

ASPECTOS COMPUTACIONAIS DE UM ALGORITMO PRIMAL
DE DIREÇÕES VIÁVEIS PARA SISTEMAS DECOMPONÍVEIS.

Odival Juliano de Campos.

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIA (M.Sc.).

Aprovada por :

Presidente

RIO DE JANEIRO

ESTADO DA GUANABARA - BRASIL

SETEMBRO DE 1972.

Ao Clóvis quero exprimir meus reais agradecimentos, lembrando que sua orientação bastante valiosa, elevou-se bem além do presente trabalho.

Ao Marinho devo toda a orientação acadêmica, tendo sobremodo colaborado no desenvolvimento da pesquisa que apresentamos.

A tia Elza devo a dedicação no trabalho de monografia e, a Sô mais que a datilografia a sua amizade.

A meus pais cujo incentivo jamais pode rei desprezar.

RESUMO

Visando otimizar-se um sistema decomponível não-linear, estuda-se um algoritmo primal de direções viáveis.

No decurso desse estudo formalizam-se e demonstram-se alguns resultados conhecidos e desenvolvem-se interpretações econômicas.

Um critério para testar a diferenciabilidade das funções pertubação envolvidas no algoritmo é desenvolvido, podendo tal propriedade ser utilizada com vantagens computacionais.

Rapidamente se discute o conceito de ϵ -ativida de de vínculos ligado à convergência do algoritmo.

Finalmente, apresenta-se uma subrotina em FORTRAN IV que obtém, em cada iteração do algoritmo, uma direção viável.

ABSTRACT

A primal feasible directions algorithm is studied to optimize non-linear decomposable systems. In the course of these studies some known results are formalized and proved, and economical interpretations are developed.

A criterion is derived to detect the differentiability of the perturbation functions involved in the algorithm, and this property is used with some computational advantages.

The concept of ϵ -activity of constraints is discussed, and related to the convergence of the algorithm.

Finally, a FORTRAN IV subroutine is presented capable of finding a feasible direction in each iteration of the algorithm.

ÍNDICE

<u>CAP. I</u>	<u>INTRODUÇÃO</u>	1
<u>CAP. II</u>	O PROBLEMA GERAL DE ALOCAÇÃO DE RECURSOS, MÉTODOS, MANIPULAÇÃO E TEOREMA DE EQUIVA- LENCIA	11
<u>SEÇÃO 1</u>	- Problema geral de alocação de recursos, métodos e interpretações	11
<u>SEÇÃO 2</u>	- Manipulação	21
<u>SEÇÃO 3</u>	- Teoremas de equivalência	27
<u>CAP. III</u>	<u>ESTRATEGIA</u>	45
<u>SEÇÃO 1</u>	- Métodos das direções viáveis	47
<u>SEÇÃO 2</u>	- Teoremas e resultados fundamentais	50
<u>SEÇÃO 3</u>	- Escolha de uma boa direção viável para o problema mestre. Otimização Unidire- cional	67
<u>SEÇÃO 4</u>	- Algoritmo primal de direções viáveis	80
<u>CAP. IV</u>	<u>ALGUNS ASPECTOS COMPUTACIONAIS</u>	83
<u>SEÇÃO 1</u>	- Diferenciabilidade das funções pertu- bação	84
<u>SEÇÃO 2</u>	- Procedimento de redefinição	104
<u>SEÇÃO 3</u>	- Convergência	107

<u>CAP. V</u>	IMPLEMENTAÇÃO	111
<u>SECÇÃO 1</u>	- O algoritmo para implementação	111
<u>SECÇÃO 2</u>	- Descrição das subrotinas	116
<u>SECÇÃO 3</u>	- Exemplo	150
<u>CAP. VI</u>	CONCLUSÕES	155
<u>APÊNDICE A</u>	DEFINIÇÕES E UM TEOREMA PARA FUNÇÕES CON- CAVAS	157
<u>APÊNDICE B</u>	LISTAGEM DOS RESULTADOS COMPUTACIONAIS	160
<u>BIBLIOGRAFIA</u>	189

C A P I T U L O I

INTRODUÇÃO

O desenvolvimento de técnicas eficientes de otimização para problemas estruturados, é de grande importância por suas aplicações à Engenharia, e à Economia, razão pela qual muitos artigos e livros foram publicados na última década sobre essa matéria.

A expressão problemas estruturados, normalmente, se associa aos problemas de grande porte, os quais queremos lembrar, não se caracterizam unicamente pela sua dimensão, mas pela combinação dimensão-estrutura. Aliás, explorar a estrutura em problemas de grande porte é frequentemente uma imposição para sua resolução.

Destacamos que a presença do grande número de variáveis e restrições, pode resultar não somente da estrutura intrínseca do problema, mas também de sua representação. Como observa Geoffrion [10], problemas com poucas não-linearidades, por exemplo, podem ser expressos por programas lineares através de aproximações de funções e conjuntos, aumentando sobremaneira a dimensão do problema, possibilitando, no entanto, sua resolução por técnicas conhecidas.

Entre os mais comuns tipos de estruturas encontradas em sistemas de grande porte, destacam-se: multidivisional, combinatorial, dinâmico e estocástico.

Particularmente nos dedicaremos aos problemas de grande porte de estrutura multidivisional ou decomponível.

Problemas decomponíveis consistem numa coleção de subsistemas interrelacionados a serem otimizados. Os subsistemas podem ser, por exemplo, módulos de um sistema de engenharia, reservatórios de um sistema de abastecimento de água, departamentos ou divisões de uma organização, setores de uma economia etc.

Uma classe bastante importante de problemas denominados problemas de alocação de recursos, são englobados pelos problemas de estrutura decomponível. Nesse caso, do ponto de vista da otimização, após definir-se o sistema produtivo, pretende-se determinar a alocação de recursos que leva ao melhor desempenho do sistema, o qual pode ser medido por funções objetivo. Nesse caso, o sistema decomponível estaria caracterizado por um conjunto de subsistemas que se encontram fracamente coesos pela necessidade dos recursos escassos.

Abaixo apresentamos interessante exemplo de otimização de um sistema decomponível.

Exemplo - Problema do corte de florestas : grande número de plantações são classificadas para o corte num período m anos. Cada subsistema é uma particular plantação ou floresta. As entradas nos subsistemas são representadas pelas capacidades de corte nos diferentes anos.

Suponhamos m' plantações e m anos. Seja :

C_{ij} : produção da plantação j se o corte efetuar-se no ano i ;

r_j : número de hectares da plantação j ;

a_i : capacidade de corte medida em hectares, no ano i ;

x_{ij} : fração da plantação j que será cortada no ano i ;

O problema do corte de florestas pode ser formulado como segue:

$$\text{Maximizar} \sum_{i=1}^m \sum_{j=1}^{m'} C_{ij} x_{ij} \quad \text{s.a}$$

$$\sum_{j=1}^{m'} r_j x_{ij} \leq a_i \quad (i = 1 \dots m)$$

$$(x_{1j}, \dots, x_{mj}) \in X_j = \left\{ (x_{1j}, \dots, x_{mj}) \mid \sum_{i=1}^m x_{ij} \leq 1 \wedge \right.$$

$$\left. \wedge x_{ij} \geq 0 \quad (i = 1 \dots m) \right\} \quad (j = 1 \dots m')$$

Métodos de coordenação

A distinção de subsistemas que operam quase independentemente, compondo um sistema de grande porte decomponível, conduz, em otimização, à possibilidade de se aplicar técnicas especiais para sua resolução. Essas técnicas, são sinteticamente apresentadas no capítulo II, onde se aborda a Teoria das Equipes e os Métodos de Coordenação.

Entre os métodos de coordenação há duas opções importantes, a Dual e a Primal. Ambas visam a resolução

dos problemas de otimização de sistemas decomponíveis, que brando-os numa série de subproblemas menores associados aos subsistemas. No capítulo II, são comentados os dois métodos de coordenação mencionados, dando-se maior destaque ao primal, que será o propósito do presente trabalho, que estuda os aspectos computacionais relativos a um algoritmo primal de direções viáveis.

Não nos restringiremos ao caso linear exemplificado pelo corte de florestas, assumiremos, no entanto, algumas hipóteses decididamente importantes, e comuns no tratamento de sistemas de grande porte. Tais hipóteses proporcionam-nos série de resultados, expostos no capítulo III, necessários para o estudo de um algoritmo desenvolvido por Geoffrion, e que se apresenta como escopo do presente trabalho.

Os problemas aqui estudados pressuporão as hipóteses de compacidade, convexidade-concavidade e diferenciabilidade, bastante comuns principalmente em problemas econômicos.

Algoritmos primais de direções viáveis

Uma classe de algoritmos bastante eficazes na resolução de problemas diferenciáveis e convexos de otimização, são os algoritmos de direções viáveis definidos no capítulo III. No tipo de problemas que estamos considerando, esses algoritmos não poderão, no entanto, ser aplicados diretamente ao problema original, devido a seu porte. Para tirarmos proveito da estrutura decomponível do problema de otimização proposto (problema que denominaremos primal ou original), com ganho no trabalho computacio-

nal, recaímos num problema manipulado cuja função objetivo não é, necessariamente, diferenciável, mesmo que todas funções do problema original o sejam. Referimo-nos às funções perturbação que terão importante papel nesse trabalho (definidas no apêndice A).

Como mostra em [11], Geoffrion, baseado nas propriedades de concavidade das funções perturbação associadas aos subsistemas, pôde derivar um programa linear, visando obter uma boa direção viável de incremento local para a função objetivo do problema manipulado.

Convergência.

Quando não é garantida a diferenciabilidade da função objetivo do problema de otimização, somente em alguns casos existem algoritmos de direções viáveis, cuja convergência foi provada, tratando-se de problema aberto noutros casos.

Recente trabalho desenvolvido por Hogan [16], levou à prova da convergência para um algoritmo de Frank-Wolfe no caso não diferenciável.

As propriedades de convergência do algoritmo que nos propusemos a estudar são problemas em aberto, razão pela qual tem merecido especial atenção por vários pesquisadores.

No capítulo IV como parte de nossas pesquisas, estudaremos as implicações computacionais advindas da diferenciabilidade local de algumas funções perturbação, bem como se expõe um critério para determinar a ocorrência da diferenciabilidade local da função perturbação.

Com a finalidade de aumentar a eficiência do

algoritmo proposto por Geoffrion, alteraremos, quando possível, a direção viável obtida pelo programa linear, que fornece a direção viável. O conceito de ϵ -atividade é também introduzido às restrições do referido programa.

Implementação.

A parte computacional do referido trabalho apresentada no capítulo V, levou à elaboração de uma subrotina, que obtém a direção viável para o problema manipulado, estando incluídas algumas das modificações objetivadas no capítulo IV.

Bibliografia.

Preferimos citar os trabalhos nos pontos da dissertação onde foram utilizados. No entanto, dada a importância de alguns trabalhos faz-se mister citá-los.

Como indicação em análise convexa temos Rockafellar[32] e Stoer e Witzgall[34]. Em programação matemática de grande porte Lasdon[21], além dos importantes trabalhos de Geoffrion[10],[11],[12] e[13].

Tratando mais dos aspectos específicos aos métodos de decentralização na resolução de problemas de allocação de recursos, citamos a tese de Jennergren[17]. Algoritmos de direções viáveis podem ser encontrados detalhados para programação linear como em não linear convexa em Zoutendijk[38]; na parte de algoritmos são também importantes os trabalhos de Polak[29] e Zangwill[37].

Notação.

O conjunto dos números naturais (não incluindo o número zero) e dos números reais serão aqui denotados respectivamente por \mathbb{N} e \mathbb{R} .

Os vetores serão matrizes colunas de números reais. O superíndice t associado a um vetor, indica que o mesmo foi transposto, passando a representar uma matriz linha de números reais. O superíndice t associado a uma matriz qualquer, indicará que estamos considerando a sua transposta.

Se x e y são vetores do \mathbb{R}^n , $x \leq y$ ($x \geq y$) indica que cada componente de x é menor ou igual (maior ou igual) à correspondente componente de y . Denotaremos $x > y$ ($x < y$) quando cada componente de x é maior (menor) que a correspondente componente de y .

O interior de um conjunto A será indicado por $\text{Int } A$, e a fronteira do mesmo conjunto por $\text{Front } A$.

Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Por $f(x)$, $x \in \mathbb{R}^n$, indicaremos o valor da função f no ponto x . Se, no entanto, quizermos explicitar que estamos considerando x como variável, denotaremos $f(\cdot)$. Por outro lado, se não houver qualquer preocupação em precisar qualquer dos dois sentidos acima, simplesmente denotaremos f . O gradiente da função f num ponto $x \in \mathbb{R}^n$ se existir, será um vetor coluna indicado por $\nabla f(x)$. A i -ésima componente desse vetor será indicada por $(\frac{\partial f}{\partial x_i})(x)$. A derivada direcional da função f no ponto $x \in \mathbb{R}^n$, e na direção $z \in \mathbb{R}^n$, será indicada por $f'(x; z)$.

Seja $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Essa função é um vetor cujas componentes são funções g_i ($i=1\dots m$) definidas no \mathbb{R}^n com valores em \mathbb{R} . Ou seja,

$$g(\cdot) = \begin{bmatrix} g_1(\cdot) \\ \vdots \\ g_i(\cdot) \\ \vdots \\ g_m(\cdot) \end{bmatrix}$$

Por $\nabla g(x)$, $x \in \mathbb{R}^n$, indicaremos, se existir, a matriz cuja i -ésima linha é constituída pelas componentes de $\nabla g_i(x)$.
Isto é,

$$\nabla g(x) = \begin{bmatrix} \nabla g_1^t(x) \\ \vdots \\ \nabla g_i^t(x) \\ \vdots \\ \nabla g_m^t(x) \end{bmatrix}$$

Portanto, $\nabla g(x)$ é uma matriz $m \times n$.

Se A for uma matriz, o elemento da I -ésima linha e J -ésima coluna será indicado por

$$A(I, J)$$

O símbolo \triangleq indicará igual por definição, sendo a expressão isto é indicada sinteticamente por i.e. .

Usaremos para supremo a indicação Sup, para ínfimo, Inf, para máximo, max, e para mínimo, min.

O vetor nulo é indicado por 0.

Em relação às referências temos:

- capítulos serão numerados com algarismos romanos;
- expressões e subtítulos serão numerados com algarismos arábicos.

A referência a uma expressão ou subtítulo do mesmo capítulo, é feita colocando-se o número de referência entre parênteses (por exemplo, (13)). Referências a expressões ou subtítulos em outros capítulos, serão feitas grafando-se o número do capítulo (em algarismos romanos), um ponto, o número da secção, um ponto, e o número da expressão ou subtítulo (por exemplo, III.3.44). Referências bibliográficas são apresentadas entre colchêtes (por exemplo, Geoffrion [10]).

A seguir daremos uma breve nota a respeito do apêndice A.

Apêndice A.

Antes de iniciarmos a apresentação do presente trabalho, será interessante indicar que as definições bem como alguns teoremas que utilizaremos, podem ser encontrados no Apêndice A.

C A P Í T U L O I I

O PROBLEMA GERAL DE ALOCAÇÃO DE RECURSOS, MÉTODOS, MANIPULAÇÃO E TEOREMAS DE EQUIVALENCIA.Introdução.

Inicialmente, enuncia-se o problema geral de alocação de recursos, o qual é interpretado segundo três formas diversas de abordagens, primal, dual e teoria das equipes, optando-se pelo desenvolvimento do primal.

A secção 2 coloca o problema geral de alocação de recursos numa forma bastante adequada do ponto de vista computacional, gerando o problema mestre de alocação de recursos, possível pelo emprêgo de uma técnica de manipulação: projeção.

Conveniente é a secção que se segue, já que demonstra a equivalência entre os problemas geral de alocação de recursos e o problema mestre, no sentido dos teoremas a serem apresentados, salientando-se a perda parcial da diferenciabilidade.

SEÇÃO 1 - PROBLEMA GERAL DE ALOCAÇÃO DE RECURSOS, MÉTODOS E INTERPRETAÇÕES.

Consideremos o problema geral de alocação de

recursos, que se identifica com a otimização de um sistema de grande porte formado por subsistemas interrelacionados*, cuja estrutura particular, multidivisional, é frequentemente referida na teoria de programação matemática como decomponível. Portanto associaremos a palavra decomponível a um sistema cuja otimização é um problema da forma:

1 Maximizar $\sum_{i=1}^k f_i(x_i)$ s.a

$$\sum_{i=1}^k g_i(x_i) \leq b \quad e$$

$$(\forall i \in I_k)(x_i \in X_i)$$

$$I_k = \{1, 2, \dots, k\} ;$$

$x_i \in \mathbb{R}^{n_i}$; vetor atividade de dimensão n_i , associado com o i -ésimo subsistema;

$X_i \subset \mathbb{R}^{n_i}$, representando o conjunto dos programas de operação (conjunto ou região viável) para o i -ésimo subsistema, associado com x_i ;

* Por subsistemas interrelacionados poderemos, neste caso, entender como subsistemas quase autônomos, que se acham dependentes a um comando central, que mantém a estrutura não separável, pela distribuição de recursos necessários ao funcionamento de cada um. Se, no entanto, cada subsistema não carecesse de tais recursos para o completo funcionamento, o sistema seria completamente separável em subsistemas cada um somente dependente ao seu próprio comando central.

$f_i : X_i \rightarrow \mathbb{R}$, função objetivo do i -ésimo subsistema associada com x_i ;

$g_i(\cdot) = (g_{i1}(\cdot), \dots, g_{im}(\cdot))$, função vetorial cujo domínio é X_i , associada com x_i , $g_i : X_i \rightarrow \mathbb{R}^m$; $b \in \mathbb{R}^m$, vetor m -dimensional que denominaremos vetor-recursos cujas componentes são os recursos comuns aos subsistemas.

2 Hipóteses Adicionais.

Inicialmente iremos supor as hipóteses de concavidade, convexidade e compactidade, além de todas variáveis x_i pertencerem aos espaços euclidianos \mathbb{R}^{n_i} . Destarte, as funções f_i são supostas côncavas nos conjuntos convexos e compactos X_i , e cada componente $g_{ij}(\cdot)$ ($j=1 \dots m$) da função vetorial g_i ($i=1 \dots k$) será suposta convexa em X_i ($i=1 \dots k$). Finalmente, supomos que o programa (1) tem uma solução viável. Outras hipóteses sobre as funções serão impostas quando necessário (III.2).

As funções f_i ($i=1 \dots k$) medem a representatividade de cada subsistema, ou seja, a sua contribuição para o benefício global, sendo diretamente dependente do programa de produção do subsistema, razão pela qual devem ser convenientemente obtidas.

O sistema opera com dois tipos de restrições. O primeiro, expresso pela desigualdade vetorial, $\sum_{i=1}^k g_i(x_i) \leq b$, advém do uso comum de certos recursos pelos subsistemas, sendo por isso chamada restrição de acoplamento ou de corporação. Podemos então interpretar $g_i(\cdot)$ ($i=1 \dots k$) como função vetorial que mede o nível de recurso de acoplamento (suposto vetorialmente), necessário

para manter o subsistema operando segundo o programa x_i . Conseqüentemente, a restrição de acoplamento exprimirá, via vetor-recursos, a disponibilidade dos recursos de utilidade comum. O segundo tipo de restrições associa-se aos conjuntos X_i ($i=1\dots k$). O conjunto X_i contém os pontos nos quais o i -ésimo subsistema pode operar, por isso é frequentemente denominado conjunto (ou região) viável (ou de operação), sendo particular para cada subsistema, razão pela qual é genéricamente representado. Na teoria da alocação de recursos é associado com as restrições e recursos divisionais (ou dos subsistemas), sendo dependente da tecnologia do subsistema.

A resolução do problema (1) pode ser encarado do ponto de vista da teoria das equipes. Definamos equipe: uma equipe, como estamos acostumados a pensar, refere-se a um grupo de membros de uma organização, que controlam determinadas variáveis em campos de ação diversos, e sob diferentes informações tomam suas decisões, estando, no entanto, ligados pelo mesmo objetivo. Uma equipe acerca-se de dois problemas fundamentais:

- 1º) Quais as possíveis opções de estruturas informacionais;
- 2º) Como deve cada membro atuar de modo a maximizar a comum preferência.

Uma equipe, ao contrário do que ocorre nos métodos de coordenação (a serem posteriormente discutidos), não se preocupa com a estrutura particular da função objetivo. Preocupa-se, entretanto, em derivar explicitamente o custo da transferência de informações entre seus membros na busca do ótimo. Citamos Marchak e Radner [25] como im

portante referência é Varaiya^[35], que possui vasta bibliografia sobre o assunto.

