	0.50	2.00	2.00	2.00	1.00
$\underline{\underline{S}}_t$	2.0 1.0	10.4 6.3	14.2 9.5	18.8 12.0	16.6 11.9
	1.0 2.0	6.3 4.3	9.5 6.9	12.0 8.4	11.9 9.5

Resumo: a) convergência boa para \underline{A}_t , \underline{W}_t e $\underline{\underline{S}}_t$

b) Caminho subótimo: prejudicado

c) Gráfico associado: 22

d) índice de acertos: 46%

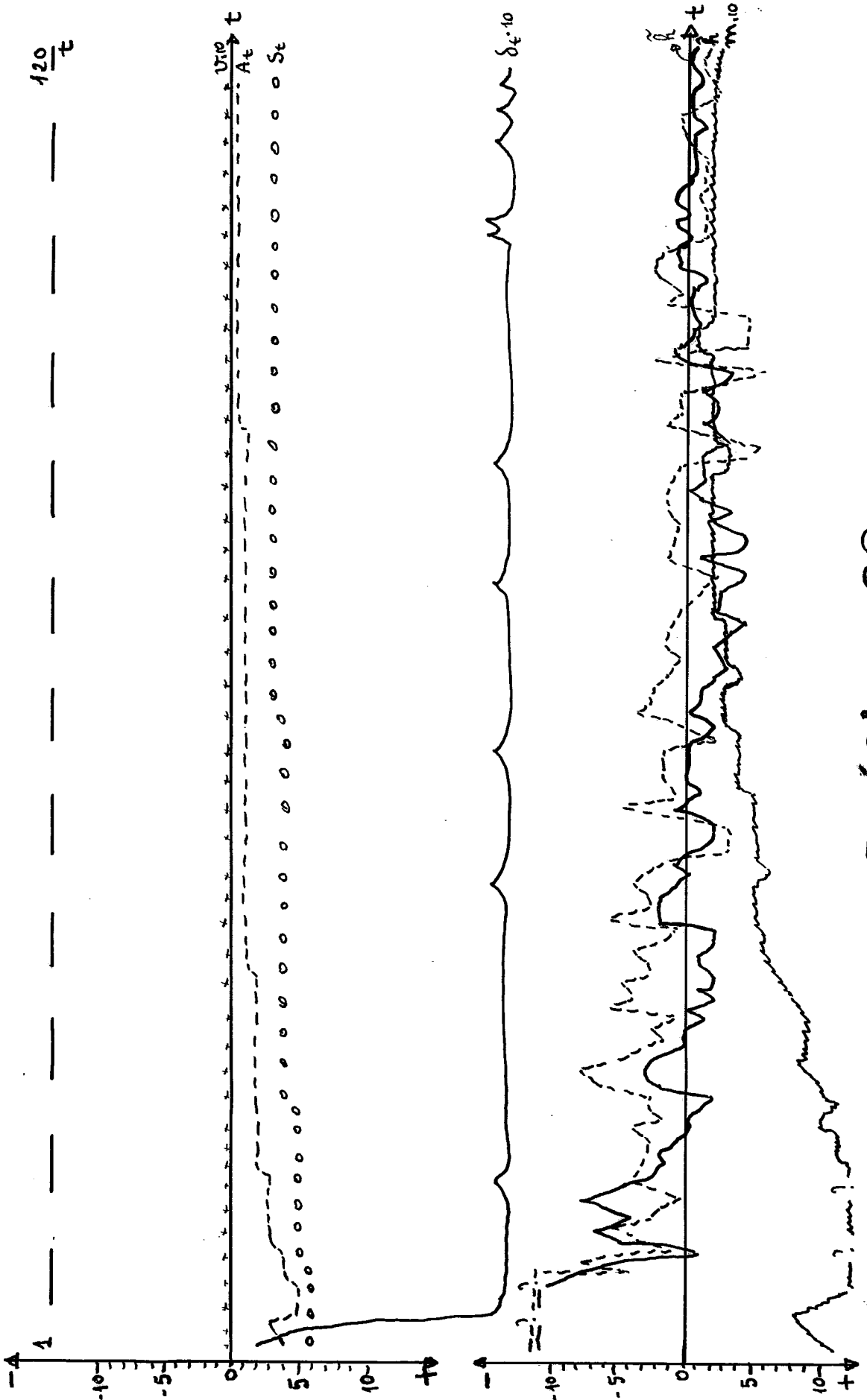


Gráfico: 22

TESTE 23

\underline{S}_t : Matriz Escalar

\underline{A}_t : Calculado pela tentativa 3

Tipo de teste: convergência

Tipo de busca: Expansão dirigida

$\bar{\sigma}^2$: Média da Diagonal principal de $\underline{\Sigma}$.

\underline{C}_S usados: Sigma 1 a 10 dos Dados (Anexo 2).

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\bar{\epsilon} = \epsilon = \bar{\sigma}^2 \text{ ou } k\bar{\sigma}^2 \quad (k \text{ constante})$$

Os valores de \hat{h} sô foram atualizados até a expansão de 25 nós.

$$\text{Cálculo de } \underline{W}_t: \underline{W}_t = \frac{1}{\|\underline{A}_t\|^2 + \bar{\epsilon}}$$

Resultados obtidos de 10 experimentos com $\underline{\Sigma} \neq \underline{\epsilon}I$

E	$\bar{\sigma}^2$	\underline{A}	t	\underline{A}_t^T	\underline{W}_t^T	g_*	\underline{W}_*
1	0.5	2 4 3 5	177	0.9 2.1 1.6 2.6	0.06 0.14 0.10 0.17	41	0.02 0.10 0.06 0.08
2	0.5		185	3.2 6.7 5.0 8.3	0.02 0.04 0.03 0.06	42	
3	0.3		195	2.9 6.1 4.4 7.5	0.02 0.05 0.04 0.06	45	
4	1.0		195	1.6 3.7 2.6 4.2	0.04 0,09 0.06 0.10	45	
5	0.3	2 4 3	185	11.7 24.3 19.5	0.01 0.02 0.02	43	-0.09 0.19 0.13
6	0.7		185	9.2 19.3 15.8	0.01 0.03 0.02	43	
7	0.5		185	4.0 8.4 7.3	0.03 0.06 0.05	44	
8	0.3	2 5	195	2.6 6.6	0.05 0.13	45	0.08 0.17
9	0.5		200	3.0 7.6	0.05 0.11	46	
10	1.0		195	2.2 5.6	0.06 0.15	45	

- Resumo:
- Em geral boa aproximação de \underline{W}_t com \underline{W}_*
 - \underline{A}_t manteve-se em geral numa relação $\approx k\bar{A}$
 - índice médio de acertos: 45.5%
 - Média dos melhores caminhos: $g_* = 44$
 - Gráfico associado: 23

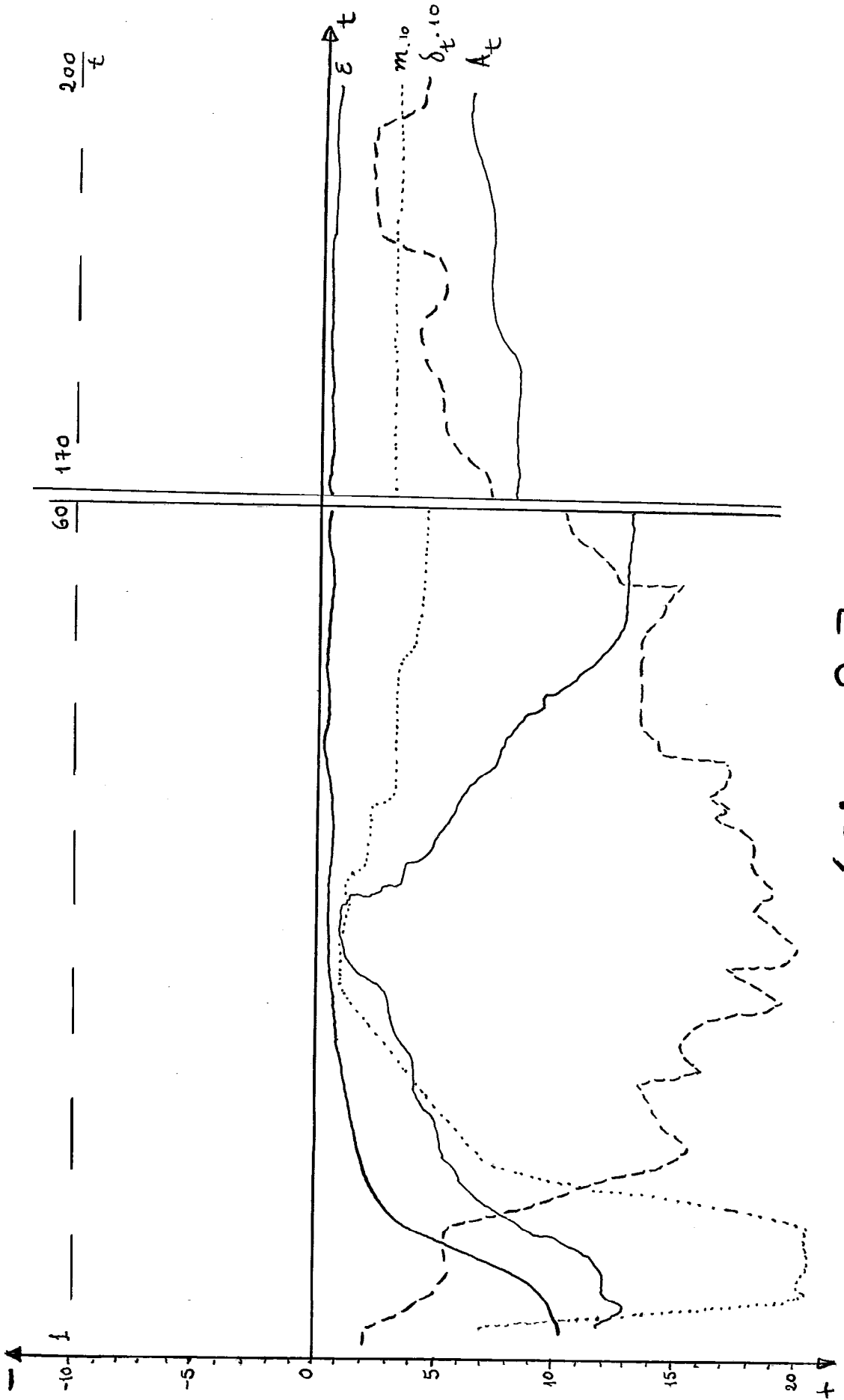


Gráfico: 23

T E S T E 24

Tipo de teste: Estabilidade

O teste foi efetuado segundo a tentativa 4, corrigindo $\underline{\underline{S}}_t$ de tal forma que

$$\underline{\underline{S}}_t^{ii} > \left(\underline{\underline{A}}_t^i\right)^2, \quad \forall i.$$

Fator de correção por linha e coluna $FC_i = \left(\underline{\underline{A}}_t^i\right)^2 + \varepsilon / \underline{\underline{S}}_t^{ii}$.

$$\varepsilon = 0.5$$

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

Tabela resumo de experimento

Exp.	1	2	3	4	
t	200	164	200	200	(PVO)
$\underline{\underline{A}}_t$	2.08	1.82	2.18	2.68	2.00
	5.20	4.58	5.46	6.78	5.00
$\underline{\underline{W}}_t$	0.09	0.05	0.07	0.05	0.08
	0.16	0.20	0.16	0.13	0.17
$\underline{\underline{S}}_t$	4.57 10.65	3.76 8.19	5.86 11.30	10.63 16.55	$\frac{\underline{\underline{A}}\underline{\underline{A}}^T + \underline{\underline{S}}}{\underline{\underline{A}}\underline{\underline{A}}^T + \underline{\underline{S}}}$
	10.65 27.00	8.19 21.02	11.30 29.78	16.55 45.96	
$\underline{\underline{\Sigma}}$	0.27 -0.04	0.55 -0.09	1.10 -0.17	2.20 -0.35	
	-0.04 0.23	-0.09 0.45	-0.17 0.91	-0.35 1.81	
g	Prej.	43	Prej.	Prej.	40
h-i;	1	-3.5	5	7.5	

Resumo: a) Prej.: prejudicado e Exp: Experimento

b) Estabilidade de $\underline{\underline{A}}_t$: razoável

c) Gráfico associado: 24

d) Índice de acertos: Prej.

