

UNIVERSIDADE FEDERAL DE SANTA CATARINA
COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO
ENGENHARIA ELÉTRICA

IDENTIFICAÇÃO DE SISTEMAS COM PARÂMETROS VARIÁVEIS
COM O TEMPO USANDO CÁLCULO VARIACIONAL

HAMILTON MEDEIROS SILVEIRA

TESE SUBMETIDA A APRECIÇÃO COM REQUISITO PARCIAL PARA
A OBTENÇÃO DO GRAU DE:
MESTRE EM CIÊNCIAS DE ENGENHARIA ELÉTRICA

FLORIANÓPOLIS
SANTA CATARINA - BRASIL
ABRIL - 1973.

ESTA TESE FOI JULGADA ADEQUADA PARA A OBTENÇÃO
DO TÍTULO DE
"MESTRE EM CIÊNCIAS"
E APROVADA EM SUA FORMA FINAL PELO ORIENTADOR E PELO CURSO
DE PÓS-GRADUAÇÃO.

Prof. Roger Pouliquen
Orientador

Polidoro São Tiago
Prof. Polidoro Ernani São Tiago Filho
Integrador do Curso de Pós-Graduação

BANCA EXAMINADORA:

Roger Pouliquen
Ernani São Tiago
Adriana

A G R A D E C I M E N T O S

O autor deseja expressar sua gratidão a todos que de uma maneira ou de outra, colaboraram na realização deste trabalho.

Especialmente:

Ao Professor ROGER POULIQUEN, pela orientação.

À ROSÂNGELA BARRETO LAUS, pelos serviços de datilografia.

Ao acadêmico ZÉLIO CASAGRANDE, pela colaboração nos serviços de computação.

À COPERTIDE por haver possibilitado o presente trabalho.

S I N O P S E

Este trabalho relaciona-se com identificação de sistemas. Muitos trabalhos existem sobre identificação de sistemas com parâmetros constantes, no entanto o problema não está totalmente resolvido. Identificação de sistemas com parâmetros - variáveis com o tempo é um problema muito difícil e poucos trabalhos existem sobre este assunto. O presente trabalho refere-se a identificação de sistemas variáveis com o tempo usando cálculo - variacional e o método do modelo. São apresentados estudos sobre tais métodos, aplicação sobre uma classe de sistemas e as principais dificuldades do uso prático destas teorias.

A B S T R A C T

This report deals with Systems Identification. There exist many works about constant parameters systems identification. However, this problem has not been completely solved.

Identification of time varying parameters systems is a problem more difficult and there are few works on this subject.

The present studies deals with time varying parameters systems using variational calculus and the model method. Such methods are studied and an application is presented using a particular class of systems.

The main difficulties found in applying such theories are also presented.

I N D I C E

Introdução	1
Identificação pelo método do modelo	3
2.1 - Princípio fundamental	3
2.2 - Fases do método do modelo	7
2.2.1 - Enunciado	7
2.2.1.1- Hipóteses de equação	8
2.2.1.2- Hipóteses de parâmetros	9
2.2.1.3- Dados práticos	10
2.2.1.4- Viabilidade de resposta	11
2.2.2 - Identificação	18
2.2.3 - Interpretação Física	18
2.2.4 - Verificação	19
2.3 - Espaço paramétrico	21
2.3.1 - Definição	21
2.3.2 - Análise na vizinhança do ponto nominal	22
2.3.3 - Forma das superfícies isodistâncias	24
2.4 - Entradas sensibilizante e esfeizante	26
2.5 - Métodos de minimização	26
2.5.1 - Métodos monovariáveis	26
2.5.2 - Métodos multivariáveis de primeira ordem	29
2.5.2.1- Métodos heurísticos	30
2.5.2.2- Métodos analíticos	30
2.5.3 - Métodos multivariáveis de segunda ordem	31
2.5.3.1- Métodos analíticos	31
2.5.3.2- Métodos heurísticos	32
2.6 - Problemas do método do modelo	32
Revisão do cálculo variacional	36
3.1 - Definição de funcional	36
3.2 - Definição de norma de uma função	36
3.3 - Definição de incremento de um funcio nal	37
3.4 - Definição de linearidade de um fun -	

	cional	37
3.5	- Definição de variação de um funcional	37
3.6	- Máximo e mínimo de um funcional	37
3.7	- Teorema fundamental do cálculo variacional	38
3.8	- Lema fundamental do cálculo variacional	38
3.9	- Problema de Otimização	39
Utilização do cálculo variacional em identificação		42
4.1	- Identificação de um sistema com parâmetros variáveis com o tempo	42
4.2	- Entrada sensibilizante e esferizante	49
Aplicação do cálculo variacional em identificação		50
5.1	- Considerações sobre identificação de uma classe de sistemas variáveis com o tempo	50
5.2	- Formulação matemática	52
5.3	- As considerações de contorno	53
5.4	- Distância Objeto-modelo	54
5.5	- Espaço de condições iniciais	59
5.6	- Custo, Algoritmos e critério de parada	59
5.7	- Instabilidade e precisão	63
5.8	- Estudo de um sistema de primeira ordem	64
5.9	- Estudo de um sistema de segunda ordem	68
Resultados e conclusões		77
6.1	- Resultados	77
6.2	- Conclusões	78

CAPÍTULO 1INTRODUÇÃO

Identificar é encontrar um modelo matemático que caracterize adequadamente o desempenho dinâmico de um sistema físico. É essencial em muitos casos de análise de controle moderno, em otimização e em métodos de projetos. Sua aplicação se estende a manutenção, medicina e outras áreas.

Em muitos casos não há informação "a priori" sobre o sistema e o modelo deverá ser determinado por testes experimentais ou definido supondo-se ter alguma forma empírica. Em outros casos a forma funcional do modelo e talvez alguns de seus parâmetros são conhecidos através da análise teórica ou de testes prévios.

Muitas técnicas são aplicadas em Identificações, como por exemplo métodos estatísticos, otimização, método do modelo, identificação da resposta impulsiva e etc ¹.

O presente trabalho se preocupa em estudar o método do modelo usando cálculo variacional para identificação de parâmetros que variam com o tempo. O problema em geral é difícil e poderemos dividi-lo em duas grandes etapas:

- a) determinação do modelo;
- b) determinação dos parâmetros.

O primeiro passo é o mais difícil, notadamente em processos pouco conhecidos, como por exemplo, um sistema biológico. Alguns sistemas de engenharia são bem conhecidos teoricamente, como por exemplo um motor de corrente contínua.

Não será pretensão deste trabalho a determinação de modelos, limitando-se apenas a tecer algumas considerações a respeito da determinação dos parâmetros de um modelo já conhecido.

Identificar processos com parâmetros cons -

tantes é um problema que já está bem estudado. Processo com parâmetros variáveis é um problema bem mais difícil de identificar, sendo possível identificar com boa precisão parâmetros que variem muito lentamente.

No segundo capítulo nos preocuparemos em apresentar as bases do método do modelo, mostrando alguns métodos de otimização e dando ênfase aos principais problemas que ocorrem na aplicação de tal método.

O terceiro capítulo refere-se a uma revisão do cálculo variacional servindo com instrução ao quarto capítulo.

Neste apresentaremos as condições de otimalidade impostas pelo cálculo variacional ao problema de identificação bem como uma idéia de aplicação desta técnica na determinação de entrada sensibilizante.

No capítulo cinco apresentaremos discussões sobre a aplicação prática do capítulo quatro em identificação. Será discutido o problema da função custo, da instabilidade e precisão. Será feito também um estudo sobre uma classe de sistemas de primeira e segunda ordens. Serão apresentadas algumas técnicas que visem contornar o problema da instabilidade tornando possível a utilização dos algoritmos.

No capítulo seis serão apresentados resultados práticos de algumas técnicas sugeridas, bem como uma conclusão sobre o presente trabalho.

CAPÍTULO 2

IDENTIFICAÇÃO PELO MÉTODO DO MODELO

Neste capítulo, vamos introduzir as principais idéias a respeito de identificação pelo método do modelo. Veremos os fundamentos de tal processo, bem como as principais dificuldades de sua aplicação prática. Para se resolver um problema de identificação, geralmente se recorre ao auxílio de um computador digital que falaremos oportunamente, dando ênfase a os princípios de cálculo que servirão para implementação em computador digital.

Procuraremos neste capítulo, simplificar ao máximo os principais conceitos, no sentido de dar, em rápida leitura, uma boa noção do método do modelo.

2.1 - PRINCÍPIO FUNDAMENTAL ²:

Identificar um processo não é muito fácil. Veremos que em alguns casos, por falta de dados, isto se torna meramente impossível.

Afinal o que é identificar?

Identificar, é fazer a melhor aproximação matemática de um processo real.

Para isto necessitaremos de um conjunto de equações, que deverão expressar matematicamente o processo físico. A este conjunto de equações denominaremos "modelo matemático".

Existem vários tipos de modelos matemáticos utilizados em identificação. Uma classe, quando estamos tratando de processos lineares, é a "resposta impulsiva". Muitos trabalhos foram publicados a esse respeito.

Entretanto, a resposta impulsiva, nada nos diz a respeito da estrutura do processo real.

Em muitas aplicações práticas, estamos interessados em conhecer a estrutura do processo real e em outros ela já é bem conhecida e desejamos conhecer os valores numéricos' de seus parâmetros.

Um de nossos propósitos neste trabalho, já que existe interesse prático, é estudar outra classe de modelos matemáticos, que poderemos chamar de "modelos paramétricos", os quais procuram representar a estrutura do processo físico. Estes modelos também dependem de "parâmetros estruturais", que simplesmente denominaremos de "parâmetros".

Ao método de identificação que utiliza modelos paramétricos, denominaremos de "método do modelo". Em outras palavras, este método estuda modelos matemáticos estruturais. Uma outra característica do método do modelo é a de que os parâmetros deverão ter interpretação física.

De agora em diante, denominaremos "objeto" ao processo real e "modelo" ao modelo paramétrico.

Identificar pelo método do modelo, é encontrar um modelo cujo desempenho seja o mais próximo possível do objeto em estudo, sendo que seus parâmetros tenham interpretação física.

A resolução do problema, em linhas gerais, se resume na suposição de um modelo cujos parâmetros são forçados a se aproximarem ao máximo dos do objeto. Esta aproximação é feita em um intervalo de tempo finito $[t_0, t_f]$ que denominaremos de "intervalo de tempo de observação".

O problema matematicamente é de otimização' e consiste em minimizar uma função custo que denominaremos de "distância de estrutura".

Definiremos distância de estrutura pela expressão abaixo:

$$D_E(O, M) = \int_{t_0}^{t_f} \| p_o(t) - p(t) \|_{A(t)}^2 dt = \int_{t_0}^{t_f} [(p_o(t) - p(t))^T A(t) (p_o(t) - p(t))] dt \quad (1)$$

onde $p_o(t)$ é o vetor parâmetro do objeto, $p(t)$ é o vetor parâmetro do modelo e $A(t)$ uma matriz peso.

Se os parâmetros forem constantes, poderemos definir a distância de estrutura por:

$$D_E(O, M) = \sum_{i=1}^N k_i (p_{o,i} - p_i)^2 \quad (2)$$

sendo k_i um fator de ponderação, $p_{o,i}$ o i -ésimo parâmetro do objeto, p_i o i -ésimo parâmetro do modelo e N o número de parâmetros. Podemos observar que tanto na (1) como na (2) $D_E(O, M) > 0$ se $O \neq M$ e $D_E(O, M) = 0$ se, $O = M$, onde O é o objeto e M o modelo.

As funções custos (1) e (2) são impossíveis de se calcular, porque desconhecemos o vetor parâmetro do objeto. Existem outras maneiras de se medir a distância de estrutura sendo o mais comum a "distância de saída".

Dizemos que um modelo é de "resposta viável" se existe um único vetor parâmetro do modelo p que torne zero a distância de saída.

Definiremos distância de saída por:

$$\begin{aligned} D_S(O, M) &= \int_{t_0}^{t_f} \frac{\| \underline{s}_o(t) - \underline{s}_m(t) \|^2}{A(t)} dt = \\ &= \int_{t_0}^{t_f} [(\underline{s}_o(t) - \underline{s}_m(t))^T A(t) (\underline{s}_o(t) - \underline{s}_m(t))] dt \end{aligned} \quad (3)$$

onde $\underline{s}_o(t)$ e $\underline{s}_m(t)$ são respectivamente os vetores saída do objeto e do modelo e $A(t)$ uma matriz de ponderação.

Se o modelo for de resposta viável, $D_S(O, M) > 0$ se $M \neq O$ e $D_S(O, M) = 0$ se $M = O$, ou seja a expressão (3) é outra maneira de se medir a distância de estrutura.

A distância de saída é muito fácil de ser obtida e é largamente utilizada na prática como critério de medição de distância de estrutura.

O diagrama da fig. 1, mostra o procedimento para minimizar $D_S(O, M)$:

Excita-se o objeto e o modelo com uma mesma função $\underline{u}(t)$. Os sinais de saída são comparados e é obtido $\xi(t)$, que é o sinal de entrada para o medidor de $D_S(O, M)$.

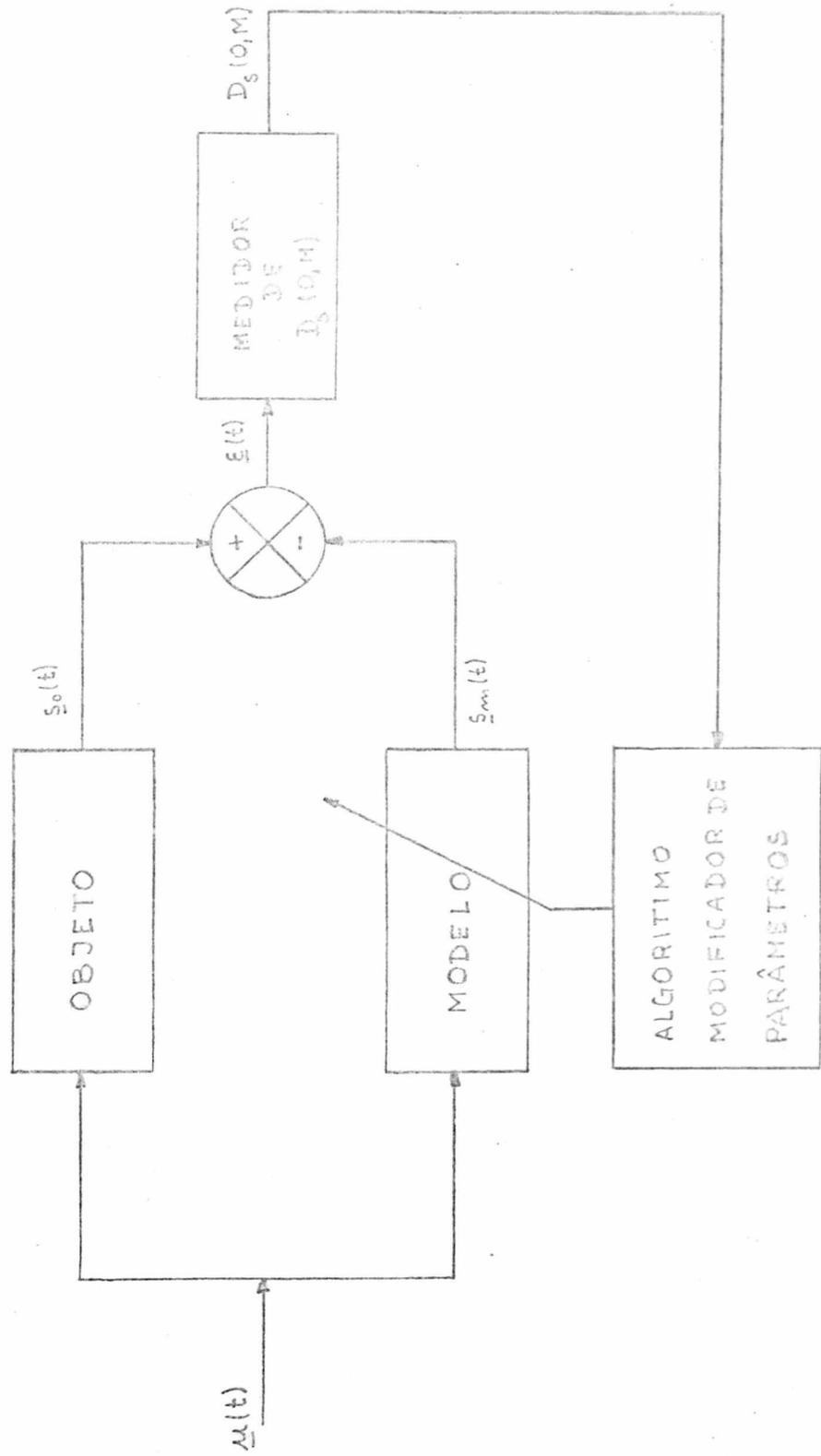


Figura 1

Findo o intervalo de observação é obtido $D_s(0, M)$. Este sinal entra no algoritmo modificador de parâmetros que atua nos parâmetros do modelo no sentido de minimizar $D_s(0, M)$. O processo é repetido e para quando $D_s(0, M)$ for mínimo.

O ideal seria encontrar parâmetros que fizessem $D_s(0, M) = 0$. No entanto isto na prática é impossível. Um dos motivos é que $y_o(t)$, que provém da medição da saída do objeto, sempre apresenta erros de medida e ruídos, tornando o mínimo de $D_s(0, M) > 0$. Isto será abordado posteriormente.

O procedimento descrito acima não é simples de ser executado e envolve o uso de computador digital. Introduzido o princípio fundamental detalharemos a idéia e apresentaremos as principais dificuldades do método do modelo.

2.2 - FASES DO MÉTODO DO MODELO:

Para encontrar um modelo cuja performance seja a mais próxima possível do desempenho do objeto, devidiremos o procedimento em quatro fases:

- Enunciado;
- Identificação;
- Interpretação física;
- Verificação.

2.2.1 - ENUNCIADO:

Nesta etapa enunciaremos o problema. Faremos hipótese sobre a equação matemática, sobre parâmetros, tomaremos os dados práticos e analisaremos a viabilidade de determinação dos parâmetros. Esta é a etapa mais difícil, por se tratar de formulação de hipóteses.

Estas hipóteses deverão ser auxiliadas por um especialista no setor do objeto de estudo que se pretende identificar. Como exemplo, poderemos citar a identificação de um sistema biológico que forçosamente exigirá nesta etapa, a assistência de um médico especialista no setor de estudo. Ele poderá nos fornecer dados que ajudarão a formular as hipóteses necessá

rias. Poderemos dividir esta etapa em quatro fases:

- Hipóteses de equação;
- Hipóteses de parâmetros;
- Dados práticos;
- Viabilidade de resposta.

No enunciado procuraremos caracterizar o modelo.

2.2.1.1 - HIPÓTESES DE EQUAÇÃO:

Aqui, procuraremos encontrar a equação matemática que rege o sistema, tendo-se sempre em mente, que os parâmetros de tal equação deverão ter significação física.

Sempre se faz uma aproximação. Determinar exatamente a estrutura do objeto é quase impossível.

Como exemplo, sabemos que todos os processos são não lineares. Em certas faixas comportam-se como quase lineares. Nestas faixas, aproximamos em engenharia a um processo puramente linear.

Assim também procedemos em identificação. Na lei de Ohm ($V = Ri$) sabemos que a resistência varia com a temperatura, porém em uma faixa estreita consideramos R "constante" o que na verdade é quase constante.

O modelo $V = Ri$ sendo R constante, em identificação é perfeitamente válido para a faixa de temperatura de R quase constante.

Se o objeto estivesse trabalhando numa faixa muito ampla de temperatura, onde deveríamos considerar a variação de resistência, deveríamos escolher outro modelo onde R fosse função da temperatura.

Como a maioria dos processos são regidos por equações diferenciais, daqui em diante sempre que nos referirmos à equação do modelo, fica explícito que se trata de equação diferencial, salvo menção em contrário.

Entendemos por "equação do modelo" dois grupos de equações; um, função dos parâmetros que descreve o processo em si e outro que descreve a saída.

Em variáveis de estado por exemplo poderemos escrever:

$$\dot{\underline{x}}(t) = \underline{b}(\underline{x}(t), \underline{p}(t), \underline{u}(t), t) \quad (4)$$

$$\underline{s}_m(t) = \underline{f}(\underline{x}(t), \underline{u}(t), t) \quad (5)$$

onde $\underline{x}(t)$ é o vetor estado do modelo, $\underline{u}(t)$ o vetor excitação, $\underline{p}(t)$ o vetor parâmetro e $\underline{s}_m(t)$ o vetor saída do modelo.

Existem outras maneiras de descrever a equação do modelo e a melhor maneira depende de cada caso em particular.

Uma equação bem simples é a que rege um processo linear de primeira ordem:

$$\dot{x}(t) = -p_1(t)x(t) + p_2(t)u(t) \quad (6)$$

$$s_m(t) = x(t) \quad (7)$$

Se é suposto que um processo tenha o modelo acima, este pode não ser o definitivo, pois as outras fases de elaboração o confirmarão ou o refutarão. As hipóteses de equações são a etapa fundamental e base para os outros passos que completam o enunciado.

2.2.1.2 - HIPÓTESES DE PARÂMETROS:

Feitas as hipóteses de equação, que sabemos ser uma aproximação das equações que realmente regem o objeto, surge então as hipóteses sobre os parâmetros.

Se supusermos que o modelo é regido pelas (6) e (7) poderemos supor várias hipóteses sobre os parâmetros. Por exemplo, poderemos supor que $p_1(t)$ e $p_2(t)$ são constantes o que im

plicaria que fariamos as hipóteses de parâmetros:

$$\forall t_1, t_2 \quad p_1(t_1) = p_1(t_2) \quad (8)$$

$$p_2(t_1) = p_2(t_2) \quad (9)$$

Citando outro exemplo, poderíamos saber por um processo de medida direta que $p_1(t) = 1$, ou em outro exemplo que $p_2(t) = 5 p_1(t)$.

As hipóteses de parâmetros melhor completam o enunciado.

2.2.1.3 - DADOS PRÁTICOS:

Como os modelos que estamos tratando são regidos por equações diferenciais, sabe-se da teoria matemática - que para a resolução de uma equação diferencial é necessário o conhecimento das condições iniciais.

Em identificação faz-se duas medidas no objeto que chamaremos de "dados práticos". A primeira é a das condições iniciais que pode ser o estado no início do intervalo de observação. A segunda é obtida medindo-se a saída do objeto $\xi_o(t)$ durante o intervalo de observação, quando submetido a uma função excitação $u(t)$ conhecida.

Aqui aparecem problemas práticos de escolha de instrumentos, de confiabilidade, de transdutores, de eliminação de ruídos e etc.

No exemplo das equações (6) e (7) se o objeto partisse do repouso $x(t_0)$ seria igual a zero dispensando-se a medida da condição inicial.

Existem problemas onde o objeto não parte do repouso sendo necessário a medição das condições iniciais. Exemplo disto é a identificação "on line" onde o objeto está constantemente em funcionamento e em dado intervalo de observação de sejamos conhecer os seus parâmetros.

Quanto maior a quantidade de dados práticos tanto melhor. Se for possível obter dois ou mais grupos de dados

práticos melhor será a identificação.

Cada grupo é caracterizado pelo intervalo ' de observação, o vetor excitação $\underline{u}(t)$ e as medidas das condições iniciais e o correspondente vetor $\underline{s}_0(t)$.

Assim, na obtenção de dois grupos excitamos o objeto com $\underline{u}_1(t)$ e medimos as condições iniciais e $\underline{s}_{01}(t)$. A seguir excitamos com outra função $\underline{u}_2(t)$ e medimos as novas condições iniciais e a nova saída $\underline{s}_{02}(t)$.

O número de grupos de dados práticos é um fator bastante importante que será abordado na fase seguinte.

2.2.1.4 - VIABILIDADE DE RESPOSTA:

Nossa preocupação nesta etapa é verificar ' se o modelo é de resposta viável com as hipóteses anteriormente estabelecidas e com os dados práticos obtidos. Em outras palavras, desejamos provar que existe um único vetor $\underline{p}(t)$ que torna $\underline{s}_0(t) = \underline{s}_m(t)$, ou seja $D_s(0, M) = 0$.

Suponhamos que um objeto tem a configuração da fig. 2, onde L é uma indutância, R uma resistência e G o ganho de um amplificador. Excitamos o sistema com uma fonte de tensão $u(t)$ e medimos este sinal com o voltmetro V. O sinal de saída é a corrente $i(t) = s_0(t)$ que é a medida pelo ampêremetro A.

Pretende-se determinar os tres parâmetros ' estruturais G, L e R.