Desde que diferentes caminhos são acionados a fim de se obter um mesmo objetivo (maximizar função objetivo ou preferência comum do sistema), parece-nos interessante a consideração do ponto de vista da Teoria das Equipes, de vez que se pode pensar na hibridização desse método com os de coordenação, ou simplesmente, utilizar resultados qualitativos de certos métodos no melhor entendimento dos caminhos gerados pelos outros.

A particular estrutura de um sistema decomponível (função objetivo e restrições de acoplamento linearmente separáveis por subsistemas), leva a possibilidade de se obter a situação global ótima (resolver-se (1)), a partir da coordenação das situações ótimas de cada subsistema (subotimizações). Os métodos ligados às técnicas de coordenação geram subproblemas parametrizados separáveis (i.e., de resoluções independentes), e se desenvolvem da idéia de determinar valores dos parâmetros, para os quais alguma contribuição das subotimizações leva ao ótimo global. A existência de tais situações conduz aos valores ótimos dos parâmetros. Tais parâmetros bem como a forma dos subproblemas derivam da particular escolha do método de coordenação. Há dois procedimentos principais para se coordenar as respostas ótimas dos subsistemas parametrizados:

- dual ;
- primal.

Nosso trabalho se preocupará mais de perto com o método de coordenação via primal, que se diferencia do dual basicamente por que, naquele são distribuídos (sem a

possibilidade de consumo extra) recursos, e neste são fixados os preços dos mesmos, deixando-se o nível de aquisição a cargo do subsistema. Passemos a uma breve descrição da coordenação via dual.

Coordenação via Dual.

O dual envolve a determinação iterativa de um vetor do \mathbb{R}^m , tais que as soluções ótimas dos subproblemas λ -parametrizados,

$$\text{Maximizar } f_i(x_i) - \lambda^t g_i(x_i) \\ x_i \in X_i$$

resolve (1). Essa forma de se tentar a solução ótima, há muito tem merecido especial atenção pelos economistas, que primitivamente concebiam uma economia miniaturizada (March e Simon [24] ; Hirshleifer [14]) com sistema de preços, referentes a uma organização. Essa forma de encarar o problema revitalizou-se com a descoberta de que os multiplicadores de Lagrange dos problemas de programação matemática, poderiam ser convenientemente interpretados como preços. O vetor m -dimensional λ em (3), é assim, frequentemente chamado e interpretado como o vetor-preços dos recursos de acoplamento, denominação que se lhe cabe natural da interpretação a seguir.

Suponhamos para efeito de interpretação, uma companhia constituída de n divisões e uma direção central. Uma determinada iteração representa a fixação, pela direção central, do preço λ , para os recursos de acoplamento. O desejo das divisões é de obter um programa x_i que maximize a eficiência de utilização de tais recursos, medida

pela diferença entre a contribuição da divisão (ou subsistema), $f_i(x_i)$, e o preço pago pelo nível de recurso comum (ou de acoplamento) requerido, $\lambda^t g_i(x_i)$. A subotimização (3) ministra, então, o programa, $x_i^* \in X_i$, mais conveniente ao subsistema. A direção central é informada dos níveis de recurso, $g_i(x_i^*)$ ($i=1\dots k$), requeridos, associando aos recursos não totalmente utilizados o preço zero, recalculando, convenientemente, os demais. A reatualização dos preços dos recursos leva, dependendo do algoritmo particularmente utilizado, à resolução de diferentes problemas, que podem ser chamados problemas diretores, pois se acham ao nível da direção central. Novos preços são fixados e novas subotimizações são efetuadas. A troca de informações direção-divisões pretende, iterativamente, obter os preços que optimizam o funcionamento da companhia, i.e., os preços ótimos.

A forma de reatualizar os preços pressupõe que a direção possa, se necessário, adquirir novas quantidades de recursos, para suprir o nível total de recursos de acoplamento, $\sum_{i=1}^k g_i(x_i^*)$, requerido numa determinada iteração. Fica desde já, evidente, que a viabilidade primal de utilização dos recursos comuns não é considerada. Diz-se em programação matemática que não é, garantidamente, mantida a viabilidade primal. Sob determinadas condições (Geoffrion [12]), a situação global ótima iterativamente conseguida por esse tipo de coordenação, não inviabiliza a restrição de disponibilidade dos recursos comuns, garantindo, assim, a viabilidade primal da solução obtida.

Dependendo da estrutura informacional (interação direção central-divisões) particularmente considerada,

derivam-se diversas teorias (centralização e decentralização) a respeito dos métodos usados.

Como cita Jennergren[17], o modelo organizacional utilizado para interpretação econômica do dual, bem como a terminologia usada, fora originalmente empregada por Baumol e Fabian[4] e também por Zschau[39]. Interpretação usada por Gale[8], Kornai e Liptak[19], Malinvaud[22] e Weitzman[36] é quase equivalente. Esses autores supõem uma economia centralmente dirigida, constituída de uma agência de planejamento e n firmas ou companhias. Ainda devendo ao mesmo pesquisador, poucos algoritmos se fundamentam na idéia de ajustamento de preços(dual).

A séria desvantagem advinda da aplicação de algoritmos do tipo ajuste de preços é que, não se garante a viabilidade das restrições de acoplamento.

Coordenação via Primal.

A par da dificuldade anteriormente exposta, foram desenvolvidos os métodos primais, i.e., aqueles que garantem a viabilidade das restrições referentes ao problema originalmente proposto.

Esses métodos de ataque ao problema original (no presente caso (1)), tiveram desenvolvimento bastante rápido a partir da descoberta do Princípio da Decomposição de Dantzig-Wolfe, visando quebrar um problema numa série de subproblemas menores. Criaram no âmbito da programação matemática, a Teoria da Decomposição que se desenvolveu sobremodo na década de 60, conduzindo cada vez mais a novos caminhos e aplicações, dada a atenção especial que atualmente se lança ao Estudo da Otimização de Sistemas de

Grande Porte.Varaiya[35]em recente pesquisa sobre a tendência da teoria da decisão em sistemas de grande porte, vem ratificar a importância desse campo da Teoria da Programação Matemática.

Nossa dissertação versará sobre o método de coordenação primal de direcionamento de recursos devido a Geoffrion [11].

A interpretação via primal tornará clara a garantia da viabilidade das soluções geradas por ele. No presents> caso, as informações transferidas para as divisões (supondo mantermos o mesmo modelo institucional que no dual), alteram-se. Neste caso, o planejador não fixa os preços, mas aloca diretamente o limite superior do nível de recurso para cada divisão (este fato determina o nome comumente usado na denominação do problema (1): Problema Geral de Alocação de Recursos). Desta forma, o subsistema é obrigado a otimizar seu funcionamento enquanto se mantém, garantidamente, a viabilidade do problema (1) face a disponibilidade de recursos comuns (representada pelo vetor-recursos b).

Da interpretação estabelecida, dois fatos podem, intuitivamente, determinar o caminho pelo qual trataremos o problema (1). Em primeiro lugar, a alocação de recursos comuns aos subsistemas leva cada subsistema à independência de funcionamento, podendo, assim, operar da melhor maneira possível, sem qualquer interação com os demais. Frisamos que isso é possível dada a estrutura decomponível do sistema. Cada subsistema, consequentemente, optimiza seu funcionamento pelos meios mais adequados. Como veremos na seção 2, esse aspecto particular poderá ser explorado aplicando

ao problema (1) a técnica de manipulação chamada projeção. Finalmente, parece não difícil notar que a alocação de recursos poderá resultar da obtenção de direções viáveis, i.e., o coordenador vê o sistema funcionando segundo determinada alocação de recursos, tenta redefinir os níveis superiores de recursos comuns para cada subsistema, levando em conta o problema fundamental:

de posse das informações divisionais a respeito do funcionamento ótimo para os níveis de recursos fixados, em quais "direções" é possível deslocar a oferta de recursos comuns, a fim de melhorar o funcionamento global do sistema.

Essa forma de raciocinar nos levará a aplicar a estratégia de direções viáveis ao problema manipulado (i.e., ao problema, neste caso, pronto para ser desacoplado) com a consequente separabilidade de cada subsistema. O processo é, portanto, iterativo e pretende levar ao ótimo global pela troca constante de informações entre as divisões e a coordenação; a coordenação procura direções sobre as quais poderá alterar as ofertas de recursos para cada divisão, a fim de melhorar a situação global. Destarte, a coordenação pelo método primal consiste em se obter k vetores do \mathbb{R}^m tais que as soluções dos subproblemas y_i -parametrizados ,

$$\begin{aligned} \text{Maximizar } f_i(x_i) \quad \text{s.a } g_i(x_i) \leq y_i \\ x_i \in X_i \end{aligned}$$

resolva (1). O vetor $y_i \in \mathbb{R}^m$ representa o nível de recursos alocados ao i -ésimo subsistema.

Neste caso, qualquer regra de parada conveniente usada, gera solução viável não pior que a original (i.e., solução viável utilizada para partida do algoritmo).

Como teremos oportunidade de verificar o algoritmo de decomposição, que será implementado, não pode ser considerado como representante de um Método de Decentralização, embora seja computacionalmente forte : diminui o trabalho computacional, ao mesmo tempo que possibilita a resolução de problemas cujo porte torna inviável a computação por métodos diretos.

A secção 2 preocupar-se-á com a manipulação do programa (1) visando derivar um programa mestre, cuja estrutura decomponível poderá, posteriormente, ser explorada pela aplicação da estratégia das direções viáveis (capítulo III).

SECÇÃO 2 - MANIPULAÇÃO.

Frequentemente sucedem desacertos práticos na resolução direta de um problema de grande porte. Tais problemas podem, dependendo da particular estrutura que englobem, ser transformados noutros equivalentes por meio de manipulações que objetivam [10 ; pg 12] :

- (a) explorar a estrutura particular do problema;
- (b) linearizar as partes não lineares do problema;
- (c) induzir separação.

Manipulando-se o problema (1) por meio da técnica de projeção, o item (a) conduz a (c), devido a estrutura decomponível do sistema. O programa obtido dessa manipulação é referido na literatura como problema mestre.

Reescrevendo o problema (1) após introduzir os vetores $y_1 \dots y_k$ de \mathbb{R}^m , vem:

4

$$\begin{aligned} & \text{Maximizar}_{x,y} \quad \sum_{i=1}^k f_i(x_i) \quad \text{s.a} \\ & \sum_{i=1}^k y_i \leq b \quad \wedge \quad (\forall i \in I_k)(x_i \in x_i \wedge g_i(x_i) \leq y_i) \end{aligned}$$

O problema (4) difere do problema (1) por acoplá-lo os subsistemas através das variáveis, y_i ($i=1 \dots k$), introduzidas, restando um problema com variáveis de acoplamento, ao invés de restrições de acoplamento. Como poderemos notar mais abaixo, y_i , poderá ser interpretado como recurso direcionado ao i -ésimo subsistema. Se $y = (y_1, \dots, y_k)$ for temporariamente fixado, obteremos um problema separável.

Projetando o problema (4) ([10; pg 11]) no espaço das variáveis y , vem:

5

$$\begin{aligned} & \text{Maximizar}_y \quad \left[\sup \left\{ \sum_{i=1}^k f_i(x_i) : \left| \dots (\forall i \in I_k)(x_i \in x_i \wedge g_i(x_i) \leq y_i) \right. \right\} \right] \text{s.a} \\ & \sum_{i=1}^k y_i \leq b \end{aligned}$$

ou ainda, observando-se a separabilidade linear da função

ção objetivo do programa (1), vem:

$$6 \quad \begin{aligned} & \text{Maximizar} \quad \sum_{i=1}^k \sup \left\{ f_i(x_i) \text{ s.a. } x_i \in X_i \wedge g_i(x_i) \leq y_i \right\} \\ & \text{s.a.} \quad \sum_{i=1}^k y_i \leq b \end{aligned}$$

A respeito do programa (6) teceremos alguns comentários. Em primeiro lugar, há de se notar a forma implícita das restrições $y_i \in Y_i$ ($i=1 \dots k$), onde:

$$Y_i \triangleq \left\{ y_i \in \mathbb{R}^m \mid (\exists x_i \in X_i) (g_i(x_i) \leq y_i) \right\} \quad (i=1 \dots k)$$

Ao invés de torná-las explícitas, pode-se optar pela forma alternada (implícita) que convenciona o valor $-\infty$ para o supremo de uma função num conjunto vazio.

A interpretação econômica anteriormente dada ao problema primal (1) fica, observando-se (6), naturalmente introduzida. O planejador fixa uma alocação de recursos dentro da disponibilidade prevista (fixação de \bar{y} tal que $\sum_{i=1}^k \bar{y}_i \leq b$); a seguir, cada subsistema optimiza seu funcionamento (obtenção de $\sup \{ f_i(x_i) \mid x_i \in X_i \wedge g_i(x_i) \leq \bar{y}_i \}$ para $i=1 \dots k$), sendo que supremos no valor $-\infty$ refletem a inadequação dos níveis de recursos distribuídos; os subsistemas comunicam à coordenação suas respostas ótimas, pondo a coordenação a par do resultado global ($\sum_{i=1}^k \sup \{ f_i(x_i) \mid x_i \in X_i \wedge g_i(x_i) \leq \bar{y}_i \}$). O recálculo das novas alocações com base nas respostas ótimas dos sub-sistemas é de tratamento mais delicado, razão pela qual, após obtermos o programa (9), passaremos a tal discussão.

Supondo-se $v_i(y_i)$ ¹ o valor ótimo² do programa parametrizado em y_i :

$$8 \quad P_{y_i}^i \quad \text{Maximizar} \quad f_i(x_i) \quad \text{s.a} \quad g_i(x_i) \leq y_i \\ x_i \in X_i$$

o problema (6) poderá ser colocado na forma equivalente (9), como se mostrará em (10),

$$9 \quad \text{Maximizar} \quad \sum_{i=1}^k v_i(y_i) \quad \text{s.a} \quad \sum_{i=1}^k y_i \leq b$$

Observação. As hipóteses (2) não garantem a existência da solução ótima de (8). Com as hipóteses de diferenciabilidade sobre as funções f_i , g_i e $y_i \in \text{Int } Y_i$ necessárias ao desenvolvimento de (III.2) e mais compactade de X_i (imposta em (2)), garante-se a existência para a solução ótima para (8), e consequentemente, valor finito para $v_i(y_i)$.

O programa (9) obtido da projeção de (1), representa nesse caso o programa mestre, sendo uma forma modificada de se apresentar o problema geral de alocação de recursos (1).

Do ponto de vista da hierarquização distinguem-se dois níveis:

¹ A função $v_i(y_i) = \sup_{x_i \in X_i} \{f_i(x_i) \text{ s.a } g_i(x_i) \leq y_i\}$ é denominada função perturbação associada com $P_{y_i}^i$, sendo, portanto, característica do subsistema considerado. Essa função será fundamental no presente trabalho.

² Valor ótimo do programa $P_{y_i}^i$ é o valor de $v_i(\cdot)$ no ponto $y_i \in \mathbb{R}^m$.

- 1) nível de coordenação, representado pelos deveres explícitos em (9);
- 2) nível de divisão, representado, implicitamente, em (9), pelas subotimizações $P_{y_i}^i$ ($i=1\dots k$).

O programa mestre obtido possui explicitamente, somente, restrições lineares sendo a função objetivo não linear geral. De forma implícita temos as restrições não lineares $y_i \in Y_i$ ($i=1\dots k$) representando a viabilidade do recurso comum alocado para o i -ésimo subsistema. Excetuando-se problemas lineares ou quadráticos, expressões explícitas para v_i ($i=1\dots k$) são praticamente impossíveis, a menos que a dimensão do vetor-recursos b seja muito pequena [21]. No caso geral, valores de v_i são dependentes da resolução dos subproblemas $P_{y_i}^i$ ($i=1\dots k$). Porém, a vantagem de (9) em relação a (1) torna-se particularmente interessante quando o número total de restrições que determinam os conjuntos X_i ($i=1\dots k$) é grande em relação ao número de restrições que acoplam os subsistemas, já que a aplicação de estratégia conveniente permite obter a solução de (1), pelas resoluções independentes dos problemas $P_{y_i}^i$ ($i=1\dots k$). Fica, portanto, decidido que nos dedicaremos a métodos que visam a resolução iterativa de (9) através de série de otimizações.

Realocação de Recursos.

Como anteriormente citamos, passamos a interpretar mais de perto a realocação de recursos com base no programa (9). Suponhamos que a resposta ótima do

subsistema seja, agora, dada na forma de produtividade marginal, i.e., variação da contribuição ótima com a variação unitária dos recursos alocados (essas quantidades são referidas, para um determinado subsistema, aos recursos comuns, componentes do vetor b). Tais quantidades capacitam o coordenador, com base na disponibilidade dos recursos comuns, a obter uma direção (pode existir uma infinidade) de realocação que conduza ao maior incremento local na contribuição global.

Observação : não interessam ao coordenador valores ótimos das contribuições dos subsistemas, mas tão somente informações que lhe permitam deduzir uma nova direção de realocação, que conduza a melhora global.

Sucessivas subotimizações e realocações prosseguem, até que os efeitos das variações marginais dos recursos comuns na contribuição ótima, sejam equiparados para todas divisões (condição de optimidade quando v_i é diferenciável, $i=1\dots k$, que provaremos no capítulo IV).

Quando as funções v_i forem diferenciáveis, as produtividades marginais do recurso j para o subsistema i em y_i , são expressas pelas componentes dos gradientes de $v_i(\cdot)$ em y_i , i.e., $\left(\frac{\partial v_i}{\partial y_{ij}}\right)(y_i)$ representará a produtividade marginal do recurso j ($j=1\dots m$) relativo ao subsistema i ($i=1\dots k$).

Acontece, no entanto, que $v_i(\cdot)$ não é necessariamente diferenciável em Y_i (Rockafellar[32]). Essa dificuldade é parcialmente removida dada a concavidade das funções v_i ($i=1\dots k$), como veremos na secção 3, levando à existência das derivadas direcionais em todas as direções, e em todos pontos interiores de Y_i . É esse fato

que permitirá o desenvolvimento do trabalho mediante a estratégia das direções viáveis. Isso mostra desde já o importante papel exercido pelas hipóteses côncavo-convexas.

A maior contribuição de Geoffrion foi derivar um programa linear explícito, que mediante algumas suposições, leva à obtenção da direção viável que maximiza o aumento inicial da função objetivo do programa (9), i.e., a direção obtida conduz ao valor máximo da derivada direcional da função objetivo de (9).

Notemos finalmente que o tratamento pode ser simplificado nos pontos onde $v_i(\cdot)$ ($i=1..k$) é diferenciável. Esses pontos não são exceções nos conjuntos Y_i ($i=1..k$), já que os pontos em que v_i ($i=1..k$) é não diferenciável constituem um subconjunto de Y_i ($i=1..k$) de medida nula. Por essa razão será exposta no capítulo IV, pesquisa visando obter a melhor direção viável de aumento da função objetivo de (9), tirando proveito da diferenciabilidade das funções v_i , onde isso ocorrer. Tal fato reduzirá consideravelmente as dimensões do programa linear utilizado na busca dessas direções viáveis.

A secção seguinte conduz-nos a resultados bastante adequados, ou seja, a equivalência entre os programas (1) e (9) no sentido dos teoremas a serem demonstrados. Embora perdida parcialmente a diferenciabilidade, resta-nos a possibilidade de aplicar estratégia conveniente (direções viáveis), que leva a uma sequência de subprogramas côncavos diferenciáveis.

SECÇÃO 3 - TEOREMAS DE EQUIVALENCIA.

Dois teoremas de fundamental importância serão aqui desenvolvidos. O teorema da equivalência propriamente dito, mostrará como se relacionam as soluções ótimas de (1), problema cuja resolução fora inicialmente proposta, e, o problema (9) resultado da projeção de (1) sobre o espaço das variáveis y . Embora este teorema seja de uso frequente, sua demonstração não se encontra na literatura geralmente citada. O segundo teorema refere-se a propriedades dos conjuntos Y_i ($i=1\dots k$) e à concavidade das funções v_i ($i=1\dots k$) nesses conjuntos [10; pg 16].

O teorema de equivalência propriamente dito será provado para o caso geral, para tanto não serão necessárias hipóteses de concavidade-convexidade e compacidade.

As hipóteses utilizadas para obtenção das propriedades dos conjuntos Y_i e funções v_i serão mencionadas nas respectivas provas.

Para a prova do teorema abaixo, explicitaremos os conjuntos Y_i associados ao programa mestre (9).

