

# DECOMPOSIÇÃO EM PROGRAMAÇÃO MATEMÁTICA

## VIA PROJEÇÃO – ANÁLISE E APLICAÇÃO

MARISTELA ROCHA BELLONI

Orientador:

Prof. Dr. José Cláudio Geromel

Dissertação apresentada no Instituto de Matemática, Estatística e Ciência da Computação, como requisito parcial para obtenção do título de Mestre em Matemática Aplicada.

Agosto/1982

UNIVERSIDADE P  
BIBLIOTECA CENTRAL

Aos meus pais  
pelo amor e  
formação que  
me deram

## AGRADECIMENTOS

*Ao terminarmos este trabalho, não poderemos esquecer a todos que me ajudaram, e em especial*

*– aos professores do IMECC e da FEC, especialmente aos meus amigos Martinez e França;*

*– ao Geromel pelo tema da tese e pelo belo trabalho de orientação;*

*– aos meus amigos que me acompanharam durante todo o curso;*

*– ao Zê, pelo seu carinho e paciência de todos os dias;*

*– ao Sr. João Baptista Esteves de Oliveira pelo lindo trabalho de datilografia.*

# I N D I C E

INTRODUÇÃO GERAL. . . . .	ix
CAP. 1 - TEORIA DA DUALIDADE. . . . .	1
1.1 - Introdução . . . . .	1
1.2 - Conceitos Básicos. . . . .	2
1.3 - Equivalência Primal/Dual . . . . .	4
1.4 - Diferenciabilidade da Função Dual. . . . .	12
1.5 - O Método Dual Clássico . . . . .	15
1.6 - Conclusão. . . . .	17
CAP. 2 - ANÁLISE DE APLICABILIDADE DA DECOMPOSIÇÃO BENDERS GENERALIZADA . . . . .	18
2.1 - Introdução . . . . .	18
2.2 - Considerações Preliminares sobre o Problema em estudo . . . . .	19
2.3 - Funções Suporte. . . . .	23
2.4 - Um Algoritmo de Relaxação. . . . .	24
2.4.1 - Decomposição Benders Generalizada como caso particular . . . . .	26
2.4.2 - Análise de aplicabilidade da decomposição Benders Generalizada . . . . .	30
2.5 - Um Algoritmo Alternativo - Método de Projeção. . . . .	31
2.5.1 - Estudo da diferenciabilidade da função perturbação generalizada. . . . .	32
2.5.2 - Análise de aplicabilidade do método de projeção. . . . .	33
2.6 - Conclusão. . . . .	34
CAP. 3 - APLICAÇÃO DA DECOMPOSIÇÃO BENDERS GENERALIZADA NO CONTROLE ÓTIMO DE CENTRAIS TÉRMICAS . . . . .	36
3.1 - Introdução . . . . .	36
3.2 - Definição do Problema em Estudo. . . . .	37
3.2.1 - A função custo de ligação. . . . .	38
3.3 - Aplicação da Técnica de Decomposição Benders Generalizada ao Problema das Usinas Térmicas . . . . .	41
3.3.1 - O método de solução do sub-problema. . . . .	47
3.3.2 - O método de solução do problema mestre . . . . .	54
3.4 - Conclusão. . . . .	56
CAP. 4 - UMA APLICAÇÃO NUMÉRICA NUM SISTEMA TÉRMICO REAL. . . . .	58
4.1 - Introdução . . . . .	58

4.2 - Dados Relativos ao Sistema Considerado . . . . .	58
4.3 - As Soluções Obtidas. . . . .	61
4.3.1 - Caso 1 . . . . .	61
4.3.2 - Caso 2 . . . . .	64
4.3.3 - Outros testes realizados . . . . .	70
4.4 - Conclusão. . . . .	72
CONCLUSÃO GERAL . . . . .	75
REFERÊNCIAS BIBLIOGRÁFICAS. . . . .	76
APÊNDICE - Listagem do Programa FORTRAN Desenvolvido. . . . .	78

## INTRODUÇÃO GERAL

Dentre os inúmeros problemas de programação matemática, existe uma classe que se caracterizam por possuírem variáveis chamadas "complicantes", no sentido de que o problema torna-se mais simples sempre que tais variáveis são fixadas temporariamente.

A partir da observação de que, em alguns casos, a fixação destas variáveis resultava em sub-problemas lineares, J. F. Benders [1962] publicou em 1962 seu método de decomposição, que foi generalizado para problemas não lineares por A. M. Geoffrion no início da década passada [1972].

Este método tem se mostrado muito eficiente na solução de certos problemas práticos importantes, como os problemas de localização [França, 1979], [Luna, 1978] e de controle ótimo de sistemas hidro-térmicos de geração de energia [Geromel & Baptistella, 1980].

Existe, entretanto, uma dificuldade que impede a aplicação desta técnica de decomposição para uma gama muito maior de problemas. Tal dificuldade é devida a necessidade da verificação de uma propriedade, conhecida como propriedade P [Geoffrion, 1972], que reduz a classe de problemas para os quais a técnica, como proposta por Geoffrion, pode ser utilizada.

No presente trabalho estudamos a decomposição Benders Generalizada a partir do conceito de *funções suporte*, que será in-

introduzido no capítulo 2, o qual permite a obtenção do método e de sua interpretação geométrica de um modo mais simples.

No entanto, veremos que nem todos os problemas que possuem variáveis "complicantes" são resolúveis deste modo, dado que a aplicabilidade do método depende da verificação da citada propriedade P. Para problemas que não verificam esta propriedade, propomos então obter sua solução por um método de tipo gradiente projetado, sempre que a função objetivo verificar determinadas condições de diferenciabilidade, que serão amplamente estudadas neste trabalho.

Após este estudo teórico, vamos considerar o problema de controle ótimo de centrais térmicas de geração de energia, que será decomposto segundo o método Benders generalizado. Aqui nosso objetivo é dar um exemplo do uso desta técnica e sobretudo quantificar a sua eficiência na solução de um problema prático.

A obtenção do método Benders generalizado e a sua aplicação na solução de problemas de controle deste tipo, porém, só é possível a partir de um estudo detalhado de alguns conceitos básicos da teoria da dualidade, tais como a equivalência primal-dual e diferenciabilidade da função dual.

O presente trabalho está assim dividido:

Capítulo 1 - Trata de conceitos básicos da teoria da dualidade, com ênfase nas condições da equivalência primal-dual e diferenciabilidade da função dual.

Capítulo 2 - Neste capítulo, o método Benders generalizado é estudado a partir do conceito de *funções suporte*. Para problemas que não verificam a propriedade P,

propomos um método de solução de tipo gradiente projetado.

Capítulo 3 - A decomposição de Benders Generalizada clássica é aplicada ao problema de controle ótimo de geração térmica de energia. Também são estudados em detalhes o problema mestre e o subproblema. Os métodos de solução para estes problemas são obtidos considerando suas estruturas particulares.

Capítulo 4 - Com o estudo desenvolvido no capítulo anterior, são apresentadas algumas soluções para o problema de controle ótimo estudado.

## CAPÍTULO 1

### TEORIA DA DUALIDADE

#### 1.1 - INTRODUÇÃO

A teoria da dualidade, quando aplicada a problemas de estruturas distintas, resulta em problemas, chamados problemas duais que, de acordo com cada uma destas estruturas, podem ser resolvidos através de, basicamente, três tipos de estratégias (Geoffrion [1970], Luna [1978]).

Freqüentemente a dualidade induz separabilidade, que permite que se resolva um problema originalmente definido no  $\mathbb{R}^n$  via solução de  $N$  problemas definidos no  $\mathbb{R}^{n_i}$ , com  $\sum_{i=1}^N n_i = n$ .

Outra manipulação muito usada na solução de problemas de grande porte é a linearização externa, que pode ser estendida (ver Capítulo 2) para funções não convexas através de aproximações não lineares.

Nosso objetivo neste capítulo, é examinar quando a função dual resultante do problema manipulado é diferenciável, para que possam ser usados métodos de tipo gradiente como estratégia de solução. Nosso interesse deve-se ao fato de que, além dos métodos de tipo gradiente serem os mais simples de serem implementados na obtenção da solução de problemas de programação matemática, o gradiente da função dual, quando existe, é

muito fácil de ser calculado.

Para atingirmos nosso objetivo, faremos um estudo neste capítulo das condições de existência, equivalência e diferenciabilidade do problema dual.

## 1.2 - CONCEITOS BÁSICOS

Consideremos inicialmente o problema abaixo, chamado *problema primal* (P):

$$\min_{x \in X} f(x) \text{ s.a } g(x) \leq 0 \quad (1)$$

onde  $X \subseteq \mathbb{R}^n$  é um conjunto compacto,  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

Associado ao problema (P) existe um outro problema, chamado *problema dual* (D), relativo às restrições  $g$ , dado por:

$$\max_{\lambda \geq 0} \phi(\lambda) \quad (2)$$

onde

$$\phi(\lambda) = \min_{x \in X} \{f(x) + \lambda'g(x)\} \quad (3)$$

é a *função dual* associada ao problema (P).

Devemos ainda evidenciar a função

$$L(x, \lambda) = f(x) + \lambda'g(x) \quad (4)$$

chamada *função lagrangeana* que desempenha papel muito importante na obtenção da solução ótima do problema (2) e na equivalência dos problemas (1) e (2) que vamos estabelecer na próxima secção.

Segue diretamente das definições (3) e (4) que o domínio de definição da função dual é

$$D = \{\lambda \mid \lambda \geq 0, \min_{x \in X} L(x, \lambda) \text{ existe}\},$$

ou seja,  $D$  é um conjunto dos vetores  $\lambda$  não negativos para os quais  $L(x, \lambda)$  tem um ínfimo finito sobre  $X$ .

Seguem também diretamente da definição de função dual os seguintes resultados, que são facilmente estabelecidos, e que são muito poderosos, conforme veremos durante o desenvolvimento deste trabalho.

TEOREMA 1.1 - A função dual  $\phi(\lambda) = \min_{x \in X} L(x, \lambda)$  é côncava sobre qualquer subconjunto convexo de seu domínio  $D$ .

PROVA - Seja  $\tilde{D}$  um subconjunto convexo de  $D$  e sejam  $\lambda_1, \lambda_2 \in \tilde{D}$ . Então para  $\forall \alpha \in [0, 1]$  temos

$$\phi(\alpha\lambda_1 + (1-\alpha)\lambda_2) = \min_{x \in X} L(x, \alpha\lambda_1 + (1-\alpha)\lambda_2). \quad (5)$$

Desde que a função lagrangeana  $L(x, \lambda)$  é obviamente linear em  $\lambda$ ,

$$\begin{aligned} \phi(\alpha\lambda_1 + (1-\alpha)\lambda_2) &= \min_{x \in X} [\alpha L(x, \lambda_1) + (1-\alpha)L(x, \lambda_2)] \\ &\geq \alpha \min_{x \in X} L(x, \lambda_1) + (1-\alpha) \min_{x \in X} L(x, \lambda_2) \\ &\geq \alpha\phi(\lambda_1) + (1-\alpha)\phi(\lambda_2), \end{aligned} \quad (6)$$

ou seja, a função dual é côncava.

É importante observar que se  $D$  for um conjunto convexo, o problema dual resulta num problema côncavo sobre um conjunto

convexo e isto nos garante que não existem pontos de ótimo distintos do ótimo global.

TEOREMA 1.2 (Teorema Fraco da Dualidade) - Seja  $x \in X$  é factível para (P) e  $\lambda$  é factível para (D), então  $f(x) \geq \phi(\lambda)$ .

PROVA - Seja  $\lambda \in D$ . Então

$$\phi(\lambda) = \min_{x \in X} \{f(x) + \lambda'g(x)\}, \quad (7)$$

de modo que

$$\phi(\lambda) \leq f(x) + \lambda'g(x), \quad \forall x \in X. \quad (8)$$

Logo, para todo  $x \in X$  factível, ou seja, tal que  $g(x) \leq 0$ , tem-se  $\phi(\lambda) \leq f(x)$  e o teorema está provado.

COROLÁRIO 1.2.1 - Qualquer solução factível de (P) é um limitante superior de (D) e qualquer solução factível de (D) é um limitante inferior de (P).

COROLÁRIO 1.2.2 - Se (P) é ilimitado, então (D) é infactível.

Se (D) é ilimitado, então (P) é infactível.

Neste ponto cabe observar que o teorema fraco da dualidade, além de ser usado na obtenção de uma série de resultados importantes, derivados da teoria da dualidade, também serve eventualmente de critério de parada de algoritmos que visam a obtenção da solução do problema dual.

### 1.3 - EQUIVALÊNCIA PRIMAL/DUAL

Muitos problemas de programação matemática, inicialmente bastante complicados do ponto de vista de solução, tornam-se

bastante simples quando manipulados através da dualidade. Entretanto, nem sempre a solução do problema dual é equivalente à solução do problema primal ao qual está associado. Isto se deve ao fato de que pode existir um "intervalo de dualidade" entre o valor ótimo do problema (P) e o valor ótimo do problema (D), ainda que ambos tenham solução. Através das definições e dos teoremas que seguem veremos como detectar a existência ou não de tal intervalo, viabilizando assim o uso da teoria da dualidade.

DEFINIÇÃO - Um ponto  $(\bar{x}, \bar{\lambda})$  com  $\bar{\lambda} \geq 0$  e  $\bar{x} \in X$  é chamado *ponto de sela do lagrangeano*  $L(x, \lambda)$  associado ao problema (P) se

- i)  $L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda})$  para todo  $x \in X$ , ou seja  $\bar{x}$  minimiza  $L(x, \bar{\lambda})$  sobre  $X$ ; (9)
- ii)  $L(\bar{x}, \bar{\lambda}) \geq L(\bar{x}, \lambda)$  para todo  $\lambda \geq 0$ , ou seja  $\bar{\lambda}$  maximiza  $L(\bar{x}, \lambda)$  sobre  $\lambda \geq 0$ .

Os pontos de sela do lagrangeano são caracterizados no teorema abaixo, que nos dá condições necessárias e suficientes para a existência de tais pontos.

TEOREMA 1.3 - Sejam  $\bar{\lambda} \geq 0$  e  $\bar{x} \in X$ . Então  $(\bar{x}, \bar{\lambda})$  é um ponto de sela de  $L(x, \lambda)$  se, e somente se,

- a)  $\bar{x}$  minimiza  $L(x, \bar{\lambda})$  sobre  $X$ ;
- b)  $g_i(\bar{x}) \leq 0$ ,  $i=1, \dots, m$ ; (10)
- c)  $\bar{\lambda}' g(\bar{x}) = 0$ ,  $i=1, \dots, m$ .

PROVA - ( $\Rightarrow$ ) A parte a) é trivial da definição de ponto de sela.

Para obtermos b), suponhamos que para algum índice  $i \in I = \{1, \dots, m\}$ ,  $g_i(\bar{x}) > 0$ . Tomando  $\lambda_i > \bar{\lambda}_i$  obtemos  $\lambda_i g_i(\bar{x}) > \bar{\lambda}_i g_i(\bar{x})$ . Então

$$\begin{aligned} f(\bar{x}) + \sum_{\substack{j \in I \\ j \neq i}} \bar{\lambda}_j g_j(\bar{x}) + \lambda_i g_i(\bar{x}) &> \\ f(\bar{x}) + \sum_{\substack{j \in I \\ j \neq i}} \bar{\lambda}_j g_j(\bar{x}) + \bar{\lambda}_i g_i(\bar{x}) &= L(\bar{x}, \bar{\lambda}) \end{aligned} \quad (11)$$

ou seja,  $(\bar{x}, \bar{\lambda})$  não seria ponto de máximo do lagrangeano em relação a  $\lambda$ .

Portanto  $g_i(\bar{x}) \leq 0$ ,  $\forall i=1, \dots, m$ .

Para provar a parte c) suponhamos que  $g_i(\bar{x}) < 0$ . Se  $\bar{\lambda}_i > 0$  bastaria tomar  $0 < \lambda_i < \bar{\lambda}_i$  para fazer subir o valor de  $L(\bar{x}, \lambda)$ , contrariando o fato de que  $\bar{\lambda}$  maximiza o lagrangeano sobre  $\lambda \geq 0$ , de modo que a condição c) deve valer.

( $\Leftarrow$ ) - Seja  $\lambda \geq 0$ . Então

$$\begin{aligned} L(\bar{x}, \lambda) &= f(\bar{x}) + \sum_{i=1}^m \lambda_i g_i(\bar{x}) \stackrel{(b)}{\leq} f(\bar{x}) \stackrel{(c)}{=} \\ &= f(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i g_i(\bar{x}) = L(\bar{x}, \bar{\lambda}), \end{aligned} \quad (12)$$

ou seja,  $\bar{\lambda}$  maximiza  $L(\bar{x}, \lambda)$  sobre  $\lambda \geq 0$ , e o teorema está provado.

Podemos agora evidenciar a importância dos pontos de sela do lagrangeano através do seguinte teorema.

TEOREMA 1.4 (Suficiência dos pontos de sela) - Se  $(\bar{x}, \bar{\lambda})$  é um ponto de sela de  $L(x, \lambda)$ , então  $\bar{x}$  resolve o problema primal (P).

PROVA - Desde que  $(\bar{x}, \bar{\lambda})$  é o ponto de sela, valem as condições a) - c) do Teorema 1.3.

Tomando  $g(x) = (g_1(x), \dots, g_m(x))'$ , da parte a) do mesmo teorema resulta

$$f(\bar{x}) + \bar{\lambda}' g(\bar{x}) \leq f(x) + \bar{\lambda}' g(x), \quad \forall x \in X. \quad (13)$$

Pela condição c),  $\bar{\lambda}' g(\bar{x}) = 0$ , de modo que

$$f(\bar{x}) \leq f(x) + \bar{\lambda}' g(x), \quad \forall x \in X. \quad (14)$$

Para todo ponto  $x \in X$  tal que  $g(x) \leq 0$ , ou seja, para todo  $x$  factível para (P) o termo  $\bar{\lambda}' g(x)$  é não positivo, de modo que obtemos da expressão anterior que

$$f(\bar{x}) \leq f(x), \quad \forall x \in X \text{ factível.}$$

Logo,  $\bar{x}$  resolve o problema primal.

DEFINIÇÃO - A função

$$v(y) = \min_{x \in X} \{f(x) / g(x) \leq y\} \quad (15)$$

definida em  $\mathbb{R}^n$ , é chamada *função de perturbação* associada a (P), onde  $y$  é o vetor de perturbação.

Seguem diretamente da definição acima que  $v(0)$  é o valor de  $f(\cdot)$  na solução de (P), e se  $y_2 \geq y_1$ , então  $v(y_2) \leq v(y_1)$ , pois se a região factível for "maior" o mínimo será menor, ou seja, a função de perturbação é não-crescente.

Além disso, é fácil verificar [Geoffrion, 1971] que a função de perturbação, sob hipóteses de convexidade de  $f(\cdot)$  e  $g(\cdot)$ , é uma função convexa. Porém, para que não ocorra o intervalo de dualidade, que inviabiliza o uso de métodos duais para a

obtenção da solução do problema primal, é necessário que o problema (P) seja estável [Geoffrion, 1971]. Esta condição deve-se ao fato de que, ainda que  $v(\cdot)$  seja convexa existem casos nos quais, por exemplo,  $v(y) \rightarrow \infty$  para  $y \rightarrow 0$  (Figura 1.1) [Geoffrion, 1971].

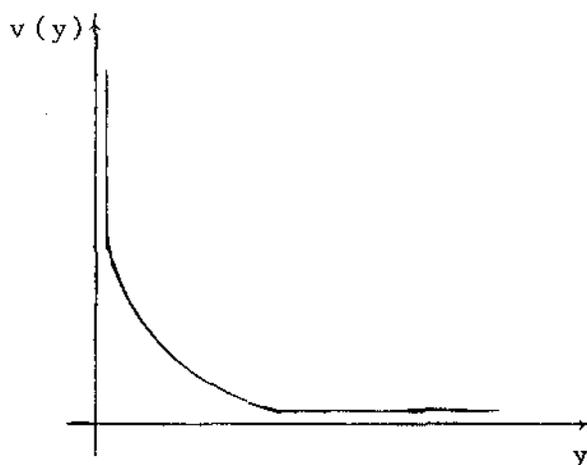


Figura 1.1 - Exemplo de função de perturbação para um problema (P) não estável.