Considerando G, L e R constantes poderemos dizer que a equação do modelo é:

$$L \frac{di(t)}{dt} + R i(t) = G u(t) \quad (10)$$

Definindo $L = p_1$, $R = p_2$, $G = p_3$ e $i(t) = z(t)$ poderemos escrever:

$$p_1(t) \dot{z}(t) + p_2(t) z(t) = p_3(t) u(t) \quad (11)$$

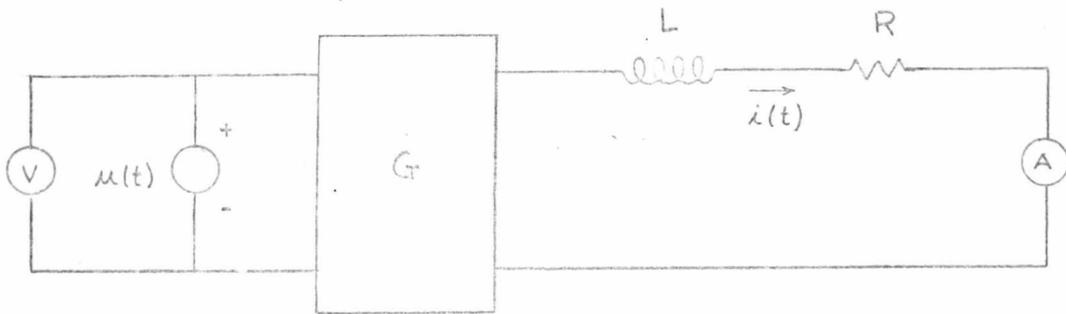


Figura 2

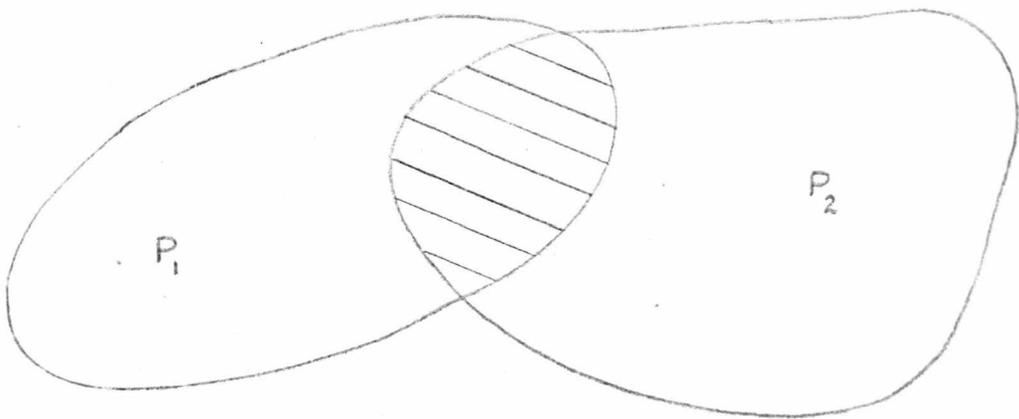


Figura 3

$$S_m(t) = x(t) \quad (12)$$

As hipóteses de parâmetros como estabelecidas anteriormente são:

$$\forall t_1, t_2 \in [t_0, t_f] \quad p_1(t_1) = p_1(t_2) = p_1 \quad (13)$$

$$p_2(t_1) = p_2(t_2) = p_2 \quad (14)$$

$$p_3(t_1) = p_3(t_2) = p_3 \quad (15)$$

Supondo que no sistema não existia corrente antes da aplicação da fonte de tensão poderemos seguramente dizer que $x(t_0) = 0$. Suponhamos ainda que obtivemos um único grupo de dados práticos.

O problema assim caracterizado, apresenta infinitas soluções para p_1 , p_2 e p_3 . É fácil de mostrar considerando a equação do sistema:

$$p_1 \dot{x}(t) + p_2 x(t) = p_3 u(t) \quad (16)$$

Se multiplicarmos a equação acima por uma constante k , matematicamente a equação não se altera:

$$k p_1 \dot{x}(t) + k p_2 x(t) = k p_3 u(t) \quad (17)$$

Poderemos então definir outros parâmetros $P_1 = k p_1$, $P_2 = k p_2$ e $P_3 = k p_3$ e a equação fica:

$$P_1 \dot{x}(t) + P_2 x(t) = P_3 u(t) \quad (18)$$

Isto demonstra que tanto poderemos determinar p_1 , p_2 e p_3 como P_1 , P_2 e P_3 , já que matematicamente as duas equações são equivalentes, ou seja admitem a mesma resposta $S_o(t) = S_m(t) = x(t)$. Como k pode ter qualquer valor, é óbvio que o problema admite infinitas soluções.

As hipóteses e dados práticos feitas para o modelo geram infinitas soluções.

Identificar o problema assim enunciado é determinar uma relação entre os parâmetros, mas jamais com as hipóteses e dados obtidos poderemos saber exatamente os valores reais. É um problema de solução impossível, sendo perfeitamente inviável a determinação da relação entre os parâmetros. Isto é o que poderemos esperar como solução do problema.

Se no entanto por medida direta soubermos o valor do fator de amplificação G por exemplo igual a 5, o problema terá mais um dado adicional e obteremos uma única solução como se pode mostrar abaixo.

A equação é:

$$p_1(t) \dot{x}(t) + p_2(t) x(t) = p_3(t) u(t) \quad (19)$$

onde a saída é $s_m(t) = x(t)$.

As hipóteses de parâmetros são:

$$\forall t_1, t_2 \in [t_0, t_f] \quad p_1(t_1) = p_1(t_2) = p_1 \quad (20)$$

$$p_2(t_1) = p_2(t_2) = p_2 \quad (21)$$

$$p_3(t) = 5 \quad (22)$$

Poderemos considerar também que $x(t_0) = 0$ e que obtivemos um único grupo de dados práticos. Em laboratório, como dados, medimos $u(t)$ e $s_0(t) = x(t)$. Como conhecemos $x(t)$, sua derivada $\dot{x}(t)$ fica então conhecida. Para dois instantes de tempo t_1 e t_2 poderemos escrever o seguinte sistema de equações:

$$\dot{x}(t_1) p_1 + x(t_1) p_2 = 5 u(t_1) \quad (23)$$

$$\dot{x}(t_2) p_1 + x(t_2) p_2 = 5 u(t_2) \quad (24)$$

onde p_1 e p_2 são incôgnitas. Se tomarmos estes instantes no período transitório, tal que $\dot{x}(t_1)$ e $\dot{x}(t_2)$ sejam diferentes de zero poderemos ter o determinante da matriz abaixo diferente de zero:

$$\begin{vmatrix} \dot{x}(t_1) & x(t_1) \\ \dot{x}(t_2) & x(t_2) \end{vmatrix} \neq 0 \quad (25)$$

o que faz com que os valores de p_1 e p_2 tenham uma única solução que satisfaz $S_0(t) = \mathcal{X}(t)$.

Daqui pode-se ver que as hipóteses e dados práticos são fatores bastante importantes no enunciado do problema e que nem todos os problemas apresentam uma única solução compatível.

Como outro exemplo onde o problema apresentaria infinitas respostas seria no caso de a equação ser:

$$p_1(t) \dot{\mathcal{X}}(t) + p_2(t) \mathcal{X}(t) = \mu(t) \quad (26)$$

$$S_m(t) = \mathcal{X}(t) \quad (27)$$

não se fazendo nenhuma hipótese de parâmetro. Portanto, não se impõe nenhuma restrição adicional. Ainda neste caso consideraremos que $\mathcal{X}(t_0) = 0$ e que obtivemos um único grupo de dados práticos medindo $S_0(t) = \mathcal{X}(t)$. Pode-se mostrar que o modelo assim enunciado apresenta infinitas soluções compatíveis de $p_1(t)$ e $p_2(t)$.

Conhecemos $\mu(t)$ e $S_0(t)$ que são dados práticos tal como no caso anterior. Se $\mathcal{X}(t)$ é conhecido também o será $\dot{\mathcal{X}}(t)$. Assim para qualquer valor de $p_1(t)$ poderemos determinar um valor compatível de $p_2(t)$, e isolando este da equação do modelo vem:

$$p_2(t) = \frac{\mu(t) - p_1(t) \dot{\mathcal{X}}(t)}{\mathcal{X}(t)} \quad (28)$$

Isto demonstra que o problema apresenta infinitas soluções, já que para qualquer $p_1(t)$ arbitrário existirá o correspondente $p_2(t)$ compatível com a equação, ou seja que satisfaz $S_0(t) = \mathcal{X}(t) = S_m(t)$.

Se neste mesmo exemplo, obtivéssemos dois grupos de dados práticos, o problema seria de solução viável.

Para isto deveremos impor a hipótese de parâmetros que $p_1(t)$ e $p_2(t)$ são os mesmos nas medidas dos dois grupos. Excitando com $\mu_1(t)$ obteríamos $S_{01}(t) = \mathcal{X}_1(t)$ e com $\mu_2(t)$ obteríamos $S_{02}(t) = \mathcal{X}_2(t)$. Conhecidos $\mathcal{X}_1(t)$ e $\mathcal{X}_2(t)$ o são também $\dot{\mathcal{X}}_1(t)$ e $\dot{\mathcal{X}}_2(t)$.

Poderemos formar o seguinte sistema de equações:

$$p_1(t) \dot{x}_1(t) + p_2(t) x_1(t) = u_1(t) \quad (29)$$

$$p_1(t) \dot{x}_2(t) + p_2(t) x_2(t) = u_2(t) \quad (30)$$

em instantes de tempo que o determinante

$$\begin{vmatrix} \dot{x}_1(t) & x_1(t) \\ \dot{x}_2(t) & x_2(t) \end{vmatrix} \neq 0 \quad (31)$$

obtermos uma única solução de $p_1(t)$ e $p_2(t)$ que satisfazem $S_{01}(t) = x_1(t) = S_{m1}(t)$ e $S_{02}(t) = x_2(t) = S_{m2}(t)$.

Como exemplo, para mostrar a influência das condições iniciais. Suponhamos que o objeto seja:

$$\ddot{y}(t) + 5 \dot{y}(t) + 6 y(t) = u(t) \quad (32)$$

$$S_0(t) = y(t) \quad (33)$$

Suponhamos ainda que o ensaio prático foi realizado com $u(t) = 0$ e tendo condições iniciais $y(t_0) = 0$ e $\dot{y}(t_0) = -2$.

A equação fica:

$$\ddot{y}(t) + 5 \dot{y}(t) + 6 y(t) = 0 \quad (34)$$

que tem somente resposta homogênea.

A equação característica será:

$$s^2 + 5s + 6 = 0 \quad (35)$$

tendo raízes -2 e -3.

A resposta terá a seguinte forma:

$$y(t) = k_1 e^{-2t} + k_2 e^{-3t} \quad (36)$$

Usando as condições iniciais determinamos que $k_1 = 1$ e $k_2 = 0$, logo a solução será:

$$S_0(t) = y(t) = e^{-2t} \quad (37)$$

Este será o resultado obtido no ensaio prático. Enunciariamos este problema da seguinte forma em identificação:

$$\ddot{x}(t) + p_1(t) \dot{x}(t) + p_2(t) x(t) = 0 \quad (38)$$

$$s_m(t) = x(t) \quad (39)$$

$$\forall t_1, t_2 \in [t_0, t_f] \quad p_1(t_1) = p_1(t_2) = p_1 \quad (40)$$

$$p_2(t_1) = p_2(t_2) = p_2 \quad (41)$$

com $x(t_0)=1$ e $\dot{x}(t_0)=-2$ e $s_0(t)=x(t)$ seria medido em laboratório sendo igual a e^{-2t} .

No entanto o processo assim enunciado apresenta infinitas soluções. Se considerarmos dois instantes t_1 e t_2 poderemos escrever:

$$p_1 \dot{x}(t_1) + p_2 x(t_1) = -\ddot{x}(t_1) \quad (42)$$

$$p_1 \dot{x}(t_2) + p_2 x(t_2) = -\ddot{x}(t_2) \quad (43)$$

Substituindo as derivadas vem:

$$-2e^{-2t_1} p_1 + e^{-2t_1} p_2 = 4e^{-2t_1} \quad (44)$$

$$-2e^{-2t_2} p_1 + e^{-2t_2} p_2 = 4e^{-2t_2} \quad (45)$$

Para qualquer valor de t_1 e t_2 o determinante de

$$\begin{bmatrix} -2e^{-2t_1} & e^{-2t_1} \\ -2e^{-2t_2} & e^{-2t_2} \end{bmatrix} \quad (46)$$

será igual a zero. Isto mostra que existem infinitos pares de valores p_1 e p_2 que satisfazem as condições do enunciado.

Se por exemplo escolhermos $p_1=10$, o correspondente valor de p_2 será igual a

$$p_2 = \frac{-\ddot{x}(t) - p_1 \dot{x}(t)}{x(t)} \quad (47)$$

$$p_2 = \frac{-4e^{-2t} + 20e^{-2t}}{e^{-2t}} = 16 \quad (48)$$

O modelo seria:

$$\ddot{x} + 10\dot{x} + 16x = 0 \quad (49)$$

e para as condições iniciais $x(t_0)=1$ e $\dot{x}(t_0)=-2$ a solução seria $S_o(t) = S_m(t) = x(t) = e^{-2t}$.

A viabilidade de resposta completa o enunciado. Pretende-se com isto verificar a unicidade de resposta, pois o bojeto efetivamente tem parâmetros bem definidos. De nada nos' adiantaria pesquisar os parâmetros do modeo se o enunciado do problema nos levar matematicamente a infinitas soluções compatíveis.

A próxima fase da elaboração do modelo é a identificação que descorreremos abaixo.

2.2.2 - IDENTIFICAÇÃO:

O enunciado em si é uma etapa qualitativa. Uma vez enunciado o modelo passaremos a uma etapa quantitativa ' que consiste em determinar os parâmetros do modelo enunciado.

A esta etapa denominaremos de identificação. Aqui em geral, se usa métodos numéricos implementados em computador digital. Veremos posteriormente que tais métodos, nada mais' são que métodos numéricos de minimização. Um pequeno estudo a respeito destes métodos será realizado mais adiante.

Nesta fase procuraremos encontrar parâmetros do modelo de tal forma que $D_s(0, M)$ seja mínimo.

2.2.3 - INTERPRETAÇÃO FÍSICA:

Foi falado anteriormente que o método do modelo se diferenciava dos demais, em virtude de estudar modelos paramétricos estruturais e os parâmetros terem significação física.

A literatura aborda outros métodos de identificação citando como exemplo de em sistemas lineares de preten

der determinar a resposta impulsiva. Realmente a determinação da resposta impulsiva se enquadra dentro de nosso conceito que identificar é encontrar um modelo matemático (resposta impulsiva) cuja performance seja a mesma da do objeto. Usa-se em alguns casos, processos estocásticos. Entretanto a resposta impulsiva nada nos diz a respeito dos parâmetros e da estrutura do modelo - tal como concebemos. Em inúmeros casos o conhecimento da resposta impulsiva é suficiente. Entretanto, em certos casos estamos interessados em saber o valor dos parâmetros do objeto, citando o caso de um amortecedor de uma máquina por exemplo, cujo valor de B (coeficiente de viscosidade) deverá estar dentro de uma determinada faixa.

Fazemos uma identificação, e se o valor de B estiver fora da faixa será então necessário fazer uma manutenção. Este é apenas um exemplo dentro de diversos para justificar a interpretação física dos parâmetros e a importância do método do modelo.

Nesta etapa de execução do modelo, verificamos se os parâmetros determinados na fase de identificação estão dentro de faixas aceitáveis. Por exemplo, no sistema elétrico abordado anteriormente jamais poderemos encontrar um valor para o resistor com sinal negativo.

Este é um fato que corriqueiramente aparece em identificação e a interpretação física dos resultados deve sempre ser levada em consideração.

Passaremos agora à última fase de elaboração do modelo que é a verificação.

2.2.4 - VERIFICAÇÃO:

É o teste final das fases anteriores. Em verdade na fase identificação jamais conseguiremos encontrar $D_s(0, M) = 0$ por motivos já abordados e que adiante retornaremos com mais detalhes.

Se o problema é de resposta viável, matematicamente $D_s(0, M) = 0$ se $0 = M$. Como é muito difícil senão im -

possível encontrar $D_s(0, M) = 0$ na prática procuramos minimizar a distância de saída de tal forma que seja menor ou igual a um valor positivo D_{s0} especificado, ou seja $D_{s(0, M)} \leq D_{s0}$.

Isto implica que existe um conjunto de vetores $p(t)$ que satisfazem $D_{s(0, M)} \leq D_{s0}$.

No fundo não procuramos em identificação o valor exato dos parâmetros do modelo e sim um conjunto de parâmetros P que façam a distância de saída ser menor ou igual a uma isodistância especificada

Obtemos na fase identificação um conjunto de respostas próximas da verdadeira para uma determinada excitação. Procuraremos na verificação observar o desempenho do modelo para outras excitações quando for possível. Se o desempenho não for bom, o que em geral ocorre, então procedemos como será descrito a seguir.

Para a excitação $u_1(t)$ identificamos um conjunto de parâmetros que denominaremos P_1 . Se for possível identificamos o mesmo modelo para outra excitação $u_2(t)$ e encontraremos outro conjunto de parâmetros que denominaremos P_2 . A melhor resposta corresponderá ao conjunto intersecção de P_1 com P_2 conforme mostra a fig. 3.

Poderemos fazer este procedimento sucessivamente no sentido de tornar o conjunto de parâmetros resposta bem pequeno. Ocorre muitas vezes na prática que dependendo da entrada, não existe intersecção, devendo-se assim escolher entradas convenientes.

Em processos com parâmetros variáveis com o tempo, o procedimento nem sempre é possível. Isto decorre do fato que para diferentes grupos de dados práticos, temos diferentes evoluções dos parâmetros variáveis com o tempo, já que os intervalos de observação são diferentes.

Assim, para a excitação $u_1(t)$ temos um determinado valor para cada parâmetro do objeto e para a excitação $u_2(t)$ outros valores, uma vez que é impossível realizar dois experimentos no mesmo tempo. Em tais casos, o problema é resolvido se tivermos certeza que para cada intervalo de observação a

evolução dos parâmetros é a mesma, constituindo esta hipótese, uma hipótese de parâmetros adicional.

Caso não seja possível fazer a hipótese acima, resolvemos o problema para um único experimento e consideramos uma resposta de parâmetros que nos dê $D_S(0, M) \leq D_{S0}$ sendo D_{S0} o menor valor possível que se pode encontrar.

2.3 - ESPAÇO PARAMÉTRICO:

2.3.1 - DEFINIÇÃO:

Denominamos de "espaço paramétrico" ao espaço de parâmetros. Objeto e modelo são comparados pela função distância de saída $D_S(0, M)$.

Se o processo apresentar parâmetros variáveis com o tempo, cada parâmetro variável poderá ser decomposto em série (potência, Fourier e etc.) e pode ser calculado por uma soma:

$$p_i(t) = \sum_{j=1}^{\infty} p_{i,j} f_j(t) \quad (50)$$

onde $p_{i,j}$ é uma constante e f_j a j -ésima função da série.

Na prática, j não precisa variar até o infinito. Tomamos um número de componentes N que bem aproxime $p_i(t)$ e podemos escrever:

$$p_i(t) \approx \sum_{j=1}^N p_{i,j} f_j(t) \quad (51)$$

O parâmetro variável com o tempo $p_i(t)$ é então decomposto em um conjunto de parâmetros constantes.

Todos os parâmetros constantes e componentes constantes de parâmetros variáveis constituem o espaço paramétrico.

A dimensão do espaço paramétrico é igual ao número de parâmetros constantes e componentes constantes de parâmetros variáveis envolvidas no problema.

2.3.2 - ANÁLISE DA VIZINHANÇA DO PONTO NOMINAL:

Denominamos de ponto nominal, o ponto no espaço paramétrico que corresponde ao valor nominal dos parâmetros estruturais do objeto.

Pretende-se estudar o desempenho de $D_E(0, M)$ em torno do ponto nominal. Para facilidade vamos supor que os parâmetros do modelo sejam constantes.

A distância da estrutura

$$D_E(0, M) = \sum_{i=1}^N k_i (p_{0,i} - p_i)^2 = \sum_{i=1}^N k_i \Delta p_i^2 \quad (52)$$

ou em forma matricial

$$D_E(0, M) = \Delta p^T A \Delta p \quad (53)$$

sendo A uma matriz definida positiva, $p_{0,i}$ e p_i os parâmetros do objeto e do modelo; N a dimensão do espaço paramétrico e k_i fatores de ponderação.

Podemos afirmar que

$$D_E(0, M) = 0 \quad \text{se e somente se} \quad M \equiv 0$$

Pretende-se em identificação minimizar $D_E(0, M)$. Se o problema for de resposta viável, isto é, apresentar um único p que zere $D_E(0, M)$, poderemos considerar a distância de saída como sendo uma medida da distância de estrutura. Para o caso de apenas uma saída pode-se escrever:

$$D_s(0, M) = D_s(e(t)) = \langle e(t), e(t) \rangle = \int_{t_0}^{t_f} (s_m(t) - s_o(t))^2 dt \quad (54)$$

Como estamos admitindo que o enunciado é de resposta viável, $D_s(0, M) = 0$ se e somente se $0 \equiv M$. O sinal $s_o(t)$, vem com ruídos e com erros de medida. Por esta razão procuramos definir $D_s(0, M)$ por uma integral, com o propósito de eliminar parte da influência dos ruídos.

Fazendo um desenvolvimento de segunda ordem de $D_s(0, M)$ para um acréscimo de p podemos escrever:

$$D_s(p+dp) = D_s(p) + \sum_{i=1}^N \frac{\partial D_s}{\partial p_i} dp_i + \frac{1}{2} \left[\sum_{i=1}^N \frac{\partial^2 D_s}{\partial p_i^2} dp_i^2 + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{\partial^2 D_s}{\partial p_i \partial p_j} dp_i dp_j \right] \quad (55)$$

Em notação matricial, escrevendo o gradiente como o vetor \underline{G} e a matriz A que será definida a seguir.

$$D_s(p+dp) - D_s(p) = \underline{G}^T \Delta p + \Delta p^T A \Delta p \quad (56)$$

Cada componentes de \underline{G} pode ser escrita:

$$G_i = \frac{\partial D_s}{\partial p_i} = 2 \int_{t_0}^{t_f} (s_m - s_0) \frac{\partial s_m}{\partial p_i} dt = 2 \langle \epsilon \cdot \nabla_i \rangle = 2 \langle \epsilon \cdot \nabla_i \rangle \quad (57)$$

As funções $\nabla_i = \frac{\partial s_m}{\partial p_i}$ são as funções sensibilidade da saída em relação aos parâmetros.

No ponto nominal $\Delta p = [p - p_0]$ e ficamos:

$$D_s(p_0 + \Delta p) = \Delta p^T A \Delta p \quad (58)$$

O hessiano:

$$\begin{aligned} \frac{\partial^2 D_s}{\partial p_i \partial p_j} &= 2 \int_{t_0}^{t_f} (s_m - s_0) \frac{\partial^2 s_m}{\partial p_i \partial p_j} dt = 2 \int_{t_0}^{t_f} \frac{\partial s_m}{\partial p_j} \frac{\partial s_m}{\partial p_i} dt = \\ &= 2 \int_{t_0}^{t_f} \frac{\partial s_m}{\partial p_i} \frac{\partial s_m}{\partial p_j} dt = 2 \int_{t_0}^{t_f} \nabla_i \nabla_j dt \end{aligned} \quad (59)$$

Assim a matriz A tem de ser definida:

$$\begin{bmatrix} \langle \nabla_1, \nabla_1 \rangle & - & - & - \\ & \ddots & & \\ - & - & \langle \nabla_N, \nabla_N \rangle \end{bmatrix} \quad (60)$$

Podemos escrever; $A = \sum Y Y^T$ o que mostra que A é positiva.
Assim:

$$D_s(p_0 + \Delta p) = \Delta p^T A \Delta p = (\Delta p^T Y)(Y^T \Delta p) = [\Delta p^T Y]^2 \quad (61)$$

o que mostra que para qualquer Δp temos uma forma quadrática. Se as funções de sensibilidade existem no ponto nominal e se se escolhe um critério quadrático, as superfícies isodistância, isto significa D_s é constante, são quadráticas fechadas. O estudo das superfícies isodistância são muito importantes em identificação porque permitem estudar o espaço paramétrico em torno do ponto nominal.

2.3.3 - FORMA DAS SUPERFÍCIES ISODISTÂNCIA:

Vimos na secção anterior que em torno do ponto nominal as superfícies isodistância são fechadas, para um critério quadrático.

Uma forma típica de tais superfícies, considerando um espaço paramétrico bidimensional está representada na fig.4, onde N é o ponto nominal.

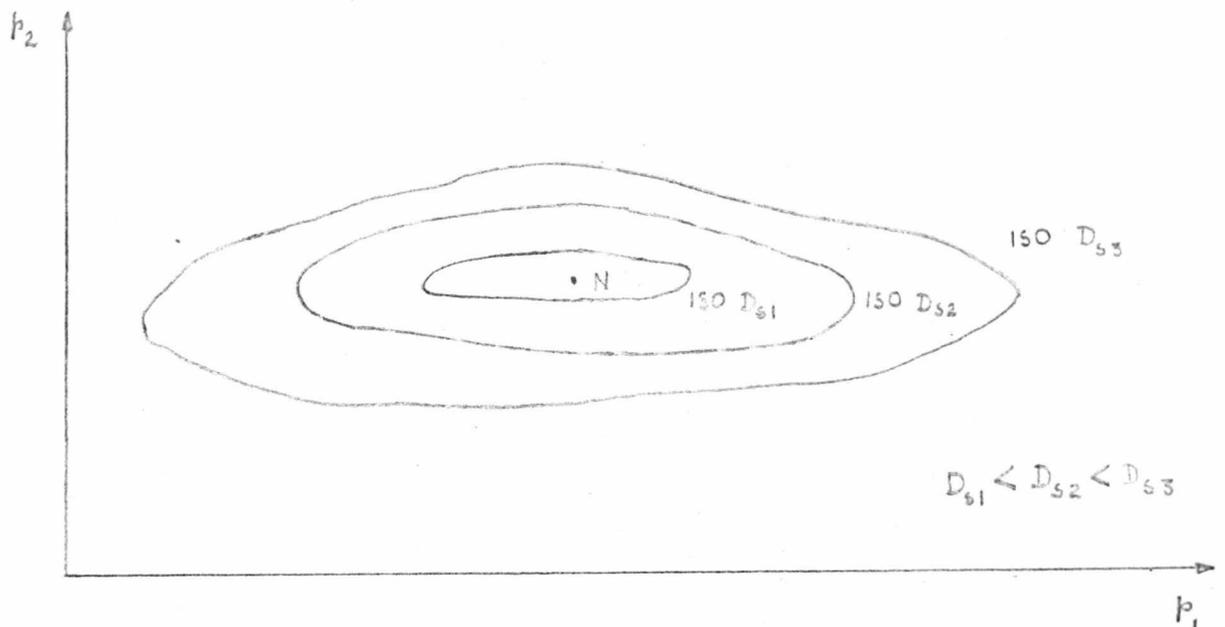


Figura 4.