10 Teorema. São válidas as quatro asserções abaixo:

- A) O programa (1) é inviável se, e somente se, (9) também o fôr;
- B) O programa (1) tem valor ótimo+∞ se, e somente se, (9) também o tiver;
- C) Se o programa (1) tiver solução ótima $x^* = (x_1^*, \dots, x_k^*)$ então (9) terá solução ótima y^* , e, x_i^* ($i=1\dots k$) será solução ótima de $P_{y_i^*}^i$ ($i=1\dots k$), sendo $g_i(x_i^*) \leq y_i^*$ ($i=1\dots k$) e $\sum_{i=1}^k y_i^* \leq b$;
- D) Se (9) tiver solução ótima y^* e x_i^* ($i=1\dots k$)

\bar{x} for solução ótima de $P_{y_i^*}^i$ ($i=1\dots k$), então $\bar{x}^* = (x_1^*, \dots, x_k^*)$ será solução ótima de (1).

Para as demonstrações abaixo seja:

$$\begin{aligned} x &= (x_1, \dots, x_k) \in \mathbb{R}^{\sum_{i=1}^{n_i}} \text{ onde } (\forall i \in I_k)(x_i \in \mathbb{R}^{n_i}) \\ x &= x_1 \times \dots \times x_k \subset \mathbb{R}^{\sum_{i=1}^{n_i}} \\ Y &= Y_1 \times \dots \times Y_k \subset \mathbb{R}^{K_m} \text{ onde } (\forall i \in I_k)(Y_i \subset \mathbb{R}^{m_i}) \\ y &= (y_1, \dots, y_k) \in \mathbb{R}^{K_m} \text{ onde } (\forall i \in I_k)(y_i \in \mathbb{R}^{m_i}) \\ \Sigma &= \left\{ y \in \mathbb{R}^{K_m} \mid \sum_{i=1}^k y_i \leq b \right\} \end{aligned}$$

Asserção A. Formalmente teremos:

$$(\exists x \in X) \left(\sum_{i=1}^k g_i(x_i) \leq b \right) \iff (\exists y \in Y) \left(\sum_{i=1}^k y_i \leq b \right)$$

ou equivalentemente,

$$(\exists x \in X) \left(\sum_{i=1}^k g_i(x_i) \leq b \right) \iff (\exists y \in Y) \left(\sum_{i=1}^k y_i \leq b \right)$$

Demonstração. \implies : seja, por hipótese, \bar{x} viável em (1), i.e.,

$$11 \quad \bar{x} \in \left\{ x \in X \mid \sum_{i=1}^k g_i(x_i) \leq b \right\}$$

Definindo $\bar{y}_i = g_i(\bar{x}_i)$ para todo $i \in I_k$, obtém-se de (11) que

$$12 \quad \sum_{i=1}^k \bar{y}_i \leq b$$

e, pela definição de \bar{y}_i ($i \in I_k$),

$$13 \quad (\forall i \in I_k)(\bar{y}_i \in Y_i)$$

Logo de (12) e (13) \bar{y} é viável em (9), ou seja,

$$\bar{y} \in \{y \in Y \mid \sum_{i=1}^K y_i \leq b\}$$

o que completa a prova da condição suficiente.

\iff : seja, por hipótese, \bar{y} tal que

$$\bar{y} \in \sum_n Y$$

Como $\bar{y} \in Y$ vem

$$14 \quad (\forall i \in I_K)(\exists \bar{x}_i \in x_i)(g_i(\bar{x}_i) \leq \bar{y}_i)$$

Da condição $\bar{y} \in \sum$ e (14) decorre

$$15 \quad \sum_{i=1}^K g_i(\bar{x}_i) \leq b$$

De (14) e (15) vem:

$$\bar{x} \in \{x \in X \mid \sum_{i=1}^K g_i(x_i) \leq b\}$$

Δ

Asserção B. Se v0a representar o valor ótimo do programa (a), teremos:

$$v01 = +\infty \iff v09 = +\infty$$

Demonstração. \implies : supondo, por hipótese, $v01 = +\infty$, temos (1) é viável. Para cada \bar{x} viável em (1) façamos:

16

$$(\forall i \in I_K) (\bar{y}_i = g_i(\bar{x}_i))$$

Da viabilidade de \bar{x} em (1) e definição de \bar{y} , vem:

$$\sum_{i=1}^K \bar{y}_i = \sum_{i=1}^K g_i(\bar{x}_i) \leq b$$

$$\bar{y} \in Y$$

Logo para cada \bar{x} viável em (1), temos:

17

$$\Theta(\bar{x}) \triangleq \{ \bar{y} \mid (\forall i \in I_K) (g_i(\bar{x}_i) = \bar{y}_i) \} \subset Y \cap \sum$$

Como por (16)

$$(\forall i \in I_K) (\bar{x}_i \text{ é viável em } P\bar{y}_i^i)$$

vem,

$$v_i(\bar{y}_i) \triangleq \sup \{ f_i(x_i) \text{ s.t. } x_i \in X_i \wedge g_i(x_i) \leq \bar{y}_i \} \triangleq f_i(\bar{x}_i)$$

para todo $i \in I_K$.

Logo para \bar{x} viável em (1),

$$\sum_{i=1}^K v_i(\bar{y}_i) \geq \sum_{i=1}^K f_i(\bar{x}_i)$$

e, portanto,

18

$$\sup \left\{ \sum_{i=1}^k v_i(\bar{y}_i) \mid \bar{y} \in \Theta(\bar{x}) \right\} \geq \sup \left\{ \sum_{i=1}^k f_i(x_i) \mid \bar{x} \text{ viável (1)} \right\} = +\infty$$

De (17) e (18), vem

$$\sup \left\{ \sum_{i=1}^k v_i(y_i) \mid y \in Y \cap \Sigma \right\} \geq \sup \left\{ \sum_{i=1}^k v_i(\bar{y}_i) \mid \bar{y} \in \Theta(\bar{x}) \right\} = +\infty$$

logo,

$$v09 = +\infty$$

\iff : como, por hipótese, $v09 = +\infty$, (9) é viável. Seja $\bar{y} = (\bar{y}_1, \dots, \bar{y}_k) \in Y \cap \Sigma$. Seja:

$$\begin{aligned} \Theta(\bar{y}) &= \left\{ x \in X \mid \left(\forall i \in I_k \right) (g_i(x_i) \leq \bar{y}_i) \right\} \\ A &= \left\{ x \in X \mid \sum_{i=1}^k g_i(x_i) \leq b \right\} \end{aligned}$$

O conjunto $\Theta(\bar{y})$ é não vazio dado que $\bar{y} \in Y$, sendo evidente a inclusão $\Theta(\bar{y}) \subset A$, para todo $\bar{y} \in Y \cap \Sigma$. Logo,

19

$$\begin{aligned} \sum_{i=1}^k v_i(\bar{y}_i) &\stackrel{\text{def}}{=} \sup \left\{ \sum_{i=1}^k f_i(x_i) \mid x \in \Theta(\bar{y}) \right\} \leq \\ &\leq \sup \left\{ \sum_{i=1}^k f_i(x_i) \mid x \in A \right\} \quad \text{para todo } \bar{y} \in Y \cap \Sigma. \end{aligned}$$

Como $Y \cap \Sigma$ é fechado, decorre de (19) :

$$\sup \left\{ \sum_{i=1}^k v_i(\bar{y}_i) \mid \bar{y} \in Y \cap \Sigma \right\} \leq \sup \left\{ \sum_{i=1}^k f_i(x_i) \mid x \in A \right\}$$

ou seja,

$v_{09} \leq v_{01}$

logo

$v_{01} = +\infty$

△

Asserção C. Se $x^o = (x_1^o, \dots, x_k^o)$ é solução ótima de (1), então existe y^o solução ótima de (9) satis fazendo $g_i(x_i^o) \leq y_i^o$ ($i=1 \dots k$) e $\sum_{i=1}^k y_i^o \leq b$, e, x_i^o é solução ótima de $P_{y_i^o}$ para todo $i \in I$.

Demonstração. Como x^o é viável em (1), fazendo para todo $i \in I_k$, $g_i(x_i^o) = y_i^o$ vem,

$$y^o \in \Sigma \cap Y$$

e, portanto, y^o é viável em (9).

a) Provemos que x_i^o é ótimo em $P_{y_i^o}$ para todo $i \in I_k$, por redução ao absurdo, i.e.,

$$(\exists i \in I_k) (\bar{x}_i \text{ é viável em } P_{y_i^o} \wedge f_i(\bar{x}_i) > f_i(x_i^o))$$

decorre que

$$x = (x_1^o, \dots, x_{i-1}^o, \bar{x}_i, x_{i+1}^o, \dots, x_k^o) \text{ é viável em (1).}$$

Como,

$$\sum_{\substack{j=1 \\ j \neq i}}^K f_j(x_j^*) + f_i(\bar{x}_i^*) > \sum_{i=1}^K f_i(x_i^*)$$

conclui-se que x^* não é solução ótima de (1), o que contradiz a hipótese. Logo, ($\forall i \in I$) (x_i^* é solução ótima de $P_{y_i^*}^i$), e portanto, $\forall i$ ($y_i^* = f_i(x_i^*)$).

b) Provemos que \bar{y}^* é solução ótima de (9) por reductio ad absurdum, i.e.,

20 $(\exists \bar{y} \in Y \cap \Sigma) (\sum_{i=1}^K v_i(\bar{y}_i) > \sum_{i=1}^K v_i(y_i^*))$

Definindo

$$A = \{x \in X \mid (\forall i \in I_K) (g_i(x_i) \leq \bar{y}_i)\} \neq \emptyset \quad \text{pois } \bar{y} \in Y$$

$$B = \{x \in X \mid \sum_{i=1}^K g_i(x_i) \leq b\}$$

já que $\bar{y} \in \Sigma$, vem:

$$A \subset B$$

logo,

21 $\sup \left\{ \sum_{i=1}^K f_i(x_i) \mid x \in B \right\} \geq \sup \left\{ \sum_{i=1}^K f_i(x_i) \mid x \in A \right\}$

Mas

$$22 \quad \text{Sup} \left\{ \sum_{i=1}^k f_i(x_i) \mid x \in A \right\} = \sum_{i=1}^k \text{Sup} \left\{ f_i(x_i) \mid x_i \in x_i \wedge g_i(x_i) \leq y_i \right\} = \\ = \sum_{i=1}^k v_i(y_i)$$

logo de (21) e (22), vem:

$$23 \quad \text{Sup} \left\{ \sum_{i=1}^k f_i(x_i) \mid x \in B \right\} \geq \sum_{i=1}^k v_i(y_i)$$

Como, pelo item A,

$$f_i(x_i^o) = v_i(y_i^o)$$

de (23), depreende-se

$$\sum_{i=1}^k f_i(x_i^o) = \sum_{i=1}^k v_i(y_i^o) > \sum_{i=1}^k v_i(y_i)$$

o que é um absurdo por (20). Logo y^o é solução ótima em (9).

Δ

Asserção D. Provemos que :

$$\left\{ \begin{array}{l} y^o \text{ é solução ótima de (9)} \\ x_i^o \text{ é solução ótima de } P_{y_i^o}^i \\ (i=1 \dots k) \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} x^o = (x_1^o, \dots, x_k^o) \\ \text{é solução ótima} \\ \text{de (9).} \end{array} \right\}$$

Demonstração. Suponhamos, por absurdo, que $x^o = (x_1^o, \dots, x_k^o)$ não seja solução ótima de (1). Ou seja,

$$(\exists x^\Delta \text{ viável em (1)}) \left(\sum_{i=1}^k f_i(x_i^\Delta) > \sum_{i=1}^k f_i(x_i^o) \right)$$

Se $y^\Delta \triangleq (g_1(x_1^\Delta), \dots, g_k(x_k^\Delta))$, segue pela viabilidade de x^Δ em (1) que:

$$y^\Delta \in \mathcal{Y}_n \quad \sum$$

e,

$$(\forall i \in I_K) (x_i^\Delta \text{ é viável em } \mathcal{P}_{y_i^\Delta})$$

Logo,

$$25 \quad \sum_{i=1}^k v_i(y_i^\Delta) > \sum_{i=1}^k f_i(x_i^\Delta)$$

Como, por hipótese, y^o é solução ótima de (6) e x_i^o ($i=1 \dots k$) solução ótima de $\mathcal{P}_{y_i^o}$ ($i=1 \dots k$), vem:

$$26 \quad \sum_{i=1}^k v_i(y_i^o) = \sum_{i=1}^k f_i(x_i^o) \geq \sum_{i=1}^k v_i(y_i^\Delta)$$

De (25) e (26)

$$\sum_{i=1}^k f_i(x_i^o) > \sum_{i=1}^k f_i(x_i^\Delta)$$

o que resulta, observando-se (24) num absurdo. Conclui-se, portanto, que x^* é solução ótima de (1).

Δ

27 Teorema. Os conjuntos Y_i são não vazios, fechados e convexos. As funções v_i são côncavas em Y_i , $i=1\dots k$, sendo $-\infty$ nos pontos não pertencentes a Y_i , $i=1\dots k$, se levarmos em conta que o supremo de uma função num conjunto vazio é $-\infty$. Finalmente, as funções v_i são não decrescentes, i.e., $\forall y^1, y^2 \in Y$ e $y^2 \geq y^1$, vem $v_i(y^2) \geq v_i(y^1)$.

A) Y_i é não vazio, $i=1\dots k$.

Demonstração. Consideremos a definição de Y_i dada na seção 2

$$Y_i \triangleq \left\{ y_i \in \mathbb{R}^m \mid (\exists x_i \in X_i) (g_i(x_i) \leq y_i) \right\}$$

Pelas hipóteses em (2), o problema (1) tem solução viável, logo,

$$(\exists \bar{x} \in X) (\forall i \in I_k) (\bar{x}_i \in X_i)$$

já que pela seção 1 $g_i(\cdot)$ é definida em X_i , seja

$$(\forall i \in I_k) (\bar{y}_i = g_i(\bar{x}_i))$$

Nessas condições,

$$(\forall i \in I_k)(\bar{y}_i \in Y_i)$$

e, portanto,

$$(\forall i \in I_k)(Y_i \neq \emptyset)$$

Δ

B) Y_k é fechado para todo $i \in I_k$.

Demonstração. Eliminemos os índices i . Seja (y^k) uma sequência qualquer em Y , convergente para um ponto y^∞ . Provemos que $y^\infty \in Y$. Por definição de Y a cada $y^k \in Y$ pode-se associar um vetor $x^k \in X$ tal que $g(x^k) \leq y^k$. Assim procedendo-se, constroi-se uma sequência (x^k) em X . Como X é compacto, tomando-se subsequências de (x^k) se necessário, pode-se considerar que

$$\lim_{k \rightarrow \infty} x^k = x^\infty \in X$$

Como g é contínua em X e $g(x^k) \leq y^k$ para todo $k \in \mathbb{N}$ vem

$$\lim_{k \rightarrow \infty} g(x^k) = g(x^\infty) \leq y^\infty$$

Logo, $y^\infty \in Y$.

Δ

Da proposição anterior concluímos que

$$(\forall i \in I_k) (Y_i \text{ é fechado})$$

Esse resultado é importante e será usado para garantir a compactade do conjunto viável do problema mestre restrito. Além disso como provaremos em III.2.18,

$$y_i \in \text{Int } Y_i \iff (\exists x_i \in X_i) (g_i(x_i) < y_i)$$

Essa propriedade e mais o fato de Y_i ser fechado, conduz a

$$y_i \in \text{Front } Y_i \iff (\forall x_i \in X_i) (\exists j \in I_m) (g_{ij}(x_i) = y_{ij})$$

c) Y_i é convexo para todo $i \in I_k$.

Demonstração. Sejam y_1 e y_2 pertencentes a Y .

$$y_1 \in Y_i \implies (\exists x' \in X_i) (g_i(x') \leq y_1)$$

$$y_2 \in Y_i \implies (\exists x^2 \in X_i) (g_i(x^2) \leq y_2)$$

Como g_i é convexa em X_i convexo, vem:

$$(\forall \theta \in [0,1]) (g_i(\theta x' + (1-\theta)x^2) \leq \theta g_i(x') + (1-\theta)g_i(x^2))$$

De (28) e (29), vem:

$$(\forall \theta \in [0,1]) (g_i(\theta x' + (1-\theta)x^2) \leq \theta y_1 + (1-\theta)y_2)$$

Como X_i é convexo,

$$x = \theta x' + (1-\theta)x'' \in X_i$$

Logo,

$$(\forall \theta \in [0,1]) (\theta y_1 + (1-\theta)y_2 \in Y_i)$$

Finalmente, temos:

$$(\forall i \in I_K) (Y_i \text{ é convexo})$$

Δ

D) As funções v_i são não decrescentes.

Demonstração. Da definição de v_i segue que se $y^2 \geq y'$, temos

$$A = \{x_i \in X_i \mid g_i(x_i) \leq y'\} \subset B = \{x_i \in X_i \mid g_i(x_i) \leq y^2\}$$

logo

$$\sup \{f_i(x_i) \mid x_i \in B\} \geq \sup \{f_i(x_i) \mid x_i \in A\}$$

ou seja

$$v_i(y^2) \geq v_i(y')$$

Δ

Logo as funções v_i são não decrescentes no sentido anteriormente exposto.

E) As funções v_i são côncavas nos conjuntos Y_i .

Demonstração. A definição de v_i anteriormente mencionada é dada abaixo:

$$v_i(y_i) = \sup \{ f_i(x_i) \text{ s.a } x_i \in X_i \wedge g_i(x_i) \leq y_i \}$$

para $y_i \in Y_i$.

Já que o valor de v_i pode ser $+\infty$ em Y_i , será adequado usar o hipografo de v_i para provarmos a concavidade da função v_i . A definição de hipografo é dada abaixo:

$$\Psi = \{ (y_i, \mu_i) \in \mathbb{R}^m \times \mathbb{R} \mid (y_i \in Y_i) \wedge \mu_i \leq v_i(y_i) \}$$

A concavidade de v_i pode, então, ser depreendida da convexidade de seu hipografo. Eliminemos os índices i . Provemos, portanto, a convexidade do hipografo de v , ou seja, dados (y^1, μ^1) e (y^2, μ^2) em Ψ e dado $\theta \in [0,1]$, provemos que

$$\theta(y^1, \mu^1) + \bar{\theta}(y^2, \mu^2) \in \Psi$$

onde $\bar{\theta} = 1-\theta$, ou, equivalentemente,

$$(\theta y^1 + \bar{\theta} y^2, \theta \mu^1 + \bar{\theta} \mu^2) \in \Psi$$

ou, finalmente, pela definição de hipografo de v ,

$$v(\theta y^1 + \bar{\theta} y^2) \geq \theta \mu^1 + \bar{\theta} \mu^2$$

Pela definição de função perturbação, vem:

$$v(\theta y^1 + \bar{\theta} y^2) = \sup \{ f(x) \text{ s.a. } x \in X \wedge g(x) \leq \theta y^1 + \bar{\theta} y^2 \}$$

Como X é convexo, decorre:

$$v(\theta y^1 + \bar{\theta} y^2) = \sup_{x^1, x^2 \in X} \{ f(\theta x^1 + \bar{\theta} x^2) \text{ s.a. } g(\theta x^1 + \bar{\theta} x^2) \leq \theta y^1 + \bar{\theta} y^2 \},$$

(usando a convexidade de g em X e a concavidade de f em X)

$$\begin{aligned} & \sup_{x^1, x^2 \in X} \{ \theta f(x^1) + \bar{\theta} f(x^2) \text{ s.a. } g(x^1) \leq y^1 \wedge g(x^2) \leq y^2 \} = \\ & = \theta \sup_{x^1 \in X} \{ f(x^1) \text{ s.a. } g(x^1) \leq y^1 \} + \bar{\theta} \sup_{x^2 \in X} \{ f(x^2) \text{ s.a. } g(x^2) \leq y^2 \} = \\ & = \theta v(y^1) + \bar{\theta} v(y^2) \end{aligned}$$

Da definição de hipografo de v , temos:

$$v(\theta y^1 + \bar{\theta} y^2) \geq \theta \mu^1 + \bar{\theta} \mu^2$$

Logo, as funções perturbação v_i associadas aos subproblemas P_i são côncavas nos conjuntos Y_i ..

Os teoremas anteriores demonstraram a equivalência entre (1) e (9).

Pode, no entanto parecer que considerarmos (9) com intuito, em parte, de obter condições de optimalidade mais convenientes. Tal fato não sucede, dado que (1) por ser um programa diferenciável, leva a condições de optimalidade expressas mais convenientemente se usarmos gradientes.