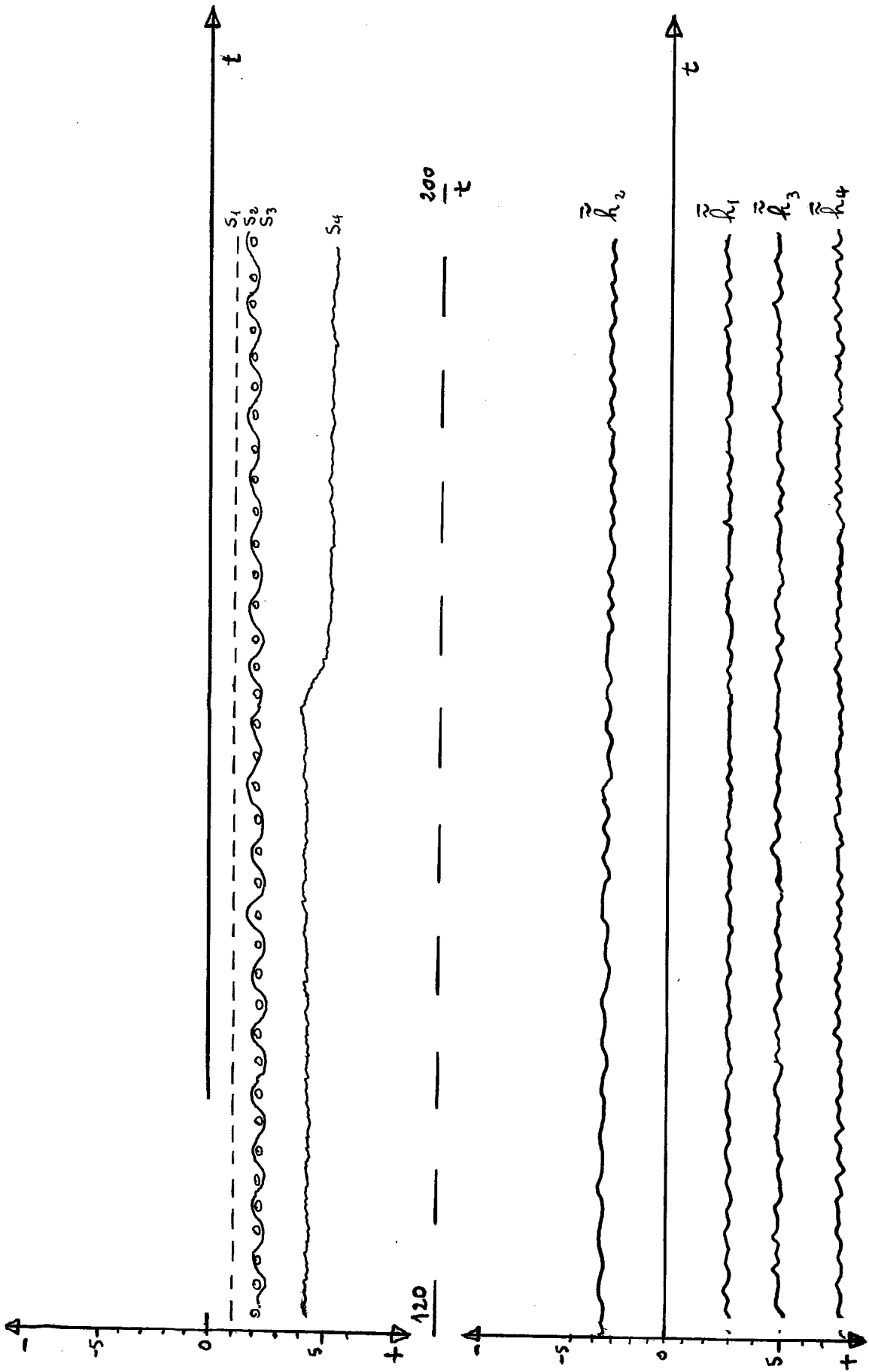


Gráfico: 24

T E S T E 25

Tipo de teste: convergência

Teste efetuado com dimensão 4 para \underline{A}_t , \underline{S}_t e \underline{W}_t (dimensão=componentes)

Tentativa: 0

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3.0$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

$$\underline{C} = \text{Sigma '4' dos Dados (Anexo 2).}$$

Tabela resumo do Experimento

t	\underline{A}_t	\underline{W}_t	\underline{S}_t			
30	2.23	0.03	6.25	8.67	7.32	11.83
	4.99	0.14	8.67	21.93	14.29	24.48
	3.49	0.06	7.32	14.29	11.38	18.02
	5.87	0.03	11.83	24.48	18.02	31.37
300	2.24	0.02	6.25	9.67	7.72	12.40
	4.47	0.11	9.67	20.39	14.59	24.39
	3.30	0.05	7.72	14.59	11.45	18.17
	5.53	0.06	12.40	24.39	18.17	31.33
PVO	2.00	0.02	5.00	7.90	6.30	10.10
	4.00	0.10	7.90	16.60	12.10	20.10
	3.00	0.06	6.30	12.10	9.50	15.10
	5.00	0.07	10.10	20.10	15.10	26.00

Resumo: a) convergência boa para \underline{A}_t , \underline{W}_t e \underline{S}_t

b) Gráfico associado: 25

c) Índice de acertos: prejudicado

d) Observe-se que na iteração 30, \underline{W}_t tinha atingido a mesma precisão da iteração 300.

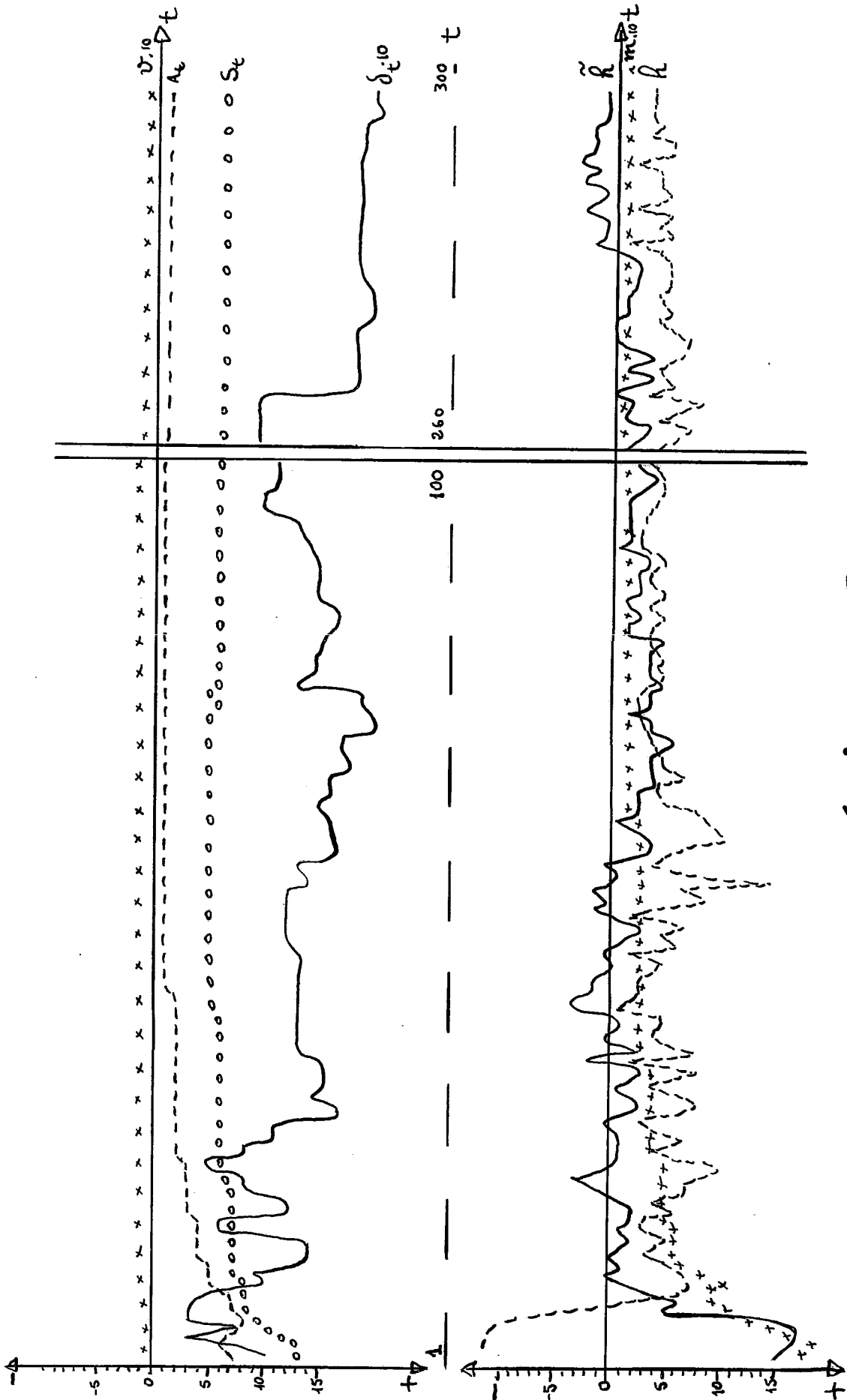


Gráfico:25

TESTE 26

Tipo de teste: convergência

Rodado com os recursos da tentativa 0

Dimensão 3 para \underline{A}_t , \underline{W}_t e \underline{S}_t

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = \tilde{h} \text{ ajustado}$$

\underline{c} : Sigma '5' dos Dados (anexo 2).

Tabela resumo do Experimento

t	\underline{A}	\underline{W}_t	\underline{S}_t		
5	0.58	-0.02	3.40	3.96	0.45
	1.45	0.64	3.96	7.40	1.83
	3.04	0.65	0.45	1.83	8,83
200	2.37	0.08	6.28	12.10	7.89
	4.73	0.08	12.10	23.76	16.23
	3.56	0.12	7.89	16.23	14.17
300	2.33	0.05	5.91	11.48	7.61
	4.64	0.10	11.48	22.65	15.56
	3.46	0.11	7.61	15.56	13.16
PVO	2.00	-0.09	4.15	8.23	5.84
	4.00	0.19	8.23	16.40	11.72
	3.00	0.13	5.84	11.72	9.40

Resumo: a) Convergência boa para \underline{A}_t e \underline{S}_t

b) Não há gráfico associado ao experimento

c) Índice de acertos e caminho subótimo prejudicados.

