1. Publicação nº	2. Versão	3. Data	5. Distribuição
INPE-2508-TDL/098		*Ago., 1982*	☐ Interna ☒ Externa ☐ Restrita
4. Origem *DRH-DMC*	Programa *FRH/CEA*		

6. Palavras chaves - selecionadas pelo(s) autor(es)

MÉTODOS NUMÉRICOS EM CONTROLE ÓTIMO — *CONTROLE SUBÓTIMO*
ESTIMAÇÃO LINEAR ÓTIMA — *OTIMIZAÇÃO DE TRAJETÓRIA*

7. C.D.U.: *681.5.015.24*

8. Título *INPE-2508-TDL/098*

IMPLEMENTAÇÃO E TESTE DE UM PROCEDIMENTO SUBÓTIMO PARA A OBTENÇÃO DO CONTROLE EM PROBLEMAS DE OTIMIZAÇÃO DINÂMICA

10. Páginas: *162*

11. Última página: *D.5*

12. Revisada por

Atair Rios Neto

9. Autoria *Luiz Antonio Waack Bambace*

13. Autorizada por

Nelson de Jesus Parada
Diretor

Assinatura responsável

14. Resumo/Notas

Este trabalho consiste na implementação e teste de um procedimento de busca direta para obter soluções numéricas subótimas, por restrição no controle, em problemas de otimização em malha aberta de sistemas dinâmicos. Adotando-se para aproximar o controle uma função que é caracterizada por um vetor de parâmetros, o problema fica reduzido a um problema de otimização de parâmetros. Combina-se um critério de busca direta com um esquema de perturbação linear e o Filtro de Kalman para achar o incremento de busca em cada passo. Escolheu-se, da mesma forma que a maioria dos autores da área de métodos subótimos, um problema de transferência de órbita Terra-Marte em tempo mínimo com trajetória plana e empuxo baixo e constante para testar o procedimento proposto. Os resultados obtidos mostram que o procedimento proposto é muito competitivo, quando comparado aos mais modernos métodos de solução numérica de problemas de otimização dinâmica, tanto em termos de precisão quanto em termos de tempo de computação e espaço de memória de computador requeridos.

15. Observações *Dissertação de Mestrado em Ciências Espaciais, aprovada em 12 de abril de 1982.*

ABSTRACT

This work consists of implementation and testing of a direct search procedure to obtain suboptimal numerical solutions, by control restriction, for open loop optimization of dynamic systems problems. By taking a control function approximation which is caracterized by a vetor of parameters, the dynamic problem is reduced to one of parameter optimization. A direct search criterion plus linear perturbation are combined with Kalman Filter to find search increment in each step. Following most of the authors working whith other suboptimal procedures, which can be considered by comparasion, a plane minimum time, low thrust Earth-Mars orbit transfer problem is considered to test numerical performance. Results obtained show that in terms of acuracy, computacional time and computer memory space, the procedure is very competitive when compared to most of the moderm methods used in the numerical optimization of dynamic systems.

Aprovada pela Banca Examinadora
em cumprimento a requisito exigido
para a obtenção do Título de Mestre
em Ciência Espacial

Dr.Celso de Renna e Souza — Presidente

Dr.Atair Rios Neto — Orientador

Dr.Wladimir Borgest — Membro da Banca -convidado-

Dr.Antonio Felix Martins Neto — Membro da Banca

Eng.Décio Castilho Ceballos, MSc. — Membro da Banca

Candidato: Luiz Antonio Waack Bambace

São José dos Campos, 12 de abril de 1982

À Denise e Andrea

AGRADECIMENTOS

Ao Instituto de Pesquisas Espaciais quero deixar agradecimentos pelo suporte financeiro e pelos incentivos dados à realização deste trabalho.

Ao Dr. Atair Rios Neto pelo encorajamento em abordar o tema e pela orientação deste trabalho.

À Denise, minha esposa, pela colaboração e incentivo.

Ao Dr. Santiago Alves Tavares e aos engenheiros Agenor de Toledo Fleury, Msc., e Décio Castilho Ceballos, Msc., e a todas as pessoas que direta ou indiretamente colaboraram na execução deste trabalho.

SUMÁRIO

Pág.

LISTA DE TABELAS ..... *xi*

LISTA DE SÍMBOLOS ..... *xii*

CAPÍTULO 1 - INTRODUÇÃO 1

CAPÍTULO 2 - FUNDAMENTOS 5

2.1 - Introdução 5

2.2 - Problema de otimização 5

2.3 - Controle ótimo e controle subótimo 6

2.4 - Colocação do problema de otimização dinâmica como um problema de otimização de parâmetros 11

2.5 - Alguns aspectos da solução numérica do problema restrito de otimização 12

2.6 - Aspectos da teoria de estimação 18

CAPÍTULO 3 - PROCEDIMENTO PROPOSTO 27

3.1 - Introdução 27

3.2 - Escolha do problema associado de minimização 27

3.3 - Escolha dos parâmetros que controlam a convergência 34

3.4 - Critérios de convergência 42

3.5 - Considerações sobre a convergência 54

CAPÍTULO 4 - ALGORITMO NUMÉRICO 57

4.1 - Introdução 57

4.2 - Descrição das etapas do algoritmo numérico 59

4.2.1 - Cálculo do incremento de busca no vetor de parâmetros ... 59

4.2.2 - Análise preliminar dos resultados 63

4.2.3 - Integração numérica e geração de derivadas parciais 73

4.2.4 - Análise final dos resultados 78

CAPÍTULO 5 - TESTE NUMÉRICO DO PROCEDIMENTO 83

5.1 - Introdução 83

5.2 - Colocação do problema de transferência de órbita Terra-Marte em tempo mínimo com trajetória plana e empuxo constante e baixo 83

Pág.

5.3 - Aproximação para o controle 86
5.3.1 - Introdução 86
5.3.2 - Aproximação do controle por interpolação lineares com tempos intermediários proporcionais ao tempo final e imposição de continuidade no controle 87
5.3.3 - Aproximação do controle por segmentos de reta não obrigatoriamente contínuos nos tempos nodais e com tempos nodais a otimizar 88
5.3.4 - Aproximação do controle por arcos de parábola não obrigatoriamente contínuos nos tempos nodais e com tempos nodais a otimizar 89
5.3.5 - Aproximação do controle por trechos de valores constantes a otimizar e com intervalos de duração também a otimizar . 90
5.4 - Resumo dos resultados obtido 91
5.5 - Conclusões 100

CAPÍTULO 6 - COMENTÁRIOS E CONCLUSÕES FINAIS 101

REFERÊNCIAS BIBLIOGRÁFICAS 107

APÊNDICE A - OBTENÇÃO DAS DERIVADAS PARCIAIS DO ESTADO EM RELAÇÃO AOS PARÂMETROS PELO MÉTODO DE GERAÇÃO DE SOLUÇÕES UNITÁRIAS

APÊNDICE B - COMPLEMENTAÇÃO DA DESCRIÇÃO DO ALGORITMO NUMÉRICO

APÊNDICE C - PARTICULARIDADES DAS SOLUÇÕES COM TEMPOS INTERMEDIÁRIOS A OTIMIZAR

APÊNDICE D - FÓRMULAS PARA O CÁLCULO DO NÚMERO DE OPERAÇÕES ARITMÉTICAS REQUERIDAS PARA O CÁLCULO DO INCREMENTO DE BUSCA POR DIVERSOS MÉTODOS SUBÓTIMOS

LISTA DE TABELAS

Pág.

5.1 - Resultados iniciais e finais do teste 1 93
5.2 - Histórico do teste 1 93
5.3 - Resultados iniciais e finais do teste 2 94
5.4 - Histórico do teste 2 94
5.5 - Resultados iniciais e finais do teste 3 95
5.6 - Histórico do teste 3 95
5.7 - Resultados iniciais e finais do teste 4 96
5.8 - Histórico do teste 4 96
5.9 - Resultados iniciais e finais do teste 5 97
5.10 - Histórico do teste 5 97
5.11 - Resultados iniciais e finais do teste 6 98
5.12 - Histórico do teste 6 98
5.13 - Resultados iniciais e finais do teste 7 99
5.14 - Histórico do teste 7 99
D.1 - Parâmetros de avaliação de custo D.4

LISTA DE SÍMBOLOS

a - Vetor a ser otimizado.

$A(t), B(t)$ - Matrizes ligadas a equações diferenciais lineares.

$C(x,u,t)$ - Função implícita usada na representação de vínculos de desigualdade na variável de controle.

$D_i(t)$ - Parcela da derivada do vetor de estado $x(t)$ em relação a i-ésima componente de a, que está relacionada com a influência deste parâmetro nas condições iniciais.

E' - Estimativa linear do valor de $E(\bar{a} + \Delta\hat{a})$.

E^* - Notação simplificada para $E(\bar{a} + \Delta\hat{a})$.

$\bar{e}_a$ - Tolerância para o valor das componentes do erro nos vínculos, exigida para se dar início à fase de busca do mínimo.

$E(a)$ - Erro nos vínculos de contorno associado a um dado valor do vetor de parâmetros a ser otimizado.

E_i - I-ésima componente de $E(a)$.

e_e - Erro normalizado.

e_m - Erro numérico desprezível.

E_p - Estimativa linear, feita na iteração anterior, de $E(a)$.

e_t - Tolerância para a razão entre os efeitos da não-linearidade nas variações de $G(x(t_f,a)$ e $E(a)$ e as variações desejadas nestas funções.

F_i - Diferença entre a matriz de projeção V_i e $P_0^{-1}P_i$.

f_{ea} - Razão entre a tolerância final nas componentes de $E(a)$ e $\bar{e}_a$.

$f(x,a,t), f(x,u,t)$ - Funções ligadas ao modelos do sistema dinâmico.

$G_{\bar{a}}$ - Derivada de $G(\bar{x}(t_f),\bar{a})$ em relação ao vetor a.

$G_{\bar{a}i}$ - I-ésima componente de $G_{\bar{a}}$.

$G_{\bar{a}0}$ - Valor de $G_{\bar{a}}$ na iteração anterior.

$G(x(t_f),a)$ - Função objetivo do problema restrito de otimização.

$G(x(t_0),t_0,x(t_f),tf)$ - Função objetivo do problema original.

h^* - Vetor linha que relaciona o vetor a com uma observação destinada a controlar a posição relativa dos tempos nodais.

H - Matriz que relaciona um vetor de observações qualquer com um vetor a ser estimado em um problema de estimação linear.

H, H', H'', H''' - Matrizes de derivadas parciais de vetores de observações fictícias em relação ao vetor a.

h_i, h_i', h_i'', h_i''' - I-ésimas linhas de H, H', H'', H''', respectivamente.

$h(x)$ - Função que relaciona um vetor de observações com um vetor a ser observado.

I' - Estiamtiva linear de $I(\bar{a} + \Delta\bar{a})$.

I'' - Notação abreviada para $I(\bar{a} + \Delta\bar{a})$.

$I_{\bar{a}}$ - Derivada de $I(\bar{a})$ em relação ao vetor a.

$I(a)$ - Função de erro nos vínculos de contorno no instante inicial do problema restrito de minimização.

I_i - I-ésima componente de $I(a)$.

I_p - Previsão linear, feita na iteração anterior, para $I(a)$.

IP - Índice de desempenho.

$IP^*, IP^{**}, IP^{***}, IP_w$ - Índices de desempenho de problemas associados de minimização.

$I(x(t_0),t_0)$ - Função de erro nos vínculos de contorno no instante inicial do problema original de otimização.

j - Número de componentes da funções vetoriais I(a) e $I(x(t_0),t_0)$.

k - Número de componentes do vetor de controle.

k - Número de intervalos em que se divide o intervalo $[t_0,t_f]$ para efeito de interpolação.

K - Ganho de Kalman.

$K(a,x(t_0))$ - Relação implícita entre $x(t_0)$ e a.

K_i - Ganho de Kalman usado para se incorporar a i-ésima componente do vetor de observações z.

ℓ - Número de componentes de $M(x(t_f),a)$ e $M(x(t_f), t_f)$.

m - Número total de vínculos $(j+\ell)$.

$\dot{m}$ - Fluxo de massa.

M' - Estimativa linear de $M(x'(t_f), \bar{a} + \Delta\bar{a})$.

M" - Notação abreviada para $M(x'(t_f),\bar{a} + \Delta\bar{a})$.

$M_{\bar{a}}$ - Derivada de $M(x(t_f),a)$ em relação ao vetor a.

m_i - I-ésima componente de $M(x(t_f),a)$.

M_p - Previsão linear feita na iteração anterior para $M(x(t_f),a)$.

m_0 - Massa inicial do foguete.

n_c - Número de componentes do vetor de controle.

neqn - Número de componentes do vetor de estado do problema de otimização a ser solucionado.

n_p - Número de parâmetros.

p - Peso do algoritmo.

P - Matriz de covariância do erro na estimativa final.

$\bar{P}$ - Matriz de covariância do erro na estimativa inicial.

$P, \bar{P}, \bar{P}', \bar{P}'', \bar{P}'''$ - Matrizes de pesos de possíveis problemas associados de minimização.

$\bar{P}, P_0$ - Matriz de covariância do erro na estimativa inicial do problema fictício de estimação correspondente.

P_i - Matriz de covariância do erro na estimativa obtida após se filtrar a i-ésima componente do vetor de observações.

$\bar{P}_i$ - Matriz de covariância do erro na estimativa obtida pela filtragem da i-ésima componente do vetor de observações.

$P_j(t), P_j'(t)$ - Parcelas da derivada do vetor de estado $x(t)$ em relação à i-ésima componente de a.

$\bar{P}_j(t)$ - Derivada do vetor de estado $x(t)$ em relação à i-ésima componente do vetor a.

q - Parâmetro ligado ao nível de ruído a ser adicional para manter a estabilidade do Filtro de Kalman.

Q - Matriz de covariância do ruído de estado.

Q - Matriz de ruído a ser adicionado.

q_i - Estimativa de q associada a i-ésima componente do vetor de observações.

R - Matriz de covariância do ruído de medidas.

R, R^*, R', R'', R'''	- Matrizes de pesos do método, análogas a matrizes de covariância de ruídos de observação.
R_i, R_{ii}	- I-ésimo termo da diagonal de R.
$S(t_a, t_b)$	- Matriz de transição de estado entre t_a e t_b.
$S_\ell(t_a, t_b)$	- L-ésima linha de $S(t_a, t_b)$.
t	- Tempo.
T	- Empuxo.
$t_i, t_{j,i}$	- Tempos nodais.
$u, u(t)$	- Função de controle.
$u(a,t)$	- Função de interpolação para o controle.
v	- Erro de medidas.
V_n	- Matriz que projeta um vetor qualquer em um hiperplano perpendiculr às n primeiras linhas de H.
$X_{\bar{a}}(t)$	- Derivada do vetor x(t) em relação ao vetor a.
x	- Vetor a ser estimado.
$x, x(t)$	- Vetor de estado do sistema dinâmico.
z, z^*, z', z'', z'''	- Vetores de observações.
α	- Peso dos métodos de Willianson e Ceballos.
$\alpha_g, \alpha_\ell, \alpha_{t_3}, \alpha_{qi}, \alpha_{\beta p}, \alpha_{\bar{\beta}}$	- Pesos do método proposto.
β	- Peso do método de Ceballos.
β	- Peso do método de Willianson.
$\bar{\beta}, \beta_i, \beta_p, \bar{\beta}_{pi}, \beta_{p0}, \beta_0$	- Pesos do método proposto.
Γ	- Matriz usada para descrever a influência do ruído de estado num processo estocástico.
Γ_{t_1, t_2}	- Matriz usada para descrever a influência do ruído de estado em um "step process".
Δa	- Incremento de busca no vetor de parâmetros.

$\Delta\bar{a}$ - Incremento de busca no vetor de parâmetros obtido pelo procedimetno proposto.

$\Delta\bar{\bar{a}}$ - Valor de partida para Δa, análogo a uma estimativa inicial.

Δa_i - Estimativa de Δa obtida após se filtrar a i-ésima componente do vetor de observações.

Δa_0 - Notação alternativa para $\Delta\bar{\bar{a}}$.

$\Delta\bar{G}$ - Estiamtiva da variação da função objetivo.

θ - Constante auxiliar para o cálculo da solução de partida.

$\lambda, \lambda(t)$ - Vetor de multiplicadores de Lagrange ligados a vínculos dinâmicos.

μ - Constante heliogravitacional.

ν_i - I-ésima componente de ν.

ν_i, ν_f - Vetores de multiplicadores de Lagrange.

$\sigma_{\Delta G}$ - Covariância da estimativa da variação da função objetivo.

τ - Tempo particular.

ÍNDICES INFERIORES

$_0$ - Condição inicial de uma variável.

$_f$ - Condição final de uma variável.

$_{\bar{a}}$ - Indica derivação em relação ao vetor a em torno das condições nominais.

ESPECIAIS

- O apóstrofo colocado sobre uma variável de estado indica que esta corresponde à aplicação do controle $u(\bar{a} + \Delta\bar{a}, t)$.

- O traço sobre uma variável indica que se refere a valores nominais (controle -u(a,t)).

- O traço precedido de uma barra indica que se refere aos valores nominais da iteração anterior.

CAPÍTULO 1

INTRODUÇÃO

A otimização de sistemas dinâmicos é de grande interes se em inúmeras áreas da ciência aplicada. Neste século houve um signi ficativo progresso nos métodos de otimização dinâmica principalmente após o advento do computador digital.

A resolução do problema de controle ótimo de sistemas dinâmicos consiste na obtenção de um vetor de controle que leva um sis tema controlável à satisfação de restrições inerentes ao problema e que ao mesmo tempo otimize uma dada função objetivo.

Na hipótese de se controlar o sistema em malha aberta, a obtenção do controle ótimo implica obter uma função do tempo, dita função de controle ótimo. Já no caso de se controlar o sistema em ma lha fechada, a obtenção do controle ótimo implica conhecer uma função do tempo e do vetor de estado do sistema, dita lei de controle ótimo.

Através do cálculo variacional pode-se resolver analiti camente problemas bastante simples de determinação do controle ótimo de sistemas dinâmicos, porém os problemas de interesse prático são em sua maioria, por demasiado complexos para permitir a obtenção de uma solução analítica, exigindo assim um tratamento numérico.

Dependendo do problema abordado e dos recursos computa cionais disponíveis pode ser necessário ou não introduzir simplifica ções no problema ao atacá-lo numericamente. Ao se proceder desta forma obtém-se um controle capaz de satisfazer às restrições impostas bem co mo levar a função de controle ótimo a um valor próximo ao ótimo. Essa abordagem será dita controle subótimo.

Tomou-se como objetivo inicial a implementação e teste de um procedimento para a obtenção de soluções subótimas para o proble

ma de obtenção do controle de malha aberta de sistemas dinâmicos proposto por Rios Neto (1979). Contudo, após implementar tal procedimento e efetuar com este vários testes numéricos, percebeu-se a conveniência de introduzir pequenas modificações no procedimento original.

O procedimento aqui proposto supõe que seja adotada para aproximar o controle, uma função do tempo e de um determinado conjunto de parâmetros, u(a,t), e visa obter para estes parâmetros valores que não só façam o sistema satisfazer aos vínculos do problema, na hipótese do sistema ser controlável, como também levem o índice de desempenho ao melhor valor possível. Desta forma o procedimento abordado neste trabalho pode ser classificado como subótimo por restrição no controle. O método aqui abordado é um método direto que faz uso do Filtro de Kalman para obter o incremento de busca no vetor de parâmetros a ser otimizado. Para obter as derivadas parciais do estado em relação aos parâmetros fez-se uso de uma versão do Método de Geração de Soluções Unitárias (Bryson and Ho, 1975) proposta por Rios Neto e Ceballos (1979).

Elaborou-se um programa de computador com o intuito de testar inicialmente o procedimento proposto empregando-o na solução do problema de transferência de órbita Terra-Marte em tempo mínimo com empuxo constante e baixo e trajetória plana, que é clássico em trabalhos de teste e comparação de procedimentos de otimização dinâmica como aqueles de Tapley e Lewallen (1967), Willianson (1971), Hull e Edgeman (1975), Ceballos (1979) e Rios Neto e Ceballos (1979).

No Capítulo 2 serão apresentados os fundamentos de controle ótimo de sistemas dinâmicos, o conceito de métodos numéricos ótimos e subótimos, uma comparação entre estes dois tipos de métodos e a classificação dos métodos ótimos, além de aspectos da teoria de estimação.

No Capítulo 3 é feito o desenvolvimento analítico do método proposto.

No Capítulo 4 é desenvolvido o algoritmo numérico.

No Capítulo 5 é feito o teste numérico do procedimento.

No Capítulo 6 serão apresentadas as conclusões e sugestões para desenvolvimentos futuros, bem como comentários de ordem qualitativa que visam situar o procedimento proposto em relação aos demais procedimentos ótimos e subótimos.

CAPÍTULO 2

FUNDAMENTOS

2.1 - INTRODUÇÃO

A menos que as equações dinâmicas, a função objetivo e as equações de restrição sejam bastante simples, faz-se necessária a utilização de métodos numéricos na resolução de problemas de otimização dinâmica, via de regra mediante o emprego de computadores digitais. Assim, o desenvolvimento de métodos numéricos de controle ótimo passou a ser bastante intenso na década de 60, quando surgiram procedimentos como o Método do Gradiente e a Programação Dinâmica. Durante toda esta década, foi dada uma grande ênfase aos chamados métodos ótimos, i.e. métodos que se destinam a achar a solução do problema de otimização dinâmica na sua forma original, sem introduzir simplificação alguma. Devido à grande complexidade dos métodos ótimos, passou-se a dar ênfase no início da década de 70 à obtenção de métodos numéricos mais simples, ditos métodos subótimos, mediante a introdução de simplificações no problema original de otimização dinâmica. Pode-se destacar como um dos mais importantes trabalhos pioneiros na área o trabalho de Willianson (1971).

Neste capítulo serão apresentados os elementos teóricos fundamentais para o desenvolvimento do procedimento proposto neste trabalho.

2.2 - PROBLEMA DE OTIMIZAÇÃO

O objetivo deste trabalho é implementar e aprimorar um procedimento para solucionar problemas de otimização do controle de malha aberta de sistemas dinâmicos que possam ser representados por sistemas de equações diferenciais ordinárias. Este tipo de problema pode ser escrito como a seguir.

Minimizar:

$$IP = G(x(t_0),\ t_0,\ x(t_f),\ t_f), \tag{2.1}$$

sujeito a vínculos dinâmicos dados por:

$$\dot{x} = f(x,\ u,\ t) \tag{2.2}$$

e a vínculos de contorno (ou terminais) dados por:

$$I(x(t_0),\ t_0) = 0, \tag{2.3}$$

$$M(x(t_f),\ t_f) = 0, \tag{2.4}$$

onde:

x é o vetor de estado $n \times 1$;
u é o vetor de controle $p \times 1$;
M é o vetor de vínculos de contorno no instante final $\ell \times 1$;
I é o vetor de vínculos de contorno no instante inicial $j \times 1$;
t é o tempo;
t_f é o instante final;
t_0 é o instante inicial.

Deve-se ressaltar que os procedimentos subótimos podem também ser utilizados, com pequenas alterações, na solução de problemas onde haja vínculos de desigualdade nas variáveis de controle (Rios Neto and Ceballos, 1981).

2.3 - CONTROLE ÓTIMO E CONTROLE SUBÓTIMO

Na solução numérica de problemas de otimização dinâmica pode-se fazer uso dos chamados métodos ótimos, que se destinam a resolver problemas de otimização desta natureza na sua forma original, e os métodos subótimos que são capazes de fornecer boas soluções para o pro

blema de otimização original mediante a solução de formas simplificadas deste.

Na solução dita subótima, por restrição no controle, adota-se, para aproximar o controle, uma função previamente estabelecida, cujos argumentos são o tempo e um determinado conjunto de parâmetros, doravante chamado vetor de parâmetros a ser otimizado ou vetor de incógnitas básico, e tenta-se achar um conjunto de parâmetros que faça com que os vínculos de contorno sejam satisfeitos. Ou seja, o problema fica como colocado a seguir:

dado $u = u(a,t)$, achar o vetor a para que o sistema dinâmico representado pela Equação 2.2, parta de uma condição inicial

$$x(t_0) = x_0$$

que satisfaça à Equação 2.3 e atinja uma condição final

$$x(t_f) = x_f$$

que satisfaça à Equação 2.4 e ao mesmo tempo torne o valor do índice de desempenho IP dado pela Equação 3.1 mínimo.

A elevada complexidade numérica dos métodos ótimos frequentemente limita ou torna desvantajosa a aplicação destes métodos na obtenção de soluções de problemas de interesse prático de obtenção do controle ótimo de malha aberta.

Dependendo do problema, dos recursos computacionais disponíveis e do método escolhido estas dificuldades podem ser:

- custo de computação elevado;
- insuficiência da capacidade do computador disponível;
- necessidade de uma boa solução de partida;
- dificuldade de implementação;
- dificuldade de realização em tempo real.

Afora estes problemas, muitas vezes é preferível, na implementação do controle no sistema físico, que o resultado seja fornecido na forma de uma função. Foi observado pelo autor em alguns testes que o ajuste da solução obtida de forma discretizada por meio de métodos ótimos, por uma determinada função, levou em geral a resultados piores que aqueles obtidos com um procedimento subótimo, onde o controle era aproximado diretamente pelo mesmo tipo de função. Provavelmente tal resultado se deve ao fato de ao se aproximar o controle por uma função qualquer, mediante o emprego de um método padrão de aproximação de funções, não se levar em conta o fato de o sistema dinâmico ser mais sensível aos erros no controle durante certas fases da trajetória do que em outras. Sem mencionar diretamente tal fato diversos autores, como por exemplo Ceballos (1979), apontam como motivo para a utilização de métodos subótimos a dificuldade de implementação da solução obtida por métodos ótimos e a decorrente dificuldade de duplicação dos resultados.

De acordo com os trabalhos de Willianson (1971), Hull e Edgeman (1975) e Ceballos (1979), os métodos subótimos, por restrição de controle, requerem um baixo tempo de computação por iteração, e requerem menos memória e são de implementação mais simples que os métodos ótimos. A velocidade de convergência e a precisão dos resultados dependerão da função de interpolação escolhida no caso dos métodos subótimos. Tal escolha é o ponto crítico na utilização de tais métodos. Uma boa escolha da função de interpolação permite com frequência obter precisões comparáveis àquelas dos métodos ótimos, e não raramente os métodos subótimos convergem mais rapidamente que os métodos ótimos, quando se escolhe, para aproximar o controle, uma função de interpolação conveniente. Até mesmo funções bastante simples podem levar a resultados satisfatórios. Segundo Ceballos (1979) para que uma função de interpolação seja satisfatória é necessário que:

a) o número de parâmetros seja grande o suficiente para permitir a satisfação dos vínculos e a minimização da função objetivo, sendo portanto maior que a soma do número de vínculos de contorno mais um;

b) a função de interpolação escolhida seja compatível com a forma da função de controle ótimo, i.e, seja capaz de reproduzir as principais características desta função; este requisito está ligado à precisão dos resultados obtidos;

c) os parâmetros sejam o mais independentes possível entre si;

d) a sensibilidade do controle em relação aos parâmetros seja preferivelmente uniforme;

e) se dê preferência a funções de interpolação que apresentem uma boa correspondência entre os desvios nos parâmetros e os desvios no controle.

Ainda no tocante à escolha da função de interpolação deve-se levar em conta que o número de parâmetros e a forma da função estão ligados com o tempo de integração e o tempo de resolução das equações de obtenção do incremento de busca, de forma que é interessante, do ponto de vista de redução de custo computacional, trabalhar com o menor número possível de parâmetros.

A interpolação da função de controle por segmentos de reta tem sido apontada por diversos autores como de grande utilidade, graças à sua elevada flexibilidade, quando o conhecimento sob a forma da função de controle ótimo é escasso. Não existe nenhum critério de escolha para a forma da função de interpolação, que na prática é feita por tentativa e erro. Aconselha-se o uso da aproximação por trechos com funções simples na fase inicial, por este tipo de aproximação ser bastante poderosa, para ao menos visualizar algumas características da função de controle ótimo. Pode-se também eventualmente usar as condições matemáticas para a optimalidade (Citron, 1969) para obter algumas características qualitativas da função de controle ótimo, contudo deve-se lembrar que neste caso está assumindo-se a hipótese de variação fraca na qual se apoia a derivação de tais equações.

De acordo com Tapley e Lewallen (1967), Bryson e Ho (1975) e Fleury (1980) os métodos numéricos de otimização podem ser classificados em diretos e indiretos; sendo que serão indiretos aqueles que fazem uso das condições necessárias requeridas pela matemática para a optimalidade e serão diretos aqueles que fazem uso apenas das equações do movimento do sistema dinâmico, das equações de definição dos vínculos de contorno e da equação de definição da função objetivo, além de um determinado critério de busca.

Os métodos ótimos indiretos são ainda classificados de acordo com quais condições necessárias são satisfeitas em cada iteração por imposição do método (Tapley and Lewallen, 1967; Bryson and Ho, 1975 e Fleury, 1980) em:

a) Métodos de Vizinhança Extrema;

b) Métodos de Perturbação ou Quasi-linearização.

A diferença entre os Métodos de Perturbação e Quasi-linearização está no fato do primeiro destes métodos fazer uso das equações diferenciais do sistema de forma não linearizada para a geração da trajetória de referência, enquanto os Métodos de Quasi-linearização usam equações diferenciais linearizadas para o mesmo fim (Tapley and Lewallen, 1967).

Exemplos de métodos diretos ótimos são o Método do Gradiente Clássico e o Método do Gradiente Conjugado. Exemplos de métodos ótimos indiretos são o Método das Funções de Perturbação, o Método de Perturbação das Equações Adjuntas, o Método de Newton-Raphson Generalizado e o Método de Quasi-linearização Modificado.

Exemplos de métodos numéricos subótimos são os métodos de Willianson (1971), Hull e Edgeman (1975), Ceballos (1979) e Rios Neto e Ceballos (1979). Todos estes métodos, à exceção do método de Hull e Edgeman, que é um método indireto (de segunda ordem), são diretos e de primeira ordem.

