

Universidade Estadual de Campinas  
Faculdade de Engenharia Elétrica  
Departamento de Telemática

## Identificação e Otimização Conjuntas: Uma Abordagem Multicritério

**Autor:** Tereza Cristina Duque/Borges  
**Orientador:** Prof. Dr. Paulo A. Valente/Ferreira

Tese apresentada à Faculdade de Engenharia Elétrica da Universidade Estadual de Campinas – FEE-UNICAMP, como parte dos requisitos exigidos para a obtenção do título de MESTRE EM ENGENHARIA ELÉTRICA.

MAIO – 1992.

Este exemplar corresponde à versão final da tese defendida por Tereza Cristina Duque Borges

Julgadora em 01 de Maio de 1992

Paulo Augusto Valente Ferreira  
Orientador

A Beto, Niu, Tinho e Mau.  
Para Pacha.

“O que eu faria se eu fosse eu?”

*Clarice Lispector*

## **Agradecimentos**

Meus sinceros agradecimentos a todos que, de alguma forma, direta ou indiretamente, contribuíram para a realização deste trabalho. Em especial a:

- Prof. Paulo Augusto Valente Ferreira, pela orientação segura e exemplo de dedicação à docência e à pesquisa;
- Ao Prof. Pedro Luís Dias Peres e ao Sérgio pelas dicas sempre oportunas;
- Aos amigos Patrícia, Adauto, Ernesto, Margarida, Paulo, Concha, Reginaldo, Carlos, Menotti e Nádia, dentre outros, pelo apoio em todos os momentos;
- Aos professores da FEE que ministraram os cursos que auxiliaram no desenvolvimento deste trabalho;
- Aos que, através da CAPES, financiaram este trabalho.

## Sumário

O problema conjunto de identificação de parâmetros e otimização de sistemas é enfocado no contexto das técnicas de otimização multicritério. A validade das propriedades de convergência da abordagem dois estágios do problema integrado de estimação de parâmetros e otimização escalar é estendida para o caso vetorial sob condições apropriadas. Alguns elementos de análise de decisão são incorporados no esquema de estimação de parâmetros e otimização conjunta para refletir situações de conflito derivadas da consideração simultânea de múltiplos critérios de desempenho. Para ilustrar os principais aspectos da abordagem proposta, os problemas conjuntos de otimização e estimação de parâmetros de reatores químicos por batelada e contínuo são abordados. O trabalho discute uma implementação multinível da abordagem proposta aplicada a um reator químico contínuo. Aspectos de descentralização e concorrência das tarefas de aquisição, identificação e otimização são também consideradas.

## Abstract

The integrated system optimization and parameter estimation problem is addressed in the context of vector-valued optimization techniques. The validity of the existing convergence properties of the basic two-step approach for solving the joint problem of scalar optimization and parameter estimation is extended for the vector-valued case under appropriate conditions. Some elements of decision analysis are incorporated into the integrated optimization and parameter estimation scheme in order to reflect conflicting situations derived from the simultaneous consideration of multiple performance criteria. To illustrate the main aspects of the proposed approach, the joint problems of optimization and parameter estimation of batch and continuous type chemical reactors are considered. The work discusses a multilevel implementation of the proposed approach when applied to a continuous chemical reactor. Aspects of decentralization and concurrency of the tasks of data acquisition, identification and optimization are also considered.

# Conteúdo

Introdução Geral	1
<b>1 Identificação e Otimização Conjuntas</b>	<b>5</b>
1.1 Introdução . . . . .	5
1.2 Formulação do Problema Conjunto . . . . .	5
1.3 Técnicas de Decomposição . . . . .	7
1.3.1 Método de Coordenação pelo Modelo - Método Factível	8
1.3.2 Método de Coordenação pelo Objetivo - Método Dual- Factível . . . . .	10
1.4 Abordagem Dois Estágios . . . . .	13
1.5 Abordagem Dois Estágios Modificada . . . . .	15
1.6 Abordagens Bi-Critério . . . . .	19
1.6.1 Abordagem $\epsilon$ -restrito . . . . .	20
1.6.2 Abordagem Paramétrica . . . . .	24
1.7 Abordagem MinMax . . . . .	27
1.8 Conclusão . . . . .	27
<b>2 Identificação e Otimização Multicritério: Aspectos Teóricos</b>	<b>29</b>
2.1 Introdução . . . . .	29
2.2 Formulação do Problema . . . . .	29
2.3 Algoritmo Dois-Estágios Generalizado . . . . .	31
2.4 Convergência e Análise de Sensibilidade . . . . .	34
2.5 Aspectos do Problema Multicritério ( $P_2^k$ ) . . . . .	36
2.6 Aspectos Computacionais . . . . .	39
2.7 Conclusão . . . . .	42

---

<b>3</b>	<b>Identificação e Otimização Conjuntas de um Reator “Batch”</b>	<b>43</b>
3.1	Introdução . . . . .	43
3.2	Descrição Qualitativa . . . . .	43
3.3	Descrição Quantitativa . . . . .	44
3.4	Sistema de Controle . . . . .	46
3.5	Modelo Algébrico Simplificado . . . . .	47
3.6	Formulação do Problema Conjunto . . . . .	49
3.7	Descrição do SRB . . . . .	51
3.7.1	Processo Físico . . . . .	51
3.7.2	Identificação de Parâmetros . . . . .	54
3.7.3	Otimização . . . . .	60
3.7.4	Interface com o Usuário . . . . .	61
3.8	Resultados Computacionais . . . . .	71
3.9	Conclusão . . . . .	77
<b>4</b>	<b>Identificação e Otimização Conjuntas de um Reator Contínuo:</b>	
	<b>Abordagem Multinível</b>	<b>79</b>
4.1	Introdução . . . . .	79
4.2	Implementação Multinível – Abordagem Multicritério . . . . .	79
4.3	Reator Químico Contínuo . . . . .	83
4.4	Formulação do Problema Conjunto . . . . .	85
4.5	Aspectos da Implementação . . . . .	88
4.5	Resultados Computacionais . . . . .	89
4.6	Conclusão . . . . .	95
	<b>Conclusão Geral</b>	<b>97</b>
	<b>Bibliografia</b>	<b>99</b>
	<b>Apêndice : Variáveis e Parâmetros dos Reatores</b>	<b>103</b>

# Introdução Geral

Em controle de processos, dois problemas são de particular interesse: *identificação* e *otimização* de sistemas. O problema de identificação consiste na determinação de um modelo para o sistema usualmente através de relações entrada/saída. Quando apenas alguns parâmetros das equações que compõem o modelo são desconhecidos ou quando deseja-se compensar incertezas do modelo com o uso de parâmetros livres tem-se caracterizado o problema de *identificação ou estimação de parâmetros*. Define-se o processo de estimação como a busca de valores de parâmetros que aproximem as saídas do modelo das saídas do processo real quando ambos são submetidos as mesmas entradas. O problema de otimização, por sua vez, trata da busca de valores para as variáveis manipuladas de forma a obter o melhor desempenho do sistema de acordo com critérios estabelecidos.

Modelagem inexata, ruído nas medidas e precisão limitada dos computadores frequentemente resultam na não-unicidade da solução do problema de estimação. Além disto, o modelo matemático adotado pode ser dependente do controle. Estas características indicam a necessidade de um tratamento integrado dos problemas de identificação e otimização de sistemas. O problema conjunto consiste, usualmente, na determinação da política de controle ótimo de um dado sistema baseado em estimativas numéricas de parâmetros incertos de seu modelo matemático que, por sua vez, são dependentes do controle (Borges e Ferreira, 1991).

O método mais comumente empregado consiste na aplicação de uma estratégia dois estágios, sendo os problemas de estimação de parâmetros e otimização resolvidos recursivamente. Estimativas iniciais das entradas são aplicadas ao sistema e medidas de estado estacionário são obtidas. A minimização de um critério de comparação entre as saídas do sistema e do modelo matemático fornece uma estimativa dos parâmetros para o problema de otimização que é, então, resolvido. As entradas otimizadas são usadas na

geração da próxima iteração e o método prossegue até que nenhuma melhoria seja obtida. Propriedades de convergência desta abordagem básica foram analisadas por Vatchsevanos (1979) e Ellis e Roberts (1982) entre outros. É possível mostrar que a solução do problema integrado é obtida se as saídas e as derivadas das saídas do sistema e do modelo com relação ao controle forem iguais no ponto limite gerado pelo procedimento.

Uma modificação da abordagem dois estágios básica, na qual um termo extra é introduzido no critério de otimização para compensar as diferenças entre as derivadas das saídas do sistema e do modelo com relação ao controle, foi proposta por Roberts (1979). Vários aspectos da solução do problema conjunto baseada na abordagem dois estágios modificada têm sido investigados (Brdys e Roberts, 1986; Tatjewski e Roberts, 1987).

Esquemas integrados baseados em técnicas multinível como a abordagem dois estágios modificada foram originalmente propostos por Haimes e Wismer (1972) e McGrew e Haimes (1974) motivados pela (possível) não-unicidade da solução do problema de estimação de parâmetros. De acordo com estes primeiros trabalhos, o problema integrado é visto como um problema bicritério (otimização e estimação de parâmetros) que é solucionado através da implementação multinível das abordagens  $\epsilon$ -restrito e paramétrica para solução de problemas de otimização multicritério (Hwang e Masud, 1979). Otimização multiobjetivo é uma estratégia de tratamento de diversos problemas de tomada de decisão com múltiplos critérios como planejamento e controle de produção, transporte, problemas de sequenciamento, que podem ou não apresentar parâmetros a serem estimados.

É importante notar que a escolha do critério de desempenho determinará os parâmetros obtidos na abordagem integrada. Consequentemente, um desempenho ótimo com respeito a um determinado critério pode representar um desempenho inaceitável em termos de uma análise multiobjetivo se os vários critérios não são propriamente expressos através do problema de otimização. Desta maneira, a suposição de uma estrutura de preferências a-priori envolvendo os critérios parece ser especialmente forte num contexto de estratégias dependentes do controle. Para superar tais dificuldades, a abordagem proposta neste trabalho permite que um operador humano intervenha no processo de otimização-estimação de parâmetros, a fim de indicar o melhor curso de ação a ser seguido.

Intervenção de operadores humanos para assegurar a estabilidade e produtividade é uma prática possível se os requerimentos de tempo não são

particularmente restritivos. Tal visão do problema está de acordo com os comentários gerais sobre o papel do operador no controle de processos e otimização, discutidos em Kümmel e Seborg (1987).

O procedimento visa o uso de informações de sucessivas iterações do algoritmo de otimização-estimação de parâmetros para melhorar o desempenho do sistema como um todo. Para realizar esta tarefa, uma extensão dos resultados prévios relacionados ao problema conjunto e o uso apropriado de elementos de otimização multicritério são desenvolvidos (Ferreira e Borges, 1992). Para ilustrar os aspectos principais, a factibilidade computacional e a utilidade da abordagem proposta com relação à estratégia dois estágios básica, o problema conjunto de estimação de parâmetros e otimização de um reator químico por batelada com dinâmica não linear é considerado.

A decomposição de problemas de otimização de grande porte em estruturas hierárquicas foi discutida por Brdys e Roberts (1986) e Haimes e Li (1988) entre outros. Também no caso do problema conjunto pode-se obter estruturas multinível compostas por subproblemas interconectados. Esta característica possibilita a utilização de processamento paralelo, facilitando a implementação de estratégias on-line para identificação de parâmetros e otimização, principalmente no caso de sistemas com dinâmica rápida.

Algoritmos que utilizam multiprocessamento na solução de problemas de otimização têm sido estudados por diversos autores como Bertsekas e Tsitsiklis (1989) e Franklin e Soong (1981). Neste trabalho, propõe-se uma implementação do método do gradiente modificado (Geromel e Ferreira, 1986) para problemas de programação não linear, associado à iteração de Jacobi (Bertsekas e Tsitsiklis, 1989), visando o aproveitamento das possibilidades de processamento paralelo. A fim de explorar os aspectos de paralelismo apresentados pela extensão proposta para a abordagem dois estágios modificada, o problema conjunto de identificação de parâmetros e otimização de um reator químico contínuo é estudado (Borges e Ferreira, 1992).

Esta dissertação é composta por quatro capítulos organizados de acordo com a descrição a seguir.

## **Capítulo 1**

São apresentadas as abordagens propostas por diversos autores para tratamento do problema integrado de identificação de parâmetros e otimização escalar, bem como as técnicas de decomposição utilizadas.

**Capítulo 2**

Uma extensão da abordagem dois estágios básica para o caso de otimização multicritério é proposta, sendo analisados aspectos de convergência e sensibilidade. Também neste capítulo, são apresentadas as propostas de paralelização do método do gradiente modificado.

**Capítulo 3**

O problema conjunto de identificação de parâmetros e otimização de um reator químico por batelada é resolvido com base nos resultados do Capítulo 2. O sistema computacional implementado é descrito com detalhes.

**Capítulo 4**

A abordagem dois estágios modificada é estendida para o caso multicritério. O problema integrado de identificação de parâmetros e otimização de um reator químico contínuo é tratado com base na estratégia proposta, sendo explorados os aspectos de paralelismo apresentados pela estrutura multinível obtida.

# Capítulo 1

## Identificação e Otimização Conjuntas

### 1.1 Introdução

O reconhecimento da importância do enfoque integrado dos problemas de identificação de parâmetros e otimização de sistemas tem levado diversos autores a formularem estratégias de tratamento do problema conjunto. Neste capítulo é apresentada uma visão geral destas abordagens e da teoria de decomposição utilizada pelas mesmas para manipulação do acoplamento entre os dois problemas e para o estabelecimento de estruturas multinível visando o processamento distribuído dos subproblemas.

### 1.2 Formulação do Problema Conjunto

A classificação do problema conjunto baseia-se nas características dos parâmetros do modelo e das variáveis de estado, controle e saída do sistema<sup>1</sup>. Parâmetros invariantes no tempo definem a identificação estática, enquanto que parâmetros que variam com o tempo caracterizam a identificação dinâmica. Otimização estática é o termo empregado quando as variáveis do sistema são quantidades escalares, enquanto que o termo otimização dinâmica é usado quando as variáveis são funções do tempo.

---

<sup>1</sup>Está sendo usada a seguinte equivalência entre os termos: estado = variáveis dependentes; controle = variáveis independentes ou manipuladas; saída = variáveis medidas.

Na versão estática, o modelo do sistema pode ser representado por

$$\varphi(x, u, \theta) = 0 \quad (1.1)$$

$$g(x, u, \theta) \leq 0 \quad (1.2)$$

$$y = h(x, u, \theta) \quad (1.3)$$

onde

$x$  : vetor n-dimensional de estado.

$u$  : vetor m-dimensional de controle.

$\theta$  : vetor p-dimensional de parâmetros do sistema.

$y$  : vetor r-dimensional de saída.

$$g : \mathcal{R}^n \times \mathcal{R}^m \times \mathcal{R}^p \rightarrow \mathcal{R}^t$$

$$h : \mathcal{R}^n \times \mathcal{R}^m \times \mathcal{R}^p \rightarrow \mathcal{R}^r$$

$$\varphi : \mathcal{R}^n \times \mathcal{R}^m \times \mathcal{R}^p \rightarrow \mathcal{R}^s$$

Os valores dos parâmetros são estimados com base na comparação da saída do modelo  $y$  com a saída efetiva do sistema  $\hat{y}$  em resposta a uma dada entrada  $u^j$ . Um índice normalmente usado é o de mínimos quadrados.

$$G(x, u, \theta) = \sum_{j=1}^N [h(x^j, u^j, \theta) - \hat{y}^j]' W^j [h(x^j, u^j, \theta) - \hat{y}^j] \quad (1.4)$$

onde

$W^j$  : matriz de ponderação  $r \times r$

$N$  : número de medidas

O problema de identificação estática pode, então, ser expresso por

$$\begin{aligned} \min_{\theta} \quad & G(x, u, \theta) \\ \text{s.a.} \quad & (1.1), (1.2), (1.3) \end{aligned} \quad (1.5)$$

O problema de otimização, supondo um índice de desempenho  $f : \mathcal{R}^n \times \mathcal{R}^m \times \mathcal{R}^p \rightarrow \mathcal{R}$ , pode ser definido por

$$\begin{aligned} \min_u \quad & f(x, u, \theta) \\ \text{s.a.} \quad & (1.1), (1.2), (1.3) \end{aligned} \quad (1.6)$$

No caso dinâmico, a evolução do sistema no intervalo  $t_0$  a  $t_f$  pode ser descrita por

$$\dot{x} = \varphi(x(t), u(t), \theta(t)), \quad x(t_0) = x_0 \quad (1.7)$$

$$g(x(t), u(t), \theta(t)) \leq 0 \quad (1.8)$$

$$y(t) = h(x(t), u(t), \theta(t)) \quad (1.9)$$

A identificação de parâmetros consiste em

$$\begin{aligned} \min G &= \frac{1}{2} \int_{t_0}^{t_f} \sum_{i=1}^N \|\hat{y}^i(t) - y(t)\|_{W^i}^2 dt \\ \text{s.a.} & (1.7), (1.8), (1.9) \end{aligned} \quad (1.10)$$

onde

$$\sum_{i=1}^N \|\hat{y}^i(t) - y(t)\|_{W^i}^2 = \sum_{i=1}^N [\hat{y}^i(t) - y(t)]' W^i [\hat{y}^i(t) - y(t)] \quad (1.11)$$

Por sua vez, o problema de otimização dinâmica será

$$\begin{aligned} \min_{u(t)} J &= \int_{t_0}^{t_f} f(x(t), u(t), \theta(t)) dt \\ \text{s.a.} & (1.7), (1.8), (1.9) \end{aligned} \quad (1.12)$$

Neste estudo, apenas o caso de maior interesse prático – identificação estática, otimização estática – é abordado. São discutidas diferentes estratégias de solução para este problema conjunto. Estas estratégias têm em comum o emprego de técnicas de decomposição que são brevemente discutidas a seguir.

### 1.3 Técnicas de Decomposição

No tratamento de sistemas de grande porte ou problemas integrados é muito conveniente decompor o sistema em subsistemas menores com seus próprios objetivos e restrições. Estes subsistemas podem ser interconectados de forma hierárquica, ou seja, um subsistema de um dado nível é controlado ou coordenado por um subsistema do nível superior e coordena subsistemas do nível inferior ao seu. Na otimização dos subsistemas deve-se considerar

as interações entre os mesmos de forma a obter-se um comportamento ótimo do sistema como um todo.

Sistemas multinível utilizam variáveis de coordenação que são manipuladas por um nível de coordenação, assegurando que as soluções dos subsistemas do nível inferior convergem para um ótimo global.

Duas abordagens de decomposição são enfocadas: Método de Coordenação pelo Modelo e Método de Coordenação pelo Objetivo. O material a seguir é baseado principalmente em Wismer (1971).

### 1.3.1 Método de Coordenação pelo Modelo - Método Factível

Considere o sistema estático

$$\varphi(u, y, x) = 0 \quad (1.13)$$

onde  $\varphi$  é composto por

$$\varphi^i(u^i, y^i, x^1, x^2) = 0 \quad i = 1, 2$$

A Figura 1.1 representa a estrutura do sistema.

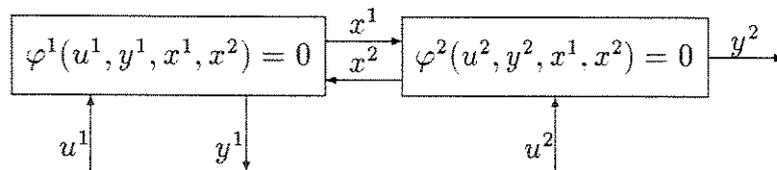


Figura 1.1: Estrutura do Sistema

Sendo,

$u^i$  : vetor de variáveis de controle do subsistema  $i$ ;

$y^i$  : vetor de variáveis de saída do subsistema  $i$ ;

$x^1$  : vetor de variáveis de interação do subsistema 1 para o 2;

$x^2$  : vetor de variáveis de interação do subsistema 2 para o 1.

O problema de otimização global é dado por

$$\begin{aligned} \min_{u,x,y} f(u, y, x) \\ \text{s.a. (1.13)} \end{aligned} \quad (1.14)$$

onde

$$f(u, y, x) = f^1(u^1, y^1, x^1) + f^2(u^2, y^2, x^2)$$

O sistema é composto por dois subsistemas que interagem entre si através das variáveis  $x^i$ . O método faz o desacoplamento dos dois sistemas, acrescentando uma restrição que fixa as variáveis de interação num certo valor  $z$ .

$$x = z \quad (1.15)$$

O problema é, então, decomposto em dois níveis. O nível inferior é responsável pela determinação de  $u, y$  que resolvem (1.14) para um certo valor  $z$  constante fornecido pelo segundo nível. Assim, tem-se

- Primeiro Nível

$$\begin{aligned} H(z) = \min_{u,y} f(u, y, z) \\ \varphi(x, y, z) = 0 \end{aligned}$$

- Segundo Nível

$$\min_z H(z)$$

O domínio para a minimização de  $f$  é

$$S_0 = \{(u, y) | \varphi(u, y, z) = 0\}$$

Entretanto,  $S_0$  pode ser vazio para alguns valores de  $z$ . Assim, é necessário definir um domínio para  $z$

$$S_1 = \{z | H(z) \text{ existe}\}$$

O primeiro nível pode, então, ser decomposto em subsistemas, pois a fixação das variáveis de acoplamento torna-os independentes.

Subsistema  $i$  ( $i = 1, 2$ ) do Primeiro Nível:

$$\begin{aligned} H^i(z) = \min_{u^i, y^i} f^i(u^i, y^i, z^i) \\ \varphi^i(u^i, y^i, z^1, z^2) = 0 \end{aligned}$$

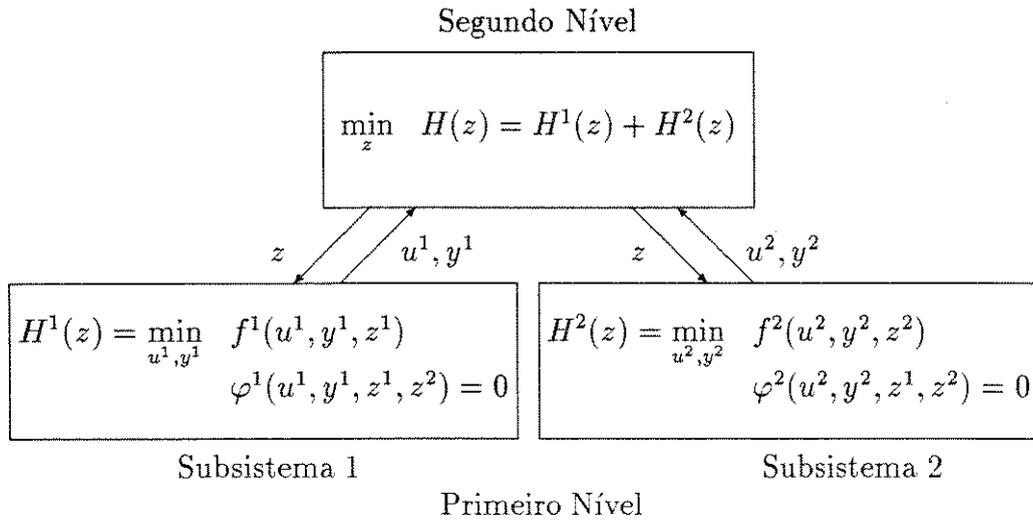


Figura 1.2: Estrutura Multinível - Coordenação pelo Modelo

O problema de segundo nível é escolher  $z$  que minimize  $H(z) = H^1(z) + H^2(z)$ . O diagrama da Figura 1.2 mostra a estrutura multinível resultante.

Neste método, as variáveis de coordenação são as variáveis  $x$  fixadas. O método é dito de coordenação pelo modelo porque acrescenta uma restrição ao modelo para possibilitar a decomposição. Outra característica do método é que o sistema pode operar com quaisquer dos valores intermediários de  $x$ ,  $u$ ,  $y$  obtidos, o que explica a denominação de método factível. Uma condição necessária para aplicação do método é que o conjunto  $S_0$  seja não vazio para vários valores de  $z$ .

### 1.3.2 Método de Coordenação pelo Objetivo - Método Dual-Factível

O método de coordenação pelo objetivo remove as interações cortando as ligações entre os subsistemas, como pode ser visto na Figura 1.3, considerando o sistema da Figura 1.1.

O corte é caracterizado pela substituição das variáveis  $x^i$  de entrada por variáveis  $z^i$ . Assim como  $u$ ,  $y$  e  $x$ , as variáveis  $z^i$  são manipuladas e arbitrárias, o que implica no desacoplamento total entre os subsistemas.

Este procedimento define os subproblemas independentes que compõem

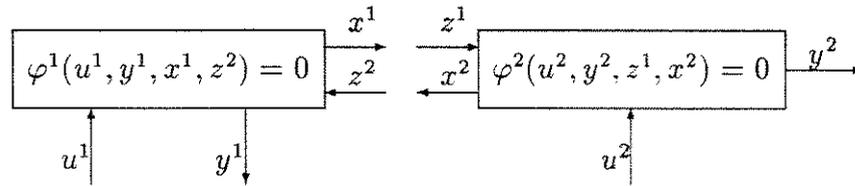


Figura 1.3: Estrutura do Sistema Desacoplado

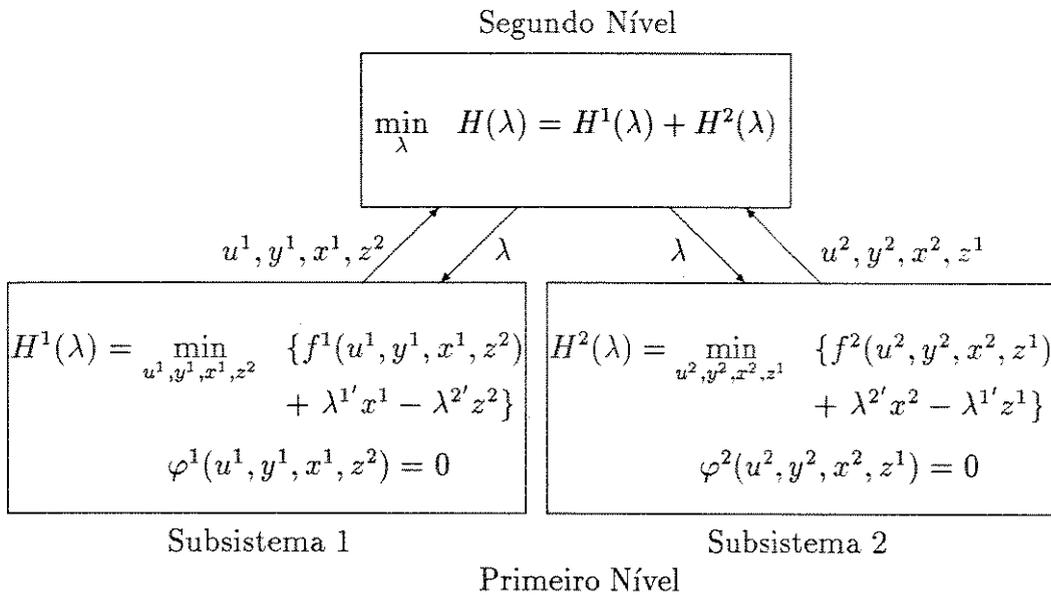


Figura 1.4: Estrutura Hierárquica - Coordenação pelo Objetivo

o primeiro nível e estabelece a função do segundo nível, que é coordenar os subsistemas do primeiro nível de forma que o princípio do balanço das interações seja obedecido. Este balanço consiste na busca de igualdade entre as variáveis  $x$  e  $z$ .

Este comportamento do sistema multinível é obtido através da alteração da função objetivo, que recebe um termo adicional representando o balanço das interações. Este termo pode ser visto como uma penalização ou como uma restrição de igualdade com um multiplicador de Lagrange associado. Assim, a função objetivo passa a ser

$$f(u, y, x, z, \lambda) = f^1(u^1, y^1, x^1) + f^2(u^2, y^2, x^2) + \lambda'(x - z) \quad (1.16)$$

onde

$$\lambda'(x - z) = \lambda^1(x^1 - z^1) + \lambda^2(x^2 - z^2)$$

O problema integrado é, então, decomposto e a estrutura hierárquica obtida encontra-se na Figura 1.4.

Os domínios para as minimizações são

$$S_0 = \{(u, y, x, z) | \varphi_1 = \varphi_2 = 0\}$$

$$S_1 = \{\lambda | H(\lambda) \text{ existe}\}$$

O termo coordenação pelo objetivo tem origem na manipulação das funções objetivo dos subproblemas de primeiro nível pelo segundo nível, através da escolha do vetor  $\lambda$ , que neste método representa as variáveis de coordenação.

Ambos os métodos estabelecem procedimentos iterativos na busca da solução ótima para o sistema como um todo. De forma genérica, é escolhido o conjunto de variáveis de coordenação (variáveis reais, multiplicadores de Lagrange) que são fixadas, produzindo os problemas de primeiro nível. O nível 2 é responsável pela determinação de valores ótimos para as variáveis de coordenação. Os níveis interagem entre si trocando os dados necessários a cada iteração.

Deve-se, entretanto, salientar que a convergência de procedimentos como os descritos nesta seção depende da existência de pontos de sela das funções lagrangeanas, o que é assegurado apenas no caso de problemas convexos (Wismer, 1971).

Nos próximos itens são detalhadas algumas das abordagens do problema conjunto que utilizam estas técnicas de decomposição.