Podemos observar que para qualquer ponto a componente do gradiente em relação a p_1 , $\frac{\partial D_s}{\partial p_1}$ é bem menor que em relação a p_2 , $\frac{\partial D_s}{\partial p_2}$. Como $G_{\lambda} = 2\langle \epsilon \cdot \nabla_{\lambda} \rangle$, concluímos que ∇_1 é menor que ∇_2 .

Esta diferença de sensibilidade, sob o ponto de vista prático, torna difícil a pesquisa do mínimo.

Comparando a figura anterior com curvas de nível de topografia, podemos associá-la a bacia de um rio, onde as linhas de maior inclinação correspondem ao percurso das águas que se deslocam em direção ao valo, ou seja, o leito do rio.

Comparativamente no nosso estudo, definimos de "vale" ao lugar geométrico dos pontos de inclinação mínima.

Um grande problema da pesquisa do mínimo, ou seja, da etapa identificação, é justamente a presença de vale no espaço paramétrico. Os algoritmos de minimização implementados em computador digital, convergem rapidamente para o mínimo em regiões de grande inclinação, sendo muito lentos em regiões de pouca inclinação. Torna-se evidente que se o problema apresenta um vale, há uma rápida convergência em direção ao vale e uma vez alcançado é extremamente difícil convergir para o ponto nominal.

O problema ideal, sob o ponto de vista de aplicação seria o de vale esférico como está representado pela figura 5, onde

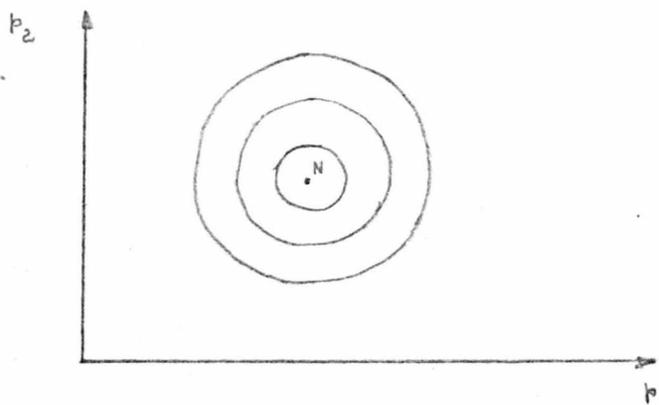


Figura 5

as linhas iso- D_3 são quase esféricas, e neste caso haveria uma rápida convergência. Notar que neste caso, as sensibilidades em relação a p_1 , e a p_2 são praticamente as mesmas.

No entanto vale esférico pouco ocorre na prática, sendo o caso comum de a sensibilidade de p_1 , ser bastante diferente de p_2 , ou seja a presença de um vale alongado.

2.4 - ENTRADAS SENSIBILIZANTE E ESFERIZANTE:

Dizemos que quando um parâmetro é pouco sensível, falta-lhe informação. O fator responsável por informações dadas aos parâmetros é o sinal de entrada.

Os coeficientes de sensibilidade permitem estabelecer uma medida da divisão de informação dada aos parâmetros. Existe em identificação, estudos que tentam, para um dado problema, determinar entrada sensibilizante e esferizante.

Dizemos que a entrada é esferizante se re parte a informação equitativamente para todos os parâmetros e que é sensibilizante se concentra a informação sobre um parâmetro desejado.

2.5 - MÉTODOS DE MINIMIZAÇÃO:

Já vimos que na etapa identificação, é necessário pesquisar um ponto do espaço paramétrico que torna $D_3(0, M) \leq D_{30}$, sendo D_{30} um número real positivo e próximo de zero. Para isto utilizaremos um algoritmo de minimização no sentido de minimizar a distância de saída. Sabe-se que $D_3(0, M)$ é uma função do vetor p por meio da saída do modelo $\hat{y}_m(t)$, então o problema de minimizar D_3 é o problema clássico em matemática da minimização de uma função de vários parâmetros.

Existem vários métodos para resolver este problema, não se pretendendo citar todos. Tentaremos classificá-los e expor o princípio de alguns deles. Tudo o que aqui vai ser dito só é válido para funções unimodais ^{2,4,7,8}.

2.5.1 - MÉTODOS MONOVARIÁVEIS:

São a base de qualquer algoritmo para funções multivariáveis. Existem vários métodos monovariáveis e o mais conhecido e eficiente é o de Fibonacci^{2,4}.

O princípio é de eliminação. Define-se um intervalo de pesquisa $[p_c, p_s]$. Seja $p_1 < p_2$, $p_1, p_2 \in [p_c, p_s]$, $D_3(0, M)$ de p_1 igual a $D_3(0, M)_1$ e $D_3(0, M)$ de p_2 igual a $D_3(0, M)_2$. Então é válido:

- a) Se $D_3(0, M)_1 > D_3(0, M)_2$, o mínimo está em $[p_1, p_s]$;
- b) Se $D_3(0, M)_2 > D_3(0, M)_1$, o mínimo está em $[p_c, p_2]$ como se pode observar na figura 6.



Figura 6.

Reduz-se assim o intervalo de pesquisa. O algoritmo de Fibonacci dá a lógica de escolha de medidas, localizando os pontos de medida equidistantes das extremidades do intervalo. A cada medida associamos um intervalo L . Se n é número de medidas pela fig. 7 podemos tirar o intervalo L_{n-1} .

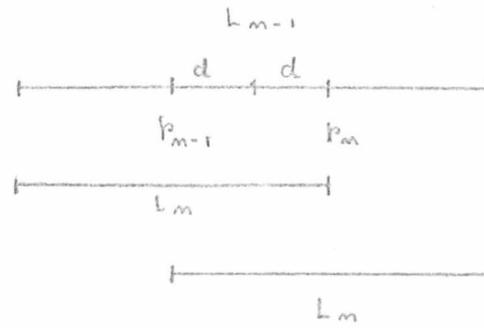


Figura 7

Assim

$$L_m = \frac{L_{m-1}}{2} + \frac{d}{2} \quad (62)$$

Podemos generalizar, e pela fig. 8, po-

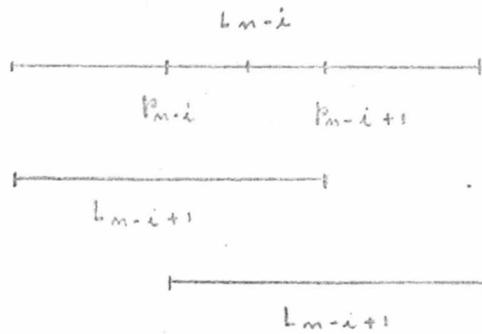


Figura 8

demos tirar

$$L_{m-i} = L_{m-i+1} + L_{m-i+2} \quad (63)$$

Da (62) podemos tirar

$$L_{m-1} = 2L_m - d \quad (64)$$

Das (64) e (63) podemos escrever os intervalos

$$L_m = 1L_m - 0d \quad (65)$$

$$L_{m-1} = 2L_m - 1d \quad (66)$$

$$L_{m-2} = 3L_m - 1d \quad (67)$$

$$L_{m-3} = 5L_m - 2d \quad (68)$$

Definindo

$$F(0) = 1 \quad (69)$$

$$F(1) = 1 \quad (70)$$

$$F(k) = F(k-1) + F(k-2) \quad (71)$$

Pode-se escrever o intervalo genérico

$$L_{m-k} = F(k+1)L_m - F(k-1)d \quad (72)$$

A (72) é a lei do algoritmo, permitindo a localização dos pontos de medida.

O intervalo inicial e o número de medidas que se deseja realizar são os dados para a utilização da (72).

2.5.2 - MÉTODO MULTIVARIÁVEIS DE PRIMEIRA ORDEM:

Seja minimizar a função $f(p)$. Representamos na fig. 9 os isoafastamentos para dois parâmetros.

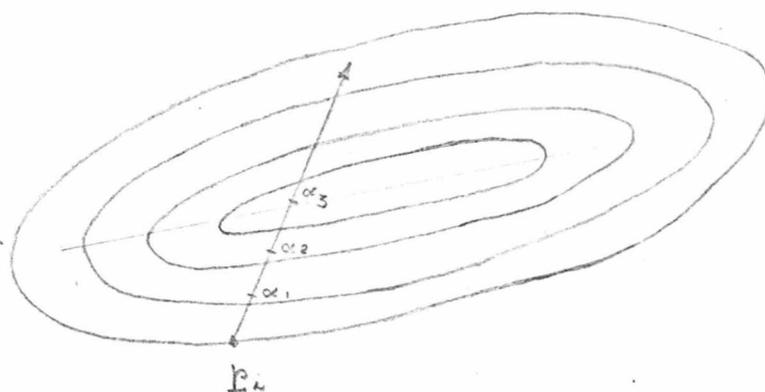


Figura 9.

Dado um ponto inicial p_1 , o princípio de

qualquer método pode ser explicado através da fig. 9. Sempre existem duas etapas:

1. procurar a melhor direção de pesquisa que minimiza D_s ;
2. progredir nesta direção (problema do passo).

Dependendo da função que queremos minimizar (analítica ou numérica), usaremos métodos analíticos ou heurísticos.

2.5.2.1 - MÉTODOS HEURÍSTICOS:

Neste caso não temos conhecimento analítico da função que desejamos minimizar, fazendo-se ensaios para encontrar a direção e progredir nela.

Um exemplo deste tipo é o método de Gauss² que pode ser visualizado na fig. 10. Consiste em variar sucessivamente cada parâmetro. A pesquisa em cada eixo é feita por um método monovariável.

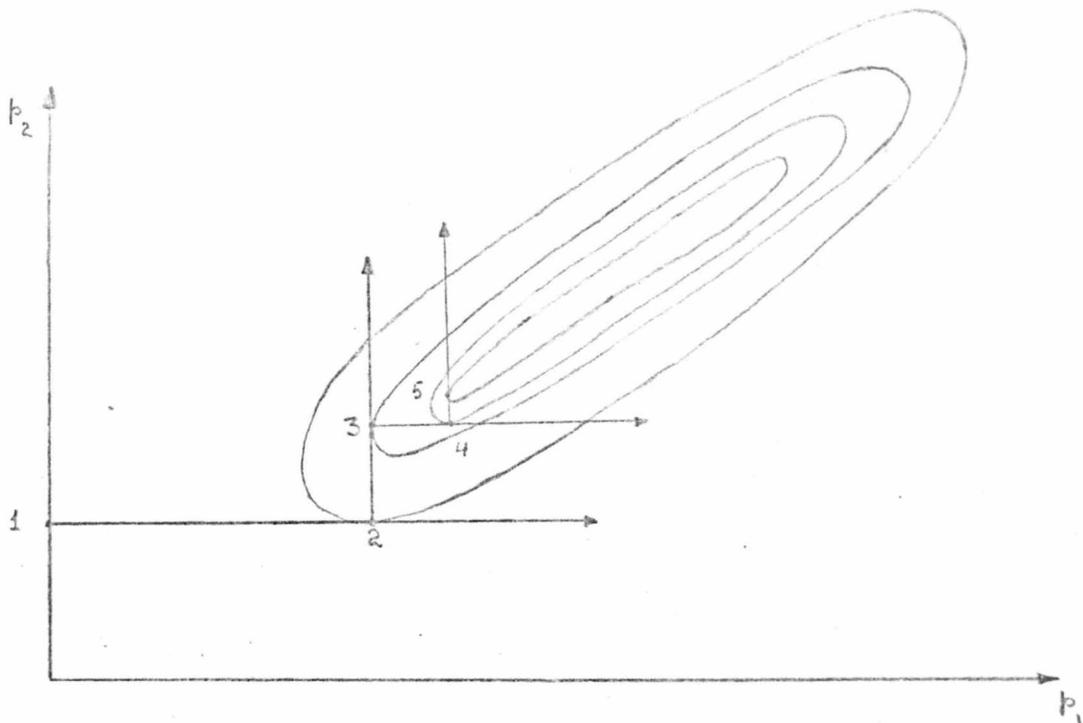


Figura 10

2.5.2.2 MÉTODOS ANALÍTICOS:

Em lugar de encontrar a melhor direção de pesquisa por ensaios, podemos utilizar as propriedades analíticas da função D_s . Um exemplo deste método é o do gradiente. A idéia natural de se procurar o mínimo é seguir a maior inclinação de D_s ⁶. Considerando o desenvolvimento de D_s de 1a. ordem podemos tirar:

$$\Delta D_s = G^T \Delta p \quad (73)$$

Para maximizar ΔD_s basta fazer Δp colinear com G porque o gradiente dá a direção de maior inclinação e assim:

$$\Delta p = -k G \quad (74)$$

onde k é uma constante positiva e G deve ser calculado. Por um algoritmo monovariável calculamos qual o melhor k que minimiza D_s na direção do gradiente. A fig. 11 mostra o método.

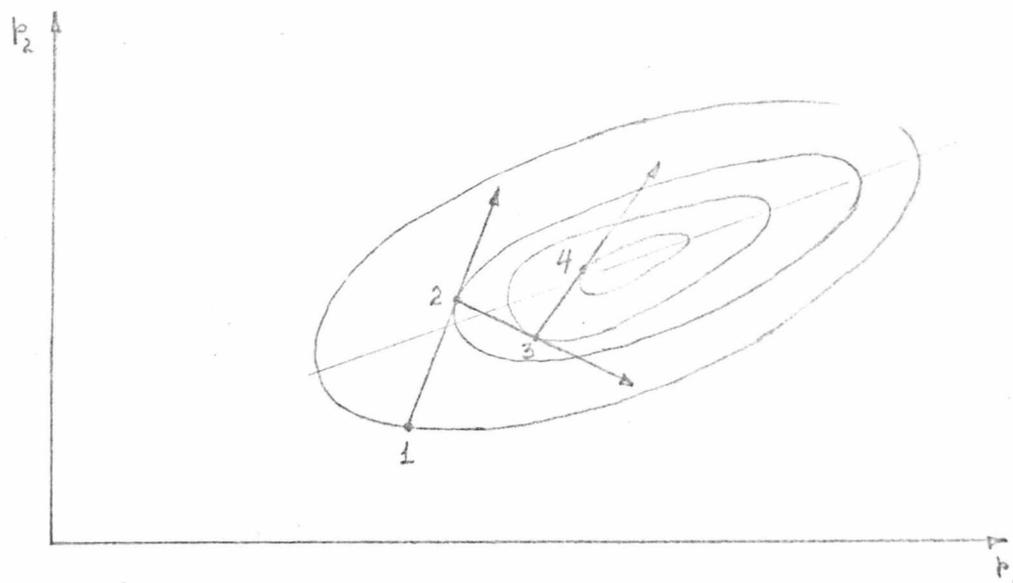


Figura 11

2.5.3 - MÉTODOS MULTIVARIÁVEIS DE SEGUNDA ORDEM:

As propriedades que vamos utilizar são as que provem do desenvolvimento da 2a. ordem de D_s no ponto nominal.

2.5.3.1 - MÉTODOS ANALÍTICOS:

Um exemplo típico é o método de Newton. Vimos que:

$$\Delta D_3 = G^T \Delta p + \Delta p^T A \Delta p \quad (75)$$

O máximo de ΔD_3 ocorre quando sua derivada é nula:

$$0 = G + 2A \Delta p \quad (76)$$

$$\Delta p = -\frac{1}{2} A^{-1} G \quad (77)$$

Esta é a lei de variação de Δp do método de Newton. Calculando-se A^{-1} e G obtem-se a variação de Δp .

Podemos também considerar a forma particular dos isoafastamentos em torno do ponto nominal. Estes isoafastamentos são quadráticos podendo-se usar as propriedades das direções conjugadas, tendo-se assim o método de mesmo nome.

2.5.3.2 - MÉTODOS HEURÍSTICOS:

Como exemplo podemos citar o método de Powell². Exporemos sobre um exemplo de 3 parâmetros deixando de lado a parte matemática. Conforme se pode ver na fig. 12, o algoritmo começa como o método de Gauss com pesquisas sobre cada parâmetro que a partir do ponto inicial A dá o ponto B. Uma pesquisa sobre a direção AB dá o ponto C, a partir do qual começa a segunda iteração sobre as novas direções d_1 , d_2 e d_3 ; d_1 e d_2 são as velhas direções d_2 e d_3 e d_3 é a direção de AB. A segunda iteração dá o ponto D e as novas direções d_1 , d_2 , d_3 onde d_1 é a velha direção d_3 . A direção d_3 . A terceira iteração dá o ponto E que é o mínimo.

2.6. - PROBLEMAS DO MÉTODO DO MODELO:

Identificar é resolver um problema bem prático. Vimos que é necessário fazer medidas no objeto. Devemos medir a excitação $u(t)$, condições iniciais e $y_0(t)$. Toda medida é acompanhada de ruídos e erros. Notadamente os ruídos impedem que $D_3(0,1)$ tenha mínimo zero. Podemos ver no exemplo de primeira

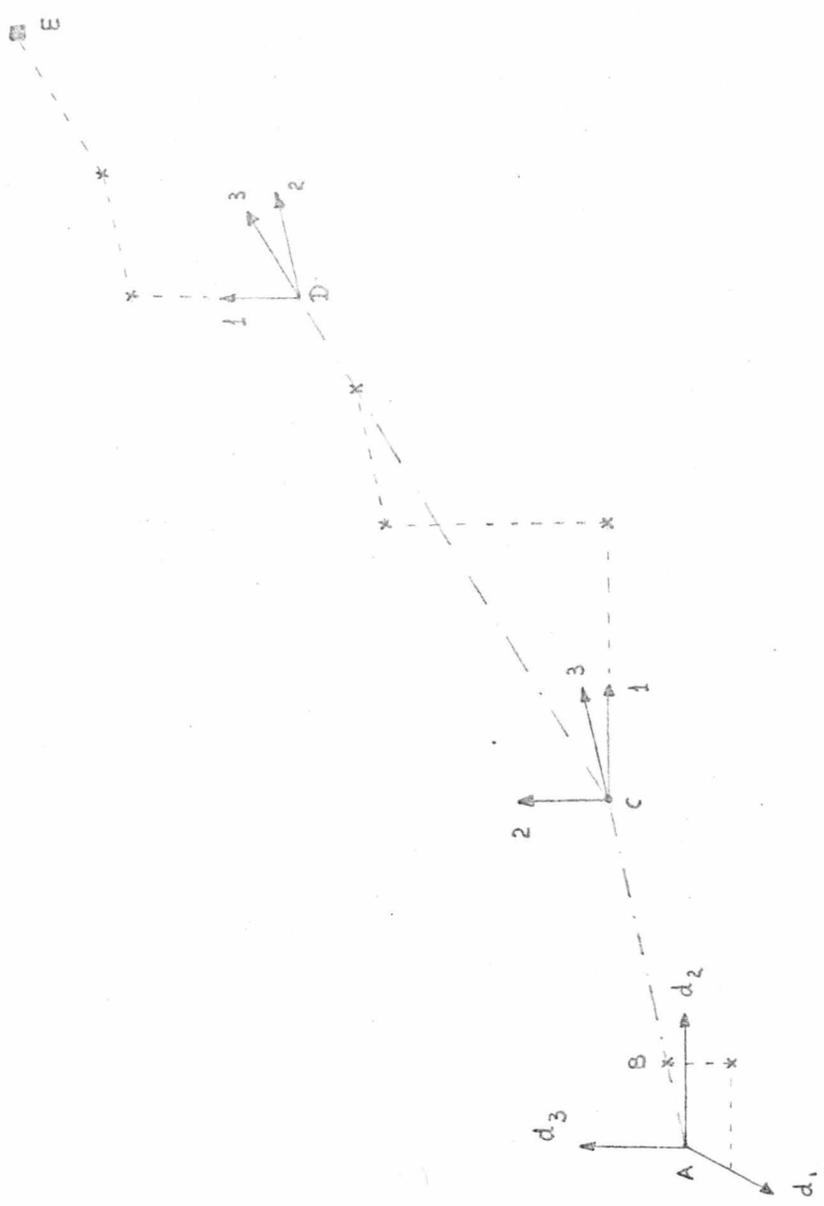


Figura 12

ordem com o parâmetro p constante e a entrada $u(t)$ sendo um salto:

$$\dot{x}(t) + p x(t) = u(t) \quad (78)$$

$$s_m(t) = x(t) \quad (79)$$

$$x(t_0) = 0 \quad (80)$$

A fig. 13 mostra a saída $s_o(t)$ medida e uma boa saída do modelo.

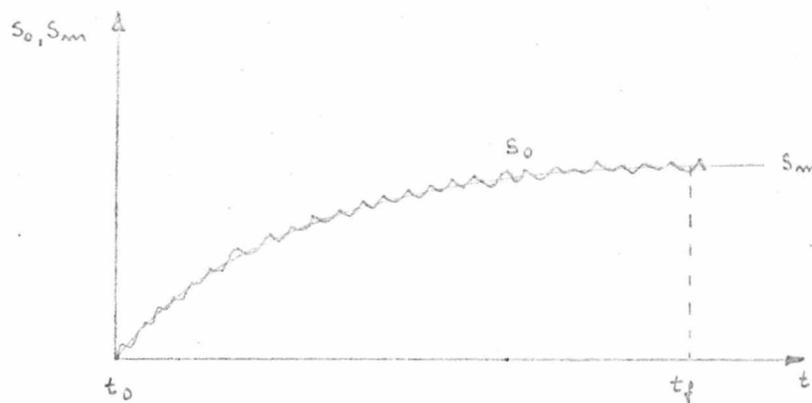


Figura 13

Vê-se pela figura que:

$$D_s(0, M) = \int_{t_0}^{t_f} (s_m(t) - s_o(t))^2 dt \neq 0 \quad (81)$$

Podemos eliminar parte dos ruídos filtrando o sinal por circuito ou numericamente.

De forma semelhante os erros de medida levam $D_s(0, M)$ ser diferente de zero para uma ótima saída do modelo. Assim não se pode nunca esperar encontrar $D_s(0, M) = 0$ e como já foi falado procura-se fazer $D_s(0, M) \leq D_{s0}$ o que implica que existem um grupo de vetores p resposta em torno no ponto nominal. Assim em identificação não se procura um vetor p ótimo mas sim um conjunto ótimo, podendo-se considerar isto como uma vantagem.

Outro problema é a existência de um vale no espaço paramétrico. Todos os métodos de minimização são de rápida convergência em pontos distantes do vale. A pesquisa do mínimo perto do ponto nominal é bastante lenta e esta é outra razão'

pela qual se procura parar o processo de pesquisa do mínimo quando $D_3(0,1) \leq D_{c_0}$.

Além do mais, cada problema requer um método ' de otimização conveniente, não existindo até o presente, um algrítimo rápido para todos os problemas.

O inconveniente principal, sob o ponto de vista de dificuldade, é o enunciado do modelo quando pouco se conhece a estrutura do objeto. Não nos cabe outra saída que a de formulação de hipóteses.

Um especialista no setor do objeto que se pretende identificar poderá ser de grande validade na formulação dessas hipóteses. Fato é que existem poucas regras para execução do enunciado do modelo e uma boa prática em identificação ajuda muito.

Como último problema poderemos citar a entrada. Deve-se escolher uma entrada que bem sensibilize os parâmetros, isto facilita a minimização e diminui o conjunto resposta. A entrada deve conter bastante informação.

CAPÍTULO 3

REVISÃO DO CÁLCULO VARIACIONAL

Cálculo variacional, bem como programação 'dinâmica em otimização pretende minimizar uma função escalar de custo.

Uma importante aplicação é a determinação 'da entrada ótima de um sistema conhecido seu modelo matemático e definindo uma função de custo que se quer minimizar.

A aplicação do cálculo variacional nestes problemas leva a resolução de um problema de duas condições de contorno.

As equações obtidas podem ser postas sob a forma de variáveis de estado, tendo-se como vantagem a fácil integração. Por outro lado, formam em geral um sistema não linear e instável e por este fato é bastante difícil a resolução do problema.

A seguir daremos algumas definições básicas, princípios do cálculo variacional e uma aplicação, para com base nisto aplicarmos cálculo variacional em problemas de identificação ^{3,5}.

3.1. - DEFINIÇÃO DE FUNCIONAL:

Um funcional J é uma regra que associa a cada função x , pertencente a uma classe Ω , um único número real.

Denominamos Ω de domínio de J .