2.4 - COLOCAÇÃO DO PROBLEMA DE OTIMIZAÇÃO DINÂMICA COMO UM PROBLEMA DE OTIMIZAÇÃO DE PARÂMETROS

Dado um problema de controle ótimo da forma:

Minimizar:

$$IP = G(x(t_0),\ t_0,\ x(t_f),\ t_f) \tag{2.5}$$

sujeito a:

$$\dot{x} = f(x,\ u,\ t) \tag{2.6}$$

$$M(x(t_f),\ t_f) = 0 \tag{2.7}$$

$$I(x(t_0),\ t_0) = 0 \tag{2.8}$$

Supondo-se que se deseja obter uma solução subótima para a função de controle u(t), mediante o emprego de uma função de forma convenientemente estabelecida u(a,t) para aproximar o controle cujos argumentos são o tempo e um vetor de parâmetros a, pode-se escolher o vetor de parâmetros de forma tal que o problema colocado pelas Equações 2.5 a 2.8 possa ser colocado como:

Minimizar:

$$IP = G(x(t_0),\ a) \tag{2.9}$$

sujeito a:

$$M(x(t_f),\ a) = 0 \qquad (\ell \text{ equações}) \tag{2.10}$$

$$I(a) = 0 \qquad (j \text{ equações}) \tag{2.11}$$

$$\dot{x} = f(x,\ a,\ t) \qquad (n \text{ equações}) \tag{2.12}$$

onde podem ser eventualmente incluídas algumas condições iniciais e o tempo final como componentes do vetor de parâmetros a ser otimizado.

Doravante chamar-se-á ao problema colocado por 2.9 a 2.12 de problema restrito de otimização ou problema básico de otimização de parâmetros.

Deve-se notar que a função de interpolação escolhida pode introduzir novos mínimos locais no problema que não existem no problema original, e também eventualmente as funções $M(x(t_f), a)$ e $I(a)$ podem ter pontos estacionários que não existam nas funções $M(x(t_f),t_f)$ e $I(x(t_0), t_0)$. Assim, escolhida uma função de interpolação com um número de parâmetros que garanta a existência de solução para o problema básico de otimização de parâmetros, independentemente do procedimento subótimo usado, é possível que não se obtenha uma solução numérica satisfatória devido ao fato de a escolha da função de interpolação introduzir novos pontos de mínimo local ou pontos estacionários nas funções $M(x(t_f), a)$ e $I(a)$ não originariamente existentes em $M(x(t_f), t_f)$ e $I(x(t_0), t_0)$. É possível, ainda, que a causa do mau comportamento numérico seja a existência de pontos estacionários nas funções $M(x(t_f),t_f)$ e $I(x(t_0), t_0)$ originais.

2.5 - ALGUNS ASPECTOS DA SOLUÇÃO NUMÉRICA DO PROBLEMA RESTRITO DE OTIMIZAÇÃO

O problema restrito de otimização é, em geral, um problema de otimização de parâmetros não-linear, onde não se conhece uma relação explícita entre o valor da função objetivo e os parâmetros, nem tampouco entre o erro nos vínculos de contorno no instante final e os parâmetros. Todavia, existem vários métodos numéricos que permitem conhecer as derivadas parciais do vetor de estado em relação aos parâmetros em torno de uma trajetória de referência, de forma a se conhecer, a partir de tais derivadas, as derivadas parciais da função objetivo em relação aos parâmetros e também as derivadas parciais do erro nos vínculos de contorno em relação aos parâmetros. Assim pode-se expandir

Willianson (1971) resolve o problema restrito de otimização empregando um algoritmo iterativo que obtém o incremento de busca a e os vetores de multiplicadores de Lagrange ν_i e ν_f pela solução a cada iteração de:

Minizar:

$$IP_w = G_{\bar{a}}\Delta a + w\Delta a^T\Delta a/2 + \nu_i^T[I_{\bar{a}}\Delta a + \alpha I(\bar{a})] +$$

$$+ \nu_f^T[M_{\bar{a}}\Delta a + \alpha M(\bar{a}, x(t_f))] \qquad (2.16)$$

onde:

α é o decréscimo desejado no erro nos vínculos de contorno;

w é um peso positivo, usado para limitar o tamanho do incremento de busca e para garantir a existência de mínimo para o IP da Equação 2.16.

Ceballos (1979) também resolve o problema básico de otimização iterativamente, obtendo o incremento de busca a pela minimização de:

$$IP = \sum_{i=1}^{np} w_i|\Delta a_i| + w_{np+1}\, G_{\bar{a}}\Delta a \qquad (2.17a)$$

sujeito a:

$$M_{\bar{a}}\Delta a + \alpha M(\bar{x}(t_f), \bar{a}) = 0$$

$$I_{\bar{a}}\Delta a + \alpha I(\bar{a}) = 0 \qquad (2.17b)$$

$$G_{\bar{a}}\Delta a + \beta(|IP| + 1) \leq 0$$

onde:

α é o decréscimo relativo desejado no erro nos vínculos de contorno;

em série de Taylor tanto a função objetivo como as equações de restrição e com tais equações obter incrementos de busca que permitam caminhar no sentido de minimizar a função objetivo e satisfazer às restrições.

Chamando-se np ao número de parâmetros, verifica-se que para expandir em série de Taylor até a primeira ordem tanto as equações de restrição como a função objetivo, precisa-se gerar numericamente np derivadas do vetor de estado em relação aos parâmetros. Para uma expansão até a segunda ordem necessita-se conhecer $np + np^2$ destas derivadas, ao passo que para se expadir até a terceira ordem necessita-se conhecer $np + np^2 + np^3$ derivadas e assim por diante.

Desta forma, o custo computacional torna na maioria das vezes desinteressante o uso de métodos de ordem elevada. Mesmo o emprego de métodos de segunda ordem, como o de Hull e Edgeman (1975), em geral se restringe a casos onde se deseja refinar soluções obtidas com métodos de primeira ordem.

Obtida uma trajetória de referência qualquer x(t) mediante o uso de um controle u(a,t), define-se:

$$M_{\bar{a}} = \frac{\partial M(\bar{x}(t_f), \bar{a})}{\partial x(t_f)} \cdot \frac{\partial \bar{x}(t_f)}{\partial a} + \frac{\partial M(\bar{x}(t_f), \bar{a})}{\partial a} \qquad (2.13)$$

$$I_{\bar{a}} = \frac{\partial I(\bar{a})}{\partial a} \qquad (2.14)$$

$$G_{\bar{a}} = \frac{\partial G(\bar{x}(t_f), \bar{a})}{\partial x(t_f)} \cdot \frac{\partial \bar{x}(t_f)}{\partial a} + \frac{\partial G(\bar{x}(t_f), \bar{a})}{\partial a} \qquad (2.15)$$

Assim, $M_{\bar{a}}$ e $I_{\bar{a}}$ são respectivamente as matrizes de derivadas dos erros nos vínculos de cortorno no instante final e inicial em relação aos parâmetros, e $G_{\bar{a}}$ é o gradiente da função objetivo.

os w_i, para $1 \leq i \leq np$ e $i \in N$, são pesos que se destinam a penalizar incrementos de busca muito grandes;

w_{np+1} é um parâmetro usado para induzir o decréscimo da função objetivo;

β é um parâmetro destinado a controlar a variação da função objetivo.

O Problema 2.17 pode ser facilmente colocado na forma de um problema de programação linear e ser resolvido pelo algoritmo Simplex. Desta forma, o método de Ceballos é uma adaptação do método de programação não-linear de Wegner (Rosen, 1961) ao problema restrito de otimização.

Hull e Edgeman (1975) adicionaram os vínculos de contorno à função objetivo por meio de multiplicadores de Lagrange, ν_i e ν_f, e expandiram a função resultante em série de Taylor até a segunda ordem na variação dos parâmetros Δa, para obter uma nova função objetivo IP_H. Impondo-se que a derivada primeira desta nova função em relação a Δa, ν_i e ν_f sejam nulas, tais autores obtiveram um sistema de equações não-lineares para o cálculo de ν_i, ν_f e do incremento de busca Δa. Tais autores empregaram a seguir o método de Newton-Raphson na solução de tais equações.

Vários outros métodos de otimização de parâmetros (Luenberger, 1969; Hestenes and Stiefel, 1962; Powell, 1964; e Rosen, 1960 e 1961), poderiam ser usados para a solução do problema restrito de otimização, como por exemplo o Método da Projeção do Gradiente, o Método do Gradiente Reduzido, os Métodos de Direções Conjugadas entre os quais se destacam o Método do Gradiente Conjugado e o Método de Powell-Fletcher-Heeves, o Método de Direções Viáveis e o Método de Powell.

O uso do Método da Projeção do Gradiente implica achar uma solução viável, ou seja, uma solução capaz de satisfazer aos vínculos, e uma vez obtida tal solução deve-se caminhar na direção da projeção do gradiente da função objetivo sobre o hiperplano tangente aos vín

culos, fazendo-se a seguir uma interação com o objetivo de corrigir os vínculos.

O uso do Método do Gradiente Reduzido implicaria usar equações que garantam a redução dos vínculos, caso não se tenha obtido uma solução viável, ou as próprias equações de restrição para calcular parte das componentes do incremento no vetor de parâmetros em função das demais, substituindo-se o resultado em $G(x(t_f), a)$ e caminhando-se na direção do gradiente desta nova função.

Pode-se conceber variantes do Método da Projeção do Gradiente que achem o incremento de busca, combinando-se um vetor que se destina a reduzir o erro nos vínculos com a projeção do gradiente da função objetivo em um hiperplano tangente à função erro nos vínculos, no ponto em questão.

A idéia de penalização também permite conceber inúmeros métodos de solução do problema restrito de otimização.

Pode-se ver que os métodos numéricos subótimos de primeira ordem são diretos e se dividem em dois tipos:

a) métodos que conjuntamente reduzem o valor da função objetivo e o erro nos vínculos de contorno e que resultam em algoritmos com pelo menos dois parâmetros de ajuste;

b) métodos que alternam fases de redução do erro nos vínculos com fases de busca do mínimo da função objetivo.

Todos os métodos que visam à redução simultânea do erro nos vínculos e da função objetivo, independentemente de serem obtidos diretamente a partir do conceito de penalização ou não, apresentam a mesma dificuldade: o acerto dos parâmetros de ajuste ou pesos do algoritmo. Para um método deste tipo fornecer bons resultados na solução de um determinado problema de otimização dinâmica é necessário que o algoritmo numérico decorrente não tenha um desempenho demasiadamente

sensível aos valores numéricos destes pesos, nem tampouco praticamente insensível aos valores destes mesmos pesos, pelo fato de em tais casos se tornar difícil a obtenção de valores para estes parâmetros que resultem em bom desempenho numérico do algoritmo. Também é desejável que se saiba como relacionar variações nos parâmetros de ajuste com mudanças no comportamento do algoritmo e ainda, se possível, que se tenham correlações que levem a bons valores de partida para tais pesos.

Estes aspectos discutidos no parágrafo anterior serviram de motivação para se tentar resolver o problema restrito de otimização pela minimização de uma norma semelhante àquela usada por Gauss para a solução do problema de estimação linear, e que está ligada ao Filtro de Kalman. Tal escolha permite que se faça uso de conceitos e métodos que são oriundos da teoria de estimação para o ajuste dos parâmetros que controlam a convergência. Reforça esta motivação a grande eficiência de um método de otimização de funções proposto por Powell (1964) que está baseado no método dos mínimos quadrados.

Rios Neto (1979) foi o primeiro a propor a utilização do Filtro de Kalman no cálculo do incremento de busca no vetor de incógnitas básico, em problemas restritos de otimização dinâmica (ver as Equações 2.9 a 2.12).

O presente trabalho não só está baseado no trabalho de Rios Neto (1979) como pretendia inicialmente ser apenas um trabalho de implementação e teste do procedimento proposto por este autor. Introduziram-se contudo, com base na experimentação numérica, pequenas modificações no procedimento original com o intuito de melhorar o desempenho numérico do algoritmo, principalmente no tocante à eficiência e à confiabilidade. Tal aumento de eficiência e confiabilidade se deve ao fato de as alterações introduzidas permitirem uma melhoria do esquema heurístico de acerto dos pesos do algoritmo, embora a custa da introdução de mais três pesos suplementares. A estratégia de busca utilizada pouco difere da apresentada no trabalho de Rios Neto, conservando-se aqui as principais hipóteses e conceitos daquele trabalho.

Conforme se verá na Seção 3.2, há uma certa analogia entre o problema de estimação linear e o problema básico de otimização de parâmetros.

2.6 - ASPECTOS DA TEORIA DE ESTIMAÇÃO

Apresentam-se nesta seção os aspectos da teoria de estimação mais relevantes para o desenvolvimento deste trabalho. Inicia-se esta seção com a apresentação do problema geral de estimação e do problema de estimação linear. Apresenta-se a seguir a norma utilizada por Gauss para a solução do problema de estimação linear, as fórmulas do Filtro de Kalman e algumas noções de filtro adaptativo.

O problema geral de estimação pode ser colocado como o problema de obter uma estimativa $\hat{x}$, do vetor aleatório x e uma matriz P de covariância do erro na estimativa $\hat{x}$, que exprimam da melhor maneira possível o comportamento do vetor aleatório x, dado que se conhece um vetor de observações z, não necessariamente com o mesmo número de componentes, relacionado com o vetor a ser estimado por:

$$z = h(x) + v \tag{2.18}$$

onde v é um vetor de ruídos com média $\bar{v}$ e matriz de covariância R conhecidas, e pode-se também conhecer eventualmente uma estimativa a priori $\bar{x}$ da variável aleatória x e a matriz $\bar{P}$ de covariância no erro na estimativa $\bar{x}$. Para o caso particular em que:

$$z = Hx + v \tag{2.19}$$

onde

v é um vetor de ruídos com média $\bar{v}$ e matriz de covariância R conhecidas;

H é uma matriz de dimensões apropriadas;

tem-se o chamado problema de estimação linear.

Existem várias soluções para o problema geral de estimação dependendo de que propriedades se exige da estimativa $\hat{x}$, de hipóteses feitas sobre a distribuição da variável aleatória x, sobre a distribuição dos ruídos de medidas v, e assim por diante. Critérios para derivar bons estimadores podem ser encontrados em vários textos didáticos da área, como por exemplo Jazwinski (1969) ou Liebelt (1967).

No caso deste trabalho o interesse se encontra em um caso particular do problema de estimação linear onde:

- o ruído de medida tem média nula e covariância R conhecida;
- se conhece uma estimativa a priori $\bar{x}$ e a matriz $\bar{P}$ de covariância do erro nesta estimativa;
- tanto o ruído de medida ν como a variável aleatória x têm o mesmo tipo de distribuição, por hipótese gaussiana ou uniforme.

Gauss solucionou este tipo de problema através do método dos mínimos quadrados ponderados minimizando:

$$IP = (\bar{x} - \hat{x})^T P^{-1}(\bar{x} - \hat{x}) + (z - H\hat{x})^T R^{-1}(z - H\hat{x}) \qquad (2.20)$$

onde:

x - é o vetor a ser estimado $n_p \times 1$;
z - é o vetor de observações $m \times 1$;
$\bar{x}$ - é a estimativa a priori de x;
P e R - são matrizes de pesos definidas positivas;

para obter expressões para o cálculo da nova estimativa de x.

No caso de P ser a matriz de covariância do erro na estimativa inicial e R ser a matriz de covariância do ruído de medida v, tais expressões de cálculo da nova estimativa serão equivalentes às do filtro de Kalman.

Vários métodos foram empregados por diversos autores para derivar as expressões do filtro de Kalman ou filtro linear ótimo, evidenciando na dedução destas diferentes propriedades deste filtro (Bryson and Ho, 1975; Jazwinski, 1969; Liebelt, 1967).

Liebelt (1967) apresenta o Filtro de Kalman através das seguintes expressões:

$$K = \bar{P}H^T(H\bar{P}H^T + R)^{-1}$$

$$\hat{x} = \bar{x} + K(z - H\bar{x}) \qquad (2.21)$$

$$P = \bar{P} - KH\bar{P}$$

que mediante o emprego de identidades de matrizes podem ser reescritas como (Jazwinski, 1969):

$$K = \bar{P}H^T(H\bar{P}H^T + R)^{-1}$$

$$\hat{x} = \bar{x} + K(z - H\bar{x}) \qquad (2.22)$$

$$P = (I - KH)\,\bar{P}(I - KH)^T + KRK^T,$$

ou ainda serem reescritas na forma como Bryson e Ho (1975) apresentam o filtro de Kalman:

$$P = \bar{P} - \bar{P}H^T(H\bar{P}H^T + R)^{-1}\,H\bar{P}$$

$$K = PH^TR^{-1} \qquad (2.23)$$

$$\hat{x} = \bar{x} + K(z - H\bar{x}).$$

Sob o ponto de vista teórico tais expressões são equivalentes, porém têm diferenças significativas em termos do condicionamento numérico. As Expressões 2.22 para a atualização da matriz de cova

riância tendem a conservar melhor a simetria e a positividade definida da nova matriz de covariância, além de fazer com que os erros nesta matriz sejam da ordem do quadrado do erro no cálculo do ganho de Kalman K. Diversos autores apresentam o Filtro de Kalman por meio de expressões diversas destas aqui apresentadas, e tais formas de apresentação foram em geral desenvolvidas com vistas a obter expressões para o filtro de Kalman de melhor condicionamento numérico. Outros exemplos de expressões do filtro de Kalman são o "filtro de raíz quadrada" e o "filtro de matriz inversa". Acredita-se que para os objetivos deste trabalho tais formas de apresentação do Filtro de Kalman sejam desnecessárias, uma vez que os erros de modelagem, linearização e propagação das equações diferenciais são provavelmente de ordem muito superior aos erros numéricos cometidos durante a aplicação do filtro. Deve-se também ter em mente que as formas mais precisas do Filtro de Kalman exigem uma maior quantidade de cálculos que aquelas mais simples.

Cabe também frisar que quando as componentes do ruído de medidas não são correlacionadas entre si e o ruído de medidas não está correlacionado com o vetor a ser estimado, pode-se aplicar o Filtro de Kalman em qualquer de suas formas de apresentação de modo sequencial, i.e, processando-se as componentes do vetor de observações uma a uma, o que elimina a necessidade de inversão de matrizes.

Outros aspectos da teoria de estimação importantes para o desenvolvimento deste trabalho são as noções de filtro adaptativo e divergência do Filtro de Kalman. Tais noções serão úteis para desenvolver um modelo para a matriz P e para determinar as medidas z, bem como para desenvolver o esquema de acerto dos parâmetros de ajuste.

Considere-se um sistema dinâmico modelado pela equação diferencial de primeira ordem

$$\dot{x}(t) = f(x(t), t) + \Gamma(t)\, w(t) \qquad (2.24)$$

onde $w(t)$ é um vetor de ruídos. Supondo-se dados uma estimativa a priori $\bar{x}(t_1)$ do vetor aleatório $x(t)$, a matriz de covariância $P(t_1)$ do erro nesta estimativa e um vetor de observações obtido num instante diferente de t_1, que será chamado de $z(t_2)$, deseja-se obter uma nova estimativa $\bar{x}(t_2)$ e uma nova matriz de covariância $P(t_2)$ do erro nesta estimativa. Chama-se a este problema de problema de estimação dinâmica. Um caso particular é o problema de estimação dinâmica em sistemas lineares, ou seja, sistemas que podem ser modelados por equações diferenciais da forma:

$$\dot{x} = A(t)\, x(t) + \Gamma(t)\, w(t) \tag{2.25}$$

onde

$A(t)$ e $\Gamma(t)$ são funções matriciais do tempo;
$w(t)$ é um ruído dependente do tempo.

Pode-se dividir a solução do problema de estimação dinâmica em duas partes. Propagar no tempo a estimativa $\bar{x}(t_1)$ e a matriz de covariância $P(t_1)$ para obter uma nova estimativa $\bar{x}(t_2)$ e uma nova matriz $\bar{P}(t_2)$, que serão consideradas respectivamente a estimativa a priori e a matriz de covariância do erro na estimativa a priori para a segunda parte do problema que é a solução de um problema estático de estimação.

No caso de se trabalhar com um estimador não-viciado, como é o caso do Filtro de Kalman, convém exigir que a estimativa a priori $\bar{x}(t_2)$ também seja não-viciada, ou seja, que possua a propriedade de:

$$E\{\bar{x}(t_2)\} = E\{x(t_2)\} \tag{2.26}$$

onde se denota por $E\{y\}$ a esperança da variável aleatória y, ou seja,

$$E\{y\} = \int_{-\infty}^{\infty} y\, p\{y\}\, dy \tag{2.27}$$

Para sistemas representados por equações diferenciais da Forma 2.25, pode-se definir uma matriz $S(t, t_0)$, chamada matriz de transição de estado com as seguintes características:

$$S(t, t_0) = \frac{\partial x(t)}{\partial x(t_0)}$$

$$S(t_0, t_0) = I \qquad (2.28)$$

$$\dot{S}(t, t_0) = A(t)\, S(t, t_0)$$

Supondo-se que o ruído tenha média nula e que $\hat{x}(t_1)$ e $\bar{x}(t_2)$ sejam não-viciadas (Jazwinski, 1969) calcula-se $\bar{x}(t_2)$ por:

$$\bar{x}(t_2) = S(t_2, t_1)\, \hat{x}(t_1) \qquad (2.29)$$

Por meio do uso do método de Lagrange de solução de equações diferenciais não homogêneas pode-se ter uma expressão para $x(t_2)$, e com a definição da matriz $\bar{P}(t_2)$ de covariância do erro na estimativa $\bar{x}(t_2)$ pode-se obter a seguinte expressão para o cálculo desta matriz:

$$\bar{P}(t_2) = S(t_2, t_1)\, P(t_1)\, S^T(t_2, t_1) + \int_{t_1}^{t_2} \int_{t_1}^{t_2} S(t_2, s)\, \Gamma(s)\, E\{w(s) w^T(r)\} \Gamma^T(r)\, S^T(t_2, r)\, dr\, ds \qquad (2.30)$$

Para o caso onde se modela o ruído de estado $w(t)$ como constante entre t_1 e t_2 e se supõe que:

$$E\{x(t_1)\, w(t)\} = 0 \quad \text{p/ } t > t_1$$

tem-se que

$$\bar{P}(t_2) = S(t_2, t_1)\, P(t_1)\, S^T(t_2, t_1) + \Gamma_{t_1, t_2}\, Q\, \Gamma^T_{t_1, t_2} \qquad (2.31a)$$

onde

$$\Gamma_{t_1;t_2} = \int_{t_1}^{t_2} S(t_2, t_1)\, \Gamma(t)dt \qquad (2.31b)$$

Um dos problemas encontrados na solução de problemas de estimação dinâmica é justamente o modelamento do ruído de estado w(t). Em geral não se tem qualquer idéia do comportamento de tal erro. Caso se adotem para tais erros valores muito grandes dar-se-á pouco crédito à estimativa a priori e muita importância à última medida, deteriorando-se a qualidade da estimativa obtida. Caso se adotem valores muito baixos para tal ruído tem-se um problema muito mais grave, a divergência do filtro, pois a matriz de covariância decresce rapidamente fazendo com que o filtro passe a ignorar as novas observações, e a estimativa obtida deixe de ser uma boa aproximação para a variável aleatória a ser estimada. Convém ter em mente que sempre existirão erros em qualquer modelo dinâmico. O fenômeno de divergência caracteriza-se pela inconsistência estatística entre os resíduos de medida definidos por:

$$r \triangleq z(t_2) - H\bar{x}(t_2) \qquad (2.32)$$

ou seja, ocorre divergência quando o quadrado do resíduo se torna significativamente maior que a sua covariância.

Jazwinski (1969) tenta contornar este problema, modelando, por simplicidade, o ruído w(t) como constante entre t_1 e t_2 e como não correlacionado com o estado inicial, e supondo que a matriz Q da Equação 2.31 seja da forma:

$$Q = q^2 I \qquad (2.33)$$

onde

q é um número real;
I é a matriz identidade;

e calcula para cada componente do vetor de ruído um valor de q_i^2 mediante a imposição de que o quadrado da i-ésima componente do resíduo seja igual à esperança do quadrado desta componente, ou seja, que

$$E(r^2) = r^2 \tag{2.34}$$

obtendo portanto q_i^2 por meio de

$$q_i^2 = \frac{z_i^2 - 2z_i^T h_i x + x^T h_i^T h_i x - h_i S(t_2,t_1) P(t_1) S^T(t_2,t_1)^T h_i - R_i}{h_i \Gamma_{t_1,t_2} \Gamma^T_{t_1,t_2} h_i^T} \tag{2.35}$$

onde

z é a i-ésima componente do vetor de observações;
h é a i-ésima linha de H, definida por meio de:

$$H^T = [h_1^T, h_2^T, h_3^T, \ldots, h_m^T]$$

ou então, caso o membro da direita de 2.35 seja negativo, por meio de

$$q_i^2 = 0$$

e recomenda que se use q^2 como:

a) o máximo dos q_i^2; ou

b) a média dos q_i^2.

Deve-se notar que para o caso em que se tem uma única media e w(t) e x(t) são gaussianos, maximizar a probabilidade de ocorrência de uma realização do resíduo equivale a dizer que:

$$E(r^2) = r^2$$

Quanto à escolha de q, verifica-se que a primeira das duas opções é claramente a favor da segurança e impede a ocorrência do

fenômeno da divergência. A segunda opção, por outro lado, permite a obtenção de estimativas melhores, porém permite também que o resíduo ao quadrado se torne vez por outra ligeiramente inconsistente com a sua covariância; quando isto ocorre tem-se uma deterioração momentânea da qualidade da estimativa obtida, e imediatamente a seguir ocorre um aumento do nível de ruído adicionado que impede a ocorrência de divergência.

CAPÍTULO 3

PROCEDIMENTO PROPOSTO

3.1 - INTRODUÇÃO

Neste capítulo descrever-se-á o procedimento proposto, que é iterativo de primeira ordem, onde se calcula o incremento de busca no vetor de parâmetros pela solução de um outro problema de minimização, dito problema associado de minimização.

Propõem-se inicialmente três formas de problema associado, mostrando-se a relação entre tais problemas e três problemas de estimação linear ótima. Discute-se brevemente sobre alguns aspectos destas possíveis formas de solução, selecionando-se uma dentre estas formas de problemas associado e a seguir discute-se a escolha dos parâmetros que controlam a convergência.

3.2 - ESCOLHA DO PROBLEMA ASSOCIADO DE MINIMIZAÇÃO

Apresentaram-se na Seção 2.5 alguns motivos pelos quais se acredita que a solução, de forma iterativa, do problema restrito de otimização dinâmica, através da minimização de uma norma quadrática semelhante àquela ligada ao Filtro de Kalman, deve dar bons resultados. Sugere-se inicialmente três formas para esta norma, que diferem entre si basicamente pela forma como se provoca o decréscimo da função objetivo.

Um possível problema associado de minimização seria:

Minimizar

$$IP^* = (\Delta\tilde{a} - \Delta\bar{a})^T \bar{P}^{-1}(\Delta\tilde{a} - \Delta\bar{a}) + (z - H\Delta\tilde{a})^T R^{-1}(z - H\Delta\tilde{a}) \qquad (3.1)$$

onde

$$\Delta\bar{a} = -p G_{\bar{a}}^T$$

$$Z^T = -[M^T(\bar{x}(t_f), \bar{a}) : I^T(\bar{a})]\rho^T$$

$$H^T = [M_{\bar{a}}^T : I_{\bar{a}}^T]$$

sendo que:

- ρ - é uma matriz diagonal principal cujos termos não-nulos são os decréscimos desejados nas respectivas componentes do erro nos vínculos de contorno;
- p - é o parâmetro que estabelece a fração do gradiente da função objetivo que será adotada como estimativa a priori;
- R e $\bar{P}$ - são matrizes definidas positivas;
- $M_{\bar{a}}$ e $I_{\bar{a}}$ - são as matrizes de derivadas parciais do erro nos vínculos de contorno em relação aos parâmetros (ver as Equação 2.13 a 2.15);
- $G_{\bar{a}}$ - é o gradiente da função objetivo (ver as Equações 3.15);
- $\Delta\hat{a}$ - é o incremento no vetor de parâmetros que minimiza a norma 3.1, por definição o incremento de busca no vetor de parâmetros.

Este problema é equivalente ao problema de:

Dada uma estimativa a priori da variável aleatória Δa

$$\Delta\bar{a} = -p G_{\bar{a}}^T = \Delta a + w \qquad (3.2)$$

onde w é um ruído de média nula e matriz de covariância $\bar{P}$

e dadas medidas z definidas por:

$$z^T = -[M^T(\bar{x}(t_f), \bar{a}) : I^T(\bar{a})]\rho^T = -(\rho\, E(\bar{a}))^T \qquad (3.3)$$

relacionadas com o vetor a ser estimado por

$$z = H\Delta a + v \tag{3.4}$$

onde

v é um ruído de média nula e matriz de covariância R, não correlacionado com w;

H é uma matriz definida por:

$$H^T = [M_{\bar{a}}^T : I_{\bar{a}}^T] \tag{3.5}$$

obter uma estimativa $\Delta\hat{a}$ da variável aleatória Δa que seja não-viciada e de variância mínima. Esta primeira forma de solução difere da solução proposta por Rios Neto (1979) apenas pelo fato de que nesta ρ é um escalar enquanto no procedimento proposto é uma matriz.

Cabe ressaltar que expandindo-se em série de Taylor o erro nos vínculos de contorno E, definido na Equação 3.3, obtém-se:

$$E(a) = E(\bar{a}) + H\Delta a + o(\Delta a^2) \tag{3.6}$$

Assim, caso a estimativa do incremento de busca satisfaça à Equação 3.4, ter-se-á uma previsão linear do novo erro nos vínculos E', dada por:

$$E' = E(\bar{a}) + H\Delta a + v \tag{3.7}$$

Desta forma pode-se associar ao ruído v os termos de ordem superior à primeira expansão em série de Taylor do erro nos vínculos de contorno somados aos erros numéricos de integração, os quais afetam não apenas o cálculo do erro nos vínculos de contorno como também o cálculo da matriz H.

O vetor de observações z definido na Equação 3.3 tem por objetivo induzir o decréscimo do erro nos vínculos de contorno, e o uso da estimativa a priori definida na Equação 3.2 tem por objetivo provo

car, quando isso for possível, um decréscimo na função objetivo, pois ao se solucionar o problema colocado pelas Equações 3.2 a 3.5 obtém-se uma estimativa $\Delta\bar{a}$ de Δa que procura satisfazer às medidas z e, na medida do possível, respeitar a informação a priori.

Outra forma de solucionar o problema restrito de otimização seria a minimização de:

$$IP^{**} = \Delta\bar{a}^T \bar{P}^{-1}\Delta\bar{a} + (z' - H'\Delta\bar{a})^T R'^{-1}(z' - H'\Delta\bar{a}) \qquad (3.8)$$

onde

$$z'^T = \left[-[M^T(\bar{x}(t_f), \bar{a}) : I^T(\bar{a})]\rho^T : - pG_{\bar{a}}\ G_{\bar{a}}^T\right]$$

$$H'^T = [M_{\bar{a}}^T : I_{\bar{a}}^T : G_{\bar{a}}^T]$$

e R' e $\bar{P}'$ são matrizes definidas positivas.