## 1.4 Abordagem Dois Estágios

Também conhecida como sequencial, esta abordagem trata os problemas em separado e de forma iterativa. O procedimento é composto pelos seguintes passos

1. O problema de identificação é resolvido para valores estimados das variáveis de controle  $u^0$  e variáveis de estado  $x^0$ , ou seja

$$\begin{aligned} \min_{\theta} \quad & G(x^0, u^0, \theta) \\ & \varphi(x^0, u^0, \theta) = 0 \\ & g(x^0, u^0, \theta) \leq 0 \end{aligned} \quad (1.17)$$

2. A solução do passo anterior  $\theta^0$  é, então, utilizada no problema de otimização para determinação de novos valores para  $x$  e  $u$ . Assim

$$\begin{aligned} \min_{u,x} \quad & f(x, u, \theta^0) \\ & \varphi(x, u, \theta^0) = 0 \\ & g(x, u, \theta^0) \leq 0 \end{aligned} \quad (1.18)$$

Este procedimento é repetido iterativamente até que nenhuma melhoria no desempenho do sistema seja obtida. Um dos problemas desta abordagem é que existem algumas condições para que a convergência para o ponto de operação ótimo ocorra. Estas condições são ilustradas a seguir, a partir de algumas hipóteses simplificadoras (Ellis e Roberts, 1982).

Considere que o problema de otimização é irrestrito e que os problemas de otimização e identificação são convexos para todos os possíveis controles  $u$ . Eliminando-se as variáveis dependentes, o problema de otimização considerando as saídas reais do processo  $\hat{y}$  pode ser descrito por

$$\min_u f(u, \hat{y}) \quad (1.19)$$

Supondo que o processo não está sujeito a ruídos, o valor da saída é função apenas do controle

$$\hat{y} = p(u) \quad (1.20)$$

Sendo a função do processo  $p$  desconhecida, é construído um modelo com parâmetros  $\theta$  livres

$$y = m(u, \theta) \quad (1.21)$$

Assim, o problema de otimização baseado no modelo é expresso por

$$\begin{aligned} \min_u \quad & f(u, y) \\ \text{s.a.} \quad & (1.21) \end{aligned} \quad (1.22)$$

Considere, ainda, que as funções objetivo em (1.19) e (1.22) são diferenciáveis e que existe apenas um ponto de mínimo global. Para que  $u^0$  seja a solução do problema de otimização é necessário e suficiente que

$$\left. \frac{df(u, \hat{y})}{du} \right|_{u=u^0} = \left. \frac{df(u, p(u))}{du} \right|_{u=u^0} = 0 \quad (1.23)$$

De forma análoga, para que  $u^*$  solucione o problema de otimização baseada no modelo, deve satisfazer

$$\left. \frac{df(u, y)}{du} \right|_{u=u^*} = \left. \frac{df(u, m(u, \theta))}{du} \right|_{u=u^*} = 0 \quad (1.24)$$

Para que os problemas expressos por (1.19) e (1.22) sejam equivalentes, isto é,  $u^0 = u^*$  deve-se determinar  $\theta$  tal que

$$\left. \frac{df(u, p(u))}{du} \right|_{u=u^*} = \left. \frac{df(u, m(u, \theta))}{du} \right|_{u=u^*} = 0 \quad (1.25)$$

Considerando,

$$\left. \frac{df(u, p(u))}{du} \right|_{u=u^*} = \left. \frac{\partial f}{\partial u} + \frac{\partial f}{\partial p} \frac{\partial p}{\partial u} \right|_{u=u^*} \quad (1.26)$$

$$\left. \frac{df(u, m(u, \theta))}{du} \right|_{u=u^*} = \left. \frac{\partial f}{\partial u} + \frac{\partial f}{\partial m} \frac{\partial m}{\partial u} \right|_{u=u^*} \quad (1.27)$$

pode-se escrever,

$$\left. \frac{\partial f}{\partial u} + \frac{\partial f}{\partial p} \frac{\partial p}{\partial u} \right|_{u=u^*} = \left. \frac{\partial f}{\partial u} + \frac{\partial f}{\partial m} \frac{\partial m}{\partial u} \right|_{u=u^*} \quad (1.28)$$

Assim, a solução baseada no modelo é correta se as condições expressas em (1.29) forem satisfeitas

$$\begin{aligned} m(u, \theta)|_{u=u^*} &= p(u)|_{u=u^*} \\ \left. \frac{dm}{du}(u, \theta) \right|_{u=u^*} &= \left. \frac{dp}{du}(u) \right|_{u=u^*} \end{aligned} \quad (1.29)$$

Portanto, o procedimento sequencial converge para o ponto de operação ótimo real se as saídas e as derivadas das saídas com relação ao controle do modelo e da planta se igualam no ponto ótimo.

É interessante observar que estas condições são idênticas ‘as utilizadas por Abastanotti et. al. (1989) no estudo de linearização de sistemas por realimentação de saída.

## 1.5 Abordagem Dois Estágios Modificada

Trata-se de uma extensão da abordagem dois estágios enfocada no item anterior. As alterações inseridas têm por objetivo assegurar a convergência para o ponto de operação ótimo. Esta estratégia permite também que se estabeleça um compromisso entre medidas das saídas da planta e uma modelagem mais precisa, optando-se por um ou outro procedimento a depender do custo da instrumentação e do custo da modelagem mais rígida. Assim, o modelo total do sistema pode ser composto por relações aproximadas para as variáveis mensuráveis e relações precisas para as saídas, cuja medição representaria um custo muito elevado ou cuja modelagem é muito simples.

Esta abordagem utiliza a decomposição por modificação do objetivo de otimização, construindo uma estrutura hierárquica com os problemas de identificação e otimização sendo desacoplados e resolvidos separadamente, porém sob controle de uma unidade de coordenação.

Considere os problemas definidos em (1.5) e (1.6). O desacoplamento entre eles é obtido pelo acréscimo das seguintes restrições de igualdade

$$v = u \quad (1.30)$$

$$\sigma = \theta \quad (1.31)$$

Substituindo-se (1.31) em (1.6), o problema de otimização passa a ser

$$\begin{aligned} \min_u \quad & f(x, u, \sigma) \\ & \varphi(x, u, \sigma) = 0 \\ & g(x, u, \sigma) \leq 0 \\ & y = h(x, u, \sigma) \end{aligned} \quad (1.32)$$

Utilizando-se as restrições de igualdade, elimina-se  $x$  e a otimização reduz-se a

$$\begin{aligned} \min_u \quad & F(u, \sigma) \\ & \psi(u, \sigma) \leq 0 \end{aligned} \quad (1.33)$$

Analogamente, substituindo-se (1.30) em (1.5), o problema de identificação de parâmetros pode ser expresso por

$$\begin{aligned} \min_{\theta} \quad & G(v, \theta) \\ & \varphi(x, v, \theta) = 0 \\ & g(x, v, \theta) \leq 0 \\ & y = h(x, u, \theta) \end{aligned} \quad (1.34)$$

Supondo-se que o vetor de parâmetros ótimo  $\theta$  encontra-se no interior da região definida pelas restrições, a condição de primeira ordem consiste em

$$\frac{\partial G(v, \theta)}{\partial \theta} = 0 \quad (1.35)$$

De acordo com (1.30), (1.31), (1.33) e (1.35) pode-se escrever a função lagrangeana do problema integrado de identificação de parâmetros e otimização.

$$L = F(u, \sigma) + \lambda'(v - u) + \mu'(\sigma - \theta) + \eta' \frac{\partial G}{\partial \theta}(v, \theta) + \xi' \psi(u, \sigma) \quad (1.36)$$

onde  $\lambda$ ,  $\mu$ ,  $\eta$ ,  $\xi$  são multiplicadores de Lagrange.

Considere que as derivadas seguintes existem e são contínuas. As condições de ponto estacionário do lagrangeano são

$$\frac{\partial L}{\partial u} = \frac{\partial F}{\partial u} - \lambda + \left[ \frac{\partial \psi}{\partial u} \right]' \xi = 0 \quad (1.37)$$

$$\frac{\partial L}{\partial \theta} = [G_{\theta\theta}] \eta - \mu = 0 \quad (1.38)$$

$$\frac{\partial L}{\partial v} = [G_{v\theta}] \eta + \lambda = 0 \quad (1.39)$$

$$\frac{\partial L}{\partial \sigma} = \frac{\partial F}{\partial \sigma} + \mu + \left[ \frac{\partial \psi}{\partial \sigma} \right]' \xi = 0 \quad (1.40)$$

onde

$$[G_{\theta\theta}]_{ij} = \frac{\partial^2 G}{\partial \theta_i \partial \theta_j} \quad (1.41)$$

$$[G_{v\theta}]_{ij} = \frac{\partial^2 G}{\partial v_i \partial \theta_j} \quad (1.42)$$

O problema de otimização modificado é obtido de (1.37)

$$\begin{aligned} \min_u \quad & F(u, \sigma) - \lambda'u \\ & \psi(u, \sigma) \leq 0 \end{aligned} \quad (1.43)$$

Reconsiderando-se as restrições de igualdade, obtém-se

$$\begin{aligned} \min_u \quad & f(x, u, \sigma) - \lambda'u \\ & \varphi(x, u, \sigma) = 0 \\ & g(x, u, \sigma) \leq 0 \end{aligned} \quad (1.44)$$

Assim em (1.44) tem-se o problema de otimização com a função objetivo modificada, o qual constitui um dos subproblemas da estrutura hierárquica a ser resolvido para valores de  $\sigma$  e  $\lambda$  fornecidos pela unidade de coordenação.

Uma expressão para o multiplicador  $\lambda$  pode ser obtida das expressões (1.37) – (1.40)

$$\lambda = [G_{v\theta}] [G_{\theta\theta}]^{-1} \left[ \frac{\partial F}{\partial \sigma} + \left[ \frac{\partial \psi}{\partial \sigma} \right]' \xi \right] \quad (1.45)$$

O problema de identificação genérico definido em (1.34) permanece inalterado e é, também, um subproblema a ser resolvido para valores de  $v$  fornecidos pela unidade de coordenação.

As variáveis de saída da planta  $\hat{y}$  podem ser divididas em variáveis mensuráveis  $\hat{y}^1$  e não mensuráveis  $\hat{y}^2$ , ou seja

$$\hat{y} = \begin{bmatrix} \hat{y}^1 \\ \hat{y}^2 \end{bmatrix} \quad (1.46)$$

Caso o número de parâmetros a serem estimados seja igual ao número de componentes de  $\hat{y}^1$ , o problema de identificação reduz-se a determinar  $\theta$ , tal que as saídas da planta  $\hat{y}^1$  e as saídas do modelo  $y^1$  sejam iguais quando excitados pela mesma entrada  $v$ , o que pode ser expresso por

$$y^1(v, \theta) = \hat{y}^1(v) \Rightarrow \theta \quad (1.47)$$

A condição de estacionariedade  $\frac{\partial G}{\partial \theta}(v, \theta) = 0$  é substituída pela condição de igualdade (1.47) na função lagrangeana (1.36) e, então, obtém-se

$$L = F(u, \sigma) + \lambda'(v - u) + \mu'(\sigma - \theta) + \eta' [y^1(v, \theta) - \hat{y}^1(v)] + \xi' \psi(u, \sigma) \quad (1.48)$$

Usando-se (1.47) e (1.45), determina-se a expressão do modificador  $\lambda$

$$\lambda = \left[ \left[ \frac{\partial y^1}{\partial v} \right]' - \left[ \frac{\partial \hat{y}^1}{\partial v} \right]' \right] \left[ \left[ \frac{\partial y^1}{\partial \theta} \right]' \right]^{-1} \left[ \frac{\partial F}{\partial \sigma} + \left[ \frac{\partial \psi}{\partial \sigma} \right]' \xi \right] \quad (1.49)$$

Na Figura 1.5, tem-se a representação da estrutura hierárquica obtida.

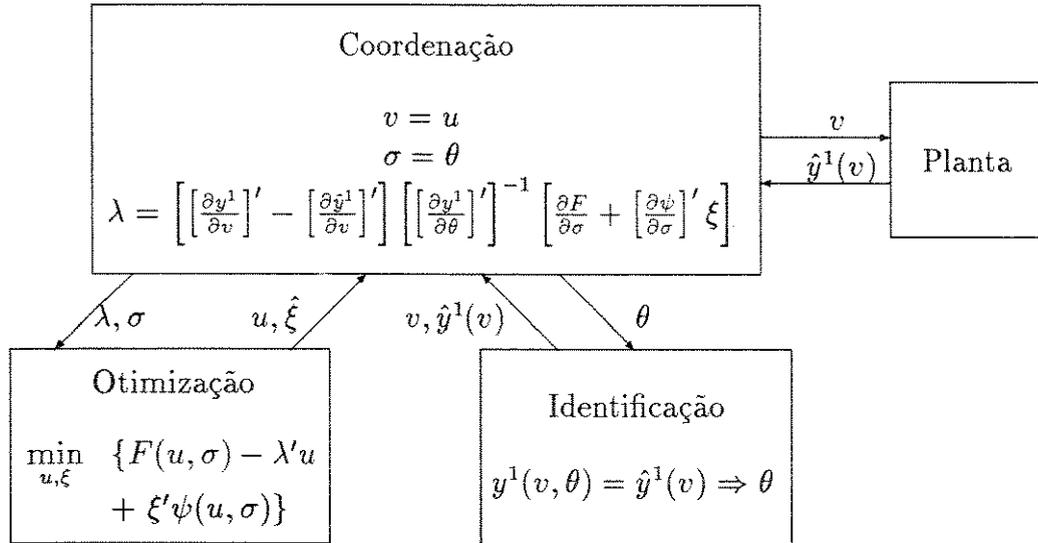


Figura 1.5: Estrutura Hierárquica - Abordagem 2-estágios Modificada

A unidade de coordenação tem por funções transferir os valores das variáveis de controle e parâmetros do modelo entre as unidades de otimização e identificação, fornecer entradas de controle para a planta, receber as saídas resultantes, as quais também são fornecidas à identificação, e calcular os modificadores  $\lambda$ .

Tendo recebido os parâmetros  $\sigma$  e modificadores  $\lambda$  estimados, a unidade de otimização determina as variáveis de controle  $u$  através do problema modificado. Por sua vez, a unidade de identificação estima os parâmetros  $\theta$  de acordo com os valores das variáveis de saída mensuráveis  $y^1(v)$  e variáveis de controle  $v$  recebidas.

O termo extra incluído na função objetivo do problema de otimização assegura que o ótimo correto será alcançado mesmo que os gradientes das variáveis de saída mensuráveis da planta e do modelo não se igualem. Isto ocorre porque os modificadores  $\lambda$  dependem da diferença entre os gradientes, penalizando o desempenho no caso de desigualdade entre eles.

Ainda com relação à convergência do procedimento iterativo, para assegurar a estabilidade do processo não se deve aplicar diretamente o controle estimado pela otimização. Assim, é aconselhável implementar a relação  $v = u$  pela seguinte equação a diferenças

$$v^i = v^{i-1} + K\{u^i - v^{i-1}\} \quad (1.50)$$

Onde  $i$  indica  $i$ -ésima iteração e  $K$  é uma matriz de ganho diagonal com elementos  $0 < K_{ii} < 1$ . Assim, previne-se alterações excessivas nas condições de operação do sistema.

Uma desvantagem desta estratégia é a necessidade de determinação de derivadas de primeira ordem no cálculo de  $\lambda$ . As derivadas das saídas do modelo não apresentam problemas, pois pode-se deduzir expressões ou perturbar as entradas do modelo e usar técnicas de diferenças finitas. Entretanto, no caso das derivadas das saídas da planta, a única opção é o uso de técnicas de diferenças finitas, o que implica em perturbar as entradas do processo e fazer medições que podem estar contaminadas por ruídos.

As experiências feitas por (Roberts, 1979; Ellis e Roberts, 1982) demonstram, entretanto, que mesmo em presença de ruído, o desempenho da estratégia dois estágios modificada é superior à estratégia dois estágios, conduzindo sempre ao ponto ótimo correto.

## 1.6 Abordagens Bi-Critério

As estratégias de tratamento do problema de identificação e otimização conjuntas tratadas nos itens anteriores são convenientes do ponto de vista de aplicação prática on-line e, além disto, consideram a dependência dos parâmetros em relação ao controle.

Por outro lado, existem abordagens que baseiam-se na caracterização do problema conjunto como um problema bi-critério, isto é

$$\begin{aligned} \min_{u, \theta} \quad & \begin{bmatrix} G(\theta) \\ f(u, \theta) \end{bmatrix} \\ & g(u, \theta) \leq 0 \\ & y = h(u, \theta) \end{aligned} \quad (1.51)$$

onde, por simplicidade, variáveis dependentes são omitidas nos modelos. Sob este enfoque, pode-se dispor da base formal desenvolvida para manipulação de problemas multiobjetivo.

As duas abordagens a serem discutidas, *ε-restrito* e *paramétrica*, não consideram a influência do controle sobre o problema de identificação. A principal motivação para o uso destas abordagens reside na possibilidade de existência de soluções múltiplas para o problema de identificação. Ambas utilizam conjuntos de medidas das saídas associadas a entradas fornecidas ao processo, podendo a função objetivo do problema de identificação ser expressa por (1.52) e sendo, assim, dependente apenas de  $\theta$ , como foi considerado em (1.51).

$$G(\theta) = \sum_{j=1}^n [h(\theta) - \hat{y}^j]' W^j [h(\theta) - \hat{y}^j] \quad (1.52)$$

Para a definição da solução do problema bi-critério expresso por (1.51) é necessário o conceito de ponto eficiente. Um ponto  $(u^*, \theta^*)$  é *eficiente* se e somente se não existe outro ponto  $(u, \theta)$  tal que

$$G(\theta) \leq G(\theta^*)$$

$$f(u, \theta) \leq f(u^*, \theta^*)$$

com a desigualdade estrita válida para pelo menos uma das expressões. Assim, qualquer ponto no qual não é possível melhorar um dos índices de desempenho sem prejuízo do outro caracteriza-se como um ponto eficiente. Na Figura 1.6, pode-se observar que o conjunto de pontos eficientes encontra-se entre os mínimos das funções objetivo.

É óbvio que a solução de (1.51) deve pertencer ao conjunto de pontos eficientes. Considerando-se que a eficácia do processo de otimização depende da fidelidade da representação do sistema real pelo modelo, define-se a solução ótima de (1.51) como o ponto eficiente  $(u^*, \theta^*)$ , com  $\theta^*$  que minimiza  $G(\theta)$ , isto é, o ponto eficiente que fornece a melhor solução para a identificação.

### 1.6.1 Abordagem $\epsilon$ -restrito

A abordagem  $\epsilon$ -restrito transforma o problema bi-critério em um problema escalar equivalente, considerando o critério de identificação de parâmetros

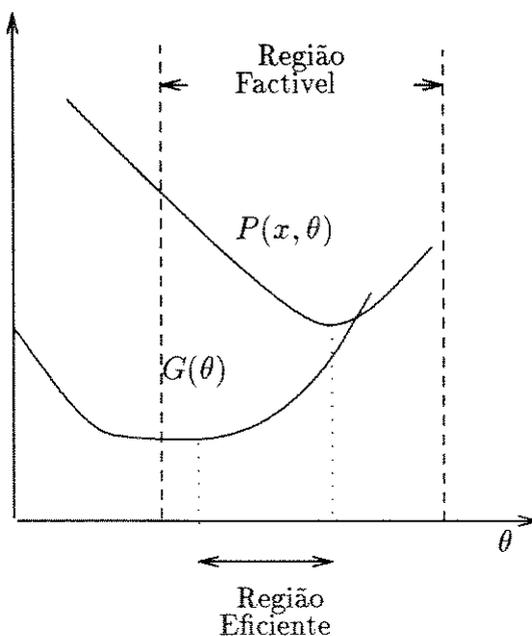


Figura 1.6: Região Eficiente

como uma restrição de desigualdade do problema de otimização. Assim

$$\begin{aligned} \min_{u, \theta} \quad & f(u, \theta) \\ & g(u, \theta) \leq 0 \\ & G(\theta) \leq \epsilon \\ & \epsilon \geq \delta > 0 \end{aligned} \quad (1.53)$$

onde  $\delta$  representa a menor tolerância ao erro de  $G(\theta)$ .

A equivalência entre os problemas (1.51) e (1.53) pode ser mostrada através do seguinte teorema

**Teorema 1.1** *Seja  $\epsilon \leq \min_{\theta} G(\theta)$ , considere  $V^* = (u^*, \theta^*)$  que resolve (1.53) e assumo que se  $V^*$  não é única, então  $V^*$  é uma solução ótima de (1.53) com valor mínimo para  $G(\theta)$ . Então  $V^*$  resolve (1.51).*

*Prova:* Por contradição, suponha que  $V^*$  não soluciona (1.51). Então, existe  $\hat{V} = (\hat{u}, \hat{\theta})$ , satisfazendo (1.51), tal que

$$f(\hat{V}) < f(V^*) \quad \text{e} \quad G(\hat{\theta}) \leq G(\theta^*) \quad (1.54)$$

ou

$$G(\hat{\theta}) < G(\theta^*) \quad e \quad f(\hat{V}) \leq f(V^*) \quad (1.55)$$

Porém, a expressão (1.54) contradiz o fato de que  $V^*$  soluciona (1.53), enquanto (1.55) nega a hipótese de que  $V^*$  é solução de (1.53) com o menor  $G(\theta)$ . Desta maneira, o teorema está demonstrado (Haimes et al, 1971).

É bom lembrar que se existe uma única solução  $\theta^*$  para o problema de identificação de parâmetros e existe  $u$  tal que  $g(u, \theta^*) \leq 0$ , pode-se resolver o problema (1.53) sequencialmente, determinando inicialmente  $\theta^*$ . A formulação expressa em (1.53) é útil quando estas condições não são satisfeitas.

O problema (1.53) é complexo. Para facilitar a implementação computacional, é conveniente decompô-lo em uma estrutura multinível. O desacoplamento entre as variáveis  $u$  e  $\theta$  para a separação dos problemas de minimização pode ser feito através do método de coordenação pelo objetivo. Considere, então,  $\theta$  como um parâmetro  $\sigma$  nos pontos onde há interação entre  $u$  e  $\theta$ , acrescentando a seguinte restrição de igualdade ao problema

$$\theta = \sigma \quad (1.56)$$

Assim, a função langrageana pode ser expressa por

$$L = f(u, \sigma) + \mu'g(u, \sigma) + \mu_0[G(\theta) - \epsilon] + \lambda'[\sigma - \theta] \quad (1.57)$$

Pode-se, então, decompor o lagrangeano em

$$L_1(u, \sigma, \mu, \lambda) = f(u, \sigma) + \mu'g(u, \sigma) + \lambda'\sigma \quad (1.58)$$

$$L_2(\theta, \mu_0, \lambda) = \mu_0[G(\theta) - \epsilon] - \lambda'\theta \quad (1.59)$$

A estrutura hierárquica obtida é composta por três níveis, como pode ser visto na Figura 1.7. O primeiro nível é composto por dois subproblemas que são coordenados pelo nível 2 através do multiplicador  $\lambda$ , procurando igualar os valores de  $\theta$  e  $\sigma$ . De acordo com Haimes e Wismer (1972), o cálculo de  $\lambda$  é feita com base em algoritmos dos tipos Gauss-Seidel e gradiente. Estes dois níveis trabalham com o valor de  $\epsilon$  definido pelo nível 3, que é recalculado quando há convergência no nível 2, até que não seja mais possível diminuir  $\epsilon$ .

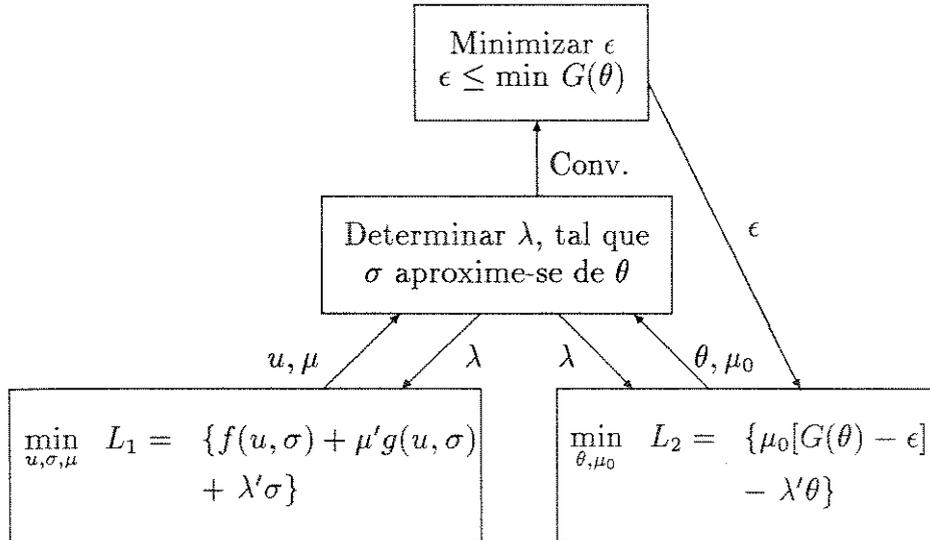


Figura 1.7: Estrutura Hierárquica - Abordagem  $\epsilon$ -restrito

### 1.6.2 Abordagem Paramétrica

A abordagem paramétrica propõe a seguinte formulação para o problema bi-critério expresso em (1.51)

$$\begin{aligned} \min_{u, \theta} \quad & [\lambda G(\theta) + (1 - \lambda)f(u, \theta)] \\ & g(u, \theta) \leq 0 \\ & 0 < \lambda < 1 \end{aligned} \quad (1.60)$$

onde  $\lambda$  é um parâmetro a ser selecionado e representa um fator de ponderação entre os critérios de identificação e otimização.

Considere a função lagrangeana associada a (1.60)

$$L = \lambda G(\theta) + (1 - \lambda)f(u, \theta) + \mu'g(u, \theta) \quad 0 < \lambda < 1 \quad (1.61)$$

O Lema citado a seguir estabelece que esta abordagem pode ser usada para encontrar os pontos eficientes de (1.51)

**Lema 1.1** *Suponha que condições de regularidade (Luenberger, 1984) são satisfeitas. Se, para algum  $0 < \lambda < 1$ ,  $(u^0, \theta^0, \mu^0)$  é um ponto de sela de (1.61), então  $(u^0, \theta^0)$  é um ponto eficiente de (1.51).*

*Prova :* McGrew e Haimes, 1974.

De acordo com resultados anteriores, a solução do problema integrado deve estar associada a  $\lambda = 1$ , ou seja, ao ponto onde apenas o critério de identificação é minimizado. Observe que o Lema 1.1 é válido apenas no intervalo aberto  $(0, 1)$  e, conseqüentemente, não garante que o ponto associado a  $\lambda = 1$  será um ponto eficiente. É possível mostrar (McGrew e Haimes, 1974) que a solução paramétrica (1.60) aproxima-se da solução do problema integrado de identificação de parâmetros e otimização, no limite, quando  $\lambda$  tende a 1 (Figura 1.8).

Também neste caso, a implementação computacional da estratégia de tratamento do problema conjunto pode ser simplificada pela utilização de uma estrutura multinível. A técnica de decomposição utilizada é a coordenação pelo modelo.

Na decomposição, adiciona-se a seguinte restrição de igualdade

$$\theta = \sigma \quad (1.62)$$

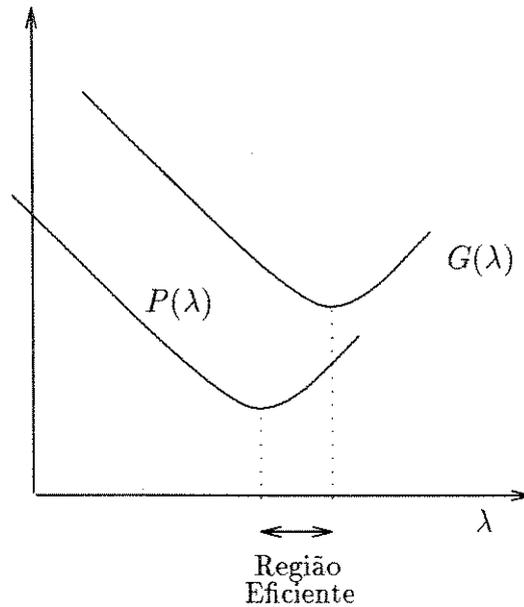


Figura 1.8: Comportamento das Funções Objetivo

Substituindo-se  $\theta$  por  $\sigma$  onde há acoplamento entre  $u$  e  $\theta$  no lagrangeano (1.61) e considerando-se a restrição (1.62), obtém-se

$$L(u, \theta, \sigma, \mu, \lambda, \zeta) = \lambda G(\theta) + (1 - \lambda)f(u, \sigma) + \mu'g(u, \sigma) + \zeta'(\sigma - \theta) \quad (1.63)$$

A função de Lagrange pode, então, ser decomposta em

$$L^1(\theta, \zeta; \sigma, \lambda) = \lambda G(\theta) + \zeta'(\sigma - \theta) \quad (1.64)$$

$$L^2(u, \mu; \sigma, \lambda) = (1 - \lambda)f(u, \sigma) + \mu'g(u, \sigma) \quad (1.65)$$

Os subsistemas de primeiro nível são compostos, assim, pela minimização das funções (1.64) e (1.65), supondo  $\sigma$  e  $\lambda$  conhecidos. O segundo nível é responsável pela determinação do  $\sigma^*$  a ser usado no primeiro nível, enquanto que o terceiro nível define o parâmetro  $\lambda$ , evoluindo a partir da proximidade de zero até a proximidade de 1.