3.2. - DEFINIÇÃO DE NORMA DE UMA FUNÇÃO:

Denominamos de norma de uma função a uma regra que associa cada função $x, x \in \Omega$ e definida em $t \in [t_0, t_f]$, um número real. A norma de x denotada por $\|x\|$, satisfaz as seguintes propriedades:

1. $\|x\| \geq 0$

(82)

$$2. \quad \|\underline{x}\| = 0 \quad \text{se e somente se} \quad \underline{x}(t) = 0 \quad \forall t \in [t_0, t_f] \quad (83)$$

$$3. \quad \|\alpha \underline{x}\| = |\alpha| \cdot \|\underline{x}\| \quad \forall \alpha \in \mathbb{R} \quad (84)$$

$$4. \quad \|\underline{x}^{(1)} + \underline{x}^{(2)}\| \leq \|\underline{x}^{(1)}\| + \|\underline{x}^{(2)}\| \quad (85)$$

3.3 - DEFINIÇÃO DE INCREMENTO DE UM FUNCIONAL:

Se J é definido para \underline{x} e $\underline{x} + \delta \underline{x}$, então o incremento de J , denotado por ΔJ é

$$\Delta J = J(\underline{x} + \delta \underline{x}) - J(\underline{x}) \quad (86)$$

3.4 - DEFINIÇÃO DE LINEARIDADE DE UM FUNCIONAL:

J é linear se e somente se satisfaz as seguintes propriedades:

$$1. \quad J(\alpha \underline{x}) = \alpha J(\underline{x}) \quad \forall \underline{x} \in \Omega \quad \forall \alpha \in \mathbb{R} \quad \text{tal que} \quad \alpha \underline{x} \in \Omega \quad (87)$$

$$2. \quad J(\underline{x}^{(1)} + \underline{x}^{(2)}) = J(\underline{x}^{(1)}) + J(\underline{x}^{(2)}) \quad \forall \underline{x}^{(1)}, \underline{x}^{(2)}, \underline{x}^{(1)} + \underline{x}^{(2)} \in \Omega \quad (88)$$

3.5 - DEFINIÇÃO DE VARIACÃO DE UM FUNCIONAL:

O incremento de um funcional pode ser escrito como:

$$\Delta J(\underline{x}, \delta \underline{x}) = \delta J(\underline{x}, \delta \underline{x}) + g(\underline{x}, \delta \underline{x}) \cdot \|\delta \underline{x}\| \quad (89)$$

onde δJ é linear em $\delta \underline{x}$ se

$$\lim_{\|\delta \underline{x}\| \rightarrow 0} \{g(\underline{x}, \delta \underline{x})\} = 0 \quad (90)$$

então J é dito ser diferenciável em \underline{x} e δJ é a variação de J em torno da função \underline{x} .

3.6 - MÁXIMO E MÍNIMO DE UM FUNCIONAL:

Um funcional J com domínio Ω tem um extremo

relativo z^* se existe $\epsilon > 0$ tal que para toda função $z \in \Omega$ que satisfaz $\|z - z^*\| < \epsilon$, o incremento de J tem o mesmo sinal.

se

$$\Delta J = J(z) - J(z^*) \geq 0 \quad (91)$$

então $J(z^*)$ é um mínimo relativo. Se

$$\Delta J = J(z) - J(z^*) \leq 0 \quad (92)$$

então $J(z^*)$ é um máximo relativo.

Se (91) se satisfaz para ϵ arbitrariamente grande então $J(z^*)$ é um mínimo global ou absoluto.

Se (92) se satisfaz para ϵ arbitrariamente grande então $J(z^*)$ é um máximo global ou absoluto.

Chamamos z^* de extremal e $J(z^*)$ de extremo.

3.7 - TEOREMA FUNDAMENTAL DO CÁLCULO VARIACIONAL:

Se z^* é um extremal então

$$\delta J(z^*, \delta z) = 0 \quad \forall \delta z \text{ admissível.} \quad (93)$$

3.8 - LEMA FUNDAMENTAL DO CÁLCULO VARIACIONAL:

Se $h(t)$ é arbitrária e $|h(t)| < \epsilon$ a integral

$$\int_{t_0}^{t_1} g(t) h(t) dt \quad (94)$$

se anula se $g(t) = 0$, em $[t_0, t_1]$.

Prova:

Provaremos por contradição:

Suponhamos que $g(t) \neq 0$. Como $h(t)$ é arbitrária podemos supor que $h(t) = k g(t)$ onde k é uma constante. Tiramos que:

$$\int_{t_0}^{t_1} g(t) h(t) dt = \int_{t_0}^{t_1} k g^2(t) dt = k \int_{t_0}^{t_1} g^2(t) dt \neq 0 \quad (95)$$

Isto ocorre porque $g^2(t)$ é uma função positiva e $g(t) \neq 0$ e então a integral $\int_{t_0}^{t_1} g^2(t) dt > 0$.

Só é possível então $\int_{t_0}^{t_1} g(t) h(t) dt = 0$ para $h(t)$ arbitrária, se $g(t) = 0$.

3.9 - PROBLEMA DE OTIMIZAÇÃO:

Vamos agora aplicar os conceitos vistos anteriormente neste capítulo para resolver o problema que a seguir enunciaremos:

Seja

$$\dot{x}(t) = b(x(t), u(t), t) \quad (96)$$

O modelo matemático de um sistema físico. Este modelo está sob a forma de variáveis de estado, onde o vetor $x(t)$ é a variável de estado e o vetor $u(t)$ é a excitação ou entrada do sistema.

Definido um critério de custo

$$J = h(x(t_f), t_f) + \int_{t_0}^{t_f} g(x(t), u(t), t) dt \quad (97)$$

pretende-se determinar a entrada ótima $u^*(t)$ que minimize o critério.

Se h for diferenciável e se $x(t_0) = x_0$ for conhecido demonstra-se que J será mínimo se forem satisfeitas as quatro condições necessárias relacionadas abaixo:

$$\dot{x}^*(t) = \frac{\partial b}{\partial x} (x^*(t), u^*(t), \Delta^*(t), t) \quad (98)$$

$$\dot{p}^*(t) = - \frac{\partial g}{\partial x} (x^*(t), u^*(t), \Delta^*(t), t) \quad (99)$$

$$0 = \frac{\partial g}{\partial u} (x^*(t), u^*(t), \Delta^*(t), t) \quad (100)$$

$$\left[\frac{\partial H}{\partial z} (z^*(t_f), t_f) - \lambda^*(t_f) \right]^T \delta z +$$

$$+ \left[\mathcal{H} (z^*(t_f), u^*(t_f), \lambda^*(t_f), t_f) + \frac{\partial H}{\partial t} (z^*(t_f), t_f) \right] \delta t_f = 0 \quad (101)$$

onde \mathcal{H} que denomina-se de Hamiltoniano é definido como:

$$\mathcal{H} (z(t), u(t), \lambda) \triangleq g(z(t), u(t), t) + \lambda^T [b(z(t), u(t), t)] \quad (102)$$

O vetor $\lambda(t)$ denomina-se de multiplicador de Lagrange. A equação (98) é chamada de equação de estado e a (99) de coestado: Os vetores $z^*(t)$, $u^*(t)$, $\lambda^*(t)$ são respectivamente os vetores ótimos de $z(t)$, $u(t)$ e $\lambda(t)$ que tornam J mínimo.

Estas equações podem ser resolvidas por computador digital. Vejamos o exemplo de um problema particular onde se conhece o tempo final t_f e o estado final z_f .

Como t_f é conhecido $\delta t_f = 0$ e o segundo termo da (101) será igual a zero. Como z_f é conhecido a variação $\delta z_f = 0$ e o primeiro termo da (101) será zero. Para a (101) ser satisfeita já que conhecemos t_f e z_f basta que

$$z^*(t_f) = z_f \quad (103)$$

que é uma condição de contorno. É condição de validade das condições necessárias que seja conhecido:

$$z(t_0) = z_0 \quad (104)$$

que é outra condição de contorno que deve ser satisfeita.

Se possível isolar $u^*(t)$ da (100) de tal forma que

$$u^*(t) = a(z^*(t), \lambda^*(t), t) \quad (105)$$

podemos substituir a (105) nas equações (98) e (99) ficando com:

$$\dot{z}^*(t) = \omega(z^*(t), \lambda^*(t), t) \quad (106)$$



$$\dot{p}^*(t) = \lambda^*(t) [u^*(t), \delta^*(t), t] \tag{107}$$

Podemos implementar no computador digital - as equações (106) e (107). Estamos em presença de um problema de duas condições de contorno dadas pela (103) e pela (104). Conhecemos $z^*(t_0)$. Devemos então determinar $\delta^*(t_0)$ de tal forma que o sistema formado pela (106) e (107) evolua para $z^*(t_f) = z_f$. Para isto é necessário um algoritmo de minimização.

Determinado $\delta^*(t_0)$ pode-se determinar $z^*(t)$ e $\lambda^*(t)$ que substituindo na (105) obtém-se a entrada ótima $u^*(t)$ que minimiza a função custo especificada.

CAPÍTULO 4

UTILIZAÇÃO DO CÁLCULO VARIACIONAL EM IDENTIFICAÇÃO.

Neste capítulo apresentaremos uma utilização do cálculo variacional em problemas de identificação pelo método do modelo.

Serão abordados aspectos teóricos concernentes ao cálculo variacional tendo como objetivo a identificação. Veremos que, com esta técnica, é possível identificar parâmetros constantes, sendo este estudo, sob o ponto de vista teórico, bem generalizado.

Neste capítulo serão determinadas condições que o cálculo variacional impõe em identificação e nos capítulos precedentes serão abordadas as aplicações das conclusões tiradas deste estudo que trataremos agora.

4.1 - IDENTIFICAÇÃO DE UM SISTEMA COM PARÂMETROS VARIÁVEIS COM O TEMPO:

O problema que se pretende resolver utilizando cálculo variacional pode ser enunciado como segue:

Dado:

$$\dot{x}(t) = b(x(t), p(t), u(t), t) \quad (108)$$

que é um sistema de equação diferenciais, sob a forma de equações de estado, que representa o modelo. O vetor $u(t)$ é o vetor excitação. O vetor $x(t)$ é o vetor estado do modelo e, $p(t)$ o vetor parâmetro.

Se $s_o(t)$ e $s_m(t)$ são respectivamente os vetores saída do objeto e do modelo quando excitados por $u(t)$, deseja-se determinar $p^*(t)$ ótimo de tal forma que o vetor saída ótima do modelo $s_m^*(t)$ se aproxime ao máximo de $s_o(t)$, também conhecido.

Para isto deve-se minimizar uma "função escalar de custo" que pretende medir a distância de estrutura, que num caso bem geral pode ser posta sob a forma abaixo, sendo t_0 e t_f os limites do intervalo de observação:

$$D_c(o, m) = \int_{t_0}^{t_f} g(s_o(t), s_m(t), p(t), t) dt \quad (109)$$

O ideal seria escolher $D_c(0, M) = D_s(0, M)$ que é o critério clássico. No entanto, tal escolha torna o problema degenerado invalidando a aplicação do C.V. Por esta razão utiliza-se para custo $D_c(0, M)$.

Considerando,

$$\underline{s}_m(t) = f(\underline{v}(t), \underline{u}(t), t) \quad (110)$$

podemos escrever:

$$\begin{aligned} D_c(0, M) &= \int_{t_0}^{t_f} k(\underline{s}_0(t), \underline{x}(t), p(t), \underline{u}(t), t) dt = \\ &= \int_{t_0}^{t_f} h(\underline{x}(t), p(t), t) dt \end{aligned} \quad (111)$$

já que $\underline{u}(t)$ e $\underline{s}_0(t)$ são conhecidos.

Da (108) podemos tirar, já que $\underline{u}(t)$ e conhecido:

$$a(\underline{x}(t), p(t), t) - \dot{\underline{x}}(t) = 0 \quad (112)$$

Podemos definir:

$$\begin{aligned} h_a(\underline{x}(t), \dot{\underline{x}}(t), p(t), \Delta(t), t) &= \\ = h(\underline{x}(t), p(t), t) + \Delta^T(t) [a(\underline{x}(t), p(t), t) - \dot{\underline{x}}(t)] \end{aligned} \quad (113)$$

Definindo:

$$D_c(0, M)_a = \int_{t_0}^{t_f} h_a(\underline{x}(t), \dot{\underline{x}}(t), p(t), \Delta(t), t) dt \quad (114)$$

e tendo em vista a (112) notamos que, se as restrições impostas pelas equações de estado são satisfeitas $D_c(0, M)_a = D_c(0, M)$ para qualquer função $\Delta(t)$.

A função $\Delta(t)$, chamada de multiplicador de Lagrange, é uma função arbitrária e sua finalidade é agrupar as restrições de estado à função custo.

O incremento de $D_c(0, M)_a$ conforme a definição 3.3, em torno de $p(t)$, $\underline{x}(t)$, $\dot{\underline{x}}(t)$ e $\Delta(t)$ será:

$$\begin{aligned}
 \Delta \mathcal{D}_c(0, M)_a &= (\underline{x}(t), \dot{\underline{x}}(t), \underline{p}(t), \underline{\Delta}(t), \delta \underline{x}(t), \delta \dot{\underline{x}}(t), \delta \underline{p}(t), \delta \underline{\Delta}(t), t) = \\
 &= \mathcal{D}_c(0, M)_a(\underline{x}(t) + \delta \underline{x}(t), \dot{\underline{x}}(t) + \delta \dot{\underline{x}}(t), \underline{p}(t) + \delta \underline{p}(t), \underline{\Delta}(t) + \delta \underline{\Delta}(t), t) - \\
 &- \mathcal{D}_c(0, M)_a(\underline{x}(t), \dot{\underline{x}}(t), \underline{p}(t), \underline{\Delta}(t), t)
 \end{aligned}
 \tag{115}$$

$$\begin{aligned}
 \Delta \mathcal{D}_c(0, M)_a &= \int_{t_0}^{t_f} h_a(\underline{x}(t) + \delta \underline{x}(t), \dot{\underline{x}}(t) + \delta \dot{\underline{x}}(t), \underline{p}(t) + \delta \underline{p}(t), \underline{\Delta}(t) + \delta \underline{\Delta}(t), t) dt - \\
 &- \int_{t_0}^{t_f} h_a(\underline{x}(t), \dot{\underline{x}}(t), \underline{p}(t), \underline{\Delta}(t)) dt
 \end{aligned}
 \tag{116}$$

Expandindo o primeiro t ermo em s erie de Taylor em torno do ponto $\underline{x}(t)$, $\dot{\underline{x}}(t)$, $\underline{p}(t)$ e $\underline{\Delta}(t)$ vem :

$$\begin{aligned}
 \Delta \mathcal{D}_c(0, M)_a &= \int_{t_0}^{t_f} \left\{ \cancel{h_a(\underline{x}(t), \dot{\underline{x}}(t), \underline{p}(t), \underline{\Delta}(t), t)} + \right. \\
 &+ \left[\frac{\partial h_a}{\partial \underline{x}}(\underline{x}(t), \dot{\underline{x}}(t), \underline{p}(t), \underline{\Delta}(t), t) \right]^T \delta \underline{x}(t) + \\
 &+ \left[\frac{\partial h_a}{\partial \dot{\underline{x}}}(\underline{x}(t), \dot{\underline{x}}(t), \underline{p}(t), \underline{\Delta}(t), t) \right]^T \delta \dot{\underline{x}}(t) + \\
 &+ \left[\frac{\partial h_a}{\partial \underline{p}}(\underline{x}(t), \dot{\underline{x}}(t), \underline{p}(t), \underline{\Delta}(t), t) \right]^T \delta \underline{p}(t) + \\
 &+ \left[\frac{\partial h_a}{\partial \underline{\Delta}}(\underline{x}(t), \dot{\underline{x}}(t), \underline{p}(t), \underline{\Delta}(t), t) \right]^T \delta \underline{\Delta}(t) + \\
 &+ \left[\text{TERMOS N. L. EM } \delta \underline{x}(t), \delta \dot{\underline{x}}(t), \delta \underline{p}(t), \delta \underline{\Delta}(t) \right] - \\
 &\left. - \cancel{h_a(\underline{x}(t), \dot{\underline{x}}(t), \underline{p}(t), \underline{\Delta}(t), t)} \right\} dt
 \end{aligned}
 \tag{117}$$

Quando $\delta \underline{z}(t)$, $\delta \dot{\underline{z}}(t)$, $\delta \underline{p}(t)$, $\delta \Delta(t)$ se aproximam de zero os termos n~ao lineares tamb~em se aproximam de zero. Pela de finic~ao 3.5, tomando os termos lineares podemos escrever:

$$\begin{aligned} \delta D_c(0, M)_a = & \int_{t_0}^{t_f} \left\{ \left[\frac{\partial h_a}{\partial \underline{z}} (\underline{z}(t), \dot{\underline{z}}(t), \underline{p}(t), \Delta(t), t) \right]^T \delta \underline{z}(t) + \right. \\ & + \left[\frac{\partial h_a}{\partial \dot{\underline{z}}} (\underline{z}(t), \dot{\underline{z}}(t), \underline{p}(t), \Delta(t), t) \right] \delta \dot{\underline{z}}(t) + \\ & + \left[\frac{\partial h_a}{\partial \underline{p}} (\underline{z}(t), \dot{\underline{z}}(t), \underline{p}(t), \Delta(t), t) \right] \delta \underline{p}(t) + \\ & \left. + \left[\frac{\partial h_a}{\partial \Delta} (\underline{z}(t), \dot{\underline{z}}(t), \underline{p}(t), \Delta(t), t) \right] \delta \Delta(t) \right\} dt \end{aligned} \quad (118)$$

O termo,

$$\int_{t_0}^{t_f} \left[\frac{\partial h_a}{\partial \dot{\underline{z}}} (\underline{z}(t), \dot{\underline{z}}(t), \underline{p}(t), \Delta(t), t) \right]^T \delta \dot{\underline{z}}(t) dt \quad (119)$$

pode ser posto em funç~ao de $\delta \underline{z}(t)$ fazendo uma integraç~ao por partes e d~a:

$$\begin{aligned} & \int_{t_0}^{t_f} \left[\frac{\partial h_a}{\partial \dot{\underline{z}}} (\underline{z}(t), \dot{\underline{z}}(t), \underline{p}(t), \Delta(t), t) \right]^T \delta \dot{\underline{z}}(t) dt = \\ & = \left[\frac{\partial h_a}{\partial \dot{\underline{z}}} (\underline{z}(t), \dot{\underline{z}}(t), \underline{p}(t), \Delta(t), t) \right]^T \delta \underline{z}(t) \Big|_{t_0}^{t_f} - \\ & - \int_{t_0}^{t_f} \frac{d}{dt} \left[\frac{\partial h_a}{\partial \dot{\underline{z}}} (\underline{z}(t), \dot{\underline{z}}(t), \underline{p}(t), \Delta(t), t) \right]^T \delta \underline{z}(t) dt \end{aligned} \quad (120)$$

Substituindo a (120) na (118) e considerando o teorema 3.7, vem:

$$\begin{aligned} \delta D_c(0, M)_a = & \left[\frac{\partial h_a}{\partial \underline{z}} (\underline{z}^*(t_f), \dot{\underline{z}}^*(t_f), \underline{p}^*(t_f), \Delta^*(t_f), t_f) \right]^T \delta \underline{z}(t_f) - \\ & - \left[\frac{\partial h_a}{\partial \dot{\underline{z}}} (\underline{z}^*(t_0), \dot{\underline{z}}^*(t_0), \underline{p}^*(t_0), \Delta^*(t_0), t_0) \right]^T \delta \underline{z}(t_0) + \end{aligned}$$

$$\begin{aligned}
& + \int_{t_0}^{t_f} \left\{ \left[\frac{\partial h_a}{\partial \underline{y}} (\underline{y}^*(t), \dot{\underline{y}}^*(t), \underline{p}^*(t), \underline{\lambda}^*(t), t) \right]^T - \right. \\
& \quad \left. - \frac{d}{dt} \left[\frac{\partial h_a}{\partial \dot{\underline{y}}} (\underline{y}^*(t), \dot{\underline{y}}^*(t), \underline{p}^*(t), \underline{\lambda}^*(t), t) \right]^T \right] \delta \underline{y}(t) + \\
& \quad + \left[\frac{\partial h_a}{\partial \underline{p}} (\underline{y}^*(t), \dot{\underline{y}}^*(t), \underline{p}^*(t), \underline{\lambda}^*(t), t) \right]^T \delta \underline{p}(t) + \\
& \quad \left. + \left[\frac{\partial h_a}{\partial \underline{\lambda}} (\underline{y}^*(t), \dot{\underline{y}}^*(t), \underline{p}^*(t), \underline{\lambda}^*(t), t) \right]^T \delta \underline{\lambda}(t) \right\} dt = 0
\end{aligned}
\tag{121}$$

Esta expressão como está apresentada, torna-se inútil se encararmos sob o ponto de vista de aplicação prática. Podemos fazer simplificações e chegar a condições que forçará $\delta D_c(0, M)_a$ se anular.

O raciocínio usado constituirá em anular os termos internos e externos ao sinal de integral tendo em vista o lema fundamental do cálculo variacional.

Existem condições que anulam os termos fora do sinal de integral, e que serão abordadas no próximo capítulo. Assim para que $\delta D_c(0, M)_a = 0$ é necessário que:

$$\begin{aligned}
& \left[\frac{\partial h_a}{\partial \underline{y}} (\underline{y}^*(t_f), \dot{\underline{y}}^*(t_f), \underline{p}^*(t_f), \underline{\lambda}^*(t_f), t_f) \right]^T \delta \underline{y}(t_f) - \\
& - \left[\frac{\partial h_a}{\partial \dot{\underline{y}}} (\underline{y}^*(t_0), \dot{\underline{y}}^*(t_0), \underline{p}^*(t_0), \underline{\lambda}^*(t_0), t_0) \right]^T \delta \underline{y}(t_0)
\end{aligned}
\tag{122}$$

Sendo possível encontrar condições que satisfazem a (122), resta assim anular os termos internos ao sinal de integral. Para isto consideremos primeiramente o termo transposto que multiplica $\delta \underline{\lambda}(t)$.

Este termo é:

$$\left[\frac{\partial h_a}{\partial \underline{\lambda}} (\underline{y}^*(t), \dot{\underline{y}}^*(t), \underline{p}^*(t), \underline{\lambda}^*(t), t) \right]^T
\tag{123}$$

Substituindo a (113) na (123) vem:

$$\underline{a}(\underline{u}^*(t), p^*(t), t) - \dot{\underline{z}}^*(t) \quad (124)$$

Identificamos a (124) como a restrição de estado, que obviamente deverá ser satisfeita, sendo então igual ao vetor nulo, o que torna o termo que multiplica $\delta \underline{\Delta}(t)$ nulo. Obtivemos assim outra condição para que $\delta \mathcal{D}_c(0, M)_a$ se anule e que é:

$$\dot{\underline{z}}^*(t) = \underline{a}(\underline{u}^*(t), p^*(t), t) \quad (125)$$

Se esta condição é satisfeita então $\mathcal{D}_c(0, M)_a = \mathcal{D}_c(0, M)$, e tudo que está sendo válido para $\mathcal{D}_c(0, M)_a$ também é válido para $\mathcal{D}_c(0, M)$. Assim, as condições que encontrarmos para minimizar $\mathcal{D}_c(0, M)_a$ serão as mesmas que minimizam $\mathcal{D}_c(0, M)$.

Consideremos agora o termo que multiplica $\delta \underline{z}(t)$:

$$\left[\frac{\partial h_a}{\partial \underline{z}}(\underline{u}^*(t), \dot{\underline{z}}^*(t), p^*(t), \underline{\Delta}^*(t), t) \right]^T - \frac{d}{dt} \left[\frac{\partial h_a}{\partial \dot{\underline{z}}}(\underline{u}^*(t), \dot{\underline{z}}^*(t), p^*(t), \underline{\Delta}^*(t), t) \right]^T \quad (126)$$

Substituindo a (113) na (126) vem:

$$\left[\frac{\partial h}{\partial \underline{z}}(\underline{u}^*(t), p^*(t), t) \right]^T + \underline{\Delta}^{*T}(t) \left[\frac{\partial a}{\partial \underline{z}}(\underline{u}^*(t), p^*(t), t) \right] - \frac{d}{dt} \left[-\underline{\Delta}^*(t) \right]^T \quad (127)$$

Como o multiplicador de Lagrange é arbitrário podemos escolhê-lo de tal forma que o coeficiente transposto de $\delta \underline{z}(t)$ seja igual ao vetor nulo, e então obtemos mais uma condição que anula $\delta \mathcal{D}_c(0, M)_a$ que é:

$$\underline{\Delta}^*(t) = - \left[\frac{\partial a}{\partial \underline{z}}(\underline{u}^*(t), p^*(t), t) \right]^T \underline{\Delta}^*(t) - \frac{\partial h}{\partial \underline{z}}(\underline{u}^*(t), p^*(t), t) \quad (128)$$

Resta pois anular o termo que multiplica $\delta p(t)$. Para isto nos basearemos no lema fundamental do cálculo variacional. Já que é possível encontrar condições que anulam os termos fora da integral e os termos internos que multiplicam $\delta \underline{\Delta}(t)$ e $\delta \underline{z}(t)$ então:

$$\delta \mathcal{D}_c(0, M)_a = \int_{t_0}^{t_f} \left[\frac{\partial h_a}{\partial p}(\underline{u}^*(t), \dot{\underline{z}}^*(t), p^*(t), \underline{\Delta}^*(t), t) \right] \delta p(t) dt = 0 \quad (129)$$

Se $p(t)$ é um vetor de n componentes, podemos escrever:

$$\delta D_c(0, M)_a = \int_{t_0}^{t_f} \left\{ \sum_{i=1}^n \frac{\partial h_a}{\partial p_i} (\underline{u}^*(t), \dot{\underline{u}}^*(t), p^*(t), \Delta^*(t), t) \delta p_i(t) \right\} dt = 0 \quad (130)$$

Como os $\delta p_i(t)$ são independentes e $\delta D_c(0, M)_a$ deve ser zero para qualquer $\delta p(t)$, podemos escolher todos os $\delta p_i(t)$ exceto $\delta p_i(t)$ iguais a zero. Assim,

$$\delta D_c(0, M)_a = \int_{t_0}^{t_f} \left\{ \frac{\partial h_a}{\partial p_i} (\underline{u}^*(t), \dot{\underline{u}}^*(t), p^*(t), \Delta^*(t), t) \delta p_i \right\} dt = 0 \quad (131)$$

Como $\delta p_i(t)$ é arbitrária, podemos aplicar o lema fundamental e então o coeficiente de $\delta p_i(t)$ deve ser nulo entre $[t_0, t_f]$.