T E S T E 27

Este teste foi rodado com θ grande ($\theta \geq 1.5$), com os recursos da tentativa 0 do capítulo V.

Tipo de teste: convergência

$$M = 5 \quad (\text{Média de 5 valores de } \underline{\underline{S}}_t)$$

$$h(1) = 40$$

$$C_i = 2.0$$

$$B_t = 3$$

$$\tilde{h}(n_t) = h \text{ ajustado}$$

$$\underline{\underline{C}} = \text{Sigma '3' dos Dados (Anexo 2).}$$

Tabela resumo do Experimento

Teste	θ	t	$\underline{\underline{A}}_t$	$\underline{\underline{W}}_t$	$\underline{\underline{S}}_t$			
a	1.5	30	2.12	-0.79	1.28	1.24	1.06	1.55
			4.84	1.25	1.24	2.71	1.58	2.35
			3.26	0.18	1.06	1.58	1.70	1.96
			5.41	0.83	1.55	2.35	1.96	3.55
b	2.0	25	1.62	-0.41	1.54	0.70	0.93	1.76
			3.33	0.86	0.70	3.93	1.75	0.35
			2.46	-0.09	0.93	1.75	1.81	1.51
			4.45	0.97	1.76	0.35	1.51	5.11
PVO			2.00	0.02	4.33	7.97	6.10	10.03
			4.00	0.10	7.97	16.20	12.03	20.03
			3.00	0.06	6.10	12.03	9.17	15.03
			5.00	0.08	10.03	20.03	15.03	25.30

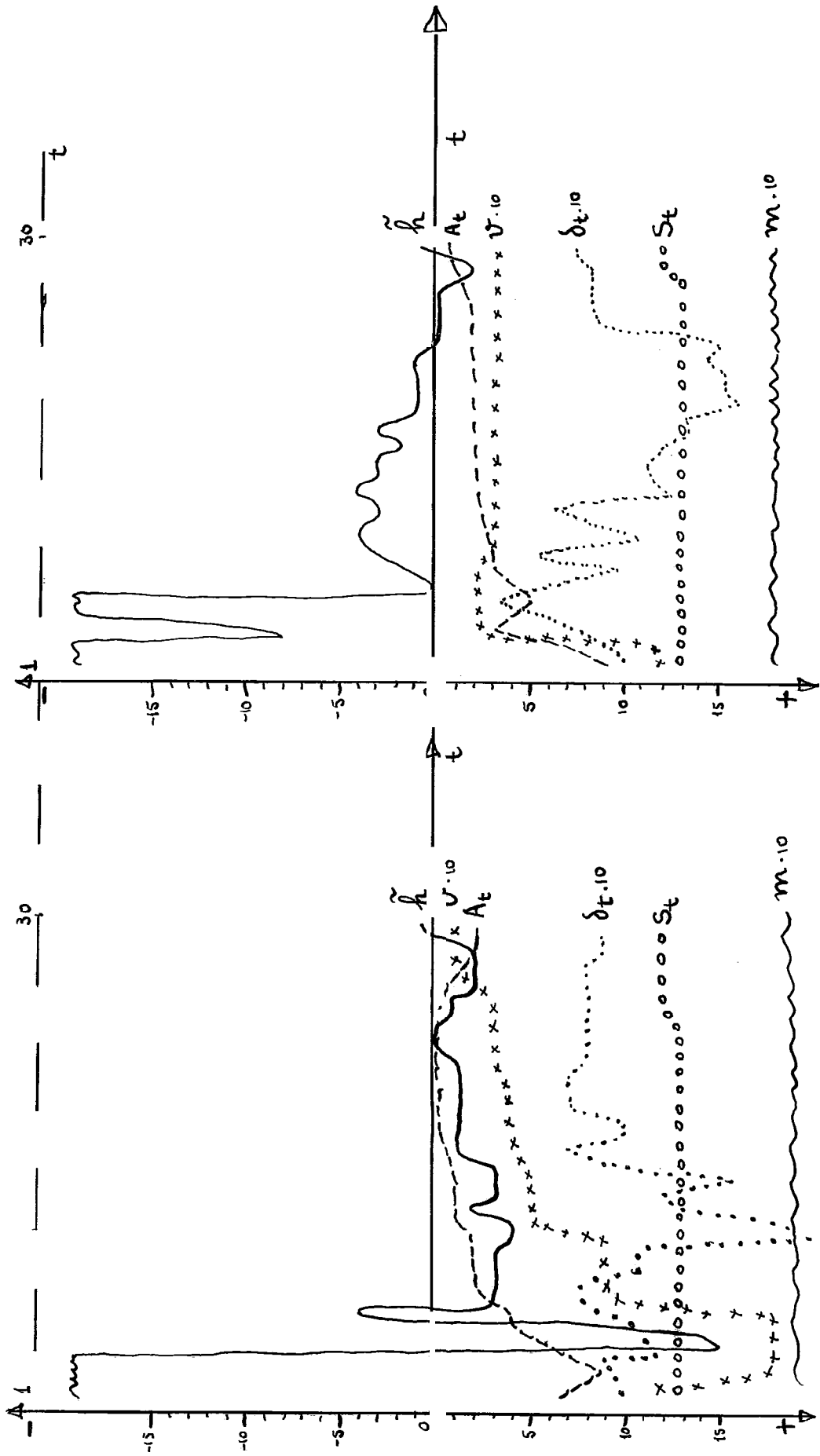
Resumo: a) $\underline{\underline{A}}_t$ convergiu razoavelmente

b) $\underline{\underline{S}}_t$ teve um viés acentuado, exigindo correção

c) $\underline{\underline{W}}_t$ não se aproxima de $\underline{\underline{W}}_*$

d) índice de acertos: prejudicado

e) Gráfico associado: 27



b) $\theta = 2.0$

a) $\theta = 1.5$

Gráfico: 27

REFERÊNCIAS BIBLIOGRÁFICAS

- 1 - ARAUJO, J. L. - Buscas Heurísticas e Processos Estocásticos, I Simpósio Nacional de Probabilidade e Estatística, Instituto de Matemática da UFRJ, Julho de 1974, pp.16-28.
 - 2 - ARAUJO, J. L. - O Problema Probabilístico do Caminho Mínimo: Uma Abordagem Através da Teoria de programação Heurística, Anais da SOBRAPO, 1974, (U01-U14).
 - 3 - ARAUJO, J. L. - Otimização de Heurísticas, VIII Simpósio de Pesquisa Operacional - Setembro de 1975.
 - 4 - ARAUJO, J. L. - Evaluating and Accelerating Heuristic Search: A Proposal, Jan. 1974, Publicação Técnica COPPE/UFRJ, Nº 1.74.
 - 5 - ARAUJO, J. L. - Anotações do Curso "Heurística 11", Junho-Setembro de 1978.
 - 6 - ARAUJO, J. L. - Eficiência de Algoritmos Adaptativos Admissíveis, XI Simpósio de Pesquisa Operacional - Outubro de 1978.
- *
- 7 - BERLINER, H. - The B Tree Search Algorithm: A Best-First Proof Procedure, Artificial Intelligence, 12(1979), 23-40.
 - 8 - DORAN, J., Michie, D. - Experiments with the Graph Traverser Program, Proc. Roy. Soc. A(1437) (Sept 1966), 235-259.

- 9 - DEMIDOVICH, B. P. and Maron, I. A. - Computational Mathematics, MIR Publishers, Moscou, 1973.
- 10 - FABIAN, V. - Stochastic Approximation of Minima with Improved Asymptotic Speed, Ann. Math. Stat., 38,n 1, Feb. 1967.
- 11 - GELPERIN, D. - On the Optimality of A ^{*}, Artificial Intelligence, 8 (1977), 69-76.
- 12 - GOMEZ, J. G. - Uma nova Formulação do Problema Clássico de Balanceamento de Linhas de Montagem e sua Solução ótima Eficiente, Tese de Doutorado, COPPE/UFRJ, 1975.
- 13 - GRAUPE, D. - Identification of Systems, Van Nostrand Reinhold Company, 1972.
- 14 - HARRIS, L. R. - The Heuristic Search under Conditions of Error, Artificial Intelligence 5(1974), 217-234.
- 15 - HART, P., Nilsson, N. and Raphael, B. - A Formal Basis for the Heuristic Determination of Minimum Cost Paths, IEEE Trans. System Sci Cybernetics (2)4 (July-1968), 100-107.
- 16 - MARTELLI, A. - On the Complexity of Admissible Search Algorithms, Artificial Intelligence, 8(1977), 1-13.
- 17 - MENDEL, J. M. - Adaptive Learning and Pattern Recognition Systems, Academic Press, 1970.
- 18 - MICHIE, D. and Ross, R. - Experiments with the Adaptive Graph Traverser, Machine Intelligence, 5(1969), 301-318.

- 19 - MILTON, R. C. and Nelder, J. A. - Statistical Computation, Academic Press, 1969.
- 20 - MOORE, E. - The Shortest Path Through a Maze, Proc. International Symp. Theory of Switching, Part II, April 1957, Harvard University Press, Cambridge, 1959, pp. 285-292.
- 21 - NEWELL, A. and Simon, H. - GPS, a Program that Simulates Human Thought, Computers and Thought (Feigenbaum, E. and Feldman, J. eds), McGraw-Hill, New York, 1963, pp. 279-293.
- 22 - NEWELL, A. and Shaw, J. C. - Empirical Explorations with the Logic Theory Machine, Proc. Western Joint Computer Conf. **11**(1957), 218-239.
- 23 - NILSSON, N. J. - Problem Solving Methods in Artificial Intelligence, McGraw-Hill, New York (1971).
- 24 - PATRICK, E. A. - Concepts of an Estimation System, an Adaptive System, and a Network of Adaptive Estimation Systems, IEEE Trans. on Systems Sc. and Cyb., V.SSC-5,n.1, Jan. 1969, pp 79-85.
- 25 - POHL, I. - First Results on the Effect of Error in Heuristic Search, Machine Intelligence, **5**(1969), 219-236.
- 26 - POHL, I. - Heuristic Search Viewed as Path Finding in a Graph, Artificial Intelligence **1**(1970), 193-204.
- 27 - RABUSKE, M. A. - Um Gerador de Números Aleatórios para IBM-

1130, Tese de Mestrado, PUC-RJ, 1972.

- 28 - SLAGLE, J. and Dixon, J. - Experiments with some Programs that Search Game Trees, J. ACM, Vol. 16, n. 2 April 1969, pp 189 - 207.
- 29 - TSYPKIN, Y. Z. - Foundations of the Theory of Learning Systems, Academic Press, 1973.
- 30 - VANDERBRUG, G. J. - Properties of Heuristic Search Strategies, Univ. Maryland, TR-226.
- 31 - VANDERBRUG, G. J. - A note on Probabilistic Analysis of Heuristic Search, Univ. Maryland, TR-283, Dec. 1973.

A N E X O 1A.1 - SISTEMAS DE ESTIMAÇÃO E ADAPTATIVOS

Devido ao grande uso dos termos 'Estimar', 'Adaptar', 'Autoadaptar' ... etc, nos diversos campos da ciência como, cibernética, psicologia, para só citar dois, é que vamos defini-los precisamente. Claro que existem casos em que o emprego de um pelo outro não causa maiores confusões.