O que faremos será construir algoritmos que resolvam (9), que, como mencionamos, é computacionalmente mais interessante, podendo-se utilizar as condições de optimalidade para (1) como regra de parada. De fato, o teorema de equivalência serve para provar que as condições de optimalidade para (1) são suficientes para (9). Ou seja, suponhamos \bar{y} candidato a solução ótima em (9). Seja \bar{x}_i ($i=1\dots k$) a solução ótima de $P_{\bar{y}_i}$. Se as condições de optimalidade de Kuhn-Tucker para (1) forem satisfeitas no ponto $\bar{x} = (\bar{x}_1, \dots, \bar{x}_k)$, pela condição suficiente do Teorema de Kuhn-Tucker [23; pg 74] temos que \bar{x} é solução ótima de (1). Decorre pela asserção C do teorema de equivalência (10) que \bar{y} é solução ótima de (9).

Conseqüentemente, desde que se obtenha uma solução ótima para (9), imediatamente temos disponível uma solução ótima para (1). Isso nos garantirá, em parte, a conveniência do método utilizado para a solução de (9).

No capítulo que segue, trataremos da base teórica do método escolhido para resolução do problema mestre (9), ressaltando que pela consideração das condições de compactade dos conjuntos X_i , originalmente não impostas por Geoffrion, obtém-se a compactade da região viável do problema mestre, como veremos.

C A P I T U L O I I I

ESTRATEGIAIntrodução.

Nas secções que se seguem será exposto um método de resolução do problema II.2.9, reescrito abaixo:

$$1 \quad \text{Maximizar } \sum_{i=1}^K v_i(y_i) \text{ s.a. } \sum_{i=1}^K y_i \leq b$$

já que pelo teorema de equivalência II.3.10, teremos a solução ótima do problema original II.1.1, a partir da resolução do problema mestre (1), via subproblemas $P_{y_i^i}$ II.2.8,

$$2 \quad P_{y_i^i} \quad \text{Maximizar } f_i(x_i) \text{ s.a. } g_i(x_i) \leq y_i \quad i=1 \dots K$$

$$x_i \in X_i$$

O método a ser utilizado, classifica-se como primal de direções viáveis já que produz uma sequência de pontos viáveis em (1).

O algoritmo de direções viáveis aplicável ao problema (1), parte de um ponto y no interior de Y , determinando uma direção viável de máximo crescimento local na função objetivo de (1). Executa-se, a seguir, maximização da função objetivo de (1) na direção obtida, encontrando-

-se um novo ponto. Para que a iteração seguinte, possa ser executada, e o processo iterativo ter continuidade, o ponto de partida para a iteração seguinte deverá também estar no interior de Y ; esta imposição que constitui limitação severa do método, será satisfeita introduzindo perturação na busca unidirecional, a ser discutida na secção 3. Isto é necessário devido a dificuldade existente no tratamento de pontos de fronteira do domínio de funções côncavas. De fato, mesmo que o conjunto das direções viáveis partindo de um ponto fronteira de Y pudesse ser descrito, não conseguíramos garantir a existência da direção de máximo crescimento local. Um exemplo em que tal direção não existe, encontra-se em [8], sendo o estudo de tais casos ainda um problema em aberto.

Portanto, $y \in \text{Int}Y$ é uma garantia de que todas direções a partir de y são viáveis em Y . Outras implicações advém dessa suposição sendo discutidas em (17) e (30).

A secção 1 define o método de direções viáveis expondo resumidamente os aspectos mais importantes, além de discutir a existência da solução ótima para o problema original.

Teoremas e resultados fundamentais são apresentados na secção 2, ressaltando-se as implicações de cada hipótese assumida, sobre as quais se estrutura a secção seguinte; aí se caracteriza a obtenção de uma boa direção viável, culminando-se com a apresentação do teorema que leva à determinação dessa direção. Ainda nessa secção, discute-se a otimização na direção obtida (otimização unidirecional ou problema mestre restrito), expondo-se um procedi

mento de eliminação dos pontos fronteira de Y, caso o ótimo do problema mestre restrito pertença a ela. Como produto da otimização unidirecional são analisados os subproblemas envolvidos, bem como as hipóteses que garantem as soluções ótimas dos mesmos.

Por fim, na secção 4 esquematiza-se o algorítmo obtido do método das direções viáveis.

SECÇÃO 1 - MÉTODO DAS DIREÇÕES VIÁVEIS. EXISTÊNCIA DAS SOLUÇÕES ÓTIMAS.

Algorítmos de direções viáveis para resolução de problemas não lineares de otimização, foram originalmente desenvolvidos por Zoutendijk [38] em 1959, e independentemente por Zukhovitskii [40] em 1962 - 1963, tendo sido precedidos pelo método dos centros e o algorítmo de Topkis e Veinott, nos quais se apresentam idealizados.

Seja o problema de otimização abaixo:

$$3 \quad \text{Maximizar } \underset{\mathbf{x}}{f^0}(\mathbf{x}) \text{ s.t. } f^i(\mathbf{x}) \leq 0 \quad i=1 \dots m$$

onde f^i ($i=0 \dots m$) são funções de \mathbb{R}^n em \mathbb{R} . Seja C a região viável, i.e.,

$$4 \quad C = \{ \mathbf{x} \in \mathbb{R}^n \mid f^i(\mathbf{x}) \leq 0, \quad i=1 \dots m \}$$

Um algorítmo que resolva (3) é dito de direções viáveis, se dado um ponto $x_i \in C$, ele determina uma semi-reta, $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{x}_i + \mu \mathbf{h}_i, \mu \geq 0 \}$, passando pelo interior (relativo) de C, e nessa semi-reta um ponto $x_{i+1} \in C$ tal

que $f^*(x_{i+1}) > f^*(x_i)$. Tal método pressupõe que $\text{Int}C \neq \emptyset$.

Algorítmos de direções viáveis surgiram da vantagem existente em se considerar apenas o comportamento local das funções critério e vínculos, e, de não se envolverem todas as funções f^i ($i=1\dots m$), em cada iteração quando se calcula o vetor $h_i \in \mathbb{R}^n$, que define a semi-resta para cada $x_i \in C$. Contudo, a eliminação das funções f^i ($i=1\dots m$) tais que $f^i(x_i) < 0$ (vínculos inativos), levava a algorítmos não convergentes, mesmo em casos totalmente diferenciáveis. Nesses casos, é possível construir exemplos nos quais fenômenos de ziguezague levam a pontos que não obedecem as condições de optimalidade de Kuhn-Tucker, embora os problemas satisfaçam a condição de qualificação de Slater.

A introdução do conceito de vínculos $f^i \varepsilon$ -ativos, i.e., $f^i(i=1\dots m)$ é dito ε -ativo em x_i , se

$$5 \quad f^i(x_i) + \varepsilon \geq 0,$$

levou Zoutendijk a algorítmos convergentes. Os algorítmos de direções viáveis aplicáveis a (3) pressupõem, para se manter as condições de convergência, certas condições. Entre tais condições, destaca-se a diferenciabilidade contínua das funções f^i ($i=0\dots m$). Tomadas, portanto, as devidas precauções, tais algorítmos podem ser empregados na resolução de II.1.1, o problema de alocação de recursos. No entanto, como já tivemos oportunidade de observar, o trabalho computacional pode crescer assustadoramente, além disso não tiraríamos proveito da estrutura particular do problema, fator fundamental para sua resolução. Por outro lado, tais algorítmos não podem ser aplicados

aos problemas (1), já que as funções v_i ($i=1\dots k$) não são necessariamente diferenciáveis em todos os pontos de Y_i ($i=1\dots k$).

A busca da direção viável que utilizaremos, aplicável a problemas em que a função objetivo é não diferenciável, deve-se a Geoffrion [11], fundamentando-se nas propriedades de concavidade das funções v_i e na relação entre vetores multiplicadores ótimos e supergradientes (serão definidos na secção 2). É por isso que voltamos a mencionar que as hipóteses de concavidade-convexidade, tem aqui papel fundamentalmente importante, sem as quais, com as ferramentas de que se dispõe atualmente, o desenvolvimento desse trabalho não poderia seguir tal linha.

O programa linear proposto por Geoffrion para busca da direção viável pressupõe perda total da diferenciabilidade das funções v_i , baseado na idéia de que o número de pontos onde a função é diferenciável, tende a diminuir na busca do ótimo.

Existência das soluções ótimas.

Com as hipóteses de compacidade dos conjuntos X_i ($i=1\dots k$) teremos a compacidade da região viável do problema de alocação de recursos II.1.1. Com as hipóteses de diferenciabilidade sobre as f_i ($i=1\dots k$) em X_i ($i=1\dots k$), que assumiremos na secção 2, temos a continuidade de f_i na região viável compacta, e, portanto, a garantia da existência da solução ótima de II.1.1. Pelo teorema II.3.10, está conseqüentemente, garantida a solução ótima do problema mestre (1) e dos subproblemas envolvidos.

A secção seguinte tratará dos teoremas e resul

tados que garantem a existência da derivada direcional das funções perturbação v_i ($i=1\dots k$), bem como dos vetores multiplicadores ótimos associados aos subproblemas $P_{\bar{y}_i}^i$ (2).

Tais fatos são fundamentais já que se proporá obter na secção 3 uma direção viável de incremento local para função objetivo de (1) com base em sua derivada direcional.

SECÇÃO 2 - TEOREMAS E RESULTADOS FUNDAMENTAIS

O primeiro teorema importante que segue, relaciona o vetor multiplicador ótimo dos subproblemas $P_{\bar{y}_i}^i$ ($\bar{y}_i \in \mathbb{R}^m$ para $i=1\dots k$) associados com as restrições $g_i(x_i) \leq \bar{y}_i$ ($i=1\dots k$), com o supergradiente da função perturbação v_i , definida no apêndice A, ($i=1\dots k$), no ponto \bar{y}_i ($i=1\dots k$).

6 Teorema: [12; pg 22]. Suponhamos que $P_{\bar{y}_i}^i$ (2) tenha solução ótima \bar{x}_i . Então $\bar{\lambda}_i \in \mathbb{R}^m$ é vetor multiplicador ótimo para $P_{\bar{y}_i}^i$ associado com as $g_i(x_i) \leq \bar{y}_i$, se, e somente se, $\bar{\lambda}_i$ é um supergradiente de v_i (.) para $y_i = \bar{y}_i$, ou seja, o par $(\bar{x}_i; \bar{\lambda}_i)$ satisfaz as condições de Kuhn-Tucker,

7

- (i) \bar{x}_i maximiza $f_i(x_i) - \bar{\lambda}_i^t [g_i(x_i) - \bar{y}_i]$ sobre x_i ;
- (ii) $\bar{\lambda}_i^t [g_i(\bar{x}_i) - \bar{y}_i] = 0$;
- (iii) $\bar{\lambda}_i \geq 0$;
- (iv) $g_i(\bar{x}_i) \leq \bar{y}_i$.

se, e somente se, $v_i(y_i) \leq v_i(\bar{y}_i) + \bar{\lambda}_i^t (y_i - \bar{y}_i)$ para todo $y_i \in \mathbb{R}^m$

Observação: a condição (iv) de (7) é dispensável pois por hipótese \bar{x}_i é ótimo para $P_{\bar{y}_i}^L$.

Demonstração. \Rightarrow : eliminemos os índices i e suponhamos que $\bar{\lambda}$ seja vetor multiplicador ótimo para $P_{\bar{y}}$ associado com $g(x) \leq \bar{y}$. Logo $(\bar{x}, \bar{\lambda})$ satisfaz as condições de optimidade de Kuhn-Tucker supra mencionadas. Da condição (i) temos:

$$8 \quad f(\bar{x}) - \bar{\lambda}^T [g(\bar{x}) - \bar{y}] \geq f(x) - \bar{\lambda}^T [g(x) - \bar{y}] \text{ para todo } x \in X$$

De (8) e (ii), vem:

$$f(\bar{x}) \geq f(x) - \bar{\lambda}^T [g(x) - \bar{y}] \text{ para todo } x \in X.$$

usando a condição (iii), para cada $y \in Y$, temos:

$$(\forall x \in X)(g(x) \leq y)(f(\bar{x}) + \bar{\lambda}^T (y - \bar{y}) \geq f(x))$$

obtendo o supremo do 2º membro para os valores de x indicados, teremos:

$$f(\bar{x}) + \bar{\lambda}^T (y - \bar{y}) \geq \sup_x \{f(x) \text{ s.t. } x \in X \wedge g(x) \leq y\}$$

logo,

$$f(\bar{x}) + \bar{\lambda}^T (y - \bar{y}) \geq v(y) \text{ para todo } y \in Y$$

já que \bar{x} é solução ótima de $P_{\bar{y}}$, temos $f(\bar{x})=v(\bar{y})$, e,
 $v(\bar{y}) + \bar{\lambda}^t(\bar{y} - \bar{y}) \geq v(y)$ para todo $y \in Y$.

Se $y \notin Y$ temos $v(y) = -\infty$, e, portanto,

$$v(y) \leq v(\bar{y}) + \bar{\lambda}^t(\bar{y} - \bar{y}) \quad \text{para todo } y \in \mathbb{R}^m,$$

o que completa a prova da condição suficiente.

\Leftarrow : supondo-se ainda a remoção dos índices i , seja $\bar{\lambda}$ um supergradiente de v em \bar{y} . Provemos que $(\bar{x}, \bar{\lambda})$ satisfaz as condições de Kuhn-Tucker (i)-(iv). Como \bar{x} é solução ótima de $P_{\bar{y}}$, vem:

$$g(\bar{x}) \leq \bar{y}$$

justificando (iv).

Sabemos que:

$$9 \quad v(y) \leq v(\bar{y}) + \bar{\lambda}^t(y - \bar{y}) \quad \text{para todo } y.$$

Façamos em (9) $y = \bar{y} + e_j$, onde $e_j \in \mathbb{R}^m$ tendo a j -ésima componente unitária e as demais nulas.

Temos,

$$v(\bar{y} + e_j) \leq v(\bar{y}) + \bar{\lambda}^t(\bar{y} + e_j - \bar{y}) = v(\bar{y}) + \bar{\lambda}_j$$

logo,

$$v(\bar{y} + e_j) - v(\bar{y}) \leq \bar{\lambda}_j$$

como $v(\cdot)$ é não decrescente, vem:

$\lambda_j > 0$

Fazendo j variar de 1 a m teremos provada a condição (iii).

Fazendo em (9) $y = g(\bar{x})$ vem:

$$10 \quad 0 = v(g(\bar{x})) - v(\bar{y}) \leq \bar{\lambda}^t [g(\bar{x}) - \bar{y}]$$

onde a igualdade ocorre, pois decrescendo \bar{y} para $g(\bar{x})$ em $P\bar{y}$ não haverá alteração na optimilidade de \bar{x} ; pela condição (iii) e viabilidade de \bar{x} em $P\bar{y}$, temos, $\bar{\lambda}^t [g(\bar{x}) - \bar{y}] \leq 0$. Essa última desigualdade e (10) levam à condição (ii). Finalmente, estabeleceremos (i) fazendo-se $y = g(x)$, em (9), para qualquer $x \in X$. Logo:

$$v(g(x)) \leq v(\bar{y}) + \bar{\lambda}^t [g(x) - \bar{y}] \quad \text{para todo } x \in X.$$

Mas $(\forall x \in X)(v(g(x)) \geq f(x))$

pois

$$x \in \{y / g(y) \leq g(x)\}$$

Este fato e a relação $v(\bar{y}) = f(\bar{x})$, levam a

$$f(\bar{x}) \geq f(x) - \bar{\lambda}^T [g(x) - \bar{y}] \quad \text{para todo } x \in X,$$

o que, observando-se (ii), leva a (i). Δ

Observação: o resultado que normalmente se obtém a partir do teorema anterior, é a equivalência entre vetor multiplicador ótimo e o negativo do supergradiente. A variação obtida pelo teorema anterior, decorre da mudança feita na formalização do enunciado do problema II.1.1, optando-se por $g(x) \leq \bar{y}$, e não $g(x) \geq \bar{y}$ como normalmente se encontra nos problema de maximização II.1.1. Optou-se pelas funções g_x convexas, dada a facilidade em interpretar o vetor y como recursos, sem qualquer complicaçāo maior.

Para esclarecer melhor este ponto, mostramos a seguir as variações obtidas na função pertubação decorrentes dessa variação no problema de optimização. Seja:

$$11 \quad \underset{x}{\text{Maximizar}} \quad f(x) \text{ s. a } g(x) \geq \underline{c}$$

$$12 \quad \underset{x}{\text{Maximizar}} \quad f(x) \text{ s. a } g(x) \geq y$$

$$13 \quad v_y(y) = \text{Sup} \left\{ f(x) \text{ s. a } x \in X \wedge g(x) \geq y \right\}$$

$$14 \quad \underset{x}{\text{Maximizar}} \quad f(x) \text{ s. a } -g(x) \leq \underline{c}$$

$$15 \quad \underset{x}{\text{Maximizar}} \quad f(x) \text{ s. a } -g(x) \leq y$$

16

$$v_2(y) = \sup \{ f(x) \text{ sujeito a } x \in X \wedge g(x) \leq y \}$$

Os problemas (11) e (14) coincidem. Vejamos qual a relação entre as funções perturbação (13) e (16), associadas aos problemas (12) e (15) pertubados.

- a) v_1 e v_2 , supostas as condições de concavidade-convexidade, são côncavas;
- b) v_1 é não crescente, e v_2 não decrescente, como provamos em II.3.27;
- c) $v_1(y) = v_2(-y)$ para todo y (imediato das definições de v_1 e v_2);
- d) nos pontos y e nas direções z em que as derivadas direcionais de v_1 e v_2 existam, temos:

$$v_1'(y; z) = v_2'(-y; -z)$$

Esta propriedade decorre da definição de derivada direcional e da propriedade (c), como abaixo verificamos:

$$\frac{v_1'(y; z)}{\theta \rightarrow +0} = \frac{v_1(y + \theta z) - v_1(y)}{\theta} = \frac{v_2(-y - \theta z) - v_2(-y)}{\theta} =$$

$$\frac{v_2(-y) + \theta(-z) - v_2(-y)}{\theta} = v_2'(-y; -z)$$

Hipóteses para garantia da existência dos vetores multiplicadores ótimos.

O teorema (6) somente relaciona as quantidades envolvidas, há, no entanto, de garantir a existência das mesmas. Abaixo apresentamos hipóteses suficientes para garantir a existência dos vetores multiplicadores ótimos associados com as restrições $g_i(x_i) \leq \bar{y}_i$ e $h_i(x_i) \geq 0$.

17 Hipótese 1 : $\bar{y}_i \in \text{Int}Y_i$ ($i = 1 \dots k$), onde Y_i , lembramos, é definido em II.2.7., i.e., $\text{Int}Y_i \neq \emptyset$.

A necessidade dessa condição fora já discutida na introdução, evitando lidar com pontos da fronteira dos conjuntos Y_i .

Essa mesma hipótese atua como fundamental na garantia dos vetores multiplicadores ótimos associados com as restrições $g_i(x_i) \leq \bar{y}_i$ dos subproblemas P_{Y_i} , o que fica claro do significado a seguir.

Seja 1 vetor do \mathbb{R}^m cujas componentes são a unidade. Temos:

$$\bar{y}_i \in \text{Int}Y_i \Rightarrow (\exists \theta > 0)(\bar{y}_i - \theta \mathbf{1} \in Y_i)$$

logo,

$$(\exists \hat{x}_i \in X_i)(g_i(\hat{x}_i) \leq \bar{y}_i - \theta \mathbf{1}) \Rightarrow (\exists \hat{x}_i \in X_i)(g_i(\hat{x}_i) < \bar{y}_i)$$

Concluímos, assim, que a hipótese 1 (17) resulta em: os vínculos g_i satisfazem a condição de qualificação de Slater sobre X_i ($i = 1 \dots k$).

No entanto, o maior interesse está em usar a condição de qualificação de Slater para garantir $y_i \in \text{Int} Y_i$. Ou seja, se

18

$$(\forall i \in I_k) (\exists \bar{x}_i \in X_i) (g_i(\bar{x}_i) < y_i)$$

então,

$$(\forall i \in I_k) (y_i \in \text{Int } Y_i)$$

sendo essa última proposição equivalente a $y \in \text{Int } Y$.