Este novo problema de estimação é equivalente a um novo problema fictício de estimação linear ótima descrito por:

dado uma estimativa a priori $\Delta\bar{a}$ da variável aleatória Δa, tal que:

$$\Delta\bar{a} = \Delta a + w = 0 \qquad (3.9)$$

onde w é um ruído de média nula e matriz de covariância $\bar{P}'$,

e dado um vetor de observações z' definido por:

$$z'^T = \left[-[M^T(\bar{x}8t_f), \bar{a}) : I^T(\bar{a})]\rho^T : -pG_{\bar{a}}\ G_{\bar{a}}^T\right] \qquad (3.10)$$

relacionado com o vetor a ser estimado a por:

$$z' = H'\Delta a + v \qquad (3.11)$$

onde

v é um ruído de média nula e matriz de covariância R' não correlacionada com o ruído w;

H' é uma matriz definida por:

$$H'^{T} = [M_{\bar{a}}^{T} : I_{\bar{a}}^{T} : G_{\bar{a}}^{T}] \qquad (3.12)$$

obter uma estimativa $\Delta\hat{a}$ da variável aleatória Δa.

Pode-se notar que o Problema Associado 3.1 difere de 3.8 basicamente pela forma de se tentar provocar decréscimos na função objetivo. Na primeira destas formas tenta-se provocar o decréscimo da função objetivo através de um termo que tem, no problema fictício de estimação linear equivalente, o significado de uma estimativa a priori, enquanto no Problema Associado 3.8 tenta-se provocar o decréscimo da função objetivo através de um termo com o significado de uma observação. Teoricamente ambas as formulações são semelhantes, caso a componente do ruído de medidas associada à componente do vetor de medidas que se destina a provocar o decréscimo da função objetivo não seja correlacionada com as demais, ou seja, os elementos da última coluna e da última linha de R' são nulos, à exceção do termo da diagonal principal.

Obviamente, dada a semelhança dos problemas associados de minimização propostos com o problema de estimação linear ótima, ao se solucionar os Problemas 3.1 e 3.8 obter-se-ão expressões para o cálculo de Δa equivalentes àquelas do Filtro de Kalman.

Impondo-se que R e R' sejam diagonais pode-se, em ambos os casos, tratar o vetor de observações sequencialmente, eliminando-se a necessidade de efetuar inversões de matrizes; tal escolha torna os Problemas Associados 3.1 e 3.8 teoricamente semelhantes, porém do ponto de vista computacional existem algumas diferenças entre estas duas formas de solução. Por um lado, a Forma 3.1 permite obter o incremento de busca com uma menor quantidade de cálculos do que a forma 3.8. Esta por sua vez apresenta a vantagem de não exigir a repetição do processamento de todas as observações, quando se efetua uma má escolha do ter

mo que visa provocar o decréscimo da função objetivo, além de permitir que se adie a escolha deste termo para depois do processamento das demais componentes do vetor de observações.

Estes aspectos serviram de motivação para o uso de uma terceira norma. Tenta-se provocar o decréscimo da função objetivo não só com um termo com o significado de uma estimativa a priori, como eventualmente com termos com o significado de medidas cuja eventual aplicação e valor serão determinados com base na análise dos resultados do processamento das componentes anteriores do vetor de observações. Esta terceira norma teria a forma:

$$IP^{***} = (\Delta\hat{a} - \Delta\bar{a})^T \bar{P}^{''-1}(\Delta\hat{a} - \Delta\bar{a}) + (z'' - H''\Delta\hat{a})^T\bar{R}''(z'' - H''\Delta\hat{a}) \quad (3.13)$$

onde

$$\Delta\bar{a} = -pG_{\bar{a}}^T$$

$$z'''^T = \left[-[M^T(\bar{x}(t_f), \bar{a}) : I^T(\bar{a})]p : z'''\right]$$

$$H'''^T = [M_{\bar{a}}^T : I_{\bar{a}}^T : H'''^T]$$

onde z''' é o vetor de decréscimo adicionais na função objetivo cujo número de componentes z_i''', obrigatoriamente negativas, será determinado com base na análise dos resultados do processamento das observações anteriores, conforme discutir-se-á no Apêndice B, e onde H''' é uma matriz de s linhas, todas iguais a $G_{\bar{a}}$.

Dada a escolha de $\bar{R}''$ como uma matriz diagonal que permite tratar as observações z'' sequencialmente, pode-se imaginar ao se resolver o Problema Associado 3.13 que se deseja de fato resolver o Problema Associado 3.1. Uma vez obtida e analisada tal solução processar-se-á ou não observações adicionais z''', definidas pelas equações 3.13, cujo objetivo é provocar um maior decréscimo da função objetivo, quan

do a análise dos resultados obtidos pela solução do Problema Associado 3.1, com um dado conjunto de parâmetros de ajuste, mostrar que se escolheu um valor muito pequeno para o parâmetro p.

Assim, pode-se dizer que a solucão básica proposta é solucionar o problema restrito de otimização pela solução do Problema Associado 3.1, sendo, quando necessário, processadas medidas adicionais z_i''' de valor negativo relacionadas com o vetor a ser estimado Δa por:

$$z_i''' = G_{\bar{a}} \Delta a + v''' \tag{3.14}$$

onde v''' é um ruído de média nula que será suposto uniformemente distribuído no intervalo $[-e_m, e_m]$. A escolha de z_i''' será discutida no Apêndice B.

Optou-se pelo emprego das equações do Filtro de Kalman na forma proposta por Bryson e Ho (1975), dada à maior familiaridade do autor com estas (Equações 2.23), que no caso de serem aplicadas sequencialmente fornecem as relações

$$\Delta a_1 = \Delta \bar{a}$$

$$\bar{P}_1 = \bar{P}$$

$$\bar{P}_{i+1} = \bar{P}_i - \bar{P}_i h_i^T (R_i + h_i \bar{P}_i h_i^T)\, h_i \bar{P}_i \tag{3.15}$$

$$K_i = \bar{P}_{i+1}\, h_i R_i^{-1}$$

$$\Delta a_{i+1} = \Delta a_i + K_i (z_i - h_i \Delta a_i)$$

onde m é igual ao número de medidas e

$$H^T = [h_1^T : h_2^T : \ldots ; h_m^T] \tag{3.16}$$

R_i é o i-ésimo termo da diagonal principal de R;

e z_i é a i-ésima componente do vetor de medidas, ou seja,

$$z_i = h_i \Delta a + v_i$$

e

$$\Delta \hat{a} = \Delta a_{m+1} \tag{3.17}$$

3.3 - ESCOLHA DOS PARÂMETROS QUE CONTROLAM A CONVERGÊNCIA

A escolha dos parâmetros que controlam a convergência está baseada nos trabalhos de Rios Neto (1979 e 1980). O objetivo de tal escolha é fazer com que se dê uma certa prioridade à redução do erro nos vínculos, enquanto todas as suas componentes não estiverem contidas numa faixa de largura $\bar{e}_a$. Após isto deve-se exigir que se caminhe no sentido de minimizar a função objetivo, sem contudo permitir que o erro nos vínculos deixe de ficar contido numa nova faixa de largura $f_{ea}.\bar{e}_a$, onde f_{ea} deve ser maior ou igual a um e $\bar{e}_a$ positivo.

Com o intuito de permitir a aplicação sequencial do Filtro de Kalman escolheu-se a matriz R do Problema Associado 3.1, a partir da hipótese de que o ruído de medidas tenha média nula, componentes não-correlacionadas e distribuição uniforme numa faixa de largura $2e_m$, como sendo:

$$R = (e_m^2/3)I \tag{3.18}$$

onde

I é a matriz identidade de dimensão $(j + \ell)(j + \ell)$;

e_m é um número da ordem da estimativa do erro global cometido na integração numérica, que será chamado erro numérico desprezível.

Escolhido e_m e calculada a matriz R, calcula-se a matriz de covariância $\bar{P}$ e a matriz de correcões do erro nos vínculos de contorno ρ. Nos demais procedimentos diretos a correção nos vínculos é fixada arbitrariamente, independentemente do decréscimo desejado na função objetivo ou do tamanho máximo do passo. No procedimento proposto introduzem-se correlações empíricas para o cálculo de ρ a partir do valor de p e da estimativa e_m dos erros de integração. Para derivar tais correlações lançar-se-á mão de conceitos e métodos da área de estimação dinâmica e filtros adaptativos.

Pode-se imaginar o Problema de Otimização 3.1, análogo ao problema de estimacão linear colocado pelas Equacões 3.2 a 3.5, como um caso particular de um problema de estimacão dinâmica, onde a matriz de transição de estado entre dois "instantes" quaisquer (na verdade entre dois valores consecutivos de um parâmetro indexador de estágios consecutivos) é a identidade e onde dada uma estimativa $\Delta\hat{\bar{a}}(t_1)$ supostamente não-viciada do incremento de busca Δa

$$\Delta\hat{\bar{a}}(t_1) = -p\bar{G}_{\bar{a}}^T \tag{3.19}$$

e também se conhece a matriz de covariância $P(t_1)$ do erro na estimativa $\Delta\hat{\bar{a}}(t_1)$, cujo valor é extraído da hipótese de que tal erro tenha componentes não-correlacionadas e que suas componentes sejam uniformemente distribuídas no intervalo $[-p|\bar{G}_{\bar{a}}|, p|\bar{G}_{\bar{a}}|]$, ou seja:

$$P(t_1) = p^2/3.\bar{G}_{\bar{a}}.\bar{G}_{\bar{a}}^T.I \tag{3.20}$$

onde I é a matriz identidade de dimensão $n_p \times n_p$.

A partir da hipótese de que o ruído de estado possa ser modelado por um "step process", ou seja, possa ser suposto constante entre t_1 e t_2, e da hipótese de que tal ruído tenha média nula, e também fazendo-se a exigência de que a estimativa a priori $\Delta\bar{a}(t_2)$, anterior ao processamento das observacões $x(t_2)$, seja não-viciada, pode-se calcular $\Delta\bar{a}(t_2)$ por meio da Equação 2.29 e a matriz $\bar{P}(t_2)$ de covarián

cia do erro na estimativa $\Delta\bar{a}(t_2)$ por meio da Equação 2.31, chegando-se a:

$$\Delta\bar{a}(t_2) = \Delta\hat{a}(t_1) = -p\underline{G}_{\bar{a}}^T$$
$$\bar{P}(t_2) = (p^2/3)\ \underline{G}_{\bar{a}}.\underline{G}_{\bar{a}}^T\ I + Q \qquad (3.21)$$

onde

I é a matriz identidade de dimensão $n_p \times n_p$;
Q é a matriz de covariância do ruído de estado de dimensão $n_p \times n_p$.

Pretende-se neste trabalho obter correlações entre o valor de p e o valor dos termos ρ_{ii} da diagonal da matriz ρ mediante o emprego de conceitos de filtro adaptativo. Deve-se notar que no caso de se utilizarem tais conceitos para o cálculo de ρ, deve-se escolher arbitrariamente a matriz Q, que não necessita ser obrigatoriamente o produto de um escalar positivo pela matriz identidade. Para a obtencão de um valor de referência que indique a ordem de grandeza de Q, supor-se-á que esta matriz é da forma $q^2.I$, e analisar-se-á as expresssões resultantes da aplicacão do filtro adaptativo de Jazwinski (1969).

Caso se conhecesse o vetor de observações z, pode-se-ia calcular, para cada medida z_i, um valor de q_i^2 que garanta a consistência estatística do resíduo de medida (Equação 2.35), por meio de (Jazwinski, 1969):

$$q_i^2 = \frac{z_i^2 + 2ph_i\underline{G}_{\bar{a}}^T z_i + p^2\underline{G}_{\bar{a}}h_i^T h_i\underline{G}_{\bar{a}}^T - e_m^2/3 - p^2h_ih_i^T\underline{G}_{\bar{a}}\underline{G}_{\bar{a}}^T/3}{h_ih_i^T} \qquad (3.22a)$$

se tal expressão não resultasse em um valor negativo de q_i^2, ou, em caso contrário, por:

$$q_i^2 = 0 \qquad (3.22b)$$

Observando-se esta expressão, pode-se ver que um valor de q_i^2 igual a $2p^2G_{\bar{a}}G_{\bar{a}}^T/3$ garante que, se tal expressão for utilizada em sentido inverso para achar z_i, existirão raízes reais e distintas para z_i. Por outro lado, convém que a matriz de covariância da estimativa a priori tenha um termo que garanta uma boa capacidade de correlação nos vínculos de contorno, mesmo quando eventualmente por impossibilidade de decrescer a função objetivo ocorrerem valores muito baixos ou nulos de p. Assim, achou-se que no caso de usar Q como uma matriz igual ao produto de um escalar positivo pela matriz identidade, uma boa escolha de q^2 seria escolher este como a média ou o máximo dos q_i^2 dados por:

$$q_i^2 = 2e_m^2/(3h_ih_i^T) + 2p^2G_{\bar{a}}G_{\bar{a}}^T/3 \qquad (3.23)$$

Como quanto maior o valor de q^2 maiores podem ser as correções nos vínculos de contorno, decidiu-se escolher q^2 como o máximo dos q_i^2 durante a fase de convergência para os vínculos, e como a média dos valores de q_i^2, durante a fase de busca de mínimo. Pode-se notar que a parcela $2p^2G_{\bar{a}}G_{\bar{a}}^T/3$ é comum a todos os q_i^2, de forma que pode-se definir e_e^2 como:

$$e_e^2 = [q^2 - 2p^2G_{\bar{a}}G_{\bar{a}}^T/3]/2 \qquad (3.24)$$

e escrever a matriz $\bar{P}(t_2)$ resultante desta escolha de Q como:

$$\bar{P}(t_2) = 2e_e^2 + p^2G_{\bar{a}}G_{\bar{a}}^T.I \qquad (3.25)$$

onde I é a matriz identidade.

A escolha da matriz Q como uma matriz diagonal qualquer implicaria ter uma matriz $\bar{P}(t_2)$ também diagonal, e por não se ter a necessidade de restringir a forma da matriz Q utilizada para a dedução da Equação 3.25, optou-se por calcular a matriz $\bar{P}$ definida na Equação 3.1 por meio de:

$$\bar{P}_{ii} = 2\beta_i^2 e_e^2 + \beta_{pi}^2 . p^2 G_{\bar{a}} G_{\bar{a}}^T \qquad p/i=1,n$$

$$\bar{P}_{i,j} = 0 \qquad p/i\neq 1 \quad e \quad i,j=1,np \tag{3.26}$$

e dividir o vetor de parâmetros a ser otimizado em dois grupos. O primeiro deles formado pelos parâmetros a_i, para os quais a correspondente componente do gradiente da função objetivo é identicamete nula; tal grupo será chamado grupo de parâmetros não-correlacionados com a função objetivo. O segundo grupo, que será chamado grupo dos parâmetros correlacionados com a função objetivo, é formado pelos demais parâmetros. Os coeficientes β_i e β_{pi} para os parâmetros pertencentes ao grupo de parâmetros não-correlacionados com a função objetivo serão calculados por:

$$\beta_i = \beta_0$$

$$\beta_{pi} = \beta_{p0} \tag{3.27}$$

e os coeficientes β_i e β_{pi} dos parâmetros pertencentes ao outro grupo de parâmetros serão calculados por:

$$\beta_i = \bar{\beta}$$

$$\beta_{pi} = \bar{\beta}_p \tag{3.28}$$

De uma certa forma, uma covariância grande indica uma grande incerteza quanto ao valor da variável, de forma que o Filtro de Kalman procurará, na medida do possível, alterar menos os termos dos parâmetros aos quais correspondam menores covariâncias e mais os parâmetros aos quais correspondam maiores covariâncias. Assim, é razoável exigir que os valores de $\bar{\beta}$, $\bar{\beta}_p$, β_0 e β_{p0} satisfaçam às seguintes desigualdades:

$$\beta_p \geqq \bar{\beta}_p$$

$$\beta_0 \geqq \bar{\beta} \tag{3.29}$$

$$\beta_{p_0} \geq 1$$

$$\beta_0 \geq 1$$

Fixados $\bar{\beta}$, β_0, β_{p_0}, deve-se exigir que $\bar{\beta}_p$ seja suficientemente grande para que seja possível existir uma raiz real para z, fato que fica assegurado pela exigência de que:

$$e_m^2/3 + h_i \bar{P} h_i^T \geq p^2 G_{\bar{a}} h_i^T h_i G_{\bar{a}}^T \qquad (3.30)$$

conforme será visto mais adiante.

De acordo com a Equação 2.32, o vetor de resíduos pode ser definido como:

$$r = z - H\Delta\bar{a} \qquad (3.31)$$

e suas componentes r_i serão dadas por:

$$r_i = z_i - h_i \Delta\bar{a} \qquad (3.32)$$

Impondo-se, da mesma forma que foi feito por Jazwinski (1969) para derivar seu filtro adaptativo, que haja a igualdade entre o quadrado das componentes r_i do vetor de resíduos e as covariâncias de tais componentes, tem-se que:

$$E\{[h_i(\Delta a - \Delta\bar{a}) + v_i]^2\} = (\rho_{ii} E_i + p h_i G_{\bar{a}}^T)^2 \qquad (3.33)$$

como porém

$$\begin{aligned} z_i &= h_i \Delta a + v_i = -\rho_{ii} E_i \\ \Delta\bar{a} &= -pG_{\bar{a}}^T = \Delta a + w \end{aligned} \qquad (3.34)$$

onde E_i é a i-ésima componente do erro nos vínculos de contorno E(a) definido na Equação 3.3, tem-se:

$$\rho_{ii}^2 E_i^2 - 2\rho_{ii} ph_i G_{\bar{a}} E_i + p^2(h_i G_{\bar{a}}^T) = E(-ph_i G_{\bar{a}}^T - h_i w + v_i + h_i p G_{\bar{a}}^T)$$

Devido às hipóteses de que v_i tenha média nula e covariância igual a $e_m^2/3$ e de que w tem média nula e matriz de covariância $\bar{P}$, e de que v_i e w são não correlacionadas, tem-se que:

$$\rho_{ii}^2 E_i^2 - 2\rho_{ii} E_i ph_i G_{\bar{a}}^T + p^2(h_i G_{\bar{a}}^T) = e_m^2/3 + h_i \bar{P} h_i^T \qquad (3.35)$$

A exigência de que os parâmetros $\bar{\beta}$, $\bar{\beta}_p$, β_0 e β_{p0} sejam tais que a Equação 3.30 seja satisfeita garante que a Equação 3.35 tenha duas raízes reais e distintas, uma negativa e outra positiva, doravante chamada $\bar{\rho}_{ii}$. Caso os modelos adotados para as distribuições das observações e da informação e priori fossem corretos esta raíz positiva seria teoricamente o maior valor de ρ_{ii} para o qual pode-se garantir que não ocorrerá divergência do Filtro de Kalman. Para não permitir que durante a fase de captura dos vínculos se deteriore muito a função objetivo, e também por segurança, para garantir que não ocorra divergência ou quebra de linearidade, resolveu-se calcular ρ_{ii} por meio de:

$$\begin{aligned} \rho_{ii} &= 1/2\ \bar{\rho}_{ii} \qquad \text{se } \bar{\rho}_{ii} \leq 2 \\ \rho_{ii} &= 1 \qquad\qquad\ \text{se } \bar{\rho}_{ii} > 2 \end{aligned} \qquad (3.36)$$

onde $\bar{\rho}_{ii}$ é a raíz positiva da Equação 3.35.

Cabe ressaltar que a ocorrência de divergência do Filtro de Kalman implica o fato de que este não levará em conta as últimas componentes do vetor de medidas ao processar-se tal vetor sequencialmente. Assim, a divergência provavelmente resultará em uma estimativa que não diminua, ou até mesmo aumente, o valor em módulo das últimas componentes do erro nos vínculos de contorno, independentemente de

ser ou não possível reduzir tais valores, uma vez que o filtro ignora, como já foi dito, as últimas observações.

A não divergência do Filtro de Kalman não implica necessariamente que a estimativa do incremento no vetor de parâmetros seja aceitável. Valores numéricos mal condicionados de $G_{\bar{a}}$, $M_{\bar{a}}$ e $I_{\bar{a}}$ podem resultar em valores das componentes do vetor de observações conflitantes entre si ou com a observação a priori. Quando isto ocorrer, embora os valores obtidos para os resíduos sejam consistentes com suas estatísticas obtêm-se incrementos de busca que não levam à redução da estimativa do erro nos vínculos ou à residução do valor da função objetivo, ou que façam com que as estimativas do erro nos vínculos deixem de ficar contidas na faixa de tolerância especificada. Pode ocorrer também que se viole a hipótese de perturbação linear, e os efeitos da não-linearidade provoquem um aumento da função objetivo ou do erro nos vínculos de contorno. Pode ocorrer ainda que a causa do mal comportamento numérico seja uma má escolha dos parâmetros de controle do algoritmo. Desta forma, é necessário efetuar-se uma série de verificações quanto à qualidade do incremento de busca no vetor de parâmetro obtido e montar um esquema heurístico de ajuste dos parâmetros de controle do algoritmo: p, $\bar{\beta}$, $\bar{\beta}_p$, β_0, β_{p0} e e_m. Tais verificações e este esquema heurístico de ajuste dos parâmetros serão abordados no Capítulo 4.

As expressões obtidas para o cálculo de ρ e P não levaram à redução do número de parâmetros de ajuste mas acredita-se que elas levem à redução da necessidade de alterações no valor destes parâmetros e do número de rejeições por má escolha dos parâmetros de ajuste. Acredita-se também que elas permitam que tais parâmetros não oscilem de ordem de grandeza e façam com que o desempenho do algoritmo numérico não seja nem muito sensível nem demasiadamente insensível ao valor dos parâmetros de ajuste. A analogia entre a solução proposta e o problema de estimação linear ótima permite estabelecer um esquema mais eficiente de ajuste dos parâmetros de controle do algoritmo.

Destaca-se aqui o fato de que não se apresenta uma prova formal de convergência do procedimento proposto neste trabalho, porém acredita-se, com base na experimentação numérica, que ele seja viável como procedimento de otimização numérica de sistemas dinâmicos.

3.4 - CRITÉRIOS DE CONVERGÊNCIA

Os critérios de convergência (ou de parada) propostos nesta Seção não são definitivos, sendo que o estabelecimento de critérios de convergência eficientes tanto sob o ponto de vista de confiabilidade como de custo adicional de computação requerido é um dos pontos em aberto para desenvolvimentos futuros.

Vários aspectos devem ser considerados no tocante ao estabelecimento de critérios de convergência satisfatórios. Dentre estes aspectos, destaca-se o fato de as matrizes de derivadas do erro nos vínculos de contorno no instante final e, eventualmente, o gradiente da função objetivo, serem obtidos a partir das derivadas do vetor de estado no instante final em relação aos parâmetros, sendo estas últimas obtidas por integração numérica. Assim sendo ocorrem incertezas significativas ao se usar um critério de convergência baseado na ortogonalidade entre o gradiente da função objetivo e o hiperplano tangente aos vínculos de contorno, ou no fato de o gradiente da função objetivo ser nulo quando a solução do problema restrito de otimização é um ponto de mínimo local da função objetivo que também satisfaz aos vínculos de contorno. A não linearidade do problema também contribui para o aumento desta incerteza. Assim, existe uma probabilidade significativa de se ter encontrado a solução, quando os vínculos de contorno são satisfeitos e a projeção do gradiente da função objetivo no hiperplano tangente às restrições for da ordem dos erros de integração. Porém, tal fato não garante que realmente se tenha convergido para a solução do problema básico de otimização de parâmetros. Deve-se ressaltar também que não se conhece de fato os erros de integração, tendo-se apenas uma estimativa eventualmente falha de sua ordem de grandeza. Por outro lado, se o problema for altamente não-linear pode-se eventualmente estar bas

tante próximo da solução sem contudo se verificar a ortogonalidade entre o gradiente da função objetivo e o hiperplano tangente às restrições, ou então ter um valor do gradiente da função objetivo quase nulo. Contudo, apesar destes problemas, o uso de critérios deste tipo é uma boa opção sob o ponto de vista de confiabilidade.

Devido, também, aos erros de integração numérica e do modelo do sistema dinâmico não tem sentido exigir erros nos vínculos de contorno muito interiores às incertezas criadas pelos erros de modelagem e integração numérica. Assim, serão considerados satisfatórios erros nos vínculos de contorno contidos numa faixa de tolerância de largura

$$\varepsilon = \bar{e}_a f_{ea} \quad , \quad f_{ea} \geq 1 \wedge \bar{e}_a > 0 \tag{3.37}$$

onde $\bar{e}_a$ e f_{ea} foram definidos na Seção 3.3,

conforme indica a própria filosofia da escolha dos parâmetros que controlam a convergência.

Do ponto de vista teórico pode-se ter duas situações distintas no tocante ao comportamento do gradiente da função objetivo no ponto do espaço de parâmetros $a = a^*$, solução do problema restrito de otimização:

a) o ponto a^*, é um ponto de mínimo global ou local da função objetivo que está na região do espaço de parâmetros onde os vínculos de contorno são satisfeitos, de forma que G, neste ponto, é nulo;

b) o ponto a^* não é um ponto de mínimo global ou local da função objetivo, de forma que o gradiente da função objetivo neste ponto não é nulo e sim normal a um hiperplano tangente aos vínculos de contorno, ou seja, o hiperplano diferencial do erro nos vínculos de contorno.

Assim, segundo Rosen (1960), ter-se-ia neste ponto que os valores exatos do gradiente da função objetivo, $G_{\bar{a}}$, e da matriz H definida em 3.5 seriam tais que:

$$[I - H^T(HH^T)^{-1} H]\, G_{\bar{a}}^T = 0 \qquad (3.38)$$

pois a matriz $V_{\ell+j}$ definida por:

$$V_{\ell+j} = [I - H^T(HH^T)^{-1}H] \qquad (3.39)$$

é, segundo este autor uma matriz que projeta um vetor qualquer do espaço de parâmetros $\mathbb{R}^{np}$, onde $np > j+\ell$, no hiperplano $J^{\ell+j}$ tangente a todas as restrições e, portanto, normal a todos os vetores h_i definidos na Equação 3.16, que são os gradientes das componentes do erro nos vínculos de contorno.

Uma desvantagem do uso de tal critério no caso de se utilizar de um algoritmo no procedimento proposto é o fato de a quantidade de cálculos requerida para determinar a matriz $V_{\ell+j}$ ser próxima à quantidade de cálculos requerida para solucionar o próprio Problema Associado de Minimização 3.1. Já no caso do método da projecão do gradiente, tal matriz é usada para a própria determinação do incremento de busca. No entanto, existem certos pontos em comum entre o procedimento proposto e o método da projeção do gradiente, para o caso em que a matriz $\bar{P}$ é o produto de um escalar positivo pela matriz identidade, conforme tentar-se-á ilustrar nas páginas seguintes, onde ficará claro que a matriz de projeção pode ser obtida pela própria rotina de aplicação do Filtro de Kalman, usando-se uma matriz $\bar{P}$ igual à identidade e um ruído de medidas nulo.

Para ilustrar tal fato introduzir-se-á uma nova notação na qual se designará por:

P_0 - matriz de covariância do erro na estimativa a priori;

P_i - para i=1, j+ℓ - matriz de covariância do erro na estimativa obtida após se filtrar a i-ésima medida;

V_n - matriz que projeta um vetor qualquer num hiperplano perpendicular a todos os h_i, definidos na Equação 3.16, tais que $i \leq n$;

K_i - ganho de Kalman usado para se filtrar a i-ésima componente do vetor de medidas:

R_i - covariância da i-ésima componente do vetor de medidas;

Δa_0 - estimativa a priori;

Δa_i - estimativa obtida após se filtrar até a i-ésima componente do vetor de medidas.

Das expressões do Filtro de Kalman, tem-se que:

$$P_1 = P_0 - P_0 h_1^T h_1 P_0 (h_1 P_0 h_1^T + R_1)^{-1} \qquad (3.40)$$

Supondo-se que R é muito pequeno face a $h_1 P_0 h_1^T$ e expandindo-se em série de Taylor até a primeira ordem

$$(h_1 P_0 h_1^T + R_1)$$

em torno de R_1 igual a zero tem-se

$$P_1 = P_0[I - h_1^T h_1 (h_1 h_1^T)^{-1}] + (h_1^T h_1)(h_1 h_1^T)^{-2} R_1 \qquad (3.41)$$

devido ao fato de P_0 ser da forma $\alpha^2 . I$. A matriz V_1 (Rosen, 1960) pode ser escrita como

$$V_1 = [I - h_1^T h_1 (h_1 h_1^T)^{-1}] \qquad (3.42)$$

portanto a primeira parcela do membro da direita da Equação 3.41 é o produto da matriz de projeção V_1 pela matriz de covariância de erro na estimativa inicial, e a segunda parcela é uma matriz de erros F_1 cujos termos são da ordem de grandeza da variância do ruído na primeira medida.