Vamos formular matematicamente o problema:

Seja x^n um vetor 1-dimensional e $\gamma^n = (X^1, \dots, X^n)$ uma sequência de n destes vetores. Suponha que $X^n \in V^1$, um espaço vetorial 1-dimensional, e que há um espaço de decisão $\Omega = \{w_i\}_{i=1}^M$ consistindo de M pontos que indexam M classes. Suponha que existem funções densidade de probabilidade (fdp) $p(X/w_i)$, $i = 1, 2, \dots, M$, chamada a pdf da classe condicional.

Denotamos $P(w_i)$ a probabilidade que x^n corresponda a classe w_i . Então a fdp de X^n pode ser escrita

$$P(X^n/B) = \sum_{i=1}^M P(X^n/w_i, B_i) P(w_i) \quad (A.1)$$

onde B_i é um conjunto de parâmetros suficiente para caracterizar $P(X^n/w_i)$ e B contém todos os parâmetros B_1, B_2, \dots, B_M , incluindo $P(w_1), \dots, P(w)$. A equação (A.1) é chamada uma mistura finita.

Conceituação 1

Estimação é obter os parâmetros em B, dados X_1, X_2, \dots, X_n .

É um problema estatístico e uma solução implica na definição precisa de conhecimentos prévios, isto é, supor que $P(X/w_i)$ é Gaussiano, unimodal ..., ou mesmo, desconhecido. O importante é que o conhecimento prévio exista.

Num sistema de estimação supervisionado (SES), o conhecimento a priori inclui classificação das amostras X_1, X_2, \dots, X_n .

Um sistema de estimação não supervisionado (SENS) pode ter uma saída. Esta saída é um ponto W_a , o índice da classe acusada de ter 'ocasionado' x^n . Um SES não pode ter saída porque a classificação das amostras é conhecida. Se tivesse saída, seria um sistema de reconhecimento.

Conceituação 2

Um sistema de estimação adaptativo (SEA) é um sistema de estimação tendo uma referência externa que muda o 'conhecimento a priori' contra o qual a estimação se realiza. Esta referência externa não está sob a influência estatística do 'conhecimento a priori' como o estão X_1, \dots, X_n .

Observação: Para ver exemplos e conceituações derivadas o leitor pode consultar PATRICK²⁴.

A.2 - SUBROTINE NDTR-SSP

Calcula $y = P(x) = P(X \leq x)$

$$X = N(0,1)$$

$$P(x) = \frac{1}{2\pi} \int_{-\infty}^x \exp(-u^2/2) du$$

ou

CALL NDTR(x,P,D)

x = escalar do qual calcula-se P(x)

P = probabilidade retorno

D = Densidade retorno

Erro Máximo: 0.000 000 7

Subrotina SUBROUTINE NDTR(X,P,D)

AX = ABS(x)

T = 1.0/(1.0+0.2316419 * AX)

D = 0.3989423 * EXP(-X * X/2.0)

**P = 1.0-D*T*(((1.330274 * T-1.821256)* T+
+ 1.781478) * T-0.3565638) * T + 0.3193815)**

IF(X) 1, 2, 2

1 P = 1.0-P

2 RETURN

END

A.3 - TEOREMA

$$E \left[\underline{\hat{A}}(n) \underline{\hat{A}}(n)^T \right] = \underline{\underline{\Sigma}} = \underline{\underline{\Sigma}} + \underline{\underline{\bar{A}}} \underline{\underline{\bar{A}}}^T$$

Prova:

$$\text{Cov} \left[\underline{\hat{A}}(n) \right] = \underline{\underline{\Sigma}} = E \left[(\underline{\hat{A}} - \underline{\bar{A}}) (\underline{\hat{A}} - \underline{\bar{A}})^T \right] = E \left[\underline{\hat{A}} \underline{\hat{A}}^T \right] - E \left[\underline{\hat{A}} \underline{\bar{A}}^T \right] - E \left[\underline{\bar{A}} \underline{\hat{A}}^T \right] + E \left[\underline{\bar{A}} \underline{\bar{A}}^T \right]$$

$$\text{Como } E[\underline{\hat{A}}] = \bar{A} \text{ e } E[\underline{\hat{A}}^T] = \bar{A}^T$$

$$\begin{aligned} \text{então } \underline{\Sigma} &= E[\underline{\hat{A}}\underline{\hat{A}}^T] - \bar{A}^T E[\underline{\hat{A}}] - \bar{A} E[\underline{\hat{A}}^T] + \bar{A}\bar{A}^T - \\ &= E[\underline{\hat{A}}\underline{\hat{A}}^T] - \bar{A}\bar{A}^T - \bar{A}\bar{A}^T + \bar{A}\bar{A}^T = E[\underline{\hat{A}}\underline{\hat{A}}^T] - \bar{A}\bar{A}^T \end{aligned}$$

$$\text{ou seja, } \underline{\Sigma} = \bar{A}\bar{A}^T + \underline{\Sigma}$$

c.q.d.

A.4 - AJUSTAMENTO DE RETAS

Para ajustar uma reta da forma

$$y = a_0 + a_1 x$$

a_0 e a_1 podem ser obtidos como segue:

$$a_0 = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{N\Sigma x^2 - (\Sigma x)^2}$$

$$a_1 = \frac{N\Sigma xy - (\Sigma x)(\Sigma y)}{N\Sigma x^2 - (\Sigma x)^2}$$

A.5 - MATRIZ DE COVARIÂNCIA

a) A matriz de Covariância pode ser obtida através da fórmula

$$\underline{\Sigma} = \underline{U}^T \underline{D} \underline{U}$$

onde \underline{U} = Matriz de transformação rotacional

$$(\underline{U} \cdot \underline{U}^T = \mathbf{I}, \text{ ou seja, } \underline{U}^T = \underline{U}^{-1})$$

$\underline{\underline{D}}$ = Matriz diagonal de autovalores

$$\underline{\underline{D}} = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_n \end{pmatrix}$$

b) Consideremos $\underline{\underline{A}} \underline{\underline{v}} = \underline{\underline{\lambda}} \underline{\underline{v}}$, onde $\underline{\underline{v}}$ representa autovalores e $\underline{\underline{\lambda}} > 0$ autovalores. Precisamos fatorar $\underline{\underline{A}}$ tal que $\underline{\underline{U}}^T \underline{\underline{U}} = \underline{\underline{A}}$.

Para obter $\underline{\underline{U}}$, podemos usar o conceito de ortogonalidade (\perp) de vetores e construí-la progressivamente, fazendo

$$\lambda_1, v_1 \quad (\text{aleatórios})$$

$$\lambda_2, v_2 \perp v_1$$

$$\lambda_3, v_3 \perp (v_1, v_2)$$

$$\lambda_4, v_4 \perp (v_1, v_2, v_3)$$

O módulo de v_1, v_2, \dots deve ser um.

A.6 - APROXIMAÇÃO ESTOCÁSTICA

A aproximação estocástica é apropriada tanto para processos de aproximação lineares como não-lineares. Sua convergência é garantida com o preenchimento de requisitos fracos e gerais estabelecidos pelo teorema de Dvoretzky²⁷.

Seja \hat{p} uma estimativa de p , vetor a ser determinado. A estimativa \hat{p} de p (no enegésimo estágio seqüencial) é deri

vada do algoritmo de aproximação estocástica, como segue:

$$\hat{P}_{n+1} = \hat{P}_n - \rho_n \phi_n \quad \forall n=1,2,3, \dots \quad (\text{A.2})$$

sendo ϕ_n uma função avaliada através de medições instantâneas do fenômeno (medidas do ruído), com valor esperado zero (dentro da precisão estabelecida para os resultados do problema).

ρ_n é uma seqüência de correções escalares.

Para a convergência de A.2, a seqüência ρ_n deve satisfazer as seguintes condições que são baseadas no teorema de Dvoretzki:

$$\left. \begin{array}{l} \lim_{n \rightarrow \infty} \sum_{k=1}^n \rho_k = \infty \\ \lim_{n \rightarrow \infty} \sum_{k=1}^n \rho_k^2 < \infty \\ \lim_{n \rightarrow \infty} \rho_n = 0 \end{array} \right\} \quad (\text{A.3})$$

Obviamente, qualquer função da forma $\rho_k = \rho_1/k$ satisfaz as exigências das equações (A.3), assim como muitas outras funções de k satisfazem.

A.7 - Métodos de Acelerar a convergência

Historicamente, o primeiro método de acelerar convergência numa aproximação estocástica foi proposto por Kestern (1958), quando propôs o número de mudanças de sinal de $(P_k - P_{k+1})$

em (A.2) como indicador de proximidade de P_0 (valor desejado).

Venter (1967), propôs estimar a inclinação da função regressão e usar este dado para acelerar a convergência e a variância assintótica do Procedimento de Robbins-Munro (RM).

Fabian¹⁰, apresenta experiências com os algoritmos de RM e Kiefer Wolfowitz (KW), obtendo melhoras significativas nos dois métodos, usando as seguintes técnicas:

- 1) Acelerar convergência pela escolha apropriada da sequência peso ρ_k em (A.2)
- 2) Tomar mais observações em cada estágio da iteração. Intuitivamente falando, maior número de informações conduz a uma solução mais rápida.

A.8 - Implementação de p (da seção 5.2.2)

Vamos reduzir o intervalo de integração de z e w para $-5,5$, pois o erro que decorre desta redução é mínimo, enquanto a economia de tempo de processamento é razoável.

Os valores de μ , a e σ^2 podem ser assim estimados:

$$\mu = \underline{W}_t^T \cdot \underline{A}_t$$

$$\alpha = 1 + \delta_t / \tilde{h}(n_t)$$

$$\sigma^2 = \underline{W}_t^T \cdot \underline{\Sigma}_t \cdot \underline{W}_t$$

As funções densidade e distribuição são tabelas de 513 pontos:

$$\text{DENS}(0,512) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad (x = -5, \dots, 5)$$

fazendo $E = 1/ \int_{-5}^5 e^{-x^2/2} dx$, então

$$\text{FDIS}(0,512) = E \int_{-5}^x e^{-x^2/2} dx \quad (x = -5, \dots, 5)$$

e

$$\text{FDIS}(x) = 1 \quad \text{para } x \geq 5.$$

O fator E difere muito pouco de 1 e é desprezado. Então, calcular p vem a ser

$$P = \int_{\Phi} \Phi(\alpha, \sigma, \mu) = \int \text{DENS}(x) * \text{FDIS} \text{ integer}(\alpha z - \mu / \sigma(\alpha - 1))$$

$$\text{onde } z = -5 + i * 10/512 \quad (i = 0, \dots, 512)$$

A.9 - Fatoração de Matrizes (MILTON¹⁹)

Seja \underline{A} uma matriz real, simétrica, positivo-definida, de ordem n. Sabemos que \underline{A} pode ser fatorada

$$\underline{A} = \underline{R}^T \cdot \underline{R} \tag{A.4}$$

onde \underline{R} é uma matriz triangular superior. A decomposição (A.4) é conhecida como decomposição de Choleski. O cálculo de \underline{R} pode ser obtido pelo Algoritmo de Decomposição Completa de Choleski (ADCC).

ADCC

Seja $r_{11} = (a_{11})^{1/2}$ e $r_{1j} = a_{1j}/r_{11}$ ($j = 2, \dots, n$),

então para $i = 2, \dots, n$,

$$r_{ii} = (a_{ii} - \sum_{k=1}^{i-1} r_{ki}^2)^{1/2},$$

$$r_{ij} = (a_{ij} - \sum_{k=1}^{i-1} r_{ki} \cdot r_{kj}) r_{ii} \quad (j = i+1, \dots, n)$$

Se $r_{ii} > 0$, então a decomposição de \underline{A} é Única.