Para provarmos a implicação acima seja, para $i \in I_k$,

$$A_i \triangleq \left\{ y_i \in \mathbb{R}^m \mid y_i^1 \leq y_i \leq y_i^2 \right\}$$

onde

$$y_i^1 = g_i(\bar{x}_i)$$

$$y_i^2 > y_i$$

O conjunto A_i possui as propriedades abaixo:

A_i é não vazio ($y_i^1 \in A_i$), além de convexo e fechado;

19 $A_i \subseteq Y_i$ (imediato);

20 $\bar{y}_i \in \text{Int } A_i$ pois $y_i^1 < \bar{y}_i$ (por hipóteses) e $\bar{y}_i < y_i^2$ (por construção). De (19) e (20) vem :

$$\bar{y}_i \in \text{Int } Y_i$$

Queremos, finalmente, antecipar que, dada a concavidade de v_i em Y_i , a hipótese $\bar{y}_i \in \text{Int } Y_i$ é suficiente para garantir a existência das derivadas direcionais da função v_i em \bar{y}_i , como mostraremos em (30).

21 Hipóteses 2 : suponhamos que os conjuntos X_i sejam caracterizados pelas funções $h_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}^{m_i}$ concavas, i.e.,

$$X_i = \{x_i \in \mathbb{R}^{n_i} / h_i(x_i) \geq \underline{o}\}$$

Os conjuntos X_i são tais que

$$(\exists x_i \in \mathbb{R}^{n_i})(h_i(x_i) > \underline{o})$$

Segue, como em (17), que as funções h_i satisfazem a condição de qualificação de Slater sobre \mathbb{R}^{n_i} .

22 Hipóteses 3 : as funções f_i e g_i são supostas definidas no \mathbb{R}^{n_i} e diferenciáveis em X_i . As funções h_i supostas diferenciáveis em X_i ($i=1\dots k$).

Essas três hipóteses e mais as hipóteses de

concavidade-convexidade (hipóteses adicionais II.1.2) garantirão a existência dos vetores multiplicadores ótimos como veremos a seguir.

23 Existência dos vetores multiplicadores associados com $g_i(x_i) \leq \bar{y}_i$.

Seja $\bar{x}_i \in \mathbb{R}^n$ solução ótima do problema $\bar{P}y_i^i$,

24 Maximizar $f_i(x_i)$ s.a $g_i(x_i) \leq \bar{y}_i$

A existência da solução ótima é garantida pela compactade dos conjuntos X_i (hipóteses II.1.2) e pela continuidade das funções f_i (hipóteses 3). Pela condição necessária do teorema de optimidade do ponto de sela de Kuhn-Tucker [23; pg 79], já que X_i é convexo, f_i côncava em X_i , g_i convexa em X_i e g_i satisfaaz a condição de qualificação de Slater sobre X_i (hipótese 1), vem:

existe $\bar{\lambda}_i \in \mathbb{R}^m$ tal que,

25 (i) \bar{x}_i maximiza $f_i(x_i) - \bar{\lambda}_i^t [g_i(x_i) - \bar{y}_i]$
sobre X_i ;

(ii) $\bar{\lambda}_i^t [g_i(\bar{x}_i) - \bar{y}_i] = 0$;

(iii) $\bar{\lambda}_i \geq 0$;

(iv) $g_i(\bar{x}_i) \leq \bar{y}_i$.

Logo, $\bar{\lambda}_i$ é vetor multiplicador ótimo associado com as restrições $g_i(x_i) \leq \bar{y}_i$.

26 Existência dos vetores multiplicadores ótimos associados com $h_i(x_i) \geq 0$.

Da condição (i) em (25) temos que \bar{x}_i resolve o problema:

27 Maximizar $\sum_i f_i(x_i) - \bar{\lambda}_i^t [g_i(x_i) - \bar{y}_i]$ s. a $h_i(x_i) \geq 0$

onde $\bar{\lambda}_i$ satisfaz as condições (ii) e (iii) em (25).

Da hipótese 3, $f_i(\cdot) + \bar{\lambda}_i^t [g_i(\cdot) - \bar{y}_i]$ e $h_i(\cdot)$ são definidas em \mathbb{R}^{m_i} , $h_i(\cdot)$ e $f_i(\cdot) + \bar{\lambda}_i^t [g_i(\cdot) - \bar{y}_i]$ são diferenciáveis em $\bar{x}_i \in X_i$, e h_i satisfaz a condição de Slater sobre \mathbb{R}^{m_i} (hipótese 2) logo pela condição necessária do teorema de Kuhn-Tucker, temos:

28 Existe $\bar{\mu}_i \in \mathbb{R}^{m_i}$ tal que,

$$(i) \nabla f_i^t(\bar{x}_i) - \bar{\lambda}_i^t \nabla g_i(\bar{x}_i) + \bar{\mu}_i^t \nabla h_i(\bar{x}_i) = 0;$$

$$(ii) \bar{\mu}_i^t [h_i(\bar{x}_i)] = 0;$$

$$(iii) \bar{\mu}_i \geq 0;$$

$$(iv) h_i(\bar{x}_i) \geq 0$$

As condições em (28) garantem, portanto, a existência dos vetores multiplicadores ótimos associados com as restrições $h_i(x_i) \geq 0$.

De (23) e (26) concluimos que os conjuntos dos vetores multiplicadores associados às restrições $g_i(x_i) \leq \bar{y}_i$ e $h_i(x_i) \geq \underline{y}_i$ de \bar{P}_i , ficam, consequentemente, caracterizados por:

29 (i) $\nabla f_i^t(\bar{x}_i) - \bar{\lambda}_i^t \nabla g_i(\bar{x}_i) + \bar{\mu}_i^t \nabla h_i(\bar{x}_i) = \underline{\omega}^t$;

(ii) $\bar{\mu}_i^t [h_i(\bar{x}_i)] = 0 \wedge \bar{\lambda}_i^t [g_i(\bar{x}_i) - \bar{y}_i] = 0$;

(iii) $\bar{\mu}_i \geq \underline{\omega} \wedge \bar{\lambda}_i \geq \underline{\omega}$;

(iv) $h_i(\bar{x}_i) \geq \underline{y}_i \wedge g_i(\bar{x}_i) \leq \bar{y}_i$.

A seguir expoem-se alguns resultados a respeito das derivadas direcionais das funções perturbação v_i .

30 Derivadas direcionais das funções perturbação

Funções côncavas possuem derivadas direcionais em todas as direções e em todos os pontos interiores ao seu domínio onde são finitas. Em tal caso, a obtenção das derivadas direcionais reduz-se a um problema de minimização. Supondo-se $v_i(\bar{y}_i)$ finita, com $\bar{y}_i \in \text{Int}Y_i$, se $v_i'(\bar{y}_i; \bar{z}_i)$ for a derivada direcional da função perturbação no ponto \bar{y}_i e na direção $\bar{z}_i \in \mathbb{R}^m$, temos: (ver apêndice A) :

31 $v_i'(\bar{y}_i; \bar{z}_i) = \min \left\{ \bar{p}_i^t \bar{z}_i / \bar{p}_i \mid \bar{p}_i^t \bar{z}_i \text{ é supergradiente de } v_i \text{ em } \bar{y}_i \right\}$

Nesse problema (31) o mínimo sempre existe, pois $\bar{p}_i^t \bar{z}_i$ é contínua (pois é linear) e o conjunto dos supergradientes de $v_i(\cdot)$ em \bar{y}_i é não vazio e compacto (além

de convexo). Já que o mínimo ocorre para um $\bar{p}_i \in \mathbb{R}^m$, vem:

$$32 \quad v_i'(\bar{y}_i; \bar{z}_i) = \bar{p}_i^t \bar{z}_i$$

Depreende-se de (32) que a direção que maximiza a derivada direcional num ponto $\bar{y}_i \in \mathbb{R}^m$ é a direção dada por um supergradiente da função perturbação em \bar{y}_i , o que o torna uma generalização do gradiente. Resulta daí que se perturbarmos as restrições $g_i(x_i) \leq \bar{y}_i$ do problema $P_{\bar{y}_i}$ na direção desse supergradiente, obteremos a maior variação local no valor ótimo do problema $P_{\bar{y}_i}$. Essa idéia é interessante e valoriza sobremodo o conceito de supergradiente. Ainda esse fato torna-se mais evidente se a função v_i for diferenciável em \bar{y}_i , o que ocorre se, e somente se, tiver um único supergradiente nesse ponto; nesse caso, para qualquer direção $\bar{z}_i \in \mathbb{R}^m (\bar{y}_i \in \text{Int}Y_i)$, vem,

$$v_i'(\bar{y}_i; \bar{z}_i) = \bar{p}_i^t \bar{z}_i$$

onde \bar{p}_i representa o único supergradiente, chamado, nesse caso particular, gradiente e indicado por $\nabla v_i(\bar{y}_i)$. Parece agora bem mais evidente que a direção que maximiza $v_i'(\bar{y}_i; \cdot)$ é a direção de $\bar{p}_i = \nabla v_i(\bar{y}_i)$. No capítulo IV exploraremos esse ponto de vista, bem como depreenderemos quando a diferenciabilidade ocorre.

O teorema que segue mostra como se obterá a derivada direcional da função perturbação v_i no ponto $\bar{y}_i \in \text{Int}Y_i$ e em qualquer direção, bem como as hipóteses que garantem a sua existência e transformação com base nos vetores multiplicadores ótimos associados com as restrições de $P_{\bar{y}_i}$.

Esses resultados serão utilizados na seção seguinte para obtenção das derivadas direcionais constantes da formulação do critério (40).

33 Teorema: [11; pg 5]. Seja \bar{x}_i solução ótima de $\bar{P}_{\bar{y}_i}$, sendo válidas as hipóteses 1, 2 e 3. Então para todo $\bar{z}_i \in \mathbb{R}^m$ $v_i^*(\bar{y}_i; \bar{z}_i)$ é igual ao valor ótimo do programa linear:

34 Minimizar $\lambda_i^t \bar{y}_i - \alpha$
 λ_i, μ_i

$$\lambda_i \geq 0 \wedge \mu_i \geq 0$$

$$\nabla f_i^t(\bar{x}_i) - \lambda_i^t \nabla g_i(\bar{x}_i) + \mu_i^t \nabla h_i(\bar{x}_i) = 0$$

$$\lambda_i^t [g_i(\bar{x}_i) - \bar{y}_i] = 0$$

$$\mu_i^t h_i(\bar{x}_i) = 0$$

Demonstração. Da hipótese 1, da existência da solução ótima para $\bar{P}_{\bar{y}_i}$, usando as observações (30) sobre funções côncavas, para uma direção $\bar{z}_i \in \mathbb{R}^m$ dada, temos:

$$v_i^*(\bar{y}_i; \bar{z}_i) = \min \left\{ \bar{y}_i^t \bar{z}_i / p_i \text{ é supergradiente de } v_i \text{ em } \bar{y}_i \right\}$$

Por outro lado, como são válidas as hipóteses 1.2 e 3 e mais as hipóteses de concavidade-convexidade, aplicando-se o teorema 4, vem:

$v_i^*(\bar{y}_i; \bar{z}_i) = \min \left\{ \lambda_i^t z_i \mid \lambda_i \text{ é um vetor multiplicador ótimo de } P_{\bar{y}_i} \text{ associado com } g_i(x_i) \leq \bar{y}_i \right\},$

ou ainda usando (29), $v_i^*(\bar{y}_i; z_i)$ é o valor ótimo do programa (34), o que completa a demonstração. Δ

Escrevendo (34) na forma matricial de programa linear ($\min\{x^t c \mid x \geq 0 \wedge x^t A = b^t\}$), temos:

$$35 \quad \text{Minimizar}_{\lambda_i, \mu_i} \begin{bmatrix} \lambda_i^t \\ \mu_i^t \end{bmatrix} \begin{bmatrix} \bar{z}_i \\ 0 \end{bmatrix} \quad \text{s.a.} \quad \lambda_i \geq 0 \wedge \mu_i \geq 0$$

$$\begin{bmatrix} \lambda_i^t & \mu_i^t \\ \hline \nabla g_i(\bar{x}_i) & g_i(\bar{x}_i) - \bar{y}_i \\ -\nabla h_i(\bar{x}_i) & 0 \end{bmatrix} \begin{bmatrix} \bar{z}_i \\ 0 \end{bmatrix} = \begin{bmatrix} \nabla f_i^t(\bar{x}_i) \\ 0 \\ h_i(\bar{x}_i) \end{bmatrix}$$

Antes de eliminarmos algumas restrições do programa linear (35), lembramos serem conhecidas as soluções ótimas \bar{x}_i dos subproblemas $P_{\bar{y}_i}$, antes de se resolver o programa (35). Nessas condições, pode-se determinar quais vínculos g_{ij} ($i = \dots k$) e h_{ij} ($i = 1 \dots k$) são inativos em \bar{x}_i ($i = 1 \dots k$), o que implica na exclusão das condições de complementaridade (29-ii) de (35), já que:

$$36 \quad g_{ij}(\bar{x}_i) - y_{ij} < 0 \implies \lambda_{ij} = 0$$

$$37 \quad h_{ij}(\bar{x}_i) > 0 \implies \mu_{ij} = 0$$

e, com elas eliminam-se os gradientes ∇g_{ij} e ∇h_{ij} dos vínculos g_{ij} e h_{ij} tais que (36) e (37), mantendo-se em $\nabla g_{ij}(\bar{x}_i)$

e $\nabla h_i(\bar{x}_i)$ somente os gradientes dos vínculos g_{ij} e h_{ij} ativos em \bar{x}_i ; ainda das condições (36) e (37) manteremos em λ_i e μ_i somente as componentes não necessariamente nulas; e, finalmente, em z_i somente as componentes correspondentes aos vínculos g_{ij} ativos em \bar{x}_i , i.e., tais que $g_{ij}(\bar{x}_i) = \bar{y}_{ij}$. As componentes z_{ij} retiradas do programa (35) receberão tratamento especial apresentando em IV.2. Um tratamento simplificado dessas componentes será visto adiante, nessa secção. Assim sendo, (35) é escrito na forma abaixo:

38

$$\text{Minimizar}_{\lambda_i, \mu_i} \begin{bmatrix} \lambda_i^t \\ \mu_i^t \end{bmatrix} \begin{bmatrix} \bar{y}_i \\ 0 \end{bmatrix} \text{ s.a. } \lambda_i \geq 0 \wedge \mu_i \geq 0$$

$$\begin{bmatrix} \lambda_i^t \\ \mu_i^t \end{bmatrix} \begin{bmatrix} \nabla g_{ij}(\bar{x}_i) \\ -\nabla h_i(\bar{x}_i) \end{bmatrix} = \nabla f_i^t(\bar{x}_i)$$

O valor ótimo de (38) nos ministrará a derivada direcional da função $v_{\bar{x}}$ no ponto $\bar{y}_i \in \text{Int}Y$ e na direção $z_i \in \mathbb{R}^m$, originalmente proposta.

Como podemos perceber, esse programa (38) não se apresenta numa forma adequada, pois a direção z_i aparece como coeficiente de custo, parametrizando (38).

Se nosso intento for maximizar a derivada direcional $v_{\bar{x}}'(\bar{y}_i; \cdot)$ teremos, praticamente, uma infinidade de direções a testar, o que, obviamente, é inadequado. Por essa razão, ao problema (38) se aplicará a técnica de manipulação chamada dualização, resultando um problema em que a direção aparece, convenientemente, como vetor variável.

Dualização do programa linear.

A dualização do programa linear (38), levará

à possibilidade de se considerar a direção z como vetor variável. Por outro lado, essa técnica de manipulação é plenamente justificável, já que pelo teorema de Dualidade em Programação Linear, o problema (38) tem solução ótima se, e somente se, o seu dual também tiver, sendo que nesse caso os valores ótimos coincidem. Esse resultado é altamente desejável o que torna satisfatória a manipulação, pois por hipótese $\bar{y}_i \in \text{Int}Y_i$ e da concavidade de v_i temos garantida a existência de $v_i^t(\bar{y}_i; z_i)$ para todo z_i , como destacamos em (30).

Seja $w_i \in \mathbb{R}^{n_i}$ vetor de variáveis duais. Dualizando (38), vem :

$$\begin{array}{l} \text{Maximizar}_{\omega_i, z_i} \nabla f_i^t(\bar{x}_i) \omega_i \quad \text{s.a} \\ \begin{bmatrix} \nabla g_i(\bar{x}_i) \\ -\nabla h_i(\bar{x}_i) \end{bmatrix} \omega_i \leq \begin{bmatrix} \bar{y}_i \\ 0 \end{bmatrix} \end{array}$$

ou, então,

$$39 \quad \text{Maximizar}_{\omega_i, z_i} \nabla f_i^t(\bar{x}_i) \omega_i \quad \text{s.a}$$

$$\nabla g_{ij}^t(\bar{x}_i) \omega_i - z_{ij} \leq 0 \quad \text{para } j \text{ tais que } g_{ij}(\bar{x}_i) = \bar{y}_i$$

$$\nabla h_{ij}^t(\bar{x}_i) \omega_i \leq 0 \quad \text{para } j \text{ tais que } h_{ij}(\bar{x}_i) = 0.$$

Portanto, encarando z_i como vetor variável, facilmente depreendemos que a solução ótima do programa

(39), conduz à obtenção da direção que maximiza $v'_i(\bar{y}_i; \cdot)$, e o valor do referido programa o valor máximo da derivada direcional.

Baseado nesse importante resultado, partimos para secção 3, que cumpre com o objetivo do presente capítulo: dado $y \in \mathbb{R}^k$ obter uma boa direção viável para o problema (1).

SECÇÃO 3 - ESCOLHA DE UMA BOA DIREÇÃO VIÁVEL PARA O PROBLEMA MESTRE. OTIMIZAÇÃO UNIDIRECIONAL.

Caracterizemos o que se entende por boa direção viável para (1). Seja $\bar{y} = (\bar{y}_1, \dots, \bar{y}_k) \in \mathbb{R}^k$. Uma direção viável de crescimento para (1) em \bar{y} é um vetor $\bar{z} = (\bar{z}_1, \dots, \bar{z}_k) \in \mathbb{R}^k$ tal que:

(i) viabilidade :

$(\exists \theta > 0)(\bar{y} + \theta \bar{z} \text{ é viável em (1)})$, i. e.,

$(\exists \theta > 0)(\sum_{i=1}^k (\bar{y}_i + \theta \bar{z}_i) \leq b)$ sendo θ real

(ii) crescimento :

$$(\exists \lambda > 0)(\forall \theta \in [0, \lambda]) \left(\sum_{i=1}^k v_i'(\bar{y}_i + \theta \bar{z}_i) \geq \sum_{i=1}^k v_i'(\bar{y}_i) \right)$$

onde λ e θ são reais.

O critério utilizado para obtenção de uma boa direção viável de acréscimo, resume-se na direção viável que maximiza a derivada direcional da função objetivo de (1), $\sum_{i=1}^k v_i'(\bar{y}_i; \cdot)$, em \bar{y}_i . Ou seja, uma direção

$z = (z_1, \dots, z_k) \in \mathbb{R}^{K^m}$, tal que resolva o problema abaixo:

40 Maximizar $\sum_{i=1}^K v_i'(\bar{y}_i; z_i)$ s.a

41 $\sum_{i=1}^K z_i \leq 0$ para j tal que $\sum_{i=1}^K \bar{y}_{ij} = b_j$

42 $|z_i| \leq 1$ para $i = 1 \dots K$ e $j = 1 \dots m$

A condição (41) garante a viabilidade da direção z , já que se $\sum_{i=1}^K \bar{y}_{ij} = b_j$ para algum j , existirá $\theta \in \mathbb{R}$ tal que,

$$\sum_{i=1}^K (\bar{y}_{ij} + \theta z_i) = \sum_{i=1}^K \bar{y}_{ij} + \theta \sum_{i=1}^K z_i \leq \sum_{i=1}^K \bar{y}_{ij} = b_j.$$

Se $\sum_{i=1}^K \bar{y}_{ij} < b_j$ para algum j , sempre existirá θ tal que $\sum_{i=1}^K (\bar{y}_{ij} + \theta z_i) \leq b_j$. Está, portanto, garantida por (41) a viabilidade da direção $z \in \mathbb{R}^{K^m}$.

A condição de crescimento, nesse caso máximo crescimento local, fica evidente da própria filosofia do programa (40).

A condição (42) leva em conta a homogeneidade de $v_i'(\bar{y}_i; z_i)$, i.e., se

$$v_i'(\bar{y}_i; z_i) = \lim_{\theta \rightarrow +0} \frac{v_i(\bar{y}_i + \theta z_i) - v_i(\bar{y}_i)}{\theta}$$

então, segue imediatamente que

43 $(\forall \varepsilon > 0) [v_i'(\bar{y}_i; \bar{z}_i) = \varepsilon v_i'(\bar{y}_i; z_i)]$,

fato que define a homogeneidade de grau 1 da derivada direcional. Assim sendo, notamos por (43) que conseguimos aumentar o valor de $v_i'(\bar{y}_i; z_i)$, aumentando-se \bar{z} (i.e., obtendo-se série de vetores paralelos de norma crescente com \bar{z}). Evidentemente isso é não desejável. Assim sendo, restringe-se z ao conjunto compacto que contém a origem, em seu interior expresso mediante (42).