Repetindo o mesmo argumento para os demais $\delta p_i(t)$ concluímos:

$$\frac{\partial h_a}{\partial p} (\underline{u}^*(t), \dot{\underline{u}}^*(t), p^*(t), \Delta^*(t), t) = 0 \quad (132)$$

Substituindo a (113) na (132) obtemos a última condição de anulação de $\delta D_c(0, M)_a$:

$$\frac{\partial h}{\partial p} (\underline{u}^*(t), p^*(t), t) + \left[\frac{\partial \Omega}{\partial p} (\underline{u}^*(t), p^*(t), t) \right] p^*(t) = 0 \quad (133)$$

As equações (122), (125), (128) e (133) constituem as condições necessárias que anulam $\delta D_c(0, M)$, ou seja condições que permitem encontrar um extremal para $D_c(0, M)$.

Definindo uma função $\mathcal{H}(\underline{u}(t), p(t), \Delta(t), t)$, chamada de Hamiltoniano, por:

$$\mathcal{H}(\underline{u}(t), p(t), \Delta(t), t) \triangleq h(\underline{u}(t), p(t), t) + \Delta^T(t) \left[\Omega(\underline{u}(t), p(t), t) \right], \quad (134)$$

podemos escrever as condições (125), (128) e (133) como segue:

$$\dot{\underline{u}}^*(t) = \frac{\partial \mathcal{H}}{\partial \underline{u}} (\underline{u}^*(t), p^*(t), \Delta^*(t), t) \quad (135)$$

$$\dot{\lambda}^*(t) = - \frac{\partial \mathcal{H}}{\partial \underline{u}} (\underline{u}^*(t), p^*(t), \Delta^*(t), t) \quad (136)$$

$$\frac{\partial \mathcal{H}}{\partial p} (\underline{u}^*(t), p^*(t), \Delta^*(t), t) = 0 \quad (137)$$

As condições necessárias que permitem achar um extremal para $D_c(0, M)$ são as (122), (135), (136) e (137).

As equações (135) e (136) são equações diferenciais e a (137) são relações algébricas. Veremos no próximo capítulo que a (122) nos fornecerá condições de contorno.

4.2 - ENTRADA SENSIBILIZANTE E ESFERIZANTE:

Outra aplicação do cálculo variacional, seria a determinação de entradas sensibilizantes ou esferizantes, que foram abordadas no capítulo 2.

Isto, não é objetivo deste trabalho e nada de muito concreto podemos adiantar, apenas ficando aqui a idéia de uma possível aplicação neste campo.

Como ponto de partida, poderíamos iniciar com um sistema de segunda ordem com parâmetros constantes e tentar encontrar uma entrada que sensibilize apenas um dos parâmetros. Para isto pode-se usar as condições (98) a (101) e as restrições seriam as equações diferenciais que regem o sistema (o modelo) e as equações diferenciais de sensibilidade que facilmente pode-se deduzir do modelo.

O problema seria minimizar então a sensibilidade de um dos parâmetros, mas a dificuldade reside em se achar uma conveniente função de custo. Para o caso de sistemas com parâmetros constantes talvez até seja possível a utilização das equações de Riccati, o que facilitaria sobremaneira a solução do problema. A pesquisa de entrada sensibilizante ou esferizante, pode ser tema de uma tese.

CAPÍTULO 5

APLICAÇÃO DE CÁLCULO VARIACIONAL EM IDENTIFICAÇÃO

Podemos aplicar as conclusões do capítulo anterior em identificação de sistemas com parâmetros variáveis com o tempo usando o método do modelo, cujas bases foram descritas no capítulo 2.

5.1 - CONSIDERAÇÕES SOBRE IDENTIFICAÇÃO DE UMA CLASSE DE SISTEMAS VARIÁVEIS COM O TEMPO:

A classe que iremos abordar é a definida abaixo:

$$p_0(t) x^{(m)}(t) + p_1(t) x^{(m-1)}(t) + \dots + p_m(t) x(t) = u(t) \quad (138)$$

onde $p_i(t)$ são parâmetros variáveis com o tempo, $x(t)$ é a saída e $u(t)$ a excitação.

Em identificação pelo método do modelo podemos associar a equação acima ao modelo de um objeto com parâmetros variáveis com o tempo. Este ficaria assim enunciado:

Hipóteses de equação:

$$p_0(t) x^{(m)}(t) + p_1(t) x^{(m-1)}(t) + \dots + p_m(t) x(t) = u(t) \quad (139)$$

Sendo a saída $s_m(t) = x(t)$.

Hipóteses de parâmetro:

$p_0(t), \dots, p_m(t)$ são parâmetros variáveis com o tempo, não tendo nenhuma relação entre si conhecidas.

Dados práticos:

O ideal seria obter vários grupos de dados práticos, ou seja, excitar o objeto com $u_1(t), \dots, u_m(t)$ e medir as correspondentes condições iniciais $x_1^{(m-1)}(t_0), \dots, x_1(t_0), \dots, x_m^{(m-1)}(t_0), \dots, x_m(t_0)$ e saídas $s_{01}(t), \dots, s_{0m}(t)$. Porém, as medidas de cada grupo são feitas em intervalos de observação dife-

rentes e nada nos assegura que os parâmetros sejam os mesmos em cada intervalo, já que são variáveis com o tempo.

Assim, o objeto varia para diferentes intervalos de observação.

Isto implica, que identificar um objeto com parâmetros variáveis com o tempo, só tem sentido em um determinado intervalo de observação e só se podendo fazer medidas em um único grupo, já que é impossível realizar dois ou mais ensaios no mesmo intervalo de tempo.

Concluindo, só é possível excitar o objeto num intervalo $[t_0, t_1]$ com uma única excitação $u(t)$ conhecida e se medir a respectiva saída.

Viabilidade de respostas:

O problema, como estruturado acima, é de caráter prático, porém admite infinitas soluções; o que significa dizer que existem infinitos vetores parâmetros que satisfazem as condições do problema. Isto decorre do fato de ser possível fazer apenas um único experimento prático, ou seja excitar com apenas uma única $u(t)$, obter a resposta no intervalo de observação considerado e medir as condições iniciais.

Para que o enunciado seja viável é necessário que façamos hipóteses de parâmetros adicionais.

Pode-se demonstrar que o enunciado será viável, se se identifica um parâmetro somente, sendo os demais conhecidos.

Fácil de mostrar, se imaginarmos um método elementar de identificação. São conhecidos $u(t)$ e $S_0(t) = x(t)$. Suponhamos que temos um processo de determinação de derivação em relação a t . Determinamos $\dot{x}(t)$, ..., $\overset{(m)}{x}(t)$ a partir de $x(t)$ no intervalo de observação. Ora, se desejamos identificar $p_m(t)$ podemos escrever:

$$p_m(t) = \frac{u(t) - p_0(t) \overset{(m)}{x}(t), \dots, - p_{m-1}(t) \dot{x}(t)}{x(t)} \quad (140)$$

onde $u(t)$ e $x(t)$, ..., $x^{(m)}(t)$ são conhecidos. Para quaisquer $p_1(t)$, ..., $p_{m-1}(t)$ vai existir um $p_m(t)$, se $x(t) \neq 0$ no intervalo de observação. Assim o problema apresenta infinitas soluções que tornam $S_m(t) = S_0(t)$. Porém, se conhecemos $p_0(t)$, $p_1(t)$, ..., $p_{m-1}(t)$, vamos obter o único $p_m(t)$ possível. Assim, o enunciado só é viável em identificação de um parâmetro variável com o tempo na classe de sistemas em estudo.

Além do mais, já é bem conhecido em identificação, que só é possível identificar com sucesso parâmetros variáveis com o tempo, se os parâmetros variam lentamente. Como regra prática, a identificação é possível, se a derivada dos parâmetros em relação ao tempo é maior que um por cento da principal constante de tempo do objeto.

Estudaremos neste capítulo somente a identificação de sistemas de primeira e de segunda ordens pertencentes a classe que estamos considerando.

Introduzidos o modelo e as restrições que de vemos observar, veremos a seguir a idéia de identificar por cálculo variacional bem como os principais problemas de ordem prática que ocorrem.

5.2 - FORMULAÇÃO MATEMÁTICA:

Enunciado um modelo, podemos aplicar os resultados do capítulo anterior. O modelo deverá ser posto sob a forma de equações de estado. Deve-se também definir uma função custo que deveremos minimizar.

Pela equação (134) calcula-se o Hamiltoniano e a seguir aplica-se nas equações (135), (136) e (137).

Como a equação (139) é uma relação algébrica se for possível (de modo geral é), isolamos $p^*(t)$ e substituímos na (135) e (136). Assim ficamos com um conjunto de equações diferenciais na forma de equações de estado, cujo vetor estado é formado pelos vetores $\underline{z}^*(t)$ e $\underline{\lambda}^*(t)$.

Falta considerar a equação (122) que fornecerá condições de contorno. Cada passo será visto com mais detalhes nos itens seguintes do presente capítulo. O descrito acima

é o procedimento clássico de utilização dos resultados do capítulo anterior.

5.3 - AS CONDIÇÕES DE CONTORNO:

As condições de contorno são estabelecidas pela (122). Substituindo a (113) na (122) tiramos:

$$[-\lambda^*(t_f)]^T \delta x(t_f) - [-\lambda^*(t_0)]^T \delta x(t_0) = 0 \quad (141)$$

Em identificação conhecemos as condições iniciais do modelo, ou seja $x(t_0)$ é conhecido.

Isto implica que sempre $\delta x(t_0) = 0$, logo o segundo membro da (141) sempre será igual a zero.

As condições iniciais deverão ser medidas no objeto. Isto, muitas vezes é difícil. Se o objeto estava em repouso, quando aplicamos $u(t)$ para medir a saída $y(t)$, então não é necessário fazer medida das condições iniciais, e podemos assegurar que $x(t_0) = 0$.

A equação (141) fica então reduzida:

$$[\lambda^*(t_f)]^T \delta x(t_f) = 0 \quad (142)$$

Para que esta condição seja satisfeita podemos apresentar 3 casos:

a. O estado final é conhecido:

Neste caso conhecemos $x(t_f)$, logo $\delta x(t_f) = 0$ e a (142) é satisfeita. O estado final poderá ser medido no objeto o que muitas vezes é difícil. Existe uma maneira, de desconhecido o estado final, forçar com que $x^*(t_f)$ se aproxime de $x(t_f)$ que veremos posteriormente.

b. O estado final é desconhecido:

Neste caso, se $\lambda^*(t_f) = 0$ satisfazemos a (142). É uma maneira usual de encontrar o mínimo em otimização, mas para identificação não é uma boa norma, pois com isso não garantimos que haja superposição dos estados finais do objeto e modelo.

c. Algumas componentes do estado final são conhecidas:

Este é um caso mixto. Se a componente $z_2^*(t_f)$ de $z^*(t_f)$ for conhecida, $S_{m_2}(t_f)$ será igual a zero. Se desconhecida, então consideramos $\lambda_2^*(t_f)$ igual a zero e assim satisfazemos a (142). Pelo mesmo motivo do caso anterior, esta não é uma boa norma em identificação.

5.4 - A DISTÂNCIA OBJETO - MODELO:

Sendo o modelo viável, poderíamos medir $D_c(0,M)$ pela distância da saída.

Assim a maneira mais clássica de se medir $D_c(0,M)$ é dada por:

$$D_c(0,M) = D_s(0,M) = \int_{t_0}^{t_f} [(\underline{s}_0(t) - \underline{s}_m(t))^T A(t) (\underline{s}_0(t) - \underline{s}_m(t))] dt \quad (143)$$

onde $A(t)$ é uma matriz de ponderação real, simétrica e semi definida positiva.

No entanto $D_c(0,M)$ definido como (143) nos leva a solução de um problema degenerativo, sob o ponto de vista do cálculo variacional, como podemos mostrar no exemplo que segue. Suponhamos que o modelo é um sistema de primeira ordem,

$$\dot{z}(t) = -p(t)z(t) + u(t) \quad (144)$$

$$S_m(t) = z(t) \quad (145)$$

onde $p(t)$ é o parâmetro $u(t)$ e $S_0(t)$ são conhecidos.

Assim, seguindo o procedimento do item 5-2,

$$D_c(0,M) = \int_{t_0}^{t_f} [(s_0(t) - z(t)) k (s_0(t) - z(t))] dt \quad (146)$$

onde k é real e positivo.

O Hamiltoniano é:

$$\mathcal{H}(z(t), p(t), \lambda(t), t) = k (s_0(t) - z(t))^2 + \lambda(t) (-p(t)z(t) + u(t)) \quad (147)$$

Aplicando nas (135), (136) e (137) vem:

$$\dot{\lambda}^*(t) = -\dot{p}^*(t) z^*(t) + u(t) \quad (148)$$

$$\dot{\lambda}^*(t) = 2 p(t) (s_0(t) - x^*(t)) + p^*(t) \lambda^*(t) \quad (149)$$

$$\lambda^*(t) x^*(t) = 0 \quad (150)$$

Como $x^*(t)$ deverá ser diferente de zero, conclui-se pela equação (150) que $\lambda^*(t) = 0$.

Assim, pela (149) concluímos:

$$s_0(t) = x^*(t) \quad (151)$$

que é uma condição evidente, mas não permite o cálculo de p^* que desejamos.

A equação (150) é $\frac{\partial \mathcal{H}}{\partial p}$ que procuramos forçar ser zero para se obter um extremal de $D_c(0, M)$. Forçando $\frac{\partial \mathcal{H}}{\partial p}$ igual a zero, matematicamente estamos também procurando um extremal para \mathcal{H} . Como \mathcal{H} é linear em relação a p , não existe derivada nula de \mathcal{H} em relação a p e o único extremal possível é para $|p|$ infinito.

O problema pode ser resolvido se pusermos uma restrição adicional:

$$\int_{t_0}^{t_f} b p^2(t) dt \leq \alpha \quad (152)$$

sendo α um número real finito, o que faz com que recaímos na (157).

Se p fosse não linear em relação a \mathcal{H} não haveria este problema. Suponhamos que o modelo seja

$$\dot{x}(t) = -p(t)x(t) + p^2(t)x^2(t) + u(t) \quad (153)$$

$$s_{in}(t) = x(t) \quad (154)$$

Usando $D_c(0, M)$ definido pela (146) definimos:

$$\mathcal{H}(x(t), p(t), \lambda(t), t) = p(t) (s_0(t) - x(t))^2 - \lambda(t) (-p(t)x(t) + p^2(t)x^2(t) + u(t)) \quad (155)$$

$$\frac{\partial \mathcal{H}}{\partial p} = -\lambda^*(t)x^*(t) - 2p^*(t)\lambda^*(t)x^*(t) = 0 \quad (156)$$

Agora podemos isolar $p^*(t)$ e substituir nas correspondentes (135) e (136) e não surge nenhum problema.

O problema é resolvido se redefinirmos $D_c(0,M)$ também em função de $p(t)$.

Definindo

$$D_c(0,M) = \int_{t_0}^{t_f} [(s_0(t) - s_m(t))^T A(t) (s_0(t) - s_m(t)) + p(t)^T R(t) p(t)] dt \quad (157)$$

onde $A(t)$ e $R(t)$ são matrizes de ponderação reais semi definidas positivas, aplicando no exemplo vem:

$$D_c(0,M) = \int_{t_0}^{t_f} [(s_0(t) - x(t)) k (s_0(t) - x(t)) + p(t) b p(t)] dt \quad (158)$$

onde k e b são reais positivos.

A distância do objeto modelo, definida na (157) não é exatamente o que desejamos minimizar, mas é um equivalente do cálculo variacional à $D_b(0,M)$.

A (157) nos afirma que além de minimizarmos a distância de saída, estamos também minimizando os parâmetros. Ora, minimizar os parâmetros não tem significado prático, pois - nosso objetivo é apenas minimizar a distância de saída.

A (157) pode ser aproximada à (143) para convenientes matrizes de ponderação $A(t)$ e $R(t)$. De fato, se $A(t)$ e $R(t)$ são escolhidas de tal forma que a influência de $p(t)^T R(t) p(t)$ seja pequena em relação a $(s_0(t) - s_m(t))^T A(t) (s_0(t) - s_m(t))$, então estamos enfatizando a minimização da distância de saída e como consequência - estamos aproximando a (157) à (143).

No exemplo se escolhessemos $k=1$ e $b=0,001$ estaríamos aproximando a (158) à (146).

Definindo $D_c(0,M)$ conforme a (157) com as restrições impostas às matrizes $A(t)$ e $R(t)$ tornam $D_c(0,M)$ viável em identificação e em cálculo variacional, como mostraremos a seguir.

Voltando ao exemplo, o Hamiltoniano será:

$$\mathcal{H}(u(t), p(t), \lambda(t), t) =$$

$$= k (s_0(t) - x(t))^2 + b p^2(t) + \lambda(t) (-p(t)x(t) + u(t)) \quad (159)$$

Aplicando nas equações (135), (136) e (137) vem:

$$\dot{x}^*(t) = -p^*(t)x^*(t) + u(t) \quad (160)$$

$$\dot{\lambda}^*(t) = 2k(s_0(t) - x^*(t)) + p^*(t)\lambda^*(t) \quad (161)$$

$$2b p^*(t) - \lambda^*(t)x^*(t) = 0 \quad (162)$$

Isolando $p^*(t)$ da (162) vem:

$$p^*(t) = \frac{1}{2b} \lambda^*(t)x^*(t) \quad (163)$$

Substituindo a (163) na (160) e (161) vem:

$$\dot{x}^*(t) = -\frac{1}{2b} \lambda^*(t)x^{*2}(t) + u(t) \quad (164)$$

$$\dot{\lambda}^*(t) = \frac{1}{2b} x^*(t)\lambda^{*2}(t) + 2k(s_0(t) - x^*(t)) \quad (165)$$

Resolvendo a (164) e (165) de tal forma que satisfaçam as condições de contorno impostas pela (122), calcula-se $p^*(t)$ substituindo-se estes resultados na (163). As (164) e (165) estão na forma de equações de estado e as variáveis envolvidas são $x^*(t)$ e $\lambda^*(t)$ sendo fácil sua resolução por métodos numéricos.

Concluimos pois que $D_c(0, M)$ conforme a 158 é viável sob o ponto de vista do cálculo variacional. Entretanto existe outra maneira de definir $D_c(0, M)$, que neste caso não representa mais uma aproximação à distância de saída clássica, sendo apenas uma função custo.

Definimos como segue:

$$D_c(0, M) = \int_{t_0}^{t_f} [p(t)^T R(t) p(t)] dt \quad (166)$$

e adicionamos às restrições de estado mais uma condição. Se o estado $\underline{x}(t)$ tem n componentes definimos:

$$\dot{x}_{(n+1)}(t) = (s_0(t) - s_{nn}(t))^T A(t) (s_0(t) - s_{nn}(t)) \quad (167)$$

com $\dot{x}_{(m+1)}(t_0) = 0$ e $\dot{x}_{(m+1)}(t_f) = 0$, onde $R(t)$ e $A(t)$ são matrizes semidefinidas positivas. Como $(\underline{s}_0(t) - \underline{s}_m(t))^T A(t) (\underline{s}_0(t) - \underline{s}_m(t))$ é $\| \underline{s}_0(t) - \underline{s}_m(t) \|^2_{A(t)}$ então $\dot{x}_{(m+1)}(t)$ só será zero se $\underline{s}_0(t) = \underline{s}_m(t)$.

Pretende-se minimizar $D_c(0, M)$, o que implica em minimizar $\int p^2(t) dt$. Se o enunciado for viável, isto é, apresentar apenas uma solução de $\dot{x}(t)$ que faz $\underline{s}_0(t) = \underline{s}_m(t)$, então poderemos concluir que o $\dot{x}(t)$, que minimiza $D_c(0, M)$ e satisfaz a (167), é o próprio vetor $p_0(t)$ parâmetro do objeto. Aplicando no exemplo vemos que $D_c(0, M)$ conforme a (166) é viável sob o ponto de vista do cálculo variacional. Se $\dot{x}_{(m+1)}(t_f) = 0$ e se o enunciado for viável, então $\delta x(t_f) = 0$ e automaticamente está satisfeita a (122).

Aplicando a (166) e a (167) no exemplo vem:

$$D_c(0, M) = \int_{t_0}^{t_f} b p^2(t) dt \quad (168)$$

$$\dot{x}_2(t) = k (s_0(t) - x_1(t))^2 \quad (169)$$

O Hamiltoniano será:

$$\begin{aligned} \mathcal{H}(x(t), p(t), \lambda(t), t) &= \\ &= b p^2(t) + \lambda_1(t) (-p(t) x_1(t) + u(t)) + \lambda_2(t) [k (s_0(t) - x_1(t))^2] \end{aligned} \quad (170)$$

Aplicando as equações (135), (136) e (137) vem:

$$\dot{x}_1^*(t) = -p^*(t) x_1^*(t) + u(t) \quad (171)$$

$$\dot{x}_2^*(t) = k (s_0(t) - x_1^*(t)) \quad (172)$$

$$\dot{\lambda}_1^*(t) = 2 k \lambda_2^*(t) (s_0(t) - x_1^*(t)) + p^*(t) \lambda_1^*(t) \quad (173)$$

$$\dot{\lambda}_2^*(t) = 0 \quad (174)$$

$$2 b p^*(t) - \lambda_1^*(t) x_1^*(t) = 0 \quad (175)$$

Isolando $p^*(t)$ vem:

$$p^*(t) = \frac{1}{2b} \lambda_1^*(t) x_1^*(t) \quad (176)$$

Substituindo $p^*(t)$ nas (171) e (173) vem:

$$\dot{x}_1^*(t) = -\frac{1}{2b} \lambda_1^*(t) x_1^*(t) + u(t) \quad (177)$$

$$\dot{x}_2^*(t) = k (s_0(t) - x_1(t)) \quad (178)$$

$$\dot{\lambda}_1^*(t) = \frac{1}{2b} \alpha_1^*(t) \lambda_1^{*2}(t) + 2R (S_0(t) - \alpha_1^*(t)) \lambda_2^*(t) \quad (179)$$

$$\dot{\lambda}_2^*(t) = 0 \quad (180)$$

Resolvendo as (177) e (180) de tal forma que satisfaçam a (122), calculamos $\lambda_1^*(t)$ e $\alpha_1^*(t)$ que substituídas na (176) dá a $p^*(t)$ procurada.

Existe outra maneira de definir $D_2(0, \infty)$ que veremos posteriormente, no estudo de um sistema de 2a. ordem.

5.5 - ESPAÇO DE CONDIÇÕES INICIAIS:

Vimos que aplicando as equações (135), (136) e (137) e fazendo as operações indicadas no item 5-2 chegamos a um conjunto de equações diferenciais na forma de estado. Aplicando a (122), estabelecemos $\lambda_i^*(t_0)$ ou $\lambda_i^*(t_1)$ ($i=1, m$).

Conhece-se $x^*(t_0)$ e condições finais. Para resolver o sistema de equações diferenciais, é necessário pesquisar $\lambda^*(t_0)$ que satisfaça as condições finais estabelecidas pela (122). Este é um problema de duas condições de contorno. Conhecido $x^*(t_0)$ é possível resolver as equações diferenciais e está resolvido o problema.

De outra forma, precisamos saber quais condições iniciais satisfazem condições finais.

Ao espaço de condições iniciais que pertencem as condições iniciais que estamos pesquisando para satisfazer as condições finais, denominamos de "espaço de condições iniciais".

Para se fazer esta pesquisa podemos usar um método de minimização. Definimos uma função escalar de custo e procuramos no espaço de condições iniciais o ponto que a minimiza. Esta função é definida de tal forma que achado o seu mínimo estamos satisfazendo as condições finais.

5.6- CUSTO, ALGORÍTIMOS E CRITÉRIO DE PARADA:

A função custo pode ser a mais variada -

possível, dependendo do problema que se está resolvendo.

Como exemplo, podemos definir a função custo em um problema que se conheça as condições finais de estado $\underline{x}(t_f)$ da seguinte forma:

$$C = \| (\underline{x}^*(t_f) - \underline{x}(t_f)) \|_{G(t)}^2 = (\underline{x}^*(t_f) - \underline{x}(t_f))^T G(t) (\underline{x}^*(t_f) - \underline{x}(t_f)) \quad (181)$$

onde $G(t)$ é uma matriz semidefinida positiva.

O problema agora é determinar as condições iniciais que minimizem C pois a prática mostrou que é muito difícil zerar a (181).

Outra maneira é definir,

$$C = \int_{t_0}^{t_f} [(\underline{s}_0(t) - \underline{s}_m(t))^T Q(t) (\underline{s}_0(t) - \underline{s}_m(t))] dt \quad (182)$$

onde $Q(t)$ é uma matriz semidefinida positiva.

Se o enunciado for variável, minimizar (182) significa forçar o estado $\underline{x}^*(t)$ se superpor ao estado do objeto em $[t_0, t_f]$, e como consequência estamos minimizando o estado final ou seja minimizando a (181).

Este foi o critério adotado neste trabalho.

Como consequência podemos ver que não é necessário conhecer $\underline{x}(t_f)$ do objeto, e isto implica em facilidade de obtenção de dados práticos, dispensando a medida do estado final.