A manipulação das restrições de desigualdade através de funções de penalização tem-se mostrado conveniente sob o ponto de vista computacional. Assim, para cada restrição pode-se associar

$$\frac{R}{\max [10^{-30}, -g_i(u, \theta)]} \quad i = 1 \dots q \quad (1.66)$$

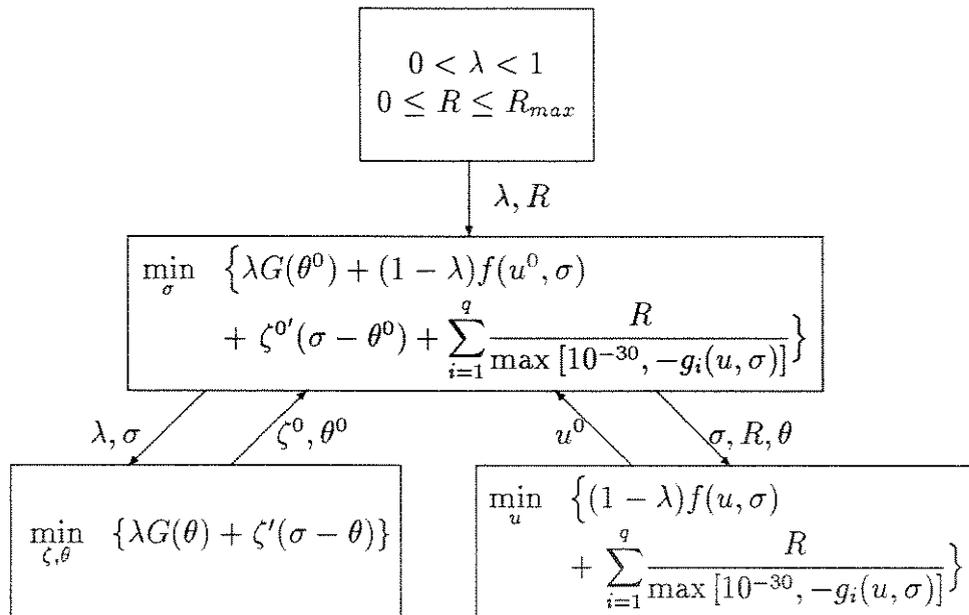


Figura 1.9: Estrutura Hierárquica - Abordagem Paramétrica

onde  $R$  é o fator de penalização, determinado pelo nível 3, variando de  $R_{max}$  até zero.

O processamento é, assim, distribuído gerando subproblemas que interagem entre si com base num procedimento que converge para a solução ótima de maneira iterativa. A estrutura hierárquica obtida é mostrada na Figura 1.9.

## 1.7 Abordagem MinMax

Trata-se de uma abordagem conservativa. Supõe que a *natureza* escolherá o pior valor possível para os parâmetros. O problema conjunto é formulado como

$$\min_u \left\{ \max_{\theta \in \Theta} f(u, \theta) \right\} \quad (1.67)$$

$$g(x, u, \theta) \leq 0$$

O conjunto  $\Theta$  é composto pelos  $\theta$ 's que resolvem o problema de identificação e é expresso por

$$\Theta = \{ \theta | G(\theta) \leq \epsilon, \quad \epsilon \geq \delta > 0 \} \quad (1.68)$$

onde  $\epsilon$  é pré-determinado.

Esta estratégia tem por objetivo otimizar o desempenho do sistema na situação mais desfavorável.

## 1.8 Conclusão

Neste capítulo foi apresentada uma visão geral das diversas abordagens do problema conjunto de identificação de parâmetros e otimização escalar, ressaltando o acoplamento existente entre os dois problemas e as técnicas de decomposição adequadas a cada abordagem.

Uma característica comum a todas estas abordagens é a existência de um único critério de desempenho. Num contexto em que os parâmetros do modelo dependam do controle aplicado, a escolha do critério induz os valores dos parâmetros estimados. É óbvio que qualquer alteração referente ao critério de desempenho fará com que os parâmetros obtidos deixem de ter validade.

No próximo capítulo, o problema conjunto de identificação de parâmetros e otimização de sistemas é abordado a luz de resultados de otimização multicritério.

## Capítulo 2

# Identificação e Otimização Multicritério: Aspectos Teóricos

### 2.1 Introdução

Problemas práticos frequentemente envolvem múltiplos critérios de desempenho. Por outro lado, a escolha do critério usualmente determina os parâmetros estimados do processo. Neste capítulo, é proposta uma extensão do algoritmo dois estágios, cuja principal característica é a utilização de otimização multicritério na solução do problema integrado. Aspectos como convergência do método, análise de sensibilidade, aplicação de alguns resultados de otimização multicritério e uma extensão do método do Gradiente Modificado (Geromel e Ferreira, 1986) para problemas de programação não-linear são também discutidos.

### 2.2 Formulação do Problema

Considere a otimização do desempenho de um sistema real, utilizando múltiplos critérios. Este problema pode ser estabelecido matematicamente

como

$$\begin{aligned} \min_x \quad & \mathcal{V}(f(x, y)) \\ & y = p(x) \\ & x \in \mathcal{X} = \{x \in \mathcal{R}^n : g(x, y) \leq 0\} \end{aligned} \quad (P)$$

onde <sup>1</sup>

$x \in \mathcal{R}^n$  : variáveis de decisão

$\mathcal{X} \subseteq \mathcal{R}^n$  : espaço de decisão

$g(\cdot) : \mathcal{X} \rightarrow \mathcal{R}^t$  : vetor t-dimensional de restrições de desigualdade

$p(\cdot) : \mathcal{X} \rightarrow \mathcal{R}^r$  : vetor r-dimensional de saídas da planta

$f(\cdot) : \mathcal{X} \rightarrow \mathcal{R}^q$  : vetor q-dimensional dos critérios de desempenho

$\mathcal{F} \subseteq \mathcal{R}^q$  : espaço dos critérios

$$\mathcal{F} = \{\zeta \in \mathcal{R}^q : \zeta = f(x), x \in \mathcal{X}\}$$

$\mathcal{V}(\cdot) : \mathcal{F} \rightarrow \mathcal{R}$  : Função Valor ou Utilidade

A Função Valor  $\mathcal{V}(\cdot)$ , implicitamente conhecida, agrega as preferências de um Decisor (Yu, 1985). Supõe-se que todas as funções são continuamente diferenciáveis e que  $\mathcal{V}(\cdot)$  é não-decrescente com relação a cada critério. Esta condição garante que uma das soluções de (P) é *eficiente* de acordo com a seguinte definição.

**Definição 2.1** *Uma decisão factível  $x^* \in \mathcal{X}$  é eficiente se, para todo  $x \in \mathcal{X}$  satisfazendo  $f(x, p(x)) \leq f(x^*, p(x^*))$ ,  $f(x, p(x)) = f(x^*, p(x^*))$ .*

O conjunto de decisões eficientes de (P) será referido como  $\mathcal{X}^*$ . No espaço dos critérios, o conjunto de soluções eficientes correspondente é denotado por  $\Gamma = \{\zeta \in \mathcal{R}^q : \zeta = f(x, p(x)), x \in \mathcal{X}^*\}$ .

Como o mapeamento  $p(\cdot)$  não é conhecido explicitamente, adota-se um modelo aproximado  $m(\cdot) : \mathcal{X} \times \Theta \rightarrow \mathcal{R}^s$ .

$$y = m(x, \theta) \quad (2.1)$$

<sup>1</sup>Assume-se que, se existem, as variáveis dependentes foram eliminadas. A notação  $x$  para variáveis de decisão foi mantida por conveniência.

Os parâmetros  $\theta \in \Theta \subseteq \mathcal{R}^r$  são determinados pela minimização de um índice de comparação entre as saídas do modelo e da planta. Um critério de mínimos quadrados é comumente usado.

$$\min_{\theta \in \Theta} \|p(x) - m(x, \theta)\|_W^2 \quad (2.2)$$

Note que o valor de  $\theta \in \Theta$  determinado por (2.2) deve ser tal que  $\mathcal{X}_\theta = \{x \in \mathcal{R}^n : g(x, m(x, \theta)) \leq 0\}$  seja não vazio.

## 2.3 Algoritmo Dois-Estágios Generalizado

Supondo-se momentaneamente que a função valor  $\mathcal{V}(\cdot)$  é conhecida, o problema ( $P$ ) pode ser solucionado seguindo-se o procedimento básico abaixo

### Algoritmo Dois-Estágios Generalizado

Obtenha algumas medidas iniciais  $p(x^i)$ ,  $i = 1, \dots, N$  ( $N \geq r$ ) e faça  $k \leftarrow N$

**Passo 1** (Estimação) Resolva o problema de estimação de parâmetros

$$\min_{\theta \in \Theta} \sum_{i=1}^k \|p(x^i) - m(x^i, \theta)\|_W^2, \quad (2.3)$$

Obtenha a solução ótima  $\theta^k \in \Theta$ .

**Passo 2** (Otimização Multicritério) Com os parâmetros estimados  $\theta^k \in \Theta$ , obtenha  $x^k$  resolvendo

$$\begin{aligned} \min_{x \in \mathcal{X}_{\theta^k}} \mathcal{V}(f(x, y)) \\ y = m(x, \theta^k) \end{aligned} \quad (P^k)$$

Aplice  $x^k$  ao processo real para obter uma medida adicional  $p(x^k)$ . Faça  $k \leftarrow k + 1$  e vá para o *Passo 1*.

Os passos de estimação e otimização são executados sequencialmente até que nenhuma melhoria no desempenho do sistema seja obtida.

Observe que o algoritmo dois estágios clássico é obtido no caso de um único critério, quando a minimização de  $\mathcal{V}(\cdot)$  é equivalente a minimização de  $f(\cdot)$  devido as hipóteses prévias sobre  $\mathcal{V}(\cdot)$ . Além disso, considere  $\{\theta^k\}$  e  $\{x^k\}$  como sendo as sequências geradas pelo algoritmo,  $\theta^* = \lim_{k \rightarrow \infty} \theta^k$  e  $x^* = \lim_{k \rightarrow \infty} x^k$ . De acordo com os resultados anteriores sabe-se que se<sup>2</sup>

$$m(x^*, \theta^*) = p(x^*) \quad (2.4)$$

$$m_x(x^*, \theta^*) = p_x(x^*) \quad (2.5)$$

então a solução  $x^* \in \mathcal{X}$  obtida pelo algoritmo dois passos clássico é também uma solução ótima do problema integrado ( $P$ ), isto é, o problema ( $P^k$ ) converge para ( $P$ ) quando  $k \rightarrow \infty$ . Considerando o problema genérico de otimização multicritério ( $P^k$ ), qualquer solução de ( $P^k$ ) é eficiente no seguinte sentido.

**Definição 2.2** Uma decisão factível  $x^k \in \mathcal{X}_{\theta^k}$  é dita ser  $\theta^k$ -eficiente se, para todo  $x \in \mathcal{X}_{\theta^k}$  satisfazendo  $f(x, m(x, \theta^k)) \leq f(x^k, m(x^k, \theta^k))$ ,  $f(x, m(x, \theta^k)) = f(x^k, m(x^k, \theta^k))$ .

O conjunto completo de soluções  $\theta^k$ -eficientes de ( $P^k$ ) será denotado por  $\mathcal{X}_{\theta^k}^*$  ( $\Gamma_{\theta^k}$  no espaço dos critérios). A relação de  $\mathcal{X}_{\theta^k}^*$  com o conjunto de soluções eficientes do problema original é estabelecido no Teorema 2.1 a partir do seguinte resultado geral para problemas multiobjetivos.

**Lema 2.1** Seja  $x^* \in \mathcal{X}^*$ . Então  $x^*$  resolve os  $q$  problemas escalares

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & f_i(x) \\ & f_j(x) \leq f_j(x^*), \quad \forall j \neq i \end{aligned} \quad (2.6)$$

*Prova:* Chankong e Haimes, 1982.

**Teorema 2.1** Seja  $x^k$  qualquer solução  $\theta^k$ -eficiente, tal que  $p(x^k) = m(x^k, \theta^k)$  e  $p_x(x^k) = m_x(x^k, \theta^k)$ . Então  $x^k$  é também uma solução eficiente do problema ( $P$ ).

---

<sup>2</sup> $m_x = \frac{dm}{dx}$  e  $p_x = \frac{dp}{dx}$ .

*Prova:* De acordo com o Lema 2.1,  $x^k \in \mathcal{X}_{\theta^k}^*$  resolve os  $q$  problemas escalares

$$\begin{aligned} \min_{x \in \mathcal{X}_{\theta^k}} \quad & f_i(x, m(x, \theta^k)) \\ & f_j(x, m(x, \theta^k)) \leq f_j(x^k, m(x^k, \theta^k)), \quad \forall j \neq i \end{aligned} \quad (2.7)$$

Desde que as saídas e as derivadas das saídas com relação ao controle são iguais, o mesmo  $x^k$  resolve os  $q$  problemas

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & f_i(x, p(x)) \\ & f_j(x, p(x)) \leq f_j(x^k, p(x^k)), \quad \forall j \neq i \end{aligned} \quad (2.8)$$

e, assim,  $x^k \in \mathcal{X}^*$ .  $\square$

**Corolário 2.1** *Seja  $(x^*, \theta^*)$  a solução obtida através do algoritmo dois passos generalizado e suponha que as condições (2.4) e (2.5) são satisfeitas. Se  $x^*$  é  $\theta^*$ -eficiente, então  $x^*$  é também uma solução eficiente do problema  $(P)$ .*

O Teorema 2.1 estabelece formalmente que se (2.4) e (2.5) são válidas, então toda solução  $\theta^k$ -eficiente é também uma solução eficiente do problema conjunto  $(P)$ . Observe que o conjunto de decisões eficientes de  $(P^k)$  quando  $k \rightarrow \infty$  pode ser obtido, por exemplo, através do problema  $\epsilon$ -restrito (Chankong e Haimes, 1982)

$$\begin{aligned} \min_{x \in \mathcal{X}_{\theta^*}} \quad & f_i(x, m(x, \theta^*)) \\ & f_j(x, m(x, \theta^*)) \leq \epsilon_j, \quad \forall j \neq i \end{aligned} \quad (2.9)$$

onde  $\epsilon_j, \forall j \neq i$  são tais que (2.9) admite alguma solução factível.

Assim, as condições de otimalidade do algoritmo dois passos clássico podem ser estendidas para o caso multicritério. O Teorema 2.1 estabelece condições para obtenção do conjunto de decisões eficientes do problema original  $(P)$ .

O desempenho do algoritmo dois passos generalizado com respeito a convergência e análise de sensibilidade é tratado na próxima seção.

## 2.4 Convergência e Análise de Sensibilidade

A estabilidade e a sensibilidade da solução multicritério com relação a variação dos parâmetros do modelo são aspectos importantes a serem analisados. Sensibilidade em otimização multicritério é discutida por Rarig e Haimes (1983), onde é proposto um índice de dispersão como medida de incerteza.

Na obtenção do índice, Rarig e Haimes (1983) consideram o problema multicritério na forma  $\epsilon$ -restrito

$$\begin{aligned} \min_x \quad & f_1(x, \theta) \\ & f_j(x, \theta) \leq \epsilon_j, \quad j = 2, \dots, m \\ & g_l(x, \theta) \leq b_l, \quad l = 1, \dots, t \end{aligned} \quad (2.10)$$

Supondo que (2.10) admite uma solução ótima  $x^*$  e que todas as  $\epsilon$ -restrições estão ativas em  $x^*$ , pode-se expressar a dependência da solução ótima  $f_1^*$  com relação aos parâmetros  $\theta$  e aos valores de  $\epsilon$  e  $b$  como

$$f_1^* = f_1^*(\theta, \epsilon, b) \quad (2.11)$$

Rarig e Haimes (1983) supõem que a incerteza dos parâmetros é devida basicamente a aleatoriedade. Os parâmetros são tratados como variáveis aleatórias independentes com médias  $\mu_i$  e variâncias  $\sigma_i$  ( $i = 1, \dots, r$ ) conhecidas.

Usando uma expansão em série de Taylor em torno da média dos parâmetros, obtém-se

$$f_1^* = f_1^*(\mu) + \sum_{i=1}^q \left( \frac{\partial f_1^*}{\partial \theta_i} \right) (\theta_i - \mu_i) \quad (2.12)$$

Assim, o valor esperado e a variância de  $f_1^*$  podem ser expressos por

$$\mathcal{E}\{f_1^*\} = f_1^*(\mu) \quad (2.13)$$

$$\sigma^2\{f_1^*\} = \sum_{i=1}^q \left( \frac{\partial f_1^*}{\partial \theta_i} \right)^2 \sigma_i^2 \quad (2.14)$$

O índice de risco/dispersão  $\Omega$  é então definido como em (2.15). Grandes valores de  $\Omega$  indicam maior probabilidade de que a solução real possa desviar-se do valor predito.

$$\Omega = \sqrt{\sum_{i=1}^q \left( \frac{\partial f_1^*}{\partial \theta_i} \right)^2 \sigma_i^2} \quad (2.15)$$

É importante mais uma vez ressaltar que no estudo desenvolvido por Rarig e Haimes (1983), as incertezas são decorrentes da aleatoriedade dos parâmetros, enquanto que no contexto da abordagem proposta neste capítulo, os parâmetros são resultantes de estimativas baseadas em medidas. Assim sendo, as propriedades assintóticas do estimador de parâmetros a mínimos quadrados (Sorenson, 1980) podem ser utilizadas na condução de um estudo similar.

Suponha que  $k$  medidas ( $k \geq r$ )  $p^i = p(x^i)$ ,  $i = 1, \dots, k$  disponíveis estão relacionadas com as saídas do modelo  $m^i(\theta) = m(x^i, \theta)$ ,  $i = 1, \dots, k$  de acordo com

$$p^i = m^i(\theta) + v^i, \quad i = 1, \dots, k \quad (2.16)$$

onde  $v^i$  são erros de medição, representados por uma sequência de ruído branco gaussiano de média nula. Além disso, suponha que as matrizes de ponderação de (2.3) são identidades e que não são impostas restrições aos parâmetros, isto é,  $\Theta = \mathcal{R}^r$ . Por simplicidade, considere apenas medidas escalares.

Seja o estimador de parâmetros baseado nas  $k$  medidas  $p^i$ ,  $i = 1, \dots, k$

$$\theta^{k+1} = \theta^k + \kappa^k (p^k - m^k(\theta^k)) \quad (2.17)$$

onde  $\kappa^k \in \mathcal{R}^r$  é um vetor de ganho. Demonstra-se (Sorenson, 1980) que se  $\theta^k$  é uma estimativa razoável de  $\theta$  então, sob condições apropriadas, a seguinte sequência de ganho pode ser definida

$$\kappa^k = \frac{\eta^k}{\sum_{i=1}^k \|\eta^i\|^2} \quad (2.18)$$

onde  $\eta^k = m_{\theta}^k(\theta) \Big|_{\theta=\theta^k}$ . Uma desvantagem deste estimador é que o ganho  $\kappa^k$  não é ótimo, o que pode determinar uma convergência lenta do procedimento.

O cálculo de  $\theta^{k+1}$  através de (2.17) envolve apenas dados disponíveis. Assim, é possível estimar o seu efeito em termos de critério de desempenho.

Considere o problema de otimização ( $P^k$ ) na forma  $\epsilon$ -restrito

$$\begin{aligned} \min_x \quad & f_i(x, \theta^k) \\ & f_j(x, \theta^k) \leq \epsilon_j^k, \quad \forall j \neq i \\ & g_l(x, \theta^k) \leq 0, \quad l = 1, \dots, t \end{aligned} \quad (2.19)$$

Seja  $x^k$  um mínimo local estrito de (2.19) e  $\mathcal{L}^k = \{l : g_l(x, \theta^k) = 0, \quad l = 1, \dots, t\}$ . Suponha que todas as  $\epsilon^k$ -restrições estão ativas em  $x^k$ . Então,

existe uma função  $x(\epsilon, \theta)$ , numa vizinhança suficientemente pequena de  $\epsilon_j^k$  e  $\theta^k$ , que satisfaz as condições necessárias e suficientes de Kuhn–Tucker para um mínimo local estrito de (2.19) (Luenberger, 1984). É claro que  $x^k = x(\epsilon^k, \theta^k)$  e que a função  $x(\epsilon, \theta)$  pode ser caracterizada através do problema

$$\begin{aligned} \min_x \quad & f_i(x, \theta) \\ & f_j(x, \theta) \leq \epsilon_j, \quad \forall j \neq i \\ & g_l(x, \theta) \leq 0, \quad l \in \mathcal{L}^k \end{aligned} \quad (2.20)$$

A solução ótima de (2.20) pode ser escrita em função dos parâmetros  $\theta$  e dos valores  $\epsilon_j$ ,  $\forall j \neq i$  como

$$f_i = f_i(\theta, \epsilon) \quad (2.21)$$

Usando o estimador (2.17) com  $\kappa^k$  dado por (2.18), uma aproximação de primeira ordem de  $f_i(\theta_{k+1}, \epsilon)$  baseada na expansão de Taylor seria

$$f_i(\theta^{k+1}, \epsilon) - f_i(\theta^k, \epsilon) = \frac{d}{d\theta} f_i(\theta^k, \epsilon)' \kappa^k (p^k - m^k(\theta^k)) \quad (2.22)$$

Pode-se observar pela expressão (2.22) que quando  $k \rightarrow \infty$  e  $\theta^k \rightarrow \theta^*$ , a sensibilidade de  $f_i(\theta^k, \epsilon)$  com respeito a variações em  $\epsilon_j$ ,  $\forall j \neq i$  diminui progressivamente. Assim, como  $\epsilon_j$ ,  $\forall j \neq i$  geralmente está associado a relações de compromisso entre os critérios, a influência das alterações nos parâmetros do modelo nas decisões obtidas no passo de otimização multicritério diminui a medida que  $k \rightarrow \infty$ .

Note, finalmente, que (2.22) pode ser vista como uma medida do impacto da solução no passo de otimização. Neste sentido, pode ser usada como um critério adicional, através do qual a sensibilidade das soluções do Decisor pode ser avaliada e eventualmente minimizada.

## 2.5 Aspectos do Problema Multicritério ( $P_2^k$ )

Nesta seção, a aplicação de alguns resultados de otimização multicritério no contexto do problema ( $P^k$ ) é analisada. Uma discussão completa dos problemas, métodos e aplicações de programação multiobjetivo pode ser encontrada em Hwang e Masud (1979), entre outros.

No domínio da otimização multicritério, grande ênfase é dedicada ao caso bicritério. A teoria e os métodos existentes mostram-se especialmente convenientes no tratamento deste tipo de formulação. Além disso, a maioria dos problemas reais apresenta algum tipo de *dicotomia*, que pode ser expressa matematicamente através de um problema bicritério. Se suposições de convexidade são estabelecidas, algoritmos muito eficientes podem ser derivados.

Considere uma formulação bicritério do problema ( $P^k$ ) com dependência explícita dos parâmetros estimados  $\theta^k$

$$\min_{x \in \mathcal{X}_{\theta^k}} \mathcal{V}(f_1(x, \theta^k), f_2(x, \theta^k)) \quad (P_1^k)$$

e suponha que as funções  $f_i(\cdot, \theta)$ ,  $i = 1, 2$  são convexas sobre um conjunto compacto convexo  $\mathcal{X}_{\theta^k}$  para cada  $\theta \in \Theta$ . Além disso, assuma que  $\mathcal{V}(\cdot)$  é uma função valor estritamente quasi-convexa.

Seja  $\mathcal{X}_{\theta^k}^*$  o conjunto  $\theta^k$ -eficiente associado a ( $P_1^k$ ). Como o problema foi suposto convexo, o conjunto  $\mathcal{X}_{\theta^k}^*$  pode ser obtido como solução de

$$\min_{x \in \mathcal{X}_{\theta^k}} w f_1(x, \theta^k) + (1 - w) f_2(x, \theta^k) \quad (2.23)$$

quando  $w$  varia no intervalo  $[0, 1]$ . Seja  $\mathcal{X}_{\theta^k}(w)$  o conjunto de soluções ótimas de (2.23) para um dado valor de  $0 \leq w \leq 1$  e defina

$$\phi(w, \theta^k) = \min_{x \in \mathcal{X}_{\theta^k}(w)} \mathcal{V}(f(x, \theta^k)) \quad (2.24)$$

O problema ( $P_1^k$ ) pode ser formulado como

$$\min_{0 \leq w \leq 1} \phi(w, \theta^k) \quad (P_2^k)$$

Geoffrion (1967) mostra que, sob as suposições acima, a função real  $\phi(\cdot, \theta^k)$  é unimodal e que o problema unidimensional ( $P_2^k$ ) pode ser eficientemente solucionado através de procedimentos de busca baseados na sequência de Fibonacci (Luenberger, 1984). Note que a única informação necessária para resolver ( $P_2^k$ ) é se, dados dois pontos  $w^1, w^2 \in [0, 1]$ , o valor de  $\phi(w^1, \theta^k)$  é menor, igual ou maior do que  $\phi(w^2, \theta^k)$ , o que é garantido se o Decisor for capaz de fazer comparações ordinais das alternativas propostas. O problema ( $P_1^k$ ) pode ser resolvido também através de procedimentos especializados para o caso bicritério (e.g., Cohon e Church, 1979; Payne et al, 1975; Ferreira e Geromel, 1990).

O formalismo adotado na seção anterior pode ser usado no desenvolvimento de algum tipo de análise de sensibilidade. Entretanto, em muitos casos, pode-se obter uma boa aproximação da curva de compromisso  $\Gamma_{\theta^k}$  do problema bicritério com pouco esforço computacional (veja, por exemplo, os mencionados Cohon e Church, 1979; Payne et al, 1975; Ferreira e Geromel, 1990). Nestes casos, as estimativas (2.17) poderiam ser usadas na predição do efeito das decisões sobre a curva de compromisso.

Finalmente, é conveniente discutir como a metodologia proposta comporta-se quando uma abordagem de critério global é adotada. Neste contexto, uma formulação muito utilizada é a de *Programação de Compromisso*, que baseia-se na minimização de  $\mathcal{V}(\cdot)$  do tipo

$$\mathcal{V}(f(x, p(x))) = \left( \sum_{i=1}^q (f_i(x, p(x)) - \underline{f}_i)^p \right)^{\frac{1}{p}} \quad (2.25)$$

ou seja, na determinação de  $x^* \in \mathcal{X}$  que minimize alguma norma  $l_p$ , com  $1 \leq p < \infty$ , entre o conjunto factível e as soluções ideais

$$\underline{f}_i = \min_{x \in \mathcal{X}} f_i(x, p(x)), \quad i = 1, \dots, q \quad (2.26)$$

Entretanto, esta formulação requer a solução de  $q$  problemas conjuntos que, possivelmente, levam a conjuntos diferentes de parâmetros. Desta maneira, o conceito de ponto ideal do problema conjunto de otimização e estimação de parâmetros parece estar mal definido. Além disso, não existe garantia de que as soluções ideais (2.26) possam ser obtidas, a não ser que (2.4) e (2.5) sejam válidas.

No contexto do problema ( $P^k$ ), uma implementação mais prática de (2.25) – (2.26) seria

$$\mathcal{V}(f(x, \theta^k)) = \left( \sum_{i=1}^q (f_i(x, \theta^k) - \underline{f}_i(\theta^k))^p \right)^{\frac{1}{p}} \quad (2.27)$$

onde

$$\underline{f}_i(\theta^k) = \min_{x \in \mathcal{X}_{\theta^k}} f_i(x, \theta^k), \quad i = 1, \dots, q \quad (2.28)$$

Esta estratégia elimina a necessidade de solução de  $q$  problemas conjuntos e fornece soluções ideais significativas a cada iteração do algoritmo dois passos generalizado. No caso particular em que  $(x^*, \theta^*)$  soluciona o problema conjunto e

$$m(x^i, \theta^*) = p(x^i), \quad i = 1, \dots, q \quad (2.29)$$

$$m_x(x^i, \theta^*) = p_x(x^i), \quad i = 1, \dots, q \quad (2.30)$$

onde  $x^i = \arg \min_{x \in \mathcal{X}_{\theta^*}} f_i(x, \theta^*)$ , os valores  $f_i(\theta^k)$  convergem para as soluções ideais (2.26) e o problema (2.27) converge para o problema original (2.25). Este comportamento é esperado quando os parâmetros são independentes do controle e o modelo é uma representação precisa do processo. No Capítulo 3 é apresentada uma aplicação baseada nestes resultados.

## 2.6 Aspectos Computacionais

Qualquer que seja a estratégia a ser adotada para resolver o problema multicritério ( $P^k$ ), a utilização de uma técnica eficiente para problemas de programação não-linear é um aspecto importante. Neste sentido, o método do Gradiente Modificado proposto por Geromel e Ferreira (1986) aliado a uma estratégia de solução paralela do problema dual associado baseada na interação de Jacobi (Bertsekas e Tsitsiklis, 1989) apresenta-se como uma opção bastante interessante.

Considere o problema genérico de programação não-linear

$$\begin{aligned} \min_{\underline{\theta} \leq \theta \leq \bar{\theta}} f(\theta) \\ A\theta = 0 \end{aligned} \quad (2.31)$$

Este método parte de uma solução factível  $\theta^0 \in \mathcal{R}^n$  para a geração de uma sequência  $\{\theta^k\}_0^\infty$  que converge para  $\theta^*$ , satisfazendo as condições de Kuhn-Tucker do problema (Luenberger, 1984). Cada elemento da sequência é gerado a partir do elemento imediatamente anterior de acordo com

$$\theta^{k+1} = \theta^k - \beta s(\theta^k) \quad (2.32)$$

onde  $s(\theta^k)$  é uma direção usável associada a  $\theta^k$  e  $\beta \in R_+$  é o tamanho do passo, tal que  $\theta^{k+1}$  é factível. Em Ferreira (1983) é utilizado

$$s(\theta) = \arg \min_{\xi} \left\{ \frac{1}{2} \|\nabla f(x) - \xi\|^2 \text{ s.a. } A\xi = 0, S(\theta)\xi \leq 0 \right\} \quad (2.33)$$

onde  $A \in \mathcal{R}^{m \times n}$  ( $m < n$ ), ( $\text{rank}(A) = m$ ) e  $S(\theta) \in \mathcal{R}^r \times \mathcal{R}^n$  é uma matriz sinal dada por

$$S(\theta) = \{S(\theta)_{ij}\} = \begin{cases} 1, & \theta_j = \underline{\theta}_j \\ -1, & \theta_j = \bar{\theta}_j \\ 0, & \text{caso contrário} \end{cases} \quad (2.34)$$

sendo  $r$  igual ao número de restrições de desigualdade ativas na solução corrente  $\theta$ .