A garantia do valor ótimo finito para (40) repousa na condição $\bar{y}_i \in \text{Int}Y_i$ ($i = 1 \dots k$), já que se supõe $v_i'(\bar{y}_i)$ finita ($i = 1 \dots k$).

E, destarte, a possibilidade em se obter as derivadas direcionais máximas por (39) que vitaliza o critério proposto em (40).

Uma série de propriedades associadas com o programa (40), segue:

- (a) o programa (40) é côncavo pois $v_i'(\bar{y}_i; .)$ é côncava;
- (b) $z=0$ (i.e., a direção nula) é viável;
- (c) o valor ótimo de (40) é não negativo;
- (d) valor ótimo de (40) é nulo se, e somente se, não houver direções viáveis de crescimento;
- (e) qualquer direção viável cujo valor da função objetivo de (40) seja positivo, ministra direção viável de crescimento, e vice-versa.

Empregando o resultado (39) em (40), vem:

$$\text{Maximizar } \sum_{i=1}^k \nabla f_i^t(x_i) \omega_i \text{ s. a}$$

45

$$\nabla_{g_{ij}}^t(\bar{x}_i) w_i - z_{ij} \leq 0 \quad \text{para } i=1 \dots k \text{ e } j \text{ tal que} \\ g_{ij}(\bar{x}_i) = \bar{y}_{ij};$$

$$-\nabla_{h_{ij}}^t(\bar{x}_i) w_i \leq 0 \quad \text{para } i=1 \dots k \text{ e } j \text{ tal que} \\ h_{ij}(\bar{x}_i) = 0;$$

$$\sum_{i=1}^k z_{ij} \leq 0 \quad \text{para } j \text{ tal que } \sum_{i=1}^k \bar{y}_{ij} = b_j$$

$$|z_{ij}| \leq 1 \quad \text{para todo } i \text{ e todo } j.$$

Concluimos, observando-se (44), que de sua solução ótima obteremos uma direção viável para (1), que leva ao maior incremento local da função objetivo de (1), resultado a que nos propusemos obter nesse capítulo.

Notemos que (44) é um programa linear com blocos na diagonal e restrições de acoplamento, sendo, consequentemente, interessante o emprego de técnicas especiais para a resolução desse programa linear (por exemplo, método da decomposição de Dantzig Wolfe).

Variáveis z_{ij}

Geoffrion [11; pg 16] observa que as variáveis z_{ij} não constantes de (45) podem ser fixadas no valor -1 e retiradas do problema. Como será exposto no capítulo IV, obteremos forma diversa de manusear as variáveis z_{ij} não constantes de (45).

Notemos, finalmente, que o programa linear não inclui restrições de não negatividade como usualmente ocorre, fato facilmente contornado substituindo-se cada variável do programa linear (44) pela diferença entre duas não negativas.

Otimização unidirecional

Originalmente, i.e., em [11], não foram impostas condições de compacidade sobre os conjuntos X_i , e portanto, não se garante o resultado que abaixo depreenderemos.

Obtida a direção viável $\bar{z} = (\bar{z}_1, \dots, \bar{z}_k) \in \mathbb{R}^{km}$, deparamo-nos com o problema mestre restrito:

$$46 \quad \begin{aligned} & \text{Maximizar}_{\theta} \sum_{i=1}^k v_i (\bar{y}_i + \theta \bar{z}_i) \text{ s.a} \\ & \sum_{i=1}^k (\bar{y}_i + \theta \bar{z}_i) \leq b \end{aligned}$$

Para melhor entendermos o significado das restrições de (46), explicitemos as restrições relativas aos conjuntos Y_i , reescrevendo (46) abaixo:

$$47 \quad \begin{aligned} & \text{Maximizar}_{\theta} \sum_{i=1}^k v_i (\bar{y}_i + \theta \bar{z}_i) \text{ s.a} \\ & \sum_{i=1}^k (\bar{y}_i + \theta \bar{z}_i) \leq b \quad \wedge \\ & (\forall i \in I_k) [(\bar{y}_i + \theta \bar{z}_i) \in Y_i] \end{aligned}$$

Inicialmente, voltemos ao problema mestre (1) e provemos a compactidade de seu conjunto viável consante o lema abaixo.

48 Lema: o conjunto viável do problema mestre,

$$Y \cap \sum$$

onde $\sum = \{y \in \mathbb{R}^{km} \mid \sum_{i=1}^k y_i \leq b\}$ é compacto.

Observação: a compactidade dos conjuntos X_i foi imposta em II.1.2.

Demonstração. Como a função $g_i (i \in I_k)$ é contínua em X_i , compacto, cada uma de suas componentes $g_{ij} (j=1 \dots m)$ também o será; logo existe o mínimo para tais funções em X_i , ou seja,

49 $(\exists x_{ij} \in X_i) (g_{ij} (x_{ij}) = \min_{x_i} g_{ij} (x_i))$ para
 $i=1 \dots k$ e $j=1 \dots m$.

Se $l_i \triangleq [g_{i1} (x_{i1}), \dots, g_{im} (x_{im})] \in \mathbb{R}^m$ então

50 $(\forall i \in I_k) (\forall y_i \in Y_i) (l_i \leq y_i)$

Logo, se $L_i \triangleq (l_1, \dots, l_k) \in \mathbb{R}^{km}$ decorre para $y = (y_1, \dots, y_k) \in \mathbb{R}^{km}$ que:

51 $(\forall y \in Y) (L_i \leq y)$

Por outro lado, como $y \in \sum$ de (50) concluímos que:

52

$$(\forall i \in I_K) \quad (\forall y \in \sum) (y_i \leq b - \sum_{\substack{j=1 \\ j \neq i}}^K l_j)$$

logo se $L_2 \triangleq (b - \sum_{j=2}^K l_j, \dots, b - \sum_{j=1}^{K-1} l_j) \in \mathbb{R}^{K^m}$, decorre que,

53

$$(\forall y \in \sum) (y \leq L_2)$$

de (51) e (53) decorre, finalmente:

$$(\forall y \in Y \cap \sum) (L_1 \leq y \leq L_2)$$

ou seja, qualquer y viável no problema mestre (1), tem cada uma de suas componentes entre as correspondentes componentes dos vetores L_1 e L_2 , ou seja:

$$L_{1j} \leq y_j \leq L_{2j} \quad j = 1 \dots K^m.$$

para todo $y = (y_1, \dots, y_{K^m})$ viável no problema mestre.

Ressaltamos, no entanto, que o conjunto \mathcal{Y} tal que

$$\mathcal{Y} = \left\{ y \in \mathbb{R}^{K^m} \mid L_1 \leq y \leq L_2 \right\}$$

com o sentido anteriormente exposto, não representa, em geral, o conjunto viável do problema mestre, sendo válida a relação abaixo,

$$Y \cap \sum \subset \mathcal{Y}$$

de onde se conclui a compacidade do conjunto viável do problema mestre, dado que o mesmo é fechado (Y é fechado e \sum é fechado conduz a $Y \cap \sum$ fechado).

▲

Voltando ao problema mestre restrito (47), para \bar{y} e \bar{z} fixados, dado que:

54

$$\{y \in \mathbb{R}^{k_m} | (\bar{y} + \theta \bar{z}) \in Y\} \cap \left\{ y \in \mathbb{R}^{k_m} \mid \sum_{i=1}^k (\bar{y}_i + \theta \bar{z}_i) \leq b \right\}$$

é subconjunto fechado da região viável do problema mestre, concluimos, pelo lema (48), a compacidade do conjunto em (54).

Conseqüentemente, as restrições de (47) definem simplesmente um intervalo compacto para θ . Seja $\bar{\theta}$ o limite superior desse intervalo. Definindo $\psi(\cdot)$ por meio de

$$\psi(\theta) = \sum_{i=1}^k v_i (\bar{y}_i + \theta \bar{z}_i) \text{ para } \theta \in [0, \bar{\theta}]$$

resulta o problema equivalente a (47)

55

$$\underset{\theta}{\text{Maximizar}} \quad \psi(\theta) \text{ s.a. } \theta \in [0, \bar{\theta}]$$

A função $\psi(\cdot)$ é côncava em $[0, \bar{\theta}]$ pois é soma de funções v_i côncavas nos subconjuntos convexos de Y_i .

Dois pontos fundamentalmente importantes referentes ao problema mestre restrito (47), merecem ser analisados:

- eliminação dos pontos fronteira de Y ;
- existência da solução ótima para (47).

56 Eliminação dos pontos fronteira de Y .

O valor de $\bar{\theta}$ que limita superiormente o intervalo em (55), pode levar a um ponto $\bar{y} + \bar{\theta}\bar{z}$ pertencente à fronteira de Y . Como as condições para existência da direção viável ótima exigem que $y \in \text{Int } Y$, introduz-se uma perturbação na busca unidirecional, que consiste em escolher um ponto próximo ao ótimo, se este estiver na fronteira de Y .

O intervalo $[0, \bar{\theta}(1-\bar{\gamma})]$, onde $\bar{\gamma}$ é real positivo suficientemente pequeno, define o conjunto

$$\{y \in \mathbb{R}^k \mid y = \bar{y} + \theta(1-\bar{\gamma}) \bar{z} \text{ para } \theta \in [0, \bar{\theta}]\}$$

contido no interior de Y , já que Y é convexo. Nessas condições, como $\sum_{i=1}^k v_i(\cdot)$ é côncava em Y , é contínua no interior de Y , e portanto, contínua em qualquer subconjunto contido no interior de Y . Ou seja, $\psi(\cdot)$ é contínua no conjunto compacto $[0, \bar{\theta}(1-\bar{\gamma})]$. Conclui-se, portanto, a existência do ótimo para o problema mestre restrito $\bar{\gamma}$ -perturbado.

Assim procedendo, mesmo que o ótimo esteja na fronteira de $[0, \bar{\theta}(1-\bar{\gamma})]$, i.e.,

$$\psi(\bar{\theta}(1-\bar{\gamma})) = \max_{\theta \in [0, \bar{\theta}(1-\bar{\gamma})]} \psi(\theta)$$

não haverá problemas, e o algoritmo pode ter sequência, já que o ponto $\bar{y} + \bar{\theta}(1-\bar{\gamma})\bar{z}$ gerado pela otimização unidirecional, está sempre no interior de Y , o que garante a aplicação do programa (44) para se obter uma nova direção viável

na busca do ótimo.

Finalmente, ressaltamos que a $\bar{\epsilon}$ -perturbação não é necessária para garantir a existência do ótimo de (47), o qual sempre existe, pois o problema mestre possui como já comentamos, solução ótima, e portanto, $\Psi(\theta)$ é superiormente limitada em $[0, \bar{\theta}]$, possuindo consequentemente, solução ótima nesse intervalo.

A $\bar{\epsilon}$ -perturbação propicia, portanto, a obtenção de um ponto que esteja garantidamente no interior de Y . Esta condição foi imposta para obtermos nova direção viável no ponto encontrado da otimização unidirecional, segundo o critério considerado.

Essa forma de proceder, segundo Geoffrion [11], gera uma solução viável arbitrariamente próxima à fronteira de Y , se a solução ótima do problema mestre pertencer à fronteira de Y .

A inclusão desse procedimento parece-nos crítico para a convergência do algoritmo. No entanto, dada sua simplicidade, e as condições que garante, pode ser admitido antes de se tentar introduzir qualquer outro procedimento complexo.

Tais procedimentos terão de prever o comportamento das funções v_i nas fronteiras dos conjuntos Y_i .

Subproblemas.

Obtida a direção viável \bar{z} a otimização unidirecional leva-nos à resolução de uma sequência de subproblemas.

A solução ótima obtida desses subproblemas é garantida, dada a continuidade das funções f_i nos conjuntos

compactos

$$\{x_i \in X_i \mid g_i(x_i) \leq \bar{y}_i + \theta(1-\bar{z})\bar{z}\}$$

não vazios para quaisquer $\theta \in [0, \bar{\theta}(1-\bar{z})]$, já que

$$(\forall \theta \in [0, (1-\bar{z})\bar{\theta}]) (\bar{y}_i + \theta(1-\bar{z})\bar{z} \in Y_i)$$

Conforme [11], a forma mais fácil de resolver-se o problema unidirecional, é, possivelmente, empregando-se técnicas de programação paramétrica, visando obter as soluções dos subproblemas $P_{\bar{y}_i + \theta \bar{z}_i}^i$ θ -parametrizados,

$$\text{Maximizar } f_i(x_i) \text{ s.a. } g_i(x_i) \leq \bar{y}_i + \theta \bar{z}_i$$

aumentando-se θ de pequenos valores e ao mesmo tempo mantendo-se a optimalidade dos subproblemas, cujo valor ótimo é o valor de v_i no ponto $\bar{y}_i + \theta \bar{z}_i$. Se, no entanto, técnicas de programação paramétrica não forem disponíveis, pode-se obter bons resultados com o método de busca unidirecional de Fibonacci.

Em qualquer dos casos acima, deparamo-nos com série de subproblemas que, para cada incremento em θ , podem ser independentemente resolvidos pelos métodos que melhor se lhes ajustem.

As soluções ótimas dos subproblemas θ - parametrizados, correspondentes ao valor de θ que maximiza o pro-

blema mestre restrito (47), serão utilizadas no cálculo dos gradientes dos vínculos ativos, que compõem a matriz do programa linear.

Inicialização

Para que possamos proceder como discutimos acima, i.e., de cada otimização unidirecional gerar um ponto no interior de Y , é mister que o ponto inicial $y^* \in \mathbb{R}^{km}$, esteja no interior de Y , ou seja,

$$(\forall i \in I_k) (y^* \in \text{Int } Y) \iff (\exists x_i \in X_i) (g_i(x_i) < y_i).$$

Suponhamos que um ponto viável $y = (y_1, \dots, y_k)$ possivelmente na fronteira de Y , seja obtido por qualquer método. Mostraremos como se poderá obter nesse caso, um ponto $y^* \in \text{Int } Y$, introduzindo, se necessário, perturbação no vetor -recursos b .

Consideremos os três procedimentos abaixo.

1) suponhamos que para $i \in I_k$, $g_{ij}(x_i) < y_{ij}$ e para todo $i \in I_k - \{i\}$, ocorra $g_{ij}(x_i) = y_{ij}$. Nesse caso, podemos definir a quantidade Δy por :

$$\Delta y = \frac{1}{k} (y_{ij} - g_{ij}(x_i))$$

e fazer

$$y_{ij}^* = y_{ij} + \Delta y \quad i \in I_k - \{i\}$$

$$y_{ij}^* = y_{ij} - (k-1) \Delta y$$

obtendo,

$$57 \quad (\forall i \in I_k) (\exists x_i \in \chi_i) (g_{ij}^*(x_i) < y_{ij}^*).$$

Por outro lado, notemos que

$$\begin{aligned} \sum_{i=1}^k y_{ij}^* &= \sum_{i \in I_k - \{j\}} y_{ij}^* + y_{ij}^* = \sum_{i \in I_k - \{j\}} y_{ij} + (k-1)\Delta y + y_{ij} - (k-1)\Delta y = \\ &= \sum_{i=1}^k y_{ij} \leq b_j \end{aligned}$$

Como y é viável no problema mestre, o ponto \bar{y} tal que

$$(\forall s \in I_m - \{j\}) (\bar{y}_{js} = y_{js}^*) \quad i=1 \dots k$$

$$\bar{y}_{ij} = y_{ij}^* \quad i=1 \dots k$$

onde, lembramos $I_m = \{1, \dots, m\}$, também será viável no problema mestre.

Observação: supusemos a existência de somente um índice \bar{i} tal que $g_{\bar{i}j}(x_i) < y_{ij}^*$, no entanto, se a última expressão ocorrer para vários índices i , obtém-se um ponto y^* tal que (57) seja verificada, por procedimento análogo.

2) suponhamos agora que

$$(\forall i \in I_k) (g_{ij}^*(x_i) = y_{ij}^*)$$

Nesse caso há duas alternativas :

a) se $\sum_{i=1}^k y_{ij} = b_j$, deveremos substituir o recurso b_j (suposto positivo) por $b_j(1+w)$, definindo

$$y_{ij}^* = y_{ij} + \frac{b_j \omega}{k} \quad \omega > 0 \quad i = 1 \dots k$$

Segue que,

$$(\forall i \in I_k) (\exists x_i \in X_i) (g_{ij}(x_i) < y_{ij}^*)$$

Observamos, portanto, que a viabilidade fora inicialmente violada, requerendo-se a perturbação w . A hipótese $b_j \geq 0$ pode ser eliminada, considerando-se $b_j + |b_j|w$. No entanto, como correntemente falamos em recursos e disponibilidade de recursos, preferimos, sem perda de generalidade, supor $b \geq 0$.

b) se $\sum_{i=1}^k y_{ij} < b_j$, definimos

$$\Omega = b_j - \sum_{i=1}^k y_{ij}$$

Fazendo $y_{ij}^* = y_{ij} + \frac{\Omega}{k}$ obteremos

$$(\forall i \in I_k) (\exists x_i \in X_i) (g_{ij}(x_i) < y_{ij}^*)$$

Utilizando-se esses três procedimentos poderemos, de um ponto viável no problema mestre, obter um ponto viável no problema mestre porém no interior de Y , perturbando-se quando necessário o vetor-recursos b .

Na próxima secção apresentamos o esquema do algoritmo primal de direções viáveis.

A seguir apresentaremos simplificadamente, não incluindo as perturbações anteriormente discutidas, o esquema do algoritmo para obter, a partir da resolução do problema mestre (1), a solução ótima, $x^{\text{ótimo}}$, do problema original II.1.1. A solução ótima do problema mestre (1) será $y^{\text{ótimo}}$.

No capítulo V, detalharemos mais o esquema aqui apresentado.

PASSO 1. Seja $y^0 \in \text{Int } Y$ solução viável inicial para o problema mestre (1). Resolver $P_{y^0}^i$ ($i=1 \dots k$) por algoritmos de programação côncava, obtendo as soluções ótimas \bar{x}_i^0 .

PASSO 2. Determinar, resolvendo o programa linear (44), uma direção viável \bar{z} (a matriz dos coeficientes para o programa linear (44), bem como a determinação dos vínculos ativos, são feitas usando-se os pontos \bar{x}_i^0 obtidos no passo 1). Se o valor ótimo de (44) for zero, faça $y^{\text{ótimo}} = y^0$ e $x^{\text{ótimo}} = \bar{x}^0 = (\bar{x}_1^0, \dots, \bar{x}_k^0)$, e pare; caso contrário, seja \bar{z} a direção obtida pelo programa linear (44). Vá para o passo 3.

PASSO 3. Resolver o problema mestre restrito (47) :

$$\text{Maximizar } \sum_{i=1}^k v_i (y_i^0 + \theta \bar{z}_i) \text{ s.a } \sum_{i=1}^k (y_i^0 + \theta \bar{z}_i) \leq b$$

Seja θ^* o valor de θ que resolve (47). Vá para o passo 1 mudando y^0 por $y^0 + \theta^* \bar{z}$.

No capítulo que segue, apresentaremos alguns resultados de nossas pesquisas, ressaltando que abordaremos os problemas da diferenciabilidade local das funções perturbação v_i , consideraremos o conceito de ϵ -atividade, bem como obteremos um procedimento que visa redefinir a direção obti-

da pelo programa linear (44) , tornando-a mais conveniente sob determinado ponto de vista.

C A P I T U L O IV

ALGUNS ASPECTOS COMPUTACIONAISIntrodução.

No presente capítulo apresentaremos os resultados de nossas pesquisas, visando tirar vantagens de alguns aspectos computacionais do algoritmo de direções viáveis estudado.

Por essa razão, apresentamos na secção 1, como pode a diferenciabilidade local das funções perturbação, ser utilizada visando diminuir o número de iterações do algoritmo que resolve o problema linear, bem como o espaço de memória requerido pelo computador. Por outro lado, apresenta-se um critério de ocorrência da diferenciabilidade local facilmente verificável. Ainda se expõe, para o caso extremo em que todas as funções perturbação são diferenciáveis, um critério de optimalidade.

Na secção 2 apresentamos um procedimento, o qual chamamos procedimento de redefinição, que visa melhorar a direção viável obtida pelo programa linear, redefinindo algumas componentes da direção obtida.

Finalmente, guardamos a última secção para breves comentários a respeito da convergência do algoritmo. Aí introduziremos e discutiremos rapidamente, os conceitos de ϵ -atividade para os vínculos y_{ij} e para os vínculos g_{ij} .