Novamente das expressões do Filtro de Kalman tem-se que:

$$P_2 = P_1[I - h_2^T h_2 P_1 (h_2 P_1 h_2^T + R_2)^{-1}] \tag{3.43}$$

Usando-se o valor de P_1 fornecido pela Equação 3.41 e lembrando-se que P_0 é o produto de um escalar positivo α^2 pela matriz identidade, pode-se reescrever a Equação 3.43 como:

$$P_2 = [P_0 V_1 + F_1][I - h_2^T h_2 (P_0 V_1 + F_1)(h_2 P_0 V_1 h_2^T + R_2 + h_2 F_1 h_2^T)^{-1}] \tag{3.44}$$

Tal expressão pode ser reescrita como:

$$\begin{aligned} P_2 = {} & P_0[V_1 - V_1 h_2^T h_2 V_1 P_0 (h_2 P_0 V_1 h_2^T + R_2 + h_2 F_1 h_2^T)^{-1}] + F_1 - \\ & - (P_0 V_1 h_2^T h_2 F_1^T + F_1 h_2^T h_2 V_1 P_0 + F_1 h_2^T h_2 F_1^T)(h_2 P_0 V_1 h_2^T + R_2 + \\ & + h_2 F_1 h_2^T)^{-1} \end{aligned} \tag{3.45}$$

Supondo-se que h_1 e h_2 são linearmente independentes, e expandindo-se em série de Taylor

$$(h_2 P_0 V_1 h_2^T + R_2 + h_2 F_1 h_2^T)^{-1}$$

e $h_2 F_1 f_2^T$ iguais a zero, chega-se à seguinte expressão para P_2:

$$P_2 = P_0[V_1 - V_1 h_2^T h_2 V_1 (h_2 V_1 h_2^T)^{-1}] + F_2 \tag{3.46}$$

onde F_2 é uma matriz cujos termos são da ordem da variância do ruído de medidas v, dada por:

$$\begin{aligned} F_2 = {} & V_1 H_2^T h_2 V_1 (h_2 V_1 h_2^T)^{-2} (R_2 + h_2 F_1 h_2^T) + F_1 + \\ & + (F_1 h_2^T h_2 F_1 + P_0 V_1 h_2^T h_2 F_1 + F_1 h_2^T h_2 V_1 P_0) \times \\ & \times [(R_2 + h_2 F_1 h_2^T)(h_2 V_1 h_2^T P_0)^{-2} - (h_2 V_1 h_2^T P_0)^{-1}] \end{aligned} \tag{3.47}$$

Por propriedade da matriz de projeção, tem-se que:

$$h_2 V_1 h_2^T = (h_2 V_1)(h_2 V_1)^T$$

que permite reescrever a Equação 3.46 como:

$$P_2 = P_0\{V_1 - V_1 h_2^T h_2 V_1 [(h_2 V_1)(h_2 V_1)]^{-1}\} + F_2 \qquad (3.48)$$

Ocorre que segundo Rosen (1960):

$$V_{n+1} = V_n - V_n h_{n+1}^T h_{n+1} V_n [(h_{n+1} V_n)(h_{n+1} V_n)^T]^{-1} \qquad (3.49)$$

Desta forma tem-se que:

$$P_2 = P_0 V_2 + F_2 \qquad (3.50)$$

Desprezando-se os termos da ordem de $R_1 R_2$, R_1^2 e R_2^2 tem-se que:

$$F_2 = F_1 + (V_1 - V_2)(R_2 + h_2 F_1 h_2^T)[(h_2 V_1)(h_2 V_1)^T]^{-1} -$$

$$- (V_1 h_2^T h_2 F_1 + F_1 h_2^T h_2 V_1)[(h_2 V_1)(h_2 V_1)^T]^{-1} \qquad (3.51)$$

Repetindo-se esta última série de operações pode-se concluir que enquanto as matrizes de ruído forem suficientemente pequenas face à P_0, e que as n primeiras linhas da matriz H definida na Equação 3.5 sejam linearmente idependentes, tem-se que:

$$P_n = P_0 V_n + F_n \qquad (3.52)$$

e que enquanto os temos de segunda ordem nas matrizes de ruído F não forem relevantes face à ordem de grandeza da variância do ruído de medidas, pode-se calcular, desprezando-se termos de segunda ordem, a matriz F por meio de:

$$F_n = F_{n-1} + \frac{(V_{n-1}-V_n)(R_n+h_nF_{n-1}h_n)-(V_{n-1}h_n^Th_nF_{n-1}+F_{n-1}h_n^Th_nV_{n-1})}{(h_nV_{n-1})(h_nV_{n-1})} \qquad (3.53)$$

Pode-se notar desta expressão que, quando h_n forma um ângulo muito pequeno com alguma combinação linear das linhas anteriores de H, ocorre um aumento muito grande da diferença entre o produto da matriz de projeção V_n pela matriz de covariância do erro na estimativa inicial e a matriz de covariância P_n. Contudo, quando isto acontece têm-se, em geral, problemas numéricos tanto ao se calcular a matriz de projeção pela Relação Recursiva 3.49 como pela sua definição. No caso de se calcular a matriz de projeção recursivamente, pode-se contornar em parte esses problemas considerando-se que h_n é coincidente com tal combinação linear e fazendo-se V_n igual a V_{n-1}. Lembrando-se que a matriz H é obtida a partir de integração numérica, pode-se verificar que, dependendo do condicionamento numérico do problema abordado, a matriz de projeção correta do problema pode não ter qualquer semelhança com a matriz de projeção obtida mediante as relações recursivas propostas por Rosen (1960). No caso de P_0 ser proporcional à matriz identidade, e a matriz V'_n, dada por:

$$V'_n = P_0^{-1}P_n \qquad (3.54)$$

não ser uma boa aproximação para a matriz de projeção V_n, muito provavelmente as relações recursivas propostas por Rosen (1960) também não fornecerão uma boa aproximação para a matriz de projeção V_n.

Contudo analisando-se as Equações 3.40 a 3.53, pode-se verificar que teoricamente a matriz de covariância P_n seria a própria matriz de projeção, caso as linhas de H fossem linearmente independentes; a matriz de covariância inicial, a própria matriz identidade; e todas as medidas, consideradas exatas. Desta forma, pode-se em tese utilizar as rotinas de computação desenvolvidas para aplicar o Filtro de Kalman para executar a obtenção recursiva da matriz de projeção através das expressões propostas por Rosen (1960), necessitando-se apenas,

neste caso, adicionar um teste que condicionasse a filtragem de uma nova medida ao condicionamento numérico da matriz H.

Frequentemente não será conveniente usar matrizes de covariância do erro na estimativa inicial proporcionais a matriz identidade, e, portanto, não se terá disponível boas aproximações para a matriz de projeção. Desta forma, seria interessante dispor-se de testes de convergência empíricos, relativamente confiáveis, que dessem ao menos uma indicação de quão próximo da solução se está, para indicar se é necessário ou não calcular a matriz de projeção $V_{\ell+j}$ definida na Equação 3.38 para uma verificação mais acurada. Ressalta-se que quando a solução de problema restrito de otimização é um ponto onde os vínculos de contorno são satisfeitos e o gradiente da função objetivo é nulo, pode-se constatar tal fato diretamente. Assim, pode-se dizer que se convergiu para a solução se:

$$\bar{G}_{\bar{a}} < e_m \alpha_g \{1 + |G(\bar{x}(t_f), \bar{a})|\} \tag{3.55}$$

onde α_g, tal que $0 < \alpha_g < 1$, é um parâmetro arbitrário;

e se os vínculos de contorno forem satisfeitos, que se caracterizam por se ter:

$$\begin{aligned} |m_i| &\leq f_{ea} \cdot \bar{e}_a \qquad \text{p/ } i=1,\ell \\ |I_i| &\leq f_{ea} \cdot \bar{e}_a \qquad \text{p/ } i=1,j \end{aligned} \tag{3.56}$$

onde se define m_i e I_i por meio de:

$$\begin{aligned} M(\bar{x}(t_f), \bar{a}) &= [m_1, m_2, \ldots, m_\ell] \\ I(\bar{a}) &= [I_1, I_2, \ldots, I_j] \end{aligned} \tag{3.57}$$

Para os casos em que o gradiente da função objetivo não tende a zero nas vizinhanças da solução do problema restrito de otimi

zação, pode-se constatar eventualmente a convergência sem a necessidade de nenhum esforço adicional de computação se os parâmetros de ajuste assumirem certos valores particulares e houver mesmo assim rejeição da estimativa do incremento de busca, e os erros nos vínculos de contorno não estiverem muito próximos ao máximo valor tolerado.

No caso em que, na fase de busca do mínimo se rejeita o incremento de busca obtido, pelo fato de a estimativa do novo erro nos vínculos ligada a tal incremento ter alguma componente não contida na faixa de tolerância escolhida, pode-se dizer que provavelmente o vetor de incógnitas básico está nas vizinhanças da solução se for satisfeita a seguinte relação empírica:

$$|pG_{\bar{a}}^{T}| \leq \alpha_g \, e_m\{1 + |G(\bar{x}(t_f), \bar{a})|\} \tag{3.58}$$

e se nenhuma componente do erro nos vínculos de contorno estiver demasiadamente próxima do máximo valor permitido, o que também de uma forma empírica pode ser caracterizado por:

$$\begin{aligned} |m_i| &< e_a^* \qquad p/\ i=1,\ell \\ |I_i| &< e_a^* \qquad p/\ i=1,j \end{aligned} \tag{3.59}$$

onde e_a^* é definido por:

$$e_a^* = \max\{(\sqrt{3}/3)e_m,\ (f_{ea}.\bar{e}_a - 3e_m)$$

ou seja, quando não se consegue um incremento de busca ao qual correspondam estimativas lineares do erro nos vínculos de contorno, que permaneçam contidas dentro da faixa de tolerância mesmo com valores muito pequenos de p, e quando não se tem erros nos vínculos muito próximos ao máximo tolerado, provavelmente houve convergência.

No caso em que se rejeita, na fase de busca do mínimo, a estiamtiva do incremento de busca por não levar a uma redução da estimativa linear do valor da função objetivo, fato que se caracteriza por:

$$\underline{G}_{\hat{a}}\Delta\hat{a} > 0 \tag{3.60}$$

pode-se considerar que se convergiu para a solução se:

$$e_m^2/3 + h_i\bar{P}h_i^T - p^2\underline{G}_{\hat{a}}h_i^T h_i \underline{G}_{\hat{a}}^T < \alpha_g e_m^2 \tag{3.61}$$

ou seja, a covariância das componentes da estimativa a priori que influem no valor da função objetivo já estão muito próximas do mínimo valor destas para o qual pode-se garantir a não-ocorrência de divergência e mesmo assim não se consegue reduzir a função objetivo (ver a Equação 3.30).

Considera-se também um forte indício de convergência a obtenção de um incremento de busca que além de levar a uma estimativa linear da função objetivo maior do que o valor obtido desta função, faz com que se viole os vínculos de contorno, i.e, não se respeite a Equação 3.56, mesmo quando a Desigualdade 3.64 é satisfeita. Desta forma, convém aplicar nestes casos o teste de ortogonalidade entre o gradiente da função objetivo e a matriz de projeção, antes de alterar os parâmetros e tentar obter uma nova estimativa.

Na prática, frequentemente rejeitar-se-ão iterações pelo fato de o incremento de busca ou não levar à redução da função objetivo ou não levar a uma estimativa do erro nos vínculos de contorno que satisfaça à Equação 3.56, mesmo quando os parâmetros de ajuste estejam muito longe dos valores limites especificados pelas Equações 3.58 e 3.61. Uma boa política, nestes casos, seria verificar a ortogonalidade entre o gradiente da função objetivo e o hiperplano tangente às restrições após um número previamente estabelecido de rejeições consecutivas do resultado do processo de cálculo do incremento de busca, considerando-se que achou-se a solução se os vínculos forem satisfeitos, i.e, se a Equação 3.56 é satisfeita e:

$$V_{\ell+j}\ G_{\bar{a}}^{T} < e_m \cdot \alpha_g \{1 + |G(\bar{x}(t_f), \bar{a})|\} \qquad (3.62)$$

onde $V_{\ell+j}$ é a matriz de projeção definida na Equação 3.39.

Para melhorar esta possível política, tentou-se estabelecer um critério adicional para os casos em que se rejeita a estimativa do incremento de busca por não levar à redução da estimativa linear do valor da função objetivo. Com tal intuito decidiu-se recorrer à estatística e ao modelo obtido, pela resolução do Problema Associado 3.1, para a média do incremento de busca e para a covariância do erro em tal incremento, com o objetivo de testar a hipótese de a estimativa linear da variação da função objetivo ser negativa quando se obteve para esta um valor positivo. Caso, dentro de um certo nível de significância, se rejeitasse a hipótese de a variação real da função objetivo ser negativa quando se obtivesse para esta um valor positivo, ter-se-ia um forte indício de convergência e desta forma seria aconselhável verificar imediatamente se a Desigualdade 3.62 é satisfeita.

A partir da matriz P de covariância do erro na estimativa $\Delta\hat{a}$ do incremento de busca no vetor de parâmetros e da suposição de que a estimativa $\Delta\hat{a}$ é não-viciada, pode-se escrever a média $\Delta\hat{G}$ da variação da função objetivo e a variância $\sigma^2_{\Delta G}$ de tal informação como:

$$\begin{aligned} \Delta\hat{G} &= G_{\bar{a}}\Delta\hat{a} \\ \sigma^2_{\Delta G} &= E[(G_{\bar{a}}\Delta a - G_{\bar{a}}\Delta\hat{a})^2] = G_{\bar{a}}\ P\ G_{\bar{a}}^{T} \end{aligned} \qquad (3.63)$$

Assim, pela desigualdade de Tchebycheff tem-se que

$$\text{prob}\ \{|G_{\bar{a}}\Delta a - G_{\bar{a}}\Delta\hat{a}| > k\sigma_{\Delta G}\} \leq 1/k^2 \qquad (3.64)$$

onde prob{y} indica a probabilidade de ocorrer o evento y. Desta forma pode-se afirmar que se a distribuição for simétrica ter-se-á:

$$\text{prob}\ \{\ G_{\bar{a}}\Delta a < 0,\ \text{dado}\ G_{\bar{a}}\Delta\hat{a} > 0\ \} \leq \frac{G_{\bar{a}}\, P\, G_{\bar{a}}^{T}}{2(G_{\bar{a}}\Delta\hat{a})^2} \triangleq \Omega \qquad (3.65)$$

Tal expressão tem alguns aspectos interessantes. Em primeiro lugar afirma que a probabilidade de se ter convergido quando se rejeita a estimativa $\Delta\hat{a}$ por não provocar o decréscimo da função objetivo é tanto maior quanto menor for o produto da matriz de covariância do erro na estimativa final pelo gradiente da função objetivo. Assim, tal probabilidade será alta se P for o produto de uma matriz qualquer pela matriz de projeção, pois neste caso $\sigma^2_{\Delta G}$ deve tender a zero nas vizinhanças da solução do problema restrito de otimização. Outro ponto interessante é que a probabilidade de se ter convergido será tanto maior quanto maior o acréscimo previsto na função objetivo, ou seja, aumentos muito grandes no valor da função objetivo são fortes indícios de convergência.

Comparar Ω, definido na Equação 3.65, com um determinado nível de significância preestabelecido, ϕ, em casos onde se obtinha um incremento de busca para o qual $G_{\bar{a}}\Delta\hat{a}$ fosse positivo, revelou-se útil e capaz de apontar a convergência quando se utilizavam valores de β_0 e β_{p0} iguais a um e valores de $\bar{\beta}$ e $\bar{\beta}_p$ menores ou iguais a um, independentemente de se usar ou não uma matriz de covariância do erro na estimativa a priori proporcional à identidade. Ao usar valores de β_0 e β_{p0} maiores, tal critério deixava frequentemente de apontar a existência de uma probabilidade não-nula de se ter convergido, até mesmo quando de fato se tinha convergido. Com base nos resultados experimentais, é recomendável que, caso na fase de busca do mínimo se rejeite a estimativa $\Delta\hat{a}$ do incremento de busca por $G_{\bar{a}}\Delta\hat{a}$ ser positivo, verifique-se por meio da Equação 3.62 se houve ou não convergência antes de se tentar uma reiteração se:

$$\frac{G_{\bar{a}}\, P\, G_{\bar{a}}^{T}}{2(G_{\bar{a}}\Delta\hat{a})^2} < 1 \qquad (3.66)$$

3.5 - CONSIDERAÇÕES SOBRE A CONVERGÊNCIA

Para que se caminhe no sentido de solucionar o problema restrito de otimização, é necessário que o erro nos vínculos de contorno diminua de iteração em iteração, até que suas componentes fiquem contidas numa faixa de tolerância de largura comparável aos erros de integração e às imprecisões do modelo do sistema dinâmico. É desejável que a função objetivo diminua enquanto isto não acontecer, porém, não se pode exigir que isto aconteça devido ao fato de enquanto os vínculos de contorno não forem satisfeitos existe a possibilidade de se obterem valores da função objetivo menores que o valor desta na solução do problema restrito de otimização. Após a satisfação dos vínculos de contorno ou terminais, deve-se exigir que a função objetivo diminua de iteração em iteração, sem contudo permitir que o erro nos vínculos de contorno saia da faixa de tolerância final desejada.

Verifica-se que ao satisfazer às medidas z definidas na Equação 3.10, sem se violar a hipótese de perturbação linear, provoca-se um decréscimo do erro nos vínculos de contorno. Por sua vez, a estimativa a priori provoca, sempre que possível, o decréscimo da função objetivo. Pode acontecer, porém, que a hipótese de perturbação linear seja violada, ou então que o Filtro de Kalman calcule um incremento de busca incapaz de provocar a redução do erro nos vínculos de contorno, ou mantê-los na faixa de tolerância especificada ou provocar o decréscimo da função objetivo. Tais fatos se devem a um mal condicionamento numérico do gradiente da função objetivo e da matriz H definida na Equação 3.5, ou então a uma má escolha dos parâmetros de ajuste do algoritmo. Por estes motivos necessita-se, para garantir a convergência, de um esquema apropriado de verificações e de ajuste dos parâmetros que comandam a convergência. Tal esquema é descrito no capítulo seguinte.

Outro ponto a ser considerado é o eventual emprego de fases de restauração, quando alguma componente do erro nos vínculos de contorno estiver próxima ao máximo valor tolerado, bem como o uso de faixas de tolerância para o erro nos vínculos de contorno sucessivamen

te menores, alterando-se, neste caso, fases de busca do mínimo da função objetivo com fases de redução do erro nos vínculos de contorno. Tais tópicos também são discutidos no capítulo a seguir.

CAPÍTULO 4

ALGORITMO NUMÉRICO

4.1 - INTRODUÇÃO

O algoritmo numérico deve satisfazer, de acordo com os recursos computacionais disponíveis, os requisitos de rapidez, compacidade e confiabilidade; deve ter ainda um grau de comunicabilidade capaz de permitir que qualquer operador obtenha soluções de um novo problema com um mínimo de treinamento na utilização do método.

O algoritmo numérico proposto é iterativo e parte de um valor inicial qualquer do vetor de parâmetros a ser otimizado, ao qual corresponda uma dada trajetória do vetor de estado do sistema dinâmico e um dado valor da função objetivo, para tentar obter um novo vetor de parâmetros ao qual corresponda uma nova trajetória do vetor de estado do sistema dinâmico e um novo valor da função objetivo mais próximos dos valores destes sobre a solução do problema restrito de otimização de parâmetros. A geração das trajetórias do sistema dinâmico é feita aqui por integração numérica direta.

O procedimento adota um esquema de perturbação no sentido de que os incrementos no vetor de parâmetros devem ser de ordem de grandeza tal que se possa desprezar os termos de ordem superior à primeira, na expansão em série de Taylor, da função objetiva, dos vínculos de contorno e dinâmicos, uma vez que a determinação do vetor de parâmetros está baseada nesta hipótese.

Para determinar a matriz de derivadas do erro nos vínculos de contorno no instante final em relação aos parâmetros, $M_{\bar{a}}$, e eventualmente para a determinação do gradiente da função objetivo, $G_{\bar{a}}$, necessita-se conhecer as derivadas do estado final em relação aos parâmetros, $x_{\bar{a}}(t_f)$. Dois são os métodos numéricos disponíveis na literatura para a geração de $x_{\bar{a}}(t_f)$ (Bryson and Ho, 1975): derivação numérica di

reta e geração de soluções unitárias. Optou-se pela utilização do segundo destes métodos devido às vantagens no tocante à precisão numérica. Na Sub-seção 4.2.4 dá-se uma breve idéia sobre estes métodos, e faz-se uma breve comparação entre estes métodos no tocante a aspectos como: custo computacional, dificuldade de implementação e precisão numérica. Apresenta-se de forma detalhada, no Apêndice A, o método de geração de derivadas parciais do estado em relação aos parâmetros escolhido. A derivação das equações utilizadas para a obtenção das derivadas parciais do estado em relação aos parâmetros está baseada no trabalho de Rios Neto e Ceballos (1979). Foi interessante apresentar neste apêndice uma extensão dos resultados de tais autores para problemas onde existam vínculos de desigualdade sobre as variáveis de controle.

Na Subseção 4.2.1 faz-se um resumo da sequência de cálculos necessária para obter o incremento no vetor de parâmetros.

Nas Subseções 4.2.2 e 4.2.4 aborda-se a análise de resultados, cuja função é verificar se o incremento de busca obtido é satisfatório. Na primeira destas subseções trata-se da parte da análise de resultados que independe do resultado da integração numérica, ao passo que na Subseção 4.2.4 trata-se da parte do processo de análise dos resultados que depende do resultado da integração numérica, e é onde se estabelecerá definitivamente se o incremento de busca é satisfatório ou não. Juntamente com estes processos de análise apresenta-se o esquema de correção e ajuste dos valores de β_0, β_{p0}, $\bar{\beta}$, $\bar{\beta}_p$ e p.

Na Subseção 4.2.3 aborda-se a integração numérica e a geração das derivadas parciais.

Não se discutem as vantagens e desvantagens, do ponto de vista numérico, entre as diferentes formas de apresentação das equações do Filtro de Kalman descritas no Capítulo 2, pois uma discussão mais detalhada sobre o assunto foge ao escopo deste trabalho.

4.2 - DESCRIÇÃO DAS ETAPAS DO ALGORITMO NUMÉRICO

Pode-se dividir o algoritmo numérico utilizado em cada iteração nas seguintes etapas principais distintas:

- cálculo do incremento de busca no vetor de parâmetros;
- análise preliminar dos resultados;
- ajuste dos parâmetros que comandam a convergência;
- integração numérica e geração de derivadas parciais do vetor de estado em relação aos parâmetros;
- análise final dos resultados;
- atualização do vetor de parâmetros, da matriz de derivadas parciais do erro nos vínculos em relação aos parâmetros e do gradiente da função objetivo.

Neste capítulo, descrevem-se apenas as verificações de caráter geral, que pressupõem a inexistência de vínculos de desigualdade entre os parâmetros do funcional. Tal não ocorre por exemplo nos casos em que se permite que os tempos nodais sejam parâmetros a otimizar totalmente independentes entre si, pois neste caso o integrador pressupõe que os tempos nodais estão ordenados, de forma que se deve rejeitar incrementos de busca que levem à alteração da posição relativa dos tempos nodais. As alterações introduzidas no algoritmo para permitir o tratamento deste tipo de funcional são descritas no Apêndice B.

4.2.1 - CÁLCULO DO INCREMENTO DE BUSCA NO VETOR DE PARÂMETROS

Para efetuar o cálculo de incremento no vetor de parâmetros é necessário conhecer os valores do gradiente da função objetivo, $G_{\bar{a}}$, das matrizes de derivadas do erro nos vínculos de contorno em relação aos parâmetros, $M_{\bar{a}}$ e $I_{\bar{a}}$, e o erro nos vínculos de contorno, $I(\bar{a})$ e $M(\bar{x}(t_f), \bar{a})$, correspondentes a um dado valor inicial $\bar{a}$ do vetor de parâmetros a ser otimizado, ao qual corresponde uma trajetória $\bar{x}(t)$, dita trajetória nominal ou de referência, e também valores satisfatórios de β_0, β_{p0}, $\bar{\beta}$, $\bar{\beta}_p$, p e e_m,

onde

β_0, β_{p0}, $\bar{\beta}$, $\bar{\beta}_p$ - são os coeficientes de segurança para o cálculo da matriz $\bar{P}$ de covariância do erro na estimativa a priori do incremento de busca no vetor de parâmetros;

p - é o parâmetro que estabelece uma relação de proporcionalidade entre o gradiente da função objetivo e a estimativa a priori do incremento de busca no vetor de parâmetros;

e_m - é o erro numérico despresível, que deve ser da ordem dos erros envolvidos na integração numérica.

Conhecidos estes parâmetros, calcula-se a estimativa a priori do incremento de busca do vetor de parâmtros por meio de:

$$\Delta\bar{a} = -pG_{\bar{a}} \tag{4.1}$$

onde

$G_{\bar{a}}$ - é o gradiente da função objetivo, ver a Equação 2.20.

Calcula-se a seguir a matriz de covariância de ruído de medidas R por meio de:

$$R = (e_m^2/3) \cdot I \tag{4.2}$$

onde I é a matriz identidade,

e a matriz de covariância do erro na estimativa a priori $\bar{P}$ por meio de:

$$\bar{P}_{ii} = 2\beta_i^2 e_e^2 + \beta_{pi}^2 p^2 G_{\bar{a}} G_{\bar{a}}^T$$

$$\bar{P}_{ij} = 0 \quad \text{se } i \neq j \tag{4.3}$$

onde

β_i - é igual a β_0 se o parâmetro a_i pertence ao grupo de parâmetros que não afetam o valor da função objetivo, dito grupo de parâmetros não-correlacionados com a função objetivo, e igual a $\bar{\beta}$ em caso contrário;

β_{pi} - é igual a β_{p0} se o parâmetro a_i pertence ao grupo de parâmetros não-correlacionados com a função objetivo, e igual a $\bar{\beta}_p$ em caso contrário;

e_e - é igual ao máximo dos e^2_{ei} durante a fase de busca da satisfação dos vínculos e igual à média de tais valores durante a fase de busca do mínimo, onde os e^2_{ei} são definidos por:

$$e^2_{ei} = e^2_m/(3h_i h_i^T) \tag{4.4}$$

onde os h_i são definidos por

$$H^T \triangleq [M_{\bar{a}}^T : I_{\bar{a}}^T] = [h_1^T, h_2^T, \ldots, h_{\ell+j}^T] \tag{4.5}$$

Em seguida deve-se calcular os decréscimos desejados nas componentes do erro nos vínculos de contorno ρ_{ii} por meio de

$$\rho_{ii} = 1/2\ \bar{\rho}_{ii} \qquad \text{se } \bar{\rho}_{ii} \leq 2$$

$$\rho_{ii} = 1 \qquad \text{se } \bar{\rho}_{ii} > 2$$

onde $\bar{\rho}_{ii}$ é a raíz real positiva da Equação 3.35, ou seja, a solução de

$$\bar{\rho}^2_{ii}\, E_i^2 - 2\bar{\rho}_{ii} E_i p h_i G_{\bar{a}}^T + p^2 (h_i G_{\bar{a}})^2 = e^2_m/3 + h_i \bar{P} h_i^T$$

e E_i é definido por:

$$E^T(a) = [E_1, E_2, \ldots, E_{\ell+j}] = [M^T(\bar{x}(t_f), \bar{a}) : I^T(\bar{a})]$$

Calculados os $\ell+j$ valores de ρ_{ii}, podem-se calcular as $\ell+j$ componentes z_i do vetor de observações por meio de

$$z_i = -\rho_{ii} E_i \tag{4.6}$$

Uma vez obtidos o vetor de observações z, a estimativa a priori, a matriz de covariância de ruído de medidas R e a matriz de covariância do erro na estimativa a priori $\bar{P}$, utiliza-se o Filtro de Kalman, adotando-se de preferência uma forma sequencial para, a partir da estimativa a priori $\Delta\bar{a}$ dada pela Equação 4.1 e do vetor de observações z dado pela Equação 4.6, e das matrizes R e $\bar{P}$, obter uma estimativa $\Delta\hat{a}$ do incremento de busca no vetor de parâmetros Δa e a matriz P de covariância do erro nesta nova estimativa.

Com a obtenção desta estimativa do incremento de busca no vetor de parâmetros, completa-se o processo de cálculo em questão. Necessita-se agora verificar se tal incremento de busca é satisfatório no sentido de que realmente permita obter um novo vetor de parâmetros mais próximo da solução do problema restrito de otimização. O processo de análise que tem por objetivo verificar tal fato é assunto das seções subsequentes. Cabe também frisar que a escolha de $\bar{\beta}_p$ deverá ser de tal forma que garanta a existência de duas raízes reais e distintas, uma positiva e outra negativa, para as equações de determinação dos $\bar{\rho}_{ii}$ (ver a Equação 3.30).

No programa de computador desenvolvido para testes do procedimento proposto, utilizou-se de uma sub-rotina que faz a aplicação sequencial do Filtro de Kalman, usando-se as equações deste filtro na forma apresentada por Bryson e Ho (1975), (ver a Equação 2.23). Tal escolha se deve pura e simplesmente à maior familiaridade do autor com tal forma das equações do filtro de Kalman e à expectativa de que se teria condições tais que dificilmente apareceriam problemas numéricos durante a aplicação do filtro.

4.2.2 - ANÁLISE PRELIMINAR DOS RESULTADOS

Esta etapa do processo de análise de resultados tem por objetivo verificar, antes de se efetuar a integração numérica, se o incremento de busca no vetor de parâmetros a ser otimizado pode ser considerado como provavelmente satisfatório, ou seja, se provavelmente provoca uma caminhada na direção da solução do problema restrito de otimização, ou se tal incremento deve ser rejeitado, por ser provavelmente insatisfatório.

Necessita-se efetuar tais verificações devido ao fato de que uma má escolha dos parâmtros de ajuste do algoritmo ou os seguintes tipo de mal condicionamento podem invalidar o vetor de incrementos obtido: dos vetores h_i, definidos na Equação 4.5; do gradiente da função objetivo, definido na Equação 2.20; dos valores dos erros nos vínculos de contorno, $M(\bar{x}(t_f), \bar{a})$ e $I(\bar{a})$. Obviamente, dado o custo computacional necessário para a obtenção do vetor de estado do sistema dinâmico por integração numérica, não tem sentido efetuar-se tais integrações indiscriminadamente, imediatamente após a obtenção de um incremento no vetor de parâmetros via Filtro de Kalman. Daí a conveniência de executar uma análise linearizada da qualidade do incremento de busca obtido. Deve-se notar que os novos erros nos vínculos de contorno no instante inicial podem ser avaliados diretamente antes mesmo de se efetuar a integração numérica. Convém que se exija, nesta análise linearizada, que o incremento no vetor de parâmetros satisfaça condições mais rigorosas do que aquelas extritamente necessárias para garantir a convergência, de forma a diminuir ainda mais a probabilidade de efetuar integrações espúrias.

Conforme já foi colocado ao se discutir, na Seção 3.3, a escolha dos parâmetros que controlam a convergência, a busca da solução se divide em duas fases básicas.

Na primeira destas fases, doravante chamada fase de captura dos vínculos, tem-se por objetivo:

a) obter erros nos vínculos de contorno, $M(\bar{x}(t_f), \bar{a})$ e $I(\bar{a})$, tais que todas as componentes destes vetores tenham valor em módulo inferior a um valor preespecificado $\bar{e}_a$, que deve ser positivo e da ordem do erro esperado de integração;

b) diminuir, se possível, a função objetivo.

Na segunda destas fases, doravante chamada fase de busca do mínimo, tem-se por objetivo diminuir o valor da função objetivo sem permitir que o erro nos vínculos de contorno saia de uma faixa de tolerância de largura $f_{ea} \cdot \bar{e}_a$ comparável às incertezas causadas pelos erros de modelagem do sistema dinâmico e de integração.