Felizmente, a estabilidade de (P) é facilmente obtida se (P) verificar alguma das condições de regularidade [Geoffrion, 1971].

Veremos a seguir que, na verdade, basta que possamos garantir a existência de um hiperplano suporte em  $v(0)$  para que não ocorra o intervalo de dualidade, ou seja, para que os problemas primal e dual sejam equivalentes.

Devemos ainda frisar que se a função  $v(\cdot)$  não for convexa, pode ser impossível determinar o hiperplano suporte, cujo gradiente é exatamente o vetor das variáveis duais ótimas.

TEOREMA 1.5 (Karlin, 1959) - Sejam  $X \subseteq \mathbb{R}^n$ , convexo,  $f(\cdot)$  definida em  $X$ , convexa, e  $g(x) = (g_1(x), \dots, g_m(x))'$  um vetor de funções

convexas definidas em  $X$ . Suponhamos que existe  $x \in X$  tal que  $g(x) < 0$ . Se  $\bar{x}$  é um ponto no qual  $f(\cdot)$  assume seu mínimo sujeito a  $g(x) \leq 0$ ,  $x \in X$ , então existe um vetor de multiplicadores  $\bar{\lambda} \geq 0$  tal que  $(\bar{x}, \bar{\lambda})$  é um ponto de sela do lagrangeano.

Reciprocamente, se  $(\bar{x}, \bar{\lambda})$  é um ponto de sela de  $L(x, \lambda)$ , então  $\bar{x}$  minimiza  $f(\cdot)$  sujeito a  $g(x) \leq 0$ ,  $x \in X$ .

Para provarmos este teorema necessitamos do seguinte lema.

LEMA 1.1 (Teorema da Separação) [Karlin, 1959] - Sejam A e B conjuntos convexos do  $\mathbb{R}^n$ . Suponhamos que A contém pelo menos um ponto interior e que nenhum ponto interior de A está em B. Então existe um hiperplano que separa A e B, ou seja, existe um vetor  $c \neq 0$  e uma constante  $\alpha$  tal que

$$\begin{aligned} c'x &\leq \alpha, & \forall x \in A \\ c'x &\geq \alpha, & \forall x \in B. \end{aligned} \tag{16}$$

PROVA [Bazaraa, 1979].

Isto posto, podemos passar a provar o Teorema de Karlin.

Seja  $(\bar{x}, \bar{\lambda})$ , com  $\bar{x} \in X$ ,  $\bar{\lambda} \geq 0$ , pontos de sela do lagrangeano. Então pelo Teorema 1.4  $\bar{x}$  resolve o problema primal.

Para provar a outra parte do teorema, definimos os conjuntos

$$A = \{(\bar{y}, y) \mid \bar{y} \geq f(x) \text{ e } y \geq g(x), \text{ para algum } x \in X\},$$

que consiste exatamente do epígrafo de  $v(\cdot)$ , e

$$B = \{(\bar{y}, y) \mid \bar{y} \leq f(\bar{x}), y_i \leq 0, i=1, \dots, m\}$$

(ver Figura 1.2).

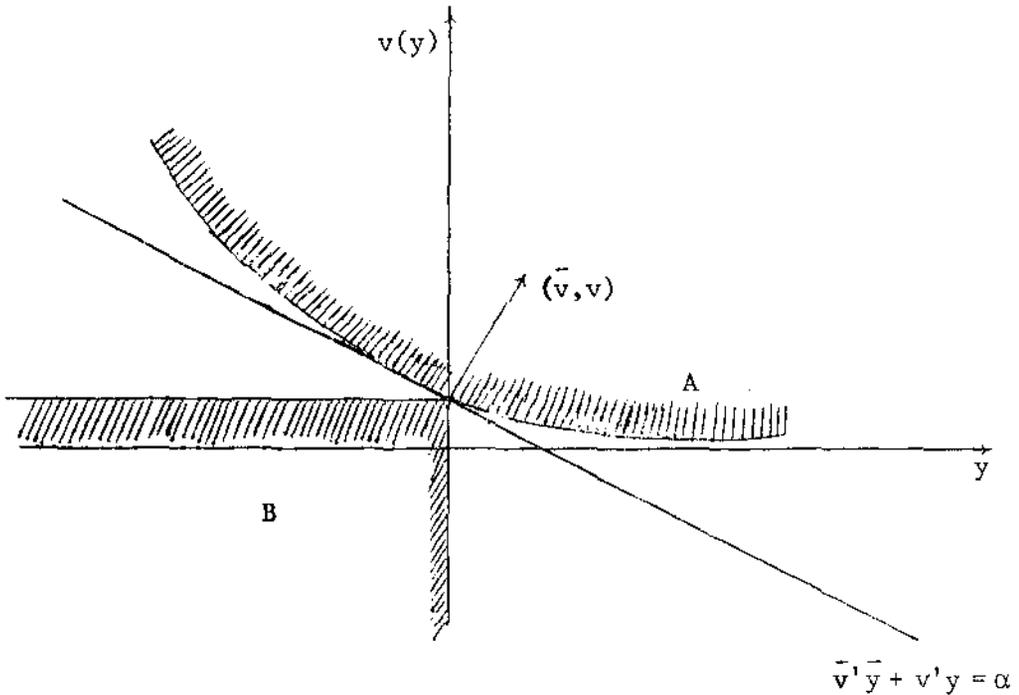


Figura 1.2 - Função de perturbação para (P) estável e hiperplano suporte em  $v(0)$ .

Vejam os que valem as hipóteses do Lema 1.1:

- i)  $\text{Int}(B) \neq \emptyset$  obviamente;
- ii) suponhamos que  $(\bar{y}, y) \in \text{Int}(B)$ . Então  $y < 0$  e

$$\begin{aligned} \bar{y} < f(\bar{x}) &\Rightarrow v(y) \geq v(0) = f(\bar{x}) > \bar{y} \Rightarrow v(y) > \bar{y} \Rightarrow \\ &\Rightarrow (\bar{y}, y) \notin \text{epígrafo de } v \Rightarrow \\ &\Rightarrow A \cap \text{Int}(B) = \emptyset. \end{aligned}$$

Logo, existe um hiperplano, de coeficientes  $(\bar{v}, v) \neq (0, 0)$ , que separa A e B. Obtemos então

$$\bar{v}'\bar{y} + v'y \geq \bar{v}'\bar{z} + v'z \tag{17}$$

para todo  $(\bar{y}, y) \in A$  e para todo  $(\bar{z}, z) \in B$ .

Se alguma das componentes de  $(\bar{v}, v)$  fosse negativa, então

a componente correspondente de  $(\bar{z}, z)$  poderia ser tomada suficientemente negativa, de modo que a desigualdade (17) não mais se verificaria. Portanto,  $(\bar{v}, v)$  tem todas as componentes não negativas.

Vejamos que  $\bar{v} > 0$ .

Desde que  $(f(\bar{x}), 0) \in B$  e  $(f(x), g(x)) \in A$ , para todo  $x \in X$ , de (17) resulta

$$\bar{v}' f(x) + v' g(x) \geq \bar{v}' f(\bar{x}) \quad \forall x \in X. \quad (18)$$

Se  $\bar{v} = 0$ , então, por (18), obteríamos

$$v' g(x) \geq 0, \quad \forall x \in X,$$

com pelo menos uma das componentes de  $v$  sendo positiva. Porém, por hipótese, existe  $x \in X$  tal que  $g(x) < 0$ , o que implica que  $\langle v, g(x) \rangle < 0$ . Logo,  $\bar{v} > 0$ .

Seja  $\bar{\lambda} = v/\bar{v}$ . Então, de (18) obtemos

$$f(x) + \bar{\lambda}' g(x) \geq f(\bar{x}), \quad \forall x \in X. \quad (19)$$

Se  $x = \bar{x}$ , obtemos  $\bar{\lambda}' g(\bar{x}) \geq 0$ . Porém,  $\bar{\lambda} \geq 0$  e

$$g(\bar{x}) \leq 0 \Rightarrow \bar{\lambda}' g(\bar{x}) \leq 0,$$

de modo que

$$\bar{\lambda}' g(\bar{x}) = 0. \quad (20)$$

De (19) e (20) obtemos que

$$L(\bar{x}, \bar{\lambda}) = f(\bar{x}) + \bar{\lambda}' g(\bar{x}) \leq L(x, \bar{\lambda}) = f(x) + \bar{\lambda}' g(x), \quad (21)$$

ou seja,  $\bar{x}$  minimiza o lagrangeano. Como  $g(\bar{x}) \leq 0$ ,  $\bar{\lambda} \geq 0$  e  $\bar{\lambda}' g(\bar{x}) = 0$ , segue do Teorema 1.3 que  $\bar{\lambda}$  maximiza o lagrangeano sobre  $\lambda \geq 0$ .

Logo,  $(\bar{x}, \bar{\lambda})$  é ponto de sela de  $L(x, \lambda)$  e está provado o teorema.

#### 1.4 - DIFERENCIABILIDADE DA FUNÇÃO DUAL

Suponhamos que, dado um problema primal (P), verificamos através dos critérios obtidos na secção anterior, que seu dual (D) lhe é equivalente. Devemos optar então por um método de solução do problema (D), de modo a obter esta solução da maneira mais eficiente possível, dentro das limitações que o problema nos impõe. Dado que geralmente os métodos de otimização no mínimo utilizam informações sobre o gradiente, vamos examinar sob que condições podemos usar tais métodos, ou seja, quando existe e como se obtém, neste caso, o gradiente da função dual. Para isto precisamos do seguinte teorema.

TEOREMA 1.6 (Danskin, 1966) - Seja

$$\psi(\lambda) = \min_{z \in S} F(z, \lambda), \quad (22)$$

onde  $F$  é uma função de classe  $C^1$  em relação à variável  $\lambda$  e  $S$  é um conjunto compacto. Então

$$D_+ \psi(\lambda, d) = \min_{z \in Z(\lambda)} \frac{d' \partial F(z, \lambda)}{\partial \lambda} \quad (23)$$

onde  $Z(\lambda)$  é o conjunto dos  $z$  que minimizam  $F(z, \lambda)$  e  $D_+ \psi(\lambda, d)$  é a derivada direcional à direita de  $\psi(\lambda)$ .

PROVA (Danskin, 1966).

De posse deste resultado, vamos retomar o problema (1) com as hipóteses adicionais de que  $f$  e  $g$  são contínuas em  $X$ . Tomando  $\psi(\lambda) = \phi(\lambda)$

função dual associada à (P), resulta imediatamente do teorema acima que:

$$D_+ \phi(\lambda, d) = \min_{x \in X(\lambda)} \frac{d' \partial L(x, \lambda)}{\partial \lambda} \quad (24)$$

Lembrando que  $D_- \phi(\lambda, d) = -D_+ \phi(\lambda, -d)$ , com (24) obtemos

$$D_- \phi(\lambda, d) = \max_{x \in X(\lambda)} \frac{d' \partial L(x, \lambda)}{\partial \lambda}. \quad (25)$$

Como se  $D_+ \phi(\lambda, d) = D_- \phi(\lambda, d)$ , para todo  $d$ , então  $\phi(\lambda)$  é diferenciável, escolhendo  $d$  como sendo as direções dos eixos coordenados, temos

$$\nabla \phi(\lambda) = \frac{\partial L(x, \lambda)}{\partial \lambda} \quad (26)$$

se o lado direito de (26) for constante em  $X(\lambda)$ .

É claro que, em particular, se  $X(\lambda)$  contiver um único elemento  $x^*$ , ou seja, se  $f(\cdot)$  for estritamente convexa,  $X(\lambda)$  é unitário, de modo que (26) se verifica, e neste caso

$$\nabla(\min_{x \in X} L(x, \lambda)) = \left. \frac{\partial L(x, \lambda)}{\partial \lambda} \right|_{x = x^*} \quad (27)$$

Finalmente, observando que

$$\frac{\partial L(x, \lambda)}{\partial \lambda} = \frac{\partial}{\partial \lambda} (f(x) + \lambda' g(x)) = g(x), \quad (28)$$

podemos resumir os resultados obtidos acima no seguinte teorema.

TEOREMA 1.7 - Se  $X \subseteq \mathbb{R}^n$  é compacto e  $f(\cdot)$  e  $g(\cdot)$  são contínuas em  $X$ , a função dual é diferenciável num ponto  $\lambda^*$  se, e somen-

te se, cada  $g_i(\cdot)$  for constante sobre  $X(\lambda^*)$ . Neste caso, as derivadas parciais de  $\phi(\cdot)$  são dadas por

$$\frac{\partial \phi}{\partial \lambda_i} \Big|_{\lambda = \lambda^*} = g_i(x), \quad \forall x \in X(\lambda^*). \quad (29)$$

Finalmente, todas estas manipulações se justificam se a partir de uma solução do problema dual pudermos evidenciar facilmente uma solução do problema primal.

TEOREMA 1.8 - Sejam  $X \subseteq \mathbb{R}^n$ , compacto,  $f(\cdot)$  e  $g(\cdot)$  contínuas em  $X$  e  $\lambda^*$  uma solução do problema dual. Suponhamos que a função dual  $\phi(\cdot)$  é diferenciável em  $\lambda^*$ . Então qualquer  $x^* \in X(\lambda^*)$  resolve o problema primal.

PROVA - Como  $\phi(\cdot)$  é diferenciável em  $\lambda^*$ , valem as seguintes condições de otimalidade:

$$a) \lambda_i^* > 0 \Rightarrow \frac{\partial \phi}{\partial \lambda_i} \Big|_{\lambda^*} = g_i(x^*) = 0, \quad x^* \in X(\lambda^*) \quad (30)$$

$$b) \lambda_i^* = 0 \Rightarrow \frac{\partial \phi}{\partial \lambda_i} \Big|_{\lambda^*} = g_i(x^*) \leq 0, \quad x \in X(\lambda^*). \quad (31)$$

Estas condições significam que  $g(x^*) \leq 0$  e  $\lambda^{*'} g(x^*) = 0$ ,  $x^* \in X(\lambda^*)$ , de modo que, pelo Teorema 1.3,  $(x^*, \lambda^*)$ , é ponto de sela do lagrangeano. Logo, pelo Teorema 1.4,  $x^*$  resolve o problema primal.

Fica então estabelecido que, sob as hipóteses do teorema anterior, se  $\lambda^*$  resolve o problema dual e  $L(x, \lambda^*)$  tem um único mínimo sobre  $X$  no ponto  $x^*$ , ou seja, se  $f(\cdot)$  for estritamente convexa, então  $x^*$  resolve o problema primal.

### 1.5 - O MÉTODO DUAL CLÁSSICO

Retomando mais uma vez o problema (P), cuja solução estamos interessados em obter, suponhamos que (P) satisfaz as hipóteses do Teorema 1.7, ou seja,  $\phi(\cdot)$  é diferenciável. Podemos então usar um método de máxima subida, adaptado às restrições  $\lambda \geq 0$  para otimizar  $\phi(\cdot)$ . O procedimento, numa iteração genérica  $p$ , será o seguinte:

#### ALGORITMO 1.1

PASSO 1 - Para  $\lambda = \lambda_p$ , resolva o problema

$$\min_{x \in X} \{f(x) + \lambda_p' g(x)\}. \quad (32)$$

Seja  $x(\lambda_p)$  a solução. Vá para o Passo 2.

PASSO 2 - Obtenha o valor da função dual e do seu gradiente para  $\lambda = \lambda_p$  e  $x = x(\lambda_p)$ , ou seja,

$$\begin{cases} \phi(\lambda_p) = L(x(\lambda_p), \lambda_p) \\ \nabla \phi(\lambda_p) = g(x(\lambda_p)). \end{cases} \quad (33)$$

Se  $\phi(\lambda_p) \geq \gamma$ , com  $\gamma$  suficientemente grande, pare. O problema (P) é infactível.

Se  $\nabla \phi(\lambda_p) - \epsilon \leq 0$ , pare. A solução  $\lambda_p$  é  $\epsilon$ -ótima e sua solução primal correspondente é  $x(\lambda_p)$ , com valor objetivo  $\phi(\lambda_p)$ . Caso contrário, vá para o Passo 3.

PASSO 3 - Defina uma nova direção de busca  $s_p$ , dada por:

$$s_p^i = \begin{cases} \left. \frac{\partial \phi}{\partial \lambda^i} \right|_{\lambda = \lambda_p}, & \text{se } \lambda_p^i > 0, \quad i = 1, \dots, m \\ \text{m\`ax} \left\{ 0, \left. \frac{\partial \phi}{\partial \lambda^i} \right|_{\lambda = \lambda_p} \right\}, & \text{se } \lambda_p^i = 0, \quad i = 1, \dots, m \end{cases} \quad (34)$$

O novo vetor  $\lambda_{p+1}$  é dado por

$$\lambda_{p+1} = \lambda_p + \alpha_p s_p, \quad (35)$$

onde o passo  $\alpha_p$  deve ser tal que

$$\phi(\lambda_{p+1}) > \phi(\lambda_p). \quad (36)$$

O passo  $\alpha_p$  pode ser obtido maximizando

$$h(\alpha) = \phi(\lambda_p + \alpha s_p), \quad (37)$$

sujeito às restrições  $\alpha_p \geq 0$  e  $\lambda_{p+1} \geq 0$ . Faça  $p = p+1$  e volte ao Passo 1.

Devemos observar, quanto ao procedimento apresentado acima, que nada foi dito sobre como resolver o problema (32). Ocorre que em muitas situações a função lagrangeano, que é otimizada no Passo 1, é separável nas variáveis  $x$ , de modo que a solução de (32) pode ser obtida através da solução de  $N$  problemas definidos no  $\mathbb{R}^{n_i}$ , com

$$\sum_{i=1}^N n_i = n,$$

simplificando consideravelmente o trabalho computacional nes-

ta etapa. Portanto, a solução do problema lagrangeano depende essencialmente da natureza do problema que desejamos resolver, e a decisão sobre qual método devemos usar para obtermos tal solução deve ser tomada depois de estudarmos o problema em questão.

Os outros passos do método dual apresentado também podem, e devem, ser modificados sempre que a natureza do problema a ser resolvido o permitir visando um desempenho melhor do mesmo.

#### 1.6 - CONCLUSÃO

Neste capítulo estabelecemos alguns resultados fundamentais da teoria da dualidade, como equivalência primal-dual e diferenciabilidade da função dual. Embora a literatura sobre estes temas seja farta, optamos por introduzir a tese através deste capítulo, dada a grande importância de todos os conceitos estabelecidos acima na obtenção do Método Benders Generalizado e também na solução de problemas de controle ótimo de centrais térmicas de geração de energia, conforme veremos nos próximos capítulos.

## CAPÍTULO 2

### ANÁLISE DE APLICABILIDADE DA DECOMPOSIÇÃO BENDERS GENERALIZADA

#### 2.1 - INTRODUÇÃO

A partir da publicação por A. M. Geoffrion [Geoffrion, 1972] da decomposição Benders Generalizada, vários autores demonstraram sua eficiência na solução de certos problemas práticos importantes, tais como os problemas de localização [França, 1979], [Luna, 1978] e problemas de controle ótimo de sistemas térmicos de geração de energia (vide Capítulos 3 e 4) [Geromei e Baptistella, 1981].

No entanto, várias dificuldades ainda existem, as quais impedem a aplicação desta técnica de decomposição para uma gama muito maior de problemas de programação matemática. A principal delas é que os problemas em consideração devem possuir necessariamente uma chamada propriedade P [Geoffrion, 1972], que é restritiva e que traduz a necessidade destes problemas terem uma estrutura particular.

Sem esta propriedade a técnica de decomposição Benders Generalizada, como inicialmente proposta, é inteiramente inviabilizada, pois a técnica de solução do "problema mestre" se mostra inoperante.

Neste capítulo, todas as hipóteses de aplicabilidade da decomposição Benders Generalizada serão explicitadas e analisadas em detalhes a partir do conceito de *função suporte*, que será introduzido, permitindo a obtenção mais simples de certas interpretações geométricas [Belloni & Geromel, 1981].