SECÇÃO 1 - DIFERENCIABILIDADE DAS FUNÇÕES PERTUBAÇÃO.

Nesta secção pretende-se mostrar como se pode utilizar a diferenciabilidade das funções pertubação v_i com vantagem computacional, desenvolvendo-se em seguida um critério visando detetar sua ocorrência, nos casos em que os subproblemas são resolvidos por algoritmos de direções viáveis.

Vimos que as funções pertubação definidas no capítulo II, eram diferenciáveis num ponto \bar{y}_i interior ao seu domínio, onde eram finitas, se, e somente se, tivessem um único supergradiente nesse ponto. Pelo teorema III.2.6 um único supergradiente equivale a existência de um único vetor multiplicador ótimo associado às restrições $g_i(x_i) \leq \bar{y}_i$ de $P_{\bar{y}_i}^i$, desde que $P_{\bar{y}_i}^i$ tenha solução ótima (o que se garante com as hipóteses do capítulo III já que $\bar{y}_i \in \text{Int } Y_i$).

Portanto, caso a função $v_i(\cdot)$ seja diferenciável em $\bar{y}_i \in \text{Int } Y_i$, vem :

$$(\forall z_i \in \mathbb{R}^m) (v_i'(\bar{y}_i; z_i) = \nabla v_i(\bar{y}_i)^t z_i)$$

ou ainda,

$$1 \quad (\forall z_i \in \mathbb{R}^m) (v_i'(\bar{y}_i; z_i) = \lambda_i^t z_i)$$

onde λ_i é o único supergradiente de $v_i(\cdot)$ em \bar{y}_i .

Suponhamos ser I o conjunto dos índices das funções v_i diferenciáveis em \bar{y}_i .

Utilizando-se o mesmo critério para obtenção de uma boa direção viável exposto em III.3.40 e levando em conta o conjunto I , vem :

2

$$\text{Maximizar} \sum_{i \notin I}^k v_i'(\bar{y}_i; z_i) + \sum_{i \in I} v_i'(\bar{y}_i; z_i) \text{ s.a}$$

$$\sum_{i=1}^k z_{ij} \leq 0 \quad \text{para } j \text{ tais que } \sum_{i=1}^k \bar{y}_{ij} = b_j$$

$$|z_{ij}| \leq 1 \quad \text{para } i=1 \dots k \text{ e } j=1 \dots m$$

Usando-se (1) em (2) para as funções diferenciáveis e III.2.30 em (2) para as não diferenciáveis, decorre :

3

$$\text{Maximizar} \sum_{i \notin I}^t \nabla f_i^t(\bar{x}_i) \omega_i + \sum_{i \in I} \lambda_i^t z_i \quad \text{s.a}$$

$$\nabla g_{ij}^t(\bar{x}_i) \omega_i - z_{ij} \leq 0 \quad \text{para } i \notin I \text{ e } j \text{ tais que } g_{ij}^t(\bar{x}_i) = \bar{y}_{ij}$$

$$-\nabla h_{ij}^t(\bar{x}_i) \omega_i \leq 0 \quad \text{para } i \notin I \text{ e } j \text{ tais que } h_{ij}^t(\bar{x}_i) \geq 0$$

$$\sum_{i=1}^k z_{ij} \leq 0 \quad \text{para } j \text{ tal que } \sum_{i=1}^k \bar{y}_{ij} = b_j$$

$$|z_{ij}| \leq 1 \quad \text{para } i=1 \dots k \text{ e } j=1 \dots m.$$

Para um sistema constituído de dois subsistemas, se v_1 for não diferenciável e v_2 diferenciável em \bar{y}_1 e \bar{y}_2 respectivamente, a forma matrical das restrições em (3) será :

$\underline{0}^t$	$\nabla f_1^t(\bar{x}_1)$	λ_2^t	q^t	← coeficientes de custo
-------------------	---------------------------	---------------	-------	-------------------------

$$\begin{array}{c|c|c|c|c}
-1 & \nabla g_1(\bar{x}_1) & & & 0 \\
1 & 0 & & & \underline{1} \\
-1 & 0 & & & \underline{1} \\
0 & -\nabla h_1(\bar{x}_1) & & & 0 \\
\hline & 0 & I & 0 & \underline{1} \\
& & -I & 0 & \underline{1} \\
\hline I & 0 & x & 0 & 0
\end{array} \leq \begin{array}{c|c|c|c|c}
& & & & 0 \\
& & & & \underline{1} \\
& & & & \underline{1} \\
& & & & 0 \\
& & & & \underline{1} \\
& & & & \underline{1} \\
& & & & 0
\end{array}$$

↑ segundo membro ↑
 matriz dos coeficientes

Vemos claramente pela figura anterior, que todas as colunas relativas às componentes de w_2 podem ser suprimidas.

Fica, portanto, patente a diminuição do número de restrições do programa linear (3) em relação a III.3.44, tornando menor o tempo de computação gasto na resolução do problema linear de obtenção da direção viável. Por outro lado, a diminuição do número de colunas, anteriormente citada, leva à diminuição no espaço de memória necessário para a resolução do problema linear.

Ressaltamos que a diferenciabilidade não deve ser forçada, i.e., se $v_i(\cdot)$ é não diferenciável no ponto \bar{y}_i , não deve haver deslocamento em qualquer direção a fim de restabelecer a diferenciabilidade. Exemplos simples mostram que caminhar sobre pontos de não diferenciabilidade, pode ser mais adequado (ver ε -atividade no fim deste capítulo).

Frisamos, no entanto, que se a diferenciabilidade for verificada, o procedimento acima mencionado, pode ser utilizado com vantagens computacionais.

Caso todas as funções v_i sejam diferenciáveis em \bar{y}_i , o programa (3) reduz-se ao programa (5) abaixo :

5 Maximizar $\sum_{i=1}^k \lambda_i^t z_i$ s.a

$$\sum_{i=1}^k z_{ij} \leq 0 \quad \text{para } j \text{ tais que } \sum_{i=1}^k \bar{y}_{ij} = b_j$$

$$|z_{ij}| \leq 1 \quad \text{para } i=1 \dots k \text{ e } j=1 \dots m.$$

Nesse caso, i.e., quando todas as funções v_i são localmente diferenciáveis, como havíamos citado em II.2

durante a análise econômica da realocação de recursos, é possível obter-se um teorema a ser utilizado como condição de optimalidade.

Critério de Optimalidade no Caso Diferenciável.

Jennergren [17] cita rapidamente um critério de optimalidade no caso em que todas as funções v_i são localmente diferenciáveis, não lhe dedicando maior interesse. Com base nesse critério, derivamos uma condição necessária e suficiente de optimalidade, a qual passamos a discutir.

6 Teorema : suponhamos que v_i ($i=1\dots k$) seja diferenciável em \bar{y}_i ($i=1\dots k$). Então se

$$\begin{array}{ll} M = \{ 1, \dots, m \} & J = \left\{ j \mid \sum_{i=1}^k \bar{y}_{ij} = b_j \right\} \\ I = \{ 1, \dots, k \} & Z = \left\{ z \in \mathbb{R}^{km} \mid z \text{ é viável em } \text{II.2.9} \right\} \end{array}$$

para cada $j \in J$

7 $\left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i)$ são iguais para todo $i \in I_k$ e

8 $\left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) = 0$ para todo $j \notin J$ e todo $i \in I_k$

se, e somente se, $\bar{y} = (\bar{y}_1, \dots, \bar{y}_k)$ for solução ótima para o problema mestre II.2.9 .

Demonstração. Seja $V(y)$ a função objetivo do problema mestre II.2.9 . Como \bar{y} é ótimo em II.2.9 se, e somente se,

$v'(y; z) \leq 0$, para qualquer $z \in Z$, o teorema (6) reduz-se a:

9 $\left\langle (7) \text{ e } (8) \text{ ocorrem} \right\rangle \text{ se, e somente se, } \left\langle (\forall z \in Z) (v'(\bar{y}; z) \leq 0) \right\rangle$

Passemos à demonstração de (9).

\Rightarrow : da diferenciabilidade das funções v_i em \bar{y}_i decorre

$$10 \quad v'(\bar{y}; z) = \sum_{i=1}^k v'_i(\bar{y}_i; z_i) = \sum_{i=1}^k \sum_{j=1}^m \left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) z_{ij} =$$

$$= \sum_{j=1}^m \sum_{i=1}^k \left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) z_{ij} =$$

$$= \sum_{j \in J} \sum_{i=1}^k \left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) z_{ij} + \sum_{j \notin J} \sum_{i=1}^k \left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) z_{ij} \quad \text{para todo } z \in Z$$

De (7) vem:

$$11 \quad \sum_{j \in J} \sum_{i=1}^k \left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) z_{ij} = \sum_{j \in J} \left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) \sum_{i=1}^k z_{ij} \quad \text{para todo } z \in Z$$

Como v_i ($i \in I_k$) é não decrescente, vem:

$$12 \quad (\forall i \in I_k) \left(\left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) \geq 0 \right)$$

De (11) e (12) e da condição de viabilidade para a direção z , i.e., $\sum_{i=1}^k z_{ij} \leq 0$ e $j \in J$, vem:

$$13 \quad (\forall z \in Z) \left(\sum_{j \in J} \sum_{i=1}^k \left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) z_{ij} \leq 0 \right)$$

De (8) vem :

$$14 \quad (\forall z \in \mathbb{R}^{km}) \left(\sum_{j \notin J} \sum_{i=1}^k \left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) z_{ij} = 0 \right)$$

Finalmente, levando (13) e (14) em (10), teremos :

$$(\forall z \in \mathbb{Z}) (V'(\bar{y}; z) \leq 0)$$

o que conclui a demonstração da condição suficiente.

\Leftarrow : por hipótese temos :

$$15 \quad (\forall z \in \mathbb{Z}) (V'(\bar{y}; z) \leq 0)$$

De (10) e (15) vem :

$$16 \quad (\forall z \in \mathbb{Z}) \left(\sum_{j \notin J} \sum_{i=1}^k \left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) z_{ij} + \sum_{j \in J} \sum_{i=1}^k \left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) z_{ij} \leq 0 \right)$$

Para todo $j' \notin J$ e todo $i' \in I_k$ e θ positivo e suficientemente pequeno, a direção z tal que

$$17 \quad \begin{aligned} z_{i'j'} &= \theta > 0 \\ z_{ij} &= 0 \quad \forall i \in I_k - \{i'\} \text{ e } \forall j \in M - \{j'\}, \end{aligned}$$

é viável em II.2.9 no ponto \bar{y} , i.e., $z \in \mathbb{Z}$. De (17) em (16), vem :

$$(\forall i \in I_k) (\forall j \notin J) \left(\theta \frac{\partial v_i}{\partial y_{i'j'}} (\bar{y}_i) \leq 0 \right)$$

e, portanto,

$$18 \quad (\forall i \in I_k) (\forall j \notin J) \left(\frac{\partial v_i}{\partial y_{ij}} (\bar{y}_i) \leq 0 \right).$$

De (12) e (18), temos :

$$19 \quad (\forall i \in I_k) (\forall j \notin J) \left(\left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) = 0 \right)$$

O resultado (19) pode ser usado em (16), obtendo-se :

$$20 \quad \sum_{j \in J} \sum_{i=1}^k \left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) z_{ij} < 0.$$

Para quaisquer $i'', i' \in I$ e $j \in J$ e θ suficientemente pequeno, a direção z tal que

$$21 \quad z_{i''j} = \theta$$

$$z_{i'j} = -\theta$$

$$z_{ij} = 0 \quad \text{para todo } i \in I_k - \{i', i''\} \text{ e todo } j \in M - \{j\}$$

é viável em II.2.9, i.e., $z \in Z$.

Substituindo-se (21) em (20), vem

$$\left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) \theta + \left(\frac{\partial v_{i''}}{\partial y_{ij}} \right) (\bar{y}_{i''}) (-\theta) = 0$$

ou, para cada $j \in J$

$$\left(\frac{\partial v_i}{\partial y_{ij}} \right) (\bar{y}_i) = \left(\frac{\partial v_{i''}}{\partial y_{ij}} \right) (\bar{y}_{i''})$$

para todo $i, i'' \in I_k$, o que completa a demonstração da condição necessária. Δ

Para que possamos proceder como temos discutido, i.e., utilizar a diferenciabilidade local de algumas funções perturbação v_i , com vantagens computacionais na resolução do problema de obtenção da direção viável em cada iteração do algoritmo, será importante derivar um critério para determinar quando a diferenciabilidade ocorre. No que segue, discutiremos uma condição suficiente para diferenciabilidade local das funções perturbação v_i , bastante simples em ser verificada, resultando daí particular interesse em empregá-la como critério computacional. Para tanto, partiremos da hipótese de que os subproblemas $P_{\bar{y}_i}^i$ (onde \bar{y}_i forá obtido da otimização unidirecional) sejam resolvidos por um algoritmo de direções viáveis (por exemplo, vide Zoutendijk[38]).

Como vimos, para se garantir a diferenciabilidade da função perturbação v_i em \bar{y}_i associada às restrições $g_i(x_i) \leq \bar{y}_i$ de $P_{\bar{y}_i}^i$, temos de garantir a unicidade do vetor multiplicador ótimo associado às restrições $g_i(x_i) \leq \bar{y}_i$ de $P_{\bar{y}_i}^i$ (já que a equivalência entre vetores multiplicadores ótimos e supergradientes está garantida pelas hipóteses do capítulo III).

Para tal, coloquemos o problema $P_{\bar{y}_i}^i$ no formato abaixo, mais conveniente para as discussões que se seguirão:

$$\text{Maximizar } \underset{\mathbf{x}}{f^*(x)} \text{ s.a. } f^i(x) \geq 0 \text{ para } i=1 \dots q,$$

onde, iremos supor f^i ($i=1 \dots q$) definidas em \mathbb{R}^n com valores em \mathbb{R} . Além do mais, para podermos garantir que o algorit-

mo de direções viáveis convirja, as funções f^i serão supostas diferenciáveis continuamente no \mathbb{R}^n .

Evidentemente, (22) é equivalente a $P_{\bar{y}_i}^i$, e portanto, com as hipóteses que impusemos no capítulo III, temos garantida a existência dos vetores multiplicadores ótimos associados às restrições de (22), o qual segundo as mesmas hipóteses também possuirá solução ótima. Seja \bar{x} solução ótima de (22). Se λ_i e μ_i forem vetores multiplicadores ótimos associados às restrições $g_i(x_i) \leq \bar{y}_i$ e $h_i(x_i) \geq 0$ de $P_{\bar{y}_i}^i$ temos

23

$$\Omega \triangleq \begin{bmatrix} \lambda_i \\ \mu_i \end{bmatrix}$$

também será um vetor multiplicador ótimo de (22) e vice-versa.

Destarte, para se garantir a unicidade do vetor multiplicador ótimo λ_i associado às restrições $g_i(x_i) \leq \bar{y}_i$ de $P_{\bar{y}_i}^i$, haveremos de garantir a unicidade do vetor multiplicador ótimo Ω associado às restrições de (22).

Unicidade do vetor multiplicador ótimo Ω para (22).

Resolvendo o problema, (22) por um algoritmo de direções viáveis, procede-se com as iterações do algoritmo até que o valor ótimo do problema

24

$$\text{MaxMin}_{h \in S^*} \left\langle \nabla f^i(x), h \right\rangle$$

seja nulo, onde

$$S^* \triangleq \{ h \in \mathbb{R}^n \mid |h_j| \leq 1, j = 1 \dots n \}$$

$$J_o(x) \triangleq \{ i \in \mathbb{N} \mid f^i(x) = 0 \} \cup \{ 0 \}$$

Este último conjunto, $J_o(x)$, representa o conjunto dos índices dos vínculos ativos em x incluindo-se o índice da função objetivo.

Assim sendo, como \bar{x} é solução ótima de (22) teremos :

$$25 \quad \text{Max}_{h \in S^*} \text{Min}_{i \in J_o(\bar{x})} \langle f^i(\bar{x}), h \rangle = 0$$

A condição expressa por (25), exprime o critério de optimalidade no emprego de algoritmos de direções viáveis.

No que segue, mostraremos como a partir da resolução do problema maxmin no ponto \bar{x} , poderemos garantir a existência de um único vetor multiplicador ótimo Ω . Ou seja, de posse das informações oriundas da resolução de (24) em \bar{x} , em que condições

$$26 \quad \exists \Omega \in \mathbb{R}^q \text{ tal que}$$

$$(i) \quad \nabla f^0(\bar{x}) + \sum_{i=1}^q \Omega_i \nabla f^i(\bar{x}) = 0$$

$$(ii) \quad \Omega_i f^i(\bar{x}) = 0 \quad i = 1 \dots q;$$

$$(iii) \quad \Omega \geq 0;$$

$$(iv) \quad f^i(\bar{x}) \geq 0 \quad i = 1 \dots q.$$

A condição (iv) de (26) pode ser dispensada já que \bar{x} é, por hipótese, solução ótima de (22).

Suponhamos, sem perda de generalidade, que em

\bar{x} , somente os r primeiros vínculos f^i sejam ativos , i.e,

$$f^i(\bar{x}) = 0 \quad i = 1 \dots r$$

$$f^i(\bar{x}) > 0 \quad i=r+1 \dots q.$$

Com os vínculos ativos associaremos componentes do vetor multiplicador ótimo para (22), Δ_i ($i=1 \dots r$), não necessariamente nulas, enquanto aos vínculos inativos estão associadas as componentes nulas do vetor multiplicador ótimo.

A obtenção de condições suficientes para unicidade do vetor multiplicador ótimo $(\Delta^t, 0^t)$, resume-se em três etapas :

etapa 1 : dualização do maxmin (24) em \bar{x} ;

etapa 2 : relação entre variáveis duais do maxmin e vetores multiplicadores ótimos para (22);

etapa 3 : critério de unicidade da solução ótima de um programa linear sem restrições de não-negatividade.

Etapa 1. Dualização do maxmin (24) em \bar{x} .

Inicialmente, transformemos convenientemente a formulação do problema maxmin (24), a fim de o colocar na forma padrão de um problema equivalente de programação linear .

De (24) , temos

$$\begin{aligned} \text{Maximizar } h_0 & \quad \text{s.a} \quad h_0 = \min_{\substack{i \in J_0(x) \\ h \in S^*}} \left\langle \nabla f^i(x), h \right\rangle \end{aligned}$$

ou,

$$27 \quad \begin{array}{lll} \text{Maximizar} & h_0 & \text{s.a.} \\ & \begin{bmatrix} h_0, h \end{bmatrix} & h_0 - \langle \nabla f^i(x), h \rangle \leq 0 \quad i \in J_0(x) \\ & & |h_j| \leq 1 \quad j=1 \dots n. \end{array}$$

Dualizando-se (27), decorre :

$$\begin{array}{ll} \text{Minimizar} & \underline{0}^t u^0 + \underline{1}^t u^1 + \underline{1}^t u^2 \quad \text{s.a.} \\ & u^0 u^1 u^2 \\ & (\underline{u^0}, \underline{u^1}, \underline{u^2}) \geq \underline{0}^t \end{array}$$

$$28 \quad \begin{bmatrix} 1 & 1 & \dots & 1 & \underline{0}^t & \underline{0}^t \\ -\nabla f^0(x) & -\nabla f^1(x) & \dots & -\nabla f^n(x) & I & -I \end{bmatrix} \begin{bmatrix} u^0 \\ u^1 \\ u^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

onde u^0 é vetor de variáveis duais associadas às restrições $h_0 - \langle \nabla f^i(x), h \rangle \leq 0$, $i \in J_0(x)$, u^1 vetor de variáveis duais associadas às restrições $h_j \leq 1$, $j=1 \dots n$, u^2 vetor de variáveis duais associadas às restrições $h_j \geq -1$, $j=1 \dots n$ e $\underline{1} \in \mathbb{R}^m$ com todas as componentes iguais a unidade.