A.10 - Demonstração do Teorema 4.2

Vamos supor, sem perda de generalidade, L com valores decrescentes, sendo portanto L_1 o maior (único). Podemos fazer $(L_j = L_1 - C_j)$ para qualquer j , sendo $C_j \geq 0$ e $C_1 = 0$. Isto nos permite escrever

$$\begin{aligned} \sum_{i=1}^B P_i L_i &= P_1 L_1 + P_2 L_2 + \dots + P_B L_B = \\ &= P_1 (L_1 - C_1) + P_2 (L_2 - C_2) + \dots + P_B (L_B - C_B) = \\ &= L_1 (P_1 + P_2 + \dots + P_B) - P_2 C_2 - P_3 C_3 - \dots - P_B C_B. \end{aligned}$$

Como para qualquer i $P_i \geq 0$ e $C_i \geq 0$, então,

$\sum_{i=1}^B P_i L_i$ será máximo quando minimizarmos $\sum_{i=2}^B P_i C_i$, o que acontece

quando $P_i = 0$, para $i \geq 2$.

Como $\prod_{i=1}^B P_i = 1$, logo $P_1 = 1$.

c.q.d.

A N E X O 2**1 - LISTAGEM DA PROGRAMAÇÃO PARA COMPUTADOR**

Linguagem usada: FORTRAN

Computador usado para testes:

BUROUGHS - 6700, mantido pelo Núcleo de computação
Eletrônica da Universidade Federal do Rio de Janeiro.

2 - DADOS**3 - RESULTADOS EXPERIMENTAIS**

Com Expansão não Dirigida

```

SUBROUTINE AJKETA(N,X,Y,AO,A1)
DIMENSION X(50),Y(50)
SY=0.
SX=0.
SX2=0.
SXY=0.
DO 1 K=1,N
SY=SY+Y(K)
SX=SX+X(K)
SX2=SX2+X(K)*X(K)
SXY=SXY+X(K)*Y(K)
DENOM=N*SX2-SX*SX
AO=(SY*SX2-SX*SY)/DENOM
A1=(N*SXY-SX*SY)/DENOM
RETURN
END

```

1

```

10 SUBROUTINE PREVER(ABARRA,SIGMA,NV,S,AI,W)
DIMENSION ABARRA(15),SIGMA(15,15),SINV(15,15),S(15,15)
DIMENSION AI(15,30),W(15),Y(15)
WRITE(6,10)
FORMAT(/,,' MATRIZ S ESPERADA',/)
DO 11 L=1,NV
DO 1 IC=1,NV
S(L,IC)=SIGMA(L,IC)+ABARRA(L)*ABARRA(IC)
11 WRITE(6,20)(S(L,K),K=1,NV)
20 FORMAT(1X,15F5.2)
DO 51 I=1,NV
W(I)=0.
CALL INVERT(S,SINV,AI,NV,DETS)
DO 2 I=1,NV
DO 2 J=1,NV
W(I)=W(I)+ABARRA(J)*SINV(I,J)
2 WRITE(6,30)
30 FORMAT(/,1X,'VECTOR W OTIMO',/)
WRITE(6,20)(W(K),K=1,NV)
XMU=0.
DO 45 K=1,NV
Y(K)=0.
XMU=XMU+W(K)*ABARRA(K)
SIG2=0.
DO 46 L=1,NV
DO 46 K=1,NV
Y(L)=Y(L)+W(K)*SIGMA(K,L)
DO 47 K=1,NV
SIG2=SIG2+Y(K)*W(K)
47 RETURN
END

```

```

SUBROUTINE INVERT(S, SINV, A, N, DETS)
DIMENSION A(15,30), S(15,15), SINV(15,15), VPIV(15)
TOL=0.00001
EPS=0.01
NZ=2*N
1   DO 3 I=1,N
    DO 3 J=1,NZ
      3   A(I,J)=0.
        NPERM=0
        DO 40 I=1,N
          DO 40 J=1,N
            40   A(I,J)=S(I,J)
              NX=N-1
              NI=N+1
C**** FORMACAO DA MATRIZ UNITARIA
        DO 7 I=1,N
          II=I+N
          7   A(I,II)=1.
C**** GERACAO PIVOT. VERIFICAR SE COLUNA DIFERENTE DE ZERO.
        DO 23 L=1,N
          LX=L+1
          MX=0
          DO 13 I=L,N
            IF(ABS(A(I,L))-TOL)13,13,11
              11   MX=1
              13   CONTINUE
            IF(MX)15,37,15
C**** TESTE MAIOR ELEMENTO COLUNA
          15   DO 19 I=LX,N
            IF(ABS(A(L,L))-ABS(A(I,L)))17,19,19
C**** TROCA DE LINHAS
          17   DO 195 JX=L,NZ
            TEMP=A(L,JX)
            A(L,JX)=A(I,JX)
            195  A(I,JX)=TEMP
              NPERM=I-L
          19   CONTINUE
C**** FAZER PIVOT IGJAL A 1
          PIV=A(L,L)
          VPIV(L)=PIV
          DO 21 JX=L,NZ
            21   A(L,JX)=A(L,JX)/PIV
C**** METODO GAUSS JORDAN
          DO 23 K=1,N
            I=K
            IF(I-L)190,18,190
              18   I=I+1
              IF(I-N)190,190,23
            190  M=0
              DIVA=A(I,L)
              DO 210 J=L,NZ
                A(I,J)=A(I,J)-A(L,J)*DIVA
C**** VERIFICAR SE LINHA E DIFERENTE DE ZERO
            IF(I-L)20,20,189
              189  IF(J-N)191,191,210
              191  IF(ABS(A(I,J))-TOL)210,210,20
            20   M=1

```

```

210  CONTINUE
      IF(M)23,37,23
23   CONTINUE
      DO 41 I=1,N
      DO 41 J=N1,NZ
      K=J-N
41   SINV(I,K)=A(I,J)
C
C***  CALCULA DETERMINANTE
C
      DETS=1
      DO 62 I=1,N
      DETS=DETS*VP IV(I)
62   CONTINUE
      DETS=DETS*(-1)**NPERM
      RETURN
C
C***  MATRIZ INVERSA INDEFINIDA
C***  FAZ S=S+EPS*I
C
37   DO 36 I=1,N
36   S(I,I)=S(I,I)+EPS
      EPS=2.*EPS
      WRITE(6,300)EPS
300  FORMAT(1X,'EPS EM INVERT =',F12.2)
      GO TO 1
      STOP
      END

SUBROUTINE ATE(VATIL,A,ALFA,ICONTA,T)
DIMENSION VATIL(15),A(15),VAUX(15)
COMMON N,NP,NSJC,IP,IU,NV,P(2000),C(2000),WTIL(15)
COMMON HCM(4,2000)
AUX=0.
BUX=0.
PASSO=.5
IF(T.LT.9)PASSO=1.0
DO 60 K=1,NV
VAUX(K)=ALFA*(VATIL(K)-A(K))
AUX=AUX+VAUX(K)**2
60  BUX=BUX+A(K)**2
AUX=SQRT(AUX)
BUX=SQRT(BUX)
CUX=AUX/BUX
DUX=1
IF(CUX.GT.PASSO)DUX=PASSO/CUX
DO 61 K=1,NV
61  A(K)=A(K)+DUX*VAUX(K)
DO 646 KT=1,NV
646 IF(A(KT).LT.0.01)A(KT)=.01
RETURN
END

```

```

SUBROUTINE HCMAI (ABARRA, B, IX, H, VALOR, VETOR, D)
DIMENSION ABARRA(15), B(15,15), R(16), H(1000), Y(16)
DIMENSION D(1000), VETOR(15)
COMMON N, NP, NSUC, IP, IU, NV, P(2000), C(2000), WTIL(15)
COMMON HCM(4,2000)
DO 1 I=1, NV, 2
R1=RANDDM(IX)
R2=RANDDM(IX)
R(I)=SQRT(-2.*ALOG(R1))*COS(2.*3.1416*R2)
R(I+1)=SQRT(-2.*ALOG(R1))*SIN(2.*3.1416*R2)
1 CONTINUE
DO 3 I=1, NV
Y(I)=0
DO 3 J=1, I
3 Y(I)=Y(I)+B(I,J)*R(J)
DO 2 I=1, NV
2 HCM(I,N)=H(N)*(ABARRA(I)+ Y(I))
C
C*** CALCULA H CHAPEJ PEQUENO
C
VALOR=0
DO 10 I=1, NV
10 VALOR=VALOR+VETOR(I)*HCM(I,N)
IF(VALOR.LT.1.)VALOR=1.
RETURN
END

```

```

SUBROUTINE BE(SIGMA, C, N)
DIMENSION SIGMA(15,15), C(15,15)
DO 2 I=1, N
DO 2 J=1, N
2 C(I,J)=0
DO 1 I=1, N
DO 1 J=1, N
IF(J.LT.1)GO TO 1
IF(J.NE.I)GO TO 5
S=0.
L=I-1
IF(L.EQ.0)GO TO 4
DO 3 K=1, L
3 S=S+C(I,K)**2
4 C(J,I)=SQRT(SIGMA(J,I)-S)
GO TO 1
5 S=0.
L=I-1
IF(L.EQ.0)GO TO 6
DO 7 K=1, L
7 S=S+C(I,K)*C(J,K)
6 C(J,I)=(SIGMA(J,I)-S)/C(I,I)
1 CONTINUE
RETURN
END

```

```

SUBROUTINE DELTA(DA,GAMA,DT,ICONTA)
E=.2
DAA=DA
DA=DA+GAMA*(DT-DA)
IF(DA-DAA)2,3,4
2 IF(DA.LT.DAA/2.)DA=DAA/2.
IF(DA.LT.E)DA=E
3 RETURN
4 IF(DA.GT.DAA*2.)DA=DAA*2.
IF(DA.LT.E)DA=E
RETURN
END

```

```

SUBROUTINE WWTILL(A,S,W,AI,ALFA,ICONTA,DETS)
DIMENSION Y(16),A(15),SINV(15,15),W(15),S(15,15),AI(15,30)
COMMON N,NP,NSUC,IP,IU,NV,P(2000),C(2000),WTIL(15)
COMMON HCM(4,2000)
IF(ICONTA.GT.0)RETURN
DO 1 I=1,NV
Y(I)=0
CALL INVERT(S,SINV,AI,NV,DETS)
DO 2 I=1,NV
DO 2 J=1,NV
2 Y(I)=Y(I)+A(J)*SINV(I,J)
DO 4 K=1,NV
4 W(K)=W(K)+ALFA*(Y(K)-W(K))
RETURN
END

```

```

SUBROUTINE PPILHA(INIC,N,ISUC,G,HC,FTG,FTH)
DIMENSION ISUC(2000),G(2000),HC(2000)
IF(INIC.GT.0)GO TO 16
INIC=N
ISUC(N)=0
RETURN
16 NA=INIC
22 IF((G(NA)*FTG+HC(NA)*FTH).GT.(G(N)*FTG+HC(N)*FTH))GO TO 19
IANT=NA
NA=ISUC(NA)
IF(NA.NE.0)GO TO 22
ISUC(IANT)=N
ISUC(N)=0
RETURN
19 IF(NA.EQ.INIC)GO TO 23
ISUC(IANT)=N
ISUC(N)=NA
RETURN
23 INIC=N
ISUC(INIC)=NA
RETURN
END