Definindo  $\Omega(\theta) \triangleq \{\xi \in \mathcal{R}^n | S(\theta)\xi \leq 0\}$ , pode-se reescrever (2.33) como

$$s(\theta) \triangleq Proj_{N(A) \cap \Omega(\theta)} \nabla f(\theta) \quad (2.35)$$

Assim,  $s(\theta)$  é a projeção ortogonal do  $\nabla f(\theta)$  na interseção do espaço nulo de  $A$  com  $\Omega(\theta)$ . O gradiente modificado  $\nabla_m f$  é definido (Geromel e Ferreira, 1986) por

$$Proj_{N(A)} \nabla_m f = Proj_{N(A) \cap \Omega(\theta)} \nabla f(\theta) \quad (2.36)$$

Assim (2.35) pode ser reescrita como

$$s(\theta) = P \nabla_m f(\theta) \quad (2.37)$$

onde  $P = I - A'(AA')^{-1}A$  é o operador de projeção ortogonal no espaço nulo de  $A$ .

Geromel e Ferreira (1986) demonstram que

$$\nabla_m f(\theta) = \nabla f(\theta) - S(\theta)' \mu(\theta) \quad (2.38)$$

onde  $\mu(\theta)$  é solução do seguinte problema quadrático

$$\max_{\mu \geq 0} \left\{ -\frac{1}{2} \mu' Q(\theta) \mu + C(\theta)' \mu \right\} \quad (2.39)$$

onde  $Q(\theta) \triangleq S(\theta)PS(\theta)'$  e  $C(\theta) \triangleq S(\theta)P\nabla f(\theta)$ .

O problema (2.39) é equivalente a

$$\min_{\mu \geq 0} \left\{ \frac{1}{2} \mu' Q(\theta) \mu - C(\theta)' \mu \right\} \quad (2.40)$$

De acordo com o método proposto por Geromel e Ferreira (1986), a solução de 2.40 é obtida, por exemplo, utilizando-se o Método Dual Clássico (Lasdon, 1970). Entretanto, sendo  $Q$  semi-definida positiva com os elementos da diagonal principal positivos, o valor de  $\mu$  pode ser determinado através do algoritmo de Gauss-Seidel (Bertsekas e Tsitsiklis, 1989), diminuindo a complexidade computacional.

Considerando o problema (2.40) irrestrito, derivando-se a função objetivo com relação aos elementos de  $\mu$  e igualando-se a zero, obtém-se

$$-c_j + \sum_{k=1}^r q_{jk} \mu_k = 0, \quad j = 1, \dots, r \quad (2.41)$$

onde, por simplicidade, a dependência de  $Q$  e  $C$  com relação a  $\theta$  foi omitida.

Assim, a solução para a  $j$ -ésima componente de  $\mu$  é

$$\bar{\mu}_j = \mu_j - \frac{1}{q_{jj}} \left( -c_j + \sum_{k=1}^r q_{jk} \mu_k \right) \quad (2.42)$$

Observando a restrição de não negatividade<sup>3</sup>

$$\mu_j = [\bar{\mu}_j]^+ \quad (2.43)$$

Uma iteração do método de Gauss-Seidel pode, então, ser expressa por

$$\begin{aligned} \mu_j &= [\bar{\mu}_j]^+ = \left[ \mu_j - \frac{1}{q_{jj}} \left( -c_j + \sum_{k=1}^r q_{jk} \mu_k \right) \right]^+ \\ \mu_i &= \mu_i \quad \forall i \neq j \end{aligned} \quad (2.44)$$

O método de Jacobi, parte de (2.42) e define uma iteração padrão mais propícia para implementação em paralelo do cálculo das componentes do vetor  $\mu$ . No caso, para  $t = 0, 1, \dots$ ,

$$\mu_j(t+1) = [\bar{\mu}_j(t)]^+ = \left[ \mu_j(t) - \frac{\gamma}{q_{jj}} \left( -c_j + \sum_{k=1}^r q_{jk} \mu_k(t) \right) \right]^+ \quad j = 1, \dots, r \quad (2.45)$$

onde  $\gamma$  é o tamanho do passo.

Desta maneira, o seguinte procedimento numérico pode ser definido.

**Passo 1** Faça  $l = 0$  e determine  $P$ . Faça  $\mu(0) = 0$ .

**Passo 2** Calcule  $\theta^l$ ,  $Q(\theta^l)$ ,  $C(\theta^l)$ , repita (2.45) até a convergência e assumo o valor obtido como  $\mu(\theta^l)$ . Faça  $\mu(0) = \mu(\theta^l)$ .

---

<sup>3</sup>  $[\mu_i]^+ = \begin{cases} \mu_i & \text{se } \mu_i \geq 0 \\ 0 & \text{se } \mu_i < 0 \end{cases}$

**Passo 3** Determine o gradiente modificado e a direção de busca para  $\theta = \theta^l$  através de (2.38) e (2.37).

**Passo 4** Se  $\|s(\theta^l)\| \leq \epsilon$ ,  $\epsilon > 0$ , pare. Se não, vá para o próximo passo.

**Passo 5** Determine o tamanho do passo  $\beta \in R_+$  tal que

$$f(\theta^{l+1}) < f(\theta^l)$$

$$\theta^{l+1} = \theta^l - \beta s(\theta^l), \text{ factível}$$

**Passo 6** Faça  $l \leftarrow l + 1$  e vá para o *Passo 2*

É possível mostrar (Bertsekas e Tsitsiklis, 1989) que o procedimento descrito no *Passo 2* converge para a solução ótima de (2.40) se  $\gamma \in (0, \frac{2}{r})$ .

## 2.7 Conclusão

Neste capítulo, foi desenvolvida uma extensão do algoritmo dois passos clássico. A nova abordagem incorpora uma análise multicritério no tratamento do problema conjunto de estimação de parâmetros e otimização, cuja ausência no contexto de controle de processos tem sido criticada desde o início dos anos 70 (Foss, 1973).

A otimização multicritério representa uma importante ferramenta para o tratamento de problemas integrados de otimização e estimação de parâmetros. A abordagem multicritério viabiliza a incorporação de informações de compromisso entre os critérios na modelagem do sistema, uma vez que, de acordo com a estratégia proposta, os parâmetros do sistema refletirão preferências do Decisor.

O método do Gradiente Modificado aliado ao paralelismo na solução do problema dual permite o tratamento eficiente de problemas de programação não-linear. No caso do problema de identificação e otimização conjuntas, este aspecto é de vital importância tendo em vista a necessidade de resolver cada um dos problemas diversas vezes até que a solução conjunta seja obtida.

## Capítulo 3

# Identificação e Otimização Conjuntas de um Reator “Batch”

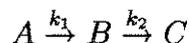
### 3.1 Introdução

Reatores químicos operando de modo contínuo têm sido frequentemente utilizados na avaliação do desempenho de abordagens do problema conjunto de identificação de parâmetros e otimização (McGrew e Haines, 1972; Roberts, 1979). Existem, entretanto, reações orgânicas complexas onde a fração de reagente convertida para o produto desejado é fundamental. Nestes casos, o processamento por batelada (“batch”) é mais indicado (Denbigh e Turner, 1984). Por outro lado, a otimização de processos químicos em geral envolve objetivos conflitantes, caracterizando-se como um problema multicritério.

Neste capítulo, o problema conjunto de identificação de parâmetros e otimização multiobjetivo de um reator por batelada é enfocado no contexto dos resultados discutidos no Capítulo 2.

### 3.2 Descrição Qualitativa

Considere um processo químico, no qual ocorrem duas reações consecutivas de primeira ordem



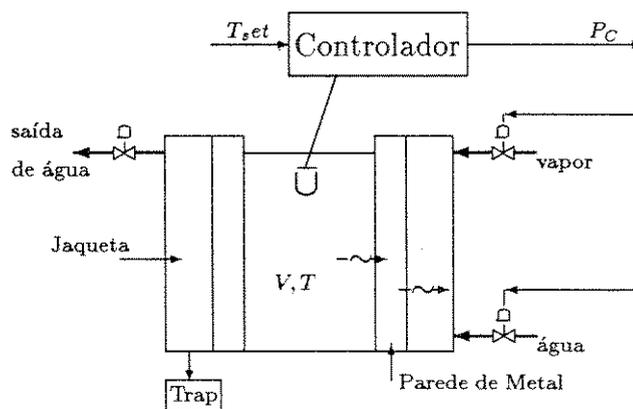


Figura 3.1: Batch Reactor

onde  $A$  é o reagente,  $B$  é o produto desejado,  $C$  é o produto indesejável, resultado da reação de  $B$ , e  $k_1$  e  $k_2$  são as taxas de reação (Luyben, 1973).

Inicialmente, o reagente  $A$  é carregado no recipiente do reator por batelada (Figura 3.1). Numa primeira fase, vapor é inserido na jaqueta que envolve o recipiente, elevando a temperatura  $T$  da massa de reação até o valor desejado. A seguir, a jaqueta é alimentada com água para remoção do calor exotérmico das reações e manutenção de um perfil de temperatura pré-estabelecido.

### 3.3 Descrição Quantitativa

A operação do reator por batelada pode ser descrita matematicamente através do modelo (Luyben, 1973) detalhado a seguir. As variáveis e os parâmetros que compõem este modelo estão descritas no Apêndice deste trabalho.

As transformações que ocorrem dentro do recipiente podem ser modeladas através de equações de continuidade, equações cinéticas e equações de energia. No caso, tem-se

- Balanços de Componentes

$$\frac{d}{dt}C_A = -k_1C_A(t), \quad C_A(0) = C_{A0} \quad (3.1)$$

$$\frac{d}{dt}C_B = k_1C_A - k_2C_B, \quad C_B(0) = 0 \quad (3.2)$$

$$k_i = \alpha_i e^{\frac{E_i}{RT}} \quad i = 1, 2 \quad (3.3)$$

- Equação de Energia (processo)

$$\frac{d}{dt}T = -\frac{\lambda_1}{\rho C_p} k_1 C_A - \frac{\lambda_2}{\rho C_p} k_2 C_B - \frac{Q_M}{V \rho C_p} \quad (3.4)$$

- Equação de Energia (parede de metal)

$$\frac{d}{dt}T_M = \frac{h_i A_i (T - T_M) - h_o A_o (T_M - T_j)}{\rho_M C_M V_M} \quad (3.5)$$

Com relação à dinâmica da jaqueta, as equações variam de acordo com a fase de processamento, na sequência abaixo

- FASE 1: Aquecimento (Suprimento de Vapor)

$$V_j \frac{d\rho_s}{dt} = w_s - w_c \quad (3.6)$$

$$\rho_s = \frac{MP_j}{R(T_j + 460)} \quad (3.7)$$

$$P_j = e^{\left(\frac{A_{vp}}{T_j + 460} + B_{vp}\right)} \quad (3.8)$$

$$W_s = C_{vs} X_s \sqrt{35 - P_j} \quad (3.9)$$

$$Q_j = -h_{os} A_o (T_j - T_M) \quad (3.10)$$

$$W_c = -\frac{Q_j}{H_s - h_c} \quad (3.11)$$

- FASE 2: Resfriamento (Suprimento de Água)

$$A_o = \left( \frac{A_{o_{total}}}{V_{j_{total}}} \right) V_j \quad (3.12)$$

$$\frac{d}{dt} V_j = F_{w0} \quad (3.13)$$

$$\frac{d}{dt} (V_j T_j) = F_{w0} T_{j0} \quad (3.14)$$

$$Q_j = h_{ow} A_o (T_M - T_j) \quad (3.15)$$

$$F_{w0} = C_{vw} X_w \sqrt{20} \quad (3.16)$$

- FASE 3 : Resfriamento (Jaqueta Cheia de Água)

$$\frac{dT_j}{dt} = \frac{F_{w0}}{V_j} (T_{j0} - T_j) + \frac{Q_j}{C_j V_j \rho_j} \quad (3.17)$$

### 3.4 Sistema de Controle

Na Figura 3.1 associa-se ao reator por batelada uma estrutura de controle de temperatura composta por válvulas de entrada de vapor e água, pelo sensor/transdutor e pelo controlador.

A temperatura da massa de reação é monitorada através do sensor, sendo convertida por um transdutor para um valor de pressão dentro da faixa 3–15 psig, de acordo com a relação

$$P_{TT} = 3 + (T - 50) \frac{12}{200} \quad (3.18)$$

onde

$P_{TT}$  : pressão de saída do transdutor de temperatura

$T \in [50, 250]^\circ F$

Um controlador PID digital recebe  $P_{TT}$  e gera a pressão  $P_C$  que é aplicada sobre as válvulas, determinando o fluxo de entrada de água e vapor na jaqueta. A expressão do controlador é a seguinte

$$P_C(t) = P_C(t - \Delta t) + k_p \{e(t) - e(t - \Delta t) + \frac{\Delta t}{\tau_i} e(t) + \frac{\tau_d}{\Delta t} [e(t) - 2e(t - \Delta t) + e(t - 2\Delta t)]\} \quad (3.19)$$

onde

$e(t)$  : diferença entre o valor medido e o valor calculado

$k_p$  : ganho proporcional

$\tau_d$  : constante de tempo derivativa

$\tau_i$  : constante de tempo integral

$\Delta t$  : intervalo de integração

As válvulas operam de acordo com as regras abaixo

Condição da Válvula de Vapor		fechada	→	aberta
Fração aberta $X_s$		0	→	1
Pressão do Controlador $P_C$ (psig)	3	9	→	15
Condição da Válvula de Água	aberta	→	fechada	
Fração Aberta $X_w$	1	→	0	

Assim, podem ser definidas as relações

$$X_s = \frac{P_C - 9}{6} \quad (3.20)$$

$$X_w = \frac{9 - P_C}{6} \quad (3.21)$$

### 3.5 Modelo Algébrico Simplificado

O modelo matemático descrito nos itens anteriores é utilizado na simulação do processo físico, descrita no item 3.7.1. Para estudo do problema conjunto de identificação e otimização, é necessário o desenvolvimento de um modelo algébrico simplificado. O controle do processo será feito com base em mudanças apropriadas do set-point de temperatura.

A partir das expressões (3.1) e (3.2), tem-se então que

$$C_A(t) = C_{A0}e^{-k_1t} \quad (3.22)$$

$$C_B(t) = \frac{k_1 C_{A0}}{k_1 - k_2} (e^{-k_2t} - e^{-k_1t}) \quad (3.23)$$

Pode-se notar que, com  $T$  constante,  $C_B$  é unimodal, ou seja, apresenta apenas um único máximo. Derivando-se  $C_B$  com relação ao tempo e igualando-se o resultado a zero, obtém-se uma expressão para o tempo ótimo de operação  $t^*$  em função da temperatura  $T$

$$t^*(T) = \frac{\ln \frac{k_2}{k_1}}{k_2 - k_1} \quad (3.24)$$

Assim, a concentração máxima de  $B$  pode ser expressa em função da temperatura  $T$ . Substituindo-se (3.24) em (3.23)

$$C_{B_{max}}(T, \alpha_1, \alpha_2, E_1, E_2) = \frac{k_1 C_{A0}}{k_1 - k_2} \left( e^{-k_2 t^*} - e^{-k_1 t^*} \right) \quad (3.25)$$

Considerando-se que as taxas de reação são expressas pela equação de Arrhenius (3.3) completa-se o modelo simplificado que é denominado neste trabalho como Modelo Não Linear.

Pode-se, ainda, utilizar uma aproximação linear da equação de Arrhenius para expressar as taxas de reação  $k_1$  e  $k_2$ . Expandindo-se a equação (3.3) em série de Taylor em torno de  $T^0$

$$k_i = k_i(T^0) + \left( \frac{\partial k_i}{\partial T} \right)_{T^0} (T - T^0) + \dots, \quad i = 1, 2$$

$$k_i = \frac{E_i k_i(T^0)}{RT^0{}^2} T + k_i(T^0) \left[ 1 + \frac{E_i}{RT^0} \right], \quad i = 1, 2$$

Fazendo  $a_i = \frac{E_i k_i(T^0)}{RT^0{}^2}$  e  $b_i = k_i(T^0) \left[ 1 + \frac{E_i}{RT^0} \right]$ ,  $i = 1, 2$  obtém-se

$$k_i(a_i, b_i, T) = a_i T + b_i, \quad i = 1, 2 \quad (3.26)$$

O modelo, neste caso, será chamado de Modelo Linear. Como pode-se observar, os parâmetros  $a_i$  e  $b_i$  ( $i = 1, 2$ ) são dependentes da temperatura e a expressão (3.26) é uma boa aproximação de (3.3) apenas em torno de  $T^0$ .

Tendo em vista que a taxa de reação é sempre um valor positivo, é necessário impor esta restrição à equação (3.26). Como  $k_i(a_i, b_i, T)$  é estritamente crescente, basta garantir que  $k_i$  é maior ou igual a zero no limite inferior do intervalo de temperatura considerado. Assim

$$k_i(a_i, b_i, T) = a_i T + b_i \geq 0, \quad i = 1, 2 \quad (3.27)$$

### 3.6 Formulação do Problema Conjunto

Considerando-se o modelo apresentado no item 3.5, supondo-se que os parâmetros  $\alpha_i$  e  $E_i$  ou  $a_i$  e  $b_i$ ,  $i = 1, 2$  são desconhecidos e dispendo-se de medidas da concentração máxima do componente  $B$  ( $C_{B_{max}}$ ) e do tempo de operação ( $t^*$ ) correspondente para determinados valores de temperatura de operação  $T$ , pode-se estabelecer o seguinte critério de mínimos quadrados

$$\min_{\theta} \frac{1}{2} \sum_{i=1}^N \{ (C_{B_{max}}(\theta, T^i) - \hat{C}_{B_{max}}^i)^2 + \lambda (t^*(\theta, T^i) - \hat{t}^{*i})^2 \} \quad (3.28)$$

onde

$N$  : número de medidas

$\theta$  : vetor de parâmetros

$\lambda$  : fator de ponderação

$\hat{C}_{B_{max}}^i$  :  $i$ -ésima medida de  $C_{B_{max}}$

$\hat{t}^{*i}$  :  $i$ -ésima medida de  $t^*$

$T^i$  :  $i$ -ésimo valor de temperatura

O problema de identificação de parâmetros pode, então, ser formulado como

$$\min_{\theta \in \Theta} G(\theta, T) \quad (3.29)$$

onde  $\Theta$  representa o conjunto de parâmetros factíveis do problema.

O desempenho econômico do reator é otimizado através de alterações no set-point da temperatura  $T$ . A condição ótima de operação, por sua vez, pode ser descrita de maneira simplificada como aquela que maximiza a concentração máxima de  $B$  e minimiza o tempo de operação.

A natureza conflitante dos dois objetivos apresentados torna necessária uma análise baseada em compromisso. Por exemplo, se  $E_1 > E_2$  a maximização de  $C_{B_{max}}$  consiste na manutenção da temperatura  $T$  no mínimo valor possível, o que implica na maximização de  $t^*$ . Assim, um pequeno aumento de  $C_{B_{max}}$  pode corresponder a um acréscimo considerável no  $t^*$ , reduzindo a eficiência do processo. Neste caso, a formulação bicritério é adequada e

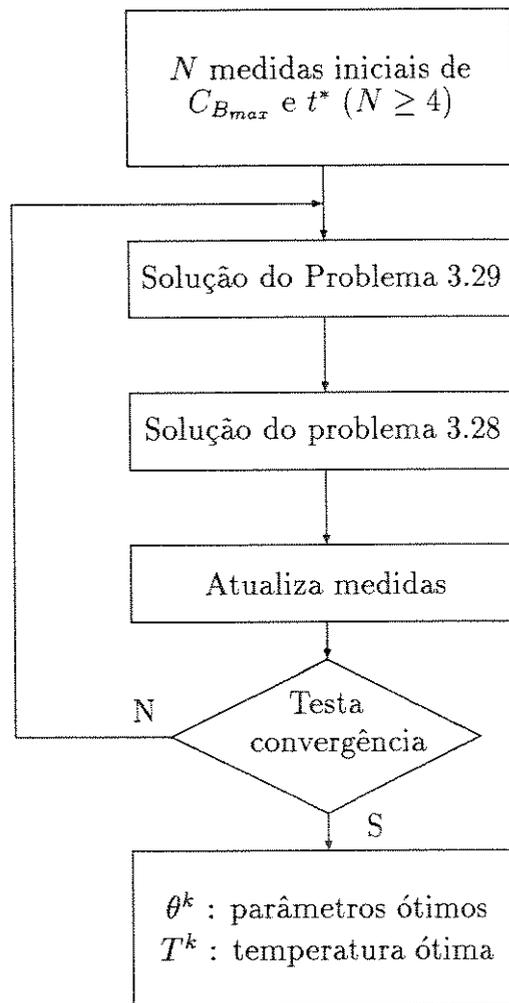


Figura 3.2: Formulação do Problema Conjunto para o Reator por Batelada

essencial para a determinação do ponto de operação do sistema. O problema de otimização pode, então, ser expresso por

$$\min_{T \in \mathcal{T}} \mathcal{V}(-C_{B_{max}}(\theta^k, T), t^*(\theta^k, T)) \quad (3.30)$$

onde  $\mathcal{T}$  é o intervalo de temperatura e  $\theta^k$  é a solução do problema (3.29).

Assim, o procedimento para a solução dos problemas de identificação de parâmetros e otimização de maneira conjunta pode ser representado pelo fluxograma da Figura 3.2.

### 3.7 Descrição do SRB

O Simulador do Reator por Batelada — SRB é o resultado da implementação efetiva do que foi discutido nos itens anteriores. Na Figura 3.3 tem-se a estrutura básica do SRB.

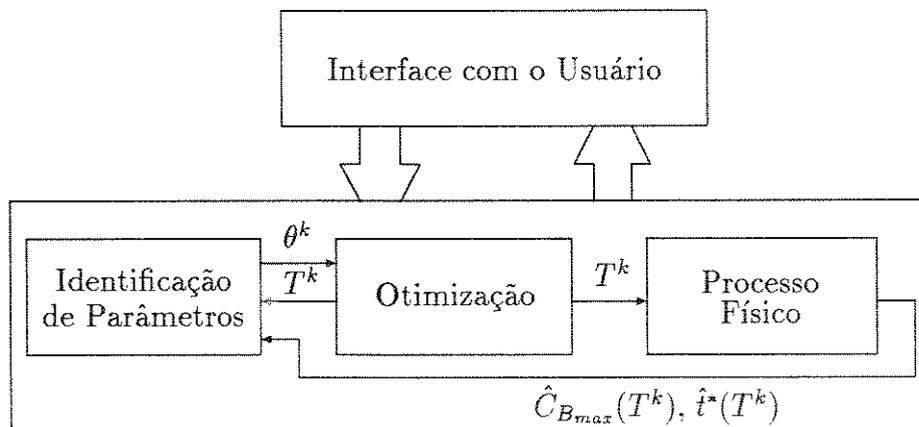


Figura 3.3: Diagrama de Blocos do SRB

O sistema foi implementado num microcomputador compatível com IBM-PC 386 com controladora de vídeo EGA, utilizando o compilador Turbo C 2.0. A seguir, são descritos os blocos componentes.

#### 3.7.1 Processo Físico

Este módulo é constituído por um conjunto de rotinas que simulam o reator por batelada e o sistema de controle utilizado. Como pode-se ver na

Figura 3.4, dada uma temperatura  $T^k$ , este bloco fornece as medidas  $\hat{C}_{B_{max}}$ ,  $\hat{t}^*$  e  $\hat{T}$  correspondentes.

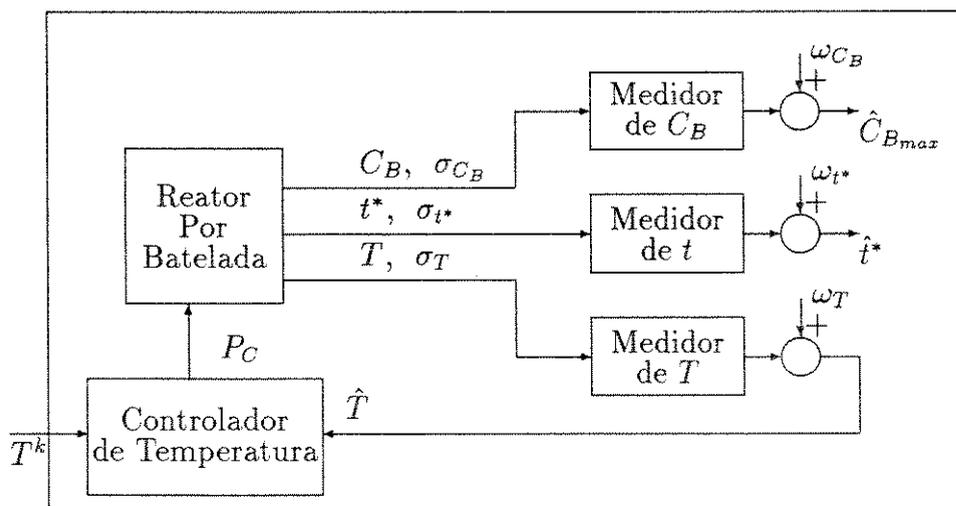


Figura 3.4: Diagrama de Blocos do Processo Físico

A simulação do reator baseia-se nas expressões do item 3.3, utilizando-se os parâmetros da Tabela 3.1.

Assim, dada a temperatura  $T^k$ , o processo físico é simulado, sendo monitorado o valor da concentração do componente  $B$  ( $C_B$ ). Quando o valor do  $C_B$  começa a decrescer, é detectado o máximo e, então, a simulação é finalizada. Os valores da concentração de  $B$  e do tempo de operação, neste ponto, são considerados como medidas de  $C_{B_{max}}$  e  $t^*$ .

As medidas obtidas podem ser contaminadas por ruído, o qual é obtido através de uma rotina de geração de ruído branco gaussiano com variância especificada pelo usuário (Figura 3.4).

Na simulação do reator, utiliza-se a integração de Euler para solução das equações diferenciais envolvidas. Durante o período de aquecimento da massa de reação, tem-se as alterações mais bruscas das variáveis do processo, sendo necessário um intervalo de integração muito pequeno. Quando a temperatura  $T$  se estabiliza, este intervalo pode ser aumentado. Os valores usados encontram-se na Tabela 3.2.

O controlador de temperatura simulado corresponde ao descrito em 3.4. O intervalo de integração utilizado obedece à Tabela 3.2. Como o controlador deve operar satisfatoriamente em toda a faixa possível de operação

$\alpha_1$	729.55 $\text{min}^{-1}$	$V_j$	18.83 $\text{ft}^3$
$\alpha_2$	6567.6 $\text{min}^{-1}$	$C_{vw}$	100 $\text{gpm}/\text{psi}^{1/2}$
$E_1$	15000 $\text{Btu}/\text{mole}$	$T_{j0}$	80° $F$
$E_2$	20000 $\text{Btu}/\text{mole}$	$A_i$	56.5 $\text{ft}^2$
$A_{vp}$	-8744.4° $R$	$\lambda_1$	-40000 $\text{Btu}/\text{mole}$
$B_{vp}$	15.70	$\lambda_2$	-50000 $\text{Btu}/\text{mole}$
$C_{A0}$	0.80 $\text{mole}/\text{ft}^3$	$C_p$	1 $\text{Btu}/\text{lb}_m - ^\circ F$
$T_0$	80° $F$	$V$	42.5 $\text{ft}^3$
$K_c$	10	$\rho$	50 $\text{lb}_m/\text{ft}^3$
$C_{vs}$	112 $\text{lb}_m/\text{min} - \text{psi}^{1/2}$	$C_M$	0.12 $\text{Btu}/\text{lb}_m - ^\circ F$
$h_{os}$	1000 $\text{Btu}/\text{hr} - ^\circ F - \text{ft}^2$	$V_M$	9.42 $\text{ft}^3$
$h_{ow}$	400 $\text{Btu}/\text{hr} - ^\circ F - \text{ft}^2$	$\rho_M$	512 $\text{lb}_m/\text{ft}^3$
$h_i$	160 $\text{Btu}/\text{hr} - ^\circ F - \text{ft}^2$	$\rho_J$	62.3 $\text{lb}_m/\text{ft}^3$
$A_o$	56.5 $\text{ft}^2$	$C_J$	1 $\text{Btu}/\text{lb}_m - ^\circ F$
$H_s - h_c$	939 $\text{Btu}/\text{lb}_m$		

Tabela 3.1: Parâmetros do Reator por Batelada

$\Delta t (T \leq T^k)$ (min)	0.002
$\Delta t (T > T^k)$ (min)	0.05

Tabela 3.2: Intervalo de Integração

[ $100^\circ F, 200^\circ F$ ] ( $[37^\circ C, 93^\circ C]$ ), optou-se por uma estratégia adaptativa, com uma tabela de escalonamento de ganhos, os quais foram obtidos utilizando-se o método do limite de estabilidade de Ziegler–Nichols (Luyben, 1973). Na Figura 3.5 é apresentado um gráfico da evolução do ganho e dos tempos derivativo e integral com a temperatura. Utilizou-se um intervalo de amostragem de 3 minutos para o controlador digital.

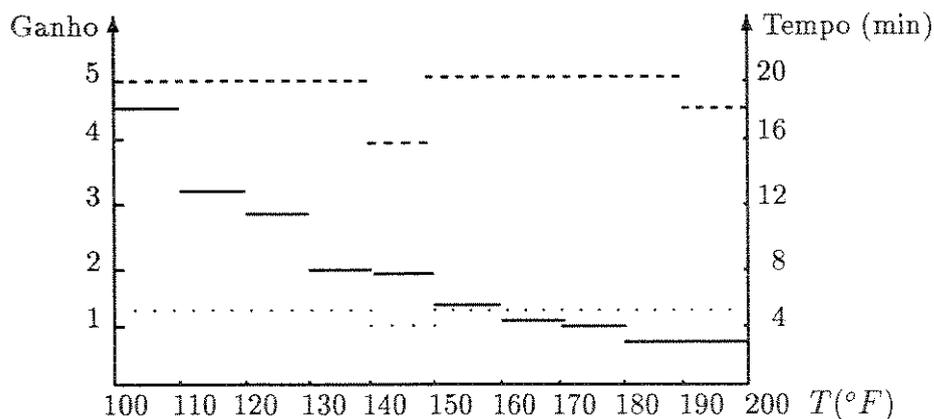


Figura 3.5: "Gain Schedule" : —  $k_p$ ; - - -  $\tau_i$ ;  $\cdots$   $\tau_d$

Como exemplo do desempenho do módulo do Processo Físico, são apresentadas, nas Figuras 3.6, 3.7, 3.8 e 3.9, as curvas da evolução das variáveis resultantes da simulação do reator na temperatura de  $150^\circ F$  ( $65.56^\circ C$ ).