Assim, denotando-se por m_i e I_i às componentes dos erros nos vínculos de contorno $M(\bar{x}(t_f),\bar{a})$ e $I(\bar{a})$, (ver a Equação 3.57), e definindo-se por M' e I' às estimativas lineares do novo erro nos vínculos de contorno e por m'_i e I'_i, respectivamente, às componentes de M' e I', ou seja, definindo-se:

$$M'^T = M^T(\bar{x}(t_f), \bar{a}) + (M_{\bar{a}}\Delta\bar{a})^T = [m'_1, m'_2, \ldots, m'_e]$$
$$I'^T = I^T(\bar{a}) + (I_{\bar{a}}\Delta\bar{a})^T = [I'_1, I'_2, \ldots, I'_e] \qquad (4.7)$$

e definindo-se

$$E^T(\bar{a}) = [M^T(\bar{x}(t_f), \bar{a}) : I^T(\bar{a})]$$
$$E'^T = [M'^T : I'^T] \qquad (4.8)$$

deve-se exigir na fase de captura dos vínculos que:

$$|m'_i| \leq |m_i| + 10e_m$$
$$|I'_i| \leq |I_i| + 10e_m \qquad (4.9a)$$

e

$$|E'| < |E(a)| \qquad (4.9b)$$

A exigência de que o módulo do erro nos vínculos de contorno diminua é necessária e suficiente para garantir a obtenção de um vetor de parâmetros a ser otimizado ao qual correspondam erros nos vínculos contidos numa faixa de tolerância especificada. A exigência adicional imposta às componentes das estimativas lineares dos erros nos vínculos de contorno é feita para minimizar os efeitos da não-linearidade. Usaram-se, neste primeiro teste, as estimativas lineares dos erros nos cínculos de contorno no instante inicial e não os novos erros, para poder avaliar melhor o motivo da rejeição da estimativa.

Caso a Desigualdade 4.9 não seja satisfeita, e o valor de e_m seja muito maior que o zero numérico, pode estar ocorrendo que:

a) a informação a priori é conflitante com o vetor de observações, ou seja, o decréscimo desejado na função objetivo é muito grande, de forma que, apesar do filtro não respeitar completamente a informação a priori e não decrescer equivalentemente o valor da função objetivo, tal informação impede a redução do erro nos vínculos de contorno ou até mesmo provoca o aumento nestes erros, devido ao fato de não se poder reduzir o erro nos vínculos de contorno e ao mesmo tempo reduzir apreciavelmente o valor da função objetivo. Para contornar tal problema e obter uma estimativa do vetor de parâmetros a ser otimizado capaz de satisfazer à Desigualdade 4.9, tem-se duas alternativas: reduzir o módulo da estimativa a priori, mediante a redução de p, sem reduzir proporcionalmente o vetor de observações; ou então aumentar a covariância da estimativa a priori sem alterar tal estimativa, mediante um aumento de $\bar{\beta}$, $\bar{\beta}_p$, β_0 e β_{p0i}.

b) as medidas z são conflitantes entre si, no sentido de que um mal condicionamento numérico das matrizes de derivadas parciais de erro nos vínculos de contorno em relação aos parâmetros, $M_{\bar{a}}$ e

$I_{\bar{a}}$, faça com que o estimador falhe ao tentar incorporar tais medidas, já que não existe um incremento no vetor de parâmetros capaz de satisfazer a todas as observações dentro do atual nível de incerteza. Para contornar tal problema deve-se aumentar a razão entre a covariância do ruído de medidas e o quadrado do vetor de medidas mediante um aumento de e_m ou uma redução de p.

Devido às dificuldades em precisar qual dos dois motivos levou à rejeição do incremento de busca obtido e ao fato de que quando todos os $\bar{\rho}_{ii}$ são maiores do que dois a redução de p pode contornas ambos os problemas decidiu-se:

a) reduzir apenas o valor de p se todos os $\bar{\rho}_{ii}$ forem maiores do que dois e repetir o cálculo do incremento de busca $\Delta\hat{a}$;

b) reduzir p e aumentar $\bar{\beta}_p$, $\bar{\beta}$, β_{p_0} e β_0 se algum dos valores de $\bar{\rho}_{ii}$ for menor que dois e repetir o cálculo de $\Delta\hat{a}$.

Uma vez que se podem obter diretamente os valores do erro nos vínculos de contorno inicial associados a um vetor de parâmetro igual a $\bar{a} + \Delta\hat{a}$, comvém verificar se:

$$|[I^T(\bar{a} + \Delta\hat{a}) : M'^T]| < |E(\bar{a})| \qquad (4.10)$$

ou seja, se o incremento de busca respeitou a hipótese de perturbação linear no tocante aos vínculos de contorno no instante inicial, ou se, em caso contrário, ainda há chances de o incremento de busca obtido levar de fato à redução do erro nos vínculos de contorno. Caso se rejeite a estimativa do incremento de busca que satisfaça à Desigualdade 3.9, por não satisfazer à Desigualdade 4.10 pode-se obter um incremento de busca satisfatório multiplicando-se o incremento de busca por um escalar α_ℓ pertencente ao intervalo (0,1) até se obter um incremento de busca que não viole à Desigualdade 4.10. Um bom valor para α_ℓ é obtido com a hipótese de que o efeito da não-linearidade se deva somente ao termo de segunda ordem, através da solução de:

$$\left|\left[\{I(a) + \alpha_\ell [I' - I(\bar{a})] + \alpha_\ell^2 [I(\bar{a} + \Delta\bar{a}) - I']\} : M'^T\right]\right| \leq \left|E(\bar{a})\right| \qquad (4.11)$$

Também pode-se obter uma estimativa de incremento de busca capaz de satisfazer à Desigualdade 4.10, diminuindo-se p e repetindo-se o cálculo do incremento de busca no vetor de parâmetros. Esta segunda alternativa pode ser vantajosa durante a fase de captura dos vínculos, quando todos os $\bar{\rho}_{ii}$ são maiores que dois, uma vez que em geral nestes casos a quebra de linearidade está ligada a decréscimos muito grandes da função objetivo. Desta forma, pode-se eventualmente, nestes casos, obter uma estimativa do incremento de busca que satisfaça à Desigualdade 4.10 sem que para isto se atrase a captura dos vínculos de contorno, de forma que o aumento da velocidade de convergência resultante compense o tempo gasto com a repetição do cálculo do incremento de busca. Assim, decidiu-se optar pela segunda alternativa, quando todos os $\bar{\rho}_{ii}$ são maiores do que dois na fase de captura dos vínculos e pela primeira delas nos demais casos, onde se multiplicará também p por α_ℓ.

Pode ocorrer que mesmo reduzindo p a valores tais que:

$$e_m^2/3 + h_i \bar{P} h_i^T - p^2 G_{\bar{a}} h_i^T h_i G_{\bar{a}}^T < e_m^2 \qquad (4.12)$$

não se consiga obter, na fase de captura dos vínculos, um incremento de busca capaz de satisfazer às Desigualdades 4.9 e 4.10. Neste caso, pode estar acontecendo que se esteja sobre um ponto estacionário do erro nos vínculos, ao qual corresponda um valor não-nulo destes erros, ou então que se tenha estimado mal os erros de integração e o valor de $\bar{e}_a$ seja muito pequeno. Admitindo-se que se esteja sobre um ponto estacionário do erro nos vínculos de contorno, é razoável que se adote, como incremento de busca o maior incremento capaz de satisfazer à Desigualdade 4.9b, uma vez que se deve tentar sair da zona de influência deste ponto singular. Admitindo-se que o erro de integração seja maior que o esperado, deve-se considerar encerrada a fase de captura dos vínculos. O programa de computador desenvolvido para teste numérico supõe que a causa do problema seja um erro na estimativa dos erros de integração, quan

do todas as componentes do erro nos vínculos de contorno forem menores do que o Cinc, onde Cinc é um parâmetro, maior do que 1, ligado às incertezas da estimativa dos erros de integração. A maneira de tratar pontos estacionários do erro nos vínculos de contorno é um dos tópicos a ser melhor estudado, visto que o tratamento aqui dado a esta situação não garante que se obtenha um novo valor para o vetor de parâmetros a ser otimizado que saia da zona de influência deste mínimo local. Uma possível maneira de contornar este problema seria adicionar temporariamente à função objetivo uma parcela que se destina à maximizar a distância a este ponto singular e proceder-se tal qual na fase de busca do mínimo, até que se consiga conjuntamente diminuir o erro nos vínculos e aumentar a distância entre o ponto singular do erro nos vínculos e a solução obtida. De qualquer forma no entanto, este tipo de problema, caso apareça, será de tratamento numérico difícil e poderá impedir que se convirja para a solução do problema restrito de otimização, partindo-se de uma dada condição inicial. Resta frisar que todos os demais métodos de otimização de parâmetros sujeito a vínculos, disponíveis na literatura, também apresentam problemas numéricos nas proximidades de pontos deste tipo.

Doravante tratar-se-á de verificações preliminares que devem ser efetuadas na fase de busca do mínimo, cujo objetivo é diminuir o valor da função objetivo sem permitir que os erros nos vínculos saiam de uma faixa de tolerância de largura $f_{ea}\ \bar{e}_a$. Esta fase tem início após todas as componentes do erro nos vínculos de contorno assumirem valores em módulo inferior a $\bar{e}_a$, ou seja, após ter-se:

$$\begin{aligned} |m_i| &< \bar{e}_a \qquad p/i=1,2,\ldots,\ell \\ |I_i| &< \bar{e}_a \qquad p/i=1,2,\ldots,j \end{aligned} \tag{4.13}$$

O primeiro passo nesta fase é verificar o desempenho do Filtro de Kalman, mediante uma análise completamente linearizada, para obter informações valiosas não só quanto à qualidade do incremento de busca como quanto à eventual necessidade de ajuste dos parâmetros de

controle de algoritmo e quanto à melhor forma de promover tais ajustes. Deve-se iniciar esta análise completamente linearizada verificando-se:

$$|m_i'| < f_{ea}\,\bar{e}_a$$
$$|I_i'| < f_{ea}\,\bar{e}_a \qquad (4.14)$$

onde m_i' e I_i' são as estimativas lineares do novo erro nos vínculos de contorno, definidas na Desigualdade 4.9, e $f_{ea}\,\bar{e}$ um parâmetro maior ou igual a um.

Caso esta condição não seja satisfeita deve-se verificar se:

$$G_{\bar{a}} < \alpha_g e_m (1 + G(\bar{x}(t_f), \bar{a})) \qquad (4.15)$$

e considerar que se convergiu caso tal desigualdade seja satisfeita. Caso não se satisfaça à Desigualdade 4.14 e 4.15 e já se tenha efetuado um número de repetições do cálculo de incremento de busca superior a um número "MAXREJ" de reiterações preespecificadas, ou se tenha que:

$$|pG_{\bar{a}}^T| < \alpha_g e_m \qquad (4.16)$$

e também que:

$$|m_i| < e_a^* \quad p/i=1,2,\ldots,\ell$$
$$|I_i| < e_a^* \quad p/i=1,2,\ldots,j \qquad (4.17)$$

onde $e_a^*\,\bar{e}$ definido por:

$$e_a^* = \max\{\sqrt{3}\,e_m/3,\ (f_{ea}\,\bar{e}_a - 3e_m)\} \qquad (4.18)$$

deve-se verificar se a Condição de Convergência 3.62

$$[I - H^T(H\,H^T)^{-1}H]\,G_{\bar{a}}^T < \alpha_g e_m(1 + |G(\bar{x}(t_f), \bar{a})|)$$

satisfeita. Caso a Desigualdade 4.14 não seja satisfeita e a Condição 3.62 não seja verdadeira ou não tenha sido testada, deve-se repetir o cálculo de incremento de busca Δa com novos valores para os parâmetros de controle do algoritmo. Neste caso, deve-se fazer com que $\bar{\beta}$ e $\bar{\beta}_p$ voltem a seus valores iniciais, e voltar à fase de captura dos vínculos sempre que f_{ea} for maior que um e a Desigualdade 4,17 não for satisfeita, pois caso contrário o procedimento pode passar a calcular incrementos de busca muito pequenos, além de se ter problemas mais frequentes com a quebra de linearidade. Já no caso da Desigualdade 4.17 ser satisfeita, deve-se aumentar $\bar{\beta}$ e $\bar{\beta}_p$ até que estes se tornem iguais a β_0 e β_{p0} respectivamente, ou até que ambos sejam iguais a um se não houver nenhum parâmetro não correlacionado com a função objetivo. Caso $\bar{\beta}$ e $\bar{\beta}_p$ sejam iguais a β_0 e β_{p0} respectivamente, ou caso não haja nenhum parâmetro não correlacionado com a função objetivo e $\bar{\beta}$ e $\bar{\beta}_p$ sejam iguais a um, deve-se diminuir p, quando as Desigualdades 4.14 e 4.18 não forem satisfeita e 4.17 o for. Quando a Desigualdade 4.17 não for satisfeita e f_{ea} for igual a um, deve-se ignorar tal fato e proceder tal qual quando esta desigualdade for satisfeita.

Obtida uma estimativa do incremento de busca no vetor de parâmetros que satisfaça à Desigualdade 4.14 deve-se verificar se a estimativa do incremento de busca no vetor de parâmetros pode provocar o decréscimo da função objetivo, verificando-se a satisfação da seguinte desigualdade:

$$G_{\bar{a}}\Delta\hat{a} < 0 \qquad (4.19)$$

Se esta desigualdade não for satisfeita, deve-se considerar que se convergiu para a solução do problema restrito de otimização se para i=1,j+ℓ tiver-se (ver a Equação 3.30):

$$e_m^2/3 + h_i\,\bar{P}\,h_i^T - p^2 G_{\bar{a}} h_i^T h_i G_{\bar{a}}^T < \alpha_g^2 e_m^2 \qquad (4.20)$$

Se a Desigualdade 4.19 não for satisfeita, deve-se verificar se a Condição de Convergência 3.62 é satisfeita, considerando-se encerrada a busca da solução do problema restrito de otimização em caso afirmativo, sempre que se tiver efetuado um número de repetições infrutíferas do cálculo do incremento de busca superior a um número máximo de rejeições previamente especificado, "MAXREJ", ou quando:

$$\frac{(\underline{G}_{\bar{a}}\, P\, \underline{G}_{\bar{a}}^{T})}{2(\underline{G}_{\bar{a}}\Delta\bar{a})^{2}} < 1 \qquad (4.21)$$

onde P é a matriz de covariância do erro na estimativa $\Delta\bar{a}$.

Caso a Desigualdade 4.19 não seja satisfeita e não se tenha constatado que se convergiu para a solução, por meio das Desigualdades 3.62, 4.15 ou 4.20 deve-se diminuir $\bar{\beta}$ e $\bar{\beta}_p$ e repetir o cálculo do incremento no vetor de parâmetros.

A seguir deve-se fazer uma verificação prévia dos efeitos da não-linearidade, uma vez que se pode calcular diretamente os novos erros nos vínculos de contorno no instante inicial. Assim, definindo-se I_i'' como as componentes do novo erro nos vínculos de contorno no instante inicial $I(\bar{a} + \Delta\bar{a})$ por meio de

$$[I_1'' , I_2'' , \ldots, I_j''] \triangleq I^{T}(\bar{a} + \Delta\bar{a})$$

deve-se exigir que:

$$|I_i''| < f_{ea}\, \bar{e}_a \qquad p/i=1,2,\ldots,j \qquad (4.22)$$

Caso esta condição não seja satisfeita, deve-se multiplicar o incremento de busca obtido por um escalar α_ℓ, pertencente ao intervalo (0,1). Pode-se achar um bom valor de α_ℓ através da hipótese de que o efeito da não-linearidade se deva somente aos termos de segunda ordem, mediante a solução da inequação:

$$|I_n + (I'_n - I_n)\alpha_\ell + (I''_n - I'_n)\alpha_\ell^2| \leq \bar{e}_a \, f_{ea} \qquad (4.23)$$

onde

- n é o índice da componente de $I(\bar{a} + \Delta\bar{a})$ que viola a Condição 4.22 a qual corresponde o menor valor de α_ℓ pertencente ao intervalo (0,1);
- I'_n é a i-ésima componente da estimativa linear do novo erro nos vínculos de contorno (ver a Desigualdade 4.9).

Obtida uma estimativa do incremento de busca que satisfaça à Desigualdade 4.23 deve-se verificar, caso a função objetivo seja não-linear e não dependa do vetor de estado no instante final, se este realmente diminuiu, ou seja, se:

$$G(\bar{a} + \Delta\bar{a}) < G(\bar{a}) \qquad (4.24)$$

e, em caso contrário, deve-se multiplicar o incremento de busca $\Delta\bar{a}$ por um escalar α_ℓ pertencente ao intervalo (0,1) até se obter um incremento de busca capaz de satisfazer a tal condição. Um bom valor de α_ℓ é o obtido a partir da hipótese de que a não-linearidade se deva apenas ao efeito do termo de segunda ordem, mediante a solução de

$$G(\bar{a}) + G_{\bar{a}}\Delta\bar{a}\alpha_\ell + [G(\bar{a} + \Delta\bar{a}) - G(\bar{a}) - G_{\bar{a}}\Delta a]\alpha_\ell^2 < G(\bar{a}) \qquad (4.25)$$

No caso de se rejeitar a estimativa do incremento de busca por não respeitar as Desigualdade 4.22 e 4.24, deve-se considerar que se convergiu para a solução quando

$$\Delta\bar{a}^T\Delta\bar{a} \max\{h_i h_i^T, \; G_{\bar{a}}G_{\bar{a}}/[(1 + G^2(\bar{x}(t_f), \bar{a}))\alpha_g^2]\} < e_m^2 \qquad (4.26)$$

onde α_g é um parâmetro que estabelece uma relação entre o erro desprezível na função objetivo e o erro numérico desprezível.

Por último, é interessante que se limite a variação das componentes de M', definido na Equação 4.7, para evitar a quebra de linearidade, verificando-se a satisfação de

$$|m'_i - m_i| \leq 2\{f_{ea}\,\bar{e}_a - |m_i|\}/3 \qquad p/i=1,2,\ldots,j \tag{4.27}$$

em caso contrário, deve-se multiplicar o incremento de busca pelo menor dos valores positivos de α_{vi} obtidos pela solução de

$$\alpha_{vi}|m'_i - m^i| = 2[f_{ea}.\bar{e}_a - |m_i|]/3 \tag{4.28}$$

Uma vez obtido um incremento de busca capaz de satisfazer todas estas condições, tem-se, como passo seguinte, a integração numérica para obter a trajetória x'(t) correspondente ao vetor de parâmetros Δa dado por

$$a' = \bar{a} + \Delta\hat{a} \tag{4.29}$$

4.2.3 - INTEGRAÇÃO NUMÉRICA E GERAÇÃO DE DERIVADAS PARCIAIS

A integração numérica foi feita com um integrador de passo fixo que faz uso do método de Runge-Kutta de 4ª ordem.

Escolheu-se, para gerar as derivadas parciais do vetor de estado em relação aos parâmetros, o método de geração de soluções unitárias. Antes de qualquer discussão a respeito deste método numérico de geração das derivadas parciais do vetor de estado em relação aos parâmetros segue-se uma descrição sucinta dos dois métodos de geração destas derivadas, citados na literatura de controle ótimo. São eles o método escolhido e o método da derivação numérica direta.

a) Método de geração numérica direta:

Consiste em gerar, além da trajetória de referência, uma trajetória adicional para cada parâmetro, onde se usa como vetor de pa

râmetro, em cada uma destas trajetórias um vetor idêntico ao usado para gerar a trajetória nominal a exceção de uma componente, que é perturbada de um acréscimo Δ^* de ordem de grandeza apropriada, suficientemente pequeno para que possa ter uma boa aproximação da derivada, porém não tão pequeno a ponto de as variações do vetor de estado se confundirem com os erros numéricos. Assim, deve-se obter tantas trajetórias perturbadas quanto for o número de componentes do vetor de parâmetros n_p.

Uma vez que as derivadas parciais do vetor de estado do sistema dinâmico no instante final em relação aos parâmetros são definidos como:

$$\frac{\partial x(t_f, a)}{\partial a_i} = \lim_{\Delta \to 0} \left[\frac{x(t_f, a + \Delta \bar{e}_i) - x(t_f, a)}{\Delta} \right] \tag{4.30}$$

onde

$\bar{e}_i$ é um vetor que tem a i-ésima componente unitária e as demais componentes nulas;

a_i é a i-ésima componente do vetor de parâmetros a ser otimizado.

pode-se aproximar tais limites por meio de:

$$\frac{\partial x(t_f, a)}{\partial a_i} \cong \frac{x(t_f, a + \Delta^* \bar{e}_i) - x(t_f, a)}{\Delta^*}$$

mediante uma escolha apropriada de Δ^*. A idéai básica deste método é aproximar as derivadas exatas mediante os declives de secantes.

b) Método de geração de soluções unitárias:

Este método faz uso das equações variacionais do sistema dinâmico para obter as derivadas parciais do estado em relação aos

parâmetros. Assim, dado um sistema dinâmico descrito pela equação diferencial

$$\dot{x} = f(x, a, t) \tag{4.31}$$

tem-se, em correspondência, um sistema de equações diferenciais variacionais, que relaciona perturbações nos valores dos parâmetros e nas condições iniciais a variações num dados instante do vetor de estado do sistema dinâmico dado por:

$$\delta\dot{x}(t) = \frac{\partial f(\bar{x}(t),\bar{a},t)}{\partial x} \delta x(t) + \frac{\partial f(\bar{x}(t),\bar{a},t)}{\partial a} \delta a \tag{4.32}$$

onde

$\bar{x}(t)$ é a trajetória correspondente aos valores nominais das condições iniciais, $\bar{x}(t_0)$, e do vetor de parâmetros, $\bar{a}$;

$x(t)$ é a trajetória correspondente aos valores perturbados das condições iniciais e do vetor de parâmetros.

$$\delta x(t) = x(t) - \bar{x}(t)$$

$$\delta a = a - \bar{a}$$

Supondo-se que $\partial f/\partial a$ é contínua, pode-se obter $\delta x(t)$ por meio de

$$\delta x(t) = S(t, t_0)\, x(t_0) + \int_{t_0}^{t} S(t,s) \frac{\partial f}{\partial a} (\bar{x}(s),\bar{a},s)ds \; \delta a \tag{4.33}$$

onde

$S(t,s)$ é a matriz de transição de estado do sistema autônomo dado por

$$\delta\dot{x}(t) = \frac{\partial f}{\partial a} (\bar{x}(t),\bar{a},t) \tag{4.34}$$

matriz esta que pode ser obtida pela integração do sistema de equações diferenciais

$$\dot{S}(t,\ t_0) = \frac{\partial f}{\partial x}\ (\bar{x}(t),\bar{a},t)\ S(t,t_0) \tag{4.35}$$

a partir da condição inicial

$$S(t_0,\ t_0) = I$$

onde I é a matriz identidade.

Como por hipótese as condições iniciais estão relacionadas com o vetor a ser estimado através de uma função conhecida que no caso mais geral pode ser representada por uma função implícita da forma:

$$K(a,\ x(t_0)) = 0 \tag{4.36}$$

pode-se obter a matriz de derivadas do estado inicial em relação aos parâmetros mediante o teorema da função implícita e relacionar $\delta x(t_0)$ e δa por meio de:

$$\delta x(t_0) = -\frac{\partial K}{\partial \bar{a}}\ (\bar{a},\bar{x}(t_0))\ \left[\frac{\partial K(\bar{a},\bar{x}(t_0))}{\partial x(t_0)}\right]^{-1} \delta a = \frac{\partial x(t_0)}{\partial a}\ \delta a \tag{4.37}$$

e escrever

$$\delta x(t) = \left[S(t,t_0)\ \frac{\partial x}{\partial a}\ (t_0) + \int_{t_0}^{t} S(t,s)\frac{\partial f}{\partial a}\ (\bar{x}(s),\bar{a},s)ds\right]\delta a =$$

$$= x_{\bar{a}}(t)\ \delta a \tag{4.38}$$

De acordo com a própria definição de derivadas, verifica-se que $x_{\bar{a}}(t)$ é a matriz de derivadas do vetor de estado em relação aos parâmetros.

Rios Neto e Ceballos (1979) obtiveram uma equação diferencial de propagação desta matriz de derivadas conforme se mostra no Apêndice A.

Comparando-se o método de geração de soluções unitárias ao método de derivação numérica direta, pode-se verificar, com a ajuda do Apêndice A, que o método de derivação numérica direta exige a execução de $n_p + 1$ integrações. Este também é o número de integrações requerido pelo método de geração de soluções unitárias, caso não se resolva propagar a inversa da matriz de transição de estado $S(t,t_0)$ para reduzir o esforço de integração. Caso se resolva propagar tal matriz, pode-se reduzir bastante o esforço de integração se os parâmetros não atuarem em todo o intervalo $[t_0, t_f]$. Tal escolha, porém, exige uma inversão de matrizes, que pode eventualmente resultar em perda de precisão de resultados ou em uma demanda muito alta de tempo computacional para sua execução, ou ainda pode ocorrer que simplesmente não seja possível executar numericamente tal inversão.

Todas as fontes de literatura de controle consultadas são unânimes em destacar o fato de que o método de geração de soluções unitárias, em relação à derivação numérica direta, é mais preciso, porém é de uma maior dificuldade de implementação.

Outro inconveniente do método de geração de soluções unitárias é exigir que se armazene uma quantidade muito grande de dados ou que se execute simultaneamente a integração das equações ligadas à obtenção das derivadas parciais do estado em relação aos parâmetros e a geração da trajetória de referência. A versão deste método, aqui utilizada, tem o último destes dois inconvenientes. Deve-se notar que desta forma ou se arrisca a realizar inutilmente a geração das derivadas parciais do estado em relação aos parâmetros correspondente a uma nova trajetória do vetor de estado ou então deve-se duplicar o esforço de cálculo da nova trajetória, gerando esta nova trajetória para verificar se o incremento de busca é satisfatório e, em caso afirmativo, ge-

rando a esta trajetória novamente para possibilitar a obtenção derivadas das parciais do vetor de estado em relação aos parâmetros.

Pode-se notar desta forma que a escolha do método de geração de derivadas parciais do vetor de estado em relação aos parâmetros é bastante delicada, e que a melhor forma de gerar tais derivadas dependerá muito do problema abordado e das precisões requeridas.

Optou-se aqui pelo uso do método de geração de soluções unitárias e por não gerar nem a matriz de transição de estado nem a sua inversa, nem ainda a nova trajetória do vetor de estado em separado da geração das derivadas para testar se o incremento de busca obtido é ou não satisfatório. Escolheu-se, dentre as várias versões do método de geração de soluções unitárias, a proposta por Rios Neto e Ceballos (1979). Tal escolha foi satisfatória do ponto de vista do desempenho numérico e trouxe dificuldades no tocante à implementação numérica como era de se esperar, pelo fato de se ter uma rotina de integração específica para cada tipo de funcional usado na aproximação da função de controle ótimo.

4.2.4 - ANÁLISE FINAL DOS RESULTADOS

Obtida uma nova trajetória, deve-se verificar se o incremento de busca obtido é realmente satisfatório. Durante a fase de captura dos vínculos deve-se verificar se o incremento de busca obtido realmente provoca o decréscimo do erro nos vínculos de contorno, a menos que se esteja sobre um ponto estacionário do erro nos vínculos onde se tenham valores insatisfatórios destes erros. Durante a fase de busca do mínimo, deve-se verificar se o incremento de busca realmente provoca o decréscimo da função objetivo sem permitir que os erros nos vínculos de contorno saiam da faixa de tolerância especificada para estes erros. Caso o incremento de busca atenda a tais condições atualiza-se o vetor de parâmetros a ser otimizado, ou seja, considera-se como o novo valor nominal do vetor de parâmetros a soma do valor nominal anterior ao incremento de busca no vetor de parâmetros a ser otimizado.

Assim, na fase de captura dos vínculos, deve-se verificar, caso não se esteja sobre um ponto estacionário do erro nos vínculos, se o módulo do erro nos vínculos realmente diminuiu, ou seja, se:

$$|[M^T(x'(t_f),\bar{a}+\Delta\bar{a}) : I^T(\bar{a}+\Delta\bar{a})]| < |[M^T(\bar{x}(t_f),\bar{a}) : I^T(\bar{a})]| \quad (4.39)$$

pnde $x'(t)$ é a trajetória correspondente ao controle $u(\bar{a}+\Delta\bar{a},t)$, obtida por integração numérica.

Deve-se considerar que se está sobre um ponto estacionário do erro nos vínculos de contorno quando não se puder encontrar, na fase de captura dos vínculos, um incremento de busca capaz de satisfazer as Desigualdades 4.9, 4.10 e 4.39, mesmo com valores de p tais que a Desigualdade 4.12 seja satisfeita.

Caso a Desigualdade 4.39 não seja satisfeita, deve-se multiplicar o incremento de busca obtido por um escalar α_ℓ pertencente ao intervalo (0,1), se algum dos valores de $\bar{\rho}_{ii}$ for menor que dois, e voltar a verificar a satisfação da Desigualdade 4.39. Caso esta não seja satisfeita e todos os $\bar{\rho}_{ii}$ sejam maiores do que dois, deve-se diminuir p e refazer o cálculo do incremento de busca no vetor de parâmetros, repetindo-se a seguir o processo de verificação. Um bom valor para α_ℓ é o obtido com a hipótese de que os termos de segunda ordem sejam responsáveis pela não-linearidade, a partir da solução de:

$$|E(\bar{a}) + \alpha_\ell[E' - E(a)] + \alpha_\ell^2[E^* - E']| < |E(\bar{a})| \quad (4.40)$$

onde

- E(a) é o vetor de erros nos vínculos de contorno correspondente ao valor nominal do vetor de parâmetros, definido na Equação 4.8;
- E' é o vetor de estimativas lineares do erro nos vínculos de contorno associado à soma do incremento de busca ao valor nominal do vetor de parâmetros, definido na Equação 4.8;

E* é o vetor de erro nos vínculos associado à soma do incremento de busca ao valor nominal do vetor de parâmetros definido por

$$E^{*T} \triangleq [M^T(x'(t_f), \bar{a} + \Delta\hat{a}) : I^T(\bar{a} + \Delta\hat{a})] \tag{4.41}$$

Na fase de busca do mínimo, deve-se verificar se todas as componentes do erro nos vínculos de contorno no instante final estão realmente contidas na faixa de tolerância especificada, ou seja, caso se defina m_i'' como as componentes do erro nos vínculos de contorno no instante final associadas a $\bar{a} + \Delta\hat{a}$, através de:

$$[m_1'' , m_2'' , \ldots, m_\ell''] \triangleq M^T(x'(t_f), \bar{a} + \Delta\hat{a}) \tag{4.42}$$

tem-se que:

$$|m_i''| < f_{ea}\ \bar{e}_a \qquad p/i=1,2,\ldots,\ell \tag{4.43}$$

Caso tal condição não seja satisfeita, deve-se multiplicar o incremento de busca obtido por um escalar α_ℓ pertencente ao intervalo (0,1), voltando-se a verificar a seguir se a Desigualdade 4.43 é satisfeita. Um bom valor de α_ℓ é o obtido mediante a hipótese dos termos de segunda ordem serem os únicos responsáveis pela não-linearidade, a partir da solução de:

$$|m_n + \alpha_\ell(m_n' - m_n) + \alpha_\ell^2(m_n'' - m_n')| \leq f_{ea}\ e_a \tag{4.44}$$

onde n é o índice da componente do erro nos vínculos de contorno no instante final, $M(x'(t_f), \bar{a} + \Delta\hat{a})$, que não satisfaz à Desigualdade 4.43 para a qual $|m_i''|$ é máximo.