```



```

SUBROUTINE MINLI(DA, HPTIL, A, HC, C, NP, NV, L, N, W, H)
DIMENSION A(15), HC(2000), C(2000), W(15), H(2000)
REAL L(2000)
XMU=0.
DO 1 I=1, NV
1 XMU=XMU+W(I)*A(I)
XKI=HPTIL-C(N)
IF(XKI.LT.1.0)XKI=1.0
PP=(HC(N)-XKI*XMU)**2/XKI**2
SP=(HC(N)-(XKI+DA)*XMU)**2/(XKI+DA)**2
L(N)=PP-SP-ALOG(1.+DA/XKI)
RETURN
END

```

```

SUBROUTINE FATCOR(PARTE, ALFA, SIMU, DENS, FDIS, PP)
DIMENSION FDIS(130), DENS(130)
AREA=PARTE/2.*(DENS(1)*FDIS(1)+DENS(129)*FDIS(129))
DO 765 I=1, 127
Z=FLOAT(I)*PARTE-5.
AV=ALFA*Z+SIMU*(ALFA-1.)
IF(AV.LT.-5.)AV=-5.
IV=(AV+5.)/PARTE
IF(IV.LT.2)IV=2
IF(IV.GT.128)IV=128
AREA=AREA+PARTE*(DENS(I)*FDIS(IV))
765 CONTINUE
PP=AREA
RETURN
END

```

```

SUBROUTINE PLOTA(NARG, MA, VETARG, ISIN1, ISIN2, ESCALA, T)
DIMENSION MA(50, 120), VETARG(9), ISIN1(9), ISIN2(9), ESCALA(9)
IC=T
IC=IC-(IC/121)*120
IF(IC.GT.120)RETURN
DO 7 I=1, NARG
L=VETARG(I)*ESCALA(I)+20
IF(L.GT.0)GO TO 4
MA(1, IC)=ISIN2(I)
GO TO 7
4 IF(L.LT.51)GO TO 6
MA(50, IC)=ISIN2(I)
GO TO 7
6 MA(L, IC)=ISIN1(I)
7 CONTINUE
RETURN
END

```

```

SUBROUTINE STE(M,S,STIL,NV,ICONTA,SA,BETA,NDUX,A)
DIMENSION S(15,15),STIL(15,15),SA(10,15,15)
DIMENSION SAUX(15,15),A(15)
ICONTA=ICONTA+1
IF(ICONTA.EQ.M)GO TO 4
DO 1 J=1,NV
DO 1 K=1,NV
1 SA(ICONTA,J,K)=STIL(J,K)
RETURN
4 MI=M-1
FLOM=FLOAT(M)
ICONTA=0
DO 3 K=1,MI
DO 3 I=1,NV
DO 3 J=1,NV
3 STIL(I,J)=(STIL(I,J)+SA(K,I,J))/FLOM
C
C*** ADAPTACAO DO S(T)
C
PASSO=.5
IF(T.LT.9.)PASSO=1
AUX=0.
BUX=0.
DO 190 I=1,NV
DO 190 J=1,NV
SAUX(I,J)=BETA*(STIL(I,J)-S(I,J))
AUX=AUX+SAUX(I,J)**2
190 BUX=BUX+S(I,J)**2
AUX=SQRT(AUX)
BUX=SQRT(BUX)
CUX=AUX/BUX
DUX=1
IF(CUX.LE.PASSO)GO TO 198
DUX=PASSO/CUX
NDUX=NDUX+1
198 CONTINUE
DO 195 I=1,NV
DO 195 J=1,NV
195 S(I,J)=S(I,J)+DUX*SAUX(I,J)
RETURN
END

```

```

C-----
C**** PROGRAMACAO E TESTE DO ESQUEMA ADAPTATIVO
C-----
      DIMENSION B(15,15),SIGMA(15,15),R(16),ABARRA(15)
      DIMENSION SA(10,15,15),IFILA(2000),SEG(50),SEH(50)
      DIMENSION WJ(15),G(2000),H(2000),IPAI(2000),A(15)
      DIMENSION HC(2000),S(15,15),W(15),STIL(15,15),EHCP(15)
      DIMENSION VML(15),AI(15,30),SD(15,15),EHCF(15),ISUC(2000)
      DIMENSION AX(15),M1(50,120),VETARG(9),ISIN1(9),ISIN2(9)
      DIMENSION FDIS(130),DENS(130),NAVE(5),ESCALA(9)
      DIMENSION NOSP(80),NACE(80)
      REAL L(2000),Y(16),D(2000)
      COMMON N,NP,NSUC,IP,IU,NV,P(2000),C(2000),WTIL(15)
      COMMON HCM(4,2000)
      DATA ITRACO,IBRANC,ISIN1,ISIN2/'-',' ','.','*','+', '=','
1'0','A','X','Y','L','1','2','3','4','5','6','7','8','9'/
C
C**** NAVE(I)=POSITIVO, SERA RODADA TENTATIVA I
C      NAVE(I)=NEGATIVO, TENTATIVA I NAO SERA RODADA.
C
      HCPQ=1.
      NAVE(1)=1
      NAVE(1)=-1
      NAVE(2)=1
      NAVE(2)=-1
      NAVE(3)=-1
      NAVE(3)=1
      NAVE(4)=1
      NAVE(4)=-1
      NAVE(5)=1
      NAVE(5)=-1
      NALTER=5
C
C**** FCORRE=POSITIVO, ENTAO RODA A-TIL COM CORRECAO
C
      FCORRE=-1.
      FCORRE=1.
C
C**** FAZ IESTAB=1 SE RODAR ESTABILIDADE
C      IESTAB=0 SE SE RODAR CONVERGENCIA
C
      IESTAB=1
      IESTAB=0
C
C**** IMPRA=1-IMPRIME GRAFICO IMPRA=0-NAO IMPRIME
C
      IMPRA=1
      IMPRA=0
C
C**** NNDS=NUMERO DE NOS A SEREM EXPANDIDOS
C
      NNDS=660
      NNDS=120
      NNDS=200
      NNDS=300
      FTG=1.
      FTH=1.

```

```

      IY=5432157797
      IX=453289
      IX=753264005
      IX=72036455
      DO 885 I=1,50
885  XR1=RANDOM(IX)
C
C**** CALCULA DENS E FDIS
C
      X=-5
      PARTE=10./128.
      DO 762 I=1,129
      PX=ABS(X)
      T=1.0/(1.0+.2316419*PX)
      U=.3989423*EXP(-X*X/2.0)
      E=1.-U*T*{(((1.330274*T-1.821256)*T+1.781478)*T-
1  0.3565638)*T+0.3193815)
      IF(X)4,2,2
4    E=1-E
2    X=X+PARTE
      FDIS(I)=E
      DENS(I)=U
762  CONTINUE
C
      DO 173 I100=1,10
      SP=1,
      DO 268 I=1,80
      NDSP(I)=0
268  NACE(I)=0
      ICONTA=0
      M=1
      HPTIL=1
      TETA=1.
      FGH=999999.
C
C*** LEITURA - IMPRESSAO - PREVISAO
C
      MN=12
      WRITE(6,950)
950  FORMAT(/,1X,'MATRIZ SIGMA',/)
      READ(5,800)NV
800  FORMAT(I6)
      DO 910 I=1,NV
      READ(5,911)(SIGMA(I,J),J=1,NV)
910  WRITE(6,911)(SIGMA(I,J),J=1,NV)
911  FORMAT(4F7.3)
      CALL BE(SIGMA,B,NV)
920  FORMAT(1X,15F5.1)
      READ(5,900)(ABARRA(I),I=1,NV)
900  FORMAT(15F5.1)
      WRITE(6,980)
980  FORMAT(/,1X,'MEDIA A',/)
      WRITE(6,900)(ABARRA(I),I=1,NV)
      WRITE(6,960)
960  FORMAT(/,1X,'MATRIZ B',/)
      DO 894 I=1,NV
894  WRITE(6,110)(B(I,J),J=1,NV)

```

```

110  FORMAT(1X,10F7,2)
      CALL PREVER(ABARRA,SIGMA,NV,S,AI,W)
      DO 78 I=1,NV
      DO 78 J=1,NV
78    SQ(I,J)=S(I,J)
      DO 55 IR=1,NV
55    WO(IR)=W(IR)
      IF(FCORRE.LT.0.)GO TO 4422
      DO 173 IZE=1,NALTER
      IF(NAVE(IZE).LT.0)GO TO 173
C
C*** LIMPA MATRIZ-GRAFICO
C
4422 CONTINUE
      IF(IMPRA.LT.1)GO TO 672
      DO 1770 J=1,120
      DO 1771 I=1,50
1771  M1(I,J)=IBRANC
1770  M1(20,J)=ITRACO
672   CONTINUE
      IPG=0
C
C*** IMPRIME CABECALHO
C
      WRITE(6,4010)
4010  FORMAT(//)
      WRITE(6,4020)NV,IX,IZE,TETA
4020  FORMAT(1X,'NV',I3,2X,'IX',I13,2X,'TENT.= ',I6,'TETA',F6.2)
      WRITE(6,40)
40    FORMAT(//,3X,'T',3X,'N',4X,'HT',5X,'H',4X,'HC',4X,'HCM',
16X,'D',6X,'A',5X,'W',5X,'G',7X,'S,,,')
C
C*** INICIALIZACAO
C
      IACE=0
      NA3=0
      NDUX=0
      NP=1
      DO 274 I=1,NV
      EHCP(I)=0.
274   EHCF(I)=0.
      H(1)=20
      H(1)=30
      H(1)=40
      COST=2
      DEL=1
      P(1)=1.
      INDSO =1
      IPRIM=1
      LFIM=1000
      LFIM=2000
      D(1)=0
      G(1)=0
      N=1
      T=1
      DO 386 I=1,LFIM
386   IFILA(I)=0

```

```

      INIC=1
      IPAI(1)=0
      ISUC(1)=0
      C(1)=2
      DO 114 K=1,NV
114   HCM(K,N)=H(N)*ABARRA(K)
C
C***  INICIALIZACAO DE S-A-DA, PARA TESTE  DE CONVERGENCIA
C
      IF(UESTAB.GT.0)GO TO 550
      DA=.5
      SCX=1
      DO 48 I=1,NV
      W(I)=1.
      A(I)=1.
      DO 49 J=1,NV
49   S(I,J)=1.
48   S(I,I)=2.
      ACRA=0.
      GO TO 552
C
C****  INICIALIZACAO DE S-A-DA , PARA TESTE DE ESTABILIDADE
C
550  CONTINUE
      NA3=50
      DO 2270 I=1,NV
      EHCF(I)=NA3*ABARRA(I)*H(I)
2270 EHCP(I)=-NA3*ABARRA(I)
      AOM=(EHCF(1)+EHCF(2))/(2.*NA3)
      AIM=(EHCP(1)+EHCP(2))/(2.*NA3)
      ACRA=50.
      DA=1
      SCX=2
      DO748 I=1,NV
      W(I)=W0(I)
      A(I)=ABARRA(I)
      DO749 J=1,NV
749  S(I,J)=S0(I,J)
748  CONTINUE
C
C****  SELECIONA NO PARA FECHAR (NP) E GERA SUCESSORES
C
552  CONTINUE
18   IF(INIC.GT.0)GO TO 12
      WRITE(6,10)
10   FORMAT('ABERTO VAZIO')
      STOP
C
C***  RECALCULA H CHAPEU PEQ. DA FILA
C
12   NOVO=INIC
      IF(T.GT.25.5)GO TO 3020
      IF(T.LT.24.5)GO TO 3043
      FTG=0.
      IPF=0
      IPF=1
      LIMG=0