### 3.7.2 Identificação de Parâmetros

O problema de identificação de parâmetros definido genericamente por (3.29) enquadra-se na classe de problemas que podem ser resolvidos através do Método do Gradiente Modificado proposto por Geromel e Ferreira (1986) e apresentado no Capítulo 2.

No SRB, a determinação do tamanho do passo  $\beta \in R_+$  é feita através do algoritmo de Fibonacci.

Pelos valores apresentados na Tabela 3.1, pode-se observar que os parâmetros  $\alpha_i$  e  $E_i$  ( $i = 1, 2$ ) do modelo não linear possuem magnitudes muito diferentes, o que dificulta o estabelecimento de critérios de convergência. Supondo que valores típicos de todas as variáveis são conhecidos, pode-se

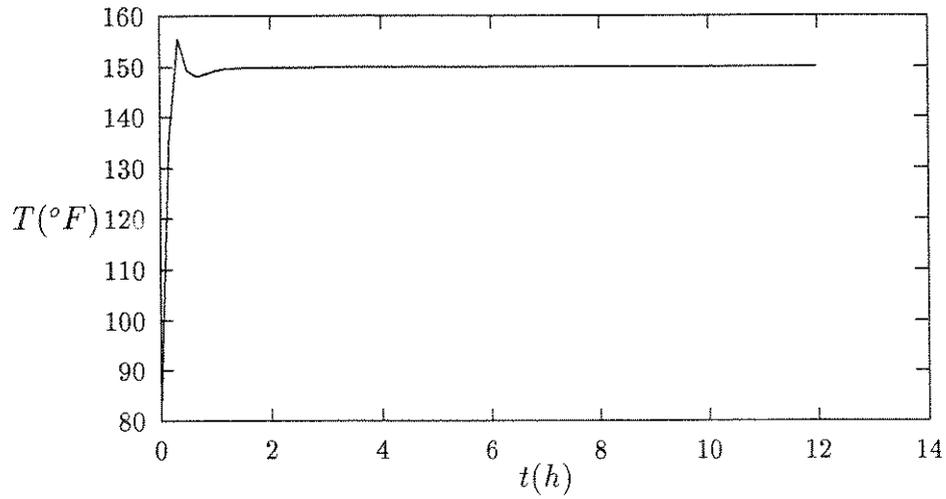


Figura 3.6: Temperatura da Reação

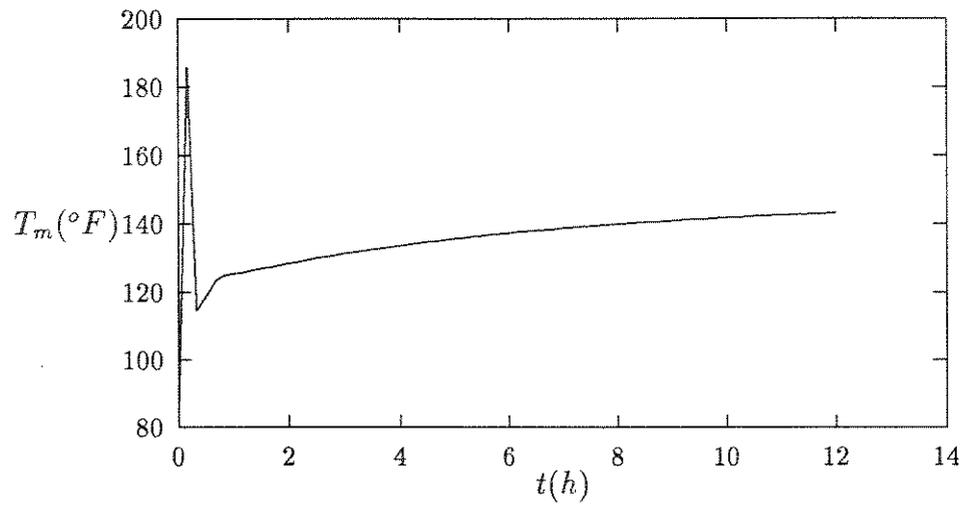


Figura 3.7: Temperatura da Parede de Metal

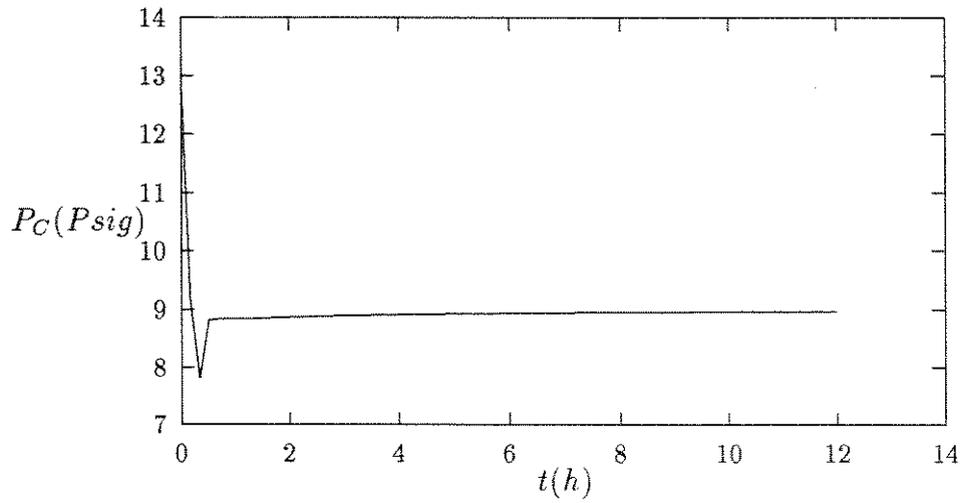


Figura 3.8: Pressão do Controlador

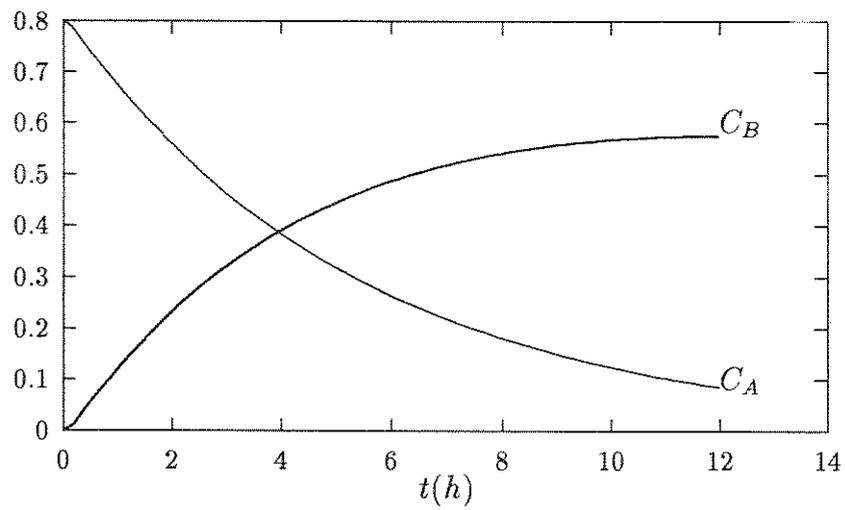


Figura 3.9: Concentrações dos Componentes A e B

transformar o problema de forma que todas as variáveis tenham a mesma ordem de grandeza.

No caso, um escalonamento é feito de forma que a região factível de todos os parâmetros seja  $[-1, 1]$ , o que é possível a partir da seguinte transformação

$$\theta^e = D\theta + c \quad (3.31)$$

onde

$\theta^e$  : vetor de parâmetros escalados

$D$  : matriz diagonal

$$\{D_{jj}\} = \frac{1}{2}(\bar{\theta}_j - \underline{\theta}_j); \quad j = 1, \dots, 4$$

$c$  : vetor

$$c_j = \frac{1}{2}(\bar{\theta}_j + \underline{\theta}_j); \quad j = 1, \dots, 4$$

$\bar{\theta}_j, \underline{\theta}_j$  : valores máximo e mínimo de  $\theta_j$ ,  $j = 1, \dots, 4$

As derivadas dos problemas original e transformado relacionam-se por

$$\nabla G(\theta^e) = D\nabla G(\theta) \quad (3.32)$$

Deve-se ressaltar que o intervalo  $[\bar{\theta}_j, \underline{\theta}_j]$  deve ser realista. Por simplicidade, no SRB este intervalo é dado por uma variação percentual em torno dos valores nominais dos parâmetros (Tabela 3.1). Assim

$$\underline{\theta} = \theta^*(1 - \Delta\%) \quad (3.33)$$

$$\bar{\theta} = \theta^*(1 + \Delta\%)$$

onde  $\Delta\%$  : variação percentual.

Considerando os modelos linear e não linear apresentados no item 3.5, pode-se particularizar o problema (3.29) de acordo com a formulação. Assim, para o caso não linear, tem-se

$$\min_{\underline{\theta} \leq \theta \leq \bar{\theta}} G(\theta, T) \quad (3.34)$$

onde  $\theta = [\alpha_1, \alpha_2, E_1, E_2]$ .

Como existem apenas restrições de desigualdade,  $P = I$  e  $s(\theta) = \nabla_m G(\theta)$ .

Utilizando-se o modelo linear, pode-se formular o problema de identificação como

$$\begin{aligned} \min_{\underline{\theta} \leq \theta \leq \bar{\theta}} G(\theta, T) \\ k_i(\theta, \underline{T}) \geq 0 \end{aligned} \quad (3.35)$$

Tem-se, então, restrições de bornes e restrições de desigualdade lineares. Estas últimas devem ser transformadas em restrições de igualdade para que o problema (3.35) se adeque ao formato do problema genérico resolvido pelo método do Gradiente Modificado. Assim, acrescentando-se as variáveis de excesso, obtém-se

$$a_i \underline{T} + b_i - v_i = 0 \quad (3.36)$$

$$v_i \geq 0, \quad i = 1, 2$$

E, portanto, o problema (3.35) pode ser reescrito como

$$\begin{aligned} \min G(\theta, T) \\ \underline{\theta} \leq \theta \leq \bar{\theta} \\ A\theta = 0 \end{aligned} \quad (3.37)$$

onde

$$\theta = [a_1 \quad a_2 \quad b_1 \quad b_2 \quad v_1 \quad v_2]$$

$$A = \begin{bmatrix} \underline{T} & 0 & 1 & 0 & -1 & 0 \\ 0 & \underline{T} & 0 & 1 & 0 & -1 \end{bmatrix}$$

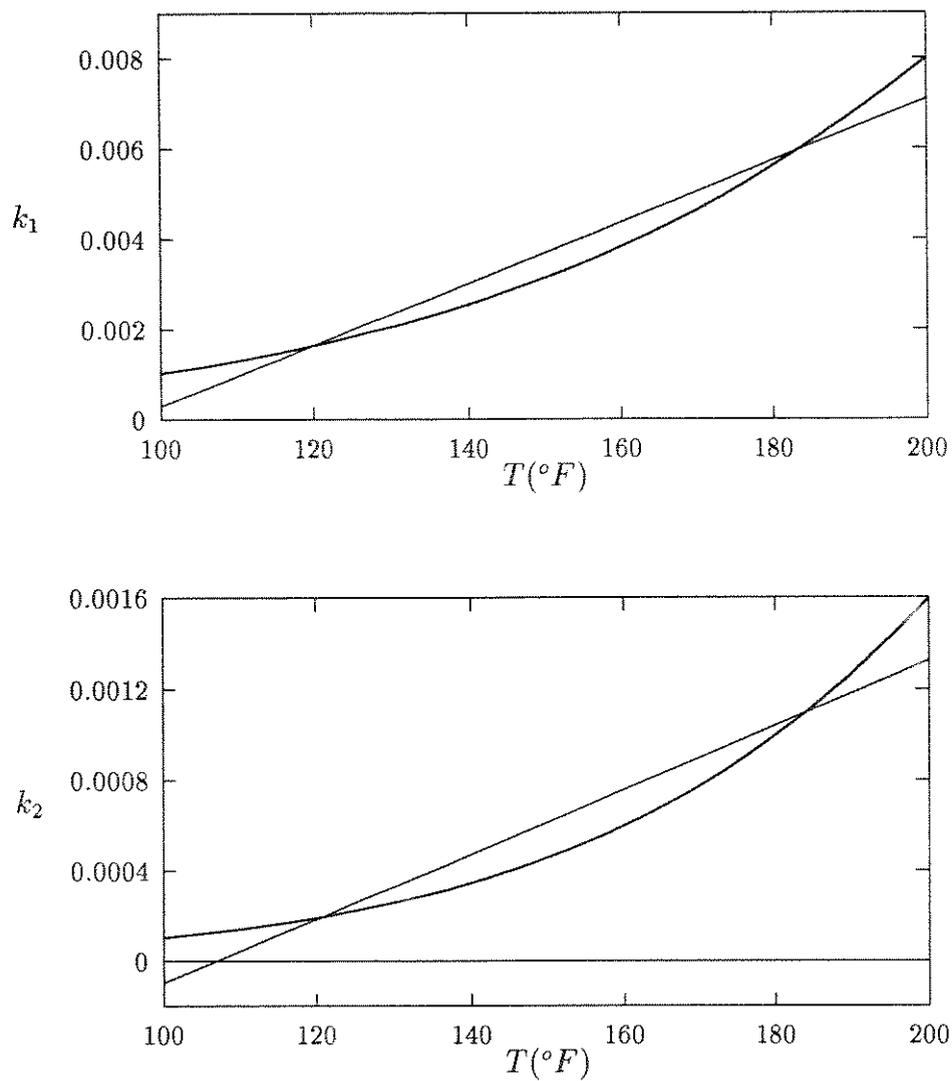
Considerando-se os valores da Tabela 3.1, pode-se obter curvas  $k_1(T)$  e  $k_2(T)$  calculados pela equação de Arrhenius para  $T \in [100^\circ F, 200^\circ F]$  e para os valores nominais dos parâmetros,  $\theta^*$ . Valores *nominais* dos parâmetros  $a_i$  e  $b_i$  ( $i = 1, 2$ ) podem ser calculados pela linearização através de mínimos quadrados das curvas citadas (Figura 3.10). A razão para o cálculo destes valores é definir uma faixa de variação de parâmetros para o problema de identificação com taxas lineares.

Para evitar problemas devido a magnitude dos parâmetros  $a_i$  ( $i = 1, 2$ ), em geral muito pequenos, pode-se utilizar a seguinte transformação

$$a_i^e = 10^2 a_i, \quad i = 1, 2 \quad (3.38)$$

o que implica em

$$k_i(\underline{T}) = a_i^e 10^{-2} \underline{T} + b_i, \quad i = 1, 2 \quad (3.39)$$

Figura 3.10: Taxas de reação  $k_1$  e  $k_2$ : modelos linear e não linear

Neste caso, também é usado o escalamento descrito para o modelo não linear, apesar da transformação acima aproximar as magnitudes dos parâmetros  $a_i$  e  $b_i$  ( $i = 1, 2$ ). Os valores limitantes dos parâmetros são definidos de maneira análoga a (3.34).

Com isto, passa-se a ter no problema (3.37)

$$\begin{aligned} \theta &= [a_1^e \ a_2^e \ b_1 \ b_2 \ v_1 \ v_2]' \\ A &= \begin{bmatrix} 10^{-2}\underline{T} & 0 & 1 & 0 & -1 & 0 \\ 0 & 10^{-2}\underline{T} & 0 & 1 & 0 & -1 \end{bmatrix} \end{aligned} \quad (3.40)$$

Pode-se, então, calcular a matriz projeção  $P$ . Para  $\underline{T} = 100^\circ F$

$$P = \begin{bmatrix} \frac{2}{3} & 0 & -\frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{2}{3} & 0 & -\frac{1}{3} & 0 & \frac{1}{3} \\ -\frac{1}{3} & 0 & \frac{2}{3} & 0 & \frac{1}{3} & 0 \\ 0 & -\frac{1}{3} & 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{2}{3} \end{bmatrix} \quad (3.41)$$

A matriz  $P$  é fixa e os elementos da diagonal principal são diferentes de zero. Para obter-se a direção  $s(\theta) = P\nabla_m G(\theta)$  é necessário apenas calcular o gradiente modificado.

No SRB, estão implementadas as expressões analíticas das derivadas parciais utilizadas no subproblema de identificação, obtidas a partir de simples, mas demorado trabalho de cálculo. Por simplicidade, estas expressões serão omitidas nesta exposição.

### 3.7.3 Otimização

O problema de otimização (3.30) pressupõe uma função valor não decrescente com as funções objetivo envolvidas. No SRB, dispõe-se de duas estratégias de otimização. A primeira delas utiliza uma função valor pré-definida do tipo

$$\begin{aligned} \mathcal{V}(-C_{B_{max}}(\theta, T), t^*(\theta, T)) &= (-C_{B_{max}}(\theta, T) + \underline{C}_{B_{max}})^2 \\ &+ \gamma(t^*(\theta, T) - \underline{t}^*)^2 \end{aligned} \quad (3.42)$$

onde  $C_{B_{max}} = \min_{T \in \mathcal{T}} -C_{B_{max}}(\theta, T)$ ,  $t^* = \min_{T \in \mathcal{T}} t^*(\theta, T)$ ,  $T \in [\underline{T}, \bar{T}]$  e  $\gamma$  é o fator de ponderação que define a relação de compromisso entre a maximização de  $C_{B_{max}}$  e a minimização de  $t^*$ . Variando-se o valor de  $\gamma$ , percorre-se a curva de soluções eficientes do problema, como será visto no item 3.8.

O problema (3.30), com a função valor definida em (3.42) é unidimensional e é resolvido no SRB através do algoritmo de Fibonacci.

A segunda estratégia implementada envolve a otimização interativa. Neste caso, a curva de trade-off para  $T \in [100^\circ F, 200^\circ F]$  é traçada considerando-se os últimos valores dos parâmetros obtidos na identificação. O usuário, então, pode selecionar um ponto desta curva, como será descrito em 3.7.4, o que equivale à escolha do set-point de temperatura que melhor reflete as preferências do usuário (Decisor) em relação aos objetivos do problema.

### 3.7.4 Interface com o Usuário

Neste item, são apresentados aspectos da interface com o usuário do SRB. O sistema é estruturado em menus sucessivos e janelas para entrada de dados. Além disto, os resultados das simulações são fornecidos em tabelas e gráficos para facilitar a compreensão dos mesmos. Na Tabela 3.3 são apresentadas as teclas e suas funções dentro dos ambientes menu ou janela.

Tecla	Função
↑	movimento para a função anterior no menu
↓	movimento para a função posterior no menu
< ENTER >	confirma opção no menu passa para o próximo item na janela de entrada de dados
< ESC >	retorno ao menu anterior

Tabela 3.3: Teclas de Função – Ambientes Menu e Janela

O SRB é ativado através do comando SRB. Inicialmente, é apresentado o menu principal reproduzido na Figura 3.11.

A opção *Modelo do Reator* possibilita a escolha do modelo do reator a ser usado. Quando selecionada, ativa um submenu como mostrado na Figura 3.12.

A segunda opção do menu principal, *Alteração das Condições Iniciais*, ativa o submenu que pode ser visto na Figura 3.13.

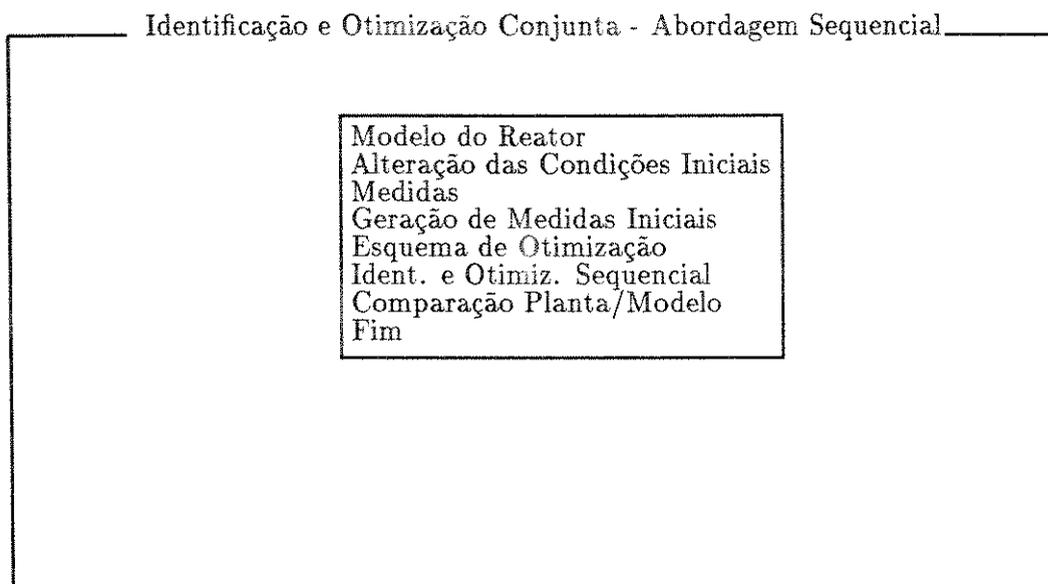


Figura 3.11: Menu Principal do SRB

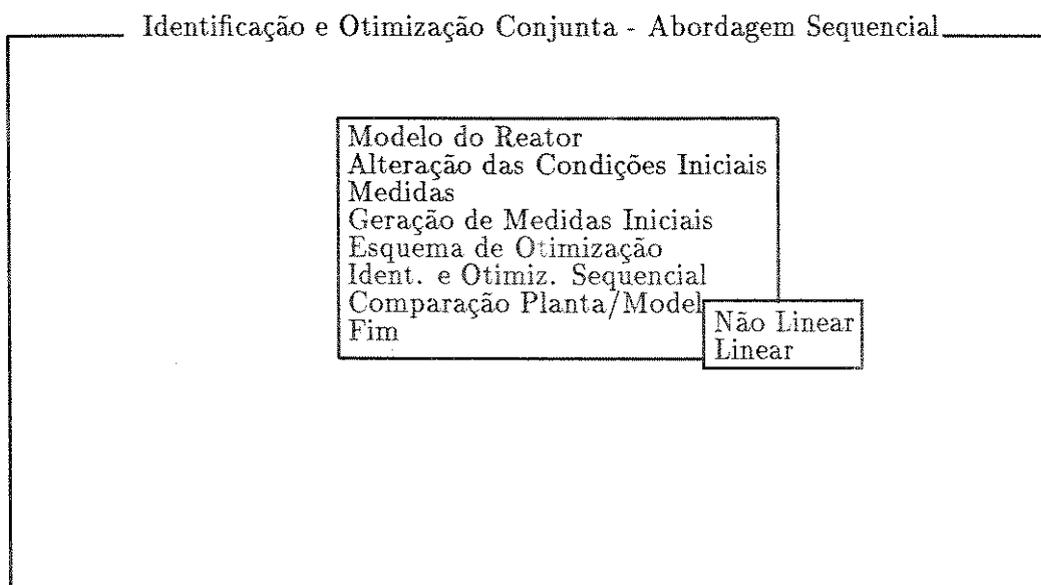


Figura 3.12: Menu de Seleção do Modelo do Reator

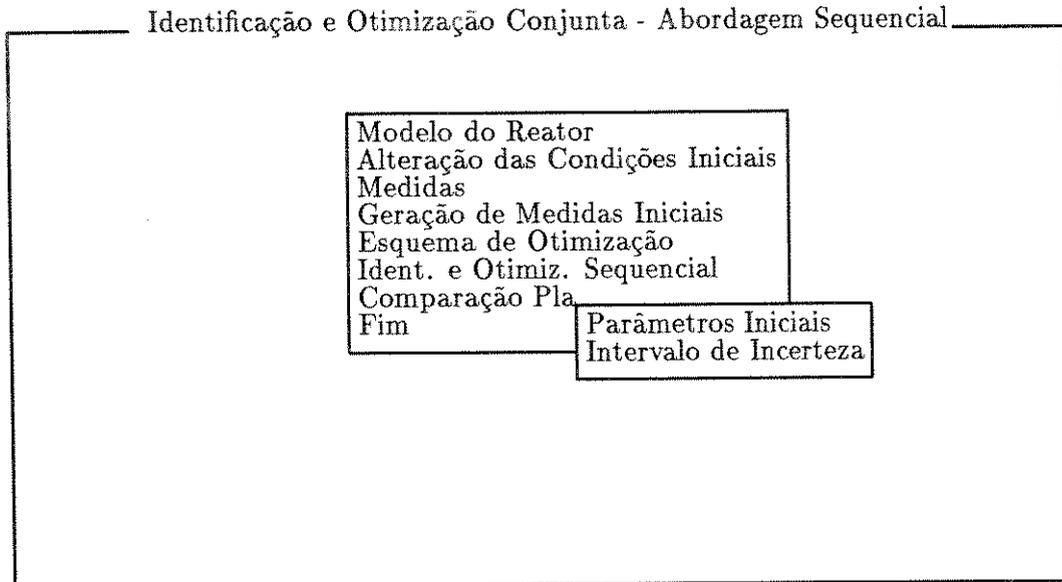


Figura 3.13: Menu de Alteração das Condições Iniciais

A opção *Parâmetros Iniciais* permite a especificação do vetor de parâmetros iniciais a serem usados na identificação. Os valores dos parâmetros devem ser factíveis. A janela para entrada dos valores varia com o modelo como pode-se ver na Figura 3.14.

Ainda no menu da Figura 3.13, a opção *Intervalo de Incerteza* abre uma janela para alteração da variação percentual em torno dos valores nominais para estabelecimento das restrições de borne (Figura 3.15).

As características das medidas utilizadas na solução do problema de identificação podem ser alteradas através da opção *Medidas*. A Figura 3.16 mostra o submenu de seleção.

O número de medidas é alterado através da janela apresentada na Figura 3.17.

A opção *Ruído* apresenta um submenu para especificação da presença de ruído nas medidas (Figura 3.18). A escolha da opção *Com Ruído* abre uma janela para alteração dos desvios-padrão para cada uma das variáveis medidas (Figura 3.19).

As medidas com as características definidas são geradas através da simulação do processo quando da seleção da opção *Geração de Medidas Iniciais*.