No caso da função objetivo depender do vetor de estado no instante final, deve-se verificar se esta realmente diminuiu, ou seja, se:

$$G(x'(t_f),\ \bar{a} + \Delta\hat{a}) < G(\bar{x}(t_f),\ \bar{a}) \tag{4.45}$$

caso não se tenha que

$$\frac{\partial G(x(t_f),\ a)}{\partial x(t_f)} \equiv 0 \tag{4.46}$$

Caso a Desigualdade 4.45 e a Equação 4.46 não se verifiquem, deve-se multiplicar o incremento de busca obtido por um escalar α_ℓ pertencente ao intervalo (0,1), tornando, a seguir, a verificar a satisfação da Desigualdade 4.45. Um bom valor de α_ℓ é o obtido a partir da hipótese de que os termos de segunda ordem sejam os únicos responsáveis pela não-linearidade, através da solução de:

$$G(\bar{x}(t_f),\ \bar{a}) + G_{\bar{a}}\Delta\bar{a}\alpha_\ell + [G(x'(t_f),\ \bar{a} + \Delta\hat{a}) - G(\bar{x}(t_f),\ \bar{a}) -$$

$$- G_{\bar{a}}\Delta\hat{a}]\alpha_\ell^2 < G(\bar{x}(t_f),\ \bar{a}) \tag{4.47}$$

Caso não seja possível achar, na fase de busca do mínimo, um incremento de busca satisfatório a despeito de:

$$\Delta^T\hat{a}\Delta\hat{a} \cdot \max\{h_i h_i^T,\ G_{\bar{a}}G_{\bar{a}}^T/[(1 + G^2(\bar{x}(t_f),\ \bar{a}))\ \alpha_\ell^2]\} < e_m^2 \tag{4.48}$$

deve-se considerar que a solução convergiu, uma vez que, em tal caso, as variações da função objetivo e do erro nos vínculos de contorno se confundem com as incertezas numéricas.

Finalmente, cabe frisar que também integram o algoritmo numérico critérios que se destinam a verificar se é ou não conveniente ajustar os parâmetros $\bar{\beta}$, $\bar{\beta}_p$ e p para aumentar a velocidade de convergência, a despeito de se ter obtido uma estimativa do incremento de busca, que satisfaça, na fase de captura dos vínculos, à Desigualdade 4.39, ou na fase de busca do mínimo, às Desigualdade 4.21, 4.43 e a 4.24 ou 4.45. Critérios estes que garantem que o incremento de busca

é satisfatório no tocante ao fato de levar a um ponto mais próximo da solução do problema restrito de otimização. Ou até mesmo se é conveniente reprocessar o cálculo do incremento de busca no vetor de parâmetros ou filtrar medidas adicionais z_i''' (ver as Equações 3.13 e 3.14) para agilizar ainda mais a convergência. Tais critérios de ajuste de $\bar{\beta}$, $\bar{\beta}_p$ e p, e os processos que podem opcionalmente ser aplicados para agilizar o algoritmo são descritos no Apêndice B.

Obtida uma estimativa que garanta a caminhada na direção da solução, pode-se adotar como o novo valor do vetor de parâmetros a soma de tal estimativa ao valor nominal anterior, fazendo-se

$$\hat{a} = \bar{a} + \Delta\hat{a}$$

e após isto, fazendo-se

$$\bar{a} = \hat{a}$$

Também é apresentada, no Apêndice C, uma descrição mais completa do algoritmo sob a forma de uma sequência de passos.

CAPÍTULO 5

TESTE NUMÉRICO DO PROCEDIMENTO

5.1 - INTRODUÇÃO

Escolheu-se o problema de transferência de órbita Terra-Marte em tempo mínimo com empuxo constante e baixo e trajetória plana para um primeiro teste do procedimento proposto, pelo fato deste problema ser clássico em trabalhos de implementação e teste de métodos numéricos de controle ótimo como os trabalhos de Willianson (1971), Hull e Edgeman (1975), Ceballos (1979) e Rios Neto e Ceballos (1979), ou em trabalhos de comparação entre métodos de otimização como o de Tapley e Lewallen (1967), no qual se analisam comparativamente os métodos ótimos diretos e os indiretos de perturbação e quasi-linearização.

5.2 - COLOCAÇÃO DO PROBLEMA DE TRANSFERÊNCIA DE ÓRBITA TERRA-MARTE EM TEMPO MÍNIMO COM TRAJETÓRIA PLANA E EMPUXO CONSTANTE E BAIXO

O problema de transferência de órbita Terra-Marte em tempo mínimo com trajetória plana e empuxo constante e baixo, que será doravante denominado simplesmente problema Terra-Marte, pode ser colocado como a seguir:

Minimizar

$$IP = t_f$$

sujeito a vínculos dinâmicos

$$\dot{x}_1(t) = x(t)$$

$$\dot{x}_2(t) = -\frac{\mu}{x_1^2(t)} + \frac{x_3^2(t)}{x_1(t)} + \frac{T \sin u(t)}{(m_0 - \dot{m}t)} \qquad (5.1)$$

$$\dot{x}_3(t) = - \frac{x_2(t)\, x_3(t)}{x_1(t)} + \frac{T \cos u(t)}{(m_0 - \dot{m}t)}$$

e a vínculos de contorno em t_f

$$x_2(t_f) = 0$$

$$x_3(t_f) - \sqrt{\frac{\mu}{x_1(t_f)}} = 0$$

$$x_1(t_f) - r_f = 0$$

e com condições iniciais preestabelecidas

$$x_1(t_0) = r_0$$

$$x_2(t_0) = 0$$

$$x_3(t_0) = \sqrt{\frac{\mu}{r_0}}$$

onde (Figura 5.1):

$x_2(t)$ é a velocidade radial;
$x_1(t)$ é o módulo do raio vetor que vai do centro do Sol ao móvel;
$x_3(t)$ é a velocidade tangencial;
T é o empuxo;
m_0 é a massa inicial do móvel;
$\dot{m}$ é a taxa de expulsão de massa pelo bocal do foguete;
t é o tempo;
$u(t)$ é o controle, definido como o ângulo entre a direção do empuxo e a velocidade do móvel;
t_0 é o tempo inicial, que será nulo;
t_f é o tempo final;
μ é a constante heliogravitacional.

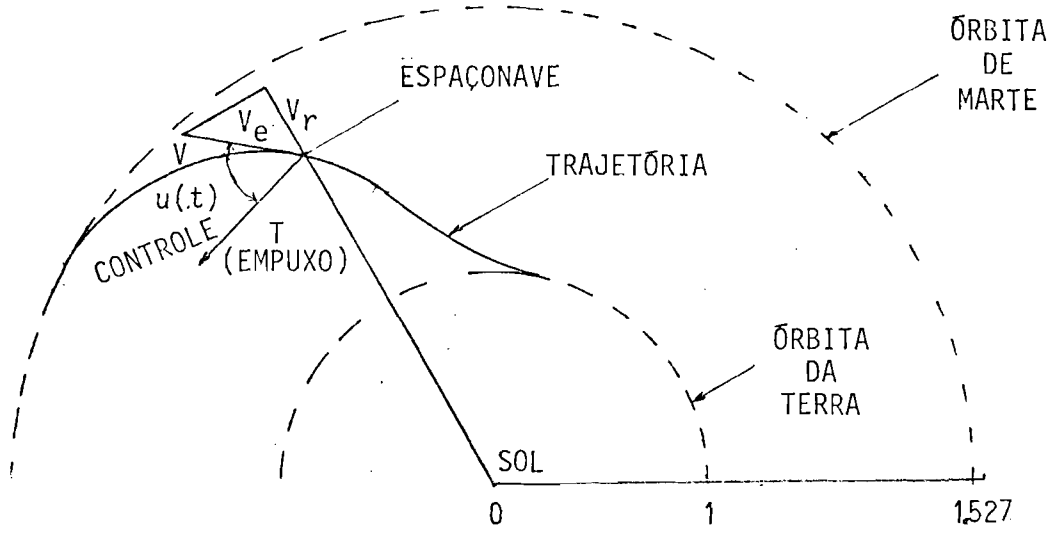


Fig. 5.1 - Transferência de órbita Terra-Marte.

Tratando o problema de forma normalizada em relação à base r_0, μ, m_0 (Bryson and Ho, 1975), têm-se os seguintes valores para os parâmetros do sistema para o caso da espaçonave "Viking II":

$\mu = 1$

$\dot{m} = 0.074800391$

$m_0 = 1$

$T = 0.14012969$

$r_0 = 1$

$r_f = 1.523679$

A solução ótima deste problema corresponde a um tempo de chegada em uma órbita circular em torno do Sol com raio igual ao da órbita de Marte de 3.31949 unidades adimensionais de tempo, e que corresponde a cerca de 193 dias ($3.31949 \times 5.022642 \times 10^6$ seg).

5.3 - APROXIMAÇÃO PARA O CONTROLE

5.3.1 - INTRODUÇÃO

A função de interpolação escolhida depende do problema a ser resolvido e deve:

a) ter um número de parâmetros suficientemente grande para permitir a satisfação dos vínculos de contorno e a minimização do índice de desempenho;

b) ser compatível com a forma da função de controle ótimo;

c) ter parâmetros que sejam o mais independentes possível entre si;

d) ser de forma tal que a sensibilidade do controle em relação aos parâmetros seja mais ou menos uniforme;

e) ser também de forma tal que aos desvios nos parâmetros não correspondam desvios no controle de ordem de grandeza muito diversa.

Ceballos (1979) afirma que os funcionais particionados apresentam boas características no tocante aos dois últimos aspectos.

Independentemente da escolha da função de interpolação adotaram-se valores de partida para os parâmetros, tais que o controle inicial fosse dado por uma reta da forma

$$u(a,t) = \pi\theta t/t_{fi}$$

e os tempos de quina, que em alguns casos eram diretamente parâmetros, obedecessem

$$t_i = t_{fi} \cdot 1/k$$

onde

t_{fi} é o valor de partida t_f;
θ é uma constante arbitrária;
k é o número de partições efetuadas;
t_i é o i-ésimo tempo nodal.

Testou-se o procedimento com vários tipos de funções de interpolação para o controle, além dos descritos a seguir, usando-se como soluções de partida valores de t_f bastante diversos (no intervalo 0,5 a 50) e em todos os casos obtiveram-se bons resultados. Decidiu-se apresentar apenas alguns dos resultados obtidos com os 4 tipos de funções de interpolação por dois motivos:

a) permitir uma comparação direta em termos de desempenho numérico entre o procedimento proposto e outros procedimentos numéricos subótimos;

b) não tornar a apresentação de resultados enfadonha.

5.3.2 - APROXIMAÇÃO DO CONTROLE POR INTERPOLAÇÃO LINEARES COM TEMPOS INTERMEDIÁRIOS PROPORCIONAIS AO TEMPO FINAL E IMPOSIÇÃO DE CONTINUIDADE NO CONTROLE

As aproximações por interpolações lineares possuem, segundo Ceballos (1979), ótimas características e são sobretudo de grande valia quando a forma da função de controle ótimo é completamente desconhecida ou quando não se possui uma formulação que se adapte às características do controle. Em particular, o tipo de interpolação linear descrito nesta subseção tem sido empregado com relativa frequência na literatura. Na interpolação em questão, subdivide-se o intervalo $[t_0, t_f]$ em k intervalos $[t_{j-1}, t_j]$, onde t_j é dado por:

$$t_j = (j - 1)t_f/k$$

e utiliza-se como parâmetros o tempo final e os valores do controle nos tempos nodais, calculando-se u(a,t) no intervalo $[t_j, t_{j+1}]$ por meio de

$$u(a,t) = (a_{j+1} - a_j)(t - t_j)/(t_{j+1} - t_j) + a_j$$

onde:

a_j - para j=1, k+1, é o valor do controle no instante $t=t_{j-1}$;
a_{k+2} - é o valor do tempo final t_f;
a_0 - é o valor do tempo inicial, que será sempre igual a zero e não é um parâmetro a ser otimizado.

Neste tipo de aproximação o número total de parâmetros n_p será dado por:

$$n_p = k + 2$$

5.3.3 - APROXIMAÇÃO DO CONTROLE POR SEGMENTOS DE RETA NÃO OBRIGATORIAMENTE CONTÍNUOS NOS TEMPOS NODAIS E COM TEMPOS NODAIS A OTIMIZAR

Neste tipo de aproximação, subdivide-se o intervalo $[t_0, t_f]$ em k subintervalos $[a_{3j+3}, a_{3j}]$, com j variando de zero a k-1, com a_0 igual a zero por definição e com a_{3k} igual a t_f, e aproxima-se o controle dentro de cada um destes intervalos por um segmento de reta, permitindo-se que o controle seja eventualmente descontínuo nos tempos nodais (a_{3j} com j=1, k-1). Escolhendo-se aqui como parâmetros a otimizar os valores do controle nos tempos nodais (dois por tempo nodal) e os próprios tempos nodais, tem-se que, no j-ésimo intervalo, pode-se obter o controle u(t) por meio de:

$$u(t) = (a_{3j-1} - a_{3j-2})(t - a_{3j-3})/(a_{3j} - a_{a3j-3}) + a_{3j-2}$$

onde

a_{3j-3} é o instante onde se começa a aproximar o controle pelo j-ésimo segmento de reta;

a_{3j} é o instante onde se deixa de aproximar o controle pelo j-ésimo segmento de reta;

a_{3j-1} é o limite pela esquerda para o valor do controle no instante a_{3j};

a_{3j-2} é o limite pela direita para o valor do controle no instante a_{3j-3}.

5.3.4 - APROXIMAÇÃO DO CONTROLE POR ARCOS DE PARÁBOLA NÃO OBRIGATORIAMENTE CONTÍNUOS NOS TEMPOS NODAIS E COM TEMPOS NODAIS A OTIMIZAR

Este tipo de aproximação é semelhante ao da subseção anterior. Divide-se, neste caso, o intervalo entre t_0 e t_f em k intervalos, onde o controle é aproximado por um arco de parábola e permite-se que o controle seja eventualmente descontínuo nos tempos nodais. Neste caso o controle u(a,t), é obtido entre os tempo nodais a_{4i+4} e a_{4i} por meio de

$$u(t) = \left[\frac{a_{4i-1} - a_{4i-3}}{(a_{4i} - a_{4i-4})^2} - \frac{a_{4i-2}}{(a_{4i} - a_{4i-4})} \right] (t - a_{4i-4})^2$$

$$a_{4i-2}(t - a_{4i-4}) + a_{4i-3}$$

onde

a_{4i-4} é o instante onde o controle passa a ser aproximado pelo i-ésimo arco de parábola;

a_{4i} é o instante onde o controle deixa de ser aproximado pelo i-ésimo arco de parábola;

a_{4i-3} é o limite pela direita para o controle no instante a_{4i-4};

a_{4i-1} é o limite pela esquerda para o controle no instante a_{4i};

a_{4i-2} é um parâmetro;

a_0 é o tempo inicial e não é um parâmetro a otimizar, tendo valor nulo.

5.3.5 - APROXIMAÇÃO DO CONTROLE POR TRECHOS DE VALORES CONSTANTES A OTIMIZAR E COM INTERVALOS DE DURAÇÃO TAMBÉM A OTIMIZAR

Neste tipo de solução, para i de um a k, aproxima-se o controle por um valor constante a otimizar a_{2i-1}, entre os tempos nodais a_{2i-2} e a_{2i} que também são parâmetros a otimizar, a exceção de a_0 que corresponde ao tempo inicial e terá valor sempre nulo.

Este tipo de função de interpolação permite que o controle seja permutado por sinais de Terra e também permite que se constate mais facilmente desvios entre o controle aplicado e o valor de projeto deste. Tais fatos implicam um aumento de confiabilidade que, caso se pretendesse realmente achar uma boa solução sob o ponto de vista de Engenharia para o problema em questão, não poderia ser ignorado. A otimização dos instantes de mudança do valor do controle permite maior flexibilidade e a redução do número de sinais a serem enviados com o intuito de alterar o valor do controle. No caso particular do problema Terra-Marte, o uso de tal tipo de aproximação acarreta uma série de problemas numéricos. Em primeiro lugar, o controle não depende dos tempos nodais de forma explícita, o que faz surgir derivadas do estado final em relação aos parâmetros com significado de tempos eventualmente muito baixas, consequentemtne os tempos nodais sofrem alterações muito pequenas, acarretando problemas numéricos ao processo de cálculo do incremento de busca. Outro ponto negativo deste tipo de solução é o fato de algumas derivadas de segunda ordem terem valor bastante alto face às derivadas de primeira ordem, destacando-se as derivadas

$$\frac{\partial x(t_f)}{\partial a_{2i-1}\partial a_{2i}}, \quad \frac{\partial x(t_f)}{\partial a_{2i}\partial a_{2i-1}}, \quad \frac{\partial x(t_f)}{\partial a_{2i-2}\partial a_{2i-1}}, \quad \frac{\partial x(t_f)}{\partial a_{2i-1}\partial a_{2i-2}}.$$

Outro aspecto negativo deste tipo de solução está relacionado com a reprodução de uma característica provavelmente importante da função de controle ótimo. Tal função que é contínua apresenta, nem trecho intermediário, uma brusca bariação do controle. Em todos os testes numéricos com os demais tipos de funcional esta característica foi reproduzida pelas funções de controle subótimas obtidas. Nos testes numéricos efetuados com o tipo de função tratado nesta subseção, pode-se dizer que, de uma certa forma, as funções de controle obtidas tentam reproduzir esta característica, e o fazem concentrando alguns patamares de duração muito curta num trecho intermediário. Esta proximidade dos tempos nodais causa uma série de problemas numéricos, dentre os quais a obtenção de valores para os novos tempos nodais que não estão ordenados e o aparecimento de derivadas do estado em relação aos parâmetros muito baixas, que fazem com que se tenha dificuldade em fazer tal concentração de segmentos se movimentar. Por último dá-se ênfase ao fato de que, para tal tipo de solução, aparece um grande número de mínimo locais.

Por todos estes aspectos, acredita-se que este tipo de função de interpolação não deve permitir a obtenção de resultados melhores do que os obtidos com este procedimento com métodos de primeira ordem. Efetuaram-se testes numéricos, empregando-se 3, 4, 5, 6, 7, 8, 9 e 10 patamares, porém resumem-se aqui os resultados obtidos com a aproximação de controle por 3 e 5 patamares (este último o melhor dentre os 8 resultados).

5.4 - RESUMO DOS RESULTADOS OBTIDOS

Nesta seção serão apresentados alguns resultados obtidos num total de 7 exemplos numéricos. Os resultados destes testes são apresentados nas Tabelas 5.1 a 5.14, onde poder-se-ão encontrar: os valores iniciais e finais dos parâmetros a serem otimizados; os valores do tempo final e do máximo valor em módulo das componentes do erro nos vínculos de contorno em algunas iterações intermediárias, bem como valores iniciais e finais destes.

Para maior simplicidade de referência e para efeito de preenchimento de tabelas define-se:

TESTE 1 - Aproximação do controle por 4 segmentos de reta com tempos nodais a otimizar e controle eventualmente descontínuo nos tempos nodais.

TESTE 2 - Aproximação do controle por 3 segmentos de reta com tempos intermediários a otimizar e controle eventualmente descontínuo nos tempos nodais.

TESTE 3 - Aproximação do controle por 2 arcos de parábola com tempos nodais a otimizar e controle eventualmente descontínuo nos tempos nodais.

TESTE 4 - Aproximação do controle por 5 segmentos de reta com imposição de continuidade ao controle e com tempos nodais proporcionais a t_f.

TESTE 5 - Aproximação do controle por 7 segmentos de reta com imposição de continuidade ao controle e com tempos nodais proporcionais a t_f.

TESTE 6 - Aproximação do controle por 3 trechos de valor constante a otimizar e com tempos nodais a otimizar.

TESTE 7 - Aproximação do controle por 5 trechos de valos constante a otimizar e com tempos nodais a otimizar.

MAXMOD - Máximo valor em módulo das componentes do erro nos vínculos de contorno.

totint - Número total de iterações.

NINT - Número de iterações.

TESTE 1

$p = 1.0000$ $\quad \beta_0 = 1.0000$ $\quad \bar{e}_a = 1.0000.10^{-3}$

$\bar{\beta} = 1.0000$ $\quad \beta_{p^0} = 2.0000$ $\quad f_{ea} = 1.0000$

$\bar{\beta}_p = 1.0000$ $\quad e_m = 1.0000.10^{-4}$ $\quad \alpha_{\bar{\beta}} = \alpha_{\bar{\beta}p} = 10/12$ $\quad$ totint = 11

TABELA 5.1

RESULTADOS INICIAIS E FINAIS DO TESTE 1

PARÂMETROS	VALORES INICIAIS	VALORES FINAIS
a_1	0.00000	.289500
a_2	1.17810	1.18379
a_3	0.85000	1.18834
a_4	1.17810	0.991843
a_5	2.35619	2.15020
a_6	1.70000	1.64033
a_7	2.35619	2.38118
a_8	3.53429	3.65197
a_9	2.55000	1.68652
a_{10}	3.53429	4.38888
a_{11}	4.71239	5.58525
$a_{12} = t_f$	3.40000	3.31943

TABELA 5.2

HISTÓRICO DO TESTE 1

NINT	MAXMOD	TF
0	.318669	3.40000
1	.233063	2.96329
2	.138704	2.95116
3	$3.48882 \cdot 10^{-2}$	3.11199
4	$9.55840 \cdot 10^{-3}$	3.32410
5	$3.69506 \cdot 10^{-4}$	3.34184
7	$2.36664 \cdot 10^{-4}$	3.33379
9	$5.48528 \cdot 10^{-4}$	3.32706
10	$3.60156 \cdot 10^{-4}$	3.32576
11	$6.51884 \cdot 10^{-4}$	3.31943

TESTE 2

$p = 1.0000$ $\quad \beta_0 = 1.0000$ $\quad \bar{e}_a = 1.0000.10^{-3}$

$\bar{\beta} = 1.0000$ $\quad \beta_{p0} = 2.0000$ $\quad f_{ea} = 1.0000$

$\bar{\beta}p = 1.0000$ $\quad e_m = 1.0000.10^{-4}$ $\quad \alpha_{\bar{\beta}=\alpha_{\bar{\beta}p}} = 10/12$ $\quad$ totint = 15

TABELA 5.3

RESULTADOS INICIAIS E FINAIS DO TESTE 2

PARÂMETROS	VALORES INICIAIS	VALORES FINAIS
a_1	0.00000	.358435
a_2	1.57080	1.37666
a_3	1.13333	1.51051
a_4	1.57080	1.41639
a_5	3.14159	3.09568
a_6	2.26666	1.68340
a_7	3.14159	4.30456
a_8	4.71239	5.63402
a_9	3.40000	3.32366

TABELA 5.4

HISTÓRICO DO TESTE 2

NINT	MAXMOD	TF
0	.318669	3.40000
1	.222261	2.95054
2	$9.93611 \cdot 10^{-2}$	2.29061
3	$6.82145 \cdot 10^{-2}$	3.25217
5	$2.21439 \cdot 10^{-4}$	3.35667
6	$2.33100 \cdot 10^{-4}$	3.35198
8	$8.41085 \cdot 10^{-4}$	3.34046
10	$5.99413 \cdot 10^{-4}$	3.33866
12	$3.42107 \cdot 10^{-4}$	3.333292
14	$8.60444 \cdot 10^{-4}$	3.32558
15	$9.78641 \cdot 10^{-4}$	3.32366

TESTE 3

$p = 1.0000$ $\quad \beta_0 = 1.0000$ $\quad \bar{e}_a = 1.0000 \cdot 10^{-2}(1.0000 \cdot 10^{-3})$

$\bar{\beta} = 1.0000$ $\quad \beta_{p0} = 4.0000$ $\quad f_{ea} = 1.0000$

$\bar{\beta}_p = 1.0000$ $\quad e_m = 1.0000.10^{-4}$ $\quad \alpha_{\bar{\beta}} = \alpha_{\bar{\beta}p} = 10/12$ $\quad$ totint = 12

OBS.: Convergiu-se com $e_a = 10^{-2}$ e reiniciou-se o processo com $e_a = 10^{-3}$.

TABELA 5.5

RESULTADOS INICIAIS E FINAIS DO TESTE 3

PARÂMETROS	VALORES INICIAIS	VALORES FINAIS
a_1	0.00000	.247226
a_2	1.38600	1.07125
a_3	2.35619	1.53430
a_4	1.70000	1.65574
a_5	2.35619	3.80928
a_6	1.38600	2.11300
a_7	4.71239	5.39833
a_8	3.40000	3.32250

TABELA 5.6

HISTÓRICO DO TESTE 3

NINT	MAXMOD	T
0	.318669	3.40000
1	.172470	2.92121
2	.111434	3.17247
3	$4.47206 \cdot 10^{-2}$	3.38230
4	$2.91976 \cdot 10^{-3}$	3.40738
5	$1.22413 \cdot 10^{-3}$	3.38990
8	$1.97272 \cdot 10^{-3}$	3.33440
10	$5.77165 \cdot 10^{-4}$	3.32692
11	$8.07622 \cdot 10^{-4}$	3.32250

TESTE 4

$p = 1.0000$ $\quad \beta_0 = 1.0000$ $\quad \bar{e}_a = 1.0000.10^{-2}$ $(1.0000.10^{-3})$

$\bar{\beta} = 1.0000$ $\quad \beta_{p0} = 1.0000$ $\quad = 1.0000$

$\bar{\beta}_p = 1.0000$ $\quad e_m = 1.0000.10^{-4}$ $\quad \alpha_{\bar{\beta}=\alpha_{\bar{\beta}p}} = 10/12$ $\quad$ totint = 14

OBS.: Convergiu-se com $\bar{e}_a = 1.0000.10^{-2}$ e reaplicou-se o algoritmo com $\bar{e}_a = 1.0000.10^{-3}$

TABELA 5.7

RESULTADOS INICIAIS E FINAIS DO TESTE 4

PARÂMETROS	VALORES INICIAIS	VALORES FINAIS
a_1	0.00000	.372881
a_2	.942478	.736625
a_3	1.88496	1.19244
a_4	2.82743	4.68391
a_5	3.76991	5.38817
a_6	4.71239	5.40535
a_7	3.40000	3.32533

TABELA 5.8

HISTÓRICO DO TESTE 4

NINT	MAXMOD	T
0	.318669	3.40000
1	.184616	2.92629
2	$8.11394 . 10^{-2}$	3.03804
4	$6.43928 . 10^{-3}$	3.44218
6	$2.24303 . 10^{-4}$	3.42719
8	$2.23497 . 10^{-4}$	3.37679
10	$6.57909 . 10^{-4}$	3.34705
12	$1.93667 . 10^{-3}$	3.34014
13	$4.67270 . 10^{-4}$	3.33139
14	$7.71669 . 10^{-4}$	3.32533

TESTE 5

$p = 1.6000$ $\quad \beta_0 = 1.0000$ $\quad \bar{e}_a = 1.0000.10^{-2}$ $(1.0000.10^{-3})$

$\bar{\beta} = 1.0000$ $\quad \beta_{p_0} = 1.0000$ $\quad f_{ea} = 1.0000$

$\bar{\beta}_p = 1.0000$ $\quad e_m = 1.0000.10^{-4}$ $\quad \alpha_{\bar{\beta}} = \alpha_{\bar{\beta}p} = 10/12$ $\quad$ totint = 17

TABELA 5.9

RESULTADOS INICIAIS E FINAIS DO TESTE 5

PARÂMETROS	VALORES INICIAIS	VALORES FINAIS
a_1	0.000000	.396428
a_2	.673198	.664570
a_3	1.34640	.904146
a_4	2.01960	1.41252
a_5	2.69279	4.41106
a_6	3.36599	5.25164
a_7	4.03919	5.25711
a_8	4.71239	5.48428
a_9	3.40000	3.32161

TABELA 5.10

HISTÓRICO DO TESTE 5

NINT	MAXMOD	T
0	.318669	3.40000
1	.148853	2.90035
2	.119369	3.08686
3	$5.60715 \cdot 10^{-2}$	3.21429
4	$8.53587 \cdot 10^{-3}$	3.38687
8	$2.86371 \cdot 10^{-4}$	3.36276
12	$1.84103 \cdot 10^{-3}$	3.32131
14	$2.84544 \cdot 10^{-3}$	3.31191
15	$1.50753 \cdot 10^{-3}$	3.33022
16	$8.82830 \cdot 10^{-5}$	3.32437
17	$6.83774 \cdot 10^{-4}$	3.32161

TESTE 6

$p = 1.0000$ $\quad \beta_0 = 1.0000$ $\quad \bar{e}_a = 1.0000.10^{-2}(1.0000.10^{-3})$

$\bar{\beta} = 1.0000$ $\quad \beta_{p0} = 1.0000$ $\quad f_{ea} = 1.0000$

$\bar{\beta}_p = 1.0000$ $\quad e_m = 1.0000.10^{-4}$ $\quad \alpha_{\bar{\beta}=\alpha_{\bar{\beta}p}} = 10/12$ $\quad$ totint = 15

OBS.: Convergiu-se com $\bar{e}_a = 1.0000.10^{-2}$, e reiniciou-se o processo com $\bar{e}_a = 1.000.10^{-3}$.

TABELA 5.11

RESULTADOS INICIAIS E FINAIS DO TESTE 6

PARÂMETROS	VALORES INICIAIS	VALORES FINAIS
a_1	0.00000	.719458
a_2	1.13333	1.00604
a_3	0.00000	1.81369
a_4	2.26666	1.77316
a_5	0.00000	5.11003
a_6	3.40000	3.36526

TABELA 5.12

NINT	MAXMOD	T
0	.597689	3.40000
1	.482852	2.53233
5	.222976	2.68377
8	$2.07637 . 10^{-3}$	3.42875
10	$4.37907 . 10^{-3}$	3.35798
11	$8.93271 . 10^{-4}$	3.37307
13	$5.21191 . 10^{-4}$	3.36646
15	$9.79355 . 10^{-4}$	3.36256

TESTE 7

$p = .8000$ $\quad \beta_0 = 1.0000$ $\quad \bar{e}_a = 1.0000.10^{-2}$ $(1.0000.10^{-3})$

$\bar{\beta} = 1.0000$ $\quad \beta_{p^0} = 4.0000$ $\quad f_{ea} = 1.0000$

$\bar{\beta}_p = 1.0000$ $\quad e_m = 1.0000.10^{-4}$ $\quad \alpha_{\bar{\beta}} = \alpha_{\bar{\beta}p} = 10/12$ $\quad$ totint = 19

OBS.: Convergiu-se com $\bar{e}_a = 1.0000.10^{-2}$ na 8ª iteração e reiniciou-se o processo com $\bar{e}_a = 1.0000.10^{-3}$.