A prática mostrou que a minimização de C é um problema bastante difícil. O sistema de equações diferenciais gerados pelo cálculo variacional é muito instável, levando facilmente C para infinito. Este fato não permite a medida de C , tornando impossível a aplicação de qualquer algoritmo de minimização. Foi necessário criar então uma outra função de custo que chamaremos custo preliminar que permitisse a utilização de algoritmos de minimização. Uma vez satisfeita a função custo preliminar, garantimos que estamos perto do mínimo e assim podemos utilizar a função custo desejada.

Com função preliminar podemos definir, se o custo desejado é dado pelo (182):

Custo preliminar é o tempo que falta para a atingir t_f quando alguma componente da saída do modelo atingir um determinado valor real D ou $-D$, sendo $D > 0$, ou seja matematicamente,

$$C_p = (t_f - \theta) \quad \theta \in [t_0, t_f] \quad (183)$$

e θ é o instante onde $s_{mi}(\theta) = D$ ou $s_{mi}(\theta) = -D$ sendo $D > 0$.

Outro critério para definir a função custo preliminar, seria:

Custo preliminar é o tempo que falta para a atingir t_f , quando alguma componente da saída do modelo é igual a $(1+D)$ ou $(1-D)$ vezes a correspondente componente da saída do objeto, sendo $1 > D > 0$.

Matematicamente,

$$C_p = (t_f - \theta) \quad \theta \in [t_0, t_f] \quad (184)$$

e θ é o instante onde $s_{mi}(\theta) = (1+D)s_{oi}(\theta)$ ou $s_{mi}(\theta) = (1-D)s_{oi}(\theta)$, $i = 1, 2, \dots, N$ e N o número de componente de s_{mi} .

Graficamente, considerando condições iniciais I e II. Para o primeiro caso (Fig. 14)

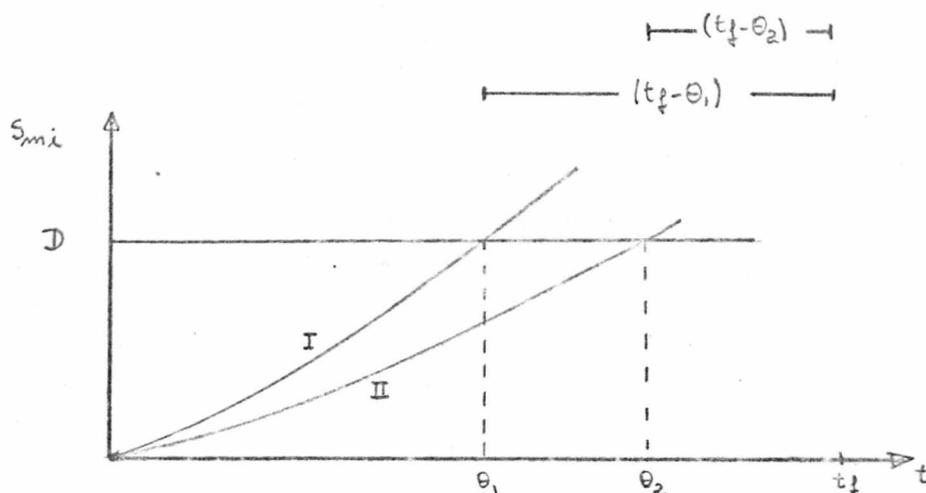


Figura 14

vemos que as condições iniciais II estão mais próximas do mínimo que as I pois $(t_f - \theta_2) < (t_f - \theta_1)$.

No segundo caso (Fig. 15),

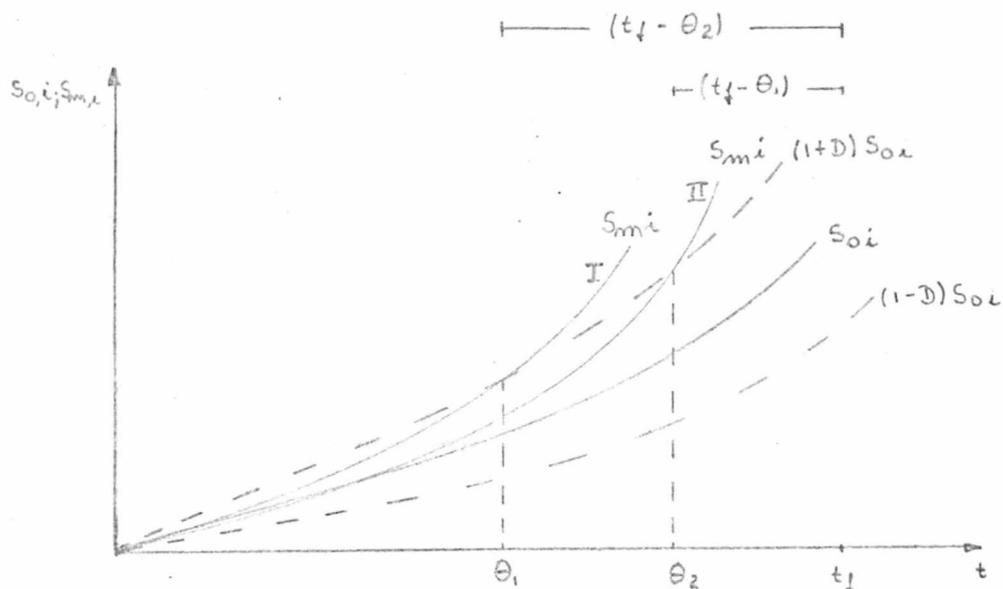


Figura 15

vemos que as condições iniciais II estão mais próximas do mínimo que as I pois $(t_f - \theta_2) < (t_f - \theta_1)$.

Os critérios preliminares permitem assim fazer medidas e a utilização de algoritmos.

Uma vez estando-se perto do mínimo, ou seja, o custo preliminar é satisfeito, podemos medir o custo pela função custo desejada, contornando-se desta forma o problema da instabilidade. O segundo critério de custo preliminar foi o que utilizamos neste trabalho.

Vemos então que a identificação de um sistema com parâmetros variáveis com o tempo, usando cálculo variacional, se reduz a encontrar um ponto no espaço de condições iniciais que minimize C .

Para minimizar C , podemos usar qualquer método de minimização, como por exemplo os expostos no capítulo 2.

Como, em virtude da instabilidade, se utiliza dois critérios de custo, o preliminar e o desejado, deve -

mos tomar alguns cuidados na utilização dos algoritmos de minimização.

Por exemplo, o método de Fibonacci utiliza comparação de medidas. Esta comparação quando feita com um mesmo critério não causa problema. Mas, se para um ponto do espaço de condições iniciais a medida é feita pelo critério preliminar e para outro é satisfeito e a medida é feita pelo custo desejado, então não podemos comparar numericamente as duas medidas e a melhor é a do segundo ponto.

No método do gradiente, precisamos conhecer a derivada parcial do custo em relação a uma componente do vetor condição inicial num determinado ponto. Esta derivada é calculada numericamente pelos valores do custo no ponto e em um ponto bem próximo. Estas duas medidas deverão ser feitas para um mesmo critério, ou o preliminar ou o desejado.

A prática mostrou que é necessário um critério que faça o algoritmo que pesquisa o mínimo de C parar.

Um critério de parada que se poderá usar é o seguinte:

A pesquisa das condições iniciais pára quando $C_k \leq \epsilon$ sendo ϵ um número real, positivo e convenientemente pequeno.

Exposto o método, no próximo item vamos falar sobre suas desvantagens e nos itens subsequentes estudar alguns modelos.

5.7- INSTABILIDADE E PRECISÃO :

Instabilidade e precisão são dois fatores que muito dificultam a pesquisa no espaço de condições iniciais do ponto que minimiza C .

De acordo com o item 5.2 as equações (135) (136) e (137) se reduzem no procedimento clássico, as equações de estado cujo vetor é formado por $y^{(k)}$ e $\lambda^{(k)}$. Estas equações são muito instáveis levando facilmente por isto então é necessário o critério preliminar.

Outra grande desvantagem é a precisão. Para um pequeno intervalo de observação $[t_0, t_f]$ é necessário que as condições iniciais tenham um grande número de algarismos significativos. Como exemplo, para um intervalo de dois segundos é necessário que as condições iniciais tenham em torno de nove algarismos significativos para uma boa resposta. Esta precisão torna a pesquisa muito demorada. Além do mais, os métodos de integração deverão ser mais precisos devendo-se usar um baixo passo de integração tornando-os também demorados.

Estes dois fatores, instabilidade e precisão, tornam a pesquisa difícil e muito demorada.

5.8 - ESTUDO DE UM SISTEMA DE PRIMEIRA ORDEM:

Tomaremos o modelo definido pela equação (144):

$$\dot{x}(t) = -p(t)x(t) + u(t) \quad (185)$$

$$s_m(t) = x(t) \quad (186)$$

onde se conhece $x(t_0)$ e $s_0(t)$.

Definindo

$$D_c(0, M) = \int_{t_0}^{t_f} [k(s_0(t) - x(t))^2 + b p^2(t)] dt \quad (187)$$

onde $k \gg b$.

Aplicando-se as condições impostas pelo cálculo variacional chega-se as equações (164) e (165):

$$\dot{\lambda}^*(t) = -\frac{1}{2b} \lambda^*(t) x^{*2}(t) + u(t) \quad (188)$$

$$\lambda^*(t) = \frac{1}{2b} x^*(t) \lambda^*(t) + 2(s_0(t) - x^*(t)) \quad (189)$$

Definindo

$$\begin{aligned} C &= \int_{t_0}^{t_f} [(s_0(t) - x(t)) q (s_0(t) - x(t))] dt = \\ &= \int_{t_0}^{t_f} q (s_0(t) - x(t))^2 dt \end{aligned} \quad (190)$$

e

$$C_p = (t_f - \theta) \quad \theta \in (t_0, t_f) \quad (191)$$

onde $1 > D > 0$ e $x(\theta) = (1+D) S_0(\theta)$ ou $x(\theta) = (1-D) S_0(\theta)$.

Com um algoritmo de minimização procura-se $\lambda^*(t)$ que faça $C \leq \epsilon$ sendo $\epsilon > 0$ e pequeno.

Achado $\lambda^*(t)$ e $x^*(t)$, $t \in [t_0, t_f]$ pode-se calcular $p^*(t)$ pela (163):

$$p^*(t) = \frac{1}{2b} X^*(t) x^*(t) \quad (192)$$

Este é o procedimento clássico porém apresenta uma desvantagem. O vetor $p^*(t)$ é calculado pela (192). Se $x(t_0) = 0$; isto é, o objeto estava em repouso antes da aplicação de $u(t)$, a (192) faz $p^*(t_0)$ também igual a zero. Ora, se o parâmetro do objeto fosse constante, por exemplo, e igual a 1, o procedimento clássico levaria obrigatoriamente $p^*(t_0) = 0$, portanto diferente de 1. Observa-se que neste caso existe um transitório no início do intervalo de observação, na evolução de $p^*(t)$. O vetor $p^*(t)$ obrigatoriamente parte de zero, cresce até atingir valores em torno de 1, continuando com este desempenho até o final do intervalo.

O transitório ocorre sempre neste caso, embora obtenhamos um valor de C bem pequeno. Isto implica que existem diferentes respostas de $p^*(t)$ que tornam C pequeno e isto é explicado pela sensibilidade dos parâmetros em relação a saída. Se a entrada $u(t)$ contiver pouca informação, a sensibilidade dos parâmetros em relação a saída é pequena, isto significando que existem vários vetores $p^*(t)$ que tornam pequeno o valor de C .

Uma maneira de atenuar o problema é utilizar outra técnica de tratamento das equações (135), (136) e (137) que descreveremos a seguir e que é diferente do procedimento clássico.

Tornando o mesmo modelo definido pela -

(185) e aplicando as equações (135), (136) e (137) e considerando a (187) chega-se as equações (160), (161) e (162):

$$\dot{x}^*(t) = -p^*(t)x^*(t) + u(t) \quad (193)$$

$$\dot{\lambda}^*(t) = 2k(s_0(t) - x^*(t)) + p^*(t)\lambda^*(t) \quad (194)$$

$$2bp^*(t) - \lambda^*(t)x^*(t) = 0 \quad (195)$$

Pelo procedimento clássico deveríamos isolar $p^*(t)$ da (195) e substituir na (193) e (194). O procedimento que propomos é o de isolar $\lambda^*(t)$ da equação (195) e substituir na (194). Assim fica eliminado $\lambda^*(t)$ e ficamos com a equação de estado e uma nova equação diferencial de parâmetro.

Assim da (195):

$$\lambda^*(t) = 2b \frac{p^*(t)}{x^*(t)} \quad (196)$$

Derivando em relação ao tempo:

$$\dot{\lambda}^*(t) = 2b \left[\frac{x^*(t)\dot{p}^*(t) - p^*(t)\dot{x}^*(t)}{x^{*2}(t)} \right] \quad (197)$$

Substituindo as equações (196) e (197) na (194) vem:

$$\dot{p}^*(t) = \frac{k}{b} (s_0(t) - x^*(t)) \frac{p^*(t)}{x^*(t)} + \frac{p^*(t)u(t)}{x^*(t)} \quad (198)$$

Substituindo a (193) na (198) vem:

$$\dot{p}^*(t) = \frac{k}{b} (s_0(t) - x^*(t)) \frac{p^*(t)}{x^*(t)} + \frac{p^*(t)u(t)}{x^*(t)} \quad (199)$$

As equações que escrevemos agora são as (193) e (199).

Tomando C definido pela (190) e C_p pela (191), o problema agora é determinar $p^*(t)$ que minimize C . A equação (199) dá diretamente $p^*(t)$ procurado. Ainda aqui apare-

ce o transitório no início do intervalo de observação quando $x(t_0) = 0$, porém levemente atenuado em comparação com o método clássico.

Outra maneira de se resolver o problema seria, considerando:

$$\dot{x}_1(t) = -p(t)x_1(t) + u(t) \quad (200)$$

$$S_{\text{min}}(t) = x_1(t) \quad (201)$$

onde se conhece $x_1(t_0)$ e $S_0(t)$.

Definindo

$$D_c(0,1) = \int_{t_0}^{t_f} b p^2(t) dt \quad (202)$$

e agrupando a restrição:

$$\dot{x}_2(t) = k(S_0(t) - x_1(t))^2 \quad \text{com} \quad x_2(t_0) = x_2(t_f) = 0. \quad (203)$$

Pelo procedimento clássico chegamos as equações (177) a (180).

Podemos da mesma forma que no caso anterior, eliminar $\lambda_1(t)$ e encontrar uma equação diferencial para o parâmetro.

Da (175) tiramos:

$$\lambda_1^*(t) = 2b \frac{p^*(t)}{x_1^*(t)} \quad (204)$$

Derivando:

$$\dot{\lambda}_1^*(t) = 2b \left[\frac{x_1^*(t) \dot{p}^*(t) - p^*(t) \dot{x}_1^*(t)}{x_1^{*2}(t)} \right] \quad (205)$$

Substituindo as (204) e (205) na (173) vem:

$$\dot{p}^*(t) = \frac{k}{b} \lambda_2^*(t) (S_0(t) - x_1^*(t)) x_1^*(t) + p^{*2}(t) + \frac{p^*(t)}{x_1^*(t)} \dot{x}_1^*(t) \quad (206)$$

Substituindo a (200) na (206) vem:

$$\dot{p}^*(t) = \frac{k}{b} \lambda_2^*(t) (S_0(t) - x_1^*(t)) x_1^*(t) + \frac{p^*(t)}{x_1^*(t)} u(t) \quad (207)$$

As equações diferenciais que vamos trabalhar são:

$$\dot{x}_1^*(t) = -p^*(t)x_1^*(t) + u(t) \quad (208)$$

$$\dot{x}_2^*(t) = p(s_0(t) - x_1^*(t)) \quad (209)$$

$$\dot{p}^*(t) = \frac{d}{dt} \lambda_2^*(t)(s_0(t) - x_1^*(t)) + \frac{p^*(t)}{x_1^*(t)} u(t) \quad (210)$$

$$\lambda_2^*(t) = 0 \quad (211)$$

Definindo:

$$C = \int_{t_0}^{t_f} q(s_0(t) - x_1(t))^2 dt \quad (212)$$

Tomando C_p pela (191), o problema agora consiste em determinar $p^*(t_0)$ e $\lambda_2^*(t_0)$ que minimizem C . A equação (210) dá diretamente $p^*(t)$ procurado. Observe-se que a pesquisa agora deverá ser de duas variáveis tornando o problema mais complexo. Como C definido pela (190) é equivalente a (209), minimizando C estamos minimizando C_p e consequentemente $S_{p_1}(t_f)$.

Poderemos concluir que a minimização da (190) satisfaz as Condições de contorno impostas pela (122).

Passaremos agora ao estudo de um modelo de segunda ordem.

5.9 - ESTUDO DE UM SISTEMA DE SEGUNDA ORDEM:

Tomaremos o modelo:

$$\ddot{x}(t) + v(t)\dot{x}(t) + p(t)x(t) = u(t) \quad (213)$$

$$S_m(t) = x(t) \quad (214)$$

onde se conhece o parâmetro $v(t)$ e $q(t)$, $u(t)$, $x(t_0)$ e $\dot{x}(t_0)$. Em equações de estado, podemos escrever:

$$\dot{x}_1(t) = x_2(t) \quad (215)$$

$$\dot{x}_2(t) = -p(t)x_1(t) - v(t)x_2(t) + u(t) \quad (216)$$

$$S_m(t) = x_1(t) \quad (217)$$

Conhecendo-se $x_1(t_0) = x(t_0)$ e $x_2(t_0) = \dot{x}(t_0)$.

Definindo,

$$D_c(0, M) = \int_{t_0}^{t_1} [k(s_0(t) - x_1(t))^2 + b p^2(t)] dt \quad (218)$$

onde $k \gg b$.

O Hamiltoniano será:

$$\begin{aligned} \mathcal{H}(x_1(t), p(t), \lambda_1(t), \lambda_2(t)) &= k(s_0(t) - x_1(t))^2 + b p^2(t) + \\ &+ \lambda_1(t) x_2(t) (-p(t) x_1(t) - v(t) x_2(t) + \mu(t)) \end{aligned} \quad (219)$$

Aplicando as equações (135), (136) e (137) vem:

$$\dot{x}_1^*(t) = x_2^*(t) \quad (220)$$

$$\dot{x}_2^*(t) = -p^*(t) x_1^*(t) - v(t) x_2^*(t) + \mu(t) \quad (221)$$

$$\dot{\lambda}_1^*(t) = 2k(s_0(t) - x_1^*(t)) + p^*(t) \lambda_2^*(t) \quad (222)$$

$$\dot{\lambda}_2^*(t) = -\lambda_1^*(t) + v(t) \lambda_2^*(t) \quad (223)$$

$$2b p^*(t) - \lambda_2^*(t) x_1^*(t) = 0 \quad (224)$$

Pelo procedimento clássico isolaremos $p^*(t)$ da (224):

$$p^*(t) = \frac{1}{2b} \lambda_2^*(t) x_1^*(t) \quad (225)$$

Substituindo a (225) nas equações (221) e (223) ficamos com as seguintes equações diferenciais cujo vetor estado é formado por $x^*(t)$ e $\lambda^*(t)$.

$$\dot{x}_1^*(t) = x_2^*(t) \quad (226)$$

$$\dot{x}_2^*(t) = -\frac{1}{2b} \lambda_2^*(t) x_1^*(t) - v(t) x_2^*(t) + \mu(t) \quad (227)$$

$$\dot{\lambda}_1^*(t) = 2k(s_0(t) - x_1^*(t)) + \frac{1}{2b} \lambda_2^*(t) x_1^*(t) \quad (228)$$

$$\dot{\lambda}_2^*(t) = -\lambda_1^*(t) + v(t) \lambda_2^*(t) \quad (229)$$

Definindo,

$$C = \int_{t_0}^{t_f} q (s_0(t) - x_1(t))^2 dt \quad (230)$$

e C_p pela (191), com um algoritmo de minimização procura-se $\lambda_1^*(t_0)$ e $\lambda_2^*(t_0)$ de tal forma que $C \leq \epsilon$ sendo ϵ positivo e pequeno.

Minimizando C satisfazemos as condições de contorno e $p^*(t)$ é calculado pela (225).

A dificuldade de encontrar $\lambda_1^*(t_0)$ e $\lambda_2^*(t_0)$ é muito grande. Em nossa pesquisa não conseguimos resolver nenhum exemplo, em virtude da grande instabilidade. Aqui também parece o transitório em $p^*(t)$ quando $x_1(t_0) = 0$ discutido no item anterior.

A função custo (218) deve ter $b \gg \delta$ para que $D_c(O, M)$ se aproxime da distância de saída. Devemos no custo minimizar a influência de $p^*(t)$. Outra maneira de se fazer isto é definir:

$$D_c(O, M) = \int_{t_0}^{t_f} [(s_0(t) - s_m(t))^T A(t) (s_0(t) - s_m(t)) + (z(t) - p(t))^T R(t) (z(t) - p(t))] dt \quad (231)$$

sendo $A(t)$ e $R(t)$ matrizes de ponderação e $z(t)$ um vetor qualquer.

Tomaremos $R(t)$ e $A(t)$ de tal forma que a influência do segundo membro seja pequena em relação ao primeiro dentro do sinal da integral, com o objetivo de aproximar a (231) com a distância de saída (143).

Esta aproximação pode ser melhorada se conhecermos aproximadamente o valor dos parâmetros do objeto. De fato, fazendo $z(t)$ igual aos parâmetros estimados, a (231) se aproximará muito mais da (143) que a (157), pois a (157) é equivalente a (231) se fizermos $z(t) = 0$.

Tomando o modelo definido pelas (215) e (217) e definindo:

$$D_c(O, M) = \int_{t_0}^{t_f} [k(s_0(t) - x_1(t))^2 + b(c(t) - p(t))^2] dt \quad (232)$$

com $b \ll k$, podemos aplicar as condições do cálculo variacional,

O Hamiltoniano será:

$$\begin{aligned} \mathcal{H}(x(t), p(t), \lambda(t), t) = & k(s_0(t) - x_1(t))^2 + b(c(t) - p(t))^2 + \\ & + \lambda_1(t) x_2(t) + \lambda_2(t) [-p(t) x_1(t) - V(t) x_2(t) + u(t)] \end{aligned} \quad (233)$$

Aplicando as equações (135), (136) e (137) vem:

$$\dot{x}_1^*(t) = x_2^*(t) \quad (234)$$

$$\dot{x}_2^*(t) = -p^*(t) x_1^*(t) - V(t) x_2^*(t) + u(t) \quad (235)$$

$$\dot{\lambda}_1^*(t) = 2k(s_0(t) - x_1^*(t)) + p^*(t) \lambda_2^*(t) \quad (236)$$

$$\dot{\lambda}_2^*(t) = -\lambda_1^*(t) + V(t) \lambda_2^*(t) \quad (237)$$

$$-2b(c(t) - p^*(t)) - \lambda_2^*(t) x_1^*(t) = 0 \quad (238)$$

Podemos isolar $p^*(t)$ da (238) e vem:

$$p^*(t) = c(t) + \frac{1}{2b} \lambda_2^*(t) x_1^*(t) \quad (239)$$

Substituindo a (239) nas (235) e (236) as equações diferenciais ficam:

$$\dot{x}_1^*(t) = x_2^*(t) \quad (240)$$

$$\dot{x}_2^*(t) = -(c(t) + \frac{1}{2b} \lambda_2^*(t) x_1^*(t)) - V(t) x_2^*(t) + u(t) \quad (241)$$

$$\dot{\lambda}_1^*(t) = 2k(s_0(t) - x_1^*(t)) + (c(t) + \frac{1}{2b} \lambda_2^*(t) x_1^*(t)) \lambda_2^*(t) \quad (242)$$

$$\dot{\lambda}_2^*(t) = -\lambda_1^*(t) + V(t) \lambda_2^*(t) \quad (243)$$

em função do vetor estado $x^*(t)$ e $\lambda^*(t)$.

Definindo C pela (230) e C_p pela (191) procuramos $\lambda_1^*(t_0)$ e $\lambda_2^*(t_0)$ tais que minimizem C e desta forma

satisfazemos as condições de contorno e $p^*(t)$ é calculado pela (239). Podemos observar que aqui não aparece o transitório em $p^*(t)$ quando temos $\lambda_1(t_0)=0$, discutido no item anterior, sendo neste caso $p^*(t_0)=c(t_0)$.

Este modo de se resolver o problema sugere um método iterativo cujos passos são descritos a seguir:

1. Supõe-se um $c(t)$ e identifica-se $p^*(t)$ que deve estar mais próximo do parâmetro do objeto que o $c(t)$ escolhido.

2. Identifica-se novamente fazendo $c(t)$ igual ao $p^*(t)$ calculado na última identificação. O novo $p^*(t)$ deverá estar mais próximo do parâmetro do objeto que o anteriormente calculado.

3. Repete-se o passo 2 até que seja satisfeito um critério de parada.

Neste procedimento, a medida que vamos identificando, mais aproximamos a (231) da distância de saída. Num caso limite em que $c(t)$ é o próximo valor do objeto, podemos observar pela (239) que $\lambda_2^*(t)=0$.

Uma das maneiras de achar um $c(t)$ inicial no item 1 é pesquisar qual o melhor $c(t)$ constante que minimize C . Assim o espaço de condições iniciais terá três variáveis, $c(t)$ constante, $\lambda_1^*(t_0)$ e $\lambda_2^*(t_0)$.

A prática mostrou que o $p^*(t)$ obtido conforme o procedimento acima, faz com que $p^*(t_0)$ seja diferente do parâmetro do objeto no instante inicial. Assim se no item 2 do método iterativo fizermos $c(t)=p^*(t)+c_0$, sendo $p^*(t)$ calculado na última identificação e c_0 uma constante, podemos atenuar o problema se procurarmos além de $\lambda_1^*(t_0)$ e $\lambda_2^*(t_0)$ também c_0 que minimizem C .