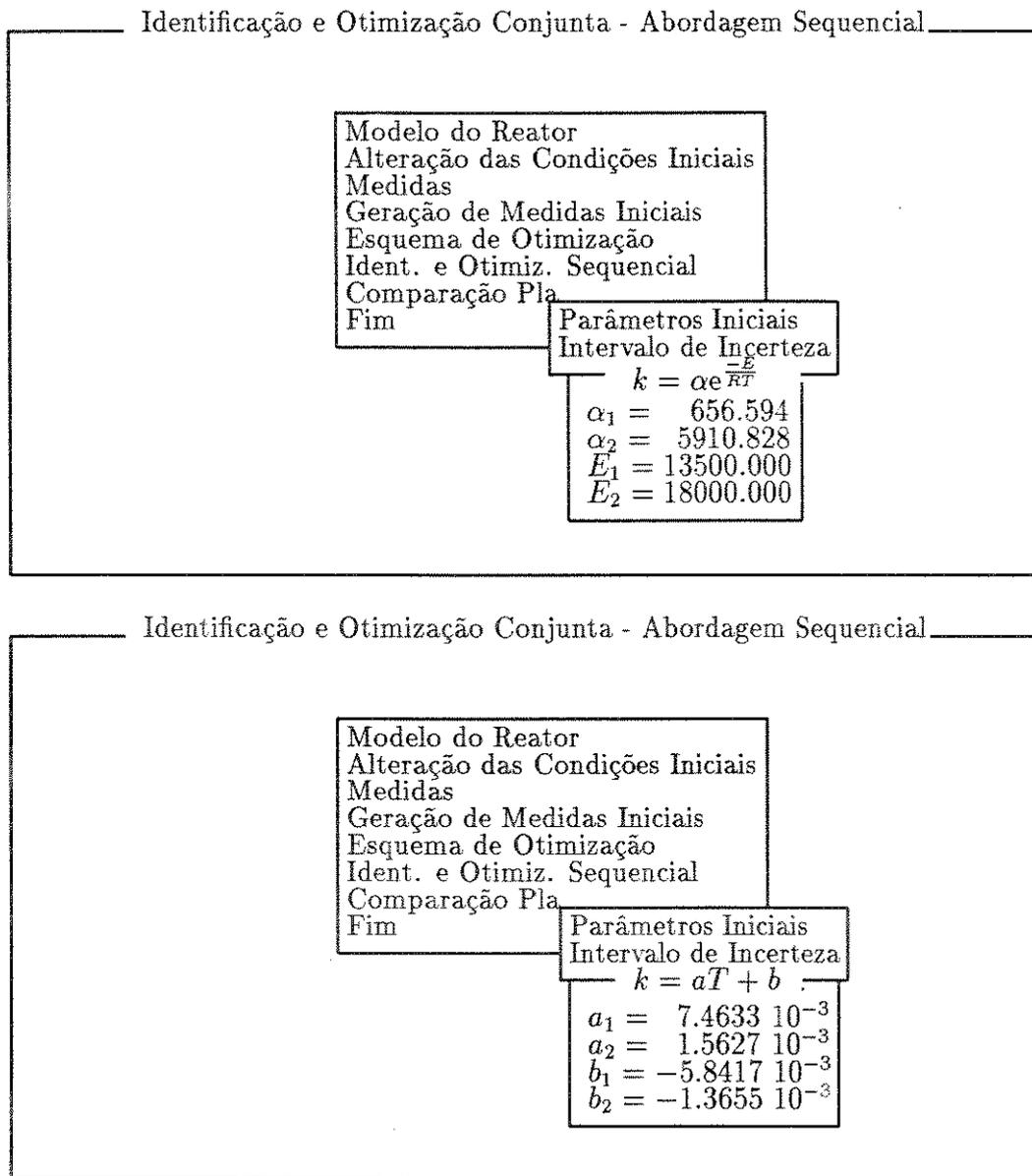


Figura 3.14: Janela para alteração dos valores iniciais dos parâmetros

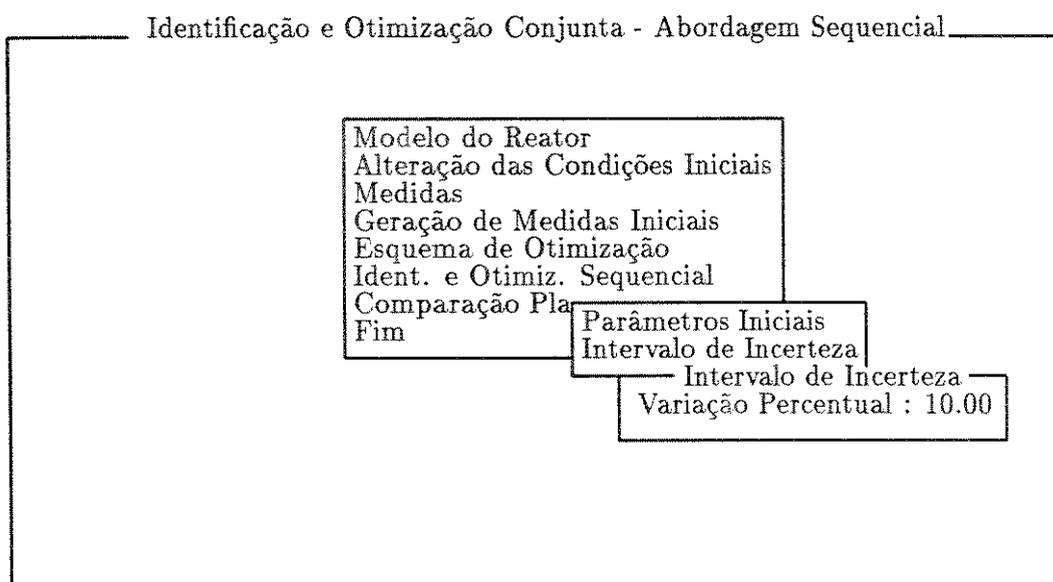


Figura 3.15: Janela para alteração do intervalo de incerteza dos parâmetros

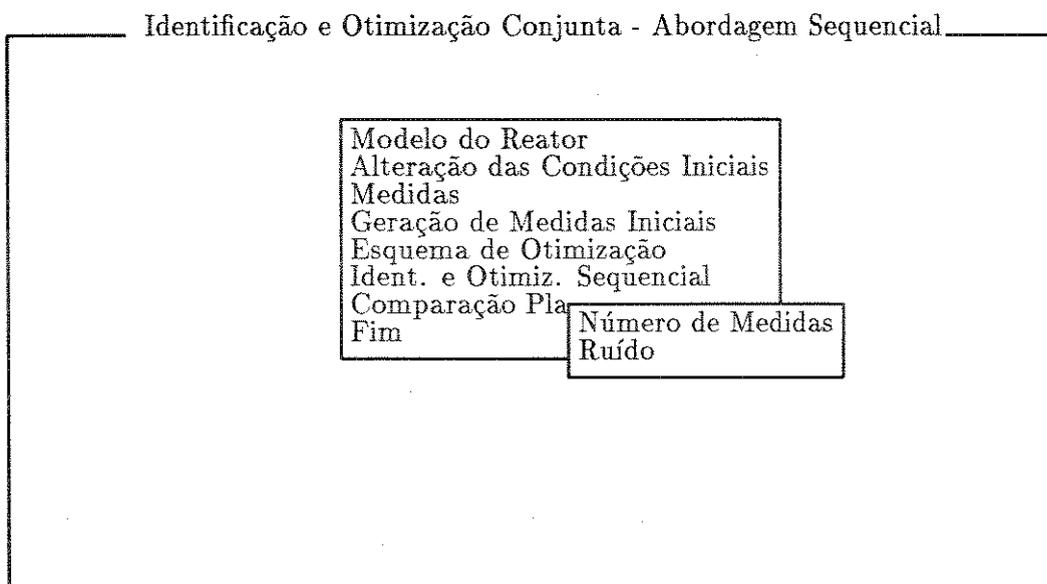


Figura 3.16: Menu de Alteração das Características das Medidas

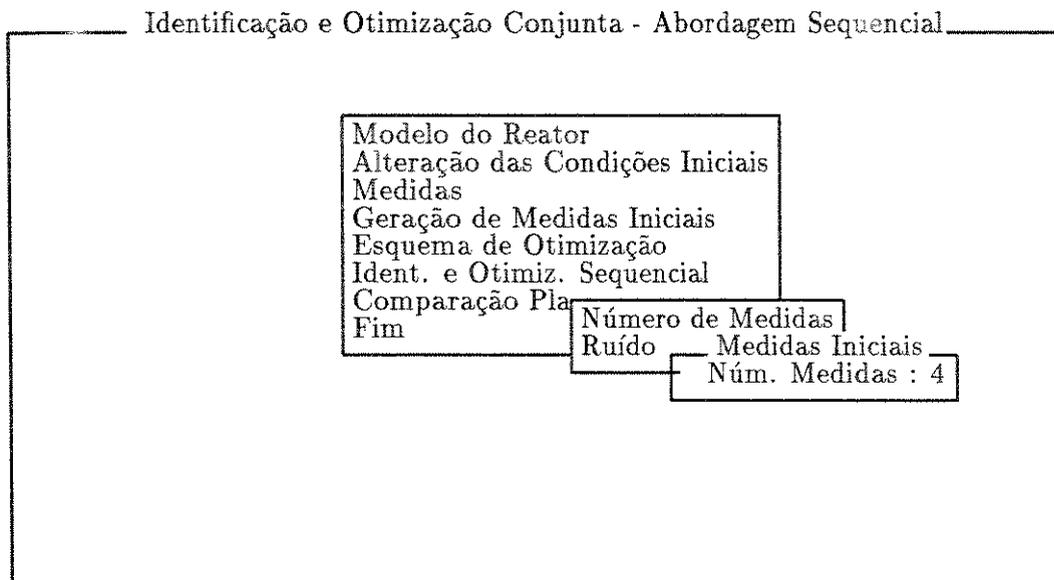


Figura 3.17: Janela de Alteração do Número de Medidas

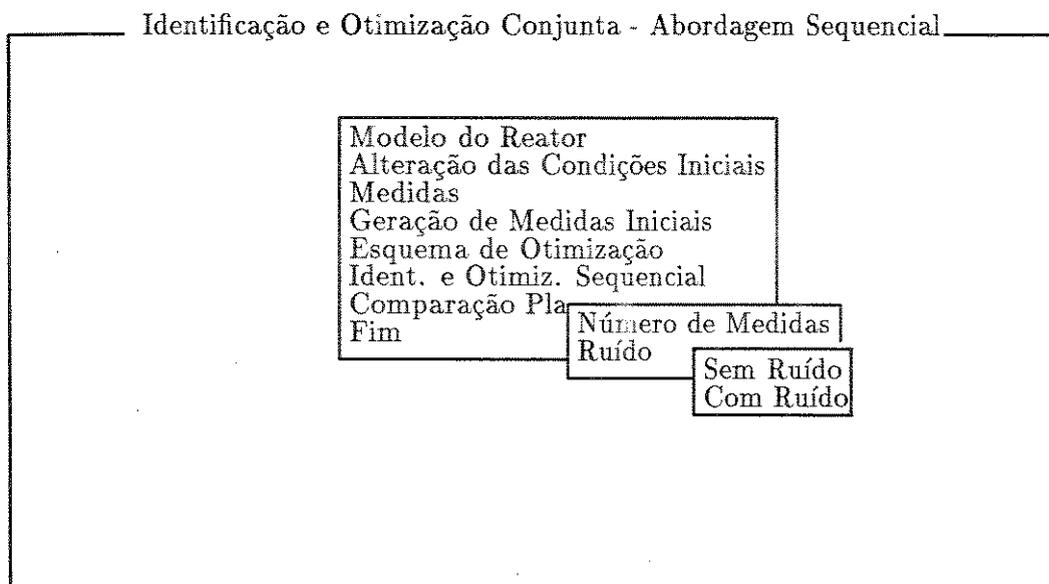


Figura 3.18: Menu para Escolha da Presença ou não de Ruído nas Medidas

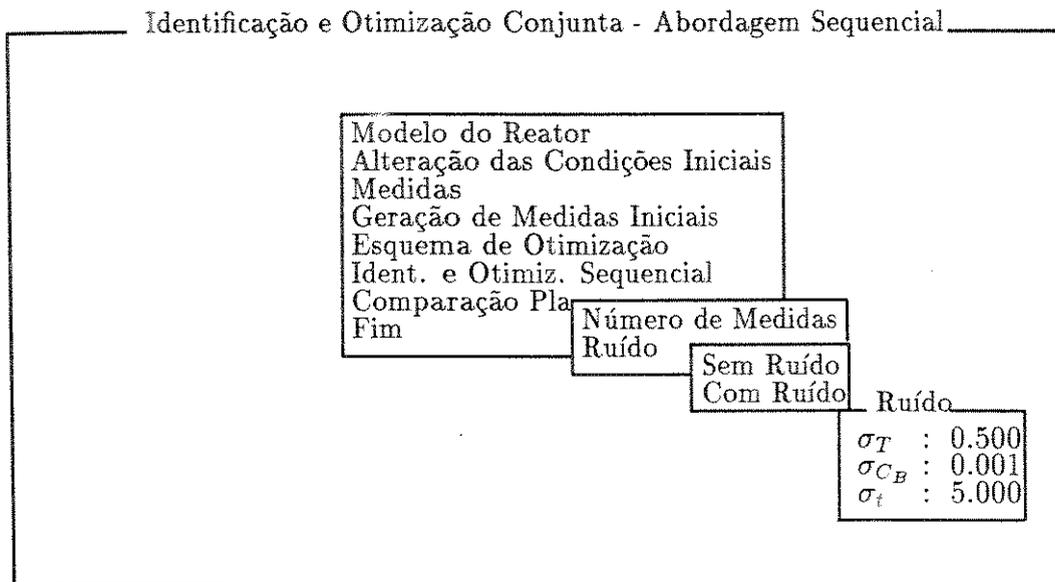


Figura 3.19: Janela para alteração dos desvios-padrão

A faixa de temperatura é dividida em intervalos iguais de acordo com o número de medidas especificado, sendo as medidas feitas nas temperaturas limites destes intervalos. Nas Figuras 3.20 e 3.21, tem-se as telas com quatro medidas sem ruído e com ruído, respectivamente.

As curvas da evolução das variáveis  $C_B$  e  $T$  são traçadas durante a simulação, a medida que os valores são gerados. A tabela com os dados gerados é editada após a finalização da tarefa.

As próximas opções do menu principal referem-se a aspectos do problema conjunto de identificação e otimização. A opção *Esquema de Otimização* fornece o submenu para seleção, como pode ser visto na Figura 3.22. A escolha da opção do submenu *Função Pré-definida* implica na abertura de uma janela para alteração do fator de ponderação  $\gamma$  da equação 3.42 (Figura 3.23).

A opção *Identificação e Otimização Sequencial* ativa o procedimento sequencial de resolução do problema conjunto. A cada iteração são fornecidos ao usuário resultados parciais. A partir das medidas e parâmetros iniciais é calculado um conjunto de parâmetros, com o qual traça-se a curva de pontos eficientes do problema de otimização para  $T \in [100^\circ F, 200^\circ F]$ . No caso

Identificação e Otimização Conjunta - Abordagem Sequencial		
$T(^{\circ}C)$	$t^*(h)$	$C_{B_{max}}$
37.78	40.71	0.6180
56.36	17.66	0.5894
74.80	8.37	0.5617
93.28	4.20	0.5347

Tecla Algo para Retornar ao Menu

Figura 3.20: Tabela de medidas sem ruído

Identificação e Otimização Conjunta - Abordagem Sequencial		
$T(^{\circ}C)$	$t^*(h)$	$C_{B_{max}}$
37.57	40.78	0.6245
55.91	17.69	0.5894
74.67	8.39	0.5680
93.34	4.19	0.5164

Tecla Algo para Retornar ao Menu

Figura 3.21: Tabela de medidas com ruído

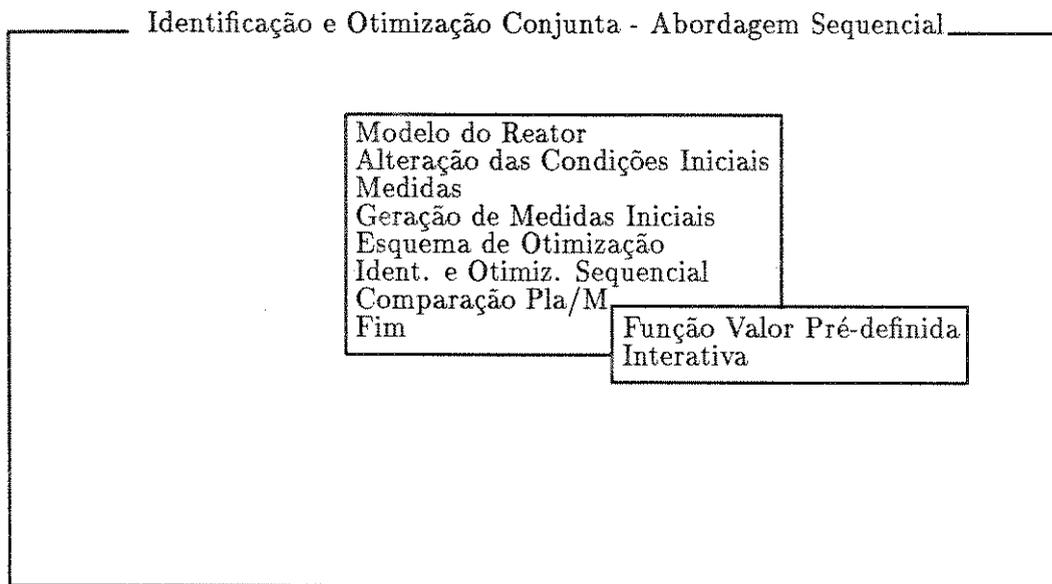


Figura 3.22: Menu de seleção do esquema de otimização

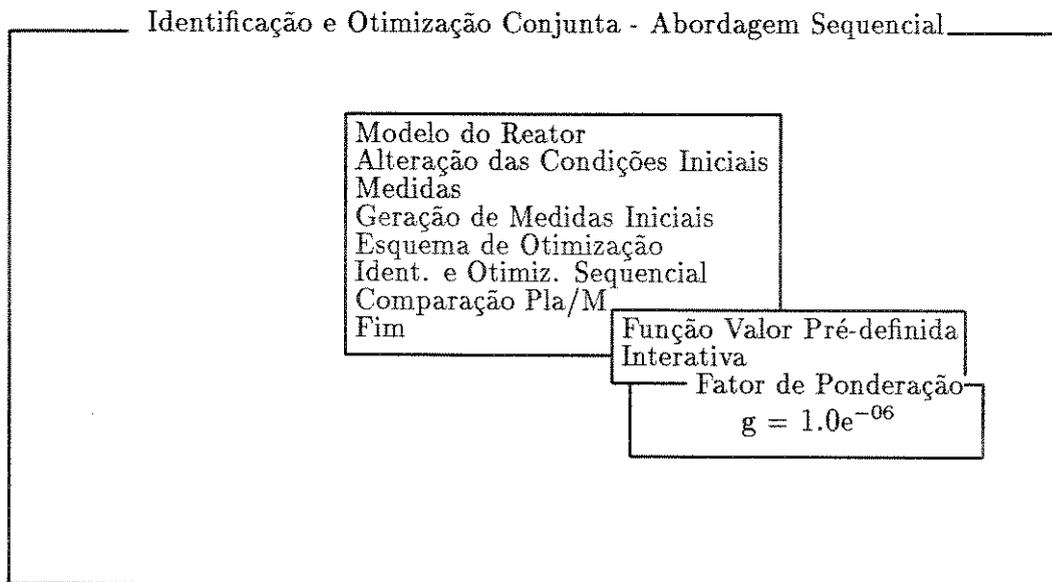


Figura 3.23: Janela de alteração do fator de ponderação

do esquema de otimização iterativo, deve-se selecionar o ponto de operação desejado, posicionando o símbolo "+" com auxílio das teclas  $\uparrow \downarrow$ . A velocidade de movimentação do "+" pode ser alterada, pressionando-se a barra de espaço do teclado. O valor do set-point da temperatura obtido com a otimização é aplicado ao simulador da planta real, gerando novas medidas de  $C_{B_{max}}$  e  $t^*$  que substituem as mais antigas.

Como pode-se ver na Figura 3.24, é indicada também a variação máxima percentual dos parâmetros ( $E\%$ ) entre iterações sucessivas. Esta informação pode ser usada para caracterizar a convergência do procedimento conjunto.

Identificação e Otimização Conjunta - Abordagem Sequencial

$\alpha_1$	$\alpha_2$	$E_1$	$E_2$	$T(^{\circ}C)$	$t^*(h)$	$\hat{t}^*(h)$	$C_{B_{max}}$	$\hat{C}_{B_{max}}$
656.616	5957.769	14880.066	19889.299	37.78	40.70	40.71	0.6180	0.6180
656.616	5957.761	14880.573	19889.686	37.78	40.72	40.71	0.6180	0.6180

E%:  $3e^{-5}$  <ENTER> Continua <ESC> Menu

Figura 3.24: Tabela de resultados parciais

O ciclo é repetido até que o usuário considere que os parâmetros convergiram, sendo que a cada interação são usados o vetor de parâmetros e o conjunto de medidas atualizados na iteração anterior.

A próxima opção no menu principal, *Comparação Planta/Modelo*, abre o submenu mostrado na Figura 3.25.

A primeira opção do submenu, *Curva de Comparação*, mostra um gráfico comparativo entre as curvas de trade-off da planta real e do modelo, utilizando o vetor de parâmetros obtido na identificação e otimização sequencial.

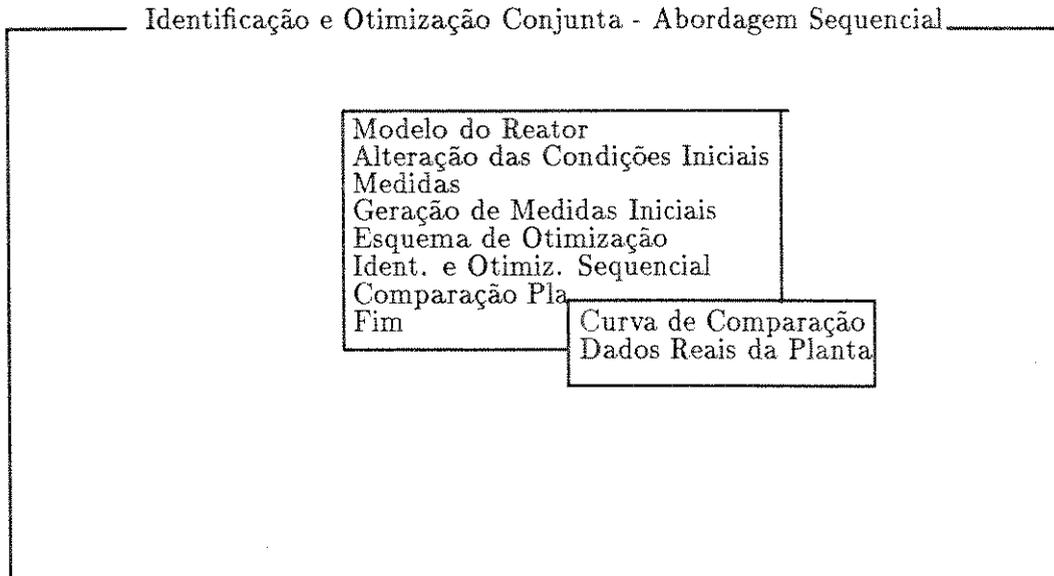


Figura 3.25: Menu de comparação modelo/planta

Esta opção serve apenas para verificar o desempenho do procedimento proposto em relação ao que poderia ser obtido com a planta real. O ponto ótimo de operação é marcado com "+".

A segunda opção do menu mostrado na Figura 3.25, *Dados Reais da Planta*, ativa a geração dos dados da planta necessários para traçar a curva de trade-off. Esta geração de dados é análoga à geração de medidas iniciais.

A última opção do menu principal, *Fim*, é para finalização do SRB.

### 3.8 Resultados Computacionais

O SRB descrito na seção anterior foi utilizado para simulação de diversos casos, fornecendo resultados para avaliação da abordagem proposta ao reator por batelada. Nos casos documentados a seguir utilizou-se uma faixa de incerteza de 10% em relação aos valores nominais ( $\Delta\% = 10\%$ ) e trabalhou-se sempre com as quatro medidas mais recentes de  $C_{B_{max}}$  e  $t^*$ . Para comparação, a Figura 3.26 mostra o comportamento da planta real e do modelo não linear em termos de  $C_{B_{max}}$  e  $t^*$ .

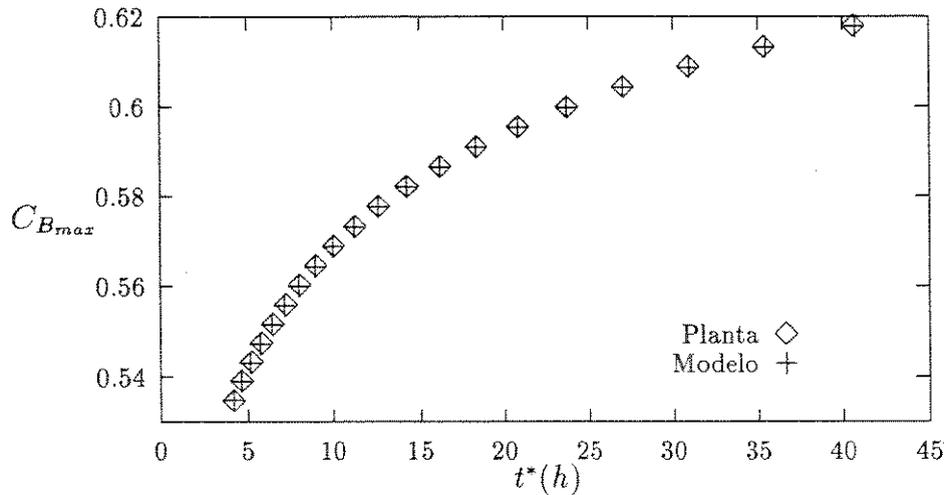


Figura 3.26: Comparação planta/modelo não linear

Pode-se observar que o modelo não linear é uma representação precisa do processo, o que possibilita o uso eficiente do algoritmo dois passos generalizado. Neste ponto, é bom lembrar que a diferença básica entre o modelo não linear e o reator simulado é o comportamento transitório da temperatura da reação. Observando-se a Figura 3.6 e considerando que no modelo não linear a temperatura da reação é constante, os maiores desvios ocorrem no início da operação, quando a concentração do componente  $B$  ainda é muito baixa. Além disso, o esquema de controle adaptativo garante o acompanhamento da referência gerada pelo modelo com um rápido tempo de subida, pouco overshoot e sem erro de regime devido a ação integral do controlador.

Considerando-se a função valor expressa em (3.42) e  $\gamma = 10^{-6}$ , o procedimento converge apenas em duas iterações para a solução correta do problema. Este comportamento é explicado pela não dependência dos parâmetros com relação ao controle no caso do modelo não linear, associada a um eficiente sistema de controle.

As soluções utópicas  $\bar{C}_{B_{max}}$  e  $\bar{t}^*$  convergem para a solução utópica do problema original,  $\bar{C}_{B_{max}} = 0.62$  e  $\bar{t}^* = 4.20h$  e, além disso, a solução final ( $T = 123.60^\circ F$ ,  $C_{B_{max}} = 0.5977$ ,  $t^* = 22.34h$ ), indicada com "\*" na Figura 3.27, é um ponto eficiente de operação da planta.

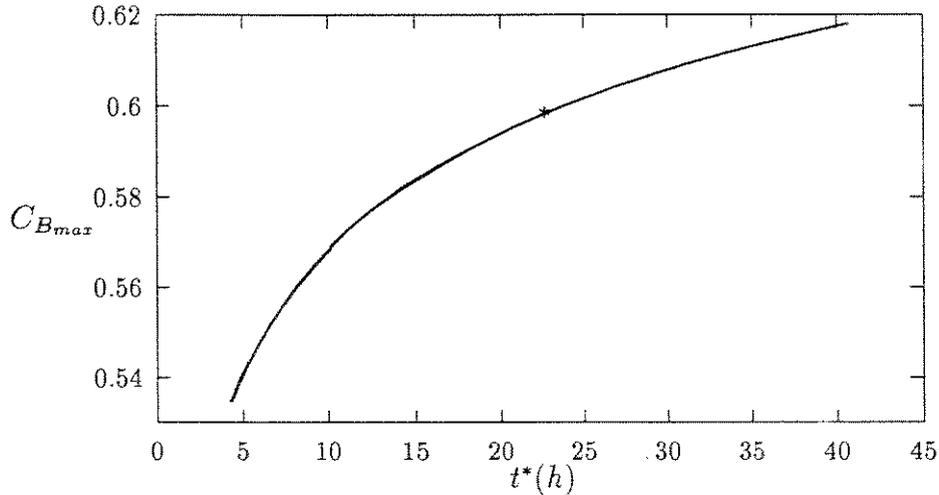


Figura 3.27: O procedimento dois estágios converge para  $\Gamma$  ("\*" : solução final)

Para aumentar as diferenças entre o modelo não linear e a planta, as medidas ( $\hat{T}, \hat{C}_{B_{max}}, \hat{t}^*$ ) foram contaminadas com ruído branco gaussiano. Os desvios padrão utilizados foram, respectivamente,  $\sigma_T = 0.5^\circ C$ ,  $\sigma_{C_{B_{max}}} = 0.001$  e  $\sigma_{t^*} = 5 \text{ min}$ . Após seis iterações, a variação percentual máxima dos parâmetros foi de  $10^{-5}$ .

Na Figura 3.28 tem-se as curvas correspondentes às iterações do algoritmo, ou seja, os conjuntos  $\theta^k$ -eficientes,  $\Gamma_{\theta^k}$ . O conjunto eficiente  $\Gamma_{\theta^k}^*$ , obtido com as medidas ruidosas é comparado ao conjunto eficiente real  $\Gamma$  na Figura 3.29. Também neste caso, a solução final, indicada com "\*", é uma aproximação muito boa de um ponto eficiente de operação da planta real.

O modelo linear do reator por batelada foi usado para simular o caso de parâmetros dependentes do controle. Inicialmente, apenas a parcela referente à minimização do tempo de operação da expressão (3.42) foi considerada. A Figura 3.30 compara o conjunto eficiente real  $\Gamma$  com o conjunto eficiente  $\Gamma_{\theta^*}$  obtido com esta estratégia. Pode-se observar que, embora o tempo de operação ideal tenha sido propriamente identificado ( $t^* = 4.23h$ ), os parâmetros resultantes são tais que uma estimativa muito grosseira do valor máximo de  $C_{B_{max}}$  é obtida.

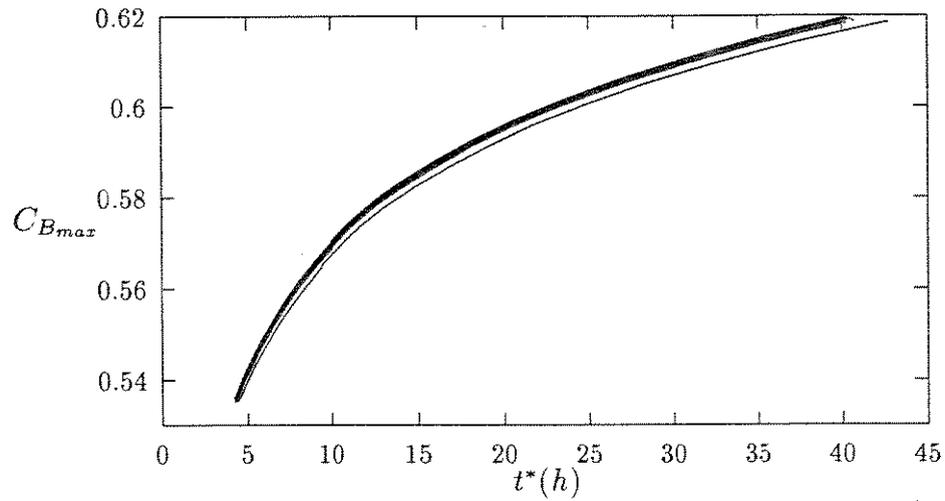
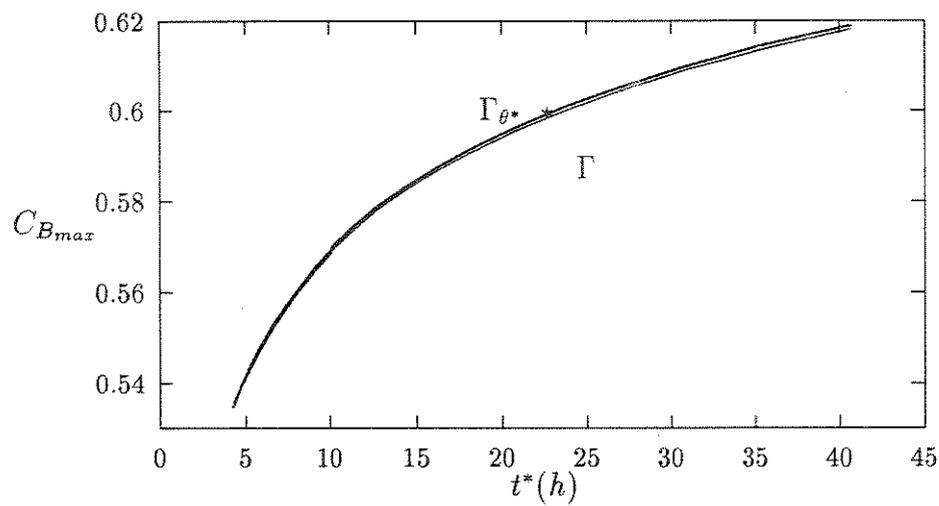


Figura 3.28: Conjuntos eficientes sob medidas ruidosas

Figura 3.29: Comparação entre  $\Gamma$  e  $\Gamma_{\theta^*}$  ("\*" : solução final)

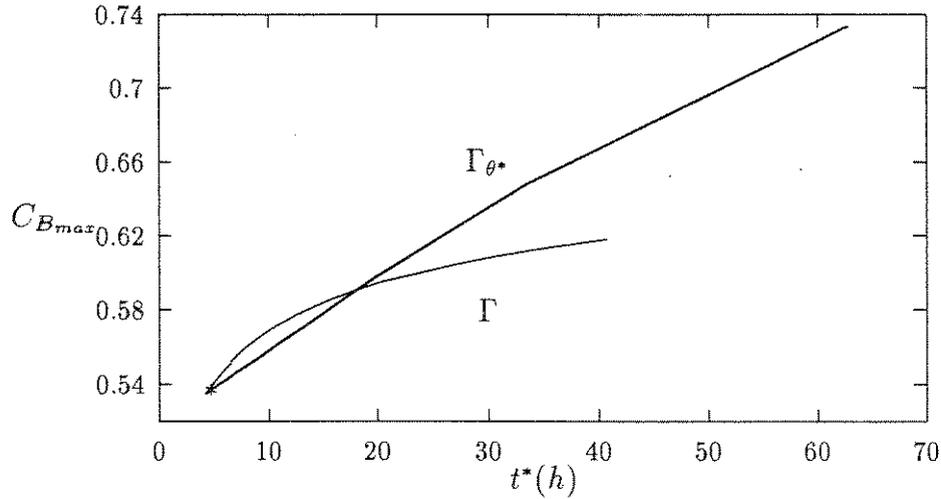


Figura 3.30: Comparação entre  $\Gamma$  e  $\Gamma_{\theta^*}$  para o caso de minimização de  $t^*$  ("\*" : solução final)

Na Figura 3.31, tem-se a comparação de  $\Gamma$  com  $\Gamma_{\theta^*}$  quando apenas a maximização de  $C_{B_{max}}$  é considerada. Neste caso, além do valor encontrado para a concentração máxima ( $C_{B_{max}} = 0.6180$ ) ser uma aproximação muito boa do valor ideal, o tempo de operação também é razoavelmente identificado. Este comportamento é consequência da dependência direta de  $C_{B_{max}}$  com relação a  $t^*$  (expressão 3.25).