TABELA 5.13

RESULTADOS INICIAIS E FINAIS DO TESTE 7

PARÂMETROS	VALORES INICIAIS	VALORES FINAIS
a_1	0.000000	.647703
a_2	.680000	.814338
a_3	0.000000	1.42172
a_4	1.360000	1.32329
a_5	0.000000	2.07461
a_6	2.040000	1.73312
a_7	0.000000	1.09381
a_8	2.720000	1.73427
a_9	0.000000	5.12474
a_{10}	3.400000	3.35146

TABELA 5.14

HISTÓRICO DO TESTE 7

NINT	MAXMOD	T
0	0.597686	3.40000
1	0.285587	1.72477
2	0.237436	2.27736
3	$9.16946 \cdot 10^{-2}$	3.42365
4	$6.68645 \cdot 10^{-3}$	3.37625
6	$9.13578 \cdot 10^{-4}$	3.37258
8	$5.70325 \cdot 10^{-3}$	3.35425
9	$4.14024 \cdot 10^{-3}$	3.36030
16	$1.43169 \cdot 10^{-3}$	3.35850
18	$1.38890 \cdot 10^{-4}$	3.35848
19	$9.13567 \cdot 10^{-3}$	3.35146

5.5 - CONCLUSÕES

Os resultados obtidos nos vários testes podem ser considerados bons dado que, a exceção dos teste 6 e 7, onde se usaram funções de interpolação provavelmente incompatíveis com a forma da função de controle ótimo obtiveram-se desvios para o valor da função objetivo bastante baixos. Ao pior dos cinco primeiros resultados corresponde um tempo final de 3.32533, ou seja, um erro no índice de desempenho de 0.176%. Erros desta ordem são, em muitas das aplicações da engenharia, inferiores aos erros de realização do controle pelo sistema, perdendo assim o seu significado como erro. No teste 1 obteve-se um tempo final ligeiramente inferior ao tempo ótimo (3.31943 contra um tempo ótimo de 3.31949), devido ao fato de não se satisfazerem exatamente os vínculos de contorno.

Também merece destaque o fato de que o procedimento proposto mostrou uma boa velocidade de convergência, tendo-se mostrado, em vários testes mais veloz que os demais procedimentos subótimos disponíveis na literatura. Em nenhum caso ele apresentou um desempenho muito inferior, em termos de velocidade de convergência, aos demais procedimentos deste tipo disponíveis na literatura.

CAPÍTULO 6

COMENTÁRIOS E CONCLUSÕES FINAIS

O procedimento proposto é bastante promissor como método de otimização dinâmica, os resultados obtidos vieram confirmar plenamente as expectativas da fase de denvolvimento. Tais resultados foram animadores tanto sob o ponto de vista de precisão numérica como de tempo gasto por iteração e do número de iterações necessárias para atingir a convergência.

Na maioria dos problemas de otimização dinâmica, o tempo de integração é responsável pela maior parte do tempo gasto por iteração. No caso dos métodos subótimos por restrição no controle o esforço de integração depende apenas da ordem do método, do número de parâmetros da função u(a,t), das características das equações diferenciais em si e do método de geração de derivadas parciais escolhido. Nos métodos subótimos, por restrição no controle de primeira ordem, necessita-se integrar um sistema de "neqn" equações diferenciais, onde "neqn" é a dimensão do vetor de estado do sistema dinâmico, para gerar a trajetória de referência, e ainda caso se opte pela derivação numérica direta ou não se gere nem a matriz de transição de estado nem a sua inversa ao se fazer uso do método de geração de soluções unitárias, necessita-se gerar um sistema suplementar de "neqn" equações para cada parâmetro a otimizar. Tais sistemas, porém, não precisam ser gerados em todo intervalo $[t_0, t_f]$. Recorda-se aqui que o método da geração de soluções unitárias exige que a integração destes sistemas de equações diferenciais deve ser feita simultaneamente com a geração da trajetória de referência. Usando-se o método da geração de soluções unitárias e gerando-se a inversa da matriz de transição do estado (Apêndice A), pode-se, em geral, reduzir bastante o tempo de integração, porém, passa-se a ter que efetuar a inversão de uma matriz de dimensão "neqn x neqn". Independentemente dos métodos escolhidos para efetuar as integrações numéricas, para gerar as derivadas parciais necessárias e para obter um

incremento de busca satisfatório, tem-se, no caso do problema Terra-Marte, que a integração numérica consome mais de 80% do tempo de computação por iteração, quando se utiliza menos de doze parâmetros.

A diferença de custo computacional, por iteração, entre os diversos métodos subótimos se deve apenas ao próprio processo de cálculo do incremento de busca e, em geral, só se torna relevante quando o número de parâmetros é maior que quatro vezes o número de componentes do vetor de estado do sistema dinâmico. Com o objetivo de fornecer uma idéia mais detalhada deste assunto, colocaram-se, no Apêndice D, fórmulas que permitem avaliar o custo computacional da obtenção do incremento de busca no vetor de parâmetros, tanto no caso do procedimento proposto, como nos casos dos parâmetros de Willianson (1971), Ceballos (1979), e também no caso das variantes do método da projeção do gradiente.

No caso dos testes apresentados no Capítulo 5, devido ao fato de se ter empregado um número baixo de parâmetros, pode-se comparar o custo computacional do emprego de diversos métodos subótimos diretamente a partir do número de iterações requeridas para se atingir a convergência. Tal se deve ao fato de o custo de cálculo do incremento de busca ser desprezível face ao custo da integração numérica.

Pode-se verificar, a partir dos testes apresentados no Capítulo 5, que o procedimento proposto apresentou uma boa velocidade de convergência se comparado à dos demais procedimentos subótimos por restrição no controle de primeira ordem, a exceção dos casos onde o controle era aproximado por trechos de valores constantes. Levando-se em conta que frequentemetne, na realização do controle, é preferível que os resultados sejam dados na forma de uma função, e que ao interpolar os resultados obtidos por métodos ótimos obtêm-se funções de controle, cujo desempenho difere significativamente da solução original; tais resultados são bastante expressivos.

Da mesma forma que no caso dos métodos subótimos, o tempo de integração é responsável pela maior parte do tempo gasto por iterações pelos métodos ótimos. Observando-se os procedimentos clássicos, verifica-se que o método do gradiente clássico requer a integração de um sistema de $2neqn^2 + 1$ equações diferenciais, onde neqn é a dimensão do vetor de estado do sistema dinâmico. O método de perturbação das equações adjuntas e o método da função de perturbação requerem a integração de um sistema de $2neqn^2 + 2$ equações para obter a matriz de transição de estado de um novo vetor formado pela justaposição do vetor de multiplicadores de Lagrange relacionados com os vínculos dinâmicos, além da geração de dois outros sistemas de neqn equações (Tapley and Lewallen, 1967).

Tapley e Lewallen (1967) afirmam que o método ótimo mais eficiente é o das funções de perturbação que, além de requerer o menor tempo por iteração dentre os métodos ótimos, é o método de convergência mais rápida. Tal método, contudo, bem como os demais métodos de perturbação, tem o inconveniente de só convergir quando os valores de partida para os multiplicadores de Lagrange, $\lambda(t)$, nos instantes terminais estão próximos aos valores corretos destes multiplicadores. Ainda segundo tais autores, atribuindo-se valor unitário para o tempo gasto por iteração pelo método das funções de perturbação na solução do problema Terra-Marte, tem-se que o método do gradiente clássico, na sua versão mais eficiente consome, para resolver tal problema, 1,5 unidades de tempo por iteração, e que o método de perturbação das equações adjuntas consome 1,15 unidades de tempo na solução deste mesmo problema. Tal discrepência se deve principalmente ao fato de estes métodos gerarem equações diferenciais de diferente condicionamento numérico.

De 1967 para cá houve uma grande evolução dos métodos que se baseiam no Método do Gradiente Conjugado, de forma que esses passaram a ser quase tão eficientes quanto o Método das funções de Perturbação, com a vantagem de não exigirem uma boa solução de partida (Miele, 1975; Miele et alii, 1979; Warren et alii, 1967; Fletcher and Reeves, 1964).

De acordo com Tapley e Lewallen (1967) e a própria experiência do autor, para obter soluções aceitáveis pelo Método das Funções de Perturbação para o problema Terra-Marte, necessita-se, conforme os valores de partida para os multiplicadores $\lambda(t)$, de 12 a 20 iterações, enquanto para resolver tal problema com o método do gradiente clássico necessita-se, conforme a versão escolhida e os valores de partida para o controle e para os parâmetros de ajuste do algoritmo, de 40 a 120 iterações. Não se têm dados sobre o desempenho do Método do Gradiente Conjugado na solução deste problema.

Considerando-se que o Método das funções de Perturbação requer um tempo de cálculo por iteração superior ao tempo gasto por um método subótimo por restrição no controle de primeira ordem, onde a função $u(a,t)$ tem nove parâmetros a otimizar, no caso do problema Terra-Marte, tem-se que o tempo computacional requerido para solucionar tal problema com o procedimento proposto é comparável ao deste método ótimo. Tal método ótimo exige uma boa solução de partida e um gasto de memória bem superior ao do procedimento proposto.

Considerando-se a facilidade de implementação do procedimento proposto se comparada a dos métodos ótimos, a precisão dos resultados obtidos, os requisitos de memória de computador exigida, a velocidade de convergência, e comparando-se estas três últimas características com aspectos dos demais métodos subótimos e ótimos, pode-se verificar que o procedimento proposto é uma boa alternativa em termos de engenharia para a solução de problemas de otimização dinâmica.

Deve-se ressaltar que a escolha do método de geração das derivadas parciais requeridas só afeta significativamente a fase final da convergência. Ressalta-se também a correspondência entre a matriz de projeção (Rosen, 1960) e a matriz de covariância do erro na estimativa final, para casos onde a matriz de covariância na estimativa inicial é proporcional à matriz identidade, e o fato de que aparentemente a introdução dos ruídos de medida diminuiram a incidência de problemas numéricos; contudo os dados disponíveis não permitem concluir que isto

realmente ocorra sempre. Ressl̃ta-se, por último, que se constatou experimentalmente o fato de que o uso de covariâncias maiores para o erro na estimativa inicial daqueles parâmetros que pouco afetam a função objetivo, permite acelerar frequentemtne a velocidade de convergência.

É necessário um estudo teórico mais aprofundado da constatação acima, e do efeito do ruído de medidas, da convergência do método em si, bem como uma comparação teórica entre o procedimento proposto e o Método da Projeção do Gradiente. Deve-se também efetuar testes numéricos mais exaustivos para confirmar a eficácia do método.

REFERÊNCIAS BIBLIOGRÁFICAS

BRYSON, A.E.; HO, Y.C. *Applied optimal control.* New York, John Wiley, 1975.

CEBALLOS, D.C. *Aproximações subótimas para o controle em problemas dinâmicos de otimização.* Tese de Mestrado em Ciência Espacial/Mecânica Orbital. São José dos Campos, INPE, 1979. (INPE-1676-TDL/019).

CITRONS, S.J. *Elements of optimal control.* New York, Holt, 1969.

DAVDSON, W.C. Variance algorithm for minimization. *The Computer Journal*, 10(4):406-410, 1968.

FLETCHER, R.; REEVES, M.C. Function minimization by conjugate gradiente. *The Computer Journal*, 7(3):149-154, 1963

FLETCHER, R.; POWELL, M.J.D. A rapid convergence method for minimization. *The Computer Journal*, 6(2):163-168, 1963.

FLEURY, A.T. *Controle ótimo de sistemas dinâmicos: métodos numéricos.* Notas de Aula. INPE. São José dos Campos, 1980, manuscrito.

HULL, D.G.; EDGEMAN, L.J. Suboptimal control using a second-order parameter optimization. *Journal of optimization and aplications*, 17 (5/6):481-491, Dec. 1975.

HESTENES, M.R.; STIEFEL, E. Method of conjugate gradients for solving linear sustems. *Journal of Research of the National Bureau of Standards*, 49(6):409-436, 1962.

JAZWINSKI, A.H. *Stochastic process and filtering theory.* New York, Academic Press, 1969. (Mathematics in Science and Engineering, vol. 64).

LIEBELT, P.B. *An introducion to optimal estimation.* Menlo Park, Adinson-Wesley, 1967.

LUENBERGER, D.G. *Optimization by space vector methods.* New York, John Wiley, 1969.

MIELE, A. Recent advances in gradiente algorithms for optimal control problems. *Journal of Optimization Theory and Aplications*, 17(5/6): 361-430, Dec. 1975.

MIELE, A.; CLOUTIER, R.J.; MOHANTY, B.P.; WU, A.K. Sequential conjugate gradient-restoration algorithm for optimal control problems with nondifferencial constraints. *International Journal of Control*, 29 (2):189-234, March 1979.

POWELL, M.J.D. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal*, 7(3):155-162, 1964.

——— An iterative method for finding stationary values of function of several variables. *The Computer Journal*, 5(2):147, 1962.

RIOS NETO, A. *Estimação linear ótima aplicada à geração de soluções numéricas subótimas em problemas de otimização de sistemas dinâmicos*. In: CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA, 5, Campinas, 1979. Anais. Campinas, Universidade Estadual de Campinas, 1979, V.D., p. 240-246.

——— *Estimação linear ótima aplicada à geração de soluções numéricas subótimas em problemas de controle de sistemas dinâmicos*. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, 3, Rio de Janeiro, 1980. Anais. Rio de Janeiro, Sociedade Brasileira de Automática, Setembro, 1980, p. 123-126.

RIOS NETO, A.; CEBALLOS, D.C. *Aproximation by polinominal arcs to generate suboptimal numerical solutions in control problems*. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, 5, Campinas, 1979. Anais. Campinas, Universidade Estadual de Campinas, 1979, V.C., p. 34-43.

——— Linear programming and suboptimal of dynamic systems control problems. In: EUROPEAN SPACE AGENCY (ESA). *Spacecraft Flight Dynamics;* Proceedings of a International Symposium held in Darmstadt, Germany, 12-22, May 1981. Paris, 1981, p. 239-243. (ESA-SP-160).

ROSEN, J.B. The gradient projection method for nonlinear programing. Part 1. Linear constraints. *Journal of Society of Industries Applied Mathematics*, 8(1):181-217, Mar. 1960.

——— The gradient projection method for nonlinear programing. Part II. Nonlinear constraints. *Journal of Society of Industries Applied Mathematics*, 9(4):514-532, Dec. 1961.

TAPLEY, B.D.; LEWALLEN, J.W. Comparasion of several numerical optimization methods. *Journal of Optimization and Applications*, 1(1):1-32, 1967.

WARREN, A.D.; LADSON, L.S.; MITTEL, S.K. The conjugate gradient method for optimal control problems. *IEE Transaction on Automatic and Control*, 12(2):132-138, April 1967.

WILLIANSON, W.E. Use of polinomial aproximations to calculate suboptimal controls. *AIAA Journal*, 9(11):2271-2273, Nov. 1971.

APÊNDICE A

OBTENÇÃO DAS DERIVADAS PARCIAIS DO ESTADO EM RELAÇÃO AOS PARÂMETROS PELO MÉTODO DE GERAÇÃO DE SOLUÇÕES UNITÁRIAS

Considere-se um sistema de equações diferenciais

$$\dot{x} = f(x,u,t) \tag{A.1}$$

com condição inicial

$$x(t_0) = x_0$$

onde u é um vetor de n componentes dado por

$$u = u(a,t)$$

Com o intuito de obter as derivadas parciais do vetor de estado x(t) em relação ao vetor de parâmetros a, no instante t, considere-se a equação diferencial variacional associada à Equação A.1 dada por

$$\delta\dot{x} = A(t)\delta x \quad B(t)\delta u$$

onde

$$A(t) = \frac{\partial f(x,u,t)}{\partial x}$$
$$B(t) = \frac{\partial f(x,u,t)}{\partial u} \tag{A.2}$$

Ao sistema autônomo dado por

$$\delta\dot{x}' = A(t)\delta x' \tag{A.3}$$

está associada uma matriz de transição de estado $S(t,\tau)$ que relaciona alterações no valor de $\Delta x'(\tau)$ com variações no valor de $\Delta x'(t)$, ou seja

$$\delta x'(t) = S(t,\tau)\delta x'(\tau) \tag{A.4}$$

Pode-se obter a matriz de transição de estado $S(t,\tau)$ mediante a integração do sistema de equações diferenciais

$$\dot{S}(t,\tau) = A(t)\ S(t,\tau) \tag{A.5}$$

sujeito à condição inicial

$$S(\tau,\tau) = I \tag{A.6}$$

onde I é a matriz identidade.

Pode-se obter diretamente a inversa da matriz $S(t,\tau)$ mediante a integração do sistema de equações diferenciais

$$\dot{S}^{-1}(t,\tau) = -S^{-1}(t,\tau)\ A(t) \tag{A.7}$$

sujeito à condição inicial

$$S(\tau,\tau) = I \tag{A.8}$$

onde I é a matriz identidade.

Supondo-se que $u(a,t)$ seja uma função contínua no intervalo $[t_i, t_k]$, que se conheça a variação do vetor de estado no instante t, pode-se escrever a variação do vetor de estado num instante t pertencente ao intervalo $[t_i, t_k]$, $\Delta x(t)$, como:

$$\delta x(t) = S(t,t_i)\Delta x(t_i) + \int_{t_i}^{t} S(t,s)\ B(s)\ \Delta u(s)\ ds \tag{A.9}$$

Tal expressão pode ser reescrita como:

$$\delta x(t\ ) = S(t,\ t_i)\,\delta x(t_i) + \int_{t_i}^{t} S(t,s)\ B(s)\ \frac{\partial u(s)}{\partial a}\ \delta a\ ds =$$

$$= S(t,t_i)\,\delta x(t_i) + \sum_{j=1}^{np}\left[\int_{t_i}^{t} S(t,s)\ B(s)\ \frac{\partial u(s)}{\partial a}\ ds\ \delta a_j\right] \qquad (A.10)$$

onde np é o número de parâmetros.

Considere-se agora os $P_j(t)$ definidos por

$$P_j(t) = \int_{t_0}^{t} S(t,s)\ B(s)\ \frac{\partial u(s)}{\partial a}\ ds \qquad (A.11)$$

Derivando-se $P_j(t)$ no tempo obtém-se

$$\dot{P}_j(t) = B(t)\ \frac{\partial u(t)}{\partial a_j} + \int_{t_i}^{t} \frac{\partial}{\partial t}\left[S(t,t_i)\right]\ S(t_i,s)\ B(s)\ \frac{\partial u(s)}{\partial a_j}\ ds =$$

$$= B(t)\ \frac{\partial u(t)}{\partial a_j} + \int_{t_i}^{t} A(t)\ S(t,t_i)\ S(t_i,s)\ B(s)\ \frac{\partial u(s)}{\partial a_j}\ ds =$$

$$= B(t)\ \frac{\partial u(t)}{\partial a_j} + A(t)\ P_j(t) \qquad (A.12)$$

Considerando-se agora um instante t_q onde ocorre uma descontinuidade no controle, tem-se a partir da imposição de continuidade do vetor de estado que:

$$dx(t_q) = \dot{x}^{(-)}(t_q)\ dt_q + \Delta x^{(-)}(t\ ) = \dot{x}^{(+)}(t_q)\ dt_q\ \Delta x^{(+)}(t_q) \qquad (A.13)$$

Assim pode-se escrever que:

$$\delta x^{(+)}(t_q) = \delta x^{(-)}(t_q) + [\dot{x}^{(-)}(t_q) - \dot{x}^{(+)}(t_q)]\, dt_q \qquad (A.14)$$

e escrever x(t), caso u(a,t) seja contínua por partes no intervalo $[t_0, t]$ como:

$$\delta x(t) = S(t,t_0)\delta x(t_0) + \sum_{i=1}^{np} P_j(t)\delta a_j + \sum_{i=1}^{nc} S(t,t_{qi})[\dot{x}^{(-)}(t_{qi}) - \dot{x}^{(+)}(t_{qi})] \qquad (A.15)$$

onde

nc é o número de instantes onde ocorrem descontinuidades;
t_{qi} são os instantes onde correm descontinuidades.

Para abreviar a notação, definem-se

$$Y(t) = S(t,t_0)\delta x(t_0)$$
$$D_i(t) = S(t,t_{qi})\,[\dot{x}^{(-)}(t_{qi}) - \dot{x}^{(+)}(t_{qi})] \qquad (A.16)$$

No caso de se aproximar cada componente do vetor de controle por várias funções contínuas do tempo e dos parâmetros, cada qual válida num intervalo $[t_{qi-1}, t_{qi}]$, onde t_{q0} é definido por t_0 e os tempos de quina não-fixos são funções do vetor de parâmetros a ser otimizado, pela própria definição de derivada, pode-se escrever as derivadas do vetor de estado x(t) em relação aos parâmetros como:

$$\frac{\partial x(t)}{\partial a_j} = P_j(t) + \sum D_i(t)\,\frac{\partial t_{qi}}{\partial a_j} + S(t,\, t_0)\,\frac{\partial x(t_0)}{\partial a_j} \qquad (A.17)$$

Deve-se ter em mente que nos trechos onde u(a,t) não depende de um determinado parâmetro a_j, ter-se-á:

$$\frac{\partial u(t)}{\partial a_j} \equiv 0 \qquad (A.18)$$

de tal forma que a Equação A.12 se reduz a

$$\dot{P}_j(t) = A(t)\, P_j(t) \qquad (A.19)$$

quando o controle u(a,t) independe do parâmetro a_j, o que equivale a dizer que conhecido $P_j(t_a)$ e dado que no intervalo $[t_a,t_b]$ a derivada $\partial u(t)/\partial a_j$ satisfaz à Equação A.18, pode-se obter $P_j(t_b)$ por meio de

$$P_j(t_b) = S(t_b,t_a)\, P_j(t_a) \qquad (A.20)$$

Desta forma, caso se propague a matriz de transição de estado, ou a sua inversa, de forma que $S(t_b,t_a)$ possa ser conhecida, pode-se dispensar a propagação de $P_j(t)$ entre t_a e t_b através do uso da Equação A.12, no caso coincidente com a Equação A.19.

Caso, por particularidades do problema, não seja vantajoso propagar a matriz de transição de estado, poder-se-ia obter $D_i(t)$ pela integração de

$$\dot{D}_i(t) = A(t)\, D_i(t) \qquad (A.21)$$

sujeito à condição inicial

$$D_i(t_{qi}) = [\dot{x}^{(-)}(t_{qi}) - \dot{x}^{(+)}(t_{qi})] \qquad (A.22)$$

De forma análoga, poder-se-iam obter apenas as n_i colunas $s_\ell(t,t_0)$ necessárias de $S(t,t_0)$, caso não fosse conveniente propagar toda a matriz de transição, mediante a integração de n_i sistemas de equações diferenciais da forma

$$\dot{s}_\ell(t,t_0) = A(t)\, s_\ell(t,t_0) \qquad (A.23)$$

sujeitos à condições iniciais

$$s_\ell(t_0,t_0) = [\Delta_{1\ell}, \Delta_{2\ell}, \Delta_{3\ell}, \ldots, \Delta_{n\ell}]$$

onde n é a dimensão do vetor de estado x(t), e Δ_{ij}, o delta de Kronecker.

A partir da Equação A.17 pode-se escrever:

$$\frac{\partial x(t)}{\partial a_j} = S(t,t_{qi})\, P_j(t_{qi}) + S(t,t_{qi})\, [\dot{x}^{(-)}(t_{qi}) - \dot{x}^{(+)}(t_{qi})] \frac{\partial t_{qi}}{\partial a_j}$$

$$+ S(t,t_{qi})\, S(t_{qi}, t_0) \frac{\partial x(t_0)}{\partial a_j} + P'_j(t) \qquad (A.24)$$

onde

$$P'_j(t) = \int_{t_{qi}}^{t} S(t,s)\, B(s) \frac{\partial u(s)}{\partial a_j}\, ds$$

que pode ser obtido pela propagação entre t_{qi} e t de

$$\dot{P}'_j(t) = B(t) \frac{\partial u(t)}{\partial a_j} + A(t)\, P'_j(t)$$

sujeita à condição inicial

$$P'_j(t_{qi}) = 0$$

Desta forma, comparando-se a Equação A.24 com a solução da Equação Diferencial A.12, que é dada por

$$P_j(t) = S(t,t_0)\, P_j(t_0) + \int_{t_0}^{t} S(t,s)\, B(s) \frac{\partial u(s)}{\partial a_j}\, ds$$

pode-se verificar que se necessita integrar um único sistema de n equações diferenciais por parâmetro a_j, de forma análoga à Equação A.12, obtendo-se $\partial x(t)/\partial a$ mediante a integração de

$$\dot{\bar{P}}_j(t) = A(t)\,\bar{P}_j(t) = B(t)\,\frac{\partial u(t)}{\partial a_j} \tag{A.25}$$

sujeito à condição inicial

$$\bar{P}_j(t_0) = \frac{\partial x(t_0)}{\partial a_j} \tag{A.26}$$

até o primeiro tempo de quina t_{q_1}, instante no qual ocorre uma descontinuidade na derivada $\partial x(t)/\partial a_j$, e onde se deve adicionar à $\bar{P}_j(t_{q_1})$ a variável $\Delta\bar{P}_j(t_{q_1})$ dada por

$$\Delta\bar{P}_j(t_{q_1}) = [\dot{x}^{(-)}(t_{q_1}) - \dot{x}^{(+)}(t_{q_1})]\,\frac{\partial t_{q_1}}{\partial a_j} \tag{A.27}$$

voltando-se a seguir a utilizar a Equação A.25 para propagar $\bar{P}_j(t)$ até o próximo tempo de quina, onde se deve adicionar uma nova parcela, $\Delta\bar{P}_j(t_{q_2})$ dada por

$$\bar{P}_j(t_{q_2}) = [\dot{x}^{(-)}(t_{q_2}) - \dot{x}^{(+)}(t_{q_2})]\,\frac{\partial t_{q_2}}{\partial a_j} \tag{A.28}$$

para voltar a utilizar a Equação A.25 para a propagação de $\bar{P}_j(t)$, e assim por diante. Resumindo-se, adiciona-se, nos instante t_{qi} posteriores a t_{q_0}, à $\bar{P}_j(t)$ parcelas dadas por

$$\Delta\bar{P}_j(t_{qi}) = [\dot{x}^{(-)}(t_{qi}) - \dot{x}^{(+)}(t_{qi})]\,\frac{\partial t_{qi}}{\partial a_j} \tag{A.29}$$

e usa-se a Equação A.25 para propagar $\bar{P}_j(t)$ entre dois instantes t_{qi} sucessivos.

A partir daí, pode-se concluir que a propagação da matriz de transição de estado ou sua inversa só permite a redução do es

forço computacional necessário para a geração das derivadas parciais do estado em relação aos parâmetros, quando o número de parâmetros a_j para os quais

$$\frac{\partial u(t)}{\partial a_j} \equiv 0$$

num intervalo $[t_a,t_b]$ e também $\bar{P}_j(t_a) \neq 0$, for superior à dimensão n do vetor de estado do sistema, $x(t)$.

Caso se imponha que se

$$\bar{P}_j(t) \neq 0 \quad e \quad \frac{\partial u(t)}{\partial a_j} = 0, \forall\ t \in [t_a,t_b]$$

tenha-se

$$\frac{\partial u(t)}{\partial a_j} = 0, \forall\ t > t_b$$

então, para qualquer instante t* posterior ao instante t_c, no qual se torna vantajoso propagar a matriz de transição de estado ou a sua inversa, vale

$$S(t_f,t^*) = S(t_f,t_c)\ S^{-1}(t^*,t_c) \qquad (A.30)$$

de forma que a geração da inversa da matriz de transição de estado permite a simplificação do algoritmo e exige uam única inversão de matrizes para a obtenção de $S(t_f,t_c)$ a partir de $S^{-1}(t_f,t_c)$.

Pode.se empregar este método praticamente sem alteração alguma em problemas onde existem vínculos de desigualdade na variável de controle da forma

$$C(x,u,t) \leq 0 \qquad (A.31)$$

onde $C(x,u,t)$ é um vetor com um número de componentes n_{vc} menor ou igual a n_c.

Neste caso deve-se calcular normalmente o controle $u(a,t)$ através da função de interpolação previamente escolhida e checar se a Equação A.31 é satisfeita, mas, caso contrário, deixar de calcular algumas das componentes do vetor de controle $\dot{o}(t)$ a partir das componentes $u_i(a,t)$ da função $u(a,t)$ proposta para o cálculo do controle e obter tais componentes a partir da solução da equação

$$C(x,u,t) = 0 \tag{A.32}$$

Devido a isto, caso n_c seja igual a n_{vc}, ter-se-á

$$\frac{\partial u}{\partial a_j} \equiv 0, \forall j \text{ e } \forall i \tag{A.33}$$

e caso $n_{vc} \leq n_c$, deve-se obter parte das derivadas $\partial u_i/\partial a_j$ diretamente e parte delas através da regra da cadeia, ou seja, supondo-se que apenas uma componente $u_k(t)$ não seja obtida a partir de $u_k(a,t)$ e sim pela solução da Equação A.32, ter-se-á:

$$\frac{\partial u_k(t)}{\partial a_j} = \sum_{\substack{i=1 \\ i\neq k}}^{nc} \frac{\partial u_k}{\partial u_i} \frac{\partial u_i}{\partial a_j} \tag{A.34}$$

Uma vez tomados todos estes cuidados podem-se usar, para a determinação das derivadas do estado em relação aos parâmetros, as expressões desenvolvidas para o caso de problemas sem vínculos vistas até agora.

APÊNDICE B

COMPLEMENTAÇÃO DA DESCRIÇÃO DO ALGORITMO NUMÉRICO

Dentro das análises preliminar e final dos resultados apresentou-se uma série de verificações que visavam garantir a convergência do algoritmo numérico. Juntamente com tais verificações apresentou-se uma série de critérios de ajuste dos pesos do algoritmo, de caráter puramente corretivo. Tais critérios não permitem identificar que a eventual causa de uma convergência lenta é uma má escolha dos pesos do algoritmo, a não ser que tais pesos levem à rejeição de uma estimativa do incremento de busca. Esses critérios também deixam em aberto a questão de como proceder nas iterações subsequentes àquelas onde se satisfez a todos os testes linearizados de convergência mas não àqueles de caráter não-linear. Neste apêndice, discutir-se-á sobre estes dois aspectos, apresentando-se conjuntametne uma série de critérios de ajuste adicional dos pesos do algoritmo e recursos extraordinários de aceleração da convergência. A seguir far-se-á um resumo do algoritmo sob a forma de uma sequência de passos. Dentre os recursos extraordinários para a aceleração da convergência, destaca-se o uso de medidas adicionais z_i''' definidas na Equação 3.14.