```

```

      NDSP(1)=1
3043 CONTINUE
      IC=T
      IRES=IC/20+1
      TE1=T/IRES
      ITE1=TE1
      DO 298 K=1,LFIM
      IF(NOVO.EQ.0)GO TO 299
      IFILA(K)=NOVO
      NOVO=ISUC(NOVO)
298 CONTINUE
299 IFILA(K)=0
      INIC=0
      DO 301 K=1,LFIM
      NT=IFILA(K)
      IF(NT.EQ.0)GO TO 3020
      HTL=0.
      DO 302 K4=1,NV
302 HTL=HTL+W(K4)*HCM(K4,NT)
      HC(NT)=HTL
      FTH=TETA
      CALL PPILHA(INIC,NT,ISUC,G,HC,FTG,FTH)
301 CONTINUE
C
C *** EXPANDE NO,,, NP      FECHA,,, NP
C
3020 NP=INIC
      IF(NP.LT.1)GO TO 1141
      IF(T.LE.25.7)GO TO 3049
3042 CONTINUE
      IF(IPF.LT.1)GO TO 1141
      IF(IPF.GT.58)GO TO 1141
      IF(NOSP(IPF).LT.LIMG)GO TO 3044
3062 IPF=IPF+1
      LIMG=LIMG+1
      IF(LIMG.GT.10)LIMG=10
      GO TO 3042
3044 CONTINUE
      KNP=INIC
      IANT=INIC
3045 KG=G(KNP)
      IF(KG.EQ.IPF)GO TO 3046
      IANT=KNP
      KNP=ISUC(KNP)
      IF(KNP.NE.0)GO TO 3045
      GO TO 3062
3046 NP=KNP
      IF(NP.EQ.INIC)GO TO 3049
      ISUC(IANT)=ISUC(NP)
      GO TO 3048
3049 CONTINUE
      INIC=ISUC(INIC)
3048 CONTINUE
      IF(D(NP).LT.0.1)IACE=IACE+1
      T=T+1
      IGR=G(NP)
      IF(IGR.LT.1)GO TO 267

```

```

IF(H(NP).LT.0.1)WRITE(6,780)NP,H(NP),G(NP),IACE,T
NOSP(IGR)=NOSP(IGR)+1
IF(D(NP).GT.,.1)GO TO 267
NACE(IGR)=NACE(IGR)+1
267 CONTINUE
IF(T.GT.NNDS)GO TO 1141
IF(H(NP).LE.0.)GO TO 673
PROBO=1.
PFOBNO=1.
ALFA=1./(T+ACRA)
BETA=ALFA
GAMA=BETA
CORR=1.
NSUC=3
IP=N+1
IU=N+NSUC
C*** SORTEIA INDICE DE SUCESSOR DTIMO
INDSO=NSUC*RANDOM(IY)
INDSO=IP+INDSO
DO 11I=IP,IU
N=N+1
IF(N.LT.LFIM)GO TO 27
as WRITE(6,80)
80 FORMAT(' FALTA ESPACO')
GO TO 73
27 IPAI(N)=NP
C(N)=COST
D(N)=DEL
C**** CALCULA H h E MINUSCULO
IF(I.NE.INDSO)GO TO 6
D(N)=0,
6 H(N)=H(NP)-C(I)+D(N)
8 G(N)=G(NP)+C(N)
IF(H(N).LT.0.)G(N)=G(N)+H(N)
IF(H(N).LT.0.)H(N)=0.
CALL HCMAI(ABARRA,B,IX,H,HC(N),W,D)
P(N)=0.
FTH=TETA
CALL PPILHA(INIC,N,ISUC,G,HC,FTG,FTH)
11 CONTINUE
C
C**** AJUSTA RETA A HCM E ESTIMA H-TIL
C
IF(T.GT.25.)GO TO 596
TTH=0.
AA0=0.
AA1=0.
DO 581 K=1,NV
KF=0
KJ=NP
DO 582 K1=1,25
IF(KJ.LT.1)GO f3 583
SEH(K1)=HCM(K,KJ)
SEG(K1)=G(KJ)
KJ=IPAI(KJ)
KF=KF+1
582 CONTINUE

```



```

583  CONTINUE
      IF(KF.LT.3)GO TO 1592
      CALL AJRETA(KF,SEG,SEH,A0,A1)
      AAO=AA0+A0/NV
      AA1=AA1+A1/NV
      IF(KF.LT.6)GO TO 581
      IF(K.EQ.1)NA3=NA3+1
      EHCF(K)=EHCF(K)+A0
      EHCP(K)=EHCP(K)+A1
581  CONTINUE
      IF(KF.LT.6)GO TO 1595
      AOM=0.
      AIM=0.
      DO 1591 K=1,NV
      AOM=AOM+EHCF(K)
1591  AIM=AIM+EHCP(K)
      AOM=AOM/NV
      AIM=AIM/NV
      IF(AIM.GT.-.5)AIM=-.5
1594  A3=AOM+AIM*G(NP)
      HPTIL=A3/ABS(AIM)
      GO TO 1596
1592  IF(NA3.LT.1)GO TO 1596
      GO TO 1594
1595  IF(NA3.GE.1)GO TO 1594
      IF(AA1.GT.-.5)AA1=-.5
      A3=AA0+AA1*G(NP)
      HPTIL=A3/ABS(AA1)
1596  CONTINUE
      IF(HPTIL.LT.1.)HPTIL=1.
C
C *** CALCULA TETA ,, CASO ESTE SEJA USADO
C
      HCPQ=HC(NP)
      IF(HC(NP).LE.1.)HCPQ=1.
      QTC=HPTIL/HCPQ
C
C *** DETERMINA PRDB-P
C
596  CONTINUE
      DO 1 I=IP,IU
1      CALL MINLI(DA,HPTIL,A,HC,C,NP,NV,L,I,W,H)
      DO 296 K=1,NSUC
2915  VML(K)=0.
      XMEN=L(IP)+1.
      NGL=0
      DO 31 I=IP,IU
      I1=I-IP+1
      IF(L(I)-XMEN)32,33,31
33   VML(I1)=NGL
      IG=IG+1
      GO TO 31
32   NGL=NGL+1
      VML(I1)=NGL
      XMEN=L(I)
      IMEL.=I
      IG=1

```

```

31  CONTINUE
    XNGL=NGL
    NOT=0
    DO 42 I=IP,IU
    I1=I-IP+1
    P(I)=0
    IF(ABS(VML(I1)-XNGL).LT.0.1)P(I)=1./FLJAT(IG)
42  CONTINUE
C
C**** DESCOBRE SUCESSOR OTIMO E CALCULA SCX (MEDIA DOS CUSTOS)
C
    SCX=0.
    DO 293 K=IP,IU
    IF(P(K).GT.0.1)IOX=K
    SCX=SCX+C(K)
293 CONTINUE
    SCX=SCX/NSUC
C
C***** CALCULA DTIL
C**** O TESTE E PROGRAM. P/ C IGUAIS
C
    SODEL=0.
    NSN=(NSUC-1)*NV
    DO 1160 K=IP,IU
    IF(P(K).GT.0.1)GO TO 1160
    DO 1161 I=1,NV
1161 SODEL=SODEL+(HCM(I,NP)-HCM(I,K))/(HCM(I,NP)-HCM(I,IOX))
1160 CONTINUE
    SODEL=SODEL/NSN
    DT=C(IOX)*(1-SODEL)
    IF((KF,GE,3).OR.(NA3.GT.0))GO TO 588
C
C*** CALCULA OCASIONALMENTE A-TIL E S-TIL
C
586 CONTINUE
    DO 527 K=1,NV
    SHH=0.
    DO 532 I=IP,IU
532 SHH=SHH+(HCM(K,NP)-HCM(K,I))/NSUC
527 AX(K)=SHH/SCX
    DO 5527 I=1,NV
    DO 5527 K=1,NV
5527 STIL(I,K)=AX(K)*AX(I)
    GO TO 235
588 CONTINUE
    IF(FCDRE.LT.0.)GO TO 4888
C
C**** CALCULA MUT,SIGT,ALFAT
C
    ALFAT=1+DA/HC(IOX)
    XMU=0.
    DO 45 K=1,NV
    Y(K)=0.
45  XMU=XMU+W(K)*A(K)
    SIG2=0.
    DO 46 LT=1,NV
    DO 46 K=1,NV

```

```

46  Y(LT)=Y(LT)+W(K)*ABS(S(K,LT)-A(K)*A(LT))
    DO 47 K=1,NV
47  SIG2=SIG2+Y(K)*W(K)
    SIG2=ABS(SIG2)
    SIGT=SQRT(SIG2)
    SIMU=XMU/SIGT
    CALL FATCOR(PARTE,ALFAT,SIMU,DENS,FDIS,PP)
    FP=PP**{NSUC-1}
    I=IOX
    GO TO (4081,4082,4083,4084,4085),IZE
C
C*** INICIO DA TENTATIVA 1 *****
C
4081 CONTINUE
    FCO=1./{1.-DA/C(I)*{1.-FP}}
    DO 6334 K=1,NV
6334 AX(K)=(HCM(K,NP)-HCM(K,I))/C(I)*FCO
    FATOR=2*HPTIL/C(I)**2*{DA*FP-HPTIL+SCX-DA}
    DO 4030 J=1,NV
    DO 4333 K=1,NV
    ST11(J,K)=(HCM(J,NP)-HCM(J,I))*(HCM(K,NP)-HCM(K,I))/C(I)**2
    STIL(J,K)=STIL(J,K)+FATOR*(S(J,K)-A(J)*A(K))
    GO TO 235
C
    FIM DA TENTATIVA 1
C
C*** INICIO DA TENTATIVA 2 *****
C
4082 CONTINUE
    IF{(HPTIL-C(I)),LT,0.2) HPTIL=HPTIL+1
    DIVISR=FP+(HPTIL-C(I)+DA)/{HPTIL-C(I)*{1.-FP}}
    DO 2238 K=1,NV
2238 AX(K)=(HCM(K,I)/(HPTIL-C(I)))/DIVISR
    DIVISR=FP+(1-FP)*{1+DA/(HPTIL-C(I))**2}
    DIV1=(HPTIL-C(I))**2
    DO 2239 J=1,NV
    DO 2239 K=1,NV
2239 STIL(J,K)=HCM(J,I),I)/(DIVISR*DIV1)
    GO TO 235
C
    FIM DA TENTATIVA 2
C
C*** INICIO DA TENTATIVA 3 *****
C
4083 CONTINUE
    DO334 K=1,NV
334 AX(K)=HCM(K,NP)/HPTIL
    DO 4033 J=1,NV
    DO 4030 K=1,NV
    STIL(J,K)=HCM(J,NP)*HCM(K,NP)/HPTIL**2
4030 CONTINUE
    GO TO 235 )*HCM(K
C
    FIM DA TENTATIVA 3
C
C*** INICIO DA TENTATIVA 4 *****
C
4084 CONTINUE
    F11=HPTIL-C(I)
    IF(F11.LE.0.5)F11=1.0