Considerando-se, então, a função valor (3.42) com fator de ponderação  $\gamma = 10^{-6}$  obtém-se uma solução de compromisso entre a minimização de  $t^*$  e a maximização de  $C_{B_{max}}$ . Os conjuntos eficientes  $\Gamma_{\theta^k}$  obtidos nas iterações do algoritmo podem ser vistos na Figura 3.32.

Na Figura 3.33 tem-se a comparação de  $\Gamma$  e  $\Gamma_{\theta^*}$  obtida nesta última simulação. Pode-se observar que a solução de compromisso ( $T = 103,44^\circ F$ ,  $C_{B_{max}} = 0.6122$ ,  $t^* = 36.32h$ ), indicada com "\*", é uma representação precisa de um ponto de operação real do reator.

Em todos os casos, pode ser observado que, de acordo com o comportamento predito, as decisões são menos influenciadas por alterações nos parâmetros do modelo, ou seja, as curvas de compromisso  $\Gamma_{\theta^k}$  convergem quando  $k \rightarrow \infty$ .

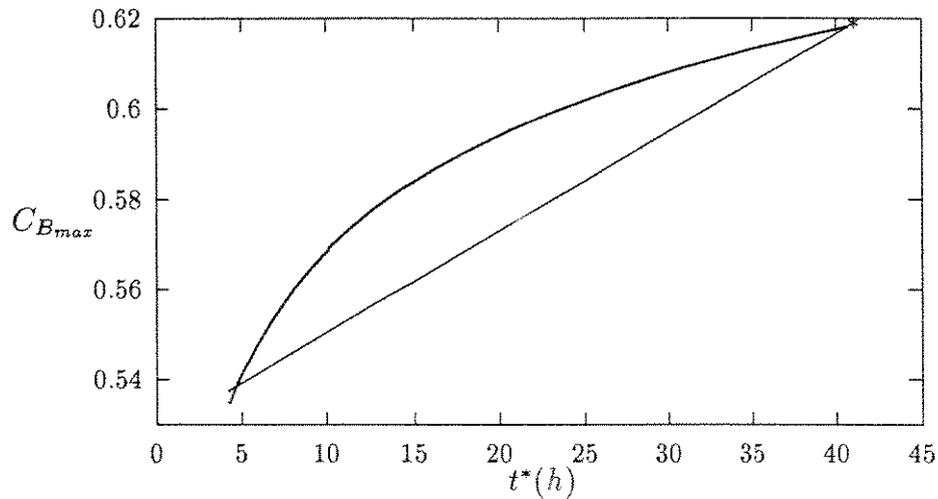


Figura 3.31: Comparação entre  $\Gamma$  e  $\Gamma_{\theta^*}$  para o caso de maximização de  $C_{B_{max}}$  ("\*" : solução final)

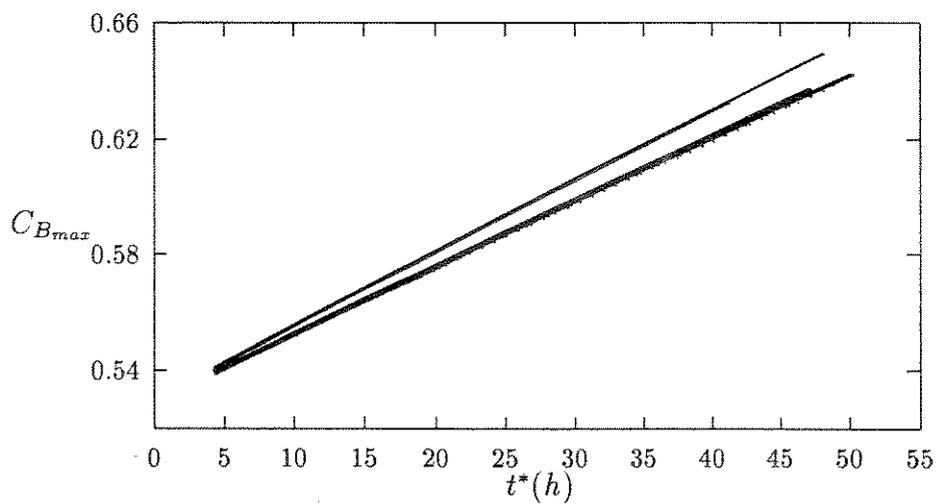


Figura 3.32: Conjuntos eficientes para solução de compromisso  $\gamma = 10^{-6}$  com parâmetros dependentes do controle (modelo linear)

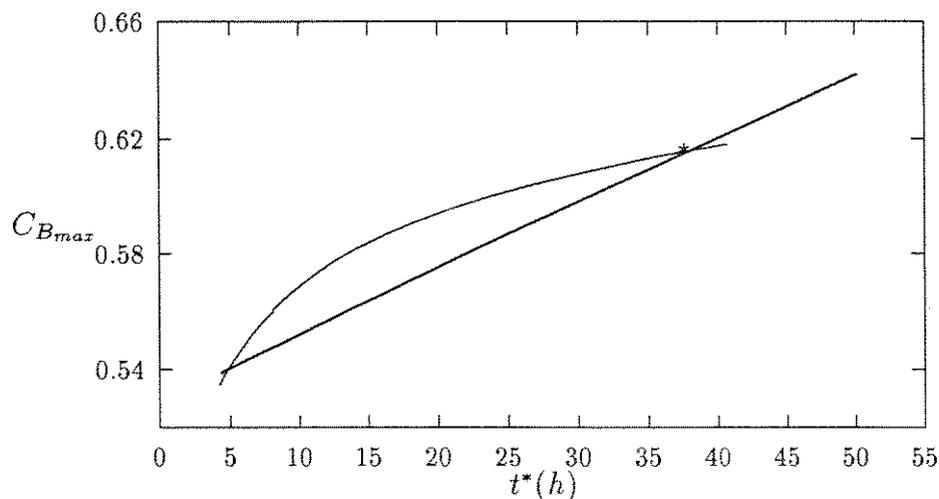


Figura 3.33: Comparação entre  $\Gamma$  e  $\Gamma_{\theta^*}$  para  $\gamma = 10^{-6}$  ("\*" : solução final)

### 3.9 Conclusão

Um reator por batelada não linear foi utilizado, neste capítulo, para ilustrar os aspectos principais da solução de problemas integrados de estimação de parâmetros e otimização multicritério.

No caso, uma formulação bicritério mostrou-se bastante conveniente, considerando-se que os critérios de concentração máxima do componente  $B$  e tempo de operação agregam as informações básicas para caracterização do desempenho do reator.

Os resultados obtidos com esta aplicação, através do SRB, atestam a factibilidade computacional e a utilidade do procedimento como uma ferramenta para síntese de políticas de controle integrado de problemas reais. O enfoque multicritério e interativo se mostra particularmente importante quando os parâmetros do modelo são dependentes do controle. Observe que no caso de reações com taxas lineares, as curvas de trade-off se aproximam em muito de retas com inclinações bem determinadas. Se, por exemplo, um critério global do tipo

$$\min_{T \in T} w C_{B_{max}}(T) + (1 - w)t^*(T), \quad 0 \leq w \leq 1 \quad (3.43)$$

---

fosse adotado, a menos que a inclinação ótima  $w^* \in [0, 1]$  fosse a-priori selecionada, a solução de (3.43) sempre indicaria um dos extremos da curva, o que pode ser inaceitável do ponto de vista prático.

## Capítulo 4

# Identificação e Otimização Conjuntas de um Reator Contínuo: Abordagem Multinível

### 4.1 Introdução

Neste Capítulo, a abordagem dois estágios modificada proposta por Roberts (1979) é estendida ao caso do problema conjunto de identificação de parâmetros e otimização multicritério. Os aspectos de paralelismo apresentados pela estrutura hierárquica obtida com o desacoplamento dos problemas de identificação e otimização são considerados. Como aplicação, a implementação multinível da solução do problema integrado de identificação e otimização multiobjetivo de um reator químico contínuo é apresentada. Para fins de comparação, o mesmo problema é resolvido com base na estratégia dois estágios generalizada proposta no Capítulo 2.

### 4.2 Implementação Multinível – Abordagem Multicritério

Seja o problema conjunto de identificação e otimização multicritério estudado no Capítulo 2. Considere, ainda, a abordagem dois estágios modificada

do problema conjunto escalar proposta por Roberts (1979) e citada no item 1.4.2 deste trabalho. O problema de otimização pode ser reescrito como

$$\begin{aligned} \min \quad & \mathcal{V}(f(x, u, \sigma)) \\ & \varphi(x, u, \sigma) = 0 \\ & g(x, u, \sigma) \leq 0 \end{aligned} \quad (4.1)$$

O problema de identificação, por sua vez, pode ser expresso por

$$\min_{\theta \in \Theta} \{G(v, \theta) = \sum_{i=1}^k \|p(v^i) - m(v^i, \theta)\|_{w_i}^2\} \quad (4.2)$$

Considerando que a solução de (4.2) é um ponto interior de  $\Theta$ , a condição necessária de primeira ordem é

$$\frac{\partial}{\partial \theta} G(v, \theta) = 0 \quad (4.3)$$

O lagrangeano do problema integrado é obtido de forma análoga ao caso escalar apresentado no item 1.4.2. Assim

$$L = \mathcal{V}(F(u, \sigma)) + \gamma'(v - u) + \mu'(\sigma - \theta) + \eta' \frac{\partial}{\partial \theta} G(v, \theta) \quad (4.4)$$

onde  $\gamma$ ,  $\mu$  e  $\eta$  são multiplicadores de Lagrange.

As condições de ponto estacionário do lagrangeano com relação a  $x$  são garantidas, resolvendo-se o problema de otimização multicritério modificado

$$\min_u \mathcal{V}(F(u, \sigma)) - \gamma'u \quad (4.5)$$

sujeito as demais restrições do problema de otimização.

Uma expressão para o multiplicador de Lagrange  $\gamma$  para o caso multicritério também pode ser obtida utilizando-se as condições de ponto estacionário.

$$\gamma = [G_{v\theta}][G_{\theta\theta}]^{-1} \left[ \frac{\partial}{\partial \sigma} \mathcal{V}(F(u, \sigma)) \right]' \quad (4.6)$$

O problema de identificação (4.4) permanece inalterado, valendo a análise feita no item 1.4.2. A estrutura hierárquica obtida é apresentada na Figura 4.1.

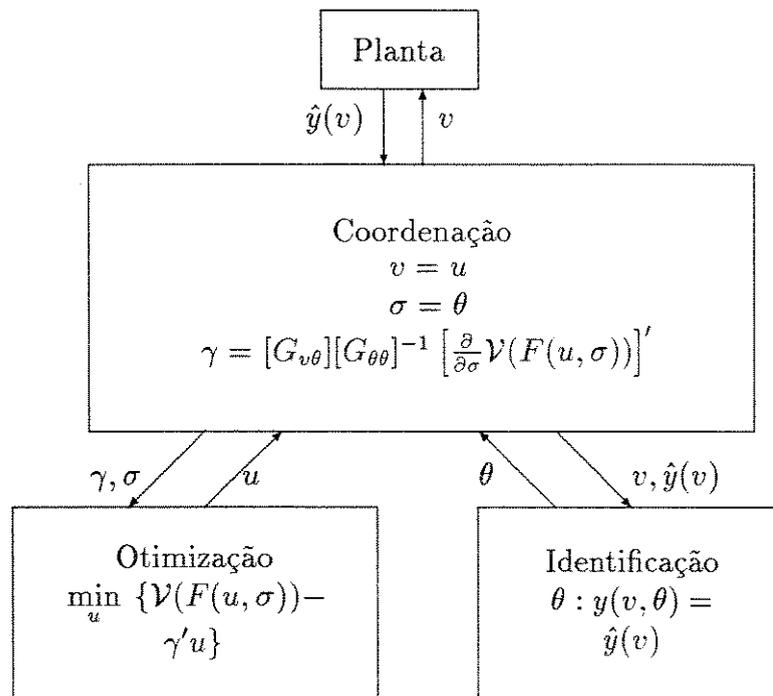


Figura 4.1: Estrutura Hierárquica – Abordagem dois estágios modificada – Caso Multicritério

Caso o número de parâmetros a serem estimados seja igual ao número de saídas, o problema de identificação reduz a determinar o vetor de parâmetros  $\theta$  tal que

$$y(v) = \hat{y}(v, \theta) \quad (4.7)$$

Neste caso, o modificador  $\gamma$  pode ser, então, expresso por

$$\gamma = \left[ \left[ \frac{\partial \hat{y}}{\partial v} \right]' - \left[ \frac{\partial y}{\partial v} \right]' \right] \left[ \left[ \frac{\partial y}{\partial \theta} \right]' \right]^{-1} \left[ \frac{\partial \mathcal{V}(F(u, \sigma))}{\partial \sigma} \right] \quad (4.8)$$

Note que se  $\gamma = 0$ , o algoritmo básico de dois estágios, apresentado no Capítulo 2, fica imediatamente caracterizado. Obviamente,  $\gamma = 0$  se as derivadas das saídas da planta e do modelo são iguais. O termo extra no subproblema de otimização atua no sentido de garantir que a função valor  $\mathcal{V}$  será minimizada, mesmo que as derivadas das saídas do modelo e da planta com relação ao controle não se igualem. Nestas circunstâncias é fácil mostrar que a solução obtida é eficiente.

**Teorema 4.1** *Suponha que  $\mathcal{V}(\cdot)$  é uma função valor estritamente crescente em relação a cada função  $F_i(\cdot)$ ,  $i = 1, \dots, m$  e que o algoritmo Dois Estágios Modificado converge para um mínimo local do problema integrado. Então a solução obtida é também localmente eficiente.*

*Prova:* Seja  $(u^*, \sigma^*)$  uma solução ótima local obtida pelo algoritmo. Neste caso,

$$\mathcal{V}(F(u^*, \sigma^*)) \leq \mathcal{V}(F(u, \sigma^*)), \quad \forall x \in \Psi_{\sigma^*} \cap \mathcal{N}(u^*, \epsilon) \quad (4.9)$$

onde  $\mathcal{N}(u^*, \epsilon) \triangleq \{u : \|u - u^*\| < \epsilon\}$  e  $\Psi_{\sigma^*} = \{u : \psi(u, \sigma^*) \leq 0\}$ .

Suponha que  $u^*$  não é localmente eficiente. Então, existe  $u^\circ \in \Psi_{\sigma^*} \cap \mathcal{N}(u^*, \epsilon)$  tal que

$$F(u^*, \sigma^*) \leq F(u, \sigma^*) \quad \text{e} \quad F(u^*, \sigma^*) \neq F(u, \sigma^*) \quad (4.10)$$

o que contradiz (4.9), dada a hipótese de crescimento de  $\mathcal{V}(\cdot)$  em relação a cada  $F_i(\cdot)$ .

Os aspectos relativos à convergência do procedimento e análise de sensibilidade abordados no Capítulo 2 com relação ao algoritmo clássico estendido ao caso vetorial são válidos também na estratégia dois estágios modificado.

Para ilustrar a factibilidade e o melhor desempenho da abordagem dois estágios modificada na solução do problema conjunto multicritério, é apresentado a seguir um estudo de caso, referente a um reator químico contínuo.

### 4.3 Reator Químico Contínuo

Os reatores contínuos são largamente utilizados na indústria química. Estes sistemas consistem basicamente de um tanque, com a inserção de reagentes e a retirada de produto ocorrendo de forma contínua. Usualmente são empregadas baterias de reatores conectados em série (Chen, 1983). Por simplicidade, neste trabalho será estudado um reator único como mostrado na Figura 4.2.

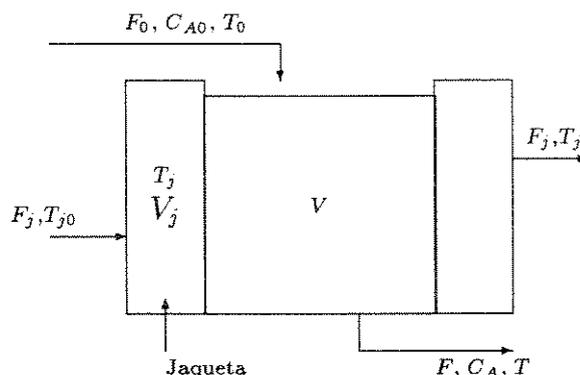


Figura 4.2: Reator químico contínuo

O líquido reagente é inserido continuamente no recipiente, no interior do qual ocorre a seguinte reação irreversível de primeira ordem



onde  $A$  é o reagente,  $B$  é o produto desejado e  $k$  é a taxa de reação (Luyben, 1973) que pode ser expressa pela equação de Arrhenius

$$k = \alpha e^{\frac{E}{RT}} \quad (4.11)$$

No Apêndice deste trabalho, as principais variáveis e parâmetros que compõem o modelo do reator são descritas.

Uma jaqueta de resfriamento envolve o recipiente para remoção do calor exotérmico da reação.

A dinâmica do processo pode ser descrita através do seguinte conjunto de equações diferenciais (Luyben, 1973)

$$\frac{dV}{dt} = F_0 - F \quad (4.12)$$

$$\frac{d}{dt}(VC_A) = F_0C_{A0} - FC_A - VkC_A \quad (4.13)$$

$$\frac{d}{dt}(VC_B) = F_0C_{B0} - FC_B + VkC_A \quad (4.14)$$

$$\frac{d}{dt}(VT) = F_0T_0 - FT + \frac{\lambda VkC_A}{\rho C_p} - \frac{UA_H}{\rho C_p}(T - T_j) \quad (4.15)$$

$$\frac{dT_j}{dt} = F_j(T_{j0} - T_j) + \frac{UA_H}{\rho_j C_j V_j}(T - T_j) \quad (4.16)$$

As variáveis de controle selecionadas são a temperatura da reação  $T$  e o fluxo de material  $F_0$ . Um controlador PI digital se encarrega de manter a temperatura no seu valor selecionado através da manipulação do fluxo de água na jaqueta  $F_j$ , supondo o volume da jaqueta constante.

$$F_j(t) = F_j(t - \Delta t) + k_p[e(t) - e(t - \Delta t) + \frac{\Delta t}{\tau_i}e(t)] \quad (4.17)$$

onde

$e(t) = T - T_{set}$ : diferença entre o valor medido e o set-point de temperatura.

$k_p$ : ganho proporcional

$\tau_i$ : constante de tempo integral

$\Delta t$ : intervalo de integração

O volume do reagente é mantido constante através do ajuste do fluxo de entrada de tal maneira que  $F = F_0$ .

As expressões (4.11)–(4.17) e os valores da tabela 4.1 são utilizadas no simulador do reator contínuo — SRC, o qual fornece as medidas necessárias para a solução do problema conjunto. Observe que pouca sofisticação é necessária neste caso em termos do sistema de controle, uma vez que as

medidas serão feitas em estado estacionário e a parte transitória não tem efeito prático sobre o desempenho econômico do reator. Com um controlador PI de ganho fixo procura-se apenas evitar erros de regime nos set-points determinados.

$\bar{F}$	40 ft <sup>3</sup> /hr	$U$	150 Btu/(hr - ft <sup>2</sup> - °R)
$\bar{V}$	48 ft <sup>3</sup>	$A$	250 ft <sup>2</sup>
$\bar{C}_{A0}$	0.50 mole/ft <sup>3</sup>	$T_{j0}$	530°R
$\bar{C}_A$	0.245 mole/ft <sup>3</sup>	$T_0$	530°R
$\bar{T}$	600°R	$\lambda$	-30000 Btu/mole
$\bar{T}_j$	594.6°R	$C_p$	0.75 Btu/lb <sub>m</sub> - °R
$\bar{F}_j$	49.9 ft <sup>3</sup> /hr	$C_j$	1.0 Btu/lb <sub>m</sub> - °R
$\bar{V}_j$	3.85 ft <sup>3</sup>	$\rho$	50 lb <sub>m</sub> /ft <sup>3</sup>
$\alpha$	7.08 10 <sup>10</sup> hr <sup>-1</sup>	$\rho_j$	62.3 lb <sub>m</sub> /ft <sup>3</sup>
$E$	30000 Btu/mole	$K_c$	4 (ft <sup>3</sup> /hr)/°R
$R$	1.99 Btu/mole°R	$T^{set}$	600°R
$k_p$	30.0	$\tau_i$	0.01

Tabela 4.1: Parâmetros do Reator Contínuo

## 4.4 Formulação do Problema Conjunto

O problema de identificação pode ser formulado, considerando-se que os parâmetros a serem estimados são a taxa de reação  $k$  e o calor exotérmico da reação  $\lambda$  e dispondo-se de medidas de  $C_A$  e  $T_j$ , cujos modelos em estado estacionário são expressos por

$$\bar{C}_A = \frac{F_0 C_{A0}}{F_0 + k\bar{V}} \quad (4.18)$$

$$\bar{T}_j = \frac{\lambda\bar{V}k\bar{C}_A - F_0 C_p \rho T_0 + (F_0 C_p \rho + UA_H)\bar{T}}{UA_H} \quad (4.19)$$

Assim, usando-se a identidade como matriz de ponderação, o problema genérico (4.2) pode ser particularizado para

$$\min_{\theta \in \Theta} \frac{1}{2} \sum_{i=1}^N \{(\bar{C}_A(v, \theta) - \hat{C}_A(v))^2 + \beta(\bar{T}_j(v, \theta) - \hat{T}_j(v))^2\} \quad (4.20)$$

onde  $\theta = [k \ \lambda]'$  e  $\Theta = [\underline{\theta}, \bar{\theta}]$ .

Para fins de otimização será utilizado o modelo estacionário do reator. De maneira a evitar a consideração explícita de restrições não-lineares, será utilizado um modelo de estado estacionário linearizado em torno do ponto de operação do reator a cada interação entre o subproblema de otimização e o nível de coordenação. Neste caso, o subproblema de otimização pode ser colocado na forma geral

$$\begin{aligned} \min_{x,u} \quad & \mathcal{V}(f(x, u, \sigma)) \\ & A(\sigma)x + B(\sigma)u = b(\sigma) \\ & \{\bar{x}(1 - \Delta) \leq x \leq \bar{x}(1 + \Delta)\} \\ & \{\bar{u}(1 - \Delta) \leq u \leq \bar{u}(1 + \Delta)\} \end{aligned} \quad (4.21)$$

onde

$$A(\sigma) = \begin{bmatrix} (-\bar{F}_o + Vk) & 0 & 0 & 0 \\ Vk & (-\bar{F}_o) & 0 & 0 \\ \left(-\frac{\lambda Vk}{\rho C_p}\right) & 0 & \left(\frac{UA_H}{\rho C_p}\right) & 0 \\ 0 & 0 & \left(-\frac{F_j}{V_j} - \frac{UA_H}{\rho_j V_j C_j}\right) & \left(\frac{T_{j0} - \bar{T}_j}{V_j}\right) \end{bmatrix}$$

$$B(\sigma) = \begin{bmatrix} 0 & (\bar{C}_A - C_{A0}) \\ 0 & (\bar{C}_B - C_{B0}) \\ 0 & (\bar{T} - T_o) \\ \left(-\frac{UA_H}{\rho_j V_j C_j}\right) & 0 \end{bmatrix}$$

$$x = [C_A \ C_B \ T_j \ F_j]'$$

$$u = [T \ F]'$$

$$b(\sigma) = \left[ -\bar{F}_o \bar{C}_A \quad -\bar{F}_o \bar{C}_B \quad -\bar{F}_o \bar{T} \quad -\frac{\bar{F}_j \bar{T}_j}{V_j} \right]'$$

e  $0 \leq \Delta \leq 1$  determina a variação de  $x$  e  $u$  em torno do ponto de operação nominal  $(\bar{x}, \bar{u})$ . Note que esta estratégia vai ao encontro de importantes necessidades práticas, pois desta forma assegura-se que não ocorrerão desvios excessivamente grandes em relação ao ponto de operação nominal, conforme discutido no Capítulo 1.

Considerando-se um dado ponto de operação, com o vetor de parâmetros correspondente, as matriz  $A$  e o vetor  $b$  são constantes e compõem um conjunto de restrições lineares. Assim, para uma função valor  $\mathcal{V}$  explícita, o problema (4.21) apresenta-se adequado à aplicação do algoritmo proposto no Capítulo 2.

A operação de qualquer reator químico é caracterizada pela existência de critérios e/ou especificações de desempenho conflitantes. Chen (1983) apresenta, para o caso de reatores contínuos, critérios como minimização do volume, maximização da concentração do produto desejado, minimização do tempo de residência, ou seja, determinação da temperatura ótima de operação, minimização do custo total e maximização do lucro, dentre outros.

Para efeito de obtenção de resultados, dois critérios foram selecionados. O primeiro deles corresponde a maximização do lucro de operação, ou seja,

$$\min \{P(x, \sigma) = -s_1 FC_B + c_1 F\} \quad (4.22)$$

onde  $s_1$  é o preço unitário de venda do produto  $B$  e  $c_1$  é o custo unitário da matéria prima. O outro critério visa garantir a obtenção de um valor objetivado  $C_B^o$  para a concentração de  $B$ ,

$$\min s_2 (C_B - C_B^o)^2 \quad (4.23)$$

onde  $s_2$  pode ser interpretado como uma penalização sobre os desvios de  $C_B$  em relação a  $C_B^o$ .

A análise multicritério possibilita a determinação de uma solução de compromisso entre os critérios fixados, de acordo com uma função valor  $\mathcal{V}$  que representa implícita ou explicitamente as preferências de um Decisor. Um passo importante na busca de uma função valor apropriada para problemas de otimização dinâmicos e/ou de grande porte é a geração de um subconjunto representativo de soluções eficientes (Haimes e Li, 1988).

O problema ponderado será utilizado para a geração dos pontos eficientes candidatos. Assim, considerando-se as restrições impostas pelo modelo adotado e  $w \in [0, 1]$ , o problema ponderado pode ser formulado como

$$\min w(-s_1 FC_B + c_1 F) + (1 - w)s_2 (C_B - C_B^o)^2 \quad (4.24)$$

O objetivo desta escolha é analisar o comportamento da solução de (4.24) e explicitar a dependência dos parâmetros a serem estimados em relação a  $w$ , e portanto em relação aos trade-offs do Decisor.

## 4.5 Aspectos da Implementação

Na implementação da abordagem dois estágios modificada, procurou-se explorar os aspectos de paralelismo apresentados pela estrutura hierárquica da Figura 4.1. Para tanto, utilizou-se uma estação de trabalho SPARC 1+ da SUN que possibilita a execução concorrente de vários processos, bem com a comunicação entre os mesmos.

Nesta aplicação, apenas um processador foi compartilhado pelos processos, explorando a capacidade de multiprocessamento da estação SUN. Esta técnica consiste na divisão do tempo da CPU entre processos de

maneira que eles estejam aparentemente sendo executados simultaneamente. Na realidade, a cada instante, apenas uma sequência de instruções está sendo processada, mas o chaveamento de um processo para outro é transparente para os mesmos. A alocação de memória e do tempo de CPU, a rotação de processos e a resolução de disputas por dispositivos periféricos são atribuições do sistema operacional.

Cada subsistema da Figura 4.1 foi implementado como um processo e as ligações entre eles equivalem a canais de comunicação. Para estabelecer a concorrência entre os processos foram utilizadas as chamadas do sistema operacional *fork()* e *execl()*. A comunicação entre a *Coordenação* e cada um dos demais processos é feita através de dois canais unidirecionais com sentidos opostos estabelecidos através da chamada *pipe()*.