As restrições em (28) podem ser desenvolvidas como mostramos a seguir,

$$\begin{bmatrix} 1 & 1 & \dots & 1 & \underline{0}^t & \underline{0}^t \\ -\nabla f^0(x) & -\nabla f^1(x) & \dots & -\nabla f^n(x) & I & -I \end{bmatrix} \begin{bmatrix} u^0 \\ u^1 \\ u^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

portanto,

$$\begin{aligned} \underline{1}^t u^0 + \underline{0}^t u^1 + \underline{0}^t u^2 &= 1 \\ -\sum_{i=0}^n \nabla f^i(x) u_i^0 + \underline{1}^t u^1 - u^2 &= 0 \end{aligned}$$

ou finalmente,

$$29 \quad \sum_{i=0}^n \nabla f^i(x) u_i^0 + u^1 - u^2 = 0 \quad \text{e} \quad \sum_{i=0}^n u_i^0 = 1.$$

Usando-se (29) o dual do maxmin (24) pode então ser escrito como segue :

$$30 \quad \begin{array}{l} \text{Minimizar} \quad \underline{0}^t u^0 + \underline{1}^t u^1 + \underline{1}^t u^2 \quad \text{s.a} \\ \underline{u^0, u^1, u^2} \\ (\underline{u^0}, \underline{u^1}, \underline{u^2}) \geq \underline{0} \\ \sum_{i=0}^n u_i^0 = 1 \\ \sum_{i=0}^n \nabla f^i(x) u_i^0 + u^1 - u^2 = 0 \end{array}$$

O maxmin (27) resolvido no ponto de ótimo \bar{x} para (22), terá valor ótimo, \underline{h}_0^* , nulo. Pelo Teorema da Dualidade em Programação Linear [9; pg 116], o problema (30), dual de (27), resolvido para $x = \bar{x}$ fornecerá :

$$31 \quad \underline{0}^t \underline{u^0} + \underline{1}^t \underline{u^1} + \underline{1}^t \underline{u^2} = 0$$

onde $(\underline{u^0}, \underline{u^1}, \underline{u^2})$ é o vetor dual ótimo.

De (31) e da restrição de não-negatividade para as variáveis duais decorre :

$$\underline{u^1} = \underline{u^2} = \underline{0}$$

e, portanto, o dual do maxmin (27) no ponto de ótimo \bar{x} poderá ser escrito como segue :

32

encontrar $\mathbf{u}^0 \geq 0$ tal que

$$\sum_{i=0}^n \mathbf{u}_i^0 = 1$$

$$\sum_{i=0}^n \nabla f^i(\bar{x}) \mathbf{u}_i^0 = 0$$

Assim sendo, a solução ótima do dual do maxmin em \bar{x} será $(\mathbf{u}^0, \underline{0}^t, \underline{0}^t)$.

Obtido o dual (32) do maxmin (27) para $x = \bar{x}$, passamos à segunda etapa, onde se relacionará as variáveis duais ótimas não necessariamente nulas, \mathbf{u}_i^0 ($i=0 \dots r$), do maxmin em $x = \bar{x}$ com as componentes não necessariamente nulas dos vetores multiplicadores ótimos, $(\Omega^t, \underline{0}^t)$, para (22).

Etapa 2. Relação entre as variáveis duais ótimas de (27) em \mathbf{u}^0 e as componentes dos vetores multiplicadores ótimos de (22) em Ω .

Os dois resultados abaixo caracterizam a relação entre as variáveis duais ótimas não necessariamente nulas do maxmin (27) no ponto \bar{x} , e as componentes não necessariamente nulas dos vetores multiplicadores ótimos para (22).

Lema 1 - Se $\mathbf{u}_0^0 = 0$ então

- A) existe mais de um vetor multiplicador ótimo associado às restrições de (22);
- B) o problema (32) tem solução não única.

Demonstrações. Antes de iniciarmos as demonstrações, lembra mos que tanto a solução ótima de (32) (garantida pelo Teorema da Dualidade em Programação Linear), quanto os vetores multiplicadores ótimos associados a (22), existem dadas as

hipóteses dos capítulos anteriores.

Asserção A. Suponhamos, como fizemos para escrever (27), que somente os primeiros r vínculos f^i de (22) sejam ativos em \bar{x} . Nesse caso, das condições de não-negatividade e complementaridade de caracterização dos vetores multiplicadores ótimos, decorre que qualquer vetor multiplicador ótimo de (22) é da forma $(\Omega^t, \underline{0}^t)$, onde $\Omega \in \mathbb{R}^n$, sendo associado às primeiras r restrições de (22). Seja $(\underline{u}^t) = (0, \underline{u}^t)$ solução de (32). Provemos que $(\Omega^t + \underline{u}^t, \underline{0}^t) \in \mathbb{R}^q$ é também um vetor multiplicador ótimo para (22).

Por hipótese, como $(\Omega^t, \underline{0}^t)$ é um vetor multiplicador ótimo para (22) vem :

$$33 \quad \nabla f^o(\bar{x}) + \sum_{i=1}^n \Omega_i^t \nabla f^i(\bar{x}) = \underline{0}$$

Como, por hipótese, $(0, \underline{u}^t)$ é solução de (32) decorre

$$34 \quad \sum_{i=1}^n u_i^t \nabla f^i(\bar{x}) = \underline{0}$$

De (33) e (34), vem :

$$\begin{aligned} \nabla f^o(\bar{x}) + \sum_{i=1}^n \Omega_i^t \nabla f^i(\bar{x}) + \sum_{i=1}^n u_i^t \nabla f^i(\bar{x}) &= \\ = \nabla f^o(\bar{x}) + \sum_{i=1}^n (\Omega_i^t + u_i^t) \nabla f^i(\bar{x}) &= \underline{0}. \end{aligned}$$

As outras condições que caracterizam os vetores multiplicadores ótimos associados a (22), i.e., não negativi

dade das componentes de $(\Omega^t, \underline{u}^t, \underline{0}^t)$ e complementaridade são facilmente verificadas.

Portanto, temos que existe mais que um vetor multiplicador ótimo para (22).

Δ

ASSERÇÃO B. Provemos que o problema (32) tem solução não única. Seja $(\Omega^t, \underline{0}^t)$ vetor multiplicador ótimo para (22). Prove mos que (K_0, K_1, \dots, K_r) tal que

$$K_0 \triangleq \frac{1}{1 + \sum_{i=1}^r \Omega_i}$$

$$K_i \triangleq \frac{\Omega_i}{1 + \sum_{i=1}^r \Omega_i} \quad i=1 \dots r$$

é solução de (32).

Usando as definições das componentes de (K_0, K_1, \dots, K_r) vem

$$\begin{aligned} \sum_{i=0}^r K_i^0 &= \frac{1}{1 + \sum_{i=1}^r \Omega_i} + \sum_{i=1}^r \frac{\Omega_i}{1 + \sum_{i=1}^r \Omega_i} = \frac{1}{1 + \sum_{i=1}^r \Omega_i} \left(1 + \sum_{i=1}^r \Omega_i \right) = \\ &= 1 \end{aligned}$$

Como $\Omega_i > 0$ ($i=1 \dots r$), vem $K_i > 0$ ($i=0 \dots r$).

Finalmente,

$$\begin{aligned} \sum_{i=0}^r \nabla f^i(\bar{x}) K_i &= \nabla f^0(\bar{x}) \frac{1}{1 + \sum_{i=1}^r \Omega_i} + \sum_{i=1}^r \nabla f^i(\bar{x}) \frac{\Omega_i}{1 + \sum_{i=1}^r \Omega_i} = \\ &= \frac{1}{1 + \sum_{i=1}^r \Omega_i} \left(\nabla f^0(\bar{x}) + \sum_{i=1}^r \nabla f^i(\bar{x}) \Omega_i \right) = 0 \end{aligned}$$

já que $(\Omega^t, \underline{0}^t)$ é um vetor multiplicador ótimo para (22). De corre, assim, que (32) tem solução não única.

Δ

Teorema. Seja ${}^4u_0^o \neq 0$. Então (32) tem solução única se, e sómente se, existir um único vetor multiplicador ótimo associado às restrições de (22).

Demonstração. \iff : já que ${}^4u_0^o \neq 0$, pelo lema 1, a demonstração deste teorema reduz-se a um simples manuseio algébrico. Seja $(\underline{\Omega}^t, \underline{u}^t)$ o único vetor multiplicador ótimo para (22). Temos :

$$\exists (\underline{\Omega}^t, \underline{u}^t) \in \mathbb{R}^q \mid \underline{\Omega} \geq 0 \wedge \nabla f^o(\bar{x}) + \sum_{i=1}^n \underline{\Omega}_i \nabla f^i(\bar{x}) = 0$$

equivale a

$$\exists (\underline{\Omega}^t, \underline{u}^t) \in \mathbb{R}^q \mid \underline{\Omega} \geq 0 \wedge \frac{1}{1 + \sum_{i=1}^n \underline{\Omega}_i} \nabla f^o(\bar{x}) + \sum_{i=1}^n \frac{\underline{\Omega}_i}{1 + \sum_{i=1}^n \underline{\Omega}_i} \nabla f^i(\bar{x}) = 0$$

equivale a

$$\exists \underline{u}^o \in \mathbb{R}^{n+1} \mid \underline{u}^o \geq 0 \wedge \sum_{i=1}^n \underline{u}_i^o \nabla f^i(\bar{x}) = 0 \wedge \sum_{i=1}^n \underline{u}_i^o = 1$$

onde

$$\underline{u}_0^o = \frac{1}{1 + \sum_{i=1}^n \underline{\Omega}_i} \quad \wedge \quad \underline{u}_i^o = \frac{\underline{\Omega}_i}{1 + \sum_{i=1}^n \underline{\Omega}_i} \quad (i=1 \dots n).$$

Do teorema acima segue que se $(\underline{u}_0^o, \underline{u}^t)$ for solução única de (32), o problema (22) tem associado um único vetor multiplicador ótimo $(\underline{\Omega}^t, \underline{u}^t) \in \mathbb{R}^q$, cujas componentes não necessariamente nulas são dadas por :

$$35 \quad \underline{\Omega}_i = \frac{\underline{u}_i^o}{\underline{u}_0^o} \quad i=1 \dots r$$

De (23) e (35) podemos facilmente obter o vetor λ_i , gradiente da função perturbação v_i em \bar{y}_i .

Portanto, o teorema anterior reduziu, caso ${}^4u_0^o \neq 0$, a unicidade do vetor multiplicador ótimo de (22) à unicidade da solução ótima de (32), que por sua vez equivale à unicida

de da solução ótima do dual do maxmin (27) no ponto \bar{x} .

Portanto, nosso problema da diferenciabilidade da função perturbação associada às restrições $g_i(x_i) \leq \bar{y}_i$ de $P_{\bar{y}}$, recai no problema de se garantir a unicidade da solução ótima do dual de um problema de programação linear sem restrições de não-negatividade. Na etapa 3 veremos como isso poderá ser feito.

Etapa 3. Critério para Unicidade da Solução Ótima do Dual de um Problema Linear sem Restrições de Não-Negatividade.

Iniciemos apresentando o teorema abaixo que pode ser encontrado em Lasdon [21; pg 161].

Teorema. Seja o programa linear abaixo:

$$\text{Minimizar } \underset{x \geq 0}{c^t x} \text{ s.a. } \bar{A}x - b = 0.$$

Consideremos a função perturbação associada a esse problema :

$$\underline{y}(y) = \inf \left\{ c^t x \text{ s.a. } x \geq 0 \text{ e } \bar{A}x - b = y \right\} .$$

Então $\underline{y}(.)$ é diferenciável em $y=0$ se o problema

36 Minimizar $\underset{x}{c^t x}$ s.a. $x \geq 0$ e $Ax = b$

tiver uma solução ótima não degenerada.

Esse teorema pode ser aplicado para garantirmos que o problema dual do maxmin, tenha solução ótima única. Ve

jamos como isso pode ser feito.

Do fato de que $\underline{v}(.)$ é diferenciável em $y=0$ se, e somente se, tiver um único subgradiente em $y=0$, e da equivalência entre vetor dual ótimo associado às restrições $Ax=b$ de (36) e subgradiente de \underline{v} em $y=0$, temos, usando o teorema anterior vem :

37 o problema dual associado a (36) tem solução ótima única se, (36) tiver uma solução ótima não degenerada.

Consideremos agora um problema na forma como é apresentado o maxmin, i.e.,

38 Maximizar $\underset{w}{c^t w}$ s.a. $A w \leq b$.

Supondo-se $b \geq 0$ (o que sempre ocorre no maxmin (27)), introduzindo-se variáveis de folga p , temos de (38) :

Maximizar $\underset{w, p}{c^t w}$ s.a. $A w + p = b$.

Mudando-se w por $w_2 - w_1$ com $w_1, w_2 \geq 0$, vem :

39 Maximizar $\underset{(w_1^t, w_2^t, p^t) \geq 0}{c^t (w_2 - w_1)}$ s.a. $A (w_2 - w_1) + p = b$

Esse problema (39) está justamente na forma de (36). Uma solução ótima não degenerada para (39) é uma solução tal que as variáveis w_1, w_2, p básicas sejam não nulas. No entanto, se (w_1^*, w_2^*, p^*) for solução ótima de (39), $(w_1^* + a, w_2^* + a, p^*)$, com $a > 0$, também será solução ótima. Desta forma, pode-se ignorar

degeneração nas variáveis w_1, w_2 , interessando somente o comportamento das variáveis p básicas.

Assim sendo, aplicando-se (37) a (39) decorre :

40 o dual de (39), ou equivalentemente, o dual de (38), tem solução ótima única se, (39) tiver u ma solução ótima tal que as variáveis de folga p básicas sejam não nulas.

Das três etapas anteriores concluímos : se $u^* \neq 0$, aplicando-se o resultado (40), para garantirmos a solução ótima do dual do problema maxmin, obtém-se o teorema a seguir.

Teorema. Se $u^* \neq 0$, a função perturbação v_i associada às restrições $g_i(x_i) \leq \bar{y}_i$ do problema $P_{\bar{y}_i}^i$, é diferenciável em \bar{y}_i se, o problema maxmin (no ponto de ótimo para $P_{\bar{y}_i}^i$) tiver uma solução ótima cujas variáveis de folga básicas são todas não nulas.

Esse teorema encerra uma condição suficiente bastante simples para a verificação da diferenciabilidade local da função perturbação v_i , podendo como tal ser utilizado como critério computacional para diferenciabilidade.

Na secção que segue, apresentaremos forma diversa da discutida em III.3, devido a Geoffrion, de se operar com as componentes z_{ij} , $i=1...k$ e $j=1...m$, da direção viável.

Obteremos um procedimento, o qual denominaremos procedimento de redefinição, por meio do qual visamos melhorar

a direção viável obtida por Geoffrion, no sentido de evitar a rápida tendência em se inviabilizar os subproblemas, ao mesmo tempo que procuraremos forçar quando possível, a alocação de recursos em níveis globais máximos (i.e., distribuição da disponibilidade total).

SEÇÃO 2 - PROCEDIMENTO DA REDEFINIÇÃO.

Em III.3 atentamos para como Geoffrion operava com as variáveis z_{ij} , excluindo do programa de obtenção, da direção viável de cada iteração do algoritmo, algumas variáveis z_{ij} . Essa eliminação, quando possível, é feita em dois níveis, fixando-se algumas z_{ij} em -1 e outras em +1, conforme discutiremos mais adiante.

No entanto, não nos restringiremos a essa forma de operar, redefinindo, após calculadas as componentes z_{ij} não previamente fixadas, aquelas componentes da direção viável inicialmente fixadas no valor -1.

O procedimento da redefinição objetivará obter direções em que, sem modificar o aumento local da função objetivo do problema mestre II.2.9, minimize a tendência de violar os vínculos ignorados na busca da direção. Procuraremos desta maneira obter direções que penetrem mais na região viável, aumentando o passo do algoritmo, além de se maximizar o total dos níveis de recursos alocados.

Fica, portanto, claro, que tal procedimento sómente poderá ser utilizado, quando existir uma infinidade de direções viáveis, que levam ao mesmo aumento local da função objetivo do problema mestre.

Descrição do Procedimento.

Na busca da direção viável z a partir de um ponto \bar{y} , apresentada em III.3.44, eliminam-se as componentes z_{ij} tais que $g_{ij}(x_i) < \bar{y}_{ij}$. Estas componentes de z podem ser previamente arbitradas, de modo a se garantir a viabilidade dos vínculos

$$41 \quad |z_{ij}| \leq 1$$

$$42 \quad \sum_{i=1}^k z_{ij} \leq 0 \quad \text{para } j \text{ tais que } \sum_{i=1}^k \bar{y}_{ij} = b_j$$

Para se obter a máxima vantagem na escolha dos valores dos z_{ij} arbitrados, devemos levar em conta os seguintes fatos:

- (a) valores pequenos para as componentes z_{ij} ($z_{ij} = -1$) dão-nos a máxima liberdade com respeito a (42) para as outras componentes de z ;
- (b) valores grandes ($z_{ij} = 1$) proporcionam as melhores soluções para os subproblemas perturbados por $\bar{y}_i + \theta z_i$, uma vez que as funções perturbação são não decrescentes.

Para apresentar o procedimento de arbitragem das componentes z_{ij} de z proposto, inicialmente definiremos os seguintes conjuntos :

$$43 \quad J(i) = \{ j \in \mathbb{N} \mid g_{ij}(\bar{x}_i) = \bar{y}_{ij} \} \quad i = 1 \dots k.$$

(referiremos-nos a esses vínculos como vínculos g_{ij} ativos em \bar{x}_i)

44

$$J = \left\{ j \in \mathbb{N} \mid \sum_{i=1}^k \bar{y}_{ij} = b_j \right\}$$

(referiremos a esses vínculos como vínculos y_{ij} ativos em \bar{y}_{ij}).

As componentes a serem arbitradas são as correspondentes a $\bar{J}(i)$ (o traço acima do conjunto indicará o seu complementar), enquanto as pertencentes a $J(i)$ correspondem às componentes da direção, a serem obtidas pelo programa linear III.3.44 .

A determinação da direção viável z será feita em duas etapas.

Etapa 1.

Fixam-se em

45

$$z_{ij} = -1$$

as componentes tais que $j \in \bar{J}(i) \cap J$, devido ao fato (a), e

46

$$z_{ij} = 1$$

as componentes tais que $j \in \bar{J}(i) \cap \bar{J}$, devido ao fato (b), uma vez que essas componentes não aparecem em (42). As demais componentes de z são , então, obtidas por III.3.44 .

Obtida a direção \bar{z} , deve-se procurar melhorar os valores fixados por (45), para tirarmos proveito do fato (b). Isso será possível sempre que para algum $j \in J$:

47

$$\sum_{i \in I(j)} \bar{z}_{ij} < L \quad I(j) = \left\{ i \in I_k \mid g_{ij}(\bar{x}_i) = \bar{y}_{ij} \right\}$$

onde L é, portanto, o número de componentes z_{ij} inicialmente fixadas em -1 para o índice j considerado em (47).

Nessas condições, passamos à etapa 2.

Etapa 2.

Fixam-se, para cada $j \in J$ onde (47) ocorre,

$$48 \quad \bar{z}_{ij} = -1 + \frac{L - \sum_{i \in I(j)} \bar{z}_{ij}}{L} \quad \text{para } i \text{ tais que } i \notin I(j).$$

Como as quantidades $L - \sum_{i \in I(j)} \bar{z}_{ij}$ são não negativas para cada $j \in J$, o valor das componentes da direção viável inicialmente fixadas em -1 , é, se possível, convenientemente redefinido, a fim de cumprir com o objetivo (b).

Observação : caso os \bar{z}_{ij} redefinidos por (48) não pertençam ao intervalo $[-1, +1]$, fixamo-los em -1 ou $+1$ conforme o valor obtido por (48) seja, respectivamente, negativo ou positivo.

Na secção 3 abordaremos sem quaisquer detalhes, que podem ser considerados num trabalho mais completo, a convergência do algoritmo primal de direções viáveis.

SEÇÃO 3 - CONVERGÊNCIA.

Quanto às propriedades de convergência do algoritmo estudado, são ainda problemas em aberto.

Parece-nos crucial para convergência do presente

algoritmo, a consideração das perturbações presentes nas otimizações unidirecionais, com a finalidade, como vimos, de se garantir que cada ponto obtido da resolução do problema mestre restrito III.3.46 esteja no interior do conjunto Y . Há, portanto, a necessidade de melhor se estudar os pontos fronteira do conjunto Y .

Por outro lado, dois conceitos podem ainda ser considerados, que repercutem positivamente na convergência do algoritmo. São eles :

- 1) ε -atividade para os vínculos y_{ij} ;
- 2) ε -atividade para os vínculos g_{ij} ;

Vínculos y_{ij} ε -ativos.

Diremos que o vínculo y_{ij} é ε -ativo se,

$$\varepsilon \geq b_j - \sum_{i=1}^k y_{ij} \quad \wedge \quad \varepsilon > 0$$

Considerar vínculos y_{ij} ε -ativos no problema de obtenção da direção viável III.3.44 para o problema mestre, significa incluir as restrições :

$$\sum_{i=1}^k z_{ij} \leq 0 \quad \text{para } j \text{ tais que } \sum_{i=1}^k z_{ij} - b_j \geq -\varepsilon \quad \wedge \quad \varepsilon > 0$$

Considerar vínculos