Em seguida, propomos um novo método de solução do problema mestre, que permite resolver por esta técnica de decomposição, problemas que não apresentam a propriedade P [Geoffrion, 1972]. Isto será feito a partir do estudo da diferenciabilidade da função objetivo do problema mestre utilizando resultados clássicos fornecidos no Capítulo 1.

## 2.2 - CONSIDERAÇÕES PRELIMINARES SOBRE O PROBLEMA EM ESTUDO

A decomposição Benders Generalizada foi proposta em [Geoffrion, 1972] para tratar problemas de programação matemática assim definidos:

$$\min_{x \in X, y \in Y} f(x, y) \text{ s.a } G(x, y) \leq 0 \quad (1)$$

onde  $X \subseteq \mathbb{R}^n$  é um conjunto convexo fechado,  $Y \subseteq \mathbb{R}^m$ ,  $f: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  e  $G: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^r$ .

Para este problema, suponhamos que seja válida a seguinte hipótese:

- o problema (1) não é convexo em  $x$  e  $y$  simultaneamente, mas fixando  $y$  ele torna-se convexo na variável  $x$ .

A variável  $y$  é chamada variável "complicante", desde que

não precisam ser feitas hipóteses sobre  $f$  e  $G$  em relação a esta variável para que a condição acima seja satisfeita. Além disso, dependendo da estrutura de cada problema a ser resolvido, podemos eventualmente obter problemas tais que, para  $y \in Y$  fixado

- i) (1) pode ser separado em um certo número de problemas independentes, cada um envolvendo um diferente subvetor  $x$ , ou
- ii) (1) assume uma estrutura para a qual são conhecidos métodos eficientes de solução.

A técnica de decomposição Benders Generalizada consiste em aproveitar esta estrutura separando o problema (1) em dois problemas interligados, a saber, um problema mestre e um subproblema, visando viabilizar a obtenção da solução de (1) através de uma seqüência de manipulações, desde a projeção de (1) no espaço das variáveis  $y$  seguida de linearizações externas e relaxação.

Numa primeira etapa obtemos então um problema na variável  $x$ , fixando  $y$ . Isto é feito através da manipulação chamada *projeção no espaço Y* [Geoffrion, 1970].

Projetando (1) no espaço das variáveis "complicantes"  $y$ , obtemos:

$$\min_y v(y) \text{ s.a } y \in Y \cap V \quad (2)$$

onde

$$v(y) = \min_{x \in X} f(x, y) \text{ s.a } G(x, y) \leq 0 \quad (3)$$

e

$$V = \{y \in \mathbb{R}^m \mid G(x,y) \leq 0, \text{ para algum } x \in X\}. \quad (4)$$

Devido a sua grande importância nesta e nas próximas seções, passaremos a nos referir ao membro direito da igualdade em (3) como sendo o problema (SP), ou seja,

$$(SP) \quad \min_{x \in X} f(x,y) \text{ s.a } G(x,y) \leq 0.$$

TEOREMA 2.1 (Projeção) - O problema (1) é infactível ou tem valor ótimo ilimitado se, e somente se, o mesmo vale para (2). Se  $(x^*, y^*)$  é ótimo para (1), então  $y^*$  tem que ser ótimo de (2). Se  $y^*$  é ótimo em (2) e  $x^*$  atinge o mínimo em (3) com  $y=y^*$  então  $(x^*, y^*)$  é ótimo em (1). Se  $\bar{y}$  é  $\epsilon_1$ -ótimo em (2) e  $\bar{x}$  é  $\epsilon_2$ -ótimo em (SP), então  $(\bar{x}, \bar{y})$  é  $(\epsilon_1 + \epsilon_2)$ -ótimo em (1).

PROVA - [A.M.Geoffrioni, 1970].

Fica claro que  $v(y)$  é o valor ótimo de (1) para  $y$  fixado e avaliar  $v(y)$  é mais fácil que resolver o problema (1).

O próximo passo consiste em caracterizar o conjunto  $V$ , definido em (4). Este conjunto consiste exatamente dos pontos  $y \in Y$  para os quais (SP) é factível. Deste modo  $Y \cap V$  pode ser pensado como sendo a projeção da região factível de (1) no espaço das variáveis  $y$ .

TEOREMA 2.2 (Representação de  $V$ ) - Suponhamos que  $X \neq \emptyset$  é convexo fechado e que  $G$  é convexa em  $X$ , para cada  $y \in Y$  fixado. Suponhamos também que o conjunto

$$Zy = \{z \in \mathbb{R}^m \mid G(x,y) \leq z, \text{ para algum } x \in X\}$$

é fechado para cada  $y$  fixado. Sob estas hipóteses,  $\forall \bar{y} \in Y$  fixado,  $\exists x \in X$  tal que  $G(x, \bar{y}) \leq 0$  se, e somente se a solução ótima  $\alpha^*$  de

$$\min_{\substack{\alpha \\ x \in X}} \alpha \text{ s.a } G(x, \bar{y}) - \alpha \cdot e \leq 0, \quad (5)$$

onde  $e \in \mathbb{R}^r$  é o vetor unitário, é não positiva.

PROVA ( $\Leftarrow$ ) Seja  $\alpha^* \leq 0$  solução de (5). Então  $G(x^*, \bar{y}) \leq \alpha^* \cdot e \leq 0$  e portanto  $\bar{y} \in V$ .

( $\Rightarrow$ ) Seja  $\bar{y} \in V$ . Então  $\exists \bar{x} \in X$  tal que  $G(\bar{x}, \bar{y}) \leq 0$ . Portanto  $\alpha = 0$  é factível para (5) e  $\alpha^* \leq 0$ , ou seja, a solução  $\alpha^*$  de (5) é não positiva.

Como  $Z_y$  é fechado e o problema (5) é convexo, seu dual lhe é equivalente. Através deste problema dual que segue podemos então caracterizar o conjunto  $V$ . De fato, o problema dual de (5) é dado por

$$\begin{aligned} \max_{\theta \geq 0} \min_{x \in X, \alpha} \left\{ \alpha + \theta' [G(x, \bar{y}) - \alpha \cdot e] \right\} = \\ \max_{\theta \geq 0} \left\{ \min_{\alpha} \alpha \left( 1 - \sum_{i=1}^r \theta_i \right) + \min_{x \in X} \theta' G(x, \bar{y}) \right\} = \\ \max_{\theta \geq 0} \min_{x \in X} \theta' G(x, \bar{y}) = \alpha^*. \\ \sum_{i=1}^r \theta_i = 1 \end{aligned}$$

Como  $\bar{y} \in V$  se e somente se  $\alpha^* \leq 0$ , resulta que  $\forall y \in Y$ ,  $y$  também está em  $V$  se, e somente se,

$$\max_{\theta \in \Theta} \min_{x \in X} \theta' G(x, y) \leq 0$$

onde

(6)

$$\Theta = \left\{ \theta \in \mathbb{R}^r \mid \theta_i \geq 0, \sum_{i=1}^r \theta_i = 1 \right\}.$$

### 2.3 - FUNÇÕES SUPORTE

DEFINIÇÃO - Seja  $f(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}$  uma função contínua em todos os pontos do seu domínio. Diremos que  $\xi(\cdot)$  é uma *função suporte* de  $f(\cdot)$  no ponto  $x \in \mathbb{R}^n$  se, e somente se,

$$\begin{cases} f(x) = \xi(x) \\ f(y) \geq \xi(y), \forall y \in \mathbb{R}^n, y \neq x. \end{cases} \quad (7)$$

Em particular, se  $f(\cdot)$  for convexa, então é claro que  $\xi(\cdot)$  pode ser escolhida linear e, conseqüentemente, existe para todo  $x \in \mathbb{R}^n$  um vetor  $\mu(x) \in \mathbb{R}^n$  tal que

$$\xi(y) = f(x) + \mu(x)'(y-x). \quad (8)$$

Além disso, se  $f(\cdot)$  for convexa e de classe  $C^1$  (admitir derivadas contínuas), então  $\forall x \in \mathbb{R}^n$ ,  $\mu(x)$  é único e é dado por

$$\mu(x) = \nabla f(x). \quad (9)$$

Se por outro lado,  $f(x)$  não for convexa, mas puder ser colocada na forma

$$f(x) = g(h(x)), \quad (10)$$

onde  $g(\cdot): \mathbb{R}^m \rightarrow \mathbb{R}$  é uma função convexa e  $h(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}^m$  é uma função qualquer, então

$$f(x) = g(\omega) \Big|_{\omega = h(x)}$$

e temos com (8)

$$\tilde{\xi}(z) = g(\omega) + \mu(\omega)'(z-\omega) \quad (11)$$

que é válida para todo  $z$  do conjunto imagem de  $h(\cdot)$ . Conseqüentemente fazendo  $z = h(y)$  obtemos

$$\xi(y) = f(x) + \mu(\omega)'[h(y)-h(x)], \quad \forall y \in \mathbb{R}^n \quad (12)$$

onde  $\mu(\omega)$  é qualquer subgradiente de  $g(\cdot)$  no ponto  $\omega = h(x)$ .

Fica claro então que, se não nos restringirmos a procurar funções suporte lineares (planos suporte), então poderemos determinar aproximações para funções  $f(\cdot)$  não convexas. Este facto é muito importante para a obtenção da decomposição Benders Generalizada, como veremos na próxima secção.

#### 2.4 - UM ALGORITMO GERAL DE RELAXAÇÃO

Seja um conjunto de  $p$  pontos  $x_j \in \mathbb{R}^n$ ,  $j = 1, \dots, p$ . Para cada  $x_j$ ,  $j = 1, \dots, p$ , determinamos uma função suporte  $\xi_j(x)$ .

Então, temos a seguinte aproximação de ordem  $p$  para a função  $f(\cdot)$  (ver Figura 2.1)

$$f(x) \approx f_p(x) = \min_{\sigma} \left\{ \sigma / \sigma \geq \xi_j(x), j=1, \dots, p \right\}. \quad (13)$$

Portanto, se quisermos resolver o problema

$$\min_{x \in X} f(x) \quad (14)$$

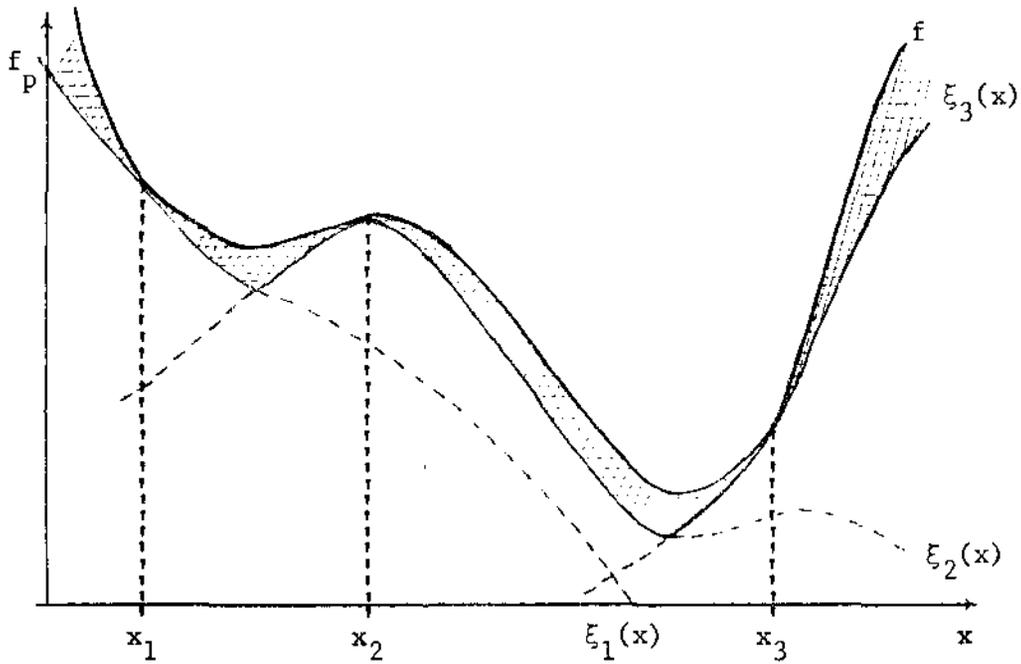


FIGURA 2.1 - Interpretação geométrica das funções suporte  $\xi_j(x)$  e da função aproximada  $f_p(x)$ .

com  $X \subseteq \mathbb{R}^n$ , podemos adotar o procedimento abaixo, denominado *relaxação* [Luna, 1978], que na  $p$ -ésima iteração pode ser descrito por

ALGORITMO 2.1

PASSO 1: Resolva o problema relaxado (ordem  $p$ )

$$\min_{x \in X} f_p(x),$$

ou seja,

$$\min_{x \in X, \sigma} \left\{ \sigma / \sigma \geq \xi_j(x), j = 1, \dots, p \right\}. \quad (15)$$

Sejam  $\sigma^*$  e  $x^*$  as soluções.

PASSO 2: Verifique se  $f(x^*) - \sigma \leq \epsilon$ , para  $\epsilon \geq 0$  arbitrariamente pequeno. Em caso afirmativo,  $x^*$  é uma solução  $\epsilon$ -óti-

ma de (14); caso contrário, vá ao passo seguinte.

PASSO 3: Faça  $x_{p+1} = x^*$ , determine  $\xi_{p+1}(x^*)$ , a função suporte no ponto  $x^*$  e volte ao passo 1, aumentando a ordem de aproximação para  $p+1$ .

Devemos observar dois fatos importantes:

1)  $\sigma^*$  é uma estimativa otimista do valor ótimo de (14).

De fato, da definição de função suporte,

$$f(x) \geq \xi_j(x), \forall x \in X \Rightarrow f(x) \geq \min_{\sigma} \left\{ \sigma / \sigma \geq \xi_j(x) \right\},$$

$j = 1, \dots, p \} = f_p(x)$ , ou seja  $f(x) \geq f_p(x)$ ,  $\forall x \in X$ .

2) O algoritmo não nos mostra como calcular as funções suporte  $\xi_j(\cdot)$ ,  $j = 1, \dots, p$ .

É claro que se  $f$  for convexa, então (8) pode ser utilizada, onde  $\mu(x_j)$  é qualquer subgradiente de  $f$  em  $x_j$  [Lasdon, 1970]. Se  $f(\cdot)$  não for convexa, mas puder ser colocada na forma (10), devemos estudar a possibilidade de tomarmos  $\xi_j(\cdot)$  como em (12).

#### 2.4.1 - Decomposição Benders Generalizada

como caso particular

Podemos agora retomar o problema (2) e evidenciar sua solução e interpretação geométrica como caso particular do procedimento apresentado acima.

Como  $v(y)$  em (3) é fácil de ser determinada para  $y = y_j$

fixado, então a solução de (1) pode ser obtida pela solução de (2). A proposta de Geoffrion é de resolvê-lo por relaxação, determinando funções suporte de  $v(y)$  e utilizando o algoritmo apresentado acima. Resta, portanto, evidenciarmos como calcular para  $v(y)$  as funções suporte  $\xi_j(y)$  em um ponto  $y = y_j \in Y \cap V$  qualquer.

TEOREMA 2.3 - A função  $v(y)$  definida em (3) admite em um ponto qualquer  $y_j \in Y \cap V$  a função suporte

$$\xi_j(y) = \min_{x \in X} \left\{ f(x, y) + \lambda_j' G(x, y) \right\}, \quad (16)$$

onde  $\lambda_j \in \mathbb{R}^r$  é a variável dual ótima do problema (SP) para  $y = y_j$ .

PROVA - Com as hipóteses de convexidade feitas sobre  $f$  e  $G$ , a teoria da dualidade permite escrever

$$v(y_j) = \max_{\lambda \geq 0} \min_{x \in X} \left\{ f(x, y_j) + \lambda' G(x, y_j) \right\}. \quad (17)$$

Como por hipótese  $\lambda_j$  é a solução dual ótima de (SP) para  $y = y_j$ , (17) implica imediatamente que  $v(y_j) = \xi_j(y_j)$ . Por outro lado, temos

$$\begin{aligned} v(y) &= \max_{\lambda \geq 0} \min_{x \in X} \left\{ f(x, y) + \lambda' G(x, y) \right\} \geq \\ &\geq \min_{x \in X} \left\{ f(x, y) + \lambda_j' G(x, y) \right\}, \quad \forall \lambda \in \mathbb{R}^r, \lambda \geq 0, \end{aligned} \quad (18)$$

e como  $\lambda_j \geq 0$ , (18) é válida em particular para  $\lambda = \lambda_j$ . Consequentemente, com (16) temos  $v(y) \geq \xi_j(y)$ ,  $\forall y \in \mathbb{R}^m$ , o que com (7) prova o teorema.

O algoritmo Benders Generalizado é então obtido imediatamente a partir do resultado do Teorema 2.3 e do Algoritmo 2.1. Iniciamos o processo com  $p = q = 0$ .

ALGORITMO 2.2

PASSO 1 - (Sub-problema) - Para  $\bar{y} \in Y \cap V$  resolva

$$\min_{x \in X} \left\{ f(x, \bar{y}) \mid G(x, \bar{y}) \leq 0 \right\} \quad (\text{SP})$$

PASSO 2 - Seja  $\lambda_*$  a solução dual ótima de (SP). Faça  $p = p+1$ ,  $\lambda_p = \lambda_*$  e obtenha a função suporte

$$\xi_p(y) = \min_{x \in X} \left\{ f(x, y) + \lambda_p' G(x, y) \right\}.$$

Calcule  $BS = \xi_p(\bar{y}) = v(\bar{y})$  (Teorema 2.3).  $BS$  é um limitante superior do valor ótimo de (2).

PASSO 3 - (Problema Mestre) - Obtenha a solução ótima de

$$\min_{y \in Y \cap V} \left\{ \sigma / \sigma \geq \xi_j(y), j = 1, \dots, p \right\}. \quad (\text{PM})$$

Seja  $y_*$  a solução de (PM) com valor objetivo  $\sigma^*$  e faça  $BI = \sigma^*$ .  $BI$  é um limitante inferior do valor ótimo de (2). Se  $BS - BI \leq \epsilon$ , para  $\epsilon$  positivo arbitrariamente pequeno, pare. A solução  $y_*$  é ótima.

PASSO 4 - Faça  $\bar{y} = y_*$  e resolva (SP). Deve ocorrer um dos seguintes fatos:

PASSO 4A - (SP) é factível.

Então vá para (2) e obtenha  $\overline{BS} = \xi_p(\bar{y})$ . Se  $\overline{BS} - BI \leq \epsilon$ , pare. A solução é ótima. Senão, se  $BS > \overline{BS}$ , faça  $BS = \overline{BS}$  e

vã para 5.

PASSO 4B - (SP) é infactível.

Então aproxime melhor  $V$  por  $V_p$  introduzindo neste último a restrição

$$\min_{x \in X} \theta'_p G(x, y) \leq 0$$

e vã para 3.

Neste ponto cabe uma rápida discussão sobre a convergência do procedimento apresentado acima. Ocorre que a grande importância deste método de decomposição deve-se exatamente ao fato de que, nas condições do problema (1), o procedimento acima tem convergência finita.

Dado que um dos objetivos deste trabalho é obter a solução de um problema real através deste método (ver Capítulos 3 e 4), optamos por apresentar apenas um dos teoremas de convergência, que trata de problemas nos quais o conjunto  $Y$  é um conjunto discreto. No entanto, a convergência do método também é finita, a menos de  $\epsilon > 0$ , para conjuntos  $Y$  de cardinalidade infinita como, por exemplo, conjuntos convexos [Geoffrion, 1972].

TEOREMA 2.4 - Nas hipóteses do problema (1) e do Teorema 2.2, suponhamos que o conjunto  $Y$  é um conjunto discreto finito. Então o procedimento dado no algoritmo 2.2 termina num número finito de passos para qualquer  $\epsilon \geq 0$  dado.

PROVA - Seja  $\epsilon \geq 0$  fixado arbitrariamente. A convergência finita é uma consequência direta da natureza finita de  $Y$  e do fa-

to de que  $\bar{y}$  não pode nunca ser repetido numa solução de (PM), no Passo 3. De fato, se  $\bar{y} \notin V$ , o Passo (4B) gera uma restrição para o problema mestre que impede que tal  $\bar{y}$  possa ser considerado novamente numa execução posterior de (PM). Se, entretanto,  $\bar{y} \in V$ , temos que o sub-problema é factível,  $v(y)$  é finito e a nova restrição  $\xi(\cdot)$  gerada no Passo (2) vai implicar na otimalidade de  $\bar{y}$  caso  $BS-BI \leq \epsilon$ , ou numa nova iteração, porém sem que  $\bar{Y}$  seja repetido. Isto prova o teorema.

#### 2.4.2 - Análise de Aplicabilidade da Decomposição

##### Benders Generalizada

O problema inicial (1) a ser resolvido foi decomposto em um sub-problema de fácil solução, que é convexo, dado que as variáveis  $y$  estão temporariamente fixadas em  $y = y_p$ , e num problema mestre, que é difícil, senão impossível de ser resolvido. Isto ocorre pois as funções

$$\xi_p(y) = \min_{x \in X} \left\{ f(x, y) + \lambda'_p G(x, y) \right\} \quad (19)$$

$$\eta_p(y) = \min_{x \in X} \theta'_p G(x, y)$$

podem ser tais que as minimizações indicadas dependam de  $y$ , sendo, portanto, impossíveis de serem calculadas explicitamente. Para contornar esta dificuldade Geoffrion [Geoffrion, 1972] propõe que (1) satisfaça a seguinte propriedade.

PROPRIEDADE P - Para todo  $\lambda \geq 0$  e  $\theta \in \Theta$ , as minimizações indicadas no cálculo de  $\xi_p(y)$  e  $\eta_p(y)$  são independentes de  $y$ .

Devemos fazer duas observações. A primeira é que a propriedade P elimina as dificuldades anteriormente citadas, permitindo que esta técnica de decomposição seja aplicável. No entanto, analisando (19), nos parece que a classe de problemas que admitem esta propriedade é muito reduzida. A segunda é que  $\xi_j(\cdot)$  ou  $\eta_j(\cdot)$  podem ser funções sem nenhuma propriedade, tal como convexidade, dificultando em muito a solução do problema mestre.

Na próxima secção vamos apresentar um método de otimização de  $v(y)$  baseado em gradiente projetado, sem no entanto ser necessária a verificação da propriedade P.

## 2.5 - UM ALGORITMO ALTERNATIVO - MÉTODO DE PROJEÇÃO

Seja novamente o problema (1), onde faremos as hipóteses iniciais de que  $f$  e  $G$  são continuamente diferenciáveis em  $x$  e  $y$ ,  $Y$  é um conjunto de restrições lineares e  $V = \mathbb{R}^m$ .

Propomos então obter a solução de (2) por um método primal de gradiente projetado no qual para  $y = y_p \in Y$  o valor da função  $v(y_p)$  é calculado pela solução de (SP) e

$$y_{p+1} = y_p - \tau_p d_p,$$

onde  $\tau_p > 0$  é um escalar arbitrariamente pequeno e  $d_p$  é a projeção ortogonal de  $\nabla v(y_p)$  sobre  $Y$ . É claro que tal procedimento só é válido se  $v(y)$  for uma função diferenciável para todo  $y \in Y$ .

### 2.5.1 - Estudo da Diferenciabilidade da Função

#### Perturbação Generalizada

A exemplo do que foi feito na secção 1.4 do capítulo anterior, vamos estabelecer as condições de diferenciabilidade de  $v(y)$ ,  $y \in Y$ , com o auxílio do Teorema de Danskin [Danskin, 1966]. Na discussão apresentada na referida secção obtivemos que dada

$$\phi(\lambda) = \min_{z \in S} F(z, \lambda),$$

sob hipóteses de diferenciabilidade da função  $F(z, \lambda)$ ,

$$\nabla(\min_{z \in S} F(z, \lambda)) = \frac{\partial F(z, \lambda)}{\partial \lambda} \Big|_{z = z^*}, \quad (20)$$

onde  $z^*$  é o único elemento de  $S$  que minimiza  $F(z, \lambda)$ .

Por um procedimento análogo pode-se mostrar que

$$\nabla(\max_{z \in S} F(z, \lambda)) = \frac{\partial F(z, \lambda)}{\partial \lambda} \Big|_{z = z^*} \quad (21)$$

onde  $z^*$  é o único elemento de  $S$  que maximiza  $F(z, \lambda)$ .

TEOREMA 2.5 - Seja  $y$  um elemento qualquer de  $Y$ . Então se o problema

$$\min_{x \in X} \left\{ f(x, y) \mid G(x, y) \leq 0 \right\} \quad (22)$$

admitir uma única solução primal  $x^*$  e uma única solução dual  $\lambda_* \geq 0$ ,  $v(y)$  é diferenciável e

$$\nabla v(y) = \frac{\partial f}{\partial y} + \frac{\partial G'}{\partial y} \lambda_* \Big|_{x = x^*} \quad (23)$$

PROVA - Como anteriormente, podemos escrever

$$\begin{aligned} v(y) &= \max_{\lambda \geq 0} \min_{x \in X} \left\{ f(x, y) + \lambda' G(x, y) \right\} \\ &= \max_{\lambda \geq 0} F(\lambda, y). \end{aligned} \tag{24}$$

Então, como  $V = \mathbb{R}^m$ , (22) é factível e  $\lambda_* < \infty$ , podemos adicionar em (24) a restrição  $\|\lambda\| \leq \rho$ , onde  $\rho$  é um escalar suficientemente grande, sem que a solução ótima seja alterada. Porém, o novo conjunto é compacto, o que nos permite escrever com (21)

$$\nabla v(y) = \nabla \left( \max_{\lambda \geq 0} F(\lambda, y) \right) \doteq \left. \frac{\partial F}{\partial y} \right|_{\lambda = \lambda_*} \tag{25}$$

Como  $X$  é compacto,  $x^* < \infty$  e então, com a mesma argumentação acima e (20) obtemos

$$\begin{aligned} \left. \frac{\partial F}{\partial y} \right|_{\lambda = \lambda_*} &= \nabla \left\{ \min_{x \in X} \left\{ f(x, y) + \lambda_*' G(x, y) \right\} \right\} \\ &= \left. \frac{\partial f}{\partial y} + \frac{\partial G'}{\partial y} \lambda_* \right|_{x = x^*} \end{aligned} \tag{26}$$

Finalmente, com (25) e (26) obtemos (23), o que prova o teorema.

### 2.5.2 - Análise de Aplicabilidade do Método de Projeção

Devemos notar que podemos impor condições ao problema em estudo para que as hipóteses do teorema anterior se verifiquem. A unicidade da solução primal de (22) é conseguida se este pro-

blema for estritamente convexo em relação a  $x$ . No entanto, isto não é suficiente para que a solução dual seja única, pois, esta depende do conjunto  $X$  para cada problema em questão. No caso em que  $X = \mathbb{R}^n$ ,  $x^*$  e  $\lambda_*$  satisfazem

$$\left. \begin{aligned} \frac{\partial f}{\partial x} + \frac{\partial G'}{\partial x} \lambda_* \end{aligned} \right|_{x = x^*} = 0 \quad (27)$$

e então, se posto  $\left( \frac{\partial G}{\partial x} \right) = r$ , que é uma condição de regularidade [Geoffrion, 1970], a unicidade de  $x^*$  implica na de  $\lambda_*$ .

Esta condição para a unicidade de  $\lambda_*$  generaliza o resultado obtido por H.P.L.Luna [Luna, 1979], onde são considerados problemas de programação matemática semelhantes ao problema (22). Naquele trabalho o autor estabelece uma condição de unicidade da solução dual para problemas nos quais a matriz jacobiana  $\frac{\partial G}{\partial x}$  é independente de  $x$  e inversível.

Por outro lado, mesmo sendo diferenciável em relação a  $y$ , a função  $v(y)$  geralmente não é convexa e, conseqüentemente, a função linear  $\xi(y) = v(y_p) + \nabla v(y_p)'(y - y_p)$  pode não ser uma função suporte de  $v(y)$  para certos pontos  $y_p \in Y$ , inviabilizando portanto a aplicação de técnicas de relaxação anteriormente estudadas para  $\xi(y)$  linear. Em [Geoffrion, 1970] é mostrado que se (1) é convexo conjuntamente nas variáveis  $x$  e  $y$ , então  $v(y)$  é convexa. No entanto, este resultado é de pouca valia no caso presente pois não fizemos nenhuma hipótese sobre as funções de (1) em relação as variáveis "complicantes"  $y$ .

## 2.6 - CONCLUSÃO

Neste capítulo analisamos as condições de aplicabilidade

da decomposição Benders Generalizada. Para tanto, o algoritmo inicialmente proposto por Geoffrion em [Geoffrion, 1972] foi obtido de uma nova maneira a partir da definição de *funções suporte* não lineares associadas a uma função dada. Este procedimento permite interpretar geometricamente as várias etapas de solução do problema mestre.

Em seguida propomos uma nova maneira de solução do problema mestre, que foi obtida pelo estudo da diferenciabilidade da função de perturbação generalizada  $v(y)$ , permitindo considerar problemas que não satisfaçam a propriedade P exigida em [Geoffrion, 1972].

## CAPÍTULO 3

### APLICAÇÃO DA DECOMPOSIÇÃO BENDERS GENERALIZADA

#### NO CONTROLE ÓTIMO DE CENTRAIS TÉRMICAS

##### 3.1 - INTRODUÇÃO

De posse dos resultados obtidos nos capítulos anteriores, vamos estudar o problema de controle ótimo de centrais térmicas de geração de energia. Tal problema difere essencialmente do problema de geração hidrelétrica, desde que envolve, além do custo operacional de geração efetiva de energia, um custo adicional, chamado custo de ligação. Este custo, conforme veremos, depende do intervalo de tempo durante o qual as térmicas estiverem desligadas (fora de operação).

Um modelo matemático que descreva este problema envolve dois tipos de variáveis de decisão: variáveis inteiras (de tipo 0-1), que indicam quais as unidades que estarão gerando energia num determinado instante  $t$ , e variáveis reais que traduzem a quantidade de energia a ser gerada por estas unidades neste mesmo instante  $t$ .

Trata-se, portanto, de um problema ao qual a técnica de decomposição Benders Generalizada é aplicável, pois é um problema de programação mista envolvendo variáveis de decisão inteiras e reais, onde as variáveis inteiras atuam como variáveis

veis "complicantes" e, como veremos, a propriedade P se verifica.

Nas próximas secções, vamos obter este modelo e a sua decomposição, de acordo com o estudo feito no Capítulo 3, em um problema mestre e um sub-problema. Também vamos explorar ao máximo as estruturas particulares destes dois problemas, visando obter métodos eficientes para suas soluções.

### 3.2 - DEFINIÇÃO DO PROBLEMA EM ESTUDO

Seja um sistema térmico constituído por N unidades geradoras, com as seguintes características de geração:

$$\left\{ \begin{array}{l} C_i(s): \text{ custo de geração de } s \text{ Kw-h em Cr\$} \\ \underline{s}_i \leq s \leq \bar{s}_i: \text{ capacidades máxima e mínima de} \\ \text{geração em Kw-h,} \end{array} \right. \quad (1)$$

para cada térmica  $i$ ,  $1 \leq i \leq N$ .

Se  $s_i^t$  é a quantidade de energia gerada pela térmica  $i$  durante o intervalo de tempo  $[t, t+1]$ , devemos satisfazer

$$\sum_{\substack{\text{têrmicas} \\ \text{em operação}}} s_i^t \geq L^t, \quad t \in [0, T-1], \quad (2)$$

onde  $L^t$  representa o déficit de produção no sistema hidrelétrico global, obtido pela diferença entre a demanda e a produção hidráulica.

Considerando um horizonte de T períodos e definindo a variável lógica

$$\mu_i^t = \begin{cases} 0, & \text{se a t\u00e9rmica } i \text{ n\u00e3o est\u00e1 em opera\u00e7\u00e3o no} \\ & \text{no instante } t; \\ 1, & \text{caso contr\u00e1rio,} \end{cases} \quad (3)$$

o problema de controle \u00f3timo das t\u00e9rmicas pode ser formulado como [Turgeon, 1978], [Baptistella & Geromel, 1980]

$$\begin{aligned} \min_{\mu \in \Lambda} \quad & \sum_{t=0}^{T-1} \sum_{i=1}^N C_i(s_i^t) \mu_i^t + \sum_{i=1}^N \Psi_i(z_i^t) \\ \text{s.a} \quad & \sum_{i=1}^N s_i^t \mu_i^t \geq L^t, \quad t = 0, 1, \dots, T-1 \\ & \underline{s}_i \leq s_i^t \leq \bar{s}_i, \end{aligned} \quad (4)$$

onde  $\Lambda$  representa um conjunto que define poss\u00edveis restri\u00e7\u00f5es operacionais sobre a entrada ou sa\u00edda de t\u00e9rmicas no sistema [Turgeon, 1978].

### 3.2.1 - A Fun\u00e7\u00e3o Custo de Liga\u00e7\u00e3o

Em [Turgeon, 1978], Turgeon, ao estudar este problema, apresenta uma solu\u00e7\u00e3o por ele obtida onde as fun\u00e7\u00f5es  $\Psi_i(\cdot)$ ,  $i=1, \dots, N$  s\u00e3o dadas como fun\u00e7\u00f5es temporais, ou seja, o custo de liga\u00e7\u00e3o de uma usina  $i$  depende essencialmente do tempo  $\Delta t$  durante o qual esta usina estava desligada (fora de opera\u00e7\u00e3o).

Fazendo  $z_i^t = \Delta t_i$ , a fun\u00e7\u00e3o custo de liga\u00e7\u00e3o, para cada t\u00e9rmica  $i$ , \u00e9 dada por

$$\Psi_i(z_i^t) = \alpha_i \left( 1 - \beta_i e^{-\gamma_i z_i^t} \right) \cdot \mu_i^t, \quad (5)$$

onde  $z_i^{t+1} = (1+z_i^t)(1-\mu_i^t)$ . A representa\u00e7\u00e3o gr\u00e1fica de (5) \u00e9 da-

dada abaixo.

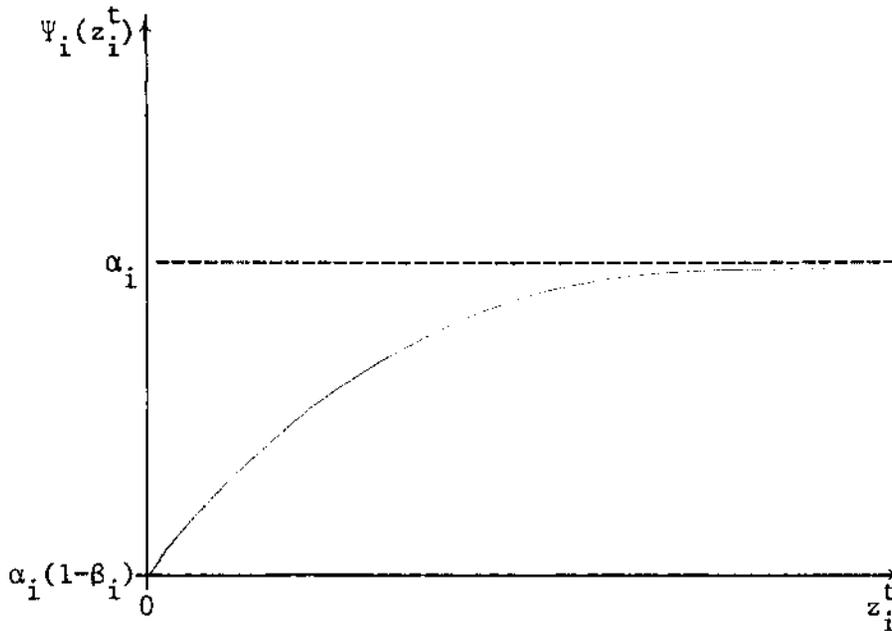


FIGURA 3.1 - Função custo de ligação

Ocorre que, se tomarmos  $\Psi_i(\cdot)$  como acima e aplicarmos a técnica de decomposição Benders Generalizada ao problema (4), o problema mestre resulta num problema de programação 0-1 não linear com  $2^{N \cdot T}$  combinações possíveis. Como a literatura atual não nos oferece um método que torna viável a obtenção da solução deste problema, optamos por resolver uma versão simplificada com as seguintes características:

- a) o custo  $\Psi_i(\cdot)$  de ligação da térmica  $i$  é tomado na situação limite, ou seja, tomando o custo máximo de ligação de cada térmica  $i$ , que ao  $z_i^t \rightarrow \infty$  resulta em  $\alpha_i$ , ou
- b) o custo  $\Psi_i(\cdot)$  de ligação da térmica  $i$  é o custo médio, dado por

$$\frac{1}{T} \int_0^T \Psi_i(\tau) d\tau.$$

Em qualquer dos casos a) ou b) acima, devemos observar que

- 1) o custo de ligação, para cada térmica  $i$ , resulta constante;
- 2) o número de combinações possíveis das variáveis inteiras cai para  $2^N$ .

De fato, redefinindo as variáveis  $\mu$  como sendo

$$\mu_i = \begin{cases} 0, & \text{se a térmica } i \text{ não está em operação } \forall t \in [0, T-1]; \\ 1, & \text{caso contrário,} \end{cases} \quad (6)$$

tomando  $\pi_i$  como sendo o custo em Cruzeiros da quantidade de combustível necessária para elevar a temperatura da térmica  $i$ , desligada no período anterior, até um valor necessário para a geração efetiva desta térmica, e indicando seu estado (ligada ou não) no período anterior por  $\bar{\mu}_i$ , temos a tabela

$\bar{\mu}_i$	$\mu_i$	$\psi_i(\mu_i)$
0	0	0
0	1	$\pi_i$
1	0	0
1	1	0

(7)

que fornece

$$\psi_i(\mu_i) = \pi_i(1 - \bar{\mu}_i)\mu_i, \quad (\text{Cr\$}) \quad (8)$$

onde  $\pi_i$  é definido segundo os casos a) ou b) anteriormente descritos.

Finalmente, definindo o conjunto  $S_i = \{s \mid s_i \leq s \leq \bar{s}_i\}$ , com (8)

podemos reescrever (4) na forma

$$\begin{aligned} \min \quad & \sum_{t=0}^{T-1} \sum_{i=1}^N C_i(s_i^t) \mu_i + \sum_{i=1}^N \psi_i(\mu_i) \\ \text{s.a.} \quad & \sum_{i=1}^N s_i^t \mu_i \geq L^t, \quad t = 0, 1, \dots, T-1 \\ & \mu \in \Lambda, \quad s_i^t \in S_i, \quad i = 1, \dots, N. \end{aligned} \quad (9)$$

Este problema não é convexo nas duas variáveis conjuntamente. Entretanto, a solução ótima global será obtida num número finito de passos, conforme veremos na próxima secção.

### 3.3 - APLICAÇÃO DA TÉCNICA DE DECOMPOSIÇÃO BENDERS

#### GENERALIZADA AO PROBLEMA DAS USINAS TÉRMICAS

Para resolvermos (9) pelo método Benders Generalizado necessitamos da hipótese de convexidade das funções  $C_i(s)$ ,  $1 \leq i \leq N$ , conforme vimos no Capítulo 2. Além disso, é importante frisar mais uma vez que o problema (9) é de programação mixta, com  $2^N$  combinações possíveis das variáveis inteiras e com  $N \cdot T$  variáveis contínuas.

Iniciamos então o processo de decomposição do problema (9), de acordo com o estudo desenvolvido no capítulo anterior, através da projeção deste problema no espaço das variáveis  $\mu$ . O problema abaixo, assim obtido, para  $\mu_i$ ,  $i=1, \dots, N$ , fixado e factível, é um problema de ótimo despacho de carga para as térmicas, para as quais  $\mu_i = 1$ , que, conforme veremos, é de fácil solução, além de ser equivalente ao problema (9), que desejamos resolver:

$$\min_{\mu \in \Lambda \cap V} \sum_{i=1}^N \Psi_i(\mu_i) + v(\mu), \quad (10)$$

onde  $v(\cdot)$ , para  $\mu = (\mu_1, \dots, \mu_N)'$  é dada por

$$v(\mu) = \min_{s_i^t \in S_i} \sum_{t=0}^{T-1} \sum_{i=1}^N C_i(s_i^t) \mu_i \quad (11)$$

$$\text{s.a. } \sum_{i=1}^N s_i^t \mu_i \geq L^t, \quad t = 0, 1, \dots, T-1$$

e  $V$  é o conjunto que deve garantir que, para todo  $\mu \in V$ , devemos ter ao menos uma solução factível para o problema (11) (ver Teorema 2.2, da representação do conjunto  $V$ ).

Felizmente, neste caso, é fácil ver que o conjunto  $V$  é dado por

$$V = \left\{ \mu_i \mid \sum_{i=1}^N \bar{s}_i \mu_i \geq L^t, \quad \forall t \in [0, T-1] \right\}. \quad (12)$$

No entanto, é sabido que, no horizonte de otimização utilizado, é grande a probabilidade que uma usina saia fora de operação devido a alguma contingência. Assim sendo, é necessário que as outras usinas em operação possam suprir a carga total do sistema.

Esta restrição adicional pode ser facilmente introduzida neste modelo, se redefinirmos o conjunto  $V$  em (12) como

$$V = \left\{ \mu_i \mid \sum_{i=1}^N \bar{s}_i \mu_i - \max_{1 \leq i \leq N} \{\bar{s}_i \mu_i\} \geq L^t, \quad t \in [0, T-1] \right\}. \quad (13)$$

Prosseguindo, devemos agora evidenciar as funções super-

te para a função definida em (11), para que possamos resolver o problema (10) por linearização externa.

Seja  $\lambda^t$  o preço da energia gerada num intervalo de tempo  $[t, t+1]$  em Cr\$/Kw-h. Então, conforme o Teorema 2.3, o problema (11) sendo convexo é equivalente a

$$v(\mu) = \max_{\lambda \geq 0} \min_{s_i^t \in S_i} \sum_{t=0}^{T-1} \sum_{i=1}^N \left\{ [C_i(s_i^t) - \lambda^t s_i^t] \mu_i + \lambda^t L^t \right\}, \quad (14)$$

onde  $\lambda = (\lambda^0, \lambda^1, \dots, \lambda^{T-1})$ .

Ainda sob as hipóteses do mesmo teorema, suponhamos que para um dado  $\mu_i^*$ ,  $i = 1, \dots, N$ , o problema max-min indicado em (14) foi resolvido, obtendo-se a solução ótima correspondente  $\lambda_*^t$ ,  $t = 0, 1, \dots, T-1$ , do preço da energia. Então,

$$\begin{aligned} v(\mu) &\geq \sum_{t=0}^{T-1} \left\{ \sum_{i=1}^N \min_{s_i^t \in S_i} [C_i(s_i^t) - \lambda_*^t s_i^t] \mu_i + \lambda_*^t L^t \right\} \\ &\geq \sum_{i=1}^N a_i^* \mu_i + b^*, \end{aligned} \quad (15)$$

onde  $a_i^*$  e  $b^*$  são dados por

$$\begin{aligned} a_i^* &= \sum_{t=0}^{T-1} \min_{s_i^t \in S_i} [C_i(s_i^t) - \lambda_*^t s_i^t], \quad i = 1, \dots, N \\ b^* &= \sum_{t=0}^{T-1} \lambda_*^t L^t. \end{aligned} \quad (16)$$

Obtivemos então, como desejávamos, uma linearização externa da função  $v(\mu)$  na qual as suas funções suporte  $\xi_j(\mu)$  são

dadas por

$$\xi_j(\mu) = \sum_{i=1}^N a_i^* \mu_i + b^*, \quad j = 1, 2, \dots, \quad (17)$$

com  $a_i^*$  e  $b^*$  definidos em (16).

Estes coeficientes,  $a_i^*$  e  $b^*$ , das funções suporte  $\xi_j(\cdot)$ ,  $j = 1, 2, \dots$ , são obviamente independentes de  $\mu$ , de modo que vale a propriedade P, discutida na secção 2.4.2 do capítulo anterior.

A partir destes resultados, podemos então resolver o problema (10) pelo seguinte procedimento computacional, que consiste exatamente da aplicação da técnica de relaxação (Algoritmo 2.1), neste problema específico.

### ALGORITMO 3.1

PASSO 1 - Seja  $\mu_i^*$ ,  $i = 1, \dots, N$  uma política de ligação factível ( $\mu^* \in \Lambda \cap V$ ) qualquer conhecida.

Forme o conjunto  $I = \{i / \mu_i^* = 1\}$  e resolva o problema de ótimo despacho:

$$\min \sum_{t=0}^{T-1} \sum_{i \in I} C_i(s_i^t) \quad (18)$$

$$\text{s.a. } \sum_{i \in I} s_i^t \geq L^t, \quad t = 0, \dots, T-1$$

$$s_i^t \in S_i, \quad i \in I.$$

Sejam  $\hat{s}_i^t$ ,  $i \in I$ ,  $t \in [0, T-1]$  e  $\lambda_*^t$ ,  $t \in [0, T-1]$ , as soluções primais e duais ótimas.

PASSO 2 - Calcule os coeficientes

$$a_i^* = \sum_{t=0}^{T-1} \min_{s_i^t \in S_i} \left\{ C_i(s_i^t) - \lambda_*^t s_i^t \right\}, \quad i=1, \dots, N$$
$$b^* = \sum_{t=0}^{T-1} \lambda_*^t L^t,$$
(19)

e obtenha o valor

$$BS = v(\mu^*) + \sum_{i=1}^N \psi_i(\mu^*) = \left\{ \sum_{i \in I} a_i^* + b^* \right\} + \sum_{i=1}^N \psi_i(\mu_i^*). \quad (20)$$

Faça  $p = 1$ ,  $a_i^p = a_i^*$ ,  $i = 1, \dots, N$ ,  $b^p = b^*$ .

BS é um limitante superior do valor ótimo de (9).

PASSO 3 - Determine a solução ótima do problema

$$\min_{\mu \in \Lambda \cap V} \left\{ \sum_{i=1}^N \psi_i(\mu_i) + \sigma(\mu) \right\},$$

onde (21)

$$\sigma(\mu) = \max_{1 \leq j \leq p} \left\{ \sum_{i=1}^N a_i^j \mu_i + b^j \right\}.$$

Seja  $\mu_i^*$ ,  $i = 1, \dots, N$  esta solução. Faça

$$BI = \sum_{i=1}^N \psi_i(\mu_i^*) + \sigma(\mu^*). \quad (22)$$

BI é um limitante inferior do valor ótimo de (9). Se  $BS - BI \leq \epsilon$  ( $\epsilon$  sendo um real positivo suficientemente pequeno), pare. A solução é ótima. Caso contrário,

redefina o conjunto I e vá ao passo seguinte.

PASSO 4 - Resolva o problema (18) e obtenha os parâmetros em (19) para o novo conjunto I. Se,

$$\overline{BS} (= \sum_{i \in I} a_i^* + b^* + \sum_{i=1}^N \Psi_i(u_i^*)) - B1 \leq \epsilon \text{ pare.}$$

A solução é ótima. Caso contrário, faça  $p = p+1$ ,  $a_i^p = a_i^*$ ,  $i = 1, \dots, N$ ,  $b^p = b^*$ . Se  $BS > \overline{BS}$  faça  $BS = \overline{BS}$  e volte ao Passo 3.

Este procedimento terá convergência finita, pois o número de variáveis lógicas factíveis é no máximo  $2^N$  e, portanto, finito (ver Teorema 2.4). Assim, ainda que tais vetores sejam gerados exhaustivamente, o processo acaba em no máximo  $2^N$  execuções do problema mestre. Conforme veremos no próximo capítulo, onde serão apresentadas algumas soluções para o problema que estamos estudando, todas as soluções ótimas ali apresentadas foram obtidas através de no máximo 5 iterações entre problema mestre e sub-problema. Isto é consequência do uso de funções suporte na solução do problema mestre. De fato, cada nova função suporte, juntamente com as anteriores, além de simplesmente determinar uma nova solução, mais próxima da solução ótima, do problema mestre, faz com que grande número de variáveis lógicas sejam excluídas definitivamente do espaço das combinações possíveis, pois cada função suporte determina um corte neste espaço. Assim, ao excluirmos uma variável  $\mu$ , estamos, na verdade, excluindo uma parte do conjunto destas variáveis.

Além disso, para qualquer fixação destas variáveis  $\mu$  pelo problema mestre (21), o problema (9) resulta em um problema convexo (18) a ser resolvido.

### 3.3.1 - O Método de Solução do Sub-problema

Suponhamos que o problema (21) foi resolvido para  $j = 1, \dots, p$ , obtendo-se como solução  $\mu_i^*$ ,  $i = 1, \dots, N$ . Então, a partir desta solução, devemos gerar uma nova linearização que, com as  $p$  anteriores, aproxime melhor o problema (10), que desejamos resolver.

Para tanto devemos calcular os coeficientes  $a_i^*$  e  $b^*$ . Conforme vimos acima, estes coeficientes são obtidos à partir da solução dual do problema (18).

Vejamos então, como podemos obter a solução deste problema (e conseqüentemente  $a_i^*$  e  $b^*$ ) através do método dual, para que possamos aproveitar melhor as particularidades inerentes ao problema em estudo.

Em primeiro lugar, devemos salientar que é lícito tentarmos resolver (18) através da obtenção do seu problema dual associado, pois trata-se de um problema convexo e a condição de regularidade de Slater é facilmente comprovada, ou seja, a equivalência primal/dual estabelecida no Capítulo 1 é verificada neste caso.

Temos então que  $\lambda_*$  é solução de

$$\max_{\lambda \geq 0} \phi(\lambda) \quad (23)$$

para  $\lambda = (\lambda^0, \lambda^1, \dots, \lambda^{T-1})'$ , onde  $\phi(\cdot)$  é a função dual associada ao problema de despacho ótimo de carga (18), que é definida por

$$\phi(\lambda) = \sum_{t=0}^{T-1} \phi(\lambda^t), \quad (24)$$

onde

$$\phi(\lambda^t) = \sum_{i=1}^N \min_{s_i^t \in S_i} [C_i(s_i^t) - \lambda^t s_i^t] \mu_i^* + \lambda^t L^t. \quad (25)$$

Observamos que, para  $\lambda^t$  fixado, o valor da função  $\phi(\cdot)$  é facilmente obtido, pois as soluções dos problemas indicados em (25) são

$$\hat{s}_i^t = \max\{0, \min\{\bar{s}_i, \bar{s}_i^t\}\}, \quad (26)$$

onde  $\bar{s}_i^t$  é a solução da equação

$$\left. \frac{dC_i(s)}{ds} \right|_{s=\bar{s}_i^t} = \lambda^t. \quad (27)$$

Suponhamos ainda que as funções  $C_i(s)$ ,  $i = 1, \dots, N$  sejam estritamente convexas e crescentes. Dado que os conjuntos  $S_i$ ,  $i = 1, \dots, N$  são compactos, a convexidade estrita das funções  $C_i(\cdot)$  resulta, conforme o Teorema 1.7, na existência do gradiente da função dual. Portanto, a função  $\phi(\cdot)$  é diferenciável, e seu gradiente pode ser obtido através da diferenciação das funções  $\phi(\cdot)$ :

$$\frac{\partial \phi}{\partial \lambda^t} = L^t - \sum_{i=1}^N \bar{s}_i^t \mu_i^*, \quad t=0,1,\dots,T-1, \quad (28)$$

que serão usados para a obtenção da solução ótima  $\lambda_*$  de (23).

Para obtermos o método de solução de (18), devemos ainda observar que para todo  $\mu \in V$ , se

$$\lambda^t \geq \max_{\mu_i=1} \left\{ \left. \frac{dC_i(s)}{ds} \right|_{s=\bar{s}_i} \right\} \quad (29)$$

então, em (26) temos  $\hat{s}_i^t = \bar{s}_i$ , para todo  $i$  tal que  $\mu_i = 1$ .

Analogamente, se

$$\lambda^t \leq \min_{\mu_i=1} \left\{ \left. \frac{dC_i(s)}{ds} \right|_{s=\underline{s}_i} \right\} \quad (30)$$

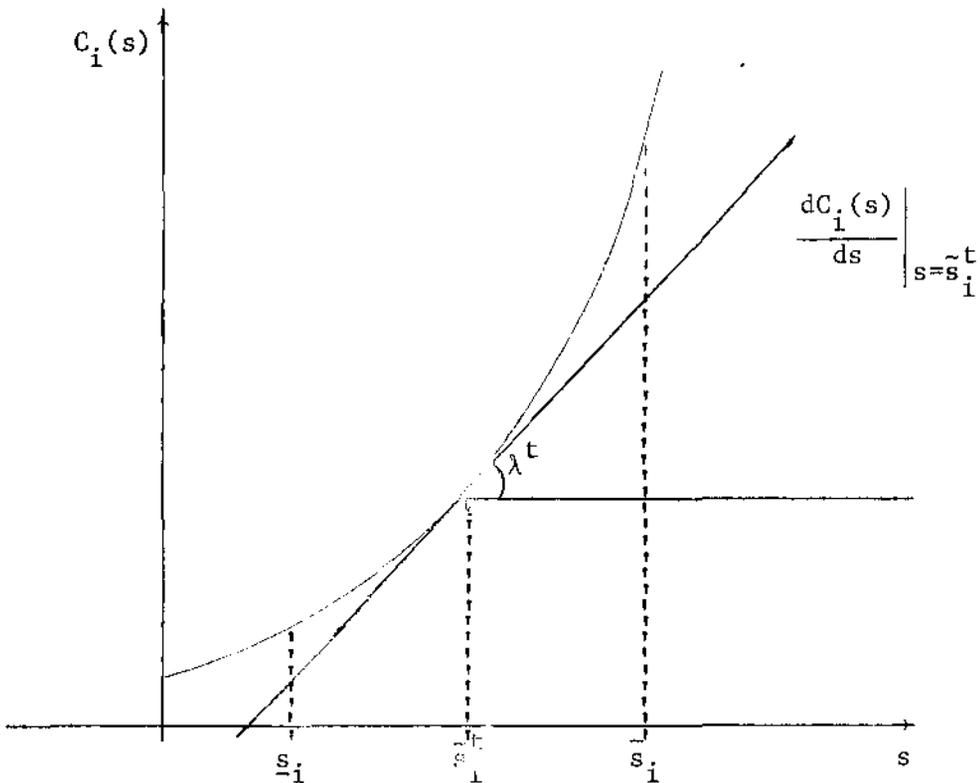


FIGURA 3.2 - Função custo de geração e sua derivada no ponto  $\tilde{s}_i^t$

então, em (26) temos  $\hat{s}_i^t = \underline{s}_i$  para todo  $i$  tal que  $\mu_i = 1$ .

A interpretação geométrica destes fatos pode ser vista na Figura 3.2.

Portanto, a solução do problema (23) é equivalente a solução de

$$\sum_{t=0}^{T-1} \max_{\underline{\lambda} \leq \lambda^t \leq \bar{\lambda}} \phi(\lambda^t), \quad (31)$$

onde

$$\underline{\lambda} = \min_{\mu_i=1} \left\{ \left. \frac{dC_i(s)}{ds} \right|_{s=\underline{s}_i} \right\} \quad (32)$$

$$\bar{\lambda} = \max_{\mu_i=1} \left\{ \left. \frac{dC_i(s)}{ds} \right|_{s=\bar{s}_i} \right\}$$

Finalmente, devemos observar que as funções  $\phi(\cdot)$  são côncavas (Teorema 1.1), diferenciáveis e definidas em conjuntos compactos. Visando aproveitar ao máximo estas informações, para através delas obter um método de solução do problema (18), vamos considerar uma aproximação quadrática para as funções  $\phi(\cdot)$ , ou seja,

$$\phi(\lambda) = a\lambda^2 + b\lambda + c. \quad (33)$$

Neste caso, a condição necessária e suficiente para que um ponto  $\lambda_*$  seja ponto de máximo é dado por

$$\left. \frac{d\phi}{d\lambda} \right|_{\lambda=\lambda_*} = 0, \quad (34)$$

ou seja,

$$\lambda_* = -\frac{b}{2a}. \quad (35)$$

Para  $\lambda_1, \lambda_2$  pertencentes ao domínio de  $\Phi(\cdot)$ , temos:

$$\left. \frac{d\Phi}{d\lambda} \right|_{\lambda=\lambda_1} = 2a\lambda_1 + b \quad (36)$$

$$\left. \frac{d\Phi}{d\lambda} \right|_{\lambda=\lambda_2} = 2a\lambda_2 + b \quad (37)$$

Subtraindo a equação (37) de (36), obtemos

$$\begin{aligned} \left. \frac{d\Phi}{d\lambda} \right|_{\lambda=\lambda_1} - \left. \frac{d\Phi}{d\lambda} \right|_{\lambda=\lambda_2} &= 2a(\lambda_1 - \lambda_2) \implies \\ \implies 2a &= \frac{\left. \frac{d\Phi}{d\lambda} \right|_{\lambda=\lambda_1} - \left. \frac{d\Phi}{d\lambda} \right|_{\lambda=\lambda_2}}{\lambda_1 - \lambda_2}. \end{aligned} \quad (38)$$

Além disso, multiplicando a equação (36) por  $\lambda_2$  e a equação (37) por  $\lambda_1$  e subtraindo as equações resultantes, obtemos

$$\begin{aligned} \lambda_2 \left. \frac{d\Phi}{d\lambda} \right|_{\lambda=\lambda_1} - \lambda_1 \left. \frac{d\Phi}{d\lambda} \right|_{\lambda=\lambda_2} &= (\lambda_2 - \lambda_1)b \implies \\ \implies b &= \frac{\lambda_1 \left. \frac{d\Phi}{d\lambda} \right|_{\lambda=\lambda_2} - \lambda_2 \left. \frac{d\Phi}{d\lambda} \right|_{\lambda=\lambda_1}}{\lambda_1 - \lambda_2}. \end{aligned} \quad (39)$$

Podemos então, substituindo (38) e (39) em (35) obter o valor  $\lambda_*$  que é dado por:

$$\lambda_* = \frac{\lambda_2 \left. \frac{d\phi}{d\lambda} \right|_{\lambda=\lambda_1} - \lambda_1 \left. \frac{d\phi}{d\lambda} \right|_{\lambda=\lambda_2}}{\left. \frac{d\phi}{d\lambda} \right|_{\lambda=\lambda_1} - \left. \frac{d\phi}{d\lambda} \right|_{\lambda=\lambda_2}}. \quad (40)$$

Ocorre que, nossas funções  $\phi(\cdot)$  não são quadráticas, mas são concavas e os  $\lambda^t$ ,  $t = 0, \dots, T-1$  devem ser tais que  $\underline{\lambda} \leq \lambda^t \leq \bar{\lambda}$ , com  $\underline{\lambda}$  e  $\bar{\lambda}$  como em (32). A diferenciabilidade de tais funções  $\phi(\cdot)$  aliadas a estas observações nos permite, a partir da expressão (40) obter um método de interpolação quadrática para aproximar  $\lambda_*$ , tomando  $\underline{\lambda}$  e  $\bar{\lambda}$  com pontos iniciais. Nossa opção por este método é devida a dois fatos básicos:

- 1) as informações sobre o comportamento das funções  $\phi(\cdot)$  (concavidade, existência de limites inferior e superior para  $\lambda^t$  e diferenciabilidade) são aproveitadas ao máximo;
- 2) o cálculo de  $\lambda_*$  envolve apenas a obtenção do gradiente das funções  $\phi(\cdot)$  e algumas operações elementares.

Podemos agora descrever em detalhes os passos do Algoritmo 3.1, referentes à obtenção da solução do sub-problema e dos coeficientes  $a_i^*$  e  $b^*$  das funções suporte  $\xi_j(\cdot)$  como segue. Inicialmente, seja  $\mu_i^*$ ,  $i = 1, \dots, N$  uma política de ligação factível ( $\mu_i^* \in \Lambda \cap V$ ) qualquer conhecida. Forme o conjunto

$$I = \{i \mid \mu_i^* = 1\},$$

faça  $t = 0$  e calcule  $\underline{\lambda}$  e  $\bar{\lambda}$  como em (32).

#### ALGORITMO 3.2

PASSO 1 - Faça  $t = t+1$ . Se  $t > T$  vá para 5. Senão faça  $k = 1$ ,

$\underline{\lambda}_k = \underline{\lambda}$ ,  $\bar{\lambda}_k = \bar{\lambda}$  e vá para 2.

PASSO 2 - Calcule

$$\bar{\lambda}_k = \frac{\bar{\lambda}_k \left. \frac{d\phi}{d\lambda} \right|_{\underline{\lambda}_k} - \underline{\lambda}_k \left. \frac{d\phi}{d\lambda} \right|_{\bar{\lambda}_k}}{\left. \frac{d\phi}{d\lambda} \right|_{\underline{\lambda}_k} - \left. \frac{d\phi}{d\lambda} \right|_{\bar{\lambda}_k}} \quad (41)$$

e vá para 3.

PASSO 3 - Se  $\left. \frac{\partial \phi}{\partial \lambda} \right|_{\tilde{\lambda}_k} = 0$  faça  $\lambda_*^t = \tilde{\lambda}_k$  e vá para 4. Senão seja

$$\tau_k = \begin{cases} 1, & \text{se } \left. \frac{\partial \phi}{\partial \lambda} \right|_{\tilde{\lambda}_k} > 0 \\ 0, & \text{se } \left. \frac{\partial \phi}{\partial \lambda} \right|_{\tilde{\lambda}_k} < 0 \end{cases} \quad (42)$$

e faça

$$\begin{cases} \underline{\lambda}_{k+1} = \tau_k \tilde{\lambda}_k + (1-\tau_k) \underline{\lambda}_k \\ \bar{\lambda}_{k+1} = (1-\tau_k) \tilde{\lambda}_k + \tau_k \bar{\lambda}_k \end{cases} \quad (43)$$

Faça  $k = k+1$  e vá para 2.

PASSO 4 - Para  $i \in I$ , calcule

$$\hat{s}_i^t = \max \left\{ 0, \min \{ \bar{s}_i, \tilde{s}_i^t \} \right\}, \quad (44)$$

onde  $\tilde{s}_i^t$  é dado por

$$\left. \frac{dC_i(s)}{ds} \right|_{s=\tilde{s}_i^t} = \lambda_*^t \quad (45)$$

e os coeficientes

$$\begin{cases} a_i^* = \sum_{t=0}^{T-1} \min_{s_i^t \in S_i} [C_i(s_i^t) - \lambda_*^t s_i^t] \\ b^* = \sum_{t=0}^{T-1} \lambda_*^t L^t \end{cases} \quad (46)$$

e  $v^*$  para 1.

PASSO 5 - Obtenha o valor de

$$BS = v(\mu^*) + \sum_{i=1}^N \psi_i(\mu_i^*) = \left\{ \sum_{i \in I} a_i^* + b^* \right\} + \sum_{i=1}^N \psi_i(\mu_i^*). \quad (47)$$

A convergência deste método é garantida, pois numa iteração genérica  $k$  temos  $\lambda_*^t \in [\underline{\lambda}_k, \bar{\lambda}_k] \subset [\underline{\lambda}_{k-1}, \bar{\lambda}_{k-1}]$ , sendo que  $\underline{\lambda}_k$  e  $\bar{\lambda}_k$  convergem quadraticamente para  $\lambda_*^t$  quando  $k \rightarrow \infty$ .

### 3.3.2 - O Método de Solução do Problema Mestre

Na secção 3.2.1 deste capítulo, optamos por simplificar nosso modelo tomando as funções custo de ligação  $\psi_i(\cdot)$ ,  $i = 1, \dots, N$  como independentes do tempo. Deste modo, além de obtermos funções custo de ligação constantes para  $t = 0, \dots, T-1$ , o número de soluções possíveis caiu de  $2^{N \cdot T}$  para  $2^N$ .

Nos exemplos reais resolvidos, que são apresentados no

próximo capítulo,  $N=10$ , ou seja, existem 1024 soluções possíveis para o problema mestre.

Verificamos então que podíamos neste caso obter a solução deste problema gerando os vetores  $\mu$  através de enumeração explícita. Isto se deve ao fato de que, como já frisamos anteriormente, a cada execução do problema mestre, os vetores  $\mu$  testados nas execuções anteriores não fazem mais parte do conjunto de soluções possíveis, ou seja, a cada iteração do problema mestre diminui o espaço de soluções viáveis.

O algoritmo que segue consiste da descrição detalhada dos passos do algoritmo 3.1 referentes à solução do problema mestre.

### ALGORITMO 3.3

PASSO 1 - Faça  $\mu_i = 1$ ,  $i = 1, \dots, N$ ,  $\sigma(\mu) = -\infty$ ,  $BI = +\infty$  e vá para 2.

PASSO 2 - Se  $\sum_{i=1}^N \bar{s}_i \mu_i \geq L^t$ ,  $\forall t = 0, \dots, T-1$ , vá para 5.  
Senão faça  $i = N$  e vá para 3.

PASSO 3 - Se  $\mu_i = 1$  faça:

1)  $\mu_i = 0$ ;

2) se  $\exists i$ ,  $i = 1, \dots, N$ , tal que  $\mu_i = 1$ , vá para 2.

Senão faça  $\mu^* = \bar{\mu}$  e volte ao programa principal.

Senão, vá para 4.

PASSO 4 - Faça  $\mu_i = 1$ ,  $i = i-1$  e vá para 3.

PASSO 5 - Para todo  $\mu_i$ ,  $i = 1, \dots, N$  tal que  $\mu_i = 1$ , calcule

$$S1 = \sum_i a_i^* \mu_i + b^*. \quad (48)$$

Se  $S1 > \sigma(\mu)$ , faça  $\sigma(\mu) = S1$  e calcule

$$S2 = \sum_i \Psi_i(\mu_i) + \sigma(\mu). \quad (49)$$

Se  $S2 < BI$ , faça  $BI = S2$ ,  $\bar{\mu} = \mu$ ,  $i = N$  e vá para 3.

Senão faça  $i = N$  e vá para 3.

Podemos agora mostrar a estrutura do algoritmo 3.1 através da Figura 3.3, onde procuramos evidenciar a interdependência dos cálculos referentes ao sub-problema e o problema mestre.

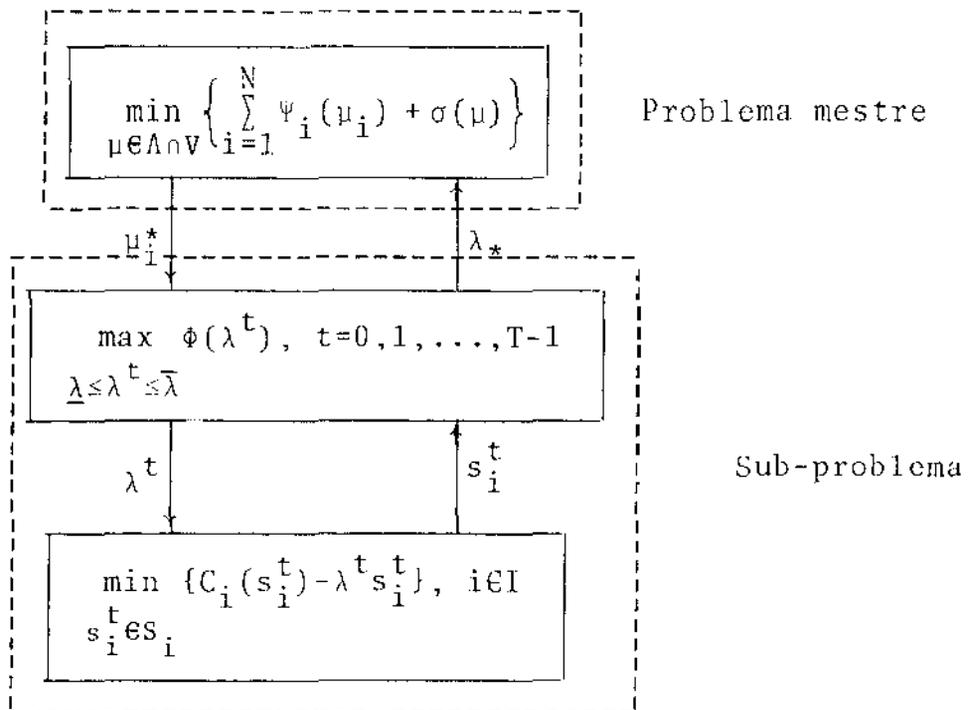


FIGURA 3.3 - Estrutura do Algoritmo 3.1

### 3.4 - CONCLUSÃO

Neste capítulo, a técnica de decomposição Benders Gene-

realizada foi aplicada ao problema de controle ótimo de usinas térmicas. Além disso, obtivemos as soluções detalhadas para o problema mestre e o sub-problema, aproveitando todas as características que são próprias ao modelo, visando maior eficiência computacional na obtenção de soluções de problemas reais, como os descritos no próximo capítulo.

## CAPÍTULO 4

### UMA APLICAÇÃO NUMÉRICA NUM SISTEMA TÉRMICO REAL

#### 4.1 - INTRODUÇÃO

O objetivo deste capítulo é apresentar alguns resultados obtidos computacionalmente para o problema de controle ótimo de usinas térmicas, apresentado no capítulo anterior.

Todos os estudos desenvolvidos naquele capítulo foram usados para a obtenção de tais resultados, dentre os quais discutiremos detalhadamente dois casos que ilustram claramente a grande eficiência computacional do método Benders Generalizado na obtenção de soluções para este problema.

Devemos salientar que esta eficiência deveu-se em grande parte ao método usado na obtenção da solução do sub-problema, ou seja, no aproveitamento total de todas as particularidades inerentes ao problema dual, evidenciadas e exploradas ao máximo na construção do Algoritmo 3.3.

#### 4.2 - DADOS RELATIVOS AO SISTEMA CONSIDERADO

As soluções que mostraremos nas próximas seções, para o problema de controle ótimo de centrais térmicas, foram obtidas através de um programa em linguagem Fortran. O computador usado em todas as execuções foi o PDP10 da Unicamp.

Tal programa foi obtido a partir dos Algoritmos 3.1, 3.2 e 3.3 apresentados no Capítulo 3. Assim, o *programa principal* consiste exatamente no método Benders Generalizado, programado em linguagem computacional e adaptado às características particulares do problema em questão. Os dados de entrada, bem como os resultados obtidos ao final da execução, também são fornecidos pelo programa principal.

As soluções do sub-problema e do problema mestre são obtidos em duas sub-rotinas, denominadas respectivamente de TERM e MESTRE, encontram-se listadas em Apêndice juntamente com o programa principal.

Todos os dados usados nos testes que seguem foram obtidos em [Turgeon, 1978]. Conforme ainda o Capítulo 3, a função de custo de geração, para cada térmica  $i$ ,  $i = 1, \dots, N$  é estritamente crescente e convexa, dada pela quadrática

$$C_i(s_i^t) = a_{i1} + a_{i2}s_i^t + a_{i3}(s_i^t)^2, \quad (1)$$

cujos coeficientes  $a_{ij}$  são alocados na matriz  $A(10 \times 3)$  da Tabela 4.1. O número de linhas desta matriz corresponde obviamente ao número de usinas a serem otimizadas em todos os testes, ou seja,  $N = 10$ .

O vetor  $Cl(i)$ ,  $i = 1, \dots, 10$ , da Tabela 4.1, contém os custos de ligação  $\pi_i$ , dados pela equação (8) do Capítulo 3, de cada térmica  $i$ . Estes custos podem ser alocados deste modo desde que os estamos tomando constantes em relação ao tempo, conforme ficou estabelecido na secção 3.2.1 do referido capí-

tulo.

TABELA 4.1 - Dados de entrada do sistema

Unidade	SINF	SSUP	A(1,1)	A(1,2)	A(1,3)	CL(1)
1	15,0	60,0	15,0	1,4000	0,00510	85,000
2	20,0	80,0	25,0	1,5000	0,00396	101,000
3	30,0	100,0	40,0	1,3500	0,00393	114,000
4	25,0	120,0	32,0	1,4000	0,00382	94,000
5	50,0	150,0	29,0	1,5400	0,00212	113,000
6	75,0	280,0	72,0	1,3500	0,00261	176,000
7	250,0	520,0	105,0	1,3954	0,00127	267,000
8	50,0	150,0	100,0	1,3285	0,00135	282,000
9	120,0	320,0	49,0	1,2643	0,00289	187,000
10	75,0	200,0	82,0	1,2136	0,00148	227,000

No caso da Tabela 4.1, para cada t ermica  $i$ ,  $i = 1, \dots, 10$  o custo de liga ao   dado por

$$\pi_i = \alpha_i. \quad (2)$$

No entanto, foram ainda efetuados testes nos quais, para cada t ermica  $i$ ,  $i = 1, \dots, 10$ .

$$\pi_i = \frac{1}{T} \left\{ \sum_{t=0}^{T-1} \alpha_i (1 - \beta_i e^{-\gamma_i t}) \right\}, \quad (3)$$

ou seja, o custo de liga ao da t ermica  $i$ ,  $i = 1, \dots, 10$ , para todos os instantes de tempo  $t$ ,  $t \in [0, T-1]$    sempre igual ao seu custo m ximo (express o (2)), ou sempre igual ao seu custo m dio (express o (3)).

A mesma tabela tamb m fornece os limites  $\underline{s}_i$  e  $\bar{s}_i$ ,  $i=1, \dots, 10$ , ou seja, as capacidades m nimas e m ximas de gera o em Kw-h de cada t ermica  $i$ , alocados respectivamente nos vetores

SINF(i) e SSUP(i).

Conforme já salientamos acima, em todos os testes realizados, o número de térmicas é de  $N=10$ , sendo que o tempo é discretizado em  $T=24$  períodos (operação diária com discretização horária).

#### 4.3 - AS SOLUÇÕES OBTIDAS

As soluções, que apresentaremos agora em detalhes, foram obtidas tomando como custo de ligação para cada térmica  $i$ ,  $i=1, \dots, 10$ ,

$$\pi_i = \alpha_i, \quad (4)$$

e o conjunto  $V$ , que garante que para  $\mu \in V$  o sub-problema tem solução factível, é dado por

$$V_1 = \left\{ \mu_i \mid \sum_{i=1}^{10} \bar{s}_i \mu_i \geq L^t, \quad t=0, \dots, T-1 \right\}. \quad (5)$$

##### 4.3.1 - CASO 1

Neste caso tomamos  $T=24$ , e com os dados de entrada fornecidos pela Tabela 4.1, obtivemos as soluções que constam da Tabela 4.3, com tempo total de CPU de 3,49s.

TABELA 4.2 - Solução térmica ótima - Caso 1

T	DEM	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
1	1459,0	60,0	80,0	100,0	120,0	150,0	0,0	504,6	0,0	244,4	200,0
2	1372,0	60,0	80,0	100,0	120,0	150,0	0,0	444,1	0,0	217,9	200,0
3	1299,0	60,0	80,0	100,0	120,0	150,0	0,0	393,4	0,0	195,6	200,0
4	1285,0	60,0	80,0	100,0	120,0	150,0	0,0	383,7	0,0	191,3	200,0
5	1271,0	60,0	80,0	100,0	120,0	150,0	0,0	374,0	0,0	187,0	200,0
6	1314,0	60,0	80,0	100,0	120,0	150,0	0,0	403,9	0,0	200,2	200,0
7	1372,0	60,0	80,0	100,0	120,0	150,0	0,0	444,1	0,0	217,9	200,0
8	1314,0	60,0	80,0	100,0	120,0	150,0	0,0	403,9	0,0	200,2	200,0
9	1271,0	60,0	80,0	100,0	120,0	150,0	0,0	374,0	0,0	187,0	200,0
10	1242,0	60,0	80,0	100,0	117,6	150,0	0,0	355,5	0,0	178,9	200,0
11	1197,0	60,0	80,0	100,0	109,1	150,0	0,0	330,1	0,0	167,7	200,0
12	1182,0	60,0	80,0	100,0	106,3	150,0	0,0	321,6	0,0	164,0	200,0
13	1154,0	60,0	80,0	100,0	101,2	149,3	0,0	306,2	0,0	157,2	200,0
14	1138,0	60,0	80,0	100,0	99,0	145,3	0,0	299,5	0,0	154,3	200,0
15	1124,0	60,0	80,0	100,0	97,0	141,8	0,0	293,6	0,0	151,7	200,0
16	1095,0	60,0	77,7	97,4	93,6	135,7	0,0	283,4	0,0	147,2	200,0
17	1066,0	60,0	74,6	94,3	90,4	129,9	0,0	273,8	0,0	143,0	200,0
18	1037,0	60,0	71,5	91,1	87,2	124,2	0,0	264,2	0,0	138,8	200,0
19	993,0	60,0	66,8	86,4	82,3	115,3	0,0	250,0	0,0	132,3	200,0
20	978,0	60,0	64,4	84,0	79,9	110,9	0,0	250,0	0,0	129,0	200,0
21	963,0	60,0	62,3	81,8	77,7	106,9	0,0	250,0	0,0	126,1	200,0
22	1022,0	60,0	69,9	89,5	85,6	121,2	0,0	259,2	0,0	136,6	200,0
23	1081,0	60,0	76,2	95,9	92,1	132,9	0,0	278,8	0,0	145,2	200,0
24	1459,0	60,0	80,0	100,0	120,0	150,0	0,0	504,6	0,0	244,4	200,0

TABELA 4.3 - Solução ótima do problema Mestre - Caso 1

1	1	1	1	1	1	0	1	0	1	1
2	1	1	1	1	1	0	1	0	1	1
3	1	1	1	1	1	0	1	0	1	1
4	1	1	1	1	1	0	1	0	1	1
5	1	1	1	1	1	0	1	0	1	1
6	1	1	1	1	1	0	1	0	1	1
7	1	1	1	1	1	0	1	0	1	1
8	1	1	1	1	1	0	1	0	1	1
9	1	1	1	1	1	0	1	0	1	1
10	1	1	1	1	1	0	1	0	1	1
11	1	1	1	1	1	0	1	0	1	1
12	1	1	1	1	1	0	1	0	1	1
13	1	1	1	1	1	0	1	0	1	1
14	1	1	1	1	1	0	1	0	1	1
15	1	1	1	1	1	0	1	0	1	1
16	1	1	1	1	1	0	1	0	1	1
17	1	1	1	1	1	0	1	0	1	1
18	1	1	1	1	1	0	1	0	1	1
19	1	1	1	1	1	0	1	0	1	1
20	1	1	1	1	1	0	1	0	1	1
21	1	1	1	1	1	0	1	0	1	1
22	1	1	1	1	1	0	1	0	1	1
23	1	1	1	1	1	0	1	0	1	1
24	1	1	1	1	1	0	1	0	1	1

Estas soluções foram obtidas após 4 iterações entre problema mestre e sub-problema. Os valores obtidos a cada iteração para  $BS = LIM.SUP.$  e  $BI = LIM.INF.$ , até a convergência, foram os seguintes:

TABELA 4.4 - Convergência do Método - Caso 1  
Número de Iterações Mestre/Sub-problema

Iterações	Lim. Sup.	Lim. Inf.
1	61999,12	59373,67
2	61018,75	60328,72
3	60914,31	60655,36
4	60736,97	60736,97

Estes resultados podem ser interpretados geometricamente no gráfico abaixo.

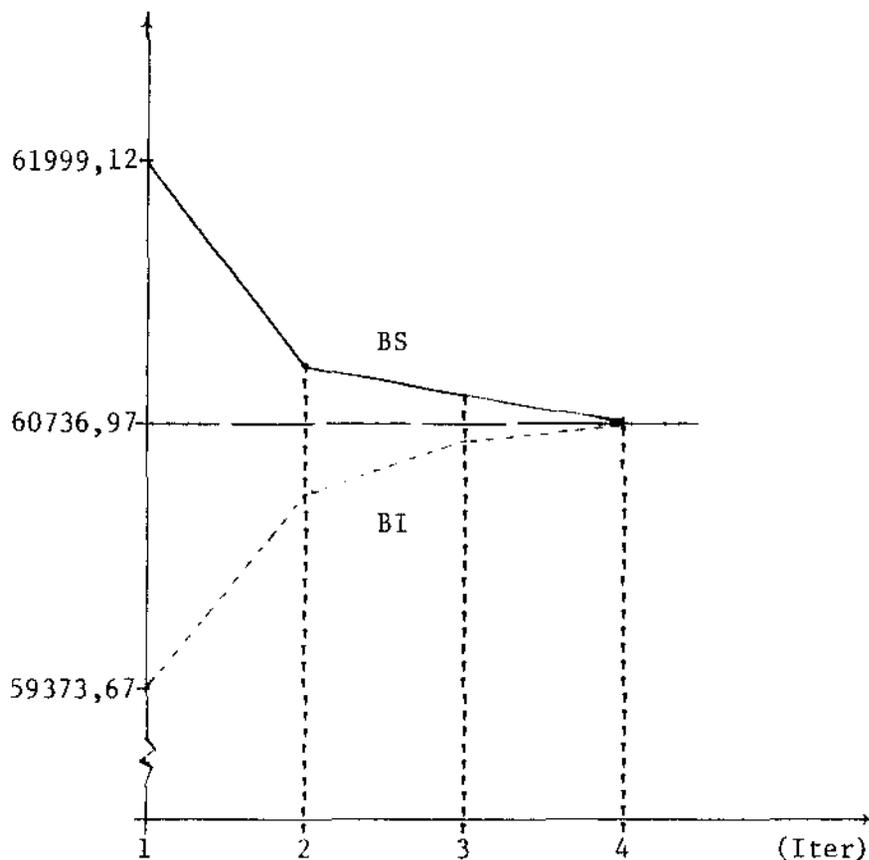


FIGURA 4.1 - Convergência do Método - Caso 1

#### 4.3.2 - CASO 2

Neste exemplo o tempo também é discretizado em  $T = 24$  períodos, as funções custo de ligação e o conjunto  $V$  são os mesmos do caso anterior. Porém, desta vez resolvemos o problema em duas etapas. Primeiramente foi resolvido o problema de controle ótimo para os 12 primeiros períodos de tempo, sempre com os dados da Tabela 4.1. Verificou-se ao final da execução que apenas uma das 10 usinas permaneceu desligada nestes 12 períodos de tempo iniciais, conforme pode ser visto nas tabelas abaixo.

TABELA 4.5 - Solução térmica ótima - Caso 2 - 1ª parte

T	DEM	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
1	1459,0	60,0	80,0	100,0	120,0	150,0	190,0	372,6	0,0	186,4	200,0
2	1372,0	60,0	80,0	100,0	110,0	150,0	170,5	332,6	0,0	168,9	200,0
3	1299,0	60,0	80,0	100,0	99,7	146,7	155,5	301,8	0,0	155,3	200,0
4	1285,0	60,0	80,0	100,0	98,1	143,7	153,2	296,9	0,0	153,1	200,0
5	1271,0	60,0	80,0	100,0	96,5	140,8	150,8	292,0	0,0	151,0	200,0
6	1314,0	60,0	80,0	100,0	101,5	149,8	158,1	307,0	0,0	157,6	200,0
7	1372,0	60,0	80,0	100,0	110,0	150,0	170,5	332,6	0,0	168,9	200,0
8	1314,0	60,0	80,0	100,0	101,5	149,8	158,1	307,0	0,0	157,6	200,0
9	1271,0	60,0	80,0	100,0	96,5	140,8	150,8	292,0	0,0	151,0	200,0
10	1242,0	60,0	77,7	97,4	93,7	135,7	146,7	283,5	0,0	147,3	200,0
11	1197,0	60,0	73,6	93,2	89,4	128,0	140,4	270,7	0,0	141,6	200,0
12	1182,0	60,0	72,2	91,9	88,0	125,5	138,3	266,4	0,0	139,7	200,0

O tempo de CPU baixou neste caso para 3,14s.

TABELA 4.6 - Solução ótima do problema Mestre - Caso 2 - 1ª parte

1	1	1	1	1	1	1	1	0	1	1
2	1	1	1	1	1	1	1	0	1	1
3	1	1	1	1	1	1	1	0	1	1
4	1	1	1	1	1	1	1	0	1	1
5	1	1	1	1	1	1	1	0	1	1
6	1	1	1	1	1	1	1	0	1	1
7	1	1	1	1	1	1	1	0	1	1
8	1	1	1	1	1	1	1	0	1	1
9	1	1	1	1	1	1	1	0	1	1
10	1	1	1	1	1	1	1	0	1	1
11	1	1	1	1	1	1	1	0	1	1
12	1	1	1	1	1	1	1	0	1	1

As soluções acima porém demandaram 5 iterações entre problema mestre e sub-problema e os valores obtidos para BS e BI até a convergencia foram os seguintes:

TABELA 4.7 - Convergência do Método - Caso 2 - 1ª parte

Número de iterações Mestre/Sub-problema

Iterações	Lim. Sup.	Lim. Sup.
1	34196,65	32981,19
2	33920,95	33414,05
3	33726,62	33633,17
4	33692,64	33683,85
5	33692,64	33692,65

Na Figura 4.2, mostramos a interpretação geométrica deste resultado.

Para obtermos a solução ótima para os 12 instantes de tempo restantes, partimos da observação de que, no resultado ob-

tido acima, apenas uma das 10 térmicas permaneceu desligada. Portanto, para  $t=13$ , e nos instantes seguintes, até  $t=24$  não precisamos considerar os custos de ligação das demais térmicas, desde que já sabemos que elas estarão ligadas, de modo que seus custos de ligação  $\pi_j$ , são todos nulos (ver equação 8 do Capítulo 3). O resultado que segue foi obtido a partir dos dados da Tabela 4.1 na qual foi feita a seguinte alteração:

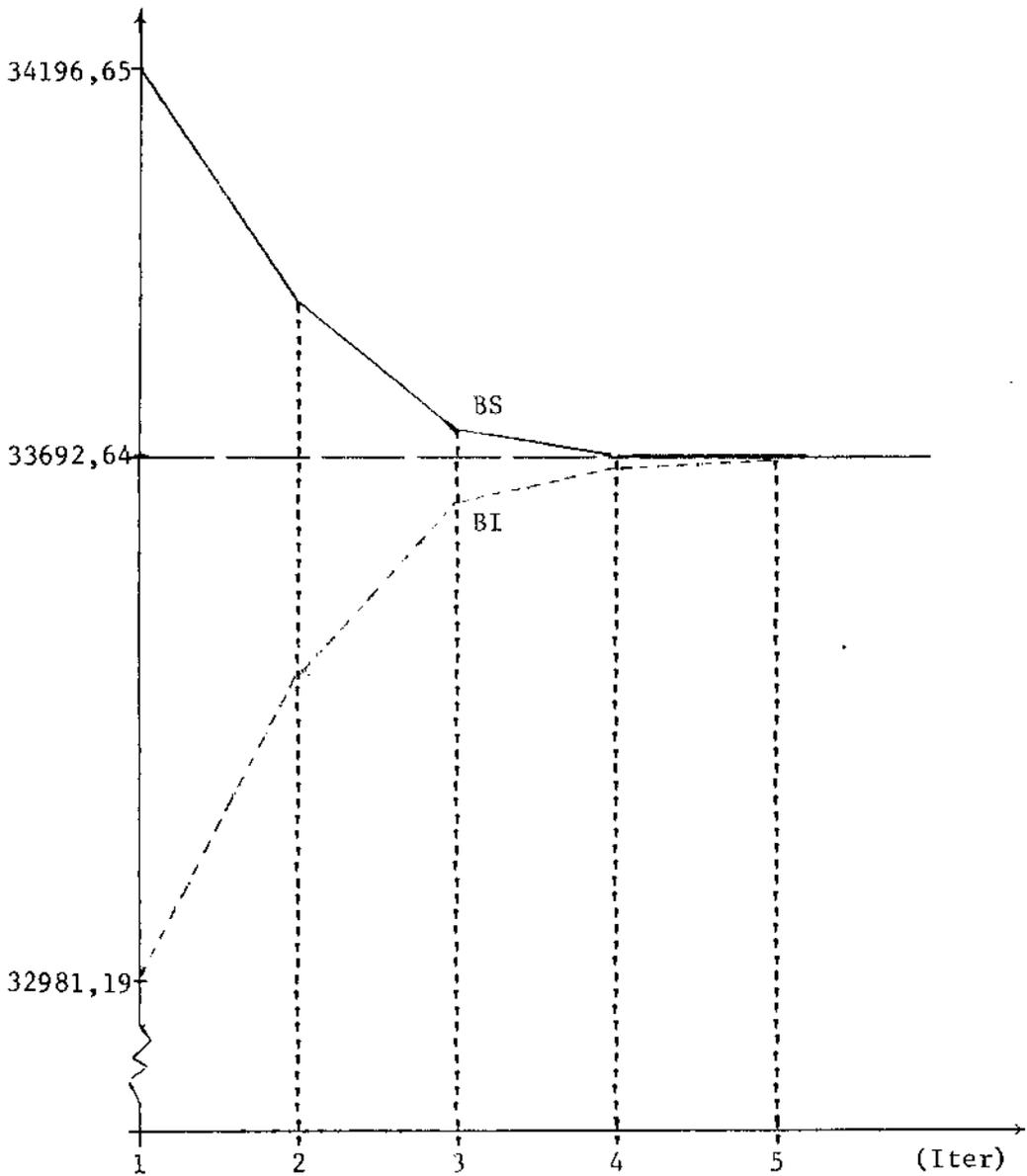


FIGURA 4.2 - Convergência do Método - Caso 2 - 1ª parte

TABELA 4.8 - Solução térmica ótima - Caso 2 - 2ª parte

T	DEM	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
13	1154,0	60,0	80,0	100,0	101,2	149,3	0,0	306,2	0,0	157,2	200,0
14	1138,0	60,0	80,0	100,0	99,0	145,3	0,0	299,5	0,0	154,3	200,0
15	1124,0	60,0	80,0	100,0	97,0	141,8	0,0	293,6	0,0	151,7	200,0
16	1095,0	60,0	77,7	97,4	93,6	135,7	0,0	283,4	0,0	147,2	200,0
17	1066,0	60,0	74,6	94,3	90,4	129,9	0,0	273,8	0,0	143,0	200,0
18	1037,0	60,0	71,5	91,1	87,2	124,2	0,0	264,2	0,0	138,8	200,0
19	993,0	60,0	66,8	86,4	82,3	115,3	0,0	250,0	0,0	132,3	200,0
20	978,0	59,8	64,4	84,0	79,9	110,9	0,0	250,0	0,0	129,0	200,0
21	963,0	58,2	62,3	81,8	77,7	106,9	0,0	250,0	0,0	126,1	200,0
22	1022,0	60,0	69,9	89,5	85,6	121,2	0,0	259,2	0,0	136,6	200,0
23	1081,0	60,0	76,2	95,9	92,1	132,9	0,0	278,8	0,0	145,2	200,0
24	1459,0	60,0	80,0	100,0	120,0	150,0	0,0	504,6	0,0	244,4	200,0

$$\begin{cases} \pi_i = 0, & i \neq 8 \\ \pi_i = 282, & i = 8 \end{cases} \quad (6)$$

Deste modo, aproveitamos as informações obtidas na primeira etapa do procedimento e o resultado que segue pode ser obtido com apenas 2,81s de CPU, (Tabelas 4.8, 4.9 e 4.10).

TABELA 4.9 - Solução ótima do problema Mestre - Caso 2 - 2ª parte

13	1	1	1	1	1	0	1	0	1	1
14	1	1	1	1	1	0	1	0	1	1
15	1	1	1	1	1	0	1	0	1	1
16	1	1	1	1	1	0	1	0	1	1
17	1	1	1	1	1	0	1	0	1	1
18	1	1	1	1	1	0	1	0	1	1
19	1	1	1	1	1	0	1	0	1	1
20	1	1	1	1	1	0	1	0	1	1
21	1	1	1	1	1	0	1	0	1	1
22	1	1	1	1	1	0	1	0	1	1
23	1	1	1	1	1	0	1	0	1	1
24	1	1	1	1	1	0	1	0	1	1

Estas soluções foram obtidas após 4 iterações entre problema mestre e sub-problema. Neste caso, os valores obtidos para BS e BI são:

TABELA 4.10 - Convergência do Método - Caso 2 - 2ª parte

Número de Iterações Mestre/Sub-problema

Iterações	Lim. Sup.	Lim. Inf.
1	28084,47	26392,48
2	27097,80	26982,16
3	27097,80	27008,04
4	27044,32	27044,32

Geometricamente, obtivemos o seguinte desempenho:

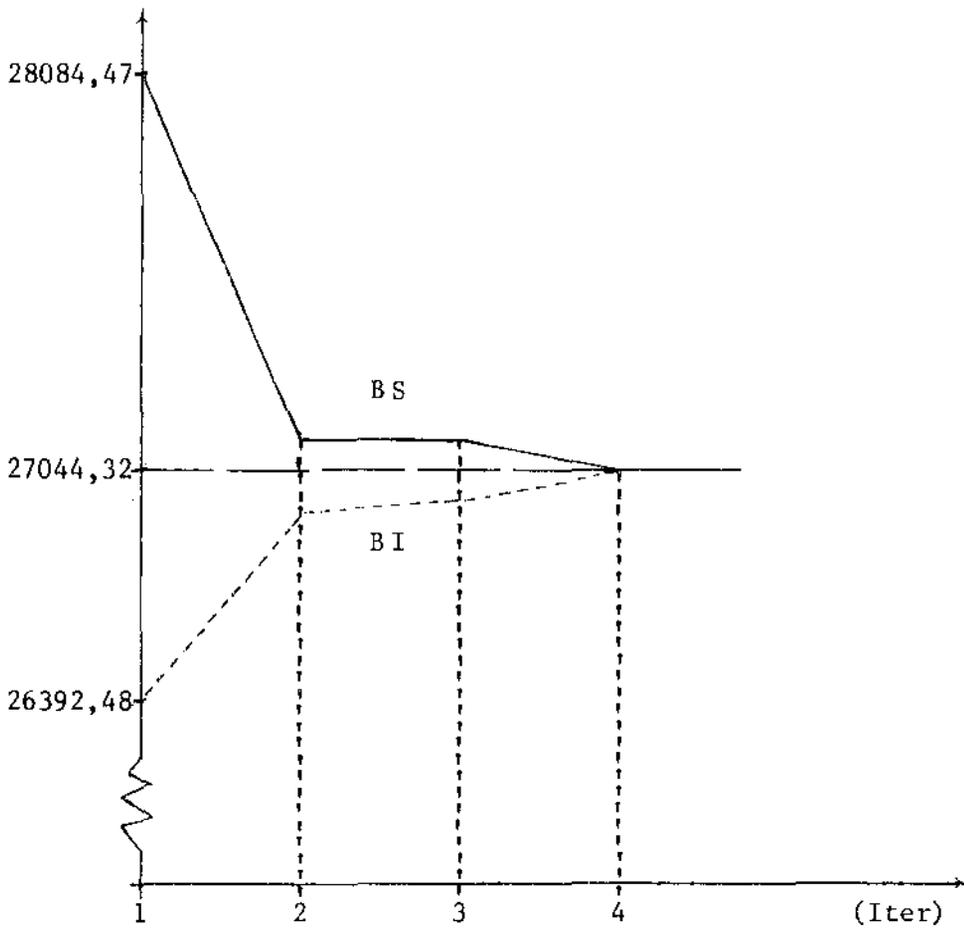


FIGURA 4.3 - Convergência do Método - Caso 2 - 2ª parte

#### 4.3.3 - Outros Testes Realizados

Nos dois casos apresentados acima consideramos os custos de ligação  $\pi_i = \alpha_i$ ,  $i = 1, \dots, 10$ , ou seja, o custo máximo de ligação de cada tērmica  $i$ , e o conjunto  $V = V_1$  dado pela expressão (5).

No entanto, foram realizados outros testes, nos moldes dos casos anteriores, alterando ora a função custo de ligação, ora o conjunto  $V$ , ou ambos. Assim, ao invés de tomarmos os custos máximos de ligação, fizemos alguns testes considerando os custos médios de ligação, dados pela expressão (3). Igualmente o

conjunto V foi alterado em alguns testes para

$$V_2 = \left\{ \mu_i \mid \sum_{i=1}^{10} \bar{s}_i \mu_i - \max_{1 \leq i \leq 10} \{ \bar{s}_i \mu_i \} \geq L^t, t \in [0, T-1] \right\}. \quad (7)$$

A restrição adicional introduzida no conjunto V, dado por  $V_2$ , conforme vimos no Capítulo 3, é devida a probabilidade que uma usina saia fora de operação por alguma contingência, no horizonte de otimização utilizado. A introdução desta restrição, evidentemente, onera o custo de geração de energia, desde que deste modo estamos operando com folga na demanda. Entretanto, esta parece ser a melhor escolha para contornarmos o problema dentre outras alternativas possíveis (ver Turgeon, 1978).

TABELA 4.11 - Performance do "Benders" nos casos resolvidos

Teste	Tempo (Períodos)	V	Custo de Ligação	Iterações Mestre/ Sub-problema	CPU Total (Seg) PDP/10
1	24	$V_1$	máximo	4	3,49
2	12	$V_1$	máximo	5	3,14
3	12	$V_1$	máximo	4	2,81
4	24	$V_2$	máximo	1	1,63
5	12	$V_2$	máximo	1	1,14
6	12	$V_2$	zero	1	1,09
7	24	$V_1$	médio	4	3,35
8	12	$V_1$	médio	5	3,23
9	12	$V_1$	médio	4	2,69
10	24	$V_2$	médio	1	1,53
11	12	$V_2$	médio	1	1,20
12	12	$V_2$	zero	1	1,34

Fazendo todas as combinações possíveis entre  $V_1$ ,  $V_2$  e as

expressões (2) e (3) para os custos de ligação obtivemos os resultados que estão resumidos na Tabela 4.11 (que inclui os casos descritos nas sub-seções anteriores), na qual evidenciamos o número de iterações entre problema mestre e sub-problema até a obtenção da convergência e o tempo total em segundos de CPU gasto em cada teste.

É importante observar, quanto a Tabela 4.11, que os testes 2 e 3 correspondem ao caso 2 descrito anteriormente e que os testes 5 e 6, 8 e 9, 11 e 12 foram feitos de maneira análoga.

Devemos ainda observar que, conforme o Algoritmo 3.1, o programa principal deve ser inicializado com uma política de ligação factível. Optamos então por inicializar o vetor  $\mu$  tomando, em todos os testes,  $\mu_i = 1, i = 1, \dots, N$ .

Ocorre que, tanto no teste 5 como no teste 11, tal solução se revelou ótima na primeira iteração (ver Tabela 4.11). Tais testes, conforme salientamos acima, correspondem a primeira parte de testes análogos ao Caso 2 descrito em 4.3.2. Assim, em ambos os casos, para obtermos as soluções correspondentes aos 12 intervalos de tempo restantes, (problemas 6 e 12) os custos de ligação foram zerados, dado que tínhamos a informação de que todas as térmicas estavam ligadas nos 12 intervalos de tempo anteriores (equação 8, Capítulo 3).

#### 4.4 - CONCLUSÃO

Neste capítulo mostramos, através de uma série de testes

computacionais, o notável desempenho do método Benders Generalizado na obtenção de soluções de problemas que, além de admitirem número muito grande de combinações possíveis das variáveis, possuem centenas de restrições operacionais.

Este desempenho é extremamente importante diante do fato de que as soluções do problema mestre foram obtidas a partir de simples comparação entre o valor da função objetivo em cada ponto factível, com o melhor valor obtido até então. É claro que tal procedimento foi possível de ser implementado pois o número de variáveis inteiras consideradas foi relativamente pequeno ( $N = 10$ ).

Na Tabela 4.12, calculamos o desempenho médio do algoritmo relativo ao problema em estudo.

TABELA 4.12 - Desempenho médio do Método de Benders

Número de variáveis inteiras (Problema Mestre) . . .	10
Número de variáveis contínuas (Sub-problema) . . .	240
Número de restrições . . . . .	528
Número médio de iterações Mestre/Sub-problema . . .	$2 \frac{2}{3}$
Tempo total médio de CPU . . . . .	2,22s

Os tempos totais de CPU verificados em cada teste, que resultaram na média obtida na Tabela 4.12, só foram possíveis devido ao método utilizado para a solução do sub-problema. Isto fica mais evidente se compararmos estes resultados com uma experiência anterior, realizada com problemas das mesmas dimensões, na qual foram gastos 20s de CPU num computador IBM-370/168 [Geromel & Baptistella, 1983].

Além disso, é importante salientar mais uma vez que o problema resolvido não é originalmente um problema convexo e, no entanto, obtivemos em todos os casos a solução ótima global.

Portanto, o método de Benders Generalizado, quando aplicável, além de viabilizar a obtenção da solução ótima global de problemas que na sua forma original não são resolúveis por métodos diretos, nem tampouco são convexos, é um método que se mostrou muito eficiente do ponto de vista computacional.

## CONCLUSÃO GERAL

Neste trabalho estudamos a decomposição Benders Generalizada visando dois aspectos: primeiramente tratamos de obter este método através do conceito de funções suporte, seguida da análise da sua aplicabilidade. Desta análise resultou um novo método de solução para problemas que não satisfaçam a propriedade P.

O segundo aspecto de nosso interesse refere-se a aplicação efetiva deste método ao problema de controle ótimo de geração térmica de energia. Para tanto, foi feito um estudo detalhado deste problema e de sua decomposição (Capítulo 3), do qual foi obtido o programa em linguagem Fortran (ver Apêndice), usado na obtenção dos resultados apresentados no Capítulo 4. Tais resultados ilustram a grande eficiência numérica deste método. De fato, problemas de grande dimensão, seja devido ao número de variáveis como ao número de restrições, puderam ser resolvidos através de um número muito pequeno de iterações e poucos segundos de CPU.

Com este trabalho mostramos, portanto, que a decomposição Benders Generalizada, além de ser um método de fácil obtenção é também muito eficiente do ponto de vista de convergência e gasto computacional.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BAPTISTELLA, L.F.B. & GEROMEL, J.C. (1980), Decomposition Approach to Problem of Unit Commitment Schedule for Hydrothermal Systems, *IEE Proc.*, Vol. 127, nº 6, pp. 250-258.
- BAZARAA, S.M. & SHETTY, C.M. (1979), *Nonlinear Programming-Theory and Algorithms*, John Wiley and Sons, Inc.
- BELLONI, M.R. & GEROMEL, J.C. (1981), Análise de Aplicabilidade de Decomposição Benders Generalizada, *Anais do XIV Simpósio Brasileiro de Pesquisa Operacional (SOBRAP0)*, Vol. 1, pp. 163-175.
- BENDERS, J.F. (1962) Partitioning Procedures for Solving Mixe - Variables Programming Problems, *Numerische Mathematik*, Vol. 4, pp. 238-252.
- DANSKIN, J.M. (1966), The Theory of Max-Min with Applications, *J.STAM Appl. Math.*, Vol. 14, nº 4, pp. 641-665.
- FRANÇA, P.M. (1979), *Problemas de Localização: Solução por decomposição*, Tese de Doutorado, Universidade Estadual de Campinas.
- GEOFFRION, A.M. (1971), Duality in Nonlinear Programming:a Simplified Application-Oriented Development, *SIAM Review* 13, pp. 1-37.
- GEOFFRION, A.M. (1972), Generalized Benders Decomposition, *Journal of Optimization Theory and Applications*, Vol. 10, nº 4, pp. 237-259.
- GEOFFRION, A.M. (1970), Elements of Large-Scale Mathematical Programming, *Management Science*, Vol. 16, nº 11.
- GEROMEL, J.C. & BAPTISTELLA, L.F.B. (1983), Optimal Operation of Thermal Systems with Start-up costs, a ser apresentado no *Third IFAC/IF0ES Symposium on Large Scale Systems: Theory and Applications*, Polônia, 11 a 15 de julho.
- GEROMEL, J.C. & BAPTISTELLA, L.F.B. (1981), A Feasible Direction Method for Large-Scale Programming with Non Convex Objective - Decomposition Approach, *Journal of Optimization Theory and Applications*, Vol. 35, nº 1.

- KARLIN, S. (1959), *Mathematical Methods and Theory in Games, Programming and Economics*, Addison-Wesley Publishing Company, Inc., Vol. 1.
- LASDON, L.S. (1970), *Optimization Theory for Large Systems*, The MacMillan Company, New York.
- LUNA, H.P.L. (1978), *Les Techniques de Décomposition-Coordination dans le Modèle Économiques D'Optimisation*, Thèse d'Etat, Université Paul Sabatier de Toulouse.
- LUNA, H.P.L. (1979), Note on Price Unicity in Economic Equilibrium Models, *Socio-Econ. Plan. Sci.*, Vol. 13, pp. 223-225.
- TURGEON, A. (1978), Optimal Scheduling of Thermal Generating Units, *IEEE A.C.*, Vol. 23, n° 6, pp. 1000-1004.

## APÊNDICE

LISTAGEM DO PROGRAMA FORTRAN DESENVOLVIDO

C PROGRAMA PRINCIPAL CUIA FUNCAO E FORNECER DADOS NECESSARIOS  
C PARA A RESOLUCAO DO PROBLEMA DAS USINAS TERMICAS

C O METODO USADO PARA A SOLUCAO DESTA PROBLEMA E O  
C 'METODO BENDERS GENERALIZADO'.

C OS DADOS NECESSARIOS PARA A EXECUCAO DO PROGRAMA SAO:

C A(1,3): MATRIZ DOS COEFICIENTES DA FUNCAO 'CUSTO DE  
C GERACAO', QUE E DEFINIDA NO SUBPROBLEMA  
C D(I): VETOR DAS DEMANDAS DE ENERGIA A CADA INS-  
C TANTE DE TEMPO  
C SINF(I): QUANTIDADE MINIMA DE ENERGIA QUE PODE SER  
C GERADA PELA USINA I  
C SSUP(I): QUANTIDADE MAXIMA DE ENERGIA QUE PODE SER  
C GERADA PELA USINA I

C ESTE PROGRAMA E DE FATO O PROPRIO METODO BENDERS  
C POIS TODOS OS PASSOS DO METODO SAO DADOS NO PRO-  
C GRAMA PRINCIPAL. JOE APENAS FORNECE OS DADOS NE-  
C CESSARIOS E COMECA A EXECUCAO EM SUAS DIVERSAS  
C ETAPAS

C DURANTE A EXECUCAO DO PROGRAMA SERAO CALCULADOS:

C B(1,NT): MATRIZ BINARIA QUE AO FINAL DA  
C EXECUCAO DAS ITERACOES QUAIS AS USINAS QUE  
C DEVERAO ESTAR LIGADAS EM CADA INSTANTE DE  
C TEMPO  
C C(1,I,J): MATRIZ DAS RESTRICOES DO PROBLE-  
C MA RESTR. AS COMPONENTES DESTA MATRIZ SAO  
C OBTIDAS DURANTE A SOLUCAO DO SUBPROBLEMA  
C ONDE OS ELEMENTOS SAO ALOCADOS TEMPORARIA-  
C MENTE NO VETOR AD(I)  
C CL(I): CUSTO LINEAR DE LIGACAO DA TERMICA I  
C S(1,NT): AO FINAL DA EXECUCAO ESTA MATRIZ  
C SERA COMO COMPONENTES A SOLUCAO TERMICA  
C DURA

C DEVER SER AINDA FORNECIDOS COMO DADOS DE ENTRADA O  
C NUMERO DE USINAS, O NUMERO DE INTERVALOS DE TEMPO  
C A SEREM CONSIDERADOS E HA PARAFETRO EPSILON QUE  
C TRAZ A TOLERANCIA QUE DESEJAMOS QUE SEJA ATENDE-  
C CIDA COMO CRITERIO DE CONVERGENCIA.

COMPON/MI/8(10,3),A(10,24),D(24),SINF(10),SSUP(10),EPS  
COSTO/FAZ/258(10,10),BFF(100),CL(10)

DIMENSIO S(10,24),AD(10),MIA(10)

REAL(20,0) T,H,EPS

8 FORMAT(21,25)

IFOR(20,2) (D(I),I=1,4)

2 FORMAT(240)

DO 3 I=1,7

3 FORM(20,2) (A(I,0),I=1,3)

FORM(20,2) (SINF(I),I=1,4)

FORM(20,2) (SSUP(I),I=1,4)

FORM(20,2) (CL(I),I=1,4)

SSO=1,I+3)

```
C      IMPRESSAO DOS DADOS DE ENTRADA
      WRITE(3,505)
505    FORMAT(30X,'DADOS DE ENTRADA',//)
      WRITE(3,510) N
510    FORMAT(1X,'NUMERO DE USINAS=',I2,//)
      WRITE(3,515) M
515    FORMAT(1X,'INTERVALOS DE TEMPO=',I2,//)
      WRITE(3,520) EPS
520    FORMAT(1X,'TOLERANCIA=',G,//)
      WRITE(3,525)
525    FORMAT(1X,'UNIDADE',4X,'SINF',6X,'SSUP',5X,
1      'A(I,1)',4X,'A(I,2)',5X,'A(I,3)',6X,'CL(I)',//)
      DO 530 I=1,N
530    WRITE(3,535) I,SINF(I),SSUP(I),A(I,1),A(I,2),A(I,3),CL(I)
535    FORMAT(3X,I2,7X,F5.1,5X,F5.1,5X,F5.1,4X,F6.1,4X,F7.5,4X,F7.3)
      WRITE(3,540)
540    FORMAT(//,1X,'NUMERO DE ITERACOES MESTRE/SUBPROBLEMA',//)
      WRITE(3,545)
545    FORMAT(1X,'ITER',5X,'LIM. SUP.',10X,'LIM. INF.',//)

C      GERACAO DO VETOR INICIAL MI(I,NT) E INICIO DO PROCESSO
C      DE OTIMIZACAO

C      INICIAMOS O PROCESSO COM A SOLUCAO FACTIVEL
C      MI(I,NT)=1,P/ TODO I E P/ TODO NT, OU SEJA,
C      TODAS AS USINAS LIGADAS EM TODOS OS INSTAN-
C      TES DE TEMPO.

      DO 10 I=1,N
      DO 10 NT=1,M
10     MI(I,NT)=1
      SE=0

C      CHAMADA DA SUBROTINA QUE RESOLVE O SUBPROBLEMA

      CALL TERA(I,M,S,AQ,0,BS)

C      ALOCACAO DOS COEFICIENTES DA P-ESIMA RESTRICAO NA MATRIZ
C      PES OU TERMINO DO PROCESSO CASO JA ESTEJA SATISFEITO
C      O CRITERIO DE CONVERGENCIA

560    NP=NP+1
      IF (NP.GT.50) GO TO 700
      DO 15 I=1,I
15     PES(NP,I)=AQ(I)
      PES(NP)=0
      IF (BS.GT.0.99) BS=BS0

C      CHAMADA DA SUBROTINA QUE RESOLVE O PROBLEMA MESTRE

      CALL MESTRE(IP,I,M,NIA,SI)

C      IMPRESSAO DO RESULTADO DA P-ESIMA ITERACAO

      WRITE(3,645) NP,BS,SI
645    FORMAT(2X,I2,3X,G,3X,G)
      IF (ABS(BS-SI).LE.EPS) GO TO 655
```

```
DO 670 I=1,N
DO 670 NT=1,N
670 XI(I,NT)=PI(I)
CALL TERM(I,I,S,AD,S,BSD)
IF (ABS(BSD-BI).GE.EPS) GO TO 655
GO TO 590

C IMPRESSAO DOS RESULTADOS

655 WRITE(3,600)
660 FORMAT(//,30X,'SOLUCAO TERMICA OTIMA',//)
690 WRITE(3,605)
605 FORMAT(1X,'I',6X,'DEP',4X,'S1',6X,'S2',6X,'S3',6X,'S4',
1 6X,'S5',6X,'S6',6X,'S7',6X,'S8',6X,'S9',6X,'S10',//)
DO 610 NT=1,N
610 WRITE(3,615) NT,S(NT),(S(I,NT),I=1,N)
615 FORMAT(1X,12,1X,11F9.1)
WRITE(3,625)
625 FORMAT(//,1X,'SOLUCAO DO PROBLEMA MESTRE',//)
DO 630 NT=1,N
630 WRITE(3,635) NT,(XI(I,NT),I=1,N)
635 FORMAT(1X,12,1X,10I4)
GO TO 2000
700 WRITE(3,690)
800 FORMAT(1X,'O NUMERO DE ITERACOES PERMITIDAS FOI ULTRAPASSADO.
1 A MELHOR APROXIMACAO OBTIDA FOI:',//)
WRITE(3,695) BS,BI
895 FORMAT(1X,'BS=',G,10X,'BI=',G,//)
GO TO 900
2000 STOP
END
```

SUBROUTINE TERC(O,N,S,A0,B,BS)

```
C      SUBROTINA QUE CALCULA O NIVEL OTIMO DE GERACAO DE ENERGIA
C      DE CADA USINA TERMICA LIGADA CONFORME O VALOR DE MI DADO.
C      NA SAIDA GATEMOS TAMBEM OS COEFICIENTES DAS NOVAS RESTRI-
C      COES DO PROBLEMA MESTRE E UM NOVO LIMITE SUPERIOR PARA V(MI).

C      ESTES RESULTADOS SAO OBTIDOS PELA OTIMIZACAO DA
C      FUNCAO DUAL DO SUBPROBLEMA, QUE E FEITA POR UM METODO
C      DE INTERPOLACAO QUADRATICA. SAO CALCULADOS INICIALMENTE OS
C      VALORES MAXIMOS E MINIMOS QUE A VARIAVEL DUAL PODE
C      ASSUMIR. ESTA OTIMIZACAO E FEITA SEPARADAMENTE NO
C      TEMPO (PORIS O VETOR MI(I,NT) E CONHECIDO), E AS
C      SOLUCOES TERMICAS TAMBEM SAO OBTIDAS SEPARADA-
C      MENTE DEVIDO A NATUREZA DO PROBLEMA.

COMMON/A1/A(10,3),MI(10,24),D(24),SINF(10),SSUP(10),EPS
COMMON/A2/RES(100,10),PRC(100),CL(10)
DIMENSION S(10,24),ALMAX(24),ALMIN(24),A0(10),ALAM(24)
DO 19 I=1,9
19  A0(I)=0.
    BS=0.
    B=0.
    DO 20 NT=1,N
      ALMAX(NT)=0.
      ALMIN(NT)=1.E+30
      DO 21 I=1,9
        IF (MI(I,NT).EQ.0.) GO TO 21
        XMI=2*A(I,1)*SINF(I)+A(I,2)
        XMA=2*A(I,3)*SSUP(I)+A(I,2)
        IF (ALMAX(NT).GE.XMA) GO TO 21
        ALMAX(NT)=XMA
        IF (ALMIN(NT).LE.XMI) GO TO 21
        ALMIN(NT)=XMI
21  CONTINUE
      Y1=ALMIN(NT)
      CALL GRAD(NT,N,Y1,GF1)
      IF (GF1.LT.EPS) GO TO 100
      Y2=ALMAX(NT)
      CALL GRAD(NT,N,Y2,GF2)
      IF (GF2.GT.EPS) GO TO 101
      IF (ABS(GF2).LE.176) GO TO 100
300  Y=(Y2+GF1-(Y1+GF2))/(GF1-GF2)
      CALL GRAD(NT,N,Y,GF)
      IF (ABS(GF).LE.EPS) GO TO 100
      IF (GF.GT.EPS) GO TO 200
      Y2=Y
      GF2=GF
      GO TO 300
200  Y1=Y
      GF1=GF
      GO TO 300
100  ALAM(NT)=Y
      DO 26 I=1,4
        SS=(ALAM(NT)-A(I,2))/(2.*A(I,3))
        AMI=SSUP(I)
        IF (SSUP(I).GT.SS) AMI=SS
        AMA=SINF(I)
        IF (SINF(I).LT.AMI) AMA=AMI
```

```
SS=AAA
AA=A(I,1)+(B(I,2)-ALAM(NT))*SS+A(I,3)*SS**2
AD(I)=AD(I)+AA
25 IF (MI(I,NT))24,24,25
S(I,NT)=SS
SS=SS+AA
GO TO 26
24 S(I,NT)=0.
26 CONTINUE
SS=SS+ALAM(NT)*D(NT)
BS=BS+ALAM(NT)*D(NT)
GO TO 60
101 *PIIC(3,70) NI
70 FORMAT(/, ' O PROBLEMA E INFACIVEL NO INSTANTE T=',I3)
60 CONTINUE
DO 4 I=1,N
IF (MI(I,NT).EQ.0) GO TO 2
BS=BS+CL(I)
2 CONTINUE
RETURN
END
```

SUBROUTINE GRAD(GT,H,Y,GFI)

C SUBROTINA AUXILIAR QUE CALCULA O VALOR DO GRADIENTE  
C DA FUNCAO DUAL DO SUBPROBLEMA.  
C

COMMON/A1/A(10,3),MI(10,24),D(24),SINF(10),SSUP(10),EPS  
GFI=0(4F)  
DO 44 I=1,H  
IF (MI(I,24).EQ.0) GO TO 44  
S=(1-A(I,2))/(2.\*A(I,3))  
AMINI=SSUP(I)  
IF (SSUP(I).GT.S) AMINI=S  
AMAXI=SINF(I)  
IF (SINF(I).LT.AMINI) AMAXI=AMINI  
S=AMAXI  
GFI=GFI+S  
44 CONTINUE  
RETURN  
END

SUBROUTINE MESTRE (NP, N, N1, NIA, RI)

```
C      SUBROTINA QUE RESOLVE O PROBLEMA MESTRE
C
C      O PROBLEMA MESTRE E OTIMIZADO POR COMPARACAO
C      DOS VALORES DA FUNCAO DE PERTURBACAO GENERALI-
C      ZADA. OS VETORES ANTERIORES SAO GERADOS POR
C      ENUMERACAO EXPLICITA.

      COM=04/A1/A(10,3), AI(10,24), D(24), SINF(10), SSUP(10), EPS
      COMMON/A2/RES(100,10), BRE(100), CL(10)
      DIMENSION NIA(10), NIANA(10)
      SI=1.E+30
      DO 1 I=1,N
1      NIA(I)=1
30      VS=0.
      DO 7 I=1,N
      IF (NIA(I).EQ.0) GO TO 2
      VS=VS+SSUP(I)
2      CONTINUE
      DO 3 NI=1,9
      IF (VS.GE.0(SI)) GO TO 3
      GO TO 100
3      CONTINUE
      SIG=-1.E+30
      DO 4 J=1,NP
      SOM1=0.
      DO 5 I=1,N
      IF (NIA(I).EQ.0) GO TO 5
      SOM1=SOM1+RES(J,I)
5      CONTINUE
      SOM1=SOM1+BRE(J)
      IF (SOM1.GE.SIG) SIG=SOM1
4      CONTINUE
      SOM2=0.
      DO 6 I=1,N
      IF (NIA(I).EQ.0) GO TO 6
      SOM2=SOM2+CL(I)
6      CONTINUE
      SOM2=SOM2+SIG
      IF (SOM2.GE.SI) GO TO 100
      SI=SOM2
      DO 11 I=1,N
11      NIAUX(I)=NIA(I)
100      I=4
7      IF (NIA(I)18,8,9
9      NIA(I)=0
      GO TO 200
8      NIA(I)=1
      I=I-1
      GO TO 7
200      DO 10 I=1,9
      IF (NIA(I).EQ.1) GO TO 300
10      CONTINUE
      DO 12 I=1,9
12      NIA(I)=NIAUX(I)
      RETURN
      END
```