Outra maneira de se resolver o problema, considerando ainda as equações (234) à (238), supondo $c(t)$ desconhecido. Precisamos aproximar $c(t)$ com $p^*(t)$ para que a (231) bem se aproxime da distância de saída. Ora, teoricamente $c(t)$ pode ser qualquer mas com isto não garantimos que a (231) seja próxima da distância de saída.

Podemos usar um artifício, mesmo desconhecendo $c(t)$, que aproxime $c(t)$ com $p(t)$, definindo a seguinte relação entre estes vetores:

$$c(t) = p(t-h) \quad (244)$$

sendo $h > 0$ e bem pequeno.

Graficamente para vetores de uma componente, as evoluções de $c(t)$ e $p(t)$ poderão ser vistas na fig. 16.

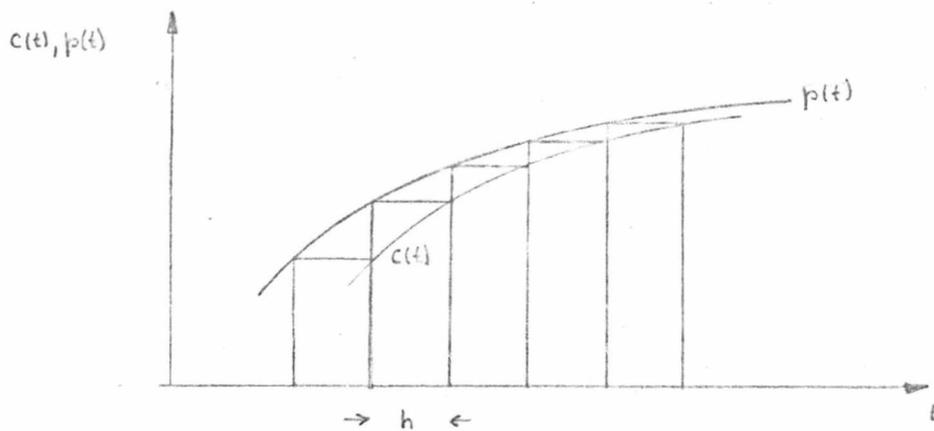


Figura 16.

Desta forma aproximamos $c(t)$ com $p(t)$ e assim aproximamos (231) da distância de saída.

Se da equação (137) pudermos encontrar:

$$p^*(t) = g(c(t), z^*(t), \lambda^*(t), t) \quad (245)$$

podemos digitalizá-la com o incremento h e substituir $p^*(t)$ por $c(t+h)$ e fica:

$$c(t+h) = g(c(t), z^*(t), \lambda^*(t), t) \quad (246)$$

que pode ser calculada conhecendo-se $c(t_0)$.

Precisamos então encontrar $c(t_0)$ e $p^*(t_0)$.

que minimizem C .

O processo de cálculo consiste em calcular $c(t)$ pela (246) substituir na (245) e calcular as equações diferenciais (135) e (136).

Se for possível escrever a (245) como:

$$c(t+h) = c(t) + \frac{1}{h} f(x^*(t), \lambda^*(t), t) \quad (247)$$

então podemos aproximar,

$$\dot{c}(t) \approx \frac{c(t+h) - c(t)}{h} \quad (248)$$

e assim podemos escrever uma equação diferencial:

$$\dot{c}(t) = \frac{1}{h} f(x^*(t), \lambda^*(t), t) \quad (249)$$

que constituem equações diferenciais adicionais.

Considerando agora as equações (234) a (238). Da (234) podemos tirar:

$$p^*(t) = c(t) + \frac{1}{2b} \lambda_2^*(t) x_1^*(t) \quad (250)$$

Fazendo

$$c(t) = p^*(t-h) \quad \text{ou seja} \quad (251)$$

$$c(t+h) = p^*(t) \quad (252)$$

podemos escrever:

$$c(t+h) = c(t) + \frac{1}{2b} \lambda_2^*(t) x_1^*(t) \quad (253)$$

Da (253) aproximamos:

$$\dot{c}(t) = \frac{1}{2bh} \lambda_2^*(t) x_1^*(t) \quad (254)$$

Substituindo a (250) nas equações (235) e (236) as equações diferenciais que devemos trabalhar ficam:

$$\dot{x}_1^*(t) = x_2^*(t) \quad (255)$$

$$\dot{x}_2^*(t) = -\left(c(t) + \frac{1}{2b} \lambda_2^*(t) x_1^*(t)\right) x_1^*(t) - V(t) x_2^*(t) + \mu(t) \quad (256)$$

$$\dot{\lambda}_1^*(t) = 2R(s_0(t) - x_1^*(t)) + (c(t) + \frac{1}{2b} \lambda_2^*(t) x_1^*(t)) \lambda_2^*(t) \quad (257)$$

$$\dot{\lambda}_2^*(t) = -\lambda_1^*(t) + v(t) \lambda_2^*(t) \quad (258)$$

$$\dot{c}(t) = \frac{1}{2bh} \lambda_2^*(t) x_1^*(t) \quad (259)$$

Devemos agora procurar $c(t_0)$, $\lambda_1^*(t_0)$ e $\lambda_2^*(t_0)$ que minimizem C . A pesquisa do mínimo ficou agora mais complexa, pois entra mais a variável $c(t_0)$. Achado o ponto que minimiza C podemos calcular $x_1^*(t)$ pela (250).

Não resta dúvida que o método bem aproxima a (231) da distância de saída. Entretanto, nada nos garante que obtenhamos um bom resultado para mínimo de C , porque estamos forçando $u(t) = p(t-h)$ que não é condição de otimalidade prevista pelo cálculo variacional.

O método é bom para o cálculo de $x_1^*(t)$ inicial do processo iterativo descrito anteriormente, servindo como primeira aproximação quando nada se conhece sobre os parâmetros do modelo.

Neste item só vimos a identificação de apenas um dos parâmetros. Com um grupo de dados é possível identificar os dois se é conhecida uma relação adicional entre os parâmetros.

Se for possível obter dois grupos de dados, ou seja nos dois intervalos de observação os parâmetros tem o mesmo desempenho, então pode-se identificar os dois parâmetros.

Considerando o mesmo modelo definido pelas (213) e (214), as restrições de estado ficariam:

$$\dot{x}_1(t) = x_2(t) \quad (260)$$

$$\dot{x}_2(t) = -p_1(t) x_1(t) - p_2(t) x_2(t) + u_1(t) \quad (261)$$

$$\dot{x}_3(t) = x_4(t) \quad (262)$$

$$\dot{x}_4(t) = -p_1(t)x_3(t) - p_2(t)x_4(t) + u_2(t) \quad (263)$$

onde se conhece $x_1(t_0)$, $x_2(t_0)$ e $S_0(t)$ de uma primeira medida e $x_3(t_0)$, $x_4(t_0)$ e $S_0(t)$ de uma segunda medida.

Basta agora definir $D_c(0, M)$ e seguir o procedimento discutido.

Apresentadas as idéias que permitem a aplicação do cálculo variacional, vamos no próximo capítulo mostrar os resultados práticos utilizando algumas delas bem como as conclusões gerais.

CAPÍTULO 6

RESULTADOS E CONCLUSÕES:

Tendo sido discutido a aplicação do cálculo variacional em identificação no capítulo anterior, passaremos a apresentar resultados práticos de alguns problemas e métodos que julgamos interessantes. Como algoritmo de minimização foi usado o método de Powell e em pesquisa unidimensional o algoritmo de Fibonacci. Para integração foi utilizado Rungue-Kutta de quarta ordem.

6.1. -RESULTADOS:

Em todos os exemplos foi usada a equação (184) para definir a função custo preliminar com D igual a 5% e a equação (182) para definir função custo desejada. O incremento de integração de Rungue Kutta foi em todos os exemplos de 5 centésimos. Os resultados que apresentaremos serão apenas alguns exemplos de sistemas de primeira ordem.

Como primeiro exemplo tomamos um objeto com parâmetro constante e igual a 1 e usamos as equações (188) e (189) para identificar. A equação (187) define y e fizemos $k=1$ e $b=0.01$. A excitação foi constante e igual a 1. A figura 17 mostra os parâmetros real e calculado e a figura 18 as saídas do objeto e modelo.

Como segundo exemplo tomamos o mesmo objeto e a mesma excitação do caso anterior e para identificar usamos as equações (193) e (199). Foi usada também a (187) com $k=1$ e $b=0.001$. A figura 19 mostra os parâmetros e a figura 20 as saídas do objeto e do modelo. Nos dois exemplos a condição inicial é zero e aparece o transitório nos parâmetros identificados nos dois casos.

Como terceiro exemplo tomamos um objeto com parâmetro igual a $0,5t$ e usamos as equações (188) e (189) para identificar. Na (187) fizemos $k=1$ e $b=0.01$. A excitação foi

igual a 1. A figura 21 mostra os parâmetros e a figura 22 mostra as saídas do objeto e modelo.

Como quarto exemplo tomamos um bojeto com parâmetro 1 até 0,7 segundos e daí em diante igual a 1,5. Usamos as equações que deduziremos a seguir. Para $D_c(0, M)$:

$$D_c(0, M) = \int_{t_0}^{t_f} [(s_0(t) - x(t))^2 k + (c(t) - p(t))^2 b] dt \quad (264)$$

O Hamiltoniano será:

$$H(x(t), p(t), \lambda(t), t) = k(s_0(t) - x(t))^2 + b(c(t) - p(t))^2 + \lambda(t)p(t)x(t) + \lambda(t)u(t) \quad (265)$$

Aplicando as condições do cálculo variacional chegamos:

$$\dot{x}^*(t) = -\dot{p}^*(t)x^*(t) + u(t) \quad (266)$$

$$\dot{\lambda}^*(t) = 2k(s_0(t) - x^*(t)) + \dot{\lambda}^*(t)p^*(t) \quad (267)$$

$$\dot{p}^*(t) = c(t) + \frac{1}{2b} \dot{\lambda}^*(t)x^*(t) \quad (268)$$

Fizemos $k=1$ e $b=0.01$. A excitação foi igual a 1 e fizemos $c(t)=0,8$. A figura 23 mostra os parâmetros e a figura 24 mostra as saídas do objeto e do modelo.

Como quinto exemplo tomamos um objeto com parâmetro 1 até 1,3 segundos e daí em diante igual a 1,5. Usamos as mesmas equações (266), (267) e (268). Fizemos $k=1$, $b=0.01$ e a excitação igual a 1. Neste exemplo procuramos as condições iniciais das (266) e (267) como também qual o melhor $c(t)$ constante que minimizam a (182). A figura 25 mostra os parâmetros e a figura 26 as saídas do objeto e do modelo.

6.2 - CONCLUSÃO:

Cumpra salientar que existem muitos traba-

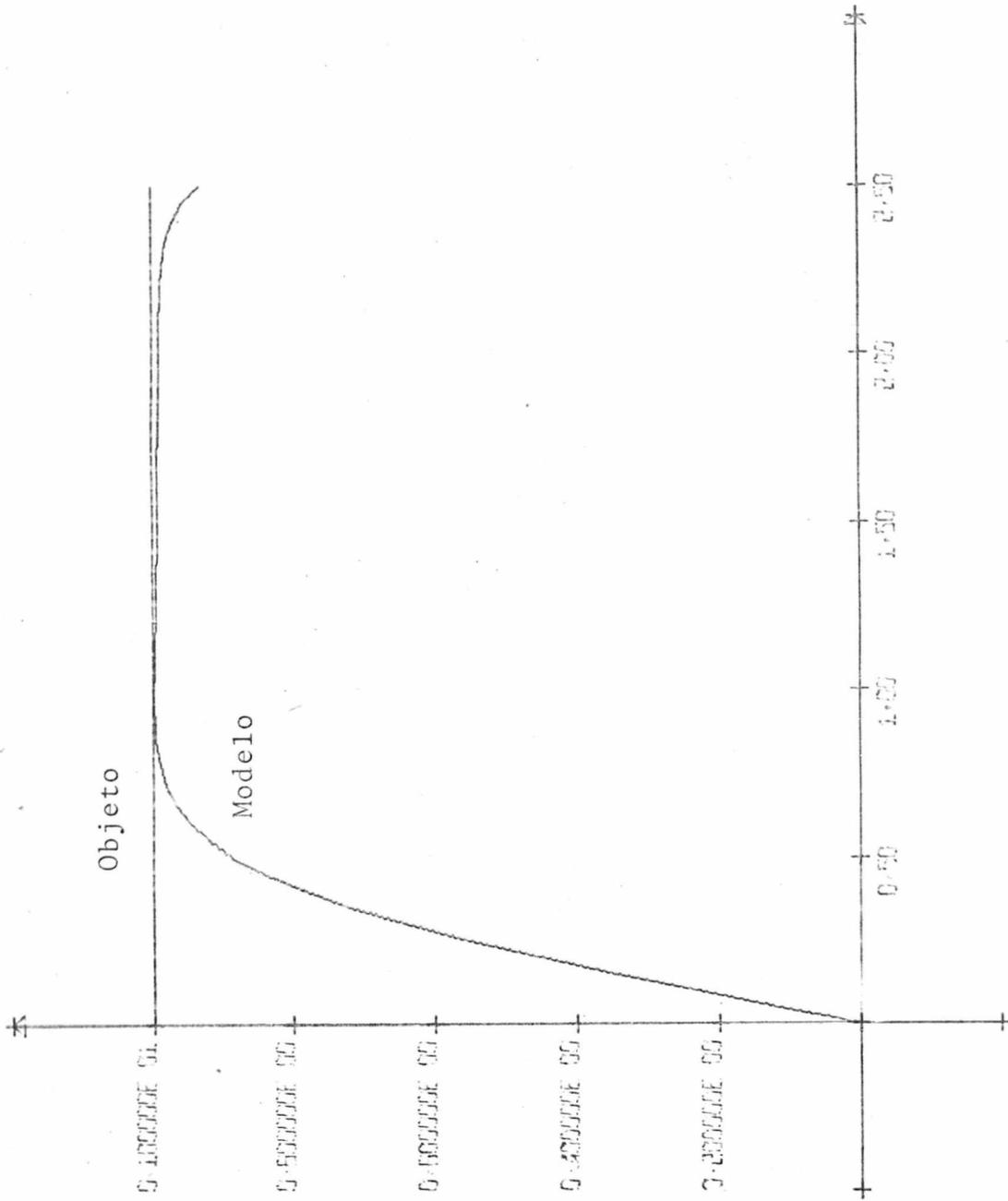
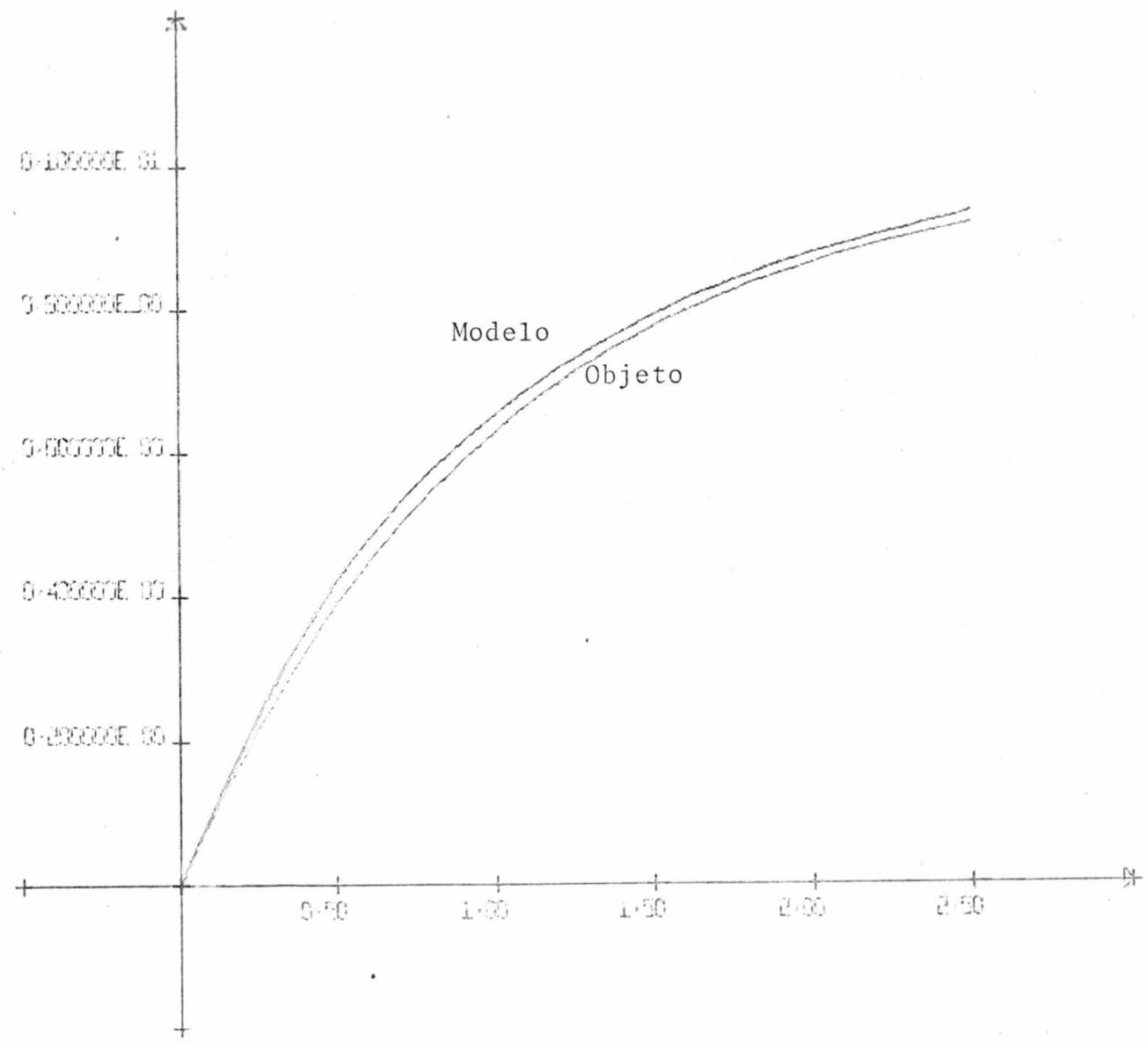