```

```

F12=F11+DA
IF(F12,LE.0.5)F12=1.0
FATOR=(F11+DA*FP)/(F11*F12)
DO 3238 K=1,NV
3238 AX(K)=FATOR*HCM(K,I)
FATOR=FP/F11**2+(1-FP)/F12**2
DO 3239 J=1,NV
DO 3239 I=1,NV
3239 STIL(J,K)=FATOR*HCM(J,I)*HCM(K,I)
GO TO 235
C FIM DA TENTATIVA 4
C
C*** INICIO DA TENTATIVA 5 *****
C
4085 CONTINUE
PROBO=FP
PROBNO=1.-FP
C FIM DA TENTATIVA 5
C
C*** DETERMINA AX E STIL COM O HPTIL
C
4888 CONTINUE
DO 160 I=1,NV
DO 161 J=1,NV
161 STIL(I,J)=0.
160 AX(I)=0.
DO 34 I=IP,IU
DX=0.
PROBA=PROBO
IF(P(I),GT.0.01)GO TO 1260
DX=DA
PROBA=PROBNO
1260 CONTINUE
XI=HPTIL-C(I)+DX
IF(XI,LT.1.)XI=1.
DO 84 K=1,NV
AX(K)=AX(K)+(HCM(K,I)/NSUC)/XI*PROBA
DO 84 J=1,NV
STIL(K,J)=STIL(K,J)+HCM(K,I)*HCM(J,I)/(NSUC*XI**2)*PROBA
84 CONTINUE
34 CONTINUE
235 CONTINUE
C
C*** ROTINAS DO ESQUEMA ADAPTATIVO
C
CALL WWTILL(A,S,W,AI,CORR,ICONTA,DETS)
CALL DELTA(DA,GAMA,DT,ICONTA)
IF(DA,GT.SCX)DA=SCX
CALL ATE(AX,A,ALFA,ICONTA,T)
CALL STE(M,S,STIL,NV,ICONTA,SA,BETA,NDUX,A)
C
C*** PREPARACAO DE ESTATISTICAS
C
IF(IMPRA,LT.1)GO TO 673
CA=0.
DO 76 I=1,NV
76 CA=CA+ABS(A(I)-ABARRA(I))

```

```

C
  CS=0.
  DO 77 I=1,NV
  DO 77 J=1,NV
77  CS=CS+ABS(S(I,J)-SD(I,J))
  CS=SQRT(CS)
C
  CM=0.
  DO 56 IR=1,NV
56  CM=CM+ABS(W(IR)-WO(IR))*ABARRA(IR)
C
  CV=0.
  DO 57 IC=1,NV
  R(IC)=0.
  DO 58 IL=1,NV
58  R(IC)=R(IC)+W(IL)-WO(IL)*SIGMA(IL,IC)
  CV=CV+R(IC)*ABS(W(IC)-WO(IC))
57  CONTINUE
C
  DHTIL=H(NP)-HPTIL
C
  DH=H(NP)-HC(NP)
C
  NARG=8
  VETARG(1)=DH
  ESCALA(1)=1.
  VETARG(2)=CV
  ESCALA(2)=10
  VETARG(3)=CM
  ESCALA(3)=10
  VETARG(4)=CA
  ESCALA(4)=1
  VETARG(5)=CS
  ESCALA(5)=1.
  VETARG(6)=DA
  ESCALA(6)=10
  VETARG(7)=DETS
  ESCALA(7)=1
  VETARG(8)=DHTIL
  ESCALA(8)=1.
  CALL PLOTA(NARG,M1,VETARG,ISIN1,ISIN2,ESCALA,T)
673 CONTINUE
  IDV=T/5.
  DV=T/5.
  DO 647 I=IP,IU
647 IF(D(I).LT.0.1)IEC=I
  IF((DV-IDV).GT.0.1)GO TO 427
  WRITE(6,50)T,NP,HPTIL,H(NP),HC(NP),HCM(1,NP),DA,A(1),
1W(1),G(NP),(S(1,J),J=1,NV),D(NP),IPAI(NP),HC(IP),HC(IU)
50  FORMAT(/,1X,F4.0,1X,I3,3(1X,F5.1),1X,F6.1,2(1X,F6.2),
12(1X,F5.2),10(1X,F6.2))
  DO 501 K=2,NV
501 WRITE(6,500)HCM(K,NP),A(K),W(K),(S(K,J),J=1,NV)
500 FORMAT(28X,F6.1,8X,F6.2,1X,F5.2,6X,10(1X,F6.2))
427 CONTINUE
  IF(H(NP).LE.0.)GO TO 73
  IF(IMPRA.LT.1)GO TO 18

```

```

IC=T
IF(IC.EQ.120)GO TO 73
IF(IC.EQ.240)GO TO 73
IF(IC.EQ.360)GO TO 73
GO TO 18

```

C

C*** DETERMINA COMPRIMENTO DO CAMINHO ENCONTRADO

C

```

73 CONTINUE
NNAC=0
NNCAM=0
MPP=NP
51 MPP=IPAI(MPP)
IF(MPP.LT.1)GO TO 52
NNCAM=NNCAM+1
IF(D(MPP).LT.0.1)NNAC=NNAC+1
GO TO 51
52 CONTINUE
G1=H(1)+NNCAM-NNAC
IF(G(NP).GE.FGH)GO TO 18
FGH=G(NP)
WRITE(6,2880)G1,G(NP),NP,H(NP),INIC
2880 FORMAT(' CAMINHO-*',F7.2,'G=',F7.2,'NP=',I6,'H=',F7.2,I6)
WRITE(6,1002)T,IACE,NNCAM,NNAC
1002 FORMAT('/',IX,'NDS EXPAND =',I6,' TOTAL ACERTOS =',I6,
1 ' NDS/CAMINHO =',I6,' ACERTOS/CAMINHO =',I6)
WRITE(6,210)(NOSP(I),I=1,60)
WRITE(6,210)(NACE(I),I=1,60)
210 FORMAT(1X,60I2)
IF(IMPRA.LT.1)GO TO 18
IF((T.EQ.120.).OR.(T.EQ.150.))GO TO 2817
GO TO 18
2817 CONTINUE
WRITE(6,1020)
1020 FORMAT('/',IX,' (H CHAS -H REAL)..ESC=1',/,IX,'
1 '* (VAR. W-W*).. ESC=10',/IX,'+ (MEDIA w-W*)..ESC=10',
2/,IX,'= (SOMA / A-A* /)..ESC=1',/,IX,'0 (SOMA / S-S* /)'
3'..ESC=1',/,IX,'A (DELTA)..ESC=10',/,IX,
4'X (DETS)..ESC=1',/,IX,'Y H(REAL)=HPTIL.ESC=1',//)
WRITE(6,4005)
DO 232 I=1,50
232 WRITE(6,1000)I,(M1(I,J),J=1,120)
1000 FORMAT(1X,I2,120A1)
WRITE(6,4005)
4005 FORMAT('1')
WRITE(6,142)
142 FORMAT('/',IX,60('()'),//)
DO 1870 J=1,120
83 1871 I=1,50
1871 M1(I,J)=IBRANC
1870 M1(20,J)=ITRACO
GO TO 18
1141 CONTINUE
270 FORMAT(1X,12F10.3)
WRITE(6,780)ADM,AIM,NP
WRITE(6,1002)T,IACE,NNCAM,NNAC
WRITE(6,210)(NOSP(I),I=1,60)

```

```
WRITE(6,210) (NACE(I),I=1,60)
WRITE(6,4005)
780  FORMAT(1X,6F12.5)
173  CONTINUE
STOP
END
```

2. Dados

4.....	Sigma: 1 ; $\bar{\sigma}^2 \approx 0.5$		
.666	-.067	.2	.067
-.067	.4	.067	.067
.2	.067	.332	.067
.067	.067	.067	.6
2. 4. 3. 5.			
4.....	Sigma: 2 ; $\bar{\sigma}^2 \approx 0.5$		
.5	.5	.5	.5
2. 4. 3. 5.			
4.....	Sigma: 3 ; $\bar{\sigma}^2 \approx 0.3$		
.333	-.033	.1	.033
-.033	.2	.033	.033
.1	.033	.166	.033
.033	.033	.033	.3
2. 4. 3. 5.			
4.....	Sigma: 4 ; $\bar{\sigma}^2 \approx 1.0$		
.1	-.1	.3	.1
-.1	.6	.1	.1
.3	.1	.5	.1
.1	.1	.1	1.
2. 4. 3. 5.			
3.....	Sigma: 5 ; $\bar{\sigma}^2 \approx 0.3$		
.15	.231	-.163	
.231	.4	-.283	
-.163	-.283	.4	
2. 4. 3.			
3.....	Sigma: 6 ; $\bar{\sigma}^2 \approx 0.7$		
.3	.462	-.326	
.462	.8	-.566	
-.326	-.566	.889	
2. 4. 3.			
3.....	Sigma: 7 ; $\bar{\sigma}^2 \approx 0.5$		
.45	.693	-.489	
.693	1.2	-.849	
-.489	-.849	1.2	
2. 4. 3.			
2.....	Sigma: 8 ; $\bar{\sigma}^2 \approx 0.3$		

0.275 -0.043

-0.043 0.226

2. 5.

2..... sigma: 9 ; $\bar{\sigma}^2 \approx 0.5$

0.55 -0.087

-0.087 0.453

2. 5.

2..... Sigma:10 ; $\bar{\sigma}^2 = 1.0$

1.1 -0.174

-0.174 0.906

2. 5.

2..... Sigma:11 ; $\bar{\sigma}^2 = 2.0$

2.20 -0.348

-0.348 1.812

2. 5.

3 - Resultados Experimentais com Expansão não dirigida

θ	Nº	$\sigma^2 \approx 0.3$		$\sigma^2 \approx 0.6$		$\sigma^2 \approx 1.0$		$\sigma^2 \approx 2.0$	
		NNE	g*	NNE	g*	NNE	g*	NNE	g*
1.0	1	446	40	584	42	553	41	-	-
	2	345	40	108	45	-	-	-	-
	3	30	46	-	-	-	-	-	-
	4	78	46	-	-	-	-	-	-
	5	108	43	-	-	-	-	-	-
1.2	1	108	42	156	44	172	45	126	44
	2	25	45	55	44	105	43	95	45
	3	55	47	165	46	125	45	170	45
	4	75	46	137	48	169	46	66	44
	5	50	45	-	-	-	-	126	46
	6	97	40	-	-	-	-	-	-
1.5	1	27	46	58	49	109	49	60	51
	2	35	51	40	49	40	44	36	48
	3	55	46	46	46	85	48	49	49
	4	35	48	55	46	90	43	-	-
	5	60	48	35	48	60	49	-	-
	6	40	49	30	49	-	-	-	-
	7	50	46	85	46	-	-	-	-
	8	40	48	60	44	-	-	-	-
	9	50	46	-	-	-	-	-	-
1.8	1	32	44	42	48	42	46	51	48
	2	28	47	-	-	-	-	-	-
	3	29	50	-	-	-	-	-	-
2.0	1	28	50	32	50	48	51	80	50
	2	30	50	34	52	32	50	42	52
	3	29	47	28	47	52	48	54	51
	4	-	-	-	-	37	51	54	51