O processo *Planta* é responsável pela simulação do reator contínuo em estudo. Recebe como entradas padrão as variáveis

de decisão do problema de otimização que atuam como set-points para os controladores e fornece medidas de estado estacionário.

O problema de identificação de parâmetros, expresso em (4.20), é solucionado pelo processo *Identificação*. As entradas padrão deste módulo são os set-points determinados pela otimização e as medidas em estado estacionário de  $C_A$  e  $T_j$ . Como saídas, tem-se as estimativas dos parâmetros  $k$  e  $\lambda$ .

O processo *Otimização* é responsável pela solução do problema modificado que, considerando-se as restrições impostas pelo modelo, pode ser expresso por

$$\min_{x,u} \{wP(x,u,\sigma) + (1-w)s_2(C_B - C_B^o)^2 - \gamma u\} \quad (4.25)$$

onde  $w \in [0,1]$  é fixado antes do início da simulação. Este processo recebe medidas do estado estacionário para composição do modelo adaptativo, os

parâmetros estimados pela *Identificação* e os multiplicadores  $\gamma$ , calculados pelo processo *Coordenação*.

Os problemas de identificação de parâmetros e otimização são resolvidos através do método do Gradiente Modificado proposto por Geromel e Ferreira (1986), aliado a uma estratégia de solução paralela do problema dual associado, baseado na iteração de Jacobi (Bertsekas, 1989), como foi mostrado no item 2.6 deste trabalho.

O processo *Coordenação* gerencia a comunicação entre os demais processos e é responsável pela atualização dos multiplicadores de Lagrange. O problema foi resolvido de forma síncrona, isto é, os processos aguardam a chegada dos resultados dos outros antes de partir para uma próxima iteração.

## 4.6 Resultados Computacionais

O problema conjunto de identificação e otimização multicritério do reator contínuo formulado no item 4.4 foi particularizado para fins de análise dos resultados, considerando-se  $s_1 = 2.5$ ,  $c_1 = 0.8$ ,  $s_2 = 10^3$  e  $C_B^o = 0.53$ .

Nas Tabelas 4.2, 4.3 e 4.4 tem-se os resultados obtidos via simulação para os valores de  $w$  indicados.

A convergência do algoritmo é determinada por uma variação percentual dos multiplicadores de Lagrange menor do que um certo limite pré-estabelecido, tipicamente  $10^{-5}$ . A partir das Tabelas 4.2, 4.3 e 4.4, pode-se verificar a convergência dos multiplicadores de Lagrange, fato que fica melhor evidenciado pelas curvas das Figuras 4.3, 4.4 e 4.5. As derivadas das saídas da planta foram obtidas por diferenças finitas, perturbando-se as variáveis de controle em 10% dos seus valores calculados.

Na Tabela 4.5 tem-se os resultados obtidos via simulação, variando-se  $w$  com incrementos de 0.2. São apresentados os valores dos set-points de temperatura e fluxo, os valores relativos a concentração  $C_B$  e ao lucro  $P$  com a operação do reator como função de  $w$ , bem como os multiplicadores de Lagrange ótimos associados.

Observa-se claramente a dependência dos parâmetros ótimos estimados com o relação aos controles, principalmente no que diz respeito à taxa de reação  $k$ , tornando assim indispensável o uso da abordagem integrada.

Note o caráter conflitante dos dois índices de desempenho, o que faz com

$\hat{C}_A$	$\hat{T}_j$	$T$	$F_o$	$k$	$\lambda$	$\gamma_1$	$\gamma_2$
0.2695	593.83	600.10	36.00	0.732	-28999.9	0.0000	0.0000
0.2545	594.12	603.34	32.40	0.891	-29000.0	2.9514	0.0356
0.2221	594.12	606.19	29.16	0.959	-29000.0	2.3714	0.0521
0.1933	594.12	608.16	26.25	1.085	-29000.1	1.9777	0.0870
0.1707	594.12	610.52	23.62	1.210	-29000.2	1.5673	0.0819
0.1480	594.12	612.63	21.28	1.313	-29000.2	1.2087	0.0709
0.1284	594.12	614.61	19.20	1.438	-29000.3	0.9329	0.0646
0.1113	594.12	616.42	17.34	1.562	-29000.5	0.7015	0.0587
0.0967	594.12	618.01	15.64	1.687	-29000.7	0.5216	0.0470
0.0842	594.12	619.36	14.08	1.811	-29000.9	0.3849	0.0347
0.0735	594.12	620.45	12.68	1.811	-29000.9	0.2808	0.0209
0.0646	594.12	621.57	11.41	1.980	-29001.2	0.2208	0.0338
0.0565	594.12	622.56	10.27	1.980	-29001.2	0.1527	0.0145
0.0497	594.12	623.21	9.24	1.980	-29001.2	0.1170	0.0216
0.0441	594.12	623.77	8.33	1.980	-29001.2	0.0894	0.0250
0.0392	594.12	624.90	7.54	1.980	-29001.2	0.0680	0.0267
0.0343	594.12	625.39	6.86	1.980	-29001.2	0.0502	0.0295
0.0308	594.12	625.87	6.25	1.980	-29001.2	0.0377	0.0274
0.0278	594.12	627.51	5.71	1.980	-29001.2	0.0280	0.0248
0.0240	594.12	629.22	5.25	1.980	-29001.2	0.0194	0.0259
0.0208	594.12	629.22	5.25	1.980	-29001.2	0.0133	0.0248

Tabela 4.2: Convergência do algoritmo dois estágios modificado -  $w = 0.0$

$\hat{C}_A$	$\hat{T}_j$	$T$	$F_o$	$k$	$\lambda$	$\gamma_1$	$\gamma_2$
0.2695	593.83	600.10	36.00	0.732	-28999.9	0.0000	0.0000
0.2545	594.12	601.68	32.40	0.891	-29000.0	1.9765	0.0238
0.2313	595.74	603.36	29.16	0.920	-29000.1	1.6309	0.0171
0.2082	597.52	605.99	26.25	0.986	-29000.1	1.4273	0.0343
0.1813	600.25	608.44	23.62	1.088	-29000.2	1.1797	0.0473
0.1573	602.87	610.72	21.31	1.214	-29000.2	0.9315	0.0602
0.1363	605.38	613.05	19.40	1.339	-29000.3	0.7166	0.0560
0.1179	607.96	615.41	17.74	1.464	-29000.5	0.5472	0.0456
0.1018	610.57	617.88	16.31	1.588	-29000.6	0.4149	0.0466
0.0876	613.28	620.53	15.08	1.758	-29000.8	0.3130	0.0443
0.0750	616.16	623.61	14.09	1.949	-29001.1	0.2290	0.0378
0.0636	619.44	628.15	13.52	2.185	-29001.5	0.1650	0.0350
0.0524	624.10	628.15	13.52	2.532	-29002.2	0.1163	0.0409
0.0524	624.10	628.15	13.52	2.532	-29002.2	0.0996	0.0101

Tabela 4.3: Convergência do algoritmo dois estágios modificado -  $w = 0.4$ 

$\hat{C}_A$	$\hat{T}_j$	$T$	$F_o$	$k$	$\lambda$	$\gamma_1$	$\gamma_2$
0.2695	593.83	600.10	36.00	0.732	-28999.9	0.0000	0.0000
0.2545	594.12	603.34	32.40	0.891	-29000.0	0.5141	0.0062
0.2221	597.22	606.19	29.16	0.959	-29000.0	0.4424	0.0097
0.1933	600.09	609.14	26.58	1.085	-29000.1	0.3730	0.0164
0.1675	603.11	613.57	25.71	1.233	-29000.2	0.3125	0.0176
0.1439	607.37	613.57	25.71	1.427	-29000.3	0.2718	0.0278
0.1439	607.37	613.57	25.71	1.427	-29000.3	0.2718	0.0104

Tabela 4.4: Convergência do algoritmo dois estágios modificado -  $w = 1.0$

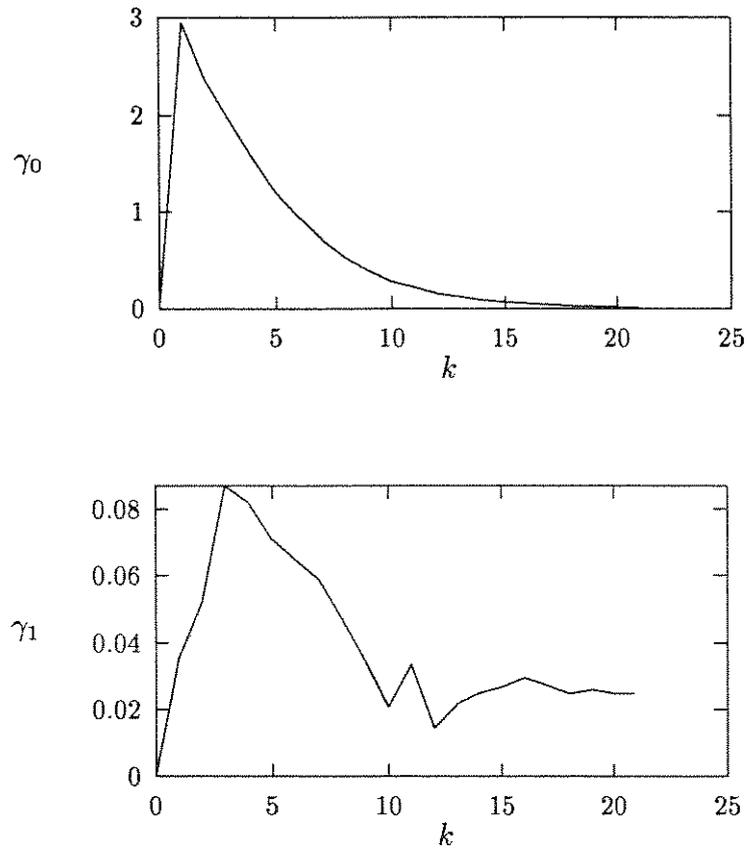
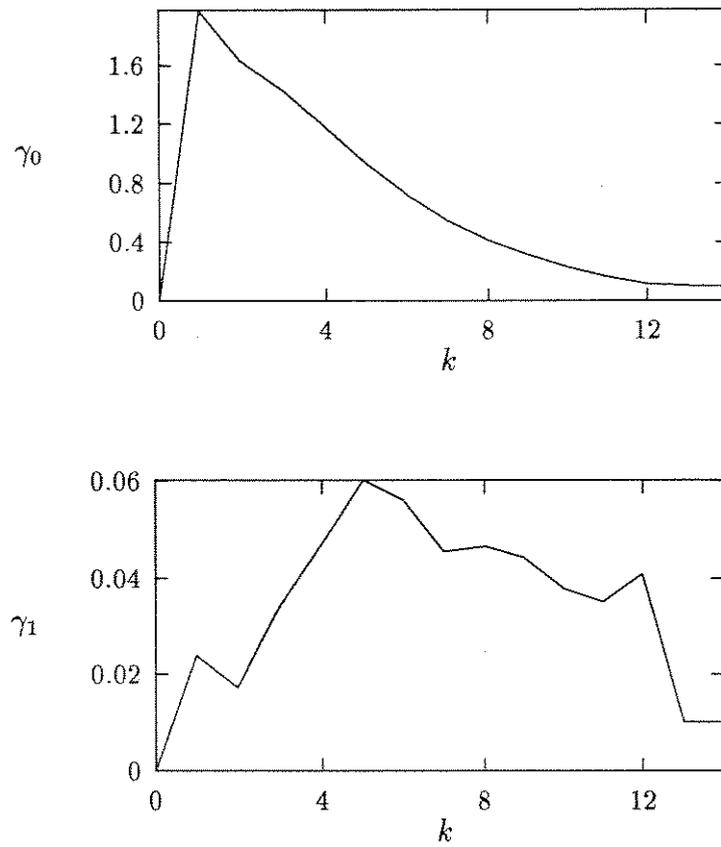
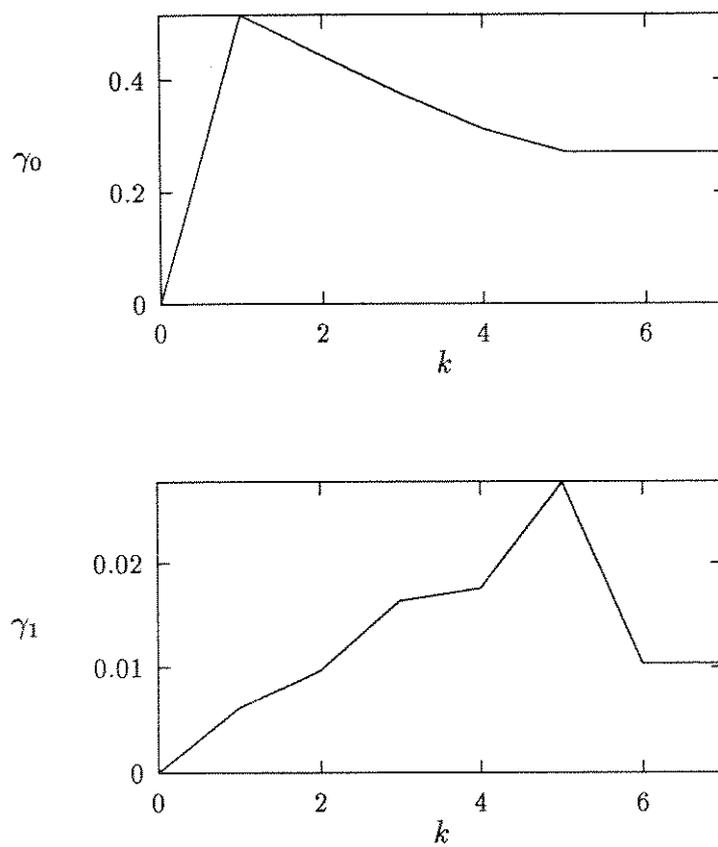


Figura 4.3: Evolução dos multiplicadores de Lagrange –  $w = 0.0$

$w$	$k$	$\lambda$	$T$	$F_o$	$P$	$C_B$	$\gamma_1$	$\gamma_2$
0.0	1.980	-29001.2	629.22	5.25	2.75	0.5292	0.0133	0.0248
0.2	2.515	-29001.7	627.46	8.09	3.98	0.5165	0.0325	0.0037
0.4	2.532	-29002.2	628.15	13.52	6.00	0.4976	0.0996	0.0001
0.6	2.171	-29001.1	624.71	16.93	6.69	0.4780	0.1698	0.0169
0.8	1.769	-29000.6	619.14	22.09	6.74	0.4421	0.2760	0.0151
1.0	1.427	-29000.3	613.57	25.71	5.54	0.4061	0.2718	0.0104

Tabela 4.5: Dependência dos parâmetros e critérios com  $w$  – Abordagem dois estágios modificada multicritério

Figura 4.4: Evolução dos multiplicadores de Lagrange -  $w = 0.4$

Figura 4.5: Evolução dos multiplicadores de Lagrange –  $w = 1.0$

que a concentração diminua progressivamente a medida que  $w$  cresce a partir de  $w = 0$ , quando apenas o critério relativo a obtenção de  $C_B^o$  é considerado, até  $w = 1$ , quando apenas o lucro é valorizado.

Finalmente, observe que a solução obtida na condição  $w = 1.0$  é *dominada* pela solução anterior ( $w = 0.8$ ). Este comportamento pode ser explicado pelo caráter não convexo de  $P$  e a eventual geração de soluções eficientes locais, neste caso. Além disso, na condição  $w = 1.0$  não se tem uma função valor (4.24) crescente com relação a cada objetivo, e portanto não se pode garantir, pelo Teorema 4.1 que a solução obtida seja eficiente.

Para fins de comparação, os casos simulados na Tabela 4.5 foram repetidos, tomando-se  $\gamma = 0$ , ou seja, utilizando-se a abordagem dois estágios generalizada, enfocada no Capítulo 2. A Tabela 4.6 contém os resultados obtidos.

$w$	$k$	$\lambda$	$T$	$F_o$	$P$	$C_B$
0.0	0.891	-29000.0	598.62	29.16	-1.14	0.2440
0.2	0.891	-29000.0	598.24	29.16	-1.09	0.2435
0.4	0.891	-29000.0	598.24	29.16	-1.09	0.2435
0.6	0.891	-29000.0	598.24	29.16	-1.09	0.2435
0.8	0.585	-29000.1	590.4	8.24	2.11	0.1376
1.0	0.850	-29000.1	597.63	19.13	1.98	0.1927

Tabela 4.6: Dependência dos parâmetros e critérios com  $w$  – Abordagem dois estágios generalizada (multicritério)

A versão modificada apresentou uma maior sensibilidade com relação à variação de  $w$ , além de fornecer soluções muito melhores do que aquelas obtidas com a abordagem clássica.

## 4.7 Conclusão

Neste Capítulo, a estratégia dois estágios modificada proposta por Roberts (1979) foi abordada no contexto das técnicas de otimização multicritério. Uma implementação multinível, explorando os aspectos de paralelismo apresentados pela estrutura hierárquica obtida com o desacoplamento dos problemas de identificação de parâmetros e otimização foi utilizada pa-

ra a solução do problema integrado formulado com base num sistema físico, representado por um reator químico contínuo.

Os resultados obtidos evidenciam a importância do enfoque integrado e multicritério em aplicações em que os parâmetros a serem estimados dependem das variáveis de controle que, por sua vez, são diretamente afetadas pela escolha do critério de desempenho. De acordo com a abordagem proposta, os parâmetros do modelo refletirão preferências e julgamentos de valor de um Decisor.

A utilização da interação de Jacobi na solução dos problemas de otimização presentes na implementação computacional da abordagem proposta apresentou resultados bastante satisfatórios.

Em todos os casos, observou-se a convergência para as soluções ótimas do problema (4.5), embora a velocidade de convergência tenha caído substancialmente com a proximidade dos autovalores de  $Q$  em relação a 0. A maior dificuldade numérica encontrada não se deveu à solução via iteração de Jacobi, mas ao cálculo da matriz de projeção  $P$  que, em função do conjunto de restrições formado, apresentou autovalores de pequeno valor, porém negativos. Nestes casos, foram necessárias algumas estratégias de pós-condicionamento numérico de maneira a obter  $Q \geq 0$ .

## Conclusão Geral

O tratamento integrado dos problemas de identificação de parâmetros e otimização tem sido defendido por diversos autores. O caráter multicritério da maioria dos problemas práticos de otimização de sistemas é também amplamente reconhecido. A inclusão da análise multicritério na abordagem do problema conjunto é, desta maneira, perfeitamente justificada.

As abordagens dois estágios e dois estágios modificada estendidas para o caso de otimização multicritério foram introduzidas neste trabalho. A convergência destes algoritmos para a solução ótima, sob condições específicas, é demonstrada e a implementação efetiva dos mesmos confirma sua factibilidade.

Como aplicações foram usados dois tipos de reatores químicos, o "batch" para o algoritmo dois estágios e o contínuo para o algoritmo modificado. Para tratamento do reator por batelada foi desenvolvido o SRB – Simulador do Reator por Batelada – que possui uma interface bastante amigável, incluindo a opção de otimização interativa, na qual o operador pode selecionar o ponto de operação desejado de acordo com critérios subjetivos. O SRC – Simulador do Reator Contínuo – explora o paralelismo apresentado pela estrutura multinível obtida através da decomposição do problema conjunto. A utilização de multiprogramação ou multiprocessamento possibilita uma maior eficiência na solução do problema como um todo, o que pode ser muito significativo no caso de sistemas físicos com dinâmica rápida. Outro aspecto relativo a um melhor desempenho na implementação dos algoritmos é a utilização da interação de Jacobi na solução do problema dual associado ao Método do Gradiente Modificado.

Esta linha de trabalho apresenta opções bastante ricas de continuação. O tratamento dos problemas conjuntos identificação e otimização multicritério nas suas versões dinâmicas é uma extensão natural dos resultados apresentados. Seria interessante também explorar a aplicação das abordagens propostas a processos físicos com dinâmica rápida. Neste ponto, é bom ressaltar a

importância do uso de estratégias mais versáteis para solução do problema de identificação de parâmetros, que segundo Kambhampati et al (1992) pode ser eficientemente resolvido apenas com informações do estado transitório da planta.

Finalmente, seria importante abordar os problemas tratados neste trabalho no contexto das técnicas de Análise de Sensibilidade (Fiacco, 1983) que consideram problemas de otimização genéricos do tipo

$$\begin{aligned} \min \quad & \mathcal{V}(f(x, \theta)) && (P^\theta) \\ & \varphi(x, \theta) = 0 \\ & g(x, \theta) \leq 0 \end{aligned}$$

onde tanto a função objetivo quanto as restrições estão sujeitas a parâmetros incertos  $\theta$ .

Os autores acreditam que o fato de se estabelecer algum tipo de comportamento associado a  $\theta$ , como discutido no Capítulo 2, pode levar a importantes particularizações dos resultados disponíveis em análise de sensibilidade e permitir caracterizações mais precisas dos conjuntos de trade-offs do problema  $(P^\theta)$  em função de  $\theta$ .

# Bibliografia

- Abastanotti, D., Colaneri, P., Geromel, J.C., Ghidoni, R., and Gardabassi, G., (1989). *Nicely Nolinear Modelling: A DF Approach with Application to "Linearized" Process Control*. Symposium on Nolinear Control Systems Design, Capri, Italy.
- Bertsekas, D.P. and Tsitsiklis, J.N., (1989). *Parallel and Distributed Computation - Numerical Methods*, Prentice-Hall.
- Borges, T. C. D. and Ferreira, P. A. V., (1991). *Identificação e Otimização Conjuntas*. Nota Interna n. 01/91, UNICAMP.
- Borges, T. C. D. e Ferreira, P. A. V., (1992). *Identificação e Otimização Conjunta: uma Abordagem Multicritério*. *Submetido ao 9º CBA*, Vitória, Espírito Santo.
- Brdys, M. and Roberts, P. D., (1986). *Optimal Structures for Steady-State Adaptive Optimizing Control of Large-Scale Industrial Processes*. *Int. J. Systems Sci.*, vol. 1.
- Chankong, V. and Haimes, Y. Y., (1982). *On the Characterization of Noninferior Solutions of the Vector Optimization Problem*. *Automatica*, vol. 18, n. 6.
- Chen, N. H., (1983). *Process Reactor Design*. Boston, Allyn & Bacon-Series in Engineering.
- Cohon, J. L., Church, R. and Sheer, D., (1979). *Generating Multiobjective Trade-offs: An Algorithm for Bicriterion Problems*. *Water Resources Research*, vol. 15.

- Denbigh, K. G. and Turner, J. C. R., (1984). *Chemical Reactor Theory: An Introduction*. 3th Edition, Cambridge University Press.
- Ellis, J. E. and Roberts, P. D. (1982). Measurement and Modelling Trade-offs for Integrated System Optimization and Parameter Estimation. *Large Scale Systems*, vol. 3.
- Ferreira, P. A. V. (1983). *um Método Primal para Solução de Problemas de Controle Ótimo de Sistemas Dinâmicos de Grande Porte*, Dissertação de Mestrado, DEE-FEC, UNICAMP.
- Ferreira, P. A. V. and Geromel, J. C., (1990). An Interactive Projection Method for Multicriteria Optimization Problems. *IEEE Trans. on Systems, Man and Cybernetics*, vol. SMC-20, n. 3.
- Ferreira, P. A. V. and Borges, T. C. D., (1992). System Modelling and Optimization Under Vector-Valued Criteria, *Proceedings of the ACC 92*, june, 24-26, Chicago, Illinois.<sup>1</sup>
- Fiacco, A. V. (1983). *Introduction to Sensitivity and Stability Analysis in Nolinear Programming*, Academic Press.
- Foss, A. S., (1973). Critique of Chemical Process Control Theory. *IEEE Transactions on Automatic Control*, december.
- Franklin, M. A. and Soong, N. L., (1981). One-dimensional Optimization on Multiprocessor Systems. *IEEE Trans. on Computers*, vol. C-30, n. 1, january.
- Geoffrion, A. M., (1967). Solving Bicriterion Mathematical Programs. *Operations Research*, vol. 5, n. 1.
- Geromel, J. C. and Ferreira, P. A. V., (1986). A Dual-Projection Method for Large Scale Systems Optimization. *Control-Theory and Advanced Technology*, vol. 2, n. 2.
- Haimes, Y. Y., Lasdon, L. S. and Wismer, D. A. (1971). On a Bicriterion Formulation of the Problems os Integrated System Identification and System Optimization. *IEEE Trans. on Systems, Man and Cybernetics*, july.

---

<sup>1</sup>Submetido em versão estendida à *Automatica*

- Haimes, Y. Y. and Wismer, D. A. (1972). A Computational Approach to the Combined Problem of Optimization and Parameter Identification. *Automatica*, vol. 8.
- Haimes, Y. Y. and Li, D., (1988). Hierarchical Multiobjective Analysis for Large-scale Systems: Review and Current Status, *Automatica*, vol. 24, n. 1.
- Hwang, C. L. and Masud, A. S. M., (1979). *Multiobjective Decision-Making Methods and Applications*. Lecture Notes in Economics and Mathematical Systems, Springer.
- Kambhupati, C. Tham, M. T., Mantague, G. A. and Morris, A. J. , (1992). Optimising Control of Fermentation Processes, *Proceedings of the IEE*, Part D, vol. 139, n. 1, january.
- Kümmel, M. and Seborg, D. (Editors), (1987). A Contemplative Stance for Chemical Process Control. *Automatica*, vol. 23, n. 6.
- Lasdon, L. (1971). *Optimization Theory for Large Systems*, McMillan.
- Luenberger, D., (1984). *Linear and Nonlinear Programming*, Addison Wesley.
- Luyben, W. D., (1973) *Process Modeling, Simulation and Control for Chemical Engineers*. McGraw-Hill Chemical Engineers Series.
- McGrew, D. R. and Haimes, Y. Y. (1974). Parametric Solution to the Joint System Identification and Optimization Problem. *Journal of Optimization Theory and Applications*, vol. 3, n. 5.
- Payne, M., Polak, E., Collins, D. C. and Miesel, W. S., (1975). An Algorithm for Bicriteria Optimization Based on the Sensitivity Function. *IEEE Trans. on Automatic Control*, vol. AC-20.
- Rarig, H. and Haimes, Y. Y., (1983) Risk/Dispersion Index Method. *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-13, n. 3.
- Roberts, P.D. (1979). An Algorithm for Steady-state System Optimization and Parameter Estimation. *International Journal of Systems Sciences*, vol. 10, n. 7.

- 
- Roberts, P.D. and Williams, T. W. C., (1981). On an Algorithm for Combined System Optimization and Parameter Estimation. *Automatica*, vol. 17, n. 1.
- Sorenson, H. W., (1980). *Parameter Estimation*. In Control and Systems Theory, vol 9, Marcel Dekker, INC.
- Tatjewski, P. and Roberts, P. D., (1987). Newton-like Algorithm for Integrated System Optimization and Parameter Estimation Technique. *Int. J. Systems Sci.*, vol. 46.
- Vachtsevanos, G. J., (1979).  $L^p$ -convergence of a Process Identification Algorithm. *Int. J. Systems Sci.* vol. 6, n. 5.
- Wismer, D. A. (1971). *Optimization Methods for Large Scale Systems ... and Applications*. McGraw-Hill.
- Yu, P. L., (1985). *Multiple-Criteria Decision Making: Concepts, Techniques and Extensions*. Plenum.

# Apêndice : Variáveis e Parâmetros dos Reatores

$C_A$  : concentração do componente A

$C_B$  : concentração do componente B

$F_0$  : fluxo de entrada do reagente

$F$  : fluxo de saída do reagente

$F_j$  : fluxo de água na jaqueta

$V$  : volume do reagente

$U$  : taxa de transferência de calor

$A_H$  : área de transferência de calor

$\lambda$  : calor exotérmico da reação

$\rho$  : densidade do reagente

$C_p$  : capacidade de calor do reagente

$k_i$  : taxa de reação

$\alpha_i$  : fator pré-exponencial

$E_i$  : energia de ativação

$R$  : constante do gás perfeito

$T$  : temperatura da massa de reação

$\lambda_1$  : calor exotérmico da reação  $A \rightarrow B$

$\lambda_2$  : calor exotérmico da reação  $B \rightarrow C$

$\rho$  : densidade da massa de reação

$C_p$  : capacidade de calor da massa de reação

$Q_M$  : Quantidade de calor do metal

$V$  : volume do recipiente

$T_M$  : temperatura da parede de metal

$T_j$  : temperatura da jaqueta

$h_i$  : coef. de transf. de calor da face interna da parede de metal

$h_o$  : coef. de transf. de calor da face externa da parede de metal

$A_i$  : área interna da parede de metal

$A_o$  : área externa da parede de metal

$\rho_M$  : densidade da parede de metal

$C_M$  : capacidade de calor da parede de metal

$V_M$  : volume da parede de metal

$T_M$  : temperatura da parede de metal

$V_j$  : volume da jaqueta

$\rho_s$  : densidade de vapor na jaqueta

$w_s$  : taxa de entrada de vapor

$w_c$  : taxa de condensação de vapor

$M$  : peso molecular do vapor

$P_j$  : pressão na jaqueta

- 
- $A_{vp}$  : constante de pressão-vapor da água  
 $B_{vp}$  : constante de pressão-vapor da água  
 $C_{vs}$  : capacidade da válvula de vapor  
 $X_s$  : fração aberta da válvula de vapor [0,1]  
 $X_w$  : fração aberta da válvula de água [0,1]  
 $Q_j$  : quantidade de calor na jaqueta  
 $h_{os}$  : entalpia do vapor na jaqueta  
 $H_s$  : entalpia do vapor de entrada  
 $h_c$  : entalpia do líquido condensado  
 $F_{w0}$  : fluxo de água  
 $T_{j0}$  : temperatura de entrada da água  
 $\rho_j$  : densidade da água na jaqueta  
 $C_j$  : capacidade de calor da água  
 $h_{ow}$  : entalpia da água na jaqueta  
 $C_{vw}$  : capacidade da válvula de água