Diversos autores da área de programação não-linear empregam em métodos de otimização de parâmetros de primeira ordem critérios de controle de passo, cujo objetivo é manter em níveis satisfatórios os efeitos da não-linearidade. Tais critérios ou exigem o conhecimento de derivadas parciais de segunda ordem, o que implica um alto custo computacional, quando se tem que achar tais derivadas via integração numérica; ou então apenas apontam a ordem de grandeza aproximada do incremento de busca. Em vista disto, decidiu.se controlar os efeitos da não-linearidade através de um processo orientado de tentativa e erro, onde se decresce o valor do peso p em caso de quebra de linearidade e aumenta-se tal valor caso os efeitos da não-linearidade se mostrem insignificantes. Ocupar-se-á a seguir da descrição deste processo.

Na fase de captura dos vínculos, deseja-se obter o mais rapidametne possível um vetor de parâmetros ao qual corresponda uma trajetória x(t) tal que os vínculos de contorno estejam contidos na faixa do ruído. Assim, convém aumentar o peso p, quando se constatar, após se obter por integração a nova trajetória nominal, que a diferença entre o valor previsto para o erro nos vínculos de contorno correspondentes a esta nova trajetória, e o valor real destes erros for pequena face à variação prevista para tais erros, ou seja, se

$$|[M^T(\bar{x}(t_f,\bar{a}):I^T(\bar{a})] - [M^T(\bar{\bar{x}}(t_f),\bar{\bar{a}}):I(\bar{\bar{a}})] - [M_{\bar{\bar{a}}_0}\Delta a)^T:(I_{\bar{\bar{a}}_0}\Delta a)^T]| \leq$$

$$|(M_{\bar{\bar{a}}_0}\Delta a)^T:(I_{\bar{\bar{a}}_0}\Delta a)^T| \qquad (B.1)$$

onde

$\bar{\bar{a}}$	- é o valor nominal anterior do vetor de parâmetros a otimizar;
$\bar{a}$	- é o valor nominal atual do vetor de parâmetros a otimizar;
Δa	- é a diferença $\bar{a} - \bar{\bar{a}}$;
$\bar{x}$	- é a trajetória correspondente a $u(\bar{a},t)$;
$\bar{\bar{x}}$	- é a trajetória correspondente a $u(\bar{\bar{a}},t)$;
I(a) e M(x(t),a)	- são os erros nos vínculos de contorno no instante inicial e final, respectivamente, correspondentes ao vetor de parâmetros a e a trajetória x(t);
$M_{\bar{\bar{a}}_0}$ e $I_{\bar{\bar{a}}_0}$	- são as matrizes de derivadas do erro nos vínculos de contorno em relação aos parâmetros em torno da trajetória x(t) e do valor nominal anterior dos parâmetros.

e se ainda nenhuma das duas pultimas iterações tiver sido rejeitada por quebra da linearidade e também se a matriz ρ recém-calculada tiver algum elemento ρ_{ii} menor do que um. Constatado que se convém aumentar p, deve-se refazer o cálculo das matrizes diagonais P e ρ e da estimativa a priori $\Delta\bar{a}$.

Recorda-se aqui que toda vez que ocorre quebra de linearidade deve-se reduzir p, e que nos casos onde não se repetiu o cálculo do incremento de busca (Capítulo 5), deve-se também multiplicar o peso p pelo escalar α_ℓ, pelo qual o incremento de busca é também multiplicado.

Na fase de busca do mínimo, é conveniente aumentar p caso:

a) não tenha ocorrido quebra de linearidade nas duas últimas iterações;

b) os desvios causados pela não-linearidade, nas componentes do erro nos vínculos de contorno, sejam menores que as correções previstas em tais erros;

c) os desvios sejam pequenos face a diferença entre o máximo valor tolerado para os desvios e o valor atual destes;

d) o desvio devido a não-linearidade no valor da função objetivo seja pequeno face a variação prevista desta função.

Tais exigências podem ser escritas como

$$|G_{\bar{a}_0}\Delta a + G(\bar{x}(t_f), \bar{a}) - G(\bar{x}(t_f, \bar{a})| < |G_{\bar{a}_0}\Delta a| \cdot e_t$$

$$|I_i(\bar{a}) - I_{pi}(\bar{a})| < e_t \min\{|I_{pi}(\bar{a}) - I(\bar{a})|, f_{ea}.\bar{e}_a - |I_i(\bar{a})|\} \quad \text{(B.2)}$$

$$\text{p/} i=1,\ell$$

$$|M_i(\bar{x}(t_f),\bar{a}) - M_{pi}(\bar{x}(t_f),\bar{a})| < e_t \min\{|M_{pi}(\bar{x}(t_f),\bar{a} -$$

$$- M(\bar{x}(t_f), \bar{a})|, f_{ea}.\bar{e}_a - |M(\bar{x}(t_f),\bar{a}|\} \qquad \text{p/} i=1,j$$

onde

$G_{\bar{a}_0}$ - é o gradiente da função objetivo em torno da trajetória nominal anterior e do valor anterior, $\bar{a}$, de a;

$G(x(t),a)$ - é o valor da função objetivo correspondente a um dado valor a do vetor de parâmetros a otimizar e a trajetória x(t);

$M_i(x(t),a)$ e $I_i(a)$ - são respectivamente a i-ésima componente de $M(x(t),a)$ e $I(a)$;

e_t - é um parâmetro escolhido convenientemente, pertencente ao intervalo (0,1);

f_{ea} e $\bar{e}_a$ - são parâmetros que especificam o máximo valor tolerado para o erro nos vínculos;

$M_{pi}(x(t),a)$ e $I_{pi}(a)$ - são respectivamente a i-ésima componente de $M(\bar{x}(t),\bar{a})$ e $I(\bar{a})$ definidos por

$$M_p(x(t_f),a) = M(\bar{x}(t_f),\bar{a}) + M_{\bar{a}_0}\Delta a$$

$$I_p(\bar{a}) = I(\bar{a}) + I_{\bar{a}_0}\Delta a \qquad (B.3)$$

Outro ponto importante é que, em geral, quando na fase de busca do mínimo se conseguem estimativas lineares do novo erro nos vínculos de contorno cujo módulo é menor que o módulo destes erros, pode-se acelerar a convergência diminuindo-se $\bar{\beta}$ e $\bar{\beta}_p$. Quando, por sua vez, todas as componentes da estimativa linear do novo erro nos vínculos de contorno têm valores em módulo inferiores àqueles do erro nos vínculos propriamente dito, já é praticamente certo que a redução de $\bar{\beta}$ e $\bar{\beta}_p$ levará a um aumento da velocidade de convergência. Para verificar tais condições deve-se ter em mente que a estimativa linear do novo erro nos vínculos é dada por:

$$E_p(x'(t_f),\bar{a} + \Delta\bar{a}) = [I_p(\bar{a} + \Delta a) : M_p(x'(t_f),a + \Delta\bar{a})] \qquad (B.4)$$

onde

$$I_p(\bar{a} + \Delta\hat{a}) = I(\bar{a}) + I_{\bar{a}}\Delta\hat{a}$$

$$M_p(x'(t_f),\bar{a} + \Delta\hat{a}) = M_{\bar{a}}\Delta a + M(\bar{x}(t_f),\bar{a})$$

e

$M_{\bar{a}}$ e $I_{\bar{a}}$ - são as matrizes de derivadas parciais dos erros nos vínculos de contorno no instante final e inicial, respectivamente, em relação aos parâmetros em torno da atual trajetória nominal e do valor nominal atual dos parâmetros;

$\Delta\hat{a}$ - é o incremento de busca no vetor de parâmetros;

$x'(t)$ - é a trajetória correspondente a $u(\bar{a} + \Delta\hat{a},t)$.

Quando se tem concomitantemente a redução dos vínculos de controno e da função objetivo, têm-se três alternativas para melhorar a velocidade de convergência:

a) aceitar o incremento de busca obtido e partir para uma nova iteração com menores valores de $\bar{\beta}$ e $\bar{\beta}_p$ se ele não violar as condições abordadas na Sub-seção 4.2.5 e ainda se:

$$\begin{aligned} &|I_{pi}(a + \Delta\bar{a})| < |I_i(\bar{a})| \qquad p/i=1, \\ &|M_{pi}(x'(t_f),\bar{a} + \Delta\bar{a}| < |M(\bar{x}(t_f),\bar{a})| \qquad p/i=1, \end{aligned} \tag{B.5}$$

onde $I_{pi}(\bar{a} + \Delta\bar{a})$, $M(x'(t_f),\bar{a} + \Delta\bar{a})$, $I_i(\bar{a})$ e $M_i(\bar{x}(t_f),\bar{a})$ são respectivamente a i-ésima componente de $I(\bar{a} + \Delta\bar{a})$, $M(x'(t_f),\bar{a} + \Delta\bar{a})$, $I(\bar{a})$ e $M(\bar{x}(t_f),\bar{a})$;

b) armazenar a estimativa obtida e repetir o cálculo do incremento de busca com menores valores de $\bar{\beta}$ e $\bar{\beta}_p$, tomando-se como incremento de busca aquele armazenado se esta nova estimativa violar alguma das condições abordadas na Sub-seção 4.2.2, ou a nova estimativa em caso contrário. Pode-se repetir eventualmente o cálculo do incremento de busca até se obter uma estimativa

insatisfatória ou até que se supere um limite máximo preespecificado de repetições destes cálculos, tomando-se como incremento de busca o último valor satisfatório obtido;

c) armazenar a estimativa obtida e filtrar uma medida adicional z_i''', definida na Equação 3.14, calculada pela solução do sistema de inequações abaixo

$$z_i''' < 0$$

$$|z_i'''| \leq |G_{\bar{a}}\Delta a|[\alpha_3 + 1/2(sng(G_{\bar{a}}\Delta a) + 1] \qquad (B.6)$$

$$|z_i''' - G_{\bar{a}}\Delta a| \leq G_{\bar{a}} P G_{\bar{a}}^T$$

onde

sng(y) é uma função que valo 1 se $y > 0$, -1 se $y < 0$ e 0 se $y = 0$,

e cuja covariância também será adotada como $e_m^2/3$. Caso esta nova estimativa seja capaz de satisfazer às condições abordadas na Subseção 4.2.3, ela será adotada como o novo incremento de busca; caso ela não seja satisfatória, o incremento de busca será a estimativa armazenada. Convém, no caso da filtragem desta medida adicional dar resultados positivos, reduzir $\bar{\beta}$ e $\bar{\beta}_p$. Também pode-se eventualmente filtrar mais de uma medida acidional z_i''', analogamente ao caso anterior.

Para completar a descrição do algoritmo numérico utilizado e para esclarecer melhor os tópicos abordados no Capítulo 4 e neste apêndice, descreve-se a seguir o algoritmo como uma sequência de passos.

1º PASSO

Leia os dados e imprima-os, faça $\Delta\hat{a} = 0$, IVINC = 0 (não se satisfizeram os vínculos de contorno). IE = 0, LINEAR = 0 e calcule R por meio da Equação 4.2. Se todos os parâmetros influem no valor da função objetivo, deve-se fornecer valor de β_0 e β_{p0} iguais a 1.

2º PASSO

Integre a trajetória correspondente a $u(\bar{a} + \Delta\hat{a}, t)$ e gere as derivadas parciais correspondentes. Se IE = 0, vá para o 5º passo.

3º PASSO

Se IVINC = 0 e a Desigualdade 4.40 for satisfeita ou se IVINC = 1 e as Desigualdades 4.44 e 4.46 forem satisfeitas, vá para o 5º passo. Caso contrário, deve-se considerar que se convergiu, e ir para o 18º passo se a Desigualdade 4.26 for satisfeita e IVINC = 1, e considerar que se estimou mal aos erros de integração, indo para o 12º passo se a Desigualdade 4.26 for satisfeita e IVINC = 0 .

4º PASSO

Faça LINEAR = 0. Se $\bar{\rho}_{ii} > 2$, $\forall$ i/1 $\leq$ i $\leq \ell + j$ e IVINC = 0, diminua p e vá para o 6º passo. Caso contrário, multiplique p e $\Delta\hat{a}$ pelo valor de α_ℓ obtido pela solução da inequação correspondente à condição violada e vá para o 2º passo.

5º PASSO

Monte a matriz H definida na Equação 3.5, faça LINEAR = LINEAR+1, IE = 1, $\bar{a} = \bar{a} + \Delta\hat{a}$, REPETE = 0, CONT = 0, NREJ = 0, e CASO = 0. Imprima os vetores a, $x(t_0)$, $x(t_f)$ e o erro nos vínculos de contorno no instante final e inicial.

6º PASSO

Calcule a e_e e as matrizes P e ρ. Vá para o 8º passo se LINEAR $\geq$ 2 ou se CASO $\neq$ 0.

7º PASSO

Se IVINC = 0, a Desigualdade B.1 é satisfeita e ∃i ∈ |N / 1 ≤ i ≤ j+ +ℓ e $\rho_{ii} < 1$ ou se IVINC = 1 e a Desigualdade B.2 é satisfeita, aumente p.

8º PASSO

Faça $\Delta\bar{a} = -pG_{\bar{a}}^{T}$. Calcule o vetor de observações z por meio da Equação 4.6, aplicando a seguir o Filtro de Kalman, de preferência sequencialmente, para obter uma estimativa $\Delta\bar{a}$ do incremento de busca no vetor de parâmetros. Faça CONTOU = CONT. Vá para o 13º passo se IVINC=1.

9º PASSO

Se a Desigualdade 4.13 for satisfeita, faça IVINC = 1, e vá para o 13º passo.

10º PASSO

Se a Desigualdade 4.9 não for satisfeita, faça CASO = 1 e diminua p. Se a Desigualdade 4.10 não for satisfeita, multiplique p pelo valor de α_{ℓ} dado pela Desigualdade 4.11 e faça CASO = -1. Vá para o 5º passo se CASO = 0.

11º PASSO

Se a Desigualdade 4.26 for satisfeita, ou se a Desigualdade 4.12 for satisfeita e CASO = 1, faça CASO = -2. Se CASO = -1 vá para o 10º passo. Se ∃ i ∈ N / 1 ≤ i ≤ j + ℓ e $\rho_{ii} < 1$, aumente $\bar{\beta}$, $\bar{\beta}_p$ e também β_0 e β_{p0} se algum parâmetro não influir no valor da função objetivo. Vá para o 6º passo.

12º PASSO

Caso os módulos das componentes do erro nos vínculos de contorno sejam todos menores que Cinc.$\bar{e}_a$ e Cinc > 1, faça IVINC = 1, $\bar{e}_a$ = Cinc. $\bar{e}_a$, Cinc = 1, p = p* e volte para o 6º passo. Caso contrário, faça IE= 0, p = p* e vá para o 6º passo.

13º PASSO

Se a Desigualdade 4.14 for satisfeita, vá para o passo seguinte. Se REPETE = 1, vá para o 17º passo. Faça NREJ = NREJ + 1. Se as Desigualdades 4.14 a 4.18 forem satisfeitas ou se NREJ > NREJmax, verifique se a Desigualdade 3.62 é satisfeita, indo para o 17º passo em caso afirmativo. Se $\bar{\beta} \geq \beta_0$ e $\bar{\beta}_p \geq \beta_{p0}$, diminua p. Se $\bar{\beta} < \beta_0$, aumente $\bar{\beta}$, e se $\bar{\beta}_p < \beta_{p0}$, aumente $\bar{\beta}_p$. Faça CASO = 2 e vá para o 6º passo.

14º PASSO

Se a Desigualdade 4.19 for satisfeita, vá para o próximo passo. Vá para o 17º passo se REPETE = 1. Faça CASO = 3 e NREJ = NREJ + 1. Se as Desigualdades 4.15, 4.20 ou 4.26 forem satisfeitas vá para o 18º passo. Se a Desigualdade 4.21 for satisfeita ou NREJ > NREJmax, verifique se a Desigauldade 3.62 é satisfetia, indo para o 18º passo em caso afirmativo. Se não houve convergência, diminua $\bar{\beta}$ e $\bar{\beta}_p$ e vá para o 6º passo.

15º PASSO

Se as Desigualdades 4.23, 4.25 e 4.28 forem satisfeitas, vá para o próximo passo. Se a Desigualdade 4.26 for satisfeita e REPETE = 0, vá para o 18º passo. Multiplique $\Delta\bar{a}$ pelo escalar α_ℓ obitdo da inequação correspondente (4.23, 4.25 ou 4.28). Se REPETE = 0, multiplique também a p por α_ℓ. Volte ao início do passo.

16º PASSO

Se IOPT = 0, NREJ = 0 e a Desigualdade B.5 for satisfeita, diminua $\bar{\beta}$ e $\bar{\beta}_p$. Vá para o segundo passo se: IOPT = 0; LINEAR $\leq$ 2; ou REPETE = 0 e $E_p(x'(t),\bar{a} + \Delta\bar{a}) > E(\bar{a})$, onde $E(\bar{a})$ é definido na Equação 3.9. Se REPETE = 1 e $G_{\bar{a}}\Delta\bar{a} < G_{\bar{a}}\, a^*$, vá para o próximo passo. Se $E(x(t),\bar{a} + \Delta\bar{a}) < E(a)$ vá para o segundo passo. Faça REPETE = 1, CONT = CONT + 1. Se CONT > CONTmax, vá para o segundo passo. Faça $\Delta a^* = \Delta\bar{a}$, $\beta^* = \bar{\beta}$, $\beta_p^* = \bar{\beta}$. Se IOPT = 1, diminua $\bar{\beta}$ e $\bar{\beta}_p$ e vá para o 6º passo. Se IOPT = 2 faça $R_{\ell+j+1} = e_m^2/3$, $h_{\ell+j+1} = G_{\bar{a}}$ e calcule z_i''' por meio das Desigualdades B.6, aplicando, a seguir, o Filtro de Kalman para filtrar esta nova medida e indo para o 13º passo.

17º PASSO

Se IOPT = 2 e CONT ≥ 2 diminua $\bar{\beta}$ e $\bar{\beta}_p$. Faça $\Delta\hat{a} = \Delta a^*$, $\bar{\beta} = \beta^*$ e $\bar{\beta}_p = \beta^*_p$.

18º PASSO

Imprima o vetor de parâmetros a otimizar a, o vetor de estado nos instantes t_0 e t_f, a função objetivo, o gradiente da função objetivo, a projeção do gradiente da função objetivo no hiperplano tangente às restrições, o erro nos vínculos de contorno, o número de iterações efetuadas e pare.

Finalmente, o algoritmo aqui descrito exige que se forneçam como dados de entrada:

a) valores de partida para a, p, $\bar{\beta}$, $\bar{\beta}_p$, β_0, β_{p_0};

b) o erro numérico desprezível e_m, e_a, f_{ea};

c) Cint, IOPT, α_g, α_3, e_t;

d) MAXrej, NREJmax;

e) razões de decréscimo de $\bar{\beta}$ e $\bar{\beta}_p$, respectivamente chamadas $\alpha_{\bar{\beta}}$ e $\alpha_{\bar{\beta}p}$, que devem ser usadas em caso de se tentar acelerar a convergência;

f) a razão usada para aumentar p quando isto for vantajoso;

g) a razão usada para diminuir p, nos casos em que há rejeição da da estimativa e não se especifica tal razão de decréscimo no próprio algoritmo;

h) valores de referência para as razões de decréscimo de $\bar{\beta}$ e $\bar{\beta}_p$, em caso de rejeição de $\Delta\hat{a}$, por não levar à redução da função objetivo; estas razões devem ser o inverso das usadas para alterar $\bar{\beta}$ e $\bar{\beta}_p$ no 13º passo. Caso, ao se alterar $\bar{\beta}$ e $\bar{\beta}_p$, se pas

se a rejeitar a estimativa do incremento de busca que exija a alteração no sentido inverso de $\bar{\beta}$ e $\bar{\beta}_p$, sabe-se que há valores satisfatórios de $\bar{\beta}$ e $\bar{\beta}_p$ entre os valores atuais e os valores usados anteriormente. Deve-se usar, então, um esquema semelhante ao de bisecção para achar valores satisfatórios para estes parâmetros.

APÊNDICE C

PARTICULARIDADES DAS SOLUÇÕES COM TEMPOS INTERMEDIÁRIOS A OTIMIZAR

Tratar-se-á aqui de particularidades específicas daquelas funções u(a,t), onde: a j-ésima componente do controle é aproximada por vários arcos de polinômio, cada qual usado para a obtenção do controle no intervalo $t_{j,i}$, $t_{j,i+1}$; pelo menos um destes tempos de quina está sujeito à variação; caso ambos os tempos de quina estejam sujeitos à variação, nenhum vínculo de igualdade relaciona a variação destes tempos de quina entre si.

Em todas as iterações os tempos de quina devem estar ordenados, obedecendo, portanto, a

$$t_{j,i} \leq t_{j,i+1} \tag{C.1}$$

onde

$t_{j,0}$ corresponde ao tempo inicial, comum a todas as componentes;
$t_{j,f}$ corresponde ao tempo final, comum a todas as componentes.

Assim sendo, numa iteração típica, têm-se inicialmente valores iniciais ou nominais de dois tempos de quina sucessivos $\bar{t}_{j,i}$ e $\bar{t}_{j,i+1}$ tais que:

$$\bar{t}_{j,i} \leq \bar{t}_{j,i+1} \tag{C.2}$$

e deseja-se não só calcular um incremento de busca no vetor de parâmetros que provoque a caminhada na direção da solução do problema restrito de otimização de parâmetros, como também não provocar a violação do vínculo da Desigualdade C.1, ou seja, deve-se ter

$$\bar{t}_{j,i} + \Delta t_{j,i} \leq \bar{t}_{j,i} + \Delta t_{jii+1} \tag{C.3}$$

Frequentemente, tal desigualdade não é respeitada, não só porque eventualmente se tenham incrementos no vetor de parâmetros muito grandes, como devido ao fato de que uma má escolha da forma da função de interpolação ou do valor inicial dos parâmetros que caracterizam esta função, para um dado trecho, pode criar a posssibilidade de que a melhor maneira de reduzir o erro nos vínculos de contorno ou o valor da função objetivo seja mover os tempos de quina que delimitam o intervalo um em direção ao outro, diminuindo assim a influência negativa da má escolha dos parâmetros, ou da função de interpolação pela redução do intervalo onde eles atuam. Outra possibilidade é que a função de controle u(t) seja de tal forma que a derivada de uma de suas componentes em relação ao tempo seja demasiadamente grande, a ponto de exigir que dois tempos de quina sucessivos fiquem muito próximos, para que seja possível obter uma boa aproximação de u(t) com a função de interpolação u(a,t) proposta. Um bom exemplo, deste caso é o da aproximação do controle por 4 segmentos de reta com tempos intermediários a otimizar (Capítulo 5).

O uso do Filtro de Kalman para a determinação do incremento de busca no vetor de parâmetros permite um controle mais fácil deste problema, uma vez que se pode alterar a covariância de algumas componentes da estimativa a priori, para impedir que os tempos de quina cruzem, ou até mesmo filtrar medidas extraordinárias para remediar tal fato, caso este ocorra. Optou-se, após algumas verificações experimentais, pela adoção de um esquema baseado na filtragem de medidas extraordinárias pelo fato de este método não interferir com o andamento do algoritmo numérico quando não ocorre problema algum.

Este método consiste em, uma vez obtido um incremento de busca tal que para dois tempos de quina sucessivos $t_{j,i}$ e $t_{j,i+1}$ não se verifique a Desigualdade B.3, aplicar medidas

$$z^* = (1/\alpha_{t_3})[t_{j,i+1} - t_{j,i}] \qquad (C.4)$$

onde α_{t_3} é um parâmetro maior ou igual a um,

com covariância

$$R^* = z^2/12 \qquad (C.5)$$

onde

$$z^* = h^* \Delta a + R^*$$

e

$$h^* \triangleq \left[\left. \frac{\partial(t_{j,i+1} - t_{j,i})}{\partial a_1} \right|_{\substack{\bar{t}_{j,i} \\ t_{j,i+1}}} , \left. \frac{\partial(t_{j,i+1} - t_{j,i})}{\partial a_2} \right|_{\substack{\bar{t}_{j,i} \\ t_{j,i+1}}} , \ldots , \frac{\partial(t_{j,i+1} - t_{j,i})}{\partial a_{np}} \right]$$

Obtida uma estimativa que não viole à Equação C.3, iniciam-se as verificações da qualidade do incremento de busca obtido, descritas no Capítulo 4.

APÊNDICE D

FÓRMULAS PARA O CÁLCULO DO NÚMERO DE OPERAÇÕES ARITMÉTICAS REQUERIDAS PARA O CÁLCULO DO INCREMENTO DE BUSCA POR DIVERSOS MÉTODOS SUBÓTIMOS

Este apêndice tem por objetivo fornecer ao leitor alguns subsídeios para uma melhor comparação entre os métodos subótimos de primeira ordem de restrição no controle, citados neste trabalho, no tocante ao custo computacional.

Uma vez que o número de integrações a serem efetuadas por iteração independe do método escolhido, e sim apenas do número de parâmetros e da dimensão do vetor de estado do sistema, as diferenças de custo computacional de uma iteração típica entre os diversos métodos subótimos por restrição no controle de primeira ordem estão diretamente relacionadas com o número de operações aritméticas requeridas para determinar o incremento de busca propriamente dito, a partir do conhecimento de uma dada trajetória de referência e das derivadas parciais correspondentes.

Em geral, o custo computacional da determinação do incremento de busca no vetor de parâmetros é bastante inferior ao da geração da trajetória de referência e das derivadas correspondentes. No entanto, isto pode deixar de ser verdadeiro, quando se tem um número de vínculos muito grande ou então um número de parâmetros muito maior que o número de vínculos.

O número de operações requeridas para a determinação do incremento de busca dependerá, no caso do procedimento proposto, do modo de aplicação das equações do Filtro de Kalman, sequencial ou não sequencial, e da versão usada de tais equações. Neste apêndice abordar-se-ão aspectos do custo computacional de apenas três versões das equações do Filtro de Kalman, justamente aquelas apresentadas no Capítulo 2.

Pelo fato de o método de Ceballos (1979) determinar, numa iteração típica, o incremento de busca por meio de um outro processo iterativo, existe uma certa dificuldade de comparar até mesmo o custo computacional da determinação do incremento de busca deste método com o dos demais métodos.

Para efeito de preenchimento da Tabela D.1 definem-se por:

K1 - o número de operações aritméticas necessárias para obter o incremento de busca utilizando-se, ao aplicar o procedimento proposto, das equações do Filtro de Kalman apresentadas por Bryson e Ho (1975) de forma não sequencial;

K2 - o número de operações aritméticas necessárias para obter o incremento de busca utilizando-se, ao aplicar o procedimento proposto, das equações do Filtro de Kalman apresentadas por Bryson e Ho (1975) de forma sequencial;

K3 - o número de operações necessárias para obter o incremento de busca utilizando-se, ao aplicar o procedimento proposto, das equações do Filtro de Kalman, tal como apresentadas por Liebelt (1967), não sequencialmente;

K4 - o número de operações necessárias para obter o incremento de busca utilizando-se, ao aplicar o método proposto, das equações do Filtro de Kalman, tal como apresentadas por Liebelt (1967), sequencialmente;

K5 - o número de operações necessárias para obter o incremento de busca mediante a aplicação do procedimento proposto, e mediante a aplicação não-sequencial das equações do Filtro de Kalman tal como apresentadas por Jazwinski (1969), Equação 2.22.

K6 - o número de operações necessárias para obter o incremento de busca mediante a aplicação do procedimento proposto, e mediante a aplicação sequencial das equações do Filtro de Kalman tal como apresentadas por Jazwinski (1969);

Nw - o número de operações necessárias para obter o incremento de busca mediante o uso do método de Willianson (1971);

Nproj - o número de operações requeridas para obter o incremento de busca via método da projeção do gradiente, obtendo-se a matriz de projeção e a matriz de correções via inversão de matrizes;

Nproj - o número de operações requeridos para obter o incremento de busca via método da projeção do gradiente, quando se usam as relações recursivas para achar a matriz de projeção, e quando se calcula o incremento de correção do erro nos vínculos juntamente com tal matriz;

C1 - o número de operações aritméticas requeridas por iteração de busca de uma solução viável pelo algoritmo Simplex, no caso de se aplicar o método de Ceballos (1979) para resolver o problema restrito de otimização;

C2 - o número de operações aritméticas requeridas por iteração de busca do mínimo do algoritmo Simplex, caso se resolva o problema restrito de otimização com o método de Ceballos (1979);

Max1 - o limite superior do número de iterações da fase de busca de uma solução viavél do Simplex no caso do método de Ceballos (1979;

Max2 - o limite superior do número de iterações da fase de busca do mínimo do Simplex do caso do método de Ceballos (1979);

np - o número de parâmetros a otimizar:

m - o número de vínculos.

Definidos tais parâmetros pode-se montar a Tabela D.1:

TABELA D.1

PARÂMETROS DE AVALIAÇÃO DE CUSTO

PARÂMETRO	EXPRESSÃO DE CÁLCULO
K1	$4np^2m + 4m^2np + 2m^3 + 3np.m - m^2/2 + m/2 - 1$
K2	$7np^2.m + 4np^2 + 5np.m + np$
K3	$4m^2.np + 2m^3 + 4m.np - m^2/2 + m/2 + np - 1$
K4	$4np^2.m - 4np^2 + 5m.np + 4np + m$
K5	$4m^2np + 2m^3 + 4m.np - m^2/2 + m/2 + np - 1$
K6	$4np.m^3 + 4np^2.m - 5np^3 - np^2 + 10np.m + m - 2np$
Nw	$2(np + m)^3 + (np + m)^2/2 - (np + m)/2 - 1$
C1	$5np + 3m + 3$
C2	$4np + 2$
Max1	$(2np + m)(2np + m - 1)(2np + m - 2)/3!$
Max2	$(2np)(2np - 1)(2np - 2)/3!$

Deve-se notar que algumas versões do Filtro de Kalman permitem obter o incremento de busca sem que para isto seja necessário obter a matriz de covariância do erro em tal incremento; tal fato é uma das causas de diferença entre os valores de K1, K2, K3, K4, K5 e K6.

Assim, no caso de se utilizar das expressões do Filtro de Kalman, proposta por Liebelt (1967) e Jazwinski (1969), a obtenção de tal matriz pode ser feita posteriormente à determinação do incremento de busca. Necessita-se efetuar 2np.m operações para obter tal matriz caso se utilize a versão do Filtro de Kalman apresentada por Liebelt (1969) de forma não-sequencial, e 2np operações adicionais para

obter tal matriz, caso se empregue sequencialmente esta versão do Filtro de Kalman. Caso se aplique a versão do Filtro de Kalman apresentada por Jazwinski, não-sequencialmente, necessita-se efetuar um total de $2np^3 + 5m.np^2$ operações pra obter tal matriz, e caso se aplique sequencialmente tais equações, $4np^2 + 2np$ operações aritméticas.