Figura 17

Figura 18



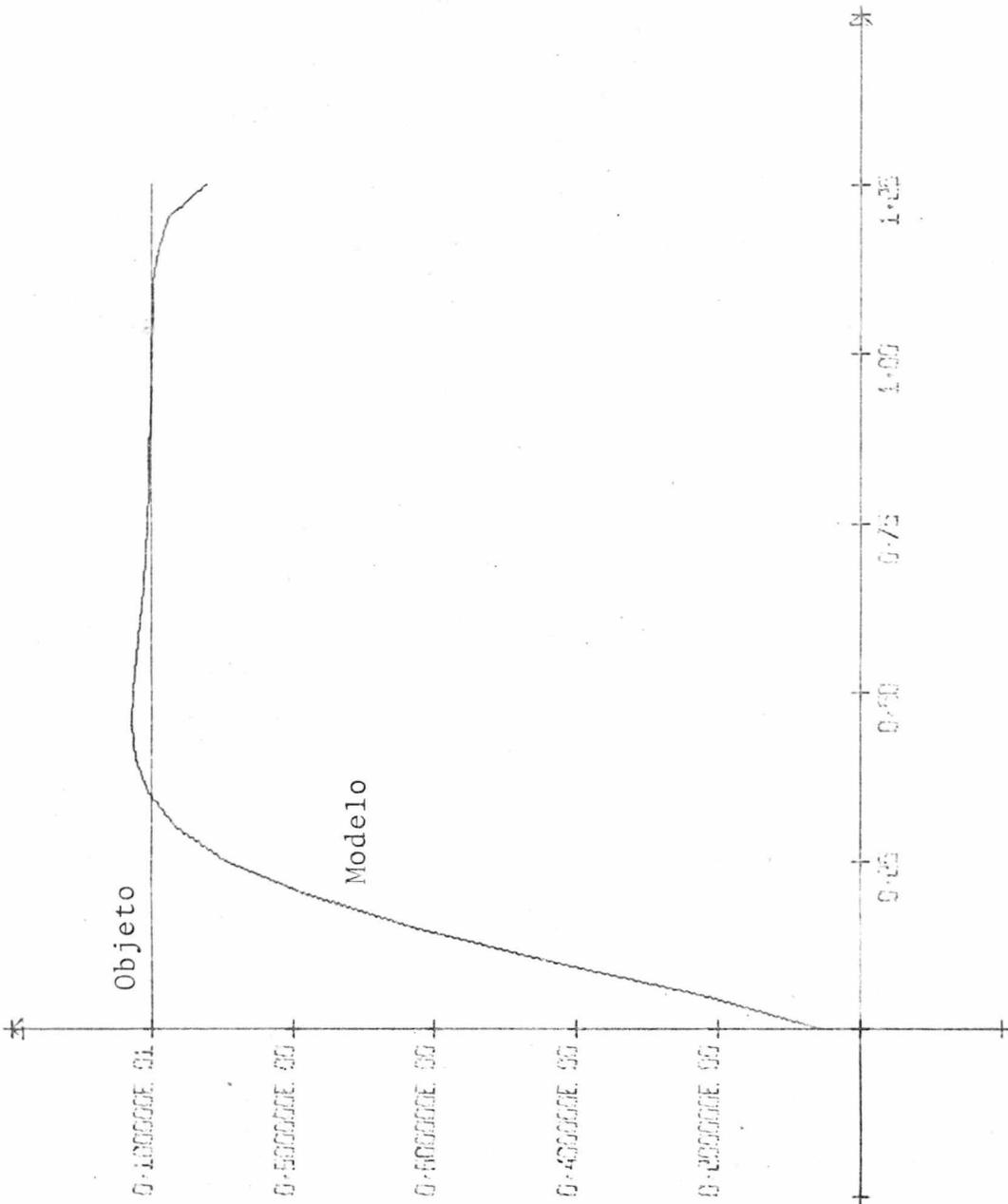


Figura 19

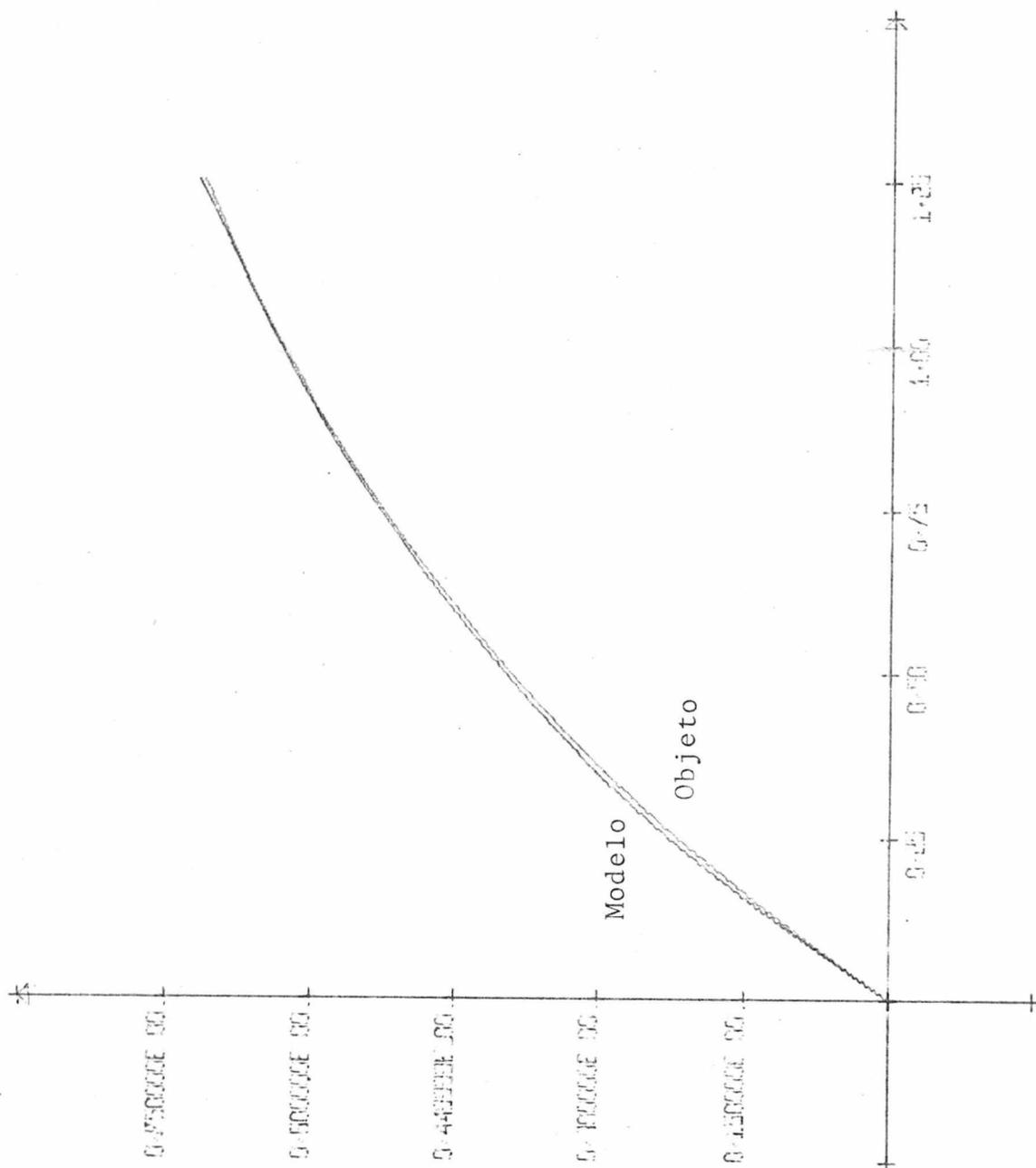


Figura 20

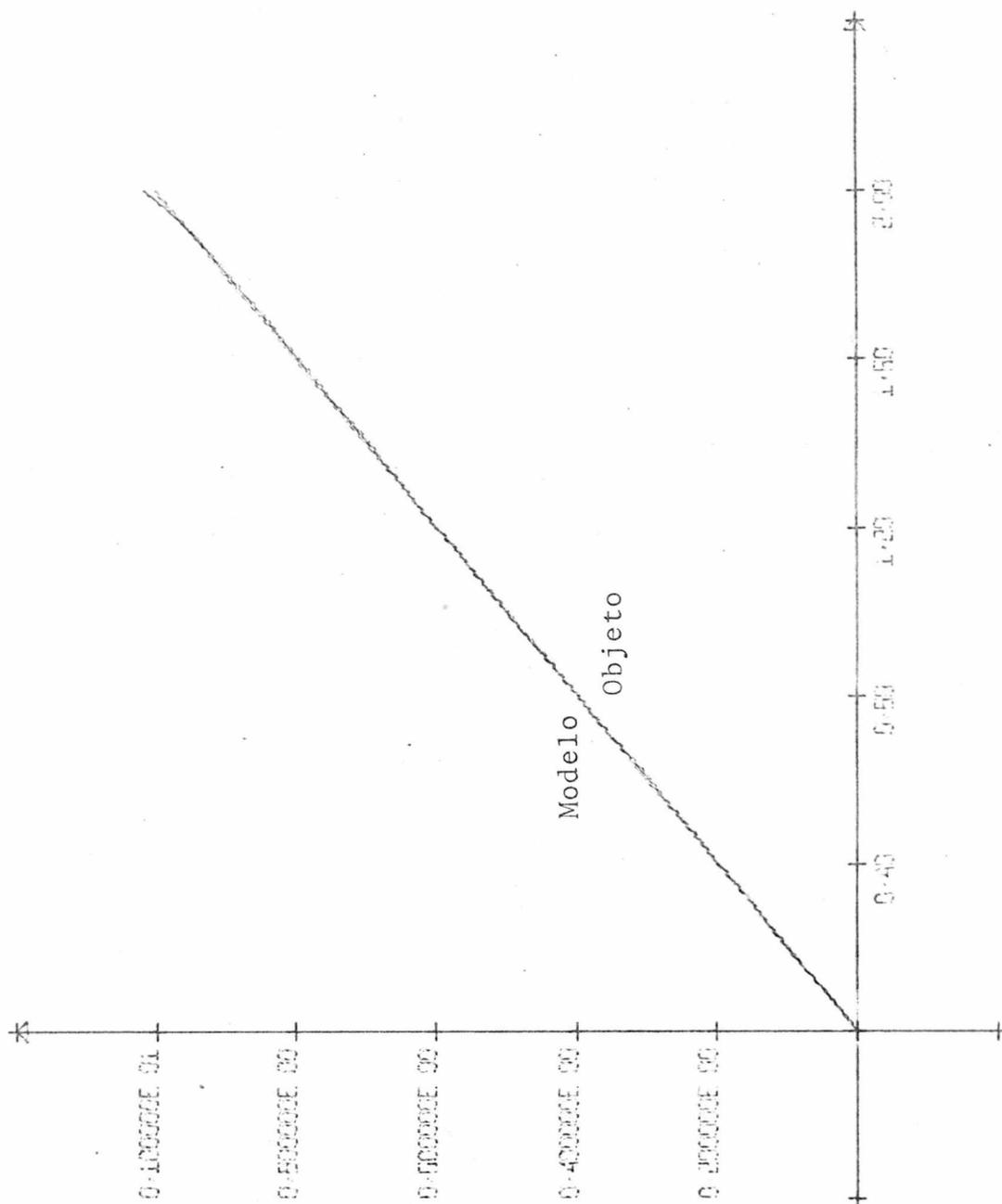
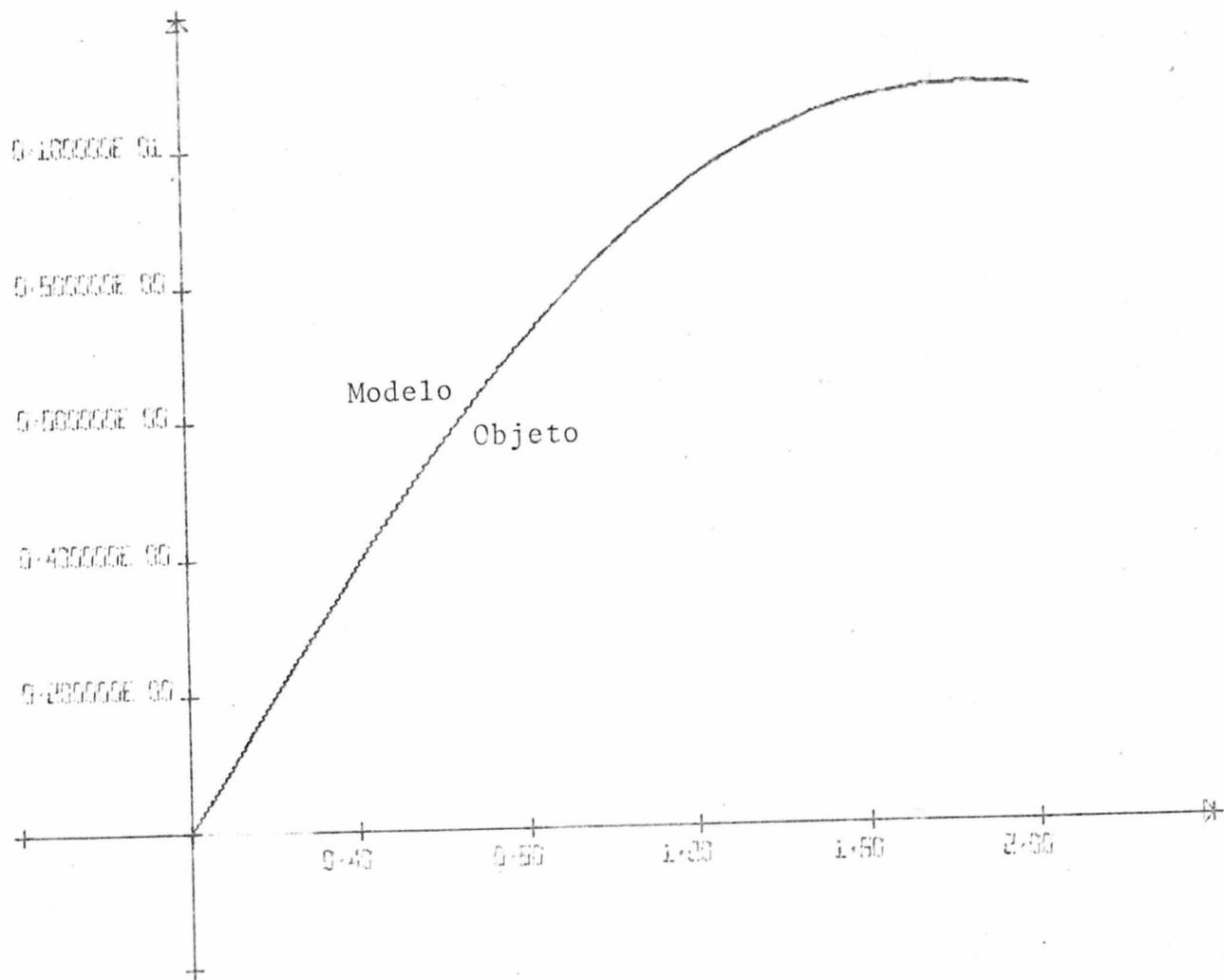


Figura 21

Figura 22



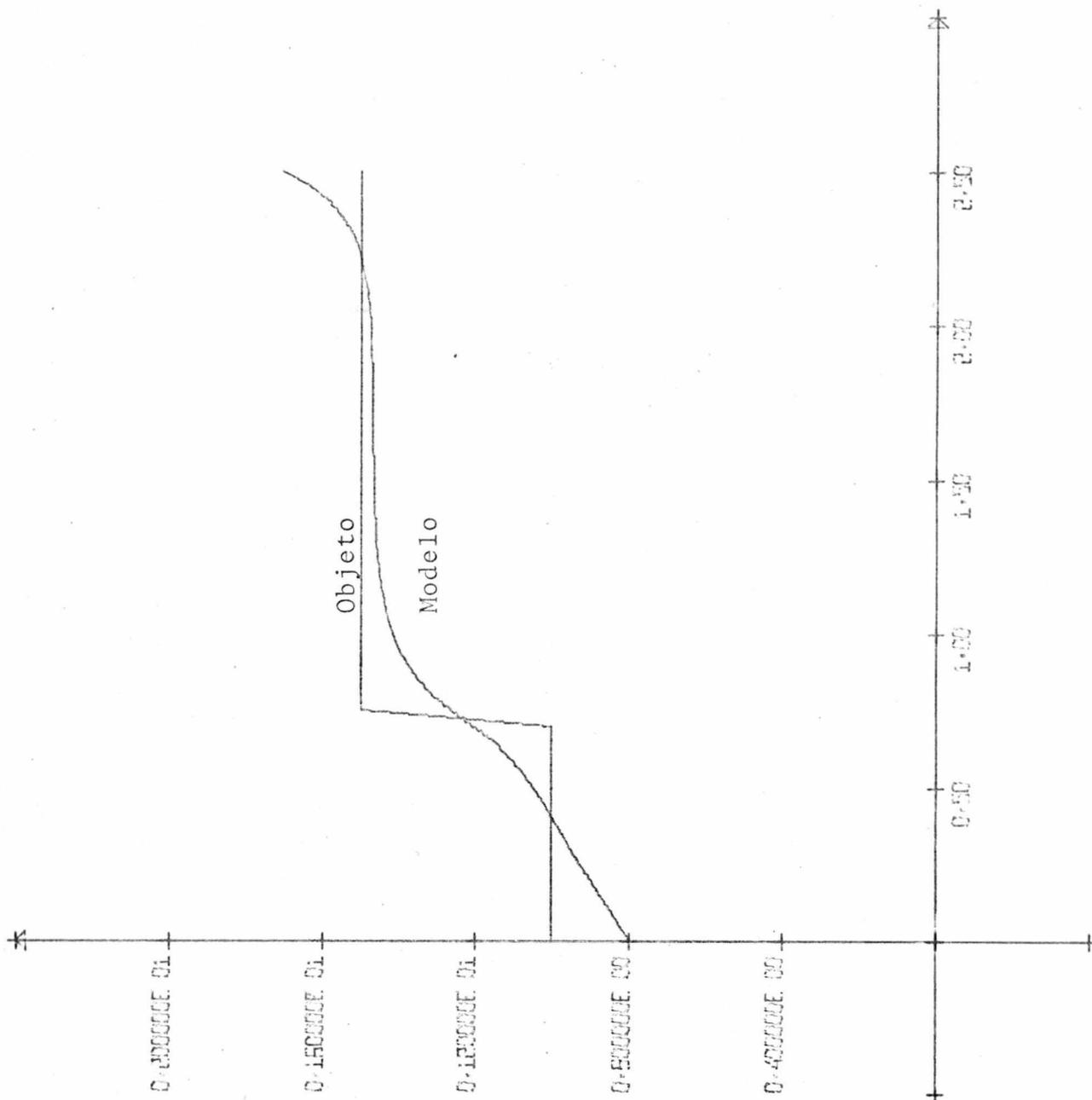


Figura 23

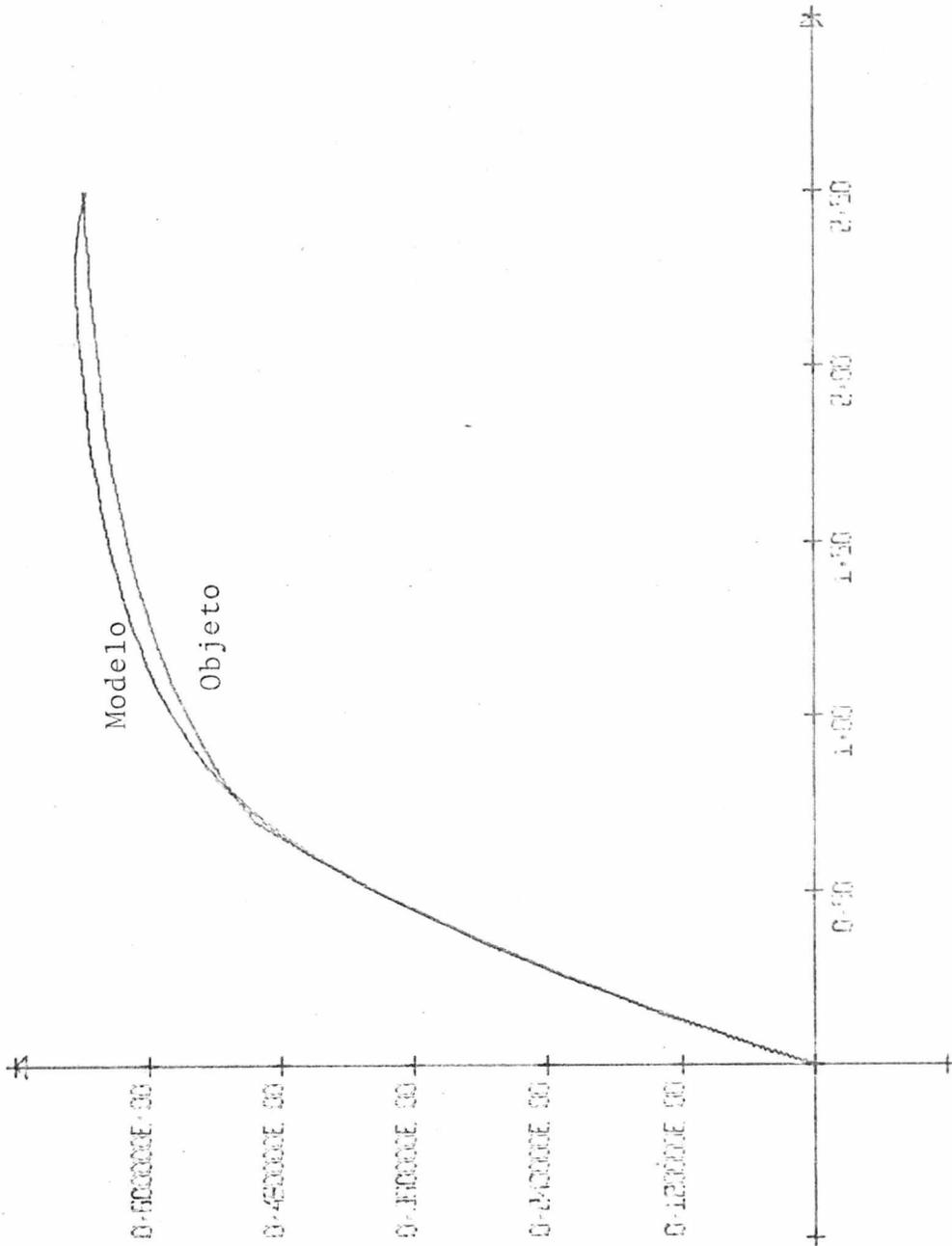


Figura 24

Figura 25

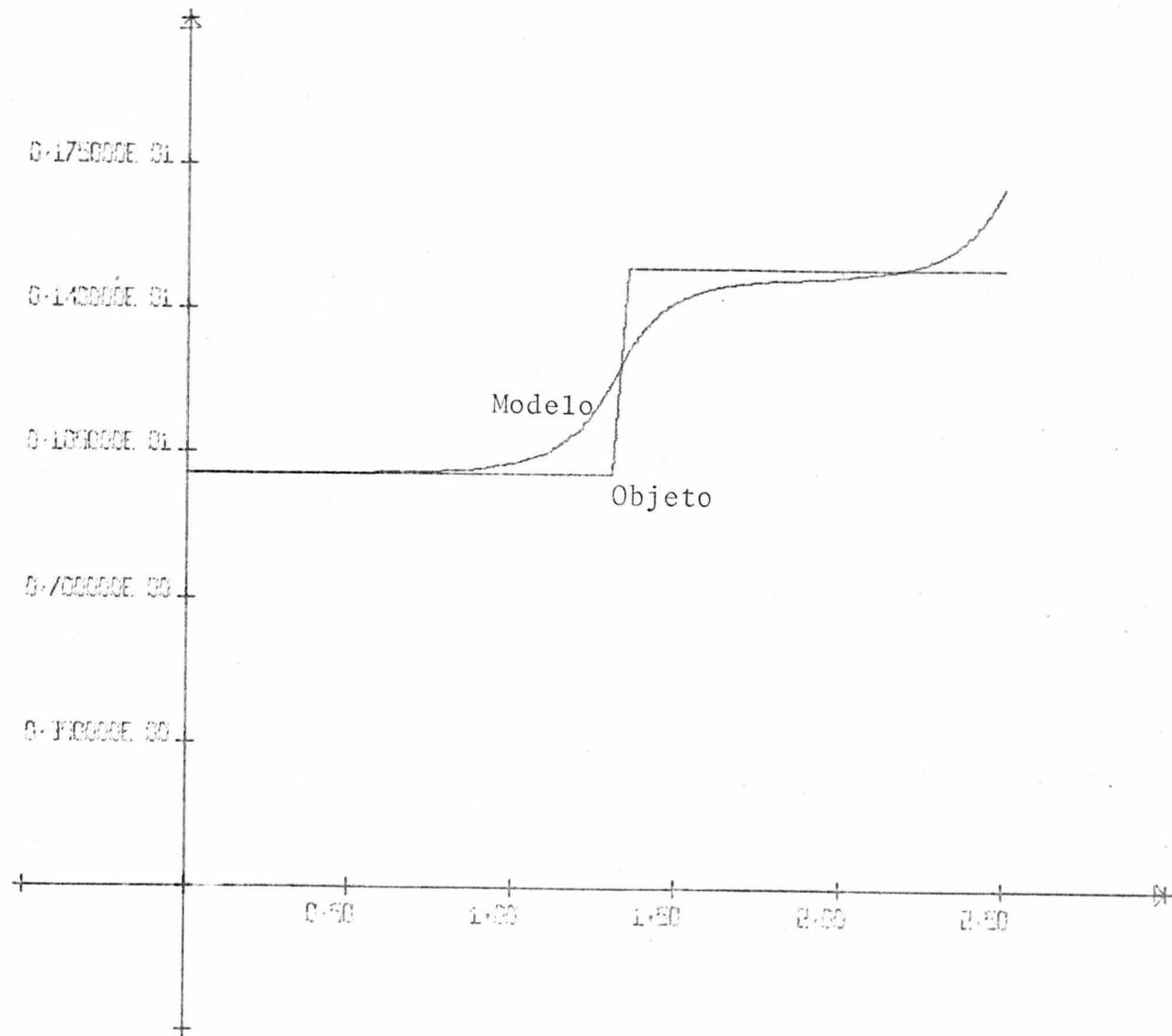
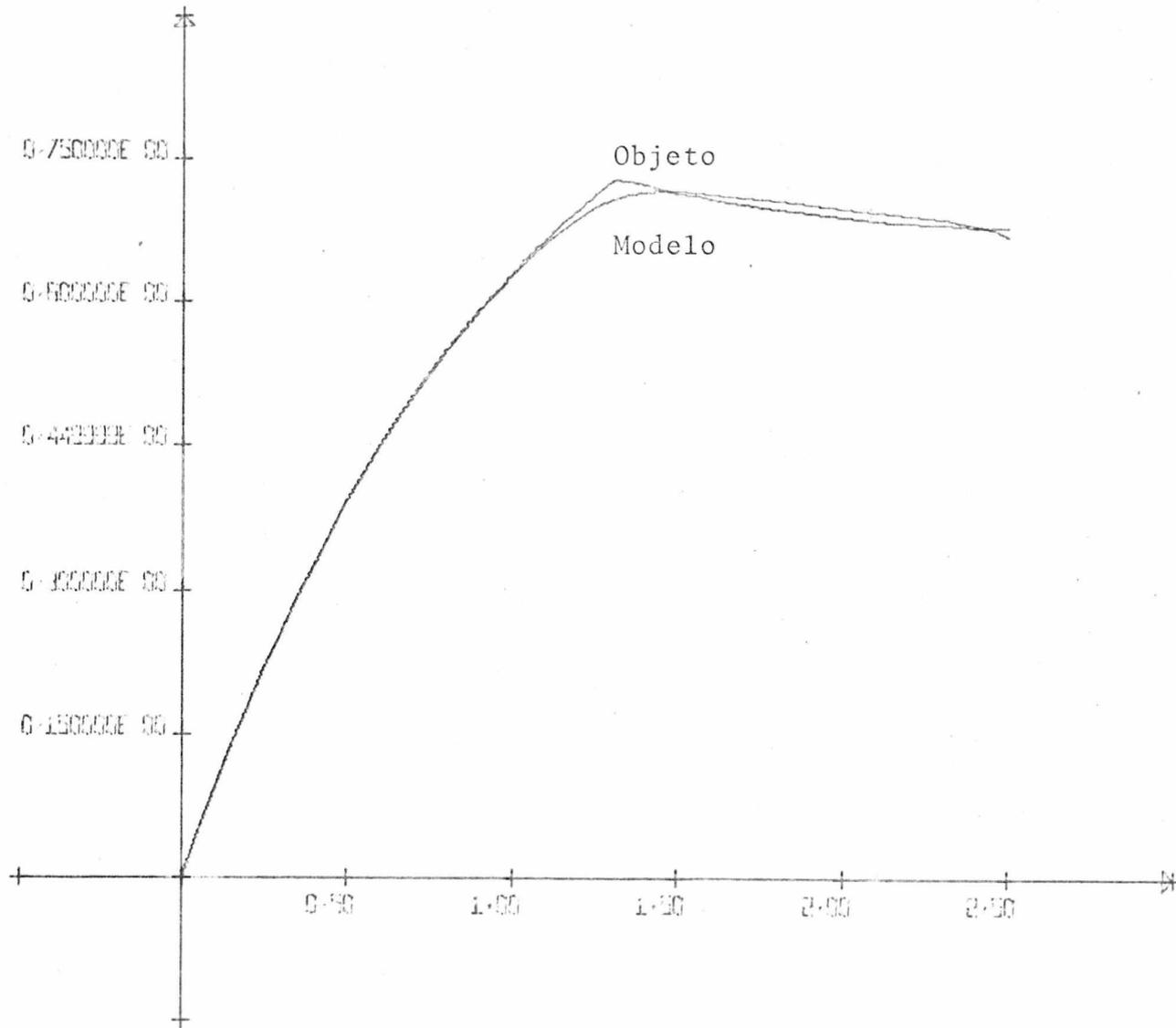


Figura 26



balhos referentes a identificação com parâmetros constantes. O problema já é bem complicado e não está completamente resolvido.

Poucos trabalhos existem a respeito de identificação de sistemas com parâmetros variáveis com o tempo. Em relação a parâmetros constantes, problemas de identificação com parâmetros variáveis são muito mais complicados de serem resolvidos. Assim é bem normal que este trabalho mostre bastante dificuldades. Apresentaremos abaixo cinco conclusões que julgamos importantes:

1. É muito difícil se fazer um modelo estrutural, notadamente em sistemas pouco conhecidos como por exemplo um sistema biológico.

2. É também difícil encontrar as condições iniciais exigidas pelo cálculo variacional quando aplicado a identificação de sistemas variáveis com o tempo.

3. Cálculo variacional é muito sensível em identificação. As condições iniciais devem ter um grande número de algarismos significativos, pois os sistemas de equações diferenciais são muito instáveis.

4. A pesquisa das condições iniciais é muito demorada, inclusive em problemas bem simples.

5. A entrada deve conter bastante informação, correndo-se o risco de minimizar o custo e se obter parâmetros bem diferentes dos do objeto. Em problemas de segunda ordem com entrada salto foi necessário considerar as duas componentes do estado como saída, para se obter um bom resultado. A função salto não contém muita informação.

6. Um grande interesse de identificação é a identificação "on line" de aplicação em sistemas adaptativos. Como cálculo variacional é muito demorado em identificação, não se pode usar esta técnica para identificação "on line".

Cálculo variacional em identificação é um problema complicadíssimo mas obteve-se resultado. Um trabalho interessante seria aplicar este método, ou programação dinâmica na determinação de entrada sensibilizante ou esferizante.

B I B L I O G R A F I A

1. NIEMAN, FISHER and SEBORG. A review of process identification and parameter estimation techniques. Inst. I. Control, 1971 , vol. 13, N° 2.
2. RICHALET, RAUT et POULIQUEN. Identification des processus par la méthode du modèle. Paris, Gordon and Breach, Théorie des Systèmes, Volume 4.
3. KIRK, DONALD E. Optimal Control Theory. New Jersey, Prentice Hall, 1970, Prnetice Hall Networks series.
4. DIXON, L.C.W. Nonlinear Optimisation. Glasgow, Bell and Bain Limited, 1972.
5. ELSGOLC, L.E. Calculus of Variations, Massachussetts, Addison - Wesley, 1962.
6. TSYPKIN Y.A. Adaptation Learning and self-learning in Control Systems, Survey Paper - I.F.A.C. Congress, Londres, 1966.
7. WILDE D. Optimum Seeking Methods, New Jersey, Prentice Hall , 1964.
8. FIACCO - Mc CORMICK. Nonlinear Programming Sequential Uncos - trained Minimization Techniques, N. York, J. Wiley & Sons, 1968.