

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA DE CAMPINAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

PROGRAMAÇÃO MATEMÁTICA DE GRANDE PORTE
UMA APLICAÇÃO À SISTEMAS HIDROTÉRMICOS
DE POTÊNCIA

Secundino Soares Filho

Orientador : Hermano de M. F. Tavares

Tese de Doutorado apresentada à Faculdade de Engenharia da Universidade Estadual de Campinas.

OUTUBRO - 1978

UNICAMP
BIBLIOTECA CENTRAL

AGRADECIMENTOS

Aos companheiros Raul, Jura, França, Rui, Akebo, Christiano, Edson, Anilton pela paciência e solidarieda de comigo.

Ao orientador e amigo Hermano Tavares, pela con fiança.

À Gêrard Authiê, o francês amigo, por tudo.

À meus pais

Elvira e Secundino.

À Geni, João e Teresa.

ÍNDICE

Apresentação.....	1
Base Matemática para Decomposição.....	2
Métodos de Decomposição e Coordenação.....	29
Método Dual Clássico.....	31
Programação Linear Generalizada.....	33
Aproximação Tangencial.....	37
Alocação de Recursos.....	40
Decomposição de Benders.....	45
Rosen.....	49
Otimização Hierárquica Aplicada a Sistemas de Potência Hidrotérmicos.....	53
Introdução.....	56
Formulação Matemática.....	57
Manipulação.....	60
Estratégia de Resolução.....	71
Interpretação Econômica.....	74
Exemplo.....	76
Comentários.....	82
Apêndice.....	85
Bibliografia.....	86

APRESENTAÇÃO

Dentro da área de Sistemas do Departamento de Engenharia Elétrica da FEC - UNICAMP, nosso interesse de pesquisa tem se voltado para a Programação Matemática. Em particular temos trabalhado em otimização de sistemas de Grande Porte, ramo relativamente recente que vem ganhando importância devido a magnitude crescente dos problemas a serem tratados.

Este trabalho está inserido neste mesmo contexto. Por um lado é desenvolvido um estudo resumindo alguns dos resultados mais importantes da Programação Matemática de Grande Porte (Capítulo II), precedido de uma reapresentação da base matemática necessária (Capítulo I). O objetivo é dar um corpo mais unificado ao tema, usando-se para isso o conceito chave de função de perturbação.

Por outro lado desenvolveu-se uma aplicação: O problema de alocação da geração em sistemas hidrotérmicos de potência. Trata-se de um problema clássico de sistemas de potência, abordado de diversas maneiras, mas onde entretanto subsistem algumas dificuldades. A principal talvez seja a dimensionalidade, que tem acarretado a necessidade de simplificar o modelamento. Métodos de decomposição se apresentam portanto promissores no tratamento deste problema. No Capítulo III apresentamos, ao lado de uma formulação mais rigorosa do problema, um método de resolução baseado em decomposição dual.

CAPÍTULO I

"BASE MATEMÁTICA PARA DECOMPOSIÇÃO"

Neste capítulo é reapresentada a teoria da dualidade em programação não linear, com ênfase especial no conceito de função de perturbação. O que se pretende mostrar é que a partir da caracterização desta função pode-se visualizar os principais resultados da programação não linear, suas condições de otimalidade (necessárias e suficientes) bem como a sequência de funcionamento dos principais métodos de decomposição e coordenação para sistemas de grande porte (cap. II). Todos os resultados aqui apresentados podem ser encontrados nos trabalhos da área, notadamente em Lasdon [1], Geoffrion [2] e Shoeffler [3]. Optou-se pela demonstração de teoremas conhecidos no sentido de dar um corpo mais acabado ao capítulo.

Seja o problema de programação matemática que também chama remos problema primal (P)

$$(P) \quad \begin{array}{l} \min f(x) \\ \text{s.a. } g(x) \leq 0 \\ x \in S \end{array}$$

onde $x \in R^n$, $S \subseteq R^n$, f é uma função definida em S e $g(x) = [g_1(x), \dots, g_m(x)]^T$ (*) é um vetor de funções definidas em S . O problema consiste em determinar um vetor $x^0 \in R^n$ que seja factível, isto é, $x^0 \in S$ e $g(x^0) \leq 0$ e que faça a função objetivo f assumir o menor valor possível. O que se desenvolverá a seguir é a análise das condições que um vetor x^0 deve satisfazer para ser solução ótima de (P).

Seja para tanto a função Lagrangeano definida no R^{n+m} associada ao problema (P)

$$L(x, u) = f(x) + u g(x)$$

onde o vetor $u \in R^m$ é chamado vetor multiplicador de Lagrange generalizado.

Um par (x^0, u^0) satisfaz as condições de otimalidade (CO) do problema (P) se e somente se

$$(CO) \quad \begin{array}{l} (1) \ x^0 \text{ minimiza } L(x, u^0) \text{ sobre } S \\ (2) \ g(x^0) \leq 0 \\ (3) \ u^0 g(x^0) = 0 \\ (4) \ u^0 \geq 0 \end{array}$$

(*) O índice T significa transposto. Será omitido de modo a simplificar a notação sempre que a operação vetorial for evidente.

TEOREMA 1

As (CO) são condições suficientes para que x^0 seja solução ótima de (P).

PROVA

As condições (1) e (2) garantem a factibilidade de x^0 em (P) pois

$$x^0 \in S$$

$$g(x^0) \leq 0$$

A condição (1) garante ademais que

$$f(x^0) + u^0 g(x^0) \leq f(x) + u^0 g(x) \quad \forall x \in S$$

Como pela condição (3) $u^0 g(x^0) = 0$ temos

$$f(x^0) \leq f(x) + u^0 g(x) \quad \forall x \in S$$

e como pela condição (4) $u^0 \geq 0$ temos que para todo $x \in S$ que satisfaz $g(x) \leq 0$, o produto $u^0 g(x) \leq 0$ e assim

$$f(x^0) \leq f(x) \quad \forall x \in S, \quad g(x) \leq 0$$

Um vetor u^0 que satisfaz as (CO) para algum x será chamado vetor multiplicador ótimo (VMO). Qualquer (VMO) caracteriza todo conjunto solução de (P) (*).

Um par (x^0, u^0) , $x^0 \in S$ e $u^0 \geq 0$, é um ponto de sela (PS) restrito da função Lagrangeano se e somente se

(PS)	(5) $L(x^0, u^0) \leq L(x, u^0)$	$\forall x \in S$
	(6) $L(x^0, u^0) \geq L(x^0, u)$	$\forall u \geq 0$

(*) Se (x^0, u^0) satisfaz (CO) e x^1 é solução ótima de (P) então pelo teorema 1 $f(x^0) = f(x^1)$ e de (1) e (3) $f(x^0) \leq f(x^1) + u^0 g(x^1)$, com o que $u^0 g(x^1) \geq 0$. Mas $u^0 \geq 0$ e $g(x^1) \leq 0$, portanto $u^0 g(x^1) = 0$ e consequentemente $L(x^0, u^0) = L(x^1, u^0)$. Assim (x^1, u^0) satisfaz (CO).

TEOREMA 2

As (CO) são equivalentes à condição de (PS)

PROVA

⇒ (1) e (5) são equivalentes. Resta então provar a equivalência entre (6) e (2), (3) uma vez que (4) é evidente.

De (6) podemos escrever

$$f(x^0) + u^0 g(x^0) \geq f(x^0) + u g(x^0) \quad \forall u \geq 0$$

ou seja

$$(u - u^0) g(x^0) \leq 0 \quad \forall u \geq 0$$

Se $g_i(x^0) > 0$, fazendo $u_j = u_j^0, \forall j \neq i$ e $u_i > u_i^0$

a relação acima não é satisfeita. Portanto

$$g(x^0) \leq 0 \quad [\text{condição (2)}]$$

Fazendo $u = 0$ pela mesma relação $u^0 g(x^0) \geq 0$. Mas como $u^0 \geq 0$ e $g(x^0) \leq 0$ temos que $u^0 g(x^0) \leq 0$. Portanto

$$u^0 g(x^0) = 0 \quad [\text{condição (3)}]$$

⇒ Por definição

$$L(x^0, u^0) = f(x^0) + u^0 g(x^0)$$

mas pela condição (3) $u^0 g(x^0) = 0$ e assim

$$L(x^0, u^0) = f(x^0)$$

Também por definição e usando o resultado anterior

$$L(x^0, u) = f(x^0) + u g(x^0) = L(x^0, u^0) + u g(x^0)$$

Como pela condição (2) $g(x^0) < 0$ temos que $u g(x^0) < 0$, $\forall u > 0$ e portanto

$$L(x^0, u) < L(x^0, u^0) \quad \forall u > 0 \quad [\text{condição (6)}]$$

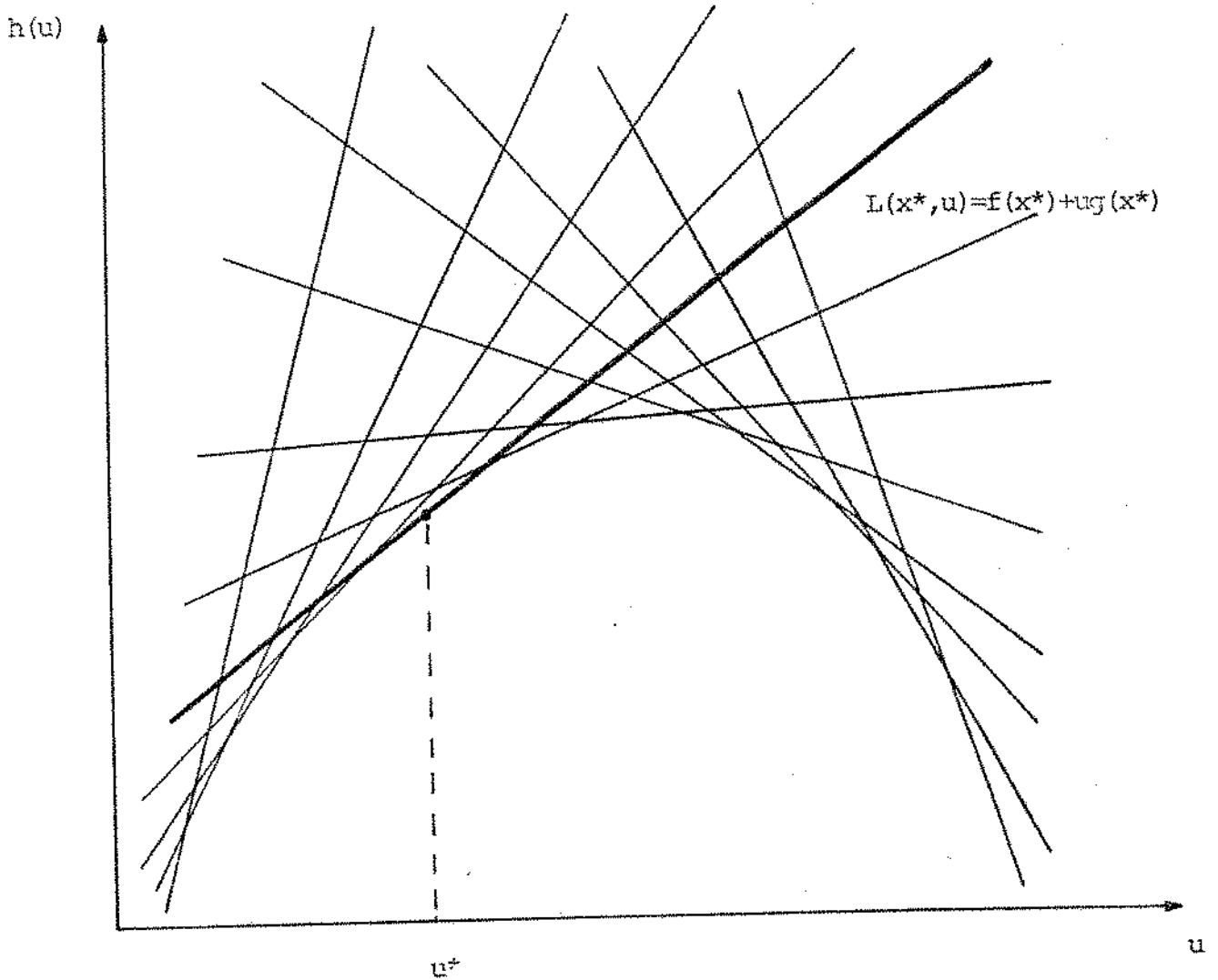
Define-se como função dual (FD) o ínfimo^(*) em $x \in S$ da função Lagrangeano para cada $u > 0$

$$(FD) \quad \boxed{h(u) = \underset{x \in S}{\text{ínfimo}} L(x, u)}$$

A função dual é côncava pois é o ínfimo ponto a ponto de uma coleção de funções (uma para cada $x \in S$) lineares em u ^(**). Graficamente

(*) O ínfimo de uma função é seu maior limitante inferior. Se a função for limitada o ínfimo será $-\infty$. A definição da função dual em termos de ínfimo e não mínimo evita a necessidade da definição de um domínio especial.

(**) É possível mostrar que ínfimo_x $\{ f_i(x) \}$ é uma função côncava se $f_i(x)$ for côncava para todo i . Isto porque uma função $f(x)$ é côncava se e somente se seu hipografo $H = \{ (y,x) / y \leq f(x) \}$ for convexo. Como a intersecção de conjuntos convexos é convexa segue que o hipografo de ínfimo_x $\{ f_i(x) \}$ é convexo e portanto caracteriza uma função côncava. A função $h(u)$ é côncava por particularização deste resultado à funções lineares.



Desta interpretação da função dual resulta uma importante propriedade utilizada nos métodos de otimização que se analisará no capítulo II qual seja que a avaliação da função dual num ponto u^* fornece automaticamente um hiperplano suporte a ela dado por

$$L(x^*, u) = f(x^*) + u g(x^*)$$

onde x^* minimiza $L(x, u^*)$ sobre S .

Associado ao problema primal (P) define-se um problema dual (D) que consiste na maximização da função dual sobre seu domínio de definição

(D)

$\begin{aligned} \max \quad & h(u) \\ u \geq & 0 \end{aligned}$

TEOREMA 3

O valor da função dual é um limitante inferior do valor da função primal, ou seja

$$h(u) \leq f(x) \quad \forall u \geq 0, x \in S, g(x) \leq 0$$

PROVA

Por definição

$$h(u) \leq f(x) + u g(x) \quad \forall x \in S$$

Como para todo $u \geq 0, g(x) \leq 0$ o produto $ug(x) \leq 0$ temos

$$h(u) \leq f(x) \quad \forall u \geq 0, x \in S, g(x) \leq 0 \quad \blacksquare$$

Dois importantes resultados derivam diretamente do teorema 3

COROLÁRIO 1

Se o valor ótimo de (D) é ilimitado então (P) é infactível e se o valor ótimo de (P) é ilimitado então (D) tem valor $-\infty, \forall u \geq 0$.

COROLÁRIO 2

Se existir x^0 factível em (P) e u^0 factível em (D) tal que $f(x^0) = h(u^0)$ então x^0 é solução ótima de (P) e u^0 é solução ótima de (D).

O corolário 2 define o que se chamará condições de dualidade (CD)

(CD)

$$(7a) \quad x^0 \in S$$

$$(7b) \quad g(x^0) \leq 0$$

$$(8) \quad u^0 \geq 0$$

$$(9) \quad f(x^0) = h(u^0)$$

TEOREMA 4 As (CO) são equivalentes às (CD)

PROVA ⇒ De (1) e (2) temos que x^0 é factível em (P) e de (4) u^0 é factível em (D). Portanto (7) e (8) são satisfeitas. Por definição

$$h(u^0) = \inf_{x \in S} L(x, u^0)$$

mas como pela condição (1) x^0 minimiza $L(x, u^0)$ sobre S temos

$$h(u^0) = f(x^0) + u^0 g(x^0)$$

e como por (3) $u^0 g(x^0) = 0$ temos finalmente

$$h(u^0) = f(x^0) \quad [\text{condição (9)}]$$

⇒ De (7b) temos $g(x^0) \leq 0$ [condição (2)] e de (8) temos $u^0 \geq 0$ [condição (4)]. Suponhamos que x^0 não satisfaz (1). Então

$$h(u^0) = \inf_{x \in S} L(x, u^0) < f(x^0) + u^0 g(x^0)$$

Mas de (9) $h(u^0) = f(x^0)$ donde

$$u^0 g(x^0) > 0$$

O que é absurdo pois $u^0 \geq 0$ e $g(x^0) \leq 0$. Então x^0 satisfaz (1).

$$h(u^0) = f(x^0) + u^0 g(x^0)$$

e como $h(u^0) = f(x^0)$ a condição (3) $u^0 g(x^0) = 0$ também é satisfeita

As condições de otimalidade (CO), de ponto de sela (PS) e de dualidade (CD) são todas equivalentes entre si como se mostrou. Um par (x, u) que as satisfaça resolve os problemas primal e dual independentemente de hipóteses de convexidade, diferenciabilidade, etc. Métodos de otimização cuja estratégia de solução baseia-se em satisfazer estas condições trabalham com uma sequência de problemas de otimização chamados problemas Lagrangeanos (PL)

$$(PL) \quad \boxed{\begin{array}{c} \min \{ f(x) + u^* g(x) \} \\ x \in S \end{array}}$$

onde $u^* \geq 0$ é um parâmetro dado. O conjunto solução de (PL) é anotado

$$X(u^*) = \left\{ x^* \in S / x^* \text{ minimiza } L(x, u^*) \right\}$$

A resolução do problema Lagrangeano (PL) satisfaz parcialmente as condições para que $x^* \in X(u^*)$ seja solução ótima de (P). A resolução de uma sequência de (PL), que corresponde a diversas avaliações da função dual $h(u)$, converge para a solução ótima de (P) e (D) quando as demais condições forem satisfeitas, quais sejam

$$\left. \begin{array}{l} (2) \quad g(x^*) \leq 0 \\ (3) \quad u^* g(x^*) = 0 \end{array} \right\} \quad (CO)$$

ou

$$(6) \quad L(x^*, u^*) \geq L(x^*, u), \forall u \geq 0 \quad \left. \right\} \quad (PS)$$

ou

$$\left. \begin{array}{l} (7b) \quad g(x^*) \leq 0 \\ (9) \quad f(x^*) = h(u^*) \end{array} \right\} \quad (CD)$$

Por outro lado cada resolução do problema Lagrangeano (PL) está associada a solução de um problema muito similar ao problema primal (P) como mostrou Everett [4]

TEOREMA 5 Se x^* resolve (PL) para $u^* \geq 0$ então x^* resolve o problema perturbado (PP)

(PP)

$\begin{aligned} \min & f(x) \\ \text{s.a} & g(x) \leq y \\ & x \in S \end{aligned}$
--

$$\left. \begin{aligned} \text{onde } y_i &= g_i(x^*) \quad \text{se } u_i^* > 0 \\ y_i &\geq g_i(x^*) \quad \text{se } u_i^* = 0 \end{aligned} \right\} \text{(a)}$$

PROVA Como x^* resolve (PL) para $u^* \geq 0$ temos

$$f(x^*) + u^* g(x^*) \leq f(x) + u^* g(x) \quad \forall x \in S$$

ou seja

$$f(x^*) \leq f(x) + u^* [g(x) - g(x^*)] \quad \forall x \in S$$

Como $u^* \geq 0$ podemos escrever que $\forall x \in S$ tal que $g(x) \leq g(x^*)$

$$f(x^*) \leq f(x)$$

o que significa que x^* resolve (PP). Observe-se que se $u_i^* = 0$ não é necessário impor $g_i(x) \leq g_i(x^*)$ o que está representado em (a).

Relacionado com (PP) define-se uma função de perturbação (FP)

(FP)	$v(y) = \text{ínfimo}_{x \in S} f(x)$	s.a. $g(x) \leq y$
	$x \in S$	

que associa a cada vetor de perturbação y o valor ótimo do problema perturbado (PP). O domínio efetivo de $v(y)$ é dado pela condição de factibilidade do problema perturbado

$$V = \left\{ y / \exists x \in S, g(x) \leq y \right\}$$

O valor da função de perturbação para $y = 0$ é o valor ótimo de (P). A seguir são provadas algumas características desta função.

TEOREMA 6 A função de perturbação é não crescente

PROVA Seja $y_1 \leq y_2$. Por definição de função de perturbação

$$v(y_1) = \text{ínfimo}_{x \in S} \left\{ f(x) / g(x) \leq y_1 \right\} = f(x_1)$$

onde $g(x_1) \leq y_1$. Por outro lado

$$v(y_2) = \text{ínfimo}_{x \in S} \left\{ f(x) / g(x) \leq y_2 \right\}$$

com o que para todo $x \in S$, $g(x) \leq y_2$ pode-se es
crever

$$v(y_2) \leq f(x)$$

Em particular para $x_1 \in S$, $g(x_1) \leq y_1 \leq y_2$

$$v(y_2) \leq f(x_1) = v(y_1)$$

TEOREMA 7

$v(y)$ é convexa sobre V convexo se $f(x)$ e $g(x)$ forem convexas sobre S convexo

PROVA

As hipóteses são ① convexidade de $f(x)$, ② convexidade $g(x)$ e ③ convexidade de S . Mostremos inicialmente a tese da convexidade de V . Se

$$y_1 \in V \Leftrightarrow \exists x_1 \in S / g(x_1) \leq y_1$$

$$y_2 \in V \Leftrightarrow \exists x_2 \in S / g(x_2) \leq y_2$$

tomando uma combinação convexa desses pontos ($0 \leq \lambda \leq 1$; $\lambda + \bar{\lambda} = 1$)

$$y = \lambda y_1 + \bar{\lambda} y_2 \geq \lambda g(x_1) + \bar{\lambda} g(x_2) \stackrel{\textcircled{2}}{\geq} g(\lambda x_1 + \bar{\lambda} x_2)$$

$$\lambda x_1 + \bar{\lambda} x_2 \stackrel{\textcircled{3}}{\in} S$$

e portanto $y \in V$.

Para se verificar a convexidade de $v(y)$, lembrando-se que uma função é convexa se e somente se tem epigrafo convexo, analisaremos a convexidade de

$$H = \left\{ (z, y) / z \geq f(x), y \geq g(x), x \in S \right\}$$

que é o epigrafo de $v(y) = \inf_{x \in S} \left\{ f(x) / g(x) \leq y \right\}$

Seja $(z_1, y_1) \in H$, $(z_2, y_2) \in H$, $0 \leq \lambda = 1 - \bar{\lambda} \leq 1$.

A combinação convexa $(z, y) = (\lambda z_1 + \bar{\lambda} z_2, \lambda y_1 + \bar{\lambda} y_2)$

é tal que

$$\lambda z_1 + \bar{\lambda} z_2 \geq \lambda f(x_1) + \bar{\lambda} f(x_2) \geq f(\lambda x_1 + \bar{\lambda} x_2) \quad \textcircled{1}$$

$$\lambda y_1 + \bar{\lambda} y_2 \geq \lambda g(x_1) + \bar{\lambda} g(x_2) \geq g(\lambda x_1 + \bar{\lambda} x_2) \quad \textcircled{2}$$

$$\lambda x_1 + \bar{\lambda} x_2 \in S \quad \textcircled{3}$$

donde se conclue que $(z, y) \in H$, H é portanto convexo e consequentemente $v(y)$ também é convexa. ■

O teorema seguinte permite uma útil interpretação geométrica associada à função de perturbação.

TEOREMA 8 x^* minimiza $L(x, u^*)$ sobre S onde $u^* \geq 0$ se e somente se $v(y) \geq v(y^*) - u^* (y - y^*) \quad \forall y \in V$ onde $y^* = g(x^*)$.

PROVA

⇒ Como x^* minimiza o Lagrangeano para $u^* \geq 0$ temos

$$f(x^*) + u^* g(x^*) \leq f(x) + u^* g(x) \quad \forall x \in S$$

ou

$$f(x) \geq f(x^*) - u^* [g(x) - g(x^*)] \quad \forall x \in S$$

como $g(x^*) = y^*$ e pelo teorema 5 $f(x^*) = v(y^*)$ temos

$$f(x) \geq v(y^*) - u^* [g(x) - y^*] \quad \forall x \in S$$

Mas como $u^* \geq 0$ segue-se que para todo $y \in V$ podemos escrever

$$f(x) \geq v(y^*) - u^*(y - y^*), \quad \forall x \in S \quad \text{tal que } g(x) \leq y$$

e tomando finalmente a expressão acima para x que minimiza $f(x)$ sujeito a $x \in S$, $g(x) \leq y$, ou seja, para x que resolve o problema perturbado com lado direito y

$$v(y) \geq v(y^*) - u^*[y - y^*] \quad \forall y \in V$$

$$= v(y) > v(y^*) - u^*[y - y^*] \quad \forall y \in V, \quad y^* = g(x^*)$$

Para x^* que resolve (PP) com lado direito $y^*=g(x^*)$ te mos $v(y^*) = f(x^*)$ e fazendo $y = g(x)$ para $x \in S$ temos

$$v(g(x)) \geq f(x^*) - u^*[g(x) - g(x^*)] \quad \forall x \in S$$

Como $v(g(x)) \leq f(x) \quad \forall x \in S$ podemos escrever

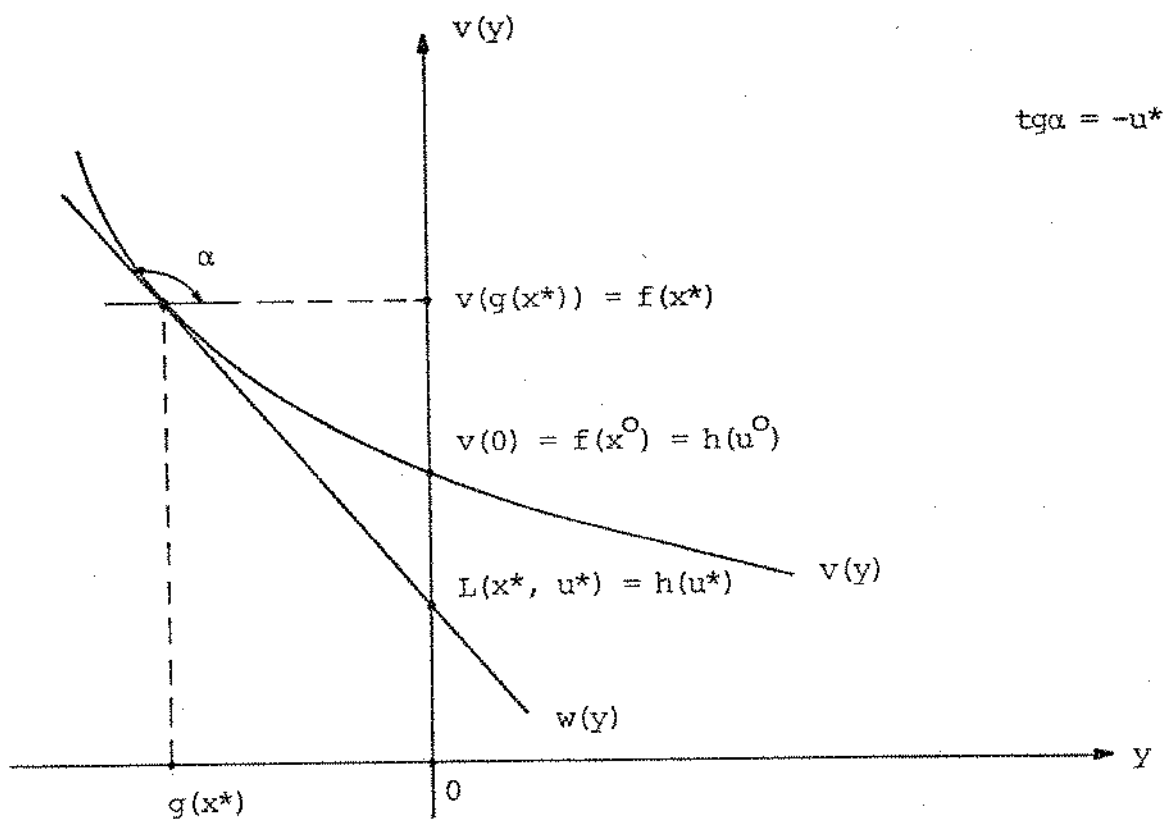
$$f(x) \geq f(x^*) - u^*[g(x) - g(x^*)] \quad \forall x \in S$$

ou

$$f(x^*) + u^* g(x^*) \leq f(x) + u^* g(x) \quad \forall x \in S$$

o que significa que x^* minimiza $L(x, u^*)$ sobre S . ■

Do teorema 8 podemos deduzir importante interpretação geo-métrica que relaciona a função de perturbação, a função dual, o problema Lagrangeano e os problemas primal e dual.



Resolver o problema Lagrangeano para $u^* > 0$ significa geometricamente determinar o hiperplano suporte à função de perturbação com inclinação $-u^*$. O ponto suporte é dado por $[f(x^*), g(x^*)]$ onde x^* é solução do problema Lagrangeano. A equação do hiperplano é dada por

$$w(y) = f(x^*) - u^* \left[y - g(x^*) \right] = L(x^*, u^*) - u^* y$$

A intersecção do hiperplano suporte com o eixo $y = 0$ fornece o valor da função dual para u^* .

$$w(0) = L(x^*, u^*) = h(u^*)$$

Assim o problema dual pode ser visualizado geometricamente como achar a inclinação $-u^0$ do hiperplano suporte à função de perturbação que maximiza a intersecção com o eixo $y = 0$. Para essa inclinação a solução do problema Lagrangeano x^0 é solução do problema primal (*)

A função de perturbação e sua interpretação geométrica joga um importante papel na compreensão dos métodos de otimização por decomposição e coordenação, como se verá no capítulo II. A moderna teoria de dualidade em programação não linear se baseia centralmente neste conceito.

A interpretação econômica do vetor multiplicador de Lagrange como preço de equilíbrio associado aos recursos y pode ser obtida da análise de sensibilidade na função de perturbação.

Se x^* resolve (PL) para $u^* > 0$ sabemos pelo teorema 8 que

$$v(y) \geq v(y^*) - u^* (y - y^*) \quad \forall y \in V$$

(*) na verdade isto nem sempre ocorre. Se o problema Lagrangeano admitir $g(x)$ não constante em $X(u^0)$ então pode-se provar [1,2] que existe pelo menos uma solução de (PL) que não é factível em (P).

onde $y^* = g(x^*)$ e $v(y^*) = f(x^*)$. Para $y = y^* + \alpha s$ onde $\alpha > 0$ é um escalar e s uma direção factível em V a partir de y^* temos

$$v(y^* + \alpha s) \geq v(y^*) - \alpha u^*s$$

ou

$$\frac{v(y^* + \alpha s) - v(y^*)}{\alpha} \geq -u^*s$$

e tomando o limite para $\alpha \rightarrow 0$ obtemos

$$D v(y^*, s) \geq -u^*s$$

onde $D v(y^*, s)$ é a derivada direcional de v no ponto y^* na direção s . Fazendo $s = e_j^{(*)}$

$$D v(y^*, e_j) \geq -u_j^*$$

o que significa que $-u_j^*$ é um limitante inferior da taxa marginal de variação no valor ótimo do problema perturbado para um acréscimo no j -ésimo recurso $(^{**})$. O vetor multiplicador de Lagrange corresponde assim a preços de equilíbrio dos recursos no sentido de que o problema (Langrangeano)

$$\left\{ \begin{array}{l} \min f(x) + u g(x) \\ x \in S \end{array} \right.$$

(*) e_j é o vetor de componentes todas zero à exceção do j -ésimo componente que é um.

(**) Assim u_j^* é uma estimativa "otimista" da redução marginal do valor ótimo de (P). No caso de $v(y)$ ser diferenciável em y^* , u_j^* é a própria redução marginal.

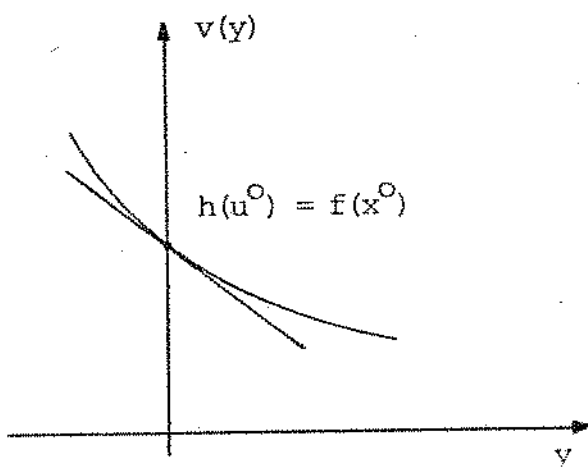
representa economicamente o mesmo problema original (P) quando é possível negociar recursos no mercado a preço u . As restrições $g(x) \leq 0$ passam a não existir na prática porém cada decisão x acarreta um custo adicional em f correspondente a negociação dos recursos necessários $ug(x)$. É possível mostrar (Everett [4]) que a elevação do preço de um recurso u_j (mantendo-se inalterado os demais) tende a reduzir a utilização desse recursos na solução do problema Lagrangeano.

Até agora nada se exigiu em termos de convexidade, diferencialidade, etc... do problema primal. Os resultados obtidos, condições suficientes de otimalidade, são pois totalmente gerais, aplicáveis a qualquer problema de programação matemática. Na análise porém das condições necessárias algumas propriedades serão exigidas no sentido de garantir a existência de vetores multiplicadores, ponto de sela, etc...

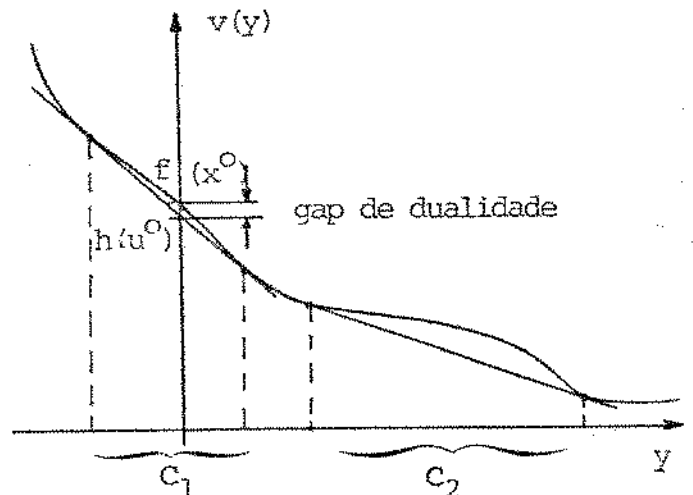
Um primeiro resultado importante decorre diretamente do teorema 8.

COROLÁRIO 3 Seja x^0 a solução ótima de (P). Então existe vetor multiplicador $u^0 \geq 0$ tal que x^0 minimiza $L(x, u^0)$ sobre S se e somente se a função de perturbação admite hiperplano suporte no ponto $v(0)$ com inclinação $-u^0$.

Para a existência de hiperplanos suportes à função de perturbação em todos os pontos $y \in V$ (e portanto em $y = 0$) uma condição necessária e suficiente é $v(y)$ ser uma função convexa, ou seja, (P) ser um programa convexo. Caso contrário existirão subconjuntos $C_i \subset V$ onde $v(y)$, $y \in C_i$, não admite hiperplano suporte.



(1) - (P) convexo, não há "gap"

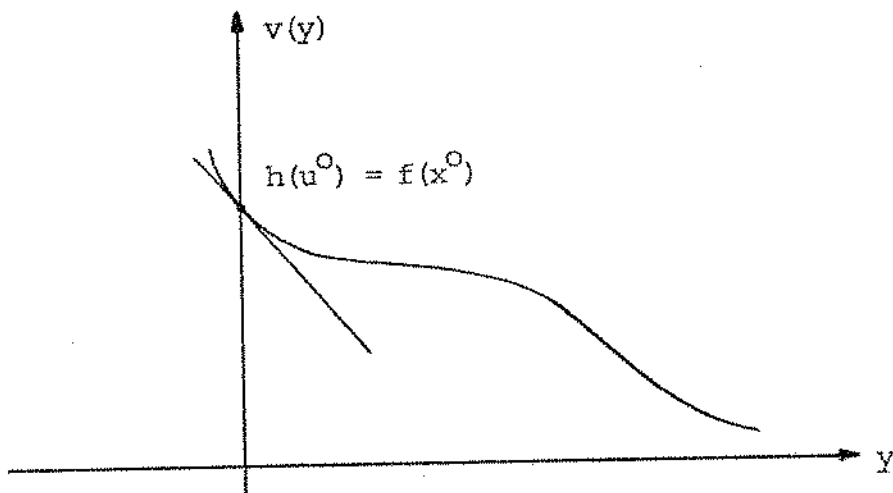


(2) - (P) não convexo, há "gap"

Métodos de otimização por resolução de problemas Lagrangeanos sucessivos não conseguem "gerar" $y \in C_i$ e portanto não convergirão se $y = 0$ estiver contido em algum C_i . Neste caso o problema primal não admite ponto de sela, existindo um "gap de dualidade" dado por

$$h(u^0) < f(x^0)$$

onde u^0 é solução ótima de (D) e x^0 é solução ótima de (P). É possível porém que não exista "gap" mesmo se (P) não for convexo.



(3) - (P) não convexo, não há "gap"

Só a convexidade, entretanto, é uma exigência ainda insuficiente para garantir a existência de vetor multiplicador ótimo (VMO). Isto porque mesmo que $v(y)$ seja convexa é possível que o hiperplano no suporte no ponto $y = 0$ tenha inclinação não finita. Para eliminar esta possibilidade, que corresponde a casos patológicos de programação matemática, define-se

(P) é estável se e somente se $v(0)$ é finita e

$$\lim_{\alpha \rightarrow 0^+} \frac{v(\alpha y) - v(0)}{\alpha \|y\|} > -\infty \quad \forall y \in V, y \neq 0$$

ou seja, a derivada direcional no ponto $y=0$ não é $-\infty$ em nenhuma direção. Com isto podemos apresentar o seguinte teorema.

TEOREMA 9

Se (P) tem solução ótima x^0 , é convexo e estável então existe (VM0) u^0 tal que (x^0, u^0) satisfaz as (CO).

PROVA

É sabido que uma função convexa admite hiperplano suporte num ponto onde é finita se e somente se sua derivada direcional não é $-\infty$ em nenhuma direção. Assim $v(y)$ sendo convexa (pois (P) é convexo) e tendo derivada direcional finita (pois (P) é estável) podemos escrever

$$v(y) \geq v(0) - u^0 y \quad \forall y \in V$$

onde $-u^0$ é a inclinação do hiperplano suporte em $v(0)$. É preciso mostrar agora que (x^0, u^0) satisfaz as (CO). Condição (2) é imediata. Fazendo $y = e_j$ obtemos

$$u_j^0 \geq v(0) - v(e_j)$$

e como $v(y)$ é não crescente $v(0) - v(e_j) \geq 0$ e portanto $u_j^0 \geq 0$ [condição (4)]. Fazendo $y = g(x^0)$ obtemos

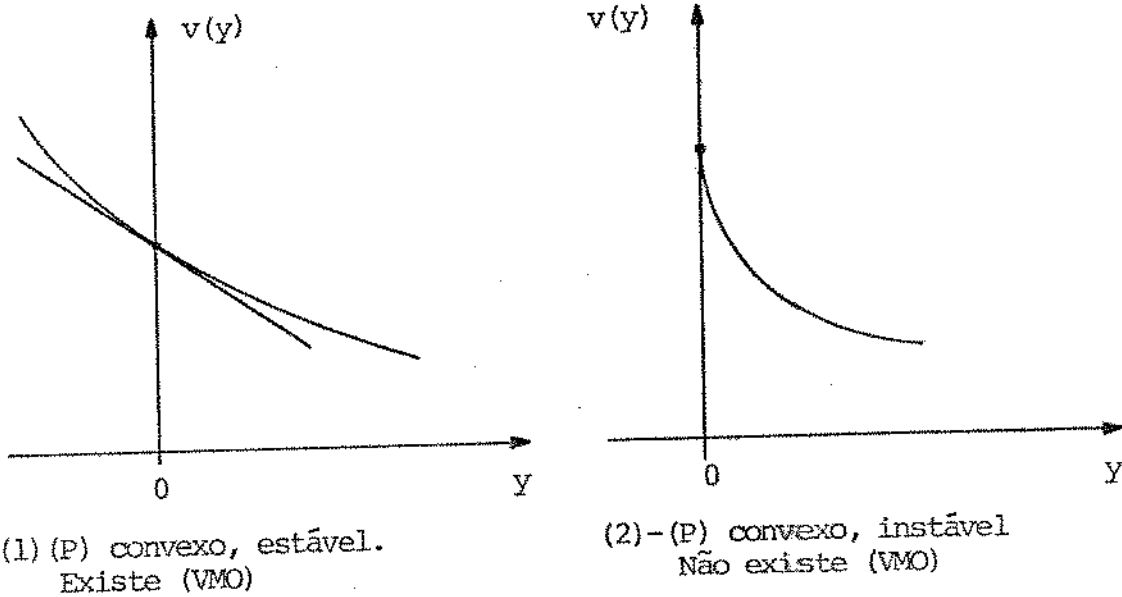
$$u^0 g(x^0) \geq v(0) - v(g(x^0))$$

Mas como $v(g(x^0)) = v(0)$ pois diminuindo o lado direito de $y = 0$ até $y = g(x^0)$ não se destrói a otimalidade de x^0 temos $u^0 g(x^0) \geq 0$. Mas de (2) e (4) $u^0 g(x^0) \leq 0$ e assim

$$u^0 g(x^0) = 0 \quad \text{[condição (3)]}$$

Finalmente pelo teorema 8 para $y^* = 0$ a condição (1) também é satisfeita

Graficamente, para um programa convexo podemos visualizar a existência ou não de (VMO) finito pela propriedade da estabilidade, lembrando que uma função convexa só pode apresentar derivada direcional infinita em pontos da fronteira de seu domínio.



No caso de problemas não convexos onde há "gap de dualidade" não existe (PS) do Lagrangeano e não se aplicam as (CO) ou (CD). Entretanto nestes casos a inclinação da função de perturbação no ponto $y = 0$ (que não define hiperplano suporte) permite caracterizar uma condição de otimalidade alternativa.

Suponhamos $S = E^n$, f e g_i de classe C_1 em S . Nestas condições é possível obter, através de expansão em Taylor e do Lema de Farkas, uma condição necessária de otimalidade para o problema primal.

Seja x^0 factível, isto é

$$g(x^0) \leq 0$$

e seja I o conjunto de índices das restrições ativas em x^0 . Assim,

$$g_i(x^0) = 0, \quad i \in I$$

$$g_i(x^0) < 0, \quad i \notin I$$

Uma direção δx com $\|\delta x\|$ arbitrariamente pequeno é factível a partir de x^0 se

$$g_i(x^0 + \delta x) \leq 0, \quad i \in I$$

Expandindo em Taylor e desprezando os termos de segunda ordem

$$g_i(x^0 + \delta x) = g_i(x^0) + \nabla g_i(x^0) \cdot \delta x \leq 0$$

e lembrando que $g_i(x^0) = 0$ para todo $i \in I$, qualquer direção factível δx a partir de x^0 se caracterizará por

$$\nabla g_i(x^0) \cdot \delta x \leq 0, \quad i \in I$$

Analisando agora a função objetivo, para que x^0 seja ótimo local é necessário que para todo δx satisfazendo a expressão acima tenhamos

$$f(x^0 + \delta x) \geq f(x^0)$$

Expandindo em Taylor e desprezando os termos de segunda ordem

$$f(x^0 + \delta x) = f(x^0) + \nabla f(x^0) \cdot \delta x \geq f(x^0)$$

ou seja

$$\nabla f(x^0) \cdot \delta x \geq 0$$

Assim, resumindo, uma condição necessária de ótimo local é para todo δx tal que

$$\nabla g_i(x^0) \cdot \delta x \leq 0, \quad i \in I$$

tenhamos

$$\nabla f(x^0) \cdot \delta x \geq 0$$

Através do Lema de Parkas a condição acima pode ser reescrita como

$$-\nabla f(x^0) = \sum_{i \in I} u_i^0 \cdot \nabla g_i(x^0)$$

com

$$u_i^0 \geq 0$$

e finalmente generalizando a expressão para incluir também as res trições não ativas (com $u_i^0 = 0$) temos

$$-\nabla f(x^0) = \sum_i u_i^0 \nabla g_i(x^0)$$

$$u_i^0 \geq 0$$

$$u_i^0 \cdot g_i(x^0) = 0$$

As condições que acabamos de obter são chamadas condições de Kuhn - Tucker (KT) e podem ser formalizadas como se segue

(KT)

$$(a) \nabla_x L(x^0, u^0) = \nabla f(x^0) + u^0 \nabla g(x^0) = 0$$

$$(b) \nabla_u L(x^0, u^0) = g(x^0) \leq 0$$

$$(c) u^0 \cdot \nabla_u L(x^0, u^0) = u^0 g(x^0) = 0$$

$$(d) u^0 \geq 0$$

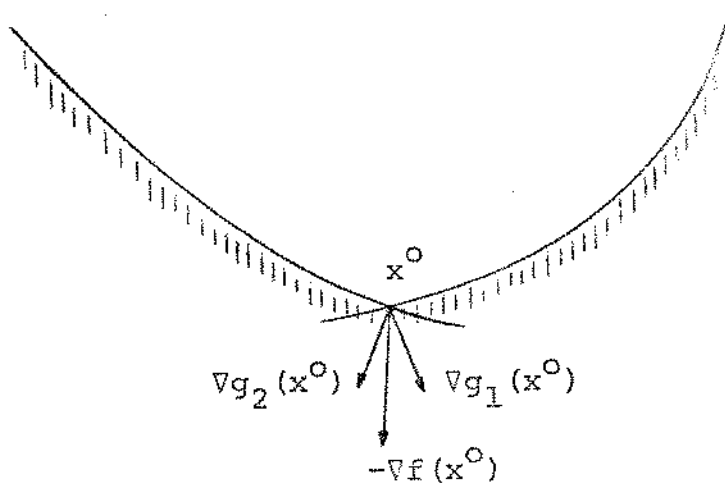
Estas condições se apresentam em termos de estacionariedade do Lagrangeano. É interessante notar que (b) (c) (d) são iguais a (2) (3) (4). A diferença entre (CO) e (KT) reside pois em substituir a minimização do Lagrangeano [condição (1)] por sua estacionariedade [condição (a)].

No caso em que $S = E^{n+}$, temos o conhecido resultado para a condição (a)

$$\left\{ \begin{array}{l} \nabla_x L(x^0, u^0) \geq 0 \\ x^0 \cdot \nabla_x L(x^0, u^0) = 0 \\ x^0 \geq 0 \end{array} \right.$$

que se apresenta análogo às condições (b) (c) (d) relativas a u , à exceção de que aqui é um mínimo restrito enquanto lá é um máximo.

Geometricamente, a interpretação das condições (KT) no caso $S = E^n$ diz que $-\nabla f(x^0)$ é uma combinação linear positiva dos $\nabla g_i(x^0)$ para $i \in I$, ou seja, deve pertencer ao cone gerado pelo gradiente das restrições ativas.



Vejamos agora a relação entre as condições (KT) e a função de perturbação. Para isso seja x^* um ótimo do problema original, é claro que

$$x^* = x^*(y)$$

A função de perturbação é por definição o valor ótimo do problema perturbado, ou seja

$$v(y) = f(x^*) = f(x^*(y))$$

tendo como gradiente

$$\nabla v(y) = \nabla f(x^*) \frac{\partial x^*}{\partial y}$$

Mas como x^* é ótimo, pelas condições (KT) podemos escrever

$$\nabla f(x^*) = - u^* \nabla g(x^*)$$

onde u^* é o vetor multiplicador de Kuhn - Tucker (VMKT). Assim sendo temos

$$\nabla v(y) = - u^* \nabla g(x^*) \frac{\partial x^*}{\partial y}$$

Seja agora por definição a função

$$h(y) = g(x^*) = g(x^*(y))$$

que tem como gradiente

$$\nabla h(y) = \nabla g(x^*) \frac{\partial x^*}{\partial y}$$

É claro que sendo I o conjunto das restrições ativas

$$h_I(y) = Y_I$$

$$h_J(y) < Y_J$$

Expandindo em Taylor a função h_I , supondo que o conjunto de restrições ativas não se altere e que o termo de primeira ordem predomina sobre os de ordem superior

$$h_I(y + \delta y) = h_I(y) + \nabla h_I \delta y = Y_I + \delta Y_I$$

ou seja

$$\nabla g_I(x^*) \frac{\delta x^*}{\delta y} \cdot \delta y = \delta y_I$$

Fazendo a partição da expressão acima em I e J temos

$$\nabla g_I(x^*) \frac{\delta x^*}{\delta y_I} \delta y_I + \nabla g_I(x^*) \frac{\delta x^*}{\delta y_J} \delta y_J = \delta y_I$$

e como $\frac{\delta x^*}{\delta y_J} = 0$ pois a solução não varia para perturbações infinitesimais nas restrições não ativas temos finalmente

$$\nabla g_I(x^*) \frac{\delta x^*}{\delta y_I} = \mathbf{1}$$

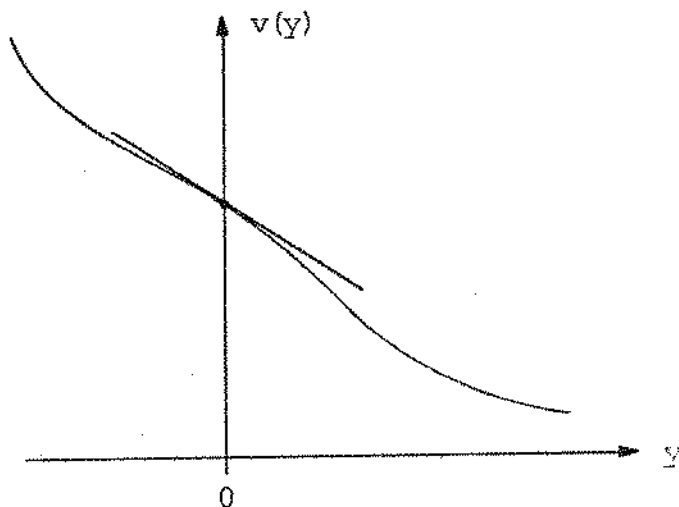
Voltando agora a expressão de $\nabla v(y)$ podemos escrevê-la fazendo a partição em I e J.

$$\nabla v(y) = - \left(u_I^* \mid u_J^* \right) \begin{pmatrix} \nabla g_I(x^*) \\ - \nabla g_J(x^*) \end{pmatrix} \begin{pmatrix} \frac{\delta x^*}{\delta y_I} \\ \frac{\delta x^*}{\delta y_J} \end{pmatrix}$$

e usando os resultados anteriores em conjunto com o fato de que $u_J^* = 0$ pelas condições (KT) temos afinal

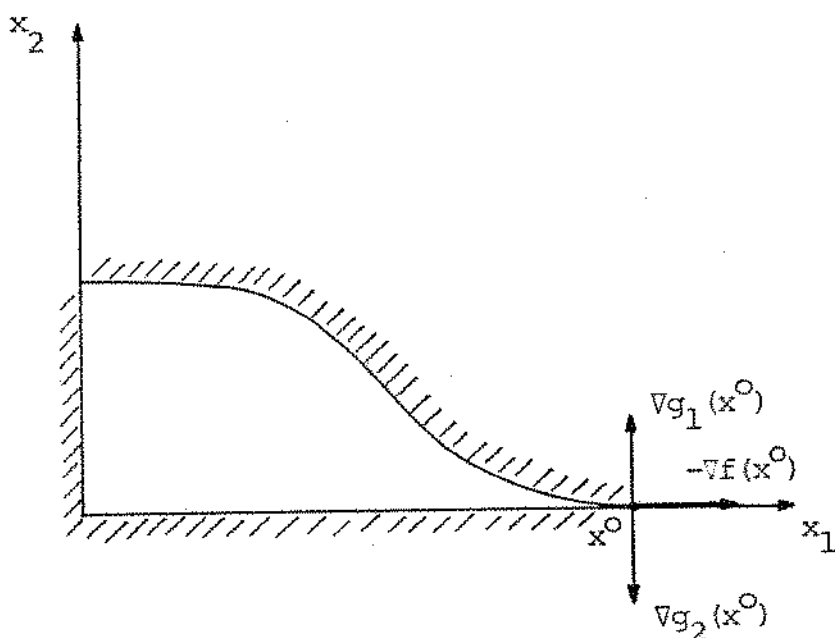
$$\nabla v(y) = - u^*$$

Com este resultado concluímos que a inclinação da função de perturbação mesmo quando não define hiperplano suporte é dada pelo (VMKT) (desde que f e g sejam de classe C_1)



(3) - (P) não convexo, estável.
Existe (VMKT)

Cabe observar que o desenvolvimento anterior é válido sob a hipótese geral de que $v(y)$ seja de classe C_1 . E ademais que certas condições de regularidade sejam satisfeitas, evitando casos patológicos onde os multiplicadores não são finitos, como no exemplo abaixo



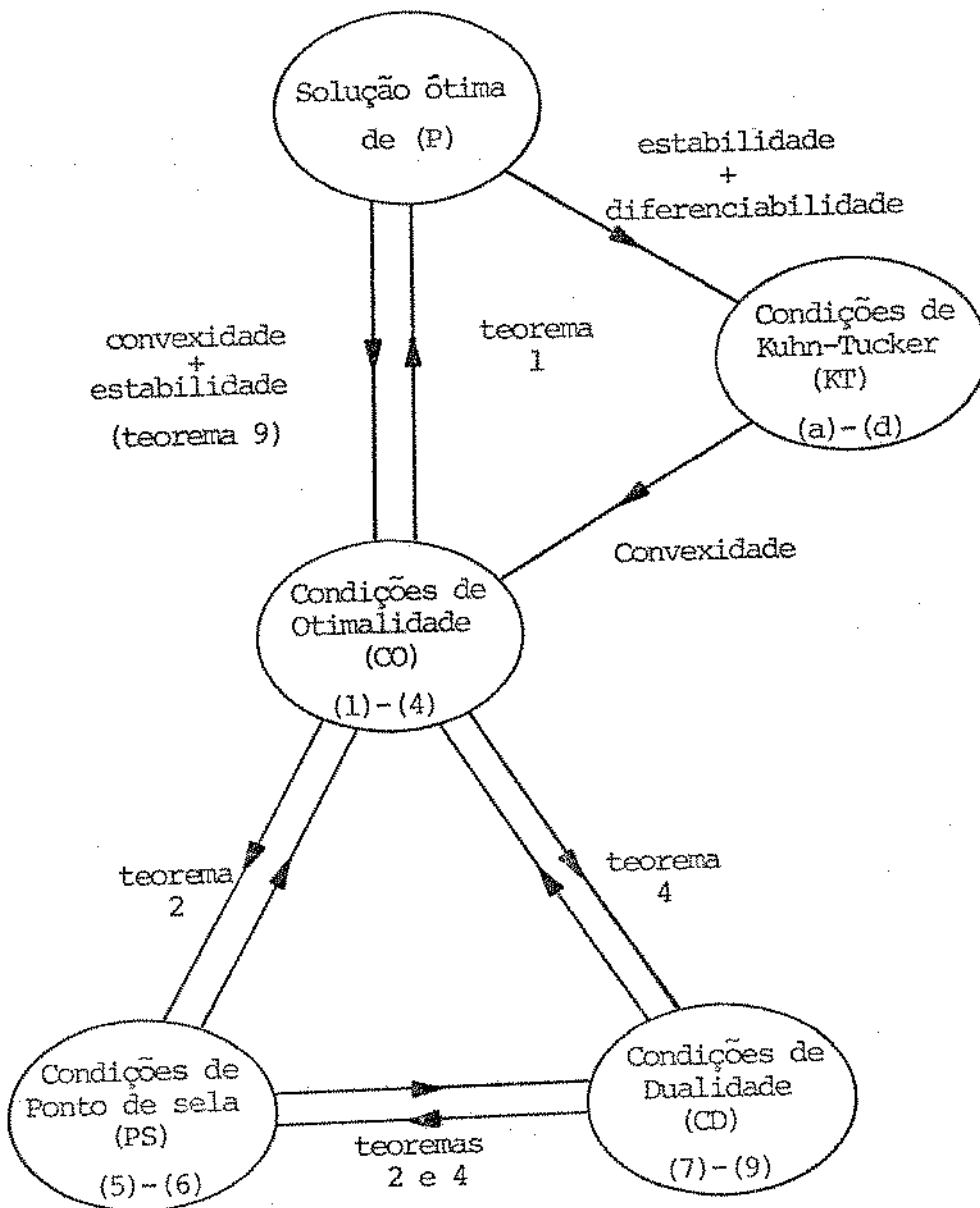
Diversas qualificações de restrição foram propostas para evitar estes casos. Algumas delas são:

- (1) Se existir um vetor Z tal que $Z \nabla g_i(x^0) < 0$ para todas as restrições ativas.
- (2) Se as restrições $g(x) \leq 0$ forem convexas e existir um ponto interior em S tal que $g(x) < 0$.
- (3) Se os gradientes das restrições ativas forem linearmente independentes em x^0 .
- (4) Se todas as restrições forem lineares.

Geoffrion [2] mostra que estabilidade é também uma qualificação de restrição, porém mais importante que as demais pois além de suficiente para garantir a não ocorrência de casos patológicos ela também é necessária.

Cabe finalmente salientar que as condições (KT) sendo satisfeitas em todo mínimo local serão uma condição suficiente de otimalidade se o problema (P) for convexo, pois neste caso admite um único mínimo, o global.

O seguinte quadro resume os principais resultados do cap. I.



CAPÍTULO II

"MÉTODOS DE DECOMPOSIÇÃO E COORDENAÇÃO"

O objetivo deste capítulo é representar de forma unificada alguns dos mais importantes métodos de otimização para programas matemáticos de grande porte. Nos anos recentes a complexidade cada vez maior dos problemas tratados pela programação matemática estimulou o aparecimento tanto de técnicas preocupadas com o aumento da eficiência computacional de métodos tradicionais (memória e tempo de processamento), cujo exemplo característico é o método simplex, como também de novos métodos onde a tônica central tem sido a decomposição.

Os métodos de decomposição e coordenação, segundo a interessante visão de Geoffrion [5], podem ser compreendidos como a junção de uma manipulação que visa transformar o problema original num problema equivalente (programa mestre) onde se evidenciam simplificações advindas da estrutura original do problema, e uma estratégia de resolução que tirando proveito dessas simplificações fornece um procedimento iterativo para a solução do programa mestre. A decomposição em subproblema de menor dimensão e passíveis de solução por técnicas conhecidas é o resultado frequente das estratégias de resolução.

Os métodos de decomposição pelo Lagrangeano são um exemplo. Restrições de acoplamento que sendo ausentes tornariam o problema original separável são afastadas via dualização e o problema Lagrangeano associado se torna assim decomponível.

Nos métodos estudados neste capítulo serão vistas as manipulações intituladas dualização, linearização (interna e externa), projeção e as estratégias de resolução chamadas otimização por partes, restrição, relaxação, direções factíveis [5].

O que se quer mostrar, sobretudo, é que a partir da interpretação desses diferentes métodos geometricamente sobre a função de perturbação é possível aclarar seus mecanismos, compará-los em termos das suas especificidades, unificá-los naquilo que têm em comum.

MÉTODO 1 - DUAL CLÁSSICO

O método dual clássico é o mais simples dos métodos de decomposição do Lagrangeano consistindo na aplicação da combinação Dualização/Direções factíveis como manipulação e estratégia de resolução. Um dos trabalhos precursores desse método é devido a Uzawa [6], tendo sido mais tarde reestudado por Lasdon [7] que forneceu interessantes interpretações econômicas.

Para sua apresentação consideremos o problema primal

$$(P) \quad \begin{array}{l} \min f(x) \\ \text{s.a } g(x) \leq 0 \\ x \in S \end{array}$$

Dualizando (P) em relação às restrições $g(x) \leq 0$ obtem-se

$$(D) \quad \begin{array}{l} \max \{h(u) = \text{ínfimo}_{x \in S} f(x) + ug(x)\} \\ u \geq 0 \end{array}$$

É possível mostrar [1], para S compacto e f, g contínuas em S, que $g(x^*)$ é um subgradiente de $h(u)$ em u^* para todo $x^* \in X(u^*)$. Em particular se $g(x^*)$ for constante sobre $X(u^*)$ então $h(u)$ é diferenciável em u^* e

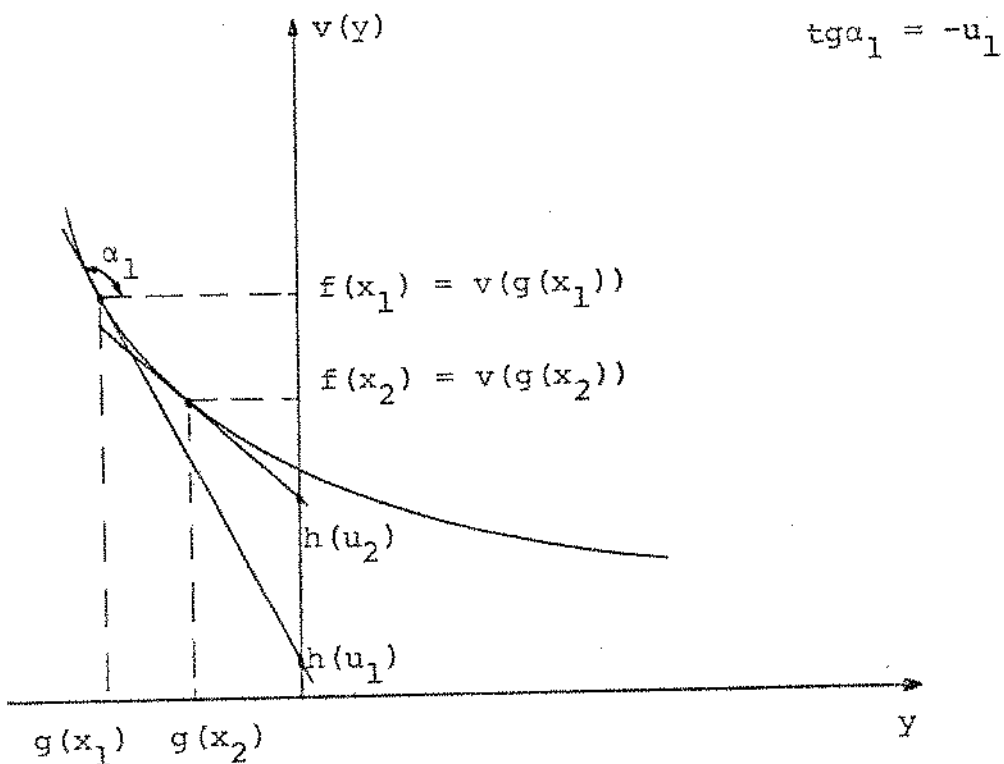
$$\nabla h(u^*) = g(x^*) \quad , \quad x^* \in X(u^*)$$

A partir desta constatação, a saber, que o gradiente da função dual é dado pela restrição, o método usa uma estratégia de resolução por direções factíveis para resolver (D).

Resumo do procedimento:

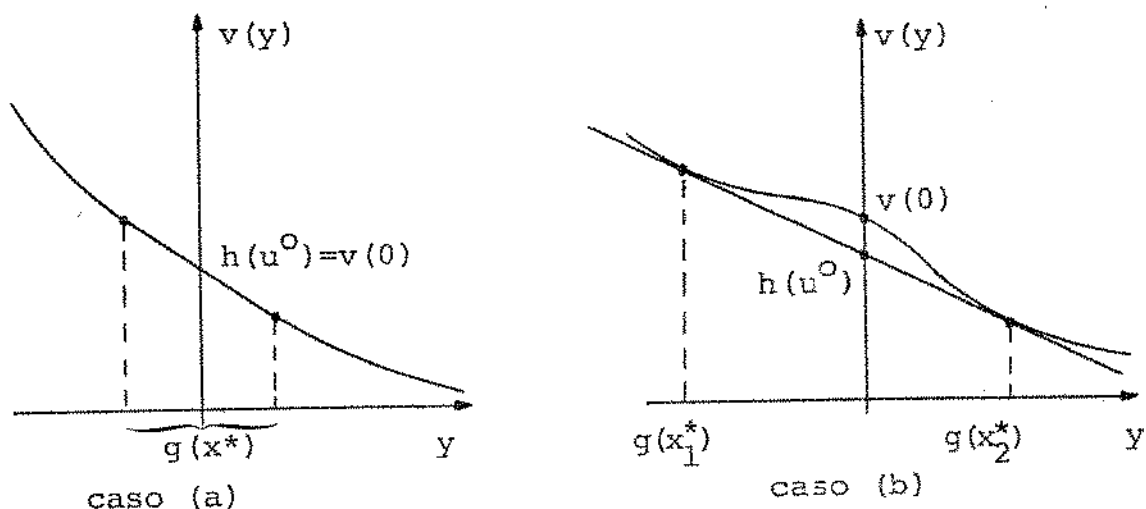
- (1) Escolha $u^* \geq 0$
- (2) Minimize $f(x) + u^* g(x)$ sobre S . Seja x^* uma solução
- (3) Se $g(x^*) \leq 0$ e $u^* g(x^*) = 0$, pare: x^* é solução ótima de (P).
Caso contrário faça $u^* = \max \{ 0, u^* + \alpha g(x^*) \}$ com $\alpha > 0$ (escalar pequeno) e volte a (2).

Geometricamente na função de perturbação



Para problemas convexos onde ademais o Lagrangeano é estritamente convexo em $x \in S$ para u^0 solução ótima de (D), o critério de otimalidade em (3) garante a convergência do método para a solução ótima de (P). Em outros casos porém é conveniente substituir o critério de parada em (3) pela maximização da função dual $h(u)$. Segundo este novo critério a convergência é garantida em todos os

casos podendo, entretanto, ocasionar dificuldades adicionais. Isto porque tanto para problemas convexos porém não estritamente (a) como para problemas não convexos com "gap de dualidade" (b) o mínimo do Lagrangeano para u^0 solução ótima de (D) é múltiplo e não g -constante.



Permanece assim a dificuldade de encontrar a solução ótima de (P) a partir da solução ótima de (D) fornecida pelo método. A distinção entre os casos (a) e (b) fica prejudicada. O que se pode garantir é que x^0 solução ótima de (P) no caso (b) não pertence a $X(u^0)$ e no caso (a) sim, sendo porém necessária sua determinação. A programação linear é um exemplo clássico do caso (a) onde portanto se utilizam métodos mais elaborados que evitam o problema da determinação de x^0 e $X(u^0)$ ao mesmo tempo em que aproveitam a estrutura linear do problema. A estrita convexidade de $f(x)$, ao garantir a unicidade do mínimo do Lagrangeano, evita o caso (a).

MÉTODO 2 - PROGRAMAÇÃO LINEAR GENERALIZADA

Desenvolvida por Dantzig e Wolfe [8] a PLG constitui um dos principais resultados da programação matemática de grande porte. É o resultado da combinação Linearização Interna/Restrição como manipulação e estratégia de resolução.

O problema primal convexo

$$(P) \quad \begin{array}{l} \min f(x) \\ \text{s.a. } g(x) \leq 0 \\ \\ x \in S \end{array}$$

é linearizado internamente por meio de uma base de linearização $\{x^j\}$ $x^j \in S \forall j$, em princípio infinita para garantir a equivalência com (P). Resulta então o programa mestre linear

$$(PM) \quad \begin{array}{l} \min \sum_j \lambda_j f(x^j) \\ \text{s.a. } \sum_j \lambda_j g(x^j) \leq 0 \\ \\ \sum_j \lambda_j = 1 \\ \\ \lambda_j \geq 0, \quad \forall j \end{array}$$

A aplicação da estratégia de restrição para resolver (PM) consiste em fixar $\lambda_j = 0 \quad \forall j \neq 1, 2, \dots, r$, obtendo-se assim um programa linear restrito (base de linearização finita)

(PL)_r

$$\min \sum_{j=1}^r \lambda_j f(x^j)$$

$$\text{s.a. } \sum_{j=1}^r \lambda_j g(x^j) \leq 0$$

$$\sum_{j=1}^r \lambda_j = 1$$

$$\lambda_j \geq 0, \quad j = 1, 2, \dots, r$$

Associado à solução ótima de (PL)_r existe um VMO $u^{r+1} \geq 0$ que é utilizado como parâmetro do problema Lagrangeano

$$\min \{f(x) + u^{r+1} g(x)\}$$

$$x \in S$$

cuja solução x^{r+1} é então introduzida na base de linearização melhorando a aproximação entre (PL)_{r+1} e (PM).

Resumo do procedimento:

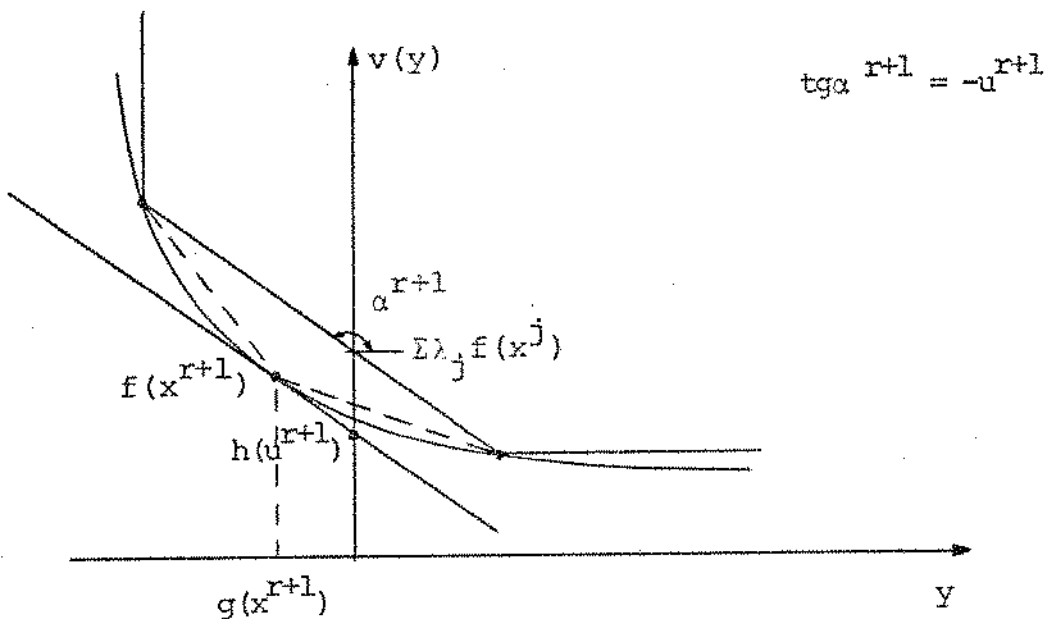
- (1) Escolha uma base de linearização $\{x^1, x^2, \dots, x^r\}$ inicial.
- (2) Resolva (PL)_r. Seja $\{\lambda_1, \lambda_2, \dots, \lambda_r\}$ a solução ótima e u^{r+1} o vetor multiplicador ótimo.
- (3) Minimize $f(x) + u^{r+1} g(x)$ sobre S. Seja x^{r+1} uma solução.

(4) Se $h(u^{r+1}) = \sum_{j=1}^r \lambda_j f(x^j)$, pare: $\sum_{j=1}^r \lambda_j x^j$ é solu

ção ótima de (P)

Caso contrário introduza x^{r+1} na base de linearização e volte a (2).

Geometricamente na função de perturbação



Em termos da função de perturbação a PLG consiste numa linearização interna de $v(y)$, a cada passo mais precisa em torno de $y=0$. Comparativamente ao método dual clássico a diferença fundamental reside no modo de determinar a sequência de vetores multiplicadores fornecidos ao problema Lagrangeano, que lá eram obtidos pelo conhecimento do gradiente (ou subgradiente) da função dual ao passo que aqui são o resultado da solução do problema linearizado.

É importante salientar que para problemas não convexos com "gap de dualidade", caso (b), a PLG também não é aplicável pois pode convergir para uma solução infactível. Para o caso (a), entretanto, ao contrário do método anterior, não ocorre dificuldade em se obter a solução primal ótima uma vez conhecida solução ótima dual, pois $(PL)_r$, fornece a cada iteração uma solução primal factível $(\sum_{j=1}^r \lambda_j x^j)$ cada vez mais próxima do ótimo. Assim, a PLG se apli

ca a problemas convexos quaisquer, em particular a programas lineares (na verdade originalmente tratou da programação linear [8]). Aliás, se (P) for linear, como a linearização é exata para uma base finita, o algoritmo converge em número finito de passos.

MÉTODO 3 - APROXIMAÇÃO TANGENCIAL

O método de aproximação tangencial de Geoffrion [9] é um procedimento dual da PLG consistindo na combinação Dualização - Linearização externa/Relaxação como manipulação e estratégia de resolução.

O problema convexo primal

$$(P) \quad \begin{array}{l} \min f(x) \\ \text{s.a. } g(x) \leq 0 \\ x \in S \end{array}$$

é dualizado obtendo-se o problema

$$(D) \quad \begin{array}{l} \max \left\{ h(u) = \inf_{x \in S} [f(x) + ug(x)] \right\} \\ u \geq 0 \end{array}$$

A seguir a função dual é linearizada externamente sendo representada por

$$h^r(u) = \inf_j \left\{ f(x^j) + u g(x^j) \right\}$$

onde x^j é solução do problema Lagrangeano para $u^j \geq 0$

$$\begin{array}{l} \min f(x) + u^j g(x) \\ x \in S \end{array}$$

A partir da representação linearizada da função dual, e lembrando que infimo é o maior limitante inferior, o problema (D) é equivalente ao problema

$$\begin{array}{l} \max z^0 \\ u, z^0 \\ \text{s.a. } z^0 \leq f(x^j) + u g(x^j) \quad x^j \in X(u^j), \forall j/u^j \geq 0 \end{array}$$

(PM)

A estratégia de resolução para (PM) é relaxação, ou seja, resolver a cada iteração o problema relaxado $(PR)_r$

$$\begin{array}{l} \max z^0 \\ u, z^0 \\ \text{s.a. } z^0 \leq f(x^j) + u g(x^j) \quad x^j \in X(u^j), u^j \geq 0, j=1,2,\dots,r \end{array}$$

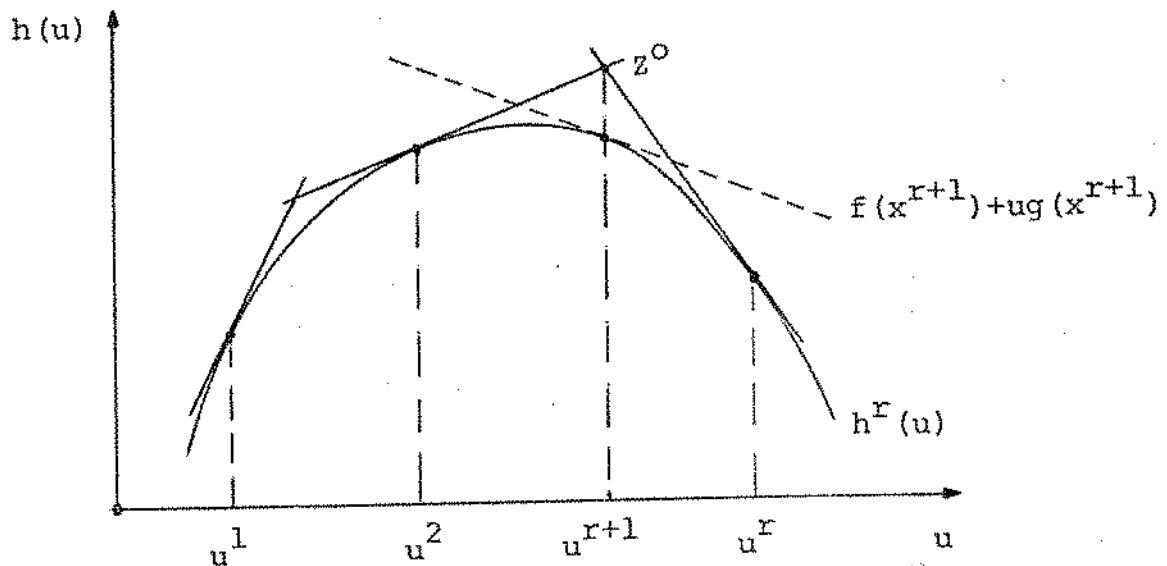
$(PR)_r$

Resumo do procedimento:

- (1) Seja $\{x^j\}$ tal que $x^j \in X(u^j)$, $u^j \geq 0$, $j = 1, 2, \dots, r$
- (2) Resolver $(PR)_r$. Seja u^{r+1} a solução ótima e z^0 o valor ótimo.
- (3) Minimizar $f(x) + u^{r+1} g(x)$ sobre S. Seja x^{r+1} uma solução

- (4) Se $h(u^{r+1}) = z^0$, pare: x^{r+1} é solução ótima de (P)
 Caso contrário introduza a restrição $z^0 \leq f(x^{r+1}) + u g(x^{r+1})$
 em $(PR)_{r+1}$ e volte a (2).

Geometricamente na função dual



A linearização externa de $h(u)$ no método da aproximação tangencial é equivalente à linearização interna de $v(y)$ na PLG. Para se ver isso seja o Lagrangeano associado a $(PL)_r$.

$$L(x, u, z^0) = \sum_{j=1}^r \lambda_j f(x^j) + u \sum_{j=1}^r \lambda_j g(x^j) + z^0 (1 - \sum_{j=1}^r \lambda_j)$$

que pode ser escrito

$$L(x, u, z^0) = \sum_{j=1}^r (f(x^j) + u g(x^j) - z^0) \lambda_j + z^0$$

A função dual associada a $L(x, u, z^0)$ será então

$$h(u, z^0) = \min_{\lambda \geq 0} L(x, u, z^0) = z^0$$

satisfeita a condição de que o mínimo do Lagrangeano exista, isto é,

$$f(x^j) + u g(x^j) - z^0 \geq 0 \quad j = 1, 2, \dots, r$$

$$u \geq 0$$

Assim o problema dual de $(PL)_r$ se escreve

$$\left\{ \begin{array}{l} \max \quad z^0 \\ u, z^0 \\ \text{s.a.} \quad z^0 \leq f(x^j) + u g(x^j) \quad j = 1, 2, \dots, r \\ u \geq 0 \end{array} \right.$$

que nada mais é que $(PR)_r$. Com isso a sequência de multiplicadores nos dois métodos será a mesma.

A equivalência entre PLG e aproximação tangencial leva a que se estendam ao segundo as observações feitas ao primeiro, em relação à convergência no caso de "gap de dualidade" e no caso linear. A título de observação, aqui a solução ótima de (P) ao fim das iterações é obtida por combinação convexa dos x^j com pesos λ_j dados pelas variáveis duais associadas a $(PR)_r$.

MÉTODO 4 - ALOCAÇÃO DE RECURSOS

Também proposto por Geoffrion [9] e estudado particularmente em Lasdon [1], o método de alocação de recursos pode ser interpretado como o oposto dos métodos "duais". Nestes últimos os preços u direcionam a utilização dos recursos no sentido de factibilizar o problema, ao passo que aqui a alocação dos recursos será feita diretamente (sendo portanto sempre factível) e os preços (variáveis duais) servirão apenas para avaliar a melhoria proveniente de

novas alocações alternativas. Em termos de manipulação e estratégia de resolução este método pode ser visto como a combinação Projeção - Linearização externa/Relaxação.

Seja o problema de programação matemática

$$(P) \quad \begin{array}{l} \min \quad \sum_{i=1}^r f_i(x_i) \\ \\ \text{s.a.} \quad \sum_{i=1}^r g_i(x_i) \leq b \\ \\ x_i \in S_i \quad i = 1, 2, \dots, r \end{array}$$

cuja estrutura é bastante frequente em problemas de grande porte (critério separável, restrições separáveis e de acoplamento). Um caso típico é a programação linear com matriz de coeficientes angular. Ao invés de dualizar em relação às restrições de acoplamento $\sum_{i=1}^r g_i(x_i) \leq b$ (método "dual") projeta-se o problema no espaço por elas definido.

Introduzindo variáveis y_i , $i = 1, 2, \dots, r$ (P) pode ser reescrito

$$\begin{array}{l} \min \quad \sum_{i=1}^r f_i(x_i) \\ \\ \text{s.a.} \quad \sum_{i=1}^r y_i \leq b \end{array}$$

$$(P') \quad \left. \begin{array}{l} g_i(x_i) \leq y_i \\ x_i \in S_i \end{array} \right\} i=1, 2, \dots, r$$

e projetando (P') no espaço dos y temos o programa mestre (PM) equivalente

$$(PM) \quad \begin{array}{l} \min \sum_{i=1}^r v_i(y_i) \\ \text{s.a.} \sum_{i=1}^r y_i \leq b \end{array}$$

onde

$$(SP)_i \quad v_i(y_i) = \min f_i(x_i) \quad \text{s.a.} \quad g_i(x_i) \leq y_i, \quad x_i \in S_i$$

Resumo do procedimento:

- (1) Escolher uma alocação de recursos $\{\bar{y}_i^{-k}\}$ inicial
- (2) Resolver $(SP)_i$, $i = 1, 2, \dots, r$. Sejam $\{\bar{x}_i^{-k}\}$ as soluções ótimas e $\{\bar{u}_i^{-k}\}$ os vetores multiplicadores ótimos
- (3) Em função dos $\{\bar{u}_i^{-k}\}$ escolher nova alocação $\{\bar{y}_i^{-k+1}\}$ de modo a melhorar (PM) e volte a (2) até que (PM) seja resolvido.

Uma possível maneira de executar (3) é baseada na visualização de $(SP)_i$ como uma função de perturbação onde \bar{u}_i^{-k} é o vetor multiplicador associado a $v_i(\bar{y}_i^{-k})$ que fornece uma medida (otimista) da taxa marginal de redução em $v_i(y_i)$ para um acréscimo em y_i

a partir de \bar{y}_i^k . Assim, a comparação entre os $\bar{u}_i^k \cdot \bar{y}_i^k$ de cada $(SP)_i$ permite direcionar a nova alocação.

Representando $v_i(y_i)$ por sua linearização externa, ou seja

$$v_i(y_i) = \supremo_j \left\{ v_i(\bar{y}_i^j) - \bar{u}_i^j (y_i - \bar{y}_i^j) \right\}$$

e aplicando relaxação, o programa mestre (PM) fica

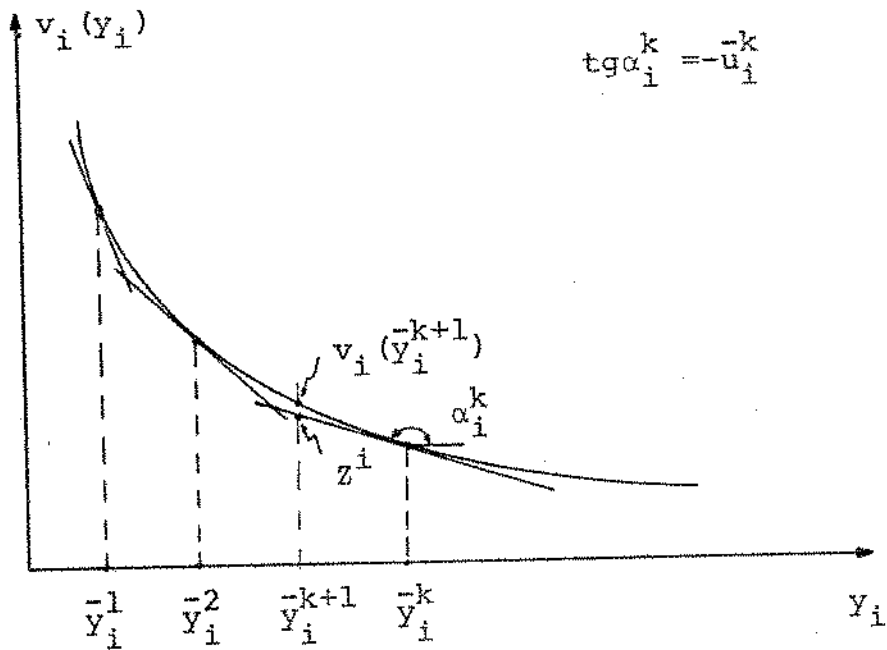
$$\begin{aligned} \min \quad & \sum_{i=1}^r \supremo_{1 \leq j \leq k} \left\{ v_i(\bar{y}_i^j) - \bar{u}_i^j (y_i - \bar{y}_i^j) \right\} \\ \text{s.a.} \quad & \sum_{i=1}^r y_i \leq b \end{aligned}$$

ou finalmente como o supremo é o menor dos limitantes superiores

$(PM)_k$

$$\begin{aligned} \min \quad & \sum_{i=1}^r z^i \\ \text{s.a.} \quad & \sum_{i=1}^r y_i \leq b \\ & z^i \geq v_i(\bar{y}_i^j) - \bar{u}_i^j (y_i - \bar{y}_i^j) \quad \begin{array}{l} i = 1, 2, \dots, r \\ j = 1, 2, \dots, k \end{array} \end{aligned}$$

Graficamente na função de perturbação



O procedimento global poderia ser resumido em:

- (1) Escolher uma alocação de recursos $\left\{ \bar{y}_i^1 \right\}$ inicial
- (2) Resolver (SP)_i, $i = 1, 2, \dots, r$. Seja $\left\{ \bar{x}_i^1 \right\}$ a solução e $\left\{ \bar{u}_i^1 \right\}$ os vetores multiplicadores ótimos. Fazer $k = 1$
- (3) Resolver (PM)_k. Seja $\left\{ \bar{y}_i^{-k+1} \right\}$ a solução e z^i os valores ótimos
- (4) Resolver (SP)_i, $i = 1, 2, \dots, r$ para a alocação $\left\{ \bar{y}_i^{-k+1} \right\}$. Seja $\left\{ \bar{x}_i^{-k+1} \right\}$ as soluções ótimas e $\left\{ \bar{u}_i^{-k+1} \right\}$ os vetores multiplicadores ótimos.
- (5) Se $v_i(\bar{y}_i^{-k+1}) = z^i$, $i = 1, 2, \dots, r$, pare: $\left\{ \bar{x}_i^{-k+1} \right\}$ é solução ótima de (P)

Caso contrário introduza em $(PM)_{k+1}$ as restrições adicionais

$$v_i^{(-k+1)} - u_i^{(-k+1)} (y_i - \bar{y}_i^{(-k+1)}) \leq z^i, \quad i = 1, 2, \dots, r \quad e$$

volte a (3).

Uma característica importante desse método é apresentar uma solução factível a cada passo, portanto utilizável do ponto de vista prático, ao contrário dos métodos "duais" infactíveis.

Do ponto de vista teórico, para que o procedimento seja viável é necessário que as funções $f_i(x_i)$ e $g_i(x_i)$ sejam convexas sobre S_i convexo (e compacto) de modo a validar a linearização externa de $v_i(y_i)$. Além disso é preciso definir o conjunto de alocações de recursos factíveis (não citado na apresentação por simplificação).

MÉTODO 5 - DECOMPOSIÇÃO DE BENDERS

Proposto por Benders [10] e analisado por Lasdon [1] e Geoffrion [5], este método se mostra particularmente interessante para problemas de programação mista. Consiste na combinação Projeção - Linearização externa/Ralaxação como manipulação e estratégia de resolução. Para sua apresentação trataremos o seguinte problema semi-linear.

(P)	Min	$cx + f(y)$
	s.a	$Ax + F(y) \leq b$
		$x \geq 0, y \in Y$

Para aproveitar a estrutura linear em x projetamos (P) sobre o espaço dos y obtendo

$$\text{Min}_{y \in V} \left\{ f(y) + \text{infimo} \left[cx \text{ s.a. } Ax \leq b - F(y) \right] \right\}$$
$$x \geq 0$$

onde

$$V = \left\{ y \in Y / x \geq 0 \text{ com } Ax \leq b - F(y) \right\}$$

O infimo dentro do minimizando é o valor ótimo do programa linear parametrizado

(PLP)

$$\begin{array}{ll} \text{Min} & cx \\ \text{s.a.} & Ax \leq b - F(y) \\ & x \geq 0 \end{array}$$

cujo dual é

(DLP)

$$\begin{array}{ll} \text{Max} & u(b - F(y)) \\ \text{s.a.} & uA \leq c \\ & u \leq 0 \end{array}$$

Sejam $\{u^j\}$ $j = 1, 2, \dots, p$ os pontos extremos e $\{u^j\}$ $j = p + 1, \dots, p+q$ os raios extremos do poliedro

$$S = \{u \leq 0 / uA \leq c\}$$

Se S é vazio (P) não tem solução factível ou tem valor ótimo ilimitado. Afastando estes casos triviais pode-se garantir pela

dualidade em programação linear que se (DLP) tiver solução ilimitada para algum y então (PLP) será infactível. Assim podemos caracterizar o conjunto V como

$$V = \{y \in Y \ / \ u^j(b - F(y)) \leq 0 \quad j = p+1, \dots, p+q\}$$

Por outro lado linearizando externamente o infimo de (PLP), ou seja, lembrando pelo teorema da dualidade em programação linear que seu valor é

$$\begin{aligned} & \text{Max } \{u^j(b - F(y))\} \\ & 1 \leq j \leq p \end{aligned}$$

e sabendo que o máximo é o menor limitante superior temos finalmente o programa mestre equivalente

(PM)	Min	$f(y) + y^0$	
		$y \in Y$	
		y^0	
	s.a.	$y^0 \geq u^j(b - F(y))$	$j = 1, 2, \dots, p$
	$0 \geq u^j(b - F(y))$	$j = p+1, \dots, p+q$	

É claro que relaxação é uma estratégia natural para resolver (PM) pois evita determinar antecipadamente os vetores $\{u^j\}$ $j = 1, 2, \dots, p+q$. O problema relaxado fica pois

(PR)	Min	$f(y) + y^0$	
		$y \in Y$	
		y^0	
	s.a.	$y^0 \geq u^j(b - F(y))$	para alguns $1 \leq j \leq p$
	$0 \geq u^j(b - F(y))$	para alguns $p+1 \leq j \leq p+q$	

Resumo do procedimento:

- (1) Resolva (PR). Seja (\bar{y}, \bar{y}^0) a solução.
- (2) Resolva (DLP) com $y = \bar{y}$

Se o valor ótimo for igual a \bar{y}^0 , pare : (\bar{y}, \bar{y}^0)

é factível em (PM) e portanto ótimo. A solução ótima de de (P) será (\bar{y}, \bar{x}) onde \bar{x} é o vetor multiplicador de (DLP).

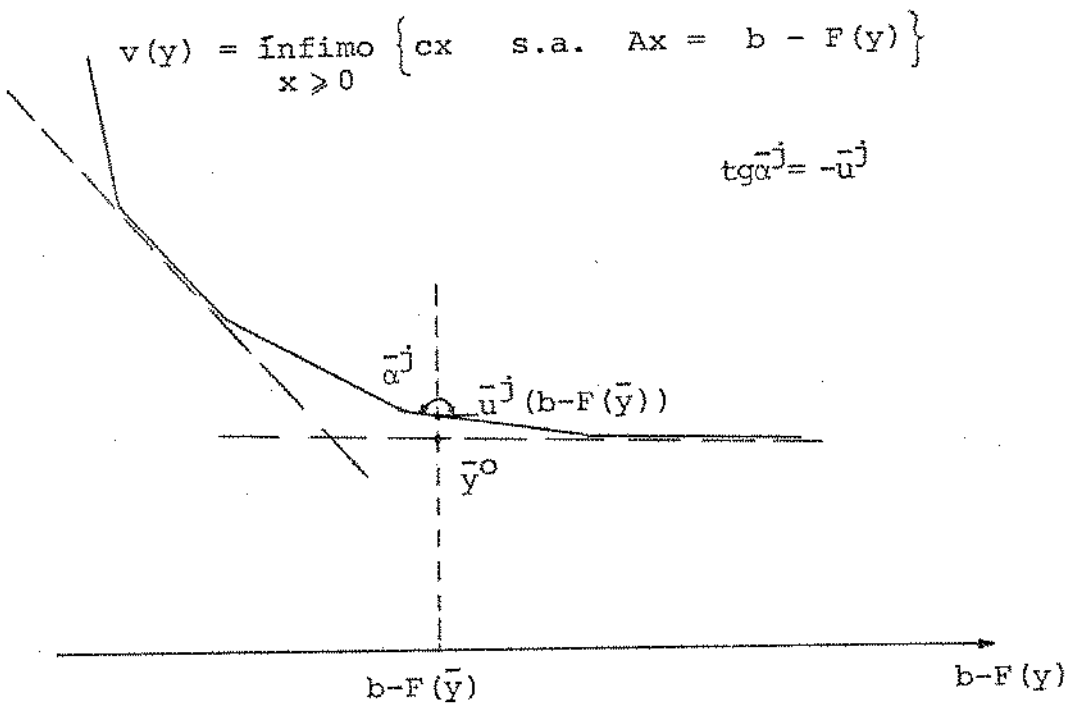
Se o valor ótimo for superior a \bar{y}^0 então uma restrição violada de (PM) é produzida do tipo:

$y^0 \geq u^j(b - F(y))$ - se o valor ótimo for finito (onde u^j é o ponto extremo solução) ou

$0 \geq u^j(b - F(y))$ - se o valor ótimo for ilimitado (onde u^j é o raio extremo).

Voltar a (1)

Graficamente na função de perturbação de (PLP)



A curva cheia representa a linearização externa da função de perturbação de (PLP). A curva tracejada é a sua relaxação. A solução de (PR) fornece um par (\bar{y}, \bar{y}^0) . Se a solução de (PLP) para $y = \bar{y}$ (obtida através de seu dual) fornecer um valor ótimo não superior a \bar{y}^0 , o processo iterativo termina. Caso contrário melhora-se a acuridade do (PR) introduzindo uma nova restrição: $y^0 \geq \bar{u}^j (b - F(y))$

Isto, é claro, quando (PLP) tem solução factível pois caso contrário a restrição a ser introduzida é $0 \geq \bar{u}^j (b - F(y))$ que reduz o domínio efetivo dos y .

As hipóteses de convexidade de $f(y)$, $F(y)$ e Y garantem o procedimento ao assegurar que (P) e (PM) serão problemas convexos. Entretanto mesmo em alguns casos, programação mixta por exemplo, o método é ainda válido.

O método de Benders aqui apresentado pode ser estendido permitindo não linearidades também em x . Isto foi desenvolvido por Geoffrion [11] com o auxílio da teoria da dualidade não linear.

MÉTODO 6 - ROSEN

O método de Rosen [12] pode ser compreendido como a conjugação Dualidade / Otimização por partes enquanto manipulação e estratégia de resolução. Para ilustrar seja o problema linear com estrutura angular

(P)

$$\begin{array}{l} \text{Min} \quad \sum_{i=1}^p C_i x_i \\ \text{s.a} \quad \sum_{i=1}^p A_i x_i = b_0 \\ \quad \quad \quad \left. \begin{array}{l} B_i x_i = b_i \\ x_i \geq 0 \end{array} \right\} \quad i = 1, 2, \dots, p \end{array}$$

Dualizando em relação à restrição de acoplamento obtemos o Lagrangeano

$$L(x, \lambda) = b_0 \cdot \lambda + \sum_{i=1}^p (C_i - \lambda A_i) x_i$$

e o problema dual resultante será

(D)

$$\text{Max}_{\lambda} \left\{ b_0 \cdot \lambda + \sum_{i=1}^p \inf_{x_i \in S_i} (C_i - \lambda A_i) x_i \right\}$$
$$\text{com } S_i = \left\{ x_i \geq 0 \ / \ B_i x_i = b_i \right\}$$

O infimo dentro do maximizando tem como valor o resultado do programa linear parametrizado.

(PLP)

$$\text{Min } (C_i - \lambda A_i) x_i$$

s.a.

$$B_i x_i = b_i$$
$$x_i \geq 0$$

Seja \bar{x}_i a solução ótima de (PLP). Para pequenas variações em λ a solução ótima de (PLP) não se altera e seu valor ótimo será dado por

$$(C_i - \lambda A_i) \bar{x}_i$$

Isto é válido enquanto o vetor de custo relativo permanecer não negativo. Matematicamente, sendo I a base ótima relativa a solução \bar{x}_i esta condição é expressa por

$$\hat{C} = (C_i - \lambda A_i) - (C_i - \lambda A_i)^I (B_i^I)^{-1} B_i \geq 0 \quad i = 1, 2, \dots, p.$$

Assim o problema (D) pode ser parcialmente representado pelo equivalente linear

$$\begin{array}{l}
 \text{Max}_{\lambda} \left\{ b_0 \cdot \lambda + \sum_{i=1}^p (C_i - \lambda A_i) \bar{x}_i \right\} \\
 \text{(PE)} \quad \text{s.a.} \\
 (C_i - \lambda A_i) - (C_i - \lambda A_i)^T (B_i^T)^{-1} B_i \geq 0 \quad i=1,2,\dots,p
 \end{array}$$

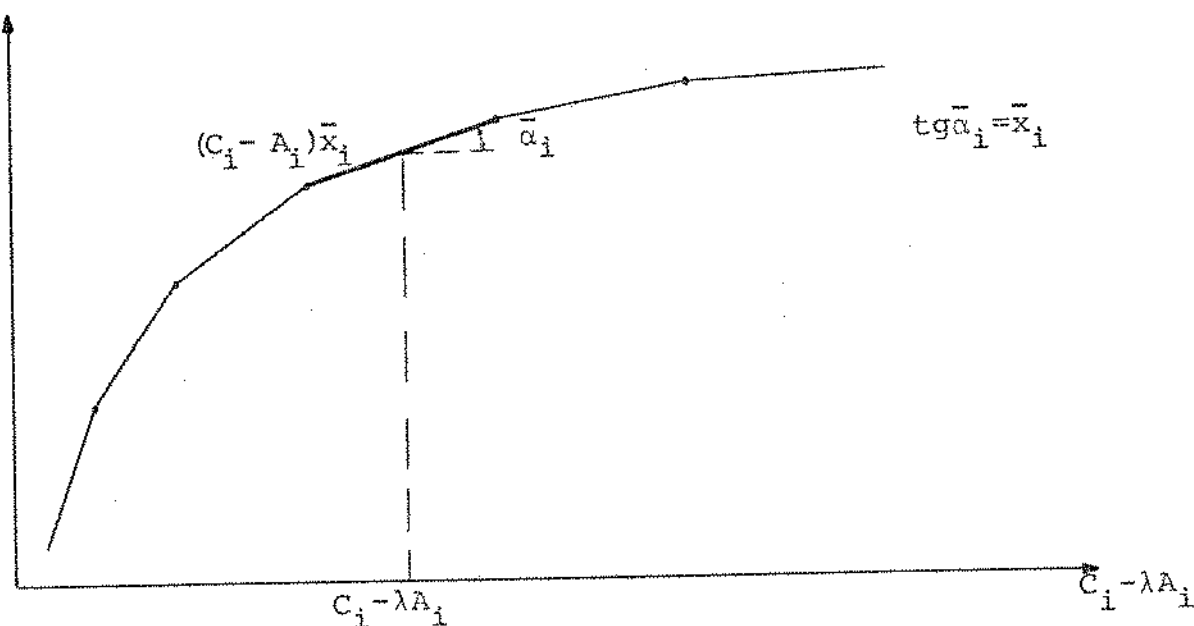
A solução de (PE) será um $\bar{\lambda}$ que saturará tantas restrições de (PE) quanto é a dimensão do vetor λ . Isto garante que $\bar{\lambda}$ será uma solução básica de (D).

A partir daí, escolhendo-se uma única restrição (das ativas) para relaxar, o que garante um procedimento tipo simplex, um novo (PE) é resolvido até que outra restrição se torne ativa.

Escrevendo o dual de (PLP)

$$\begin{array}{l}
 \text{Max } b_i u_i \\
 \text{(DLP)} \quad \text{s.a. } u_i B_i \leq C_i - \lambda A_i
 \end{array}$$

podemos interpretar o método na função de perturbação de (DLP)



O valor ótimo de cada subproblema é representado pelo trecho linear válido num certo politopo em λ . A restrição de (PE) define este politopo. A cada iteração um subproblema muda de trecho, alterando o conjunto restrição de (PE), até a otimização de (D).

Em problemas com estrutura dual angular o método se equivale a combinação Projeção / Otimização por partes. A decomposição de Rosen pode ser estendida para a classe particular de problemas não lineares, que constitui a generalização da estrutura dual angular, a saber

$$\left\{ \begin{array}{l} \text{Min} \quad \sum_{i=1}^P C_i(y) x_i + C_0(y) \\ \text{s.a.} \quad A_i(y) x_i \geq b_i(y) \quad i = 1, 2, \dots, P \end{array} \right.$$

CAPÍTULO III

"OTIMIZAÇÃO HIERÁRQUICA APLICADA A SISTEMAS DE
POTÊNCIA HIDROTÉRMICOS"

Neste capítulo estuda-se a aplicação de técnicas de otimização por decomposição e coordenação ao problema de operação de sistemas de potência hidrotérmicos. Consegue-se assim o tratamento de problemas de grande dimensão (várias unidades térmicas e hidroelétricas) com uma formulação rigorosa do modelo incluindo funções de geração hidroelétrica e custo de operação térmico não lineares, influência das variações no nível dos reservatórios sobre a potência hidroelétrica gerada, acoplamento hidroelétrico através de reservatórios em cascata com vários vales hidráulicos, demanda de carga aleatória.

Estruturas de cálculo em três níveis são obtidas e aplicadas a um exemplo ilustrativo.

LISTA DE SÍMBOLOS

- m: índice de tempo
- i: índice termoelétrico
- j: índice hidroelétrico
- M: conjunto de índices de tempo
- I: conjunto de índices termoelétricos
- J: conjunto de índices hidroelétricos
- g: geração de potência termoelétrica
- g, \bar{g} : limites de geração de potência termoelétrica
- ψ : função custo de operação termoelétrico
- h: geração de potência hidroelétrica
- ϕ : função geração de potência hidroelétrica
- x: estoque do reservatório
- x, \bar{x} : limites no estoque do reservatório
- y: entrada de água independente
- Z: entrada de água dependente
- u: turbinagem
- u, \bar{u} : limites de turbinagem
- v: vertimento
- S_j : conjunto de índices das hidroelétricas imediatamente vizinhas rio acima à hidroelétrica j
- T_j : conjunto de índices de todas hidroelétricas rio abaixo à hidroelétrica j no mesmo vale
- t_{kj} : tempo de percurso da água da hidroelétrica k até sua vizinha à jusante j
- d: demanda de carga
- F: função distribuição acumulada de d
- p: oferta de potência
- α , β : pesos

L, L', L'' : função Lagrangeano

λ, ρ : multiplicadores de Lagrange

H, H', H'' : função dual.

3.1 INTRODUÇÃO

O custo marginal desprezível da geração hidroelétrica torna a otimização de sistemas de potência hidrotérmicos um problema de como usar, num certo período de tempo, a disponibilidade de água no sistema hidráulico de modo a substituir geração termoelétrica por geração hidroelétrica. A questão básica para o decisor é encontrar um compromisso entre o ganho relativo associado com a geração hidroelétrica imediata e a expectativa de benefícios futuros advindos da estocagem, tudo medido em termos de economia de combustível térmico.

A natureza cíclica dos fluxos de água e da demanda de carga assim como a validade da modelização sugerem dividir o problema em longo e curto prazos. Para o problema de longo prazo, o qual pode ser um ano (ciclo da água), é somente possível estimar em probabilidade a vazão dos rios e a carga e isto torna o problema mais difícil. Além disso, é essencial levar em consideração variações na altura dos reservatórios as quais afetam a eficiência da geração hidroelétrica. Por outro lado para períodos de otimização menores, um dia ou uma semana, é geralmente suposto completo conhecimento dos fluxos nos rios e da carga, o que não parece ser uma hipótese razoável no caso da segunda que apresenta significantes variações aleatórias mesmo a curto prazo. Ademais é comum se supor desprezível a influência de variações na altura dos reservatórios quando o horizonte de otimização é pequeno, hipótese essa que na verdade só é válida na ausência de pequenos reservatórios de elevada relação entre entrada de água e capacidade.

Muitas abordagens têm sido sugeridas para o problema hidrotérmico nos anos recentes. Métodos variacionais [13, 14], programação dinâmica [15-17] princípio do máximo de Pontryagin [18-20] e técnicas de programação matemática [21-23] em geral têm sido usadas para resolvê-lo em diferentes formulações. A dificuldade básica que persiste é harmonizar uma modelização acurada com a possibilidade de resolver sistemas reais de grande porte. Uma manipulação usada para superar o problema da dimensionalidade é agregar o subsistema hidroelétrico trabalhando com um reservatório composto equivalente [24-25] reduzindo a dimensão do problema e viabilizando o uso das técnicas existentes. Isto é, entretanto, uma indesejável simplificação sempre que a decisão "onde" armazenar é tão importante quanto a deci

são "quanto" armazenar.

Neste trabalho propõe-se uma abordagem por decomposição ao problema hidrotérmico que possibilita tratar sistemas de grande dimensão numa formulação mais geral e precisa. É suposta demanda de carga estocástica, variações na altura do reservatório e redes hidráulicas completamente gerais com usinas em cascata, atraso de transporte da água e vertimento. Com tais hipóteses pode-se resolver acuradamente problemas hidrotérmicos de curto e médio prazo e mesmo obter soluções aproximadas para o de longo prazo*. Se obterá inicialmente um problema equivalente determinístico baseado nas características estocásticas da demanda. A seguir se aplicará decomposição por dualidade para separar todas as unidades de produção do sistema hidrotérmico. Uma estrutura hierárquica de cálculo, em três níveis, resultará dessas manipulações. As condições sobre a qual este procedimento é válido serão discutidas e uma interpretação econômica será dada. Um exemplo ilustra o procedimento.

3.2 FORMULAÇÃO DO PROBLEMA

O problema aqui tratado pode ser formulado como segue: Minimizar função objetivo abaixo que representa o custo de operação térmico global

$$J = \sum_{m \in M} \sum_{i \in I} \psi_i(g_i^m) \quad (1)$$

(*) Uma maneira de reduzir os erros da otimização advindos da aleatoriedade das vazões de água é utilizar um procedimento tipo "malha aberta adaptada" que implica na reotimização em subintervalos do horizonte de otimização levando-se em consideração as últimas realizações da vazão nos rios e empregando como política de ação não mais que as primeiras decisões.

onde ψ_i é uma função convexa (geralmente quadrática)

O intervalo operacional das unidades térmicas é limitado

$$\underline{g}_i \leq g_i^m \leq \bar{g}_i, \quad i \in I, m \in M \quad (2)$$

A equação de balanço de potência que acopla os subsistemas hidroelétrico e termoelétrico é

$$\sum_{i \in I} g_i^m + \sum_{j \in J} h_j^m = d^m, \quad m \in M \quad (3)$$

Assume-se conhecida a função distribuição acumulada contínua $F_m(\xi)$ da variável aleatória demanda de carga em cada período de tempo m e supõe-se não correlação entre diferentes períodos de tempo. Ademais, não se considera em (3) perdas de transmissão para simplificar a exposição, mas isso será levado em consideração posteriormente.

A geração de potência hidroelétrica é uma função da turbinagem e da altura do reservatório, o qual por sua vez é função do esvaziamento do reservatório.

$$h_j^m = \phi_j(u_j^m, x_j^m), \quad j \in J, m \in M \quad (4)$$

Uma vez que a taxa de eficiência da turbina é decrescente com aumentos na turbinagem e altura do reservatório, ϕ_j é considerada uma função côncava.

A dinâmica dos reservatórios é descrita por equações a diferenças

$$x_j^{m+1} = x_j^m + y_j^m + z_j^m - u_j^m - v_j^m, \quad j \in J, m \in M \quad (5)$$

A entrada de água dependente Z_j^m é a variável de acoplamento conectando unidades hidroelétricas num mesmo vale hidráulico e é expressa por

$$Z_j^m = \sum_{k \in S_j} \left(u_k^{m-t_{kj}} + v_k^{m-t_{kj}} \right) \quad , j \in J, m \in M \quad (6)$$

onde S_j é o conjunto de índices das hidroelétricas vizinhas à hidroelétrica j rio acima e t_{kj} é o intervalo de tempo de deslocamento da água entre as hidroelétricas $k \in S_j$ e j (suposto constante e inteiro).

Existem também restrições locais sobre variáveis como capacidade dos reservatórios, intervalo de operação das turbinas e outras físicas.

$$\left. \begin{aligned} x_j^m \in X_j &= \left\{ x_j^m / \underline{x}_j \leq x_j^m \leq \bar{x}_j \right\} \\ u_j^m \in U_j &= \left\{ u_j^m / \underline{u}_j \leq u_j^m \leq \bar{u}_j \right\} \\ v_j^m &\geq 0 \end{aligned} \right\} , j \in J, m \in M \quad (7)$$

Finalmente é assumida disponível a solução do problema hidrotérmico para um horizonte de tempo maior; ou que por conhecimento empírico conhece-se o estado final dos reservatórios (por exemplo, cheios no início do período de seca). Isto permite a compatibilidade da solução garantindo otimização em períodos maiores de tempo. Assim, temos

$$\left. \begin{aligned} x_j^0 \\ x_j^T \end{aligned} \right\} \text{ dados} \quad , j \in J \quad (8)$$

Nestas condições o problema global pode ser apresentado como segue: Minimizar (1) sujeito a (2)-(8) o que caracteriza um problema dinâmico estocástico de grande porte com limites nas variáveis de estado e controle.

3.3 MANIPULAÇÃO

3.3.1 Equivalente determinístico

Inicialmente se tratará o aspecto estocástico do problema. A demanda de carga aleatória presente em (3) assegura que qualquer que seja a política de geração proposta somente é possível estimar as probabilidades de deficit e superavit da potência oferecida por esta política. Em outras palavras a equação (3) não significa uma restrição rígida a ser satisfeita por todas as candidatas a solução ótima mas somente nos permite comparar, para uma dada distribuição de probabilidades de d^m , as diferentes esperanças matemáticas de deficit e superavit que cada solução apresenta. Uma abordagem clássica em programação estocástica para o problema de recursos aleatórios é o "problema de dois estágios" [26] onde se transfere a função objetiva, com fatores de penalização apropriados, o valor esperado associado a deficit e superavit das restrições aleatórias.

De acordo com isso seja p^m a oferta global de potência no período m . Assim

$$\sum_{i \in I} g_i^m + \sum_{j \in J} h_j^m = p^m, \quad m \in M \quad (9)$$

e seja J' a função objetiva modificada

$$J' = \sum_{m \in M} \left\{ \sum_{i \in I} \psi_i (g_i^m) + \alpha^m \int_0^{p^m} (\xi - p^m) dF_m(\xi) + \right. \\ \left. + \beta^m \int_{p^m}^{\infty} (\xi - p^m) dF_m(\xi) \right\} \quad (10)$$

Então minimizar (10) sujeito a (2) (4)-(9) é o equivalente determinístico do problema estocástico original. Deve-se salientar que embora inicialmente se estava interessado somente em saber os níveis ótimos de geração de potência em cada unidade de produção agora se quer também saber a oferta de potência global ótima, o que significa que se tem uma nova variável de decisão.

O peso α^m é o ganho unitário reduzido associado com superavit de potência e o peso β^m é o custo unitário associado com deficit de potência. Eles podem ser obtidos em sistemas reais e dependem da política utilizada para igualar a demanda. Aqui será suposto que eles serão o preço de exportação e importação da energia trocada com sistemas vizinhos. Outros tipos de solução para garantir o balanço de potência podem também ser considerados como geração de emergência (com custos marginais mais elevados), corte de carga, armazenamento por bombeamento e outros. A função objetiva J' expressa o compromisso entre o custo de operação corrente do sistema e o custo corretivo de satisfazer a demanda.

3.3.2 Decomposição por dualidade

Agora, para viabilizar a resolução do problema para sistemas hidrotérmicos de grande porte se aplicará decomposição por dualidade no intuito de se obter subproblemas mais simples de resolver. Uma estrutura hierárquica de cálculo fornecerá, então, a solução ótima global.

Antes disso, entretanto, deve-se observar que o problema tal qual se apresenta no momento é convexo uma vez que J' é uma função convexa (veja apêndice A) e substituindo a restrição (9) por

$$\sum_{i \in I} g_i^m + \sum_{j \in J} h_j^m \geq p^m, \quad m \in M \quad (11)$$

tem-se um domínio de factibilidade convexo (lembrar que ϕ_j são funções côncavas). Esta relaxação é uma manipulação que assegura teoricamente o procedimento, embora de um ponto de vista prático a solu

ção não se altere. Realmente, para que a solução ótima do problema relaxado admita uma estrita desigualdade em (11) em algum período de tempo é necessário que todas as unidades térmicas estejam em seus limites inferiores neste período de tempo e que as hidroelétricas não possam reduzir suas turbinagens. Além do mais como $w(p)$ é uma função decrescente de p (apêndice A) é possível demonstrar que (11) será sempre ativa .

Agora, seja o Lagrangeano relativo à restrição (11)

$$L = J' + \sum_{m \in M} \lambda^m \left\{ p^m - \sum_{i \in I} g_i^m - \sum_{j \in J} h_j^m \right\} \quad (12)$$

Para problemas convexos, como vimos no cap. I, um ponto de sela do Lagrangeano restrito é uma condição suficiente e necessária (sobre qualificação de restrição) para otimalidade. Neste caso, um mínimo de L relativamente às variáveis primais $g_i^m, p^m, u_j^m, x_j^m; i \in I, j \in J, m \in M$ sujeito às restrições (2) (4) - (8) e um máximo relativo às variáveis duais $\lambda^m; m \in M$ sujeito a $\lambda^m \geq 0^*$, define o ponto de sela do Lagrangeano, solução do problema. O método dual clássico em dois níveis é proposto para encontrar este ponto de sela, onde no nível inferior se minimiza L para λ fixo, em outras palavras se avalia a função dual

$$H(\lambda) = \min L \quad \text{s.a.} \quad (2) (4) - (8) \quad (13)$$

e no nível superior se modifica o multiplicador de Lagrange λ de modo a maximizar a função dual $H(\lambda)$. Para isso se usa o gradiente da do pelas componentes

$$\frac{\partial H(\lambda)}{\partial \lambda^m} = p^m - \sum_{i \in I} g_i^m - \sum_{j \in J} h_j^m, \quad m \in M \quad (14)$$

(*) Certamente, os multiplicadores de Lagrange na solução ótima serão estritamente positivos uma vez que as restrições (11) serão ativas.

Mas o Lagrangeano (12) é aditivamente separável

$$\begin{aligned}
 L = & \sum_{m \in M} \left\{ \sum_{i \in I} \psi_i (g_i^m) - \lambda^m \sum_{i \in I} g_i^m \right\} \\
 & + \sum_{m \in M} \left\{ \alpha^m \int_0^{p^m} (\xi - p^m) dF_m(\xi) + \beta^m \int_{p^m}^{\infty} (\xi - p^m) dF_m(\xi) + \lambda^m p^m \right\} \\
 & - \sum_{m \in M} \left\{ \lambda^m \sum_{j \in J} h_j^m \right\} \quad (15)
 \end{aligned}$$

podendo assim sua minimização, para $\tilde{\lambda}^m$ fixo, ser realizada separadamente como três subproblemas:

* Subproblema térmico

$$\min \sum_{m \in M} \left\{ \sum_{i \in I} \psi_i (g_i^m) - \tilde{\lambda}^m \sum_{i \in I} g_i^m \right\} \quad (16)$$

$$\text{s.a. } \underline{g}_i \leq g_i^m \leq \bar{g}_i \quad ; \quad i \in I, \quad m \in M$$

este problema por sua vez reduz-se a otimizações independentes uma para cada planta termoelétrica $i \in I$

$$\begin{aligned}
 \min \sum_{m \in M} \left\{ \psi_i (g_i^m) - \tilde{\lambda}^m g_i^m \right\} \\
 \text{s.a. } \underline{g}_i \leq g_i^m \leq \bar{g}_i \quad ; \quad m \in M \quad (17)
 \end{aligned}$$

cuja solução é

$$\hat{g}_i^m = \min \left\{ \bar{g}_i, \max [\underline{g}_i, \bar{g}_i^m] \right\}$$

$$\text{onde } \frac{d \psi_i(g_i^m)}{d g_i^m} \Big|_{\hat{g}_i^m} - \hat{\lambda}^m = 0 \quad ; m \in M \quad (18)$$

* Subproblema estocástico

$$\begin{aligned} \min \sum_{m \in M} \left\{ \alpha^m \int_0^{p^m} (\xi - p^m) dF_m(\xi) + \right. \\ \left. + \beta^m \int_{p^m}^{\infty} (\xi - p^m) dF_m(\xi) + \hat{\lambda}^m p^m \right\} \end{aligned} \quad (19)$$

A solução de (19) pode ser obtida analiticamente e não é difícil mostrar que \hat{p}^m que resolve (19) é dado por (veja apêndice A)

$$\hat{p}^m = F_m^{-1} \left(\frac{\beta^m - \hat{\lambda}^m}{\beta^m - \alpha^m} \right) \quad ; m \in M \quad (20)$$

e a condição para que este desenvolvimento seja válido é

$$\beta^m \geq \hat{\lambda}^m \geq \alpha^m \quad ; m \in M \quad (21)$$

* Subproblema hidroelétrico

$$\max \sum_{m \in M} \hat{\lambda}^m \sum_{j \in J} h_j^m \quad \text{s.a (4)-(8)} \quad (22)$$

Ao contrário dos subproblemas anteriores cujas soluções são analíticas, o subproblema hidroelétrico somente pode ser solucionado por métodos iterativos e assim a eficiência computacional do procedimento global dependerá basicamente dele. Na ausência das restrições (6), o que ocorre para unidades hidroelétricas em paralelo, o subproblema se reduz à otimizações independentes, uma para cada usina hidroelétrica, cuja solução pode ser obtida por técnicas conhecidas (por exemplo Programação Dinâmica). No caso geral, entretanto, a dimensionalidade torna inviável abordagens globais. Ainda aqui se aplicará decomposição por dualidade, através de dois procedimentos alternativos.

a) Decomposição Espacial

Com o intuito de desacoplar espacialmente o subproblema hidroelétrico dualiza-se (22) em relação a restrição (6). Seja portanto o Lagrangeano

$$L' = \sum_{m \in M} \left\{ \hat{\lambda}^m \sum_{j \in J} \phi_j(u_j^m, x_j^m) + \sum_{j \in J} \rho_j^m \left[\sum_{k \in S_j} (u_k^{m-t_{kj}} + v_k^{m-t_{kj}}) - z_j^m \right] \right\} \quad (23)$$

que para o multiplicador de Lagrange $\hat{\rho}_j^m$ fixo se decompõe aditivamente em sublagrangeanos, um para cada $j \in J$.

$$L'_j = \sum_{m \in M} \left\{ \hat{\lambda}^m \phi_j(u_j^m, x_j^m) + \hat{\rho}_k^{m+t_{jk}} (u_j^m + v_j^m) - \hat{\rho}_j^m z_j^m \right\} \quad (24)$$

onde k é o índice da hidroelétrica à jusante de j .

Um método dual clássico pode então ser aplicado onde uma sequência de multiplicadores de Lagrange é estabelecida, convergindo para a minimização da função dual

$$H'(\rho) = \max_{j \in J} \sum_{j \in J} L'_j \quad \text{s.a (5) (7) (8) (25)}$$

através da utilização do vetor gradiente dado pelas componentes

$$\frac{\partial H'(\rho)}{\partial \rho_j^m} = \sum_{k \in S_j} \left(u_k^{m-t_{kj}} + v_k^{m-t_{kj}} \right) - z_j^m, \quad j \in J, m \in M \quad (26)$$

A cada passo subproblemas hidroelétricos são resolvidos, um para cada $j \in J$.

$$\begin{array}{l}
 \max \sum_{m \in M} \tilde{\lambda}_j^m \phi_j(u_j^m, x_j^m) + \tilde{\rho}_k^{m+t} (u_j^m + v_j^m) - \tilde{\rho}_j^m z_j^m \\
 \text{s.a. } x_j^{m+1} = x_j^m + y_j^m + z_j^m - u_j^m - v_j^m \\
 x_j^m \in X_j \\
 u_j^m \in U_j, \quad j \in J, m \in M \\
 v_j^m \geq 0 \\
 \left. \begin{array}{l} x_j^0 \\ x_j^T \end{array} \right\} \text{ dados}
 \end{array} \quad (SH)_j \quad (27)$$

A resolução de (27) pode ser feita por Programação Dinâmica. A dificuldade de aplicação desta técnica ao problema (22), qual se ja, a questão da dimensionalidade, fica então resolvida pela decomposição espacial.

b) Decomposição temporal

A decomposição no tempo de um problema de otimização dinâmico é uma técnica já conhecida, proposta inicialmente por Lasdon [1] e Shoeffler [27] e mais recentemente por Tamura [28]. Apresenta a vantagem de permitir a solução de problemas de controle ótimo com

restrições no estado (independentes no tempo) de uma maneira simples bem como facilita o tratamento de sistemas com equação dinâmica contendo atrasos de transporte. Estas duas características es-tão presentes no subproblema hidroelétrico (22).

Antes, porém, sejam as seguintes manipulações. Uma vez que para cada reservatório $j \in J$

$$u_j^m + v_j^m = -x_j^{m+1} + x_j^m + y_j^m + z_j^m$$

e como z_j^m é dado por (6) tem-se

$$z_j^m = \sum_{k \in S_j} \left(-x_j^{m-t_{kj}+1} + x_k^{m-t_{kj}} + y_k^{m-t_{kj}} + z_k^{m-t_{kj}} \right)$$

e continuando as substituições para $z_k^{m-t_{kj}}$ obtem-se a expressão

$$z_j^m = \sum_{k \in R_j} \left(-x_k^{m-\tau_{kj}+1} + x_k^{m-\tau_{kj}} + y_k^{m-\tau_{kj}} \right)$$

onde R_j é o conjunto dos índices das hidroelétricas rio acima cuja descarga de água alcança a hidroelétrica j e τ_{kj} é o tempo de percursos acumulado de $k \in R_j$ até j .

Podemos assim reescrever o problema (22) substituindo a ex-pressão de z_j^m acima em (5) e tornando implícita as restrições $v_j^m > 0$. A formulação do problema (22) fica então

$$\left. \begin{aligned}
 & \max \sum_{m \in M} \hat{\lambda}_j^m \sum_{j \in J} \phi_j (u_j^m, x_j^m) \\
 & \text{s.a. } x_j^{m+1} \leq x_j^m + y_j^m + \sum_{k \in R_j} \left\{ -x_k^{m-\tau_{kj}+1} + \right. \\
 & \quad \left. + x_k^{m-\tau_{kj}} + y_k^{m-\tau_{kj}} \right\} - u_j^m \\
 & x_j^m \in X_j \\
 & u_j^m \in U_j \quad , j \in J, m \in M \quad (28) \\
 & \left. \begin{array}{l} x_j^0 \\ x_j^T \end{array} \right\} \text{ dados}
 \end{aligned}
 \right\}$$

Agora escrevendo o Lagrangeano relativo à restrição dinâmica de desigualdade tem-se

$$L'' = \sum_{m \in M} \left\{ \sum_{j \in J} \hat{\lambda}_j^m \phi_j (u_j^m, x_j^m) + \sum_{j \in J} \rho_j^m \left[-x_j^{m+1} + x_j^m + y_j^m + \right. \right. \\
 \left. \left. + \sum_{k \in R_j} (-x_k^{m-\tau_{kj}+1} + x_k^{m-\tau_{kj}} + y_k^{m-\tau_{kj}}) - u_j^m \right] \right\} \quad (29)$$

Para os multiplicadores $\hat{\rho}_j^m$ fixos o Lagrangeano (29) é aditivamente separável em m e j (o que significa que a decomposição é também espacial).

$$L'' = \sum_{m \in M} \sum_{j \in J} L''_{jm}$$

onde

$$L''_{jm} = \hat{\lambda}^m \phi_j(u_j^m, x_j^m) + \left[-\hat{\rho}_j^{m-1} + \hat{\rho}_j^m + \sum_{k \in T_j} (-\hat{\rho}_k^{m+\tau_{jk}})^{-1} + \right. \\ \left. + \hat{\rho}_k^{m+\tau_{jk}} \right] x_j^m - \hat{\rho}_j^m u_j^m + (\hat{\rho}_j^m + \sum_{k \in T_j} \hat{\rho}_k^{m+\tau_{jk}}) y_j^m \quad (30)$$

e T_j é o conjunto dos índices de todas as hidroelétricas à jusante de j no mesmo vale hidráulico e τ_{jk} é o intervalo de tempo acumulado de j até $k \in T_j$.

É importante observar que L''_{jm} é da forma

$$L''_{jm} = \hat{\lambda}^m \phi_j(u_j^m, x_j^m) + \Delta \hat{\rho}_j^m x_j^m - \hat{\rho}_j^m u_j^m + cte$$

e sua maximização sobre x_j, u_j é assegurada pelas condições de Kuhn-Tucker

$$\left. \begin{array}{l} \frac{\partial L''_{jm}}{\partial x_j^m} \\ \frac{\partial L''_{jm}}{\partial u_j^m} \end{array} \right\} \begin{array}{l} = 0 \quad \text{se } \underline{x}_j < x_j^m < \bar{x}_j \\ < 0 \quad \text{se } x_j^m = \underline{x}_j \\ > 0 \quad \text{se } x_j^m = \bar{x}_j \end{array} \quad (31)$$

$$\frac{\partial L''_{jm}}{\partial u_j^m} \left\{ \begin{array}{ll} = 0 & \text{se } \underline{u}_j < u_j^m < \bar{u}_j \\ < 0 & \text{se } \underline{u}_j = u_j^m \\ > 0 & \text{se } u_j^m = \bar{u}_j \end{array} \right. \quad (31)$$

Uma vez maximizado o Lagrangeano (29) sobre X, U para $\hat{\rho}$ fixo deve-se modificar o multiplicador de modo a minimizar a função dual

$$H''(\rho) = \max L'' \quad \text{sobre } X, U \quad (32)$$

e para isso se utiliza o vetor gradiente dado pelas componentes

$$\begin{aligned} \frac{\partial H''(\rho)}{\partial \rho_j^m} &= -x_j^{m+1} + x_j^m + y_j^m + \sum_{k \in R_j} (-x_k^{m-\tau_{kj}+1} + \\ &+ x_j^{m-\tau_{kj}} + y_j^{m-\tau_{kj}}) - u_j^m \end{aligned} \quad (33)$$

3.4 ESTRATÉGIA DE RESOLUÇÃO

A lógica computacional para os dois procedimentos pode ser resumida como se segue:

Passo 1: Escolha multiplicadores iniciais $\beta^m \geq \lambda^m \geq \alpha^m$, $\lambda^m \geq 0$, $m \in M$.

Passo 2: Resolva o subproblema térmico de acordo com (18) seja \hat{g}_i^m , $i \in I$, $m \in M$ a solução.

- Passo 3: Resolva o subproblema estocástico de acordo com (20) seja \hat{p}^m , $m \in M$ a solução.
- Passo 4: Escolha multiplicadores iniciais ρ_j^m (≥ 0 no caso b), $j \in J$, $m \in M$.
- Passo 5: Caso a - Resolva (27) para cada $j \in J$
Caso b - Maximize (30) sobre X_j , U_j para cada $j \in J$, $m \in M$,
Sejam \hat{x}_j^m , \hat{u}_j^m , $j \in J$, $m \in M$ a solução.
- Passo 6: Se (26) - caso a - ou (33) - caso b - for zero dentro de uma precisão estabelecida: ir ao passo 7.
Caso contrário determinar novo ρ_j^m usando, por exemplo, um método tipo gradiente conjugado

$$\rho_j^m = \rho_j^m - \epsilon^* w^m$$

onde $\epsilon > 0$ pequeno e w^m é uma direção conjugada.

Voltar ao passo 5.

- Passo 7: Se (14) é zero dentro de uma precisão estabelecida, pare:
A presente solução \hat{g}_i^m , \hat{p}^m , \hat{x}_j^m , \hat{u}_j^m ; $i \in I$, $j \in J$, $m \in M$ é ótima.
Caso contrário determinar novo λ^m usando, por exemplo, um método de gradiente

$$\lambda^m = \lambda^m + \epsilon G^m$$

onde $\epsilon > 0$ pequeno e G^m é o gradiente dado por (14).
Volte ao passo 2.

A estrutura hierárquica de cálculo se apresenta esquematizada na fig. 1.

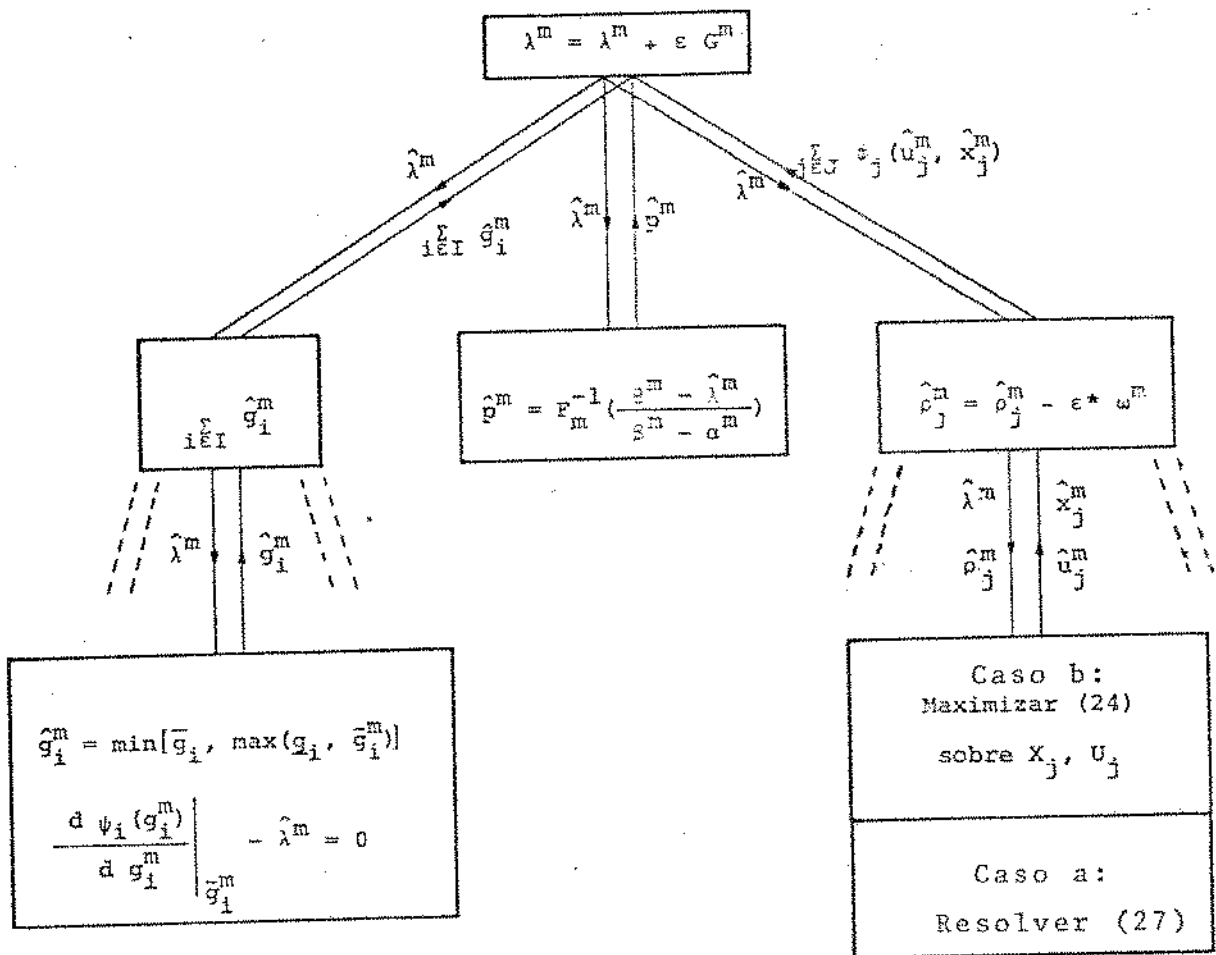


Fig.1- ESTRUTURA HIERÁRQUICA DE CÁLCULO

3.5 INTERPRETAÇÃO ECONÔMICA

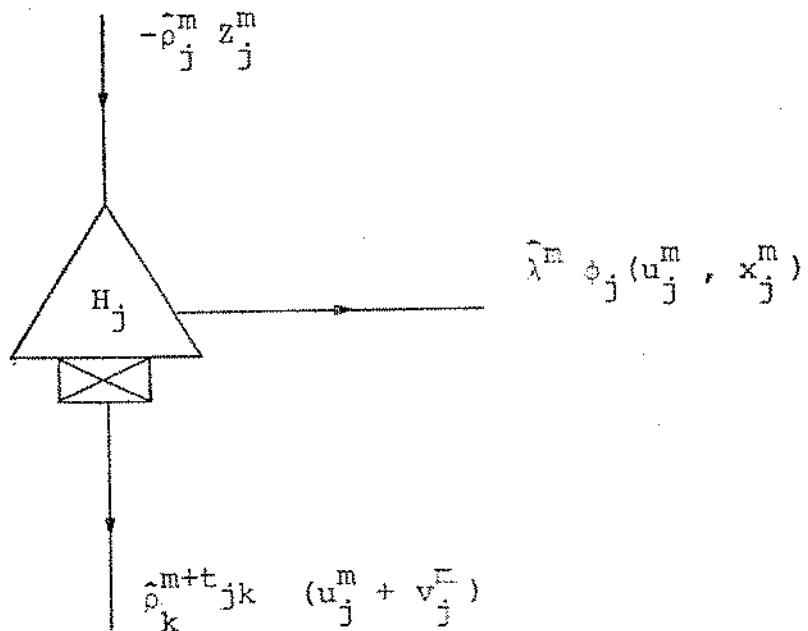
A decomposição do problema hidroelétrico, baseada na teoria de Lagrange generalizada, e a estrutura hierárquica de cálculo resultante apresentam interessantes interpretações econômicas. O acoplamento hidrotérmico, expressado pela equação (11), foi desfeito através da fixação dos multiplicadores de Lagrange λ^m que podem ser interpretados como o preço da energia^(*) no tempo, para o sistema. Quando (11) não é satisfeito deve-se mudar este preço de modo que no equilíbrio os subsistemas Termoelétrico e Hidroelétrico desacopladas fornecem requerimentos globais do sistema.

Deve-se enfatizar também que a abordagem empregada no problema estocástico, a qual acarretou outro subproblema no segundo nível, também originou uma nova variável de decisão que é a oferta global de potência. Isto significa que para um problema com recurso aleatório deve-se responder não somente "como" distribuir esse recurso mas também "quanto" do recurso é conveniente supor-se dispor. Em nosso caso para um valor fixo da energia no mercado doméstico λ^m , a comparação entre os preços de importação e exportação de energia, β^m e α^m respectivamente, define para uma dada distribuição de probabilidades da demanda de carga, uma solução de compromisso minimizado o risco de deficit elevados sem entretanto desperdiçar energia vendendo-a a preços inferiores nos mercados vizinhos. A regra dada por (20) expressa essa interpretação. A condição (21), de que o preço da energia doméstica permaneça entre os preços de importação e exportação, assegura soluções não triviais.

Por outro lado, no caso a, o acoplamento hidráulico dado por (6) foi quebrado pela fixação dos multiplicadores de Lagrange ρ_j^m que representam o valor da água ao longo da rede fluvial, no tempo. Assim o objetivo econômico no subproblema hidroelétrico (27) é

(*) Como trabalhamos com o problema discretizado no tempo (com intervalo de discretização constante) a potência média é proporcional a energia (gerada ou consumida) com o que passa a ser equivalente falarmos em "preço da energia" ou "preço da potência".

interpretado como a maximização do benefício líquido numa unidade hidroelétrica que paga* pela água que recebe de suas vizinhas rio acima $-\hat{\rho}_j^m z_j^m$ e ganha* tanto pela geração de potência elétrica $\hat{\lambda}^m \phi_j(u_j^m, x_j^m)$ como pela descarga de água na sua hidroelétrica vizinha a jusante $\hat{\rho}_k^{m+t} j_k(u_j^m + v_j^m)$.



No caso b, o acoplamento hidráulico (6) bem como a equação dinâmica (5) foram desfeitos através da fixação dos multiplicadores ρ_j^m que continuam representando o valor da água no tempo e no espaço. O objetivo econômico expresso em (30) pode ser interpretado como o interesse financeiro de um gerente responsável pela hidroelétrica j no instante m. Ele ganha $\hat{\lambda}^m \phi_j(u_j^m, x_j^m)$ devido a geração de energia, paga $-\Delta \hat{\rho}_j^m$ pela desvalorização da água armazenada no seu reservatório durante o período m, e também paga $\hat{\rho}_j^m$ pela água turbinada. O valor da água armazenada no reservatório j no instante m é dado

(*) Na verdade os multiplicadores podendo ser negativos implicam na possibilidade contrária de ganhar pela água que recebe e pagar pela água que descarrega.

pela potencialidade de geração que pode fornecer tanto na hidroelétrica j como nas demais rio abaixo, ou seja

$$\rho_j^m + \sum_{k \in T_j} \rho_k^{m+\tau_{kj}}$$

Estas interpretações econômicas podem nos ajudar a obter reduções no tempo de computação, uma vez que podemos inicializar o algoritmo hierárquico com bons multiplicadores de Lagrange. Isto é um ponto importante em métodos duais, que afeta consideravelmente a eficiência computacional.

Uma vez que o preço de equilíbrio da energia deve ser igual ao custo marginal global de operação em cada instante de tempo nós podemos estimá-lo diretamente da performance corrente do sistema. Além disso a restrição (21) garante que o preço de equilíbrio estará limitado pelos preços de exportação e importação.

Para o preço de equilíbrio da água sabemos que seu valor é expresso em termos da possibilidade potencial de gerar energia. Assim, ao nos aproximarmos do horizonte de otimização esta potencialidade é reduzida, o que significa que a água se desvaloriza no tempo. O preço da água é uma função decrescente do tempo.

3.6 EXEMPLO

Para ilustrar o método resolveu-se um exemplo com 4 usinas hidroelétricas e 2 usinas termoelétricas sob um horizonte de otimização de 12 intervalos. O sistema é mostrado na fig. 2.

USINAS TERMOELÉTRICAS

O custo incremental das usinas termoelétricas é:

$$\frac{d\psi_1}{dg_1} = 10.00 + 1.00 g_1$$

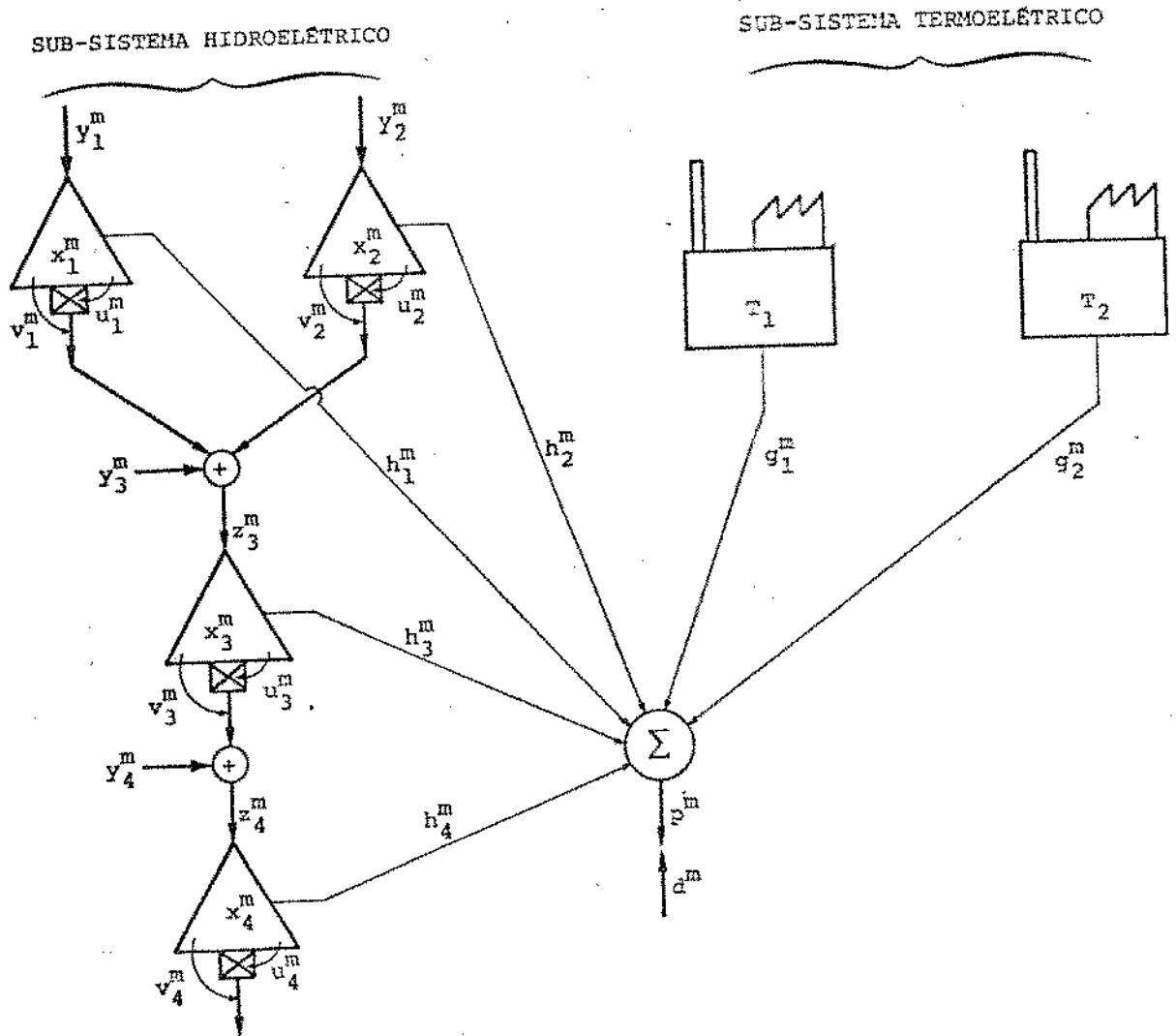


Fig.2 - SISTEMA HIDROTÉRMICO DE POTÊNCIA

$$\frac{d\psi_2}{dg_2} = - 20.00 + 1.66 g_2$$

e os limites de operação são:

$$10.0 \leq g_1 \leq 80.0$$

$$20.0 \leq g_2 \leq 80.0$$

USINAS HIDROELÉTRICAS

A função geração de potência para as usinas hidroelétricas é:

$$\phi(u, x) = c_1 x^2 + c_2 u^2 + c_3 x u + c_4 x + c_5 u + c_6$$

onde os parâmetros c_i são dados para cada hidroelétrica por:

HIDROELÉTRICA	c_1	c_2	c_3	c_4	c_5	c_6
1	-0.001	-0.1	0.01	0.40	4.0	-30.0
2	-0.001	-0.1	0.01	0.38	3.5	-30.0
3	-0.001	-0.1	0.01	0.30	3.0	-30.0
4	-0.001	-0.1	0.01	0.38	3.8	-30.0

Os limites dos reservatórios e turbinagens, assim como as condições inicial e final dos reservatórios estão dados na tabela seguinte.

HIDROELÉTRICA	\underline{x}	\bar{x}	\underline{u}	\bar{u}	x^0	x^T
1	80.0	150.0	5.0	15.0	100.0	120.0
2	60.0	120.0	6.0	15.0	80.0	70.0
3	100.0	240.0	10.0	30.0	170.0	170.0
4	70.0	160.0	13.0	25.0	120.0	140.0

REDE HIDRÁULICA

As afluências independentes em cada reservatório são dadas na tabela seguinte.

RESERVATÓRIO	INTERVALO DE TEMPO											
	1	2	3	4	5	6	7	8	9	10	11	12
1	10.	9.	8.	7.	6.	7.	8.	9.	10.	11.	12.	10.
2	8.	8.	9.	9.	8.	7.	6.	7.	8.	9.	9.	8.
3	20*	20*	10*	2.	3.	4.	3.	2.	1.	1.	1.	2.
4	20*	20*	20*	18*	0.	0.	0.	0.	0.	0.	0.	0.

(*) Estão incluídos nestes valores a condição inicial das vazões de água na rede.

O intervalo de tempo de percurso da água entre as hidroelétricas é:

TRECHO	TEMPO DE PERCURSO
$H_1 - H_3$	2
$H_2 - H_3$	3
$H_3 - H_4$	4

DEMANDA DE POTÊNCIA

A demanda de potência foi considerada uma variável aleatória com distribuição Normal com média e variância dadas a seguir.

DEMANDA DE POTÊNCIA	INTERVALO DE TEMPO											
	1	2	3	4	5	6	7	8	9	10	11	12
MÉDIA	200.	210.	220.	230.	240.	250.	240.	230.	220.	210.	210.	190.
VARIÂNCIA	5.	5.	5.	5.	5.	5.	5.	5.	5.	5.	5.	5.

e considerou-se os pesos α e β como preços de exportação e importação de energia trocada com sistemas vizinhos (constantes no tempo).

$$\alpha = 10.$$

$$\beta = 100.$$

Com estes dados e usando a técnica de decomposição no tempo para resolver o subproblema hidroelétrico otimizamos a operação do sistema hidrotérmico.

Os valores iniciais para λ e ρ foram

$$\lambda = 0.5 (\alpha + \beta) \quad ; \quad \rho = 0$$

A convergência foi obtida com 5 iterações do nível superior e número médio de 150 iterações no nível inferior. O programa requereu aproximadamente 8 minutos de CPU num computador PDP-10. Os resultados encontram-se na fig. 3.

O vertimento v_j nos reservatórios foi zero no ótimo. Este é um resultado esperado quando não há excesso de água no sistema. Realmente, como a geração não é função de v_j , não existe ganho associado com $v_j > 0$. Assim, v_j será maior que zero somente quando não for possível turbinar mais, ou seja.

$$\hat{v}_j^m (\hat{u}_j^m - \bar{u}_j) = 0$$

V_j, m

A utilização de Programação Dinâmica para resolver o subproblema hidroelétrico acarretou um tempo de processamento proibitivo mesmo para problemas de dimensão pequena. Além disso apresentou dificuldades de convergência advindas da discretização das variáveis de estado e controle. Um melhoramento na convergência pressuporia uma discretização mais fina, o que elevaria ainda mais o tempo de processamento.

Uma alternativa a ser estudada seria a utilização de técnicas especiais de Programação Dinâmica.

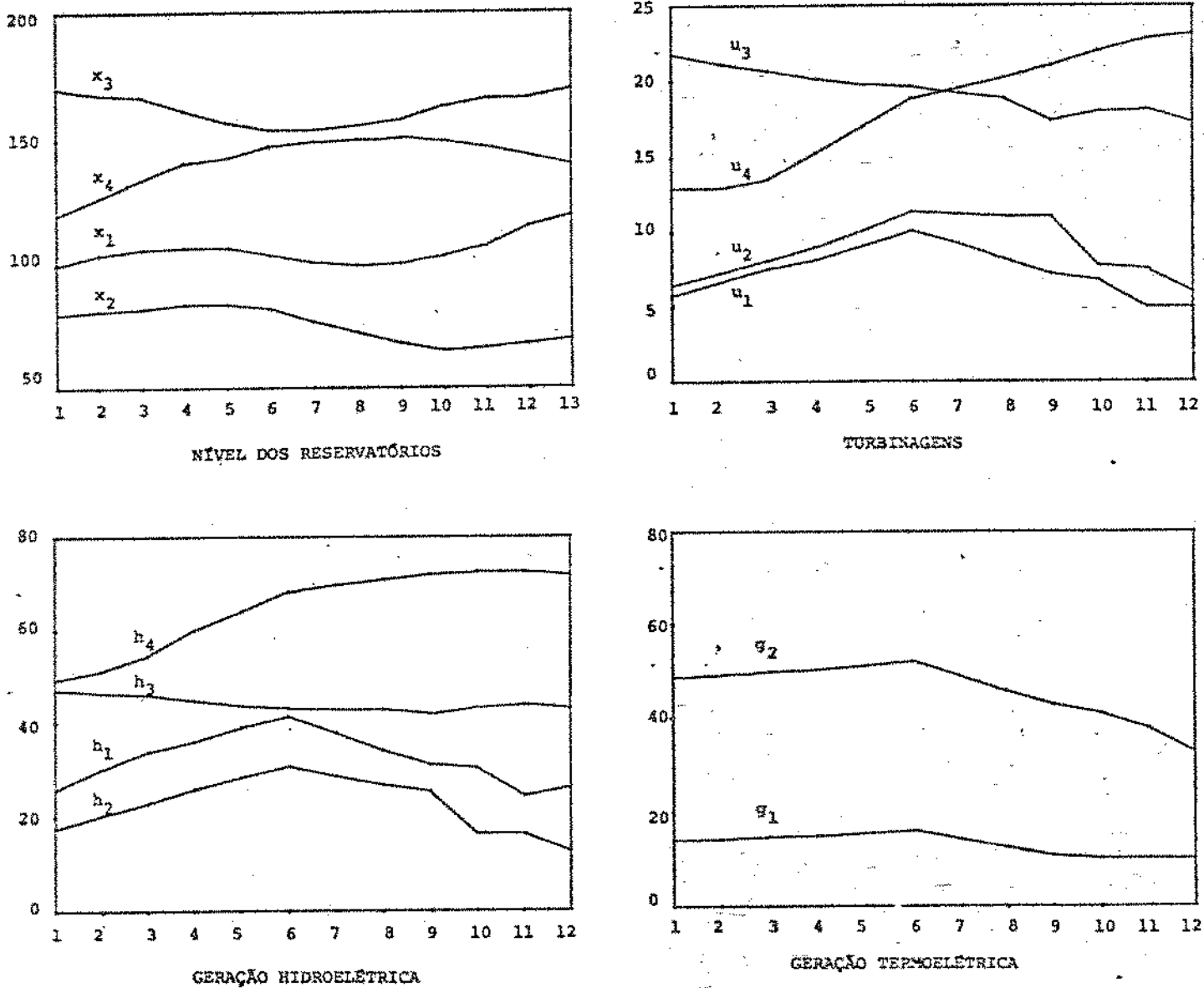
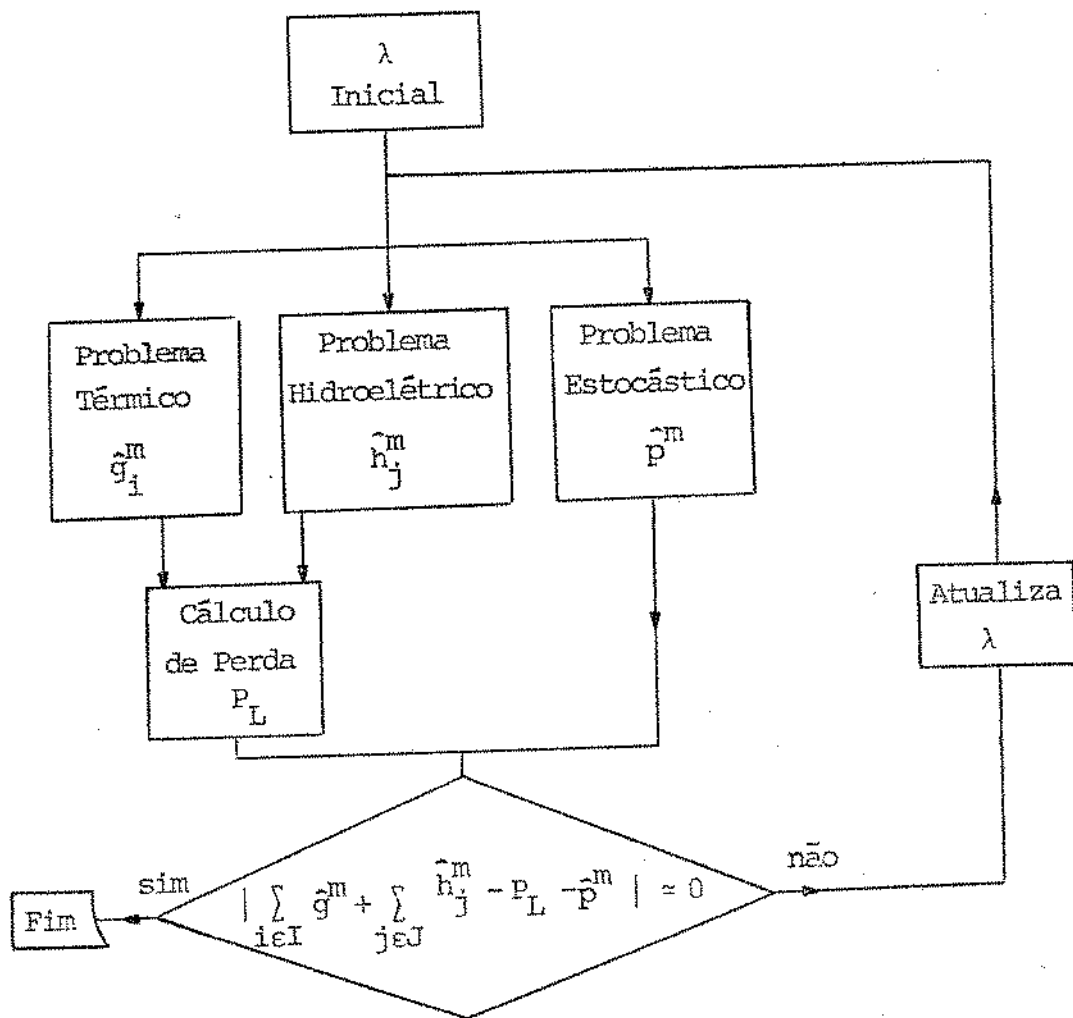


Fig. 3 - RESULTADOS

3.7 COMENTÁRIOS

É conveniente levantar alguns comentários:

1. A estrita convexidade das funções ψ_i e estrita concavidade das funções ϕ_j são condições suficiente para garantir a aplicabilidade da técnica de decomposição proposta. Além de assegurar a existência do ponto de sela do Lagrangeano garante a convergência até ele. Estas hipóteses estão de acordo com as condições reais.
2. A natureza estocática da demanda de potência elimina dificuldades eventuais de convergência no nível superior uma vez que para cada λ obtemos uma política de operação viável. Parando o procedimento iterativo antes de alcançar o balanço de potência dispomos de uma solução sub-ótima que é interessante em muitos casos.
3. Não é difícil considerar as perdas de transmissão na rede elétrica neste método. Alguns trabalhos anteriores [22, 29] têm proposto separação entre cálculo de perda e escala de geração, sugerindo procedimentos iterativos para resolver o problema global. Do mesmo modo podemos incorporar a estimação das perdas de transmissão neste método como mostra a fig. 4. A estimação da perda pode ser obtida por uma fórmula de perda [13, 23] ou, se uma precisão melhor for desejada, através de cálculos de fluxo de carga [21, 22, 29]



4. O método apresentado permite resolver problemas de grande porte com uma formulação rigorosa. É possível tratar problemas de dimensão superior ao exemplo apresentado, que serviu apenas de ilustração. De qualquer maneira na literatura levantada não são tratados problemas superiores a este a não ser através de agregação dos reservatórios. E raros são os trabalhos que admitem as não linearidades presentes no problema [30].
5. É possível considerar muitos outros fatores na formulação do problema sem acarretar complicações teóricas ou computacionais. Por exemplo: evaporação nos reservatórios, decomposição do sistema baseada na topologia da rede de transmissão, várias políticas de balanço de potência hierarquicamente empregadas, restrições de navegação, irrigação, etc...

6. Melhoramentos no tempo de computação poderão ser obtidos com mais adequadas inicializações nos multiplicadores de Lagrange assim como por melhor racionalização na programação.
7. Ao invés de fixar o estado final dos reservatórios é possível atribuir-se um custo terminal associado ao valor da água. A otimização neste caso procuraria utilizar os recursos hidráulicos do sistema no sentido de minimizar o custo presente de operação (dentro do horizonte) sem entretanto comprometer o funcionamento futuro do sistema (além do horizonte). No caso de se considerar as afluências estocásticas é imprescindível utilizar esta alternativa.
8. Outros métodos de decomposição e coordenação poderiam ser testados e comparados quanto a eficiência de resolução do problema. No caso de se trabalhar com funções não estritamente convexas, lineares por partes por exemplo, poderiam ser usados métodos baseados em linearização e relaxação. Esta parece ser uma linha de estudo interessante, uma vez que poderíamos linearizar por partes as funções ψ_i e ϕ_j e encontrar talvez mais eficientemente boas soluções.

APENDICE A

Seja $F(\xi)$ a função de distribuição de probabilidades de uma variável aleatória definida sobre o domínio $(0, \infty)$.

Seja a função

$$w(p) = \alpha \int_0^p (\xi - p) dF(\xi) + \beta \int_p^{\infty} (\xi - p) dF(\xi)$$

cuja primeira derivada é

$$\begin{aligned} \frac{d w(p)}{dp} &= \alpha \frac{d}{dp} \int_0^p (\xi - p) dF(\xi) + \beta \frac{d}{dp} \int_p^{\infty} (\xi - p) dF(\xi) \\ &= (\beta - \alpha) \frac{d}{dp} \int_0^p (\xi - p) dF(\xi) - \beta \\ &= (\beta - \alpha) F(p) - \beta \end{aligned}$$

Como $F(p)$ é uma função não decrescente $w(p)$ é uma função convexa. Ainda mais se $\beta - \alpha \geq 0$, $\alpha \geq 0$, $\beta > 0$ e como $0 \leq F(p) \leq 1$, a função $w(p)$ será também decrescente.

Finalmente, o ponto de mínimo de $w(p) + \lambda p$ deverá satisfazer

$$(\beta - \alpha) F(p) - \beta + \lambda = 0$$

ou seja

$$F(p) = \frac{\beta - \lambda}{\beta - \alpha}$$

BIBLIOGRAFIA

- (1) L. S. LASDON, "Optimization Theory for Large Scale Systems", Mac Millan, 1970.
- (2) A. M. GEOFFRION, "Duality in Nonlinear Programming: A Simplified Applications - Oriented Development", in Perspectives on Optimization, Ed. Geoffrion, Addison-Wesley, pp. 65 - 101, 1972.
- (3) J. D. SCHOEFFLER, "Static Multilevel Systems", in Optimization Methods for Large Scale Systems, Ed. Wismer, Mc Graw-Hill, 1971.
- (4) H. EVERETT, "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources", Operations Research, 11, pp. 399-417, 1963.
- (5) A. M. GEOFFRION, "Elements of Large - Scale Mathematical Programming", R- 481-PR, RAND Corporation, November, 1969.
- (6) H. UZAWA, "Iterative Methods for Concave Programming", in Studies in Linear and Nonlinear Programming, Stanford University Press, 1958.
- (7) L. S. LASDON, "A Multi-Level Technique for Optimization", Systems Research Center Report 50-64-19, Case Institute of Technology, 1964.
- (8) G. B. DANTZIG, e P. WOLFE, "The Decomposition Algorithm for Linear Programming", Operations Research, 8, January-February, pp. 101-111, 1960.
- (9) A. M. GEOFFRION, "Primal Resource - Directive Approaches for Optimizing Nonlinear Decomposable Systems", Working Paper 141, The RAND Corporation, 1968.

- (10) J. F. BENDERS, "Partitioning Procedures for Solving Mixed - Variables Programming Problems", Numerische Mathematik, 4, pp. 238 - 252, 1962.
- (11) A. M. GEOFFRION, "Generalized Benders Decomposition" Journal of Optimization Theory and Applications, vol.10, n° 4, 1972.
- (12) J. B. ROSEN, "Primal Partition Programming for Block Diagonal Matrices", Numerische Mathematik, 6, pp. 250 - 260, 1964.
- (13) L. K. KIRCHMAYER, "Economic Operation of Power Systems", John Wiley & Sons, 1958.
- (14) J. H. DRAKE, L. K. KIRCHMAYER, R. B. MAYALL, e W. WOOD , "Optimum Operation of a Hydrothermal System", AIEE Transactions on Power Apparatus and Systems, vol. PAS-80, pp.242 - 250 , Agosto, 1962.
- (15) J. D. C. LITTLE, "The use of Storage Water in a Hydroelectric System", ORSA, 1955.
- (16) B. BERNHOLTZ, e L. V. GRAHAM, "Hydrothermal Economic Scheduling - Part. I: Solution by Incremental Dynamic Programming", AIEE Transactions on Power Apparatus and Systems, vol. PAS - 79, pp. 921 - 932, Dezembro, 1960.
- (17) E. B. DAHLIN, e D.N.C. SHEN, "Application of Dynamic Programming to Optimization of Hydroelectric / Steam Power System Operation", Proc. IEE, vol. 112, n° 12, Dezembro, 1965.
- (18) E. B. DAHLIN, e D.N.C. SHEN, "Optimal Solution to the Hydro-Steam Dispatch Problem for Certain Practical Systems", IEEE Transactions on Power Apparatus and Systems, vol.PAS-85, pp. 437-458, Maio, 1966.

- (19) I. HANO, Y. TAMURA, e S. NARITA, "An Application of the Maximum Principle to the Most Economical Operation of Power Systems, IEEE Transactions on Power Apparatus and Systems, vol. PAS - 85, pp. 486 - 494, Maio, 1966.
- (20) A. TURGEONS, "Optimal Operation of a Hydro-steam Power Systems", Ph.D Tese, Cornell University, 1975.
- (21) M. RAMAMOORTY, e J. GAPOLA RAO, "Load Scheduling of Hydroelectric Thermal Generating Systems Using Nonlinear Programming Techniques", Proc. IEE, vol. 117, n^o 4, pp. 794 - 798, Abril, 1970.
- (22) A. P. BONAERT, EL-ABIAD A. H., e A. J. KOIVO, " Optimal Scheduling of Hydro-Thermal Power Systems", PICA Conference, Boston, pp. 263-270, Maio, 1971.
- (23) S. K. AGARWAL, e I. J. NAGRATH, "Optimal Scheduling of Hydro Thermal Systems", Proc. IEE, vol. 119, n^o 2, pp. 169 - 173, Fevereiro, 1972.
- (24) N. V. ARVANITIDIS, e J. ROSING, "Composite Representation of a Multireservoir Hydroelectric Power System", IEEE Transactions on Power Apparatus and Systems, vol. PAS-89, n^o 2, pp. 319 - 326, Fevereiro, 1970.
- (25) N. V. ARVANITIDIS, e J. ROSING, "Optimal Operation of Multireservoir Systems Using a Composite Representation", IEEE Transactions on Power Apparatus and Systems, vol. PAS-89, n^o 2, pp. 237 - 335, Fevereiro, 1970.
- (26) J. K. SENGUPTA, "Stochastic Programming", North Holland, 1972.
- (27) J. D. SCHOEFFLER, e L. S. LASDON, "Decemtralized Plant Control", ISA Trans., 5, pp. 175 - 183, April, 1966.

- (28) H. TAMURA, "Decentralized Optimization for Distributed-lag Models of Discret Systems", Automatica, vol. 11, pp. 593 - 602, 1975.
- (29) J. F. DOPAZO, O. A. KLITIN, G. W. STAGG, e M. WATSON "An Optimization Technique for Real and Reactive Powe Allocation", Proc. IEEE, vol. 55, n^o 11, pp. 1877 - 1885, Novembro, 1967.
- (30) T. N. SAHA, e S. A. KHAPARDE, "An Application of a Direct Method to the Optimal Scheduling of Hydro thermal System", IEEE Transactions on Power Apparatus and Systems, vol. PAS-97, n^o 3, pp. 977-983, Maio/Junho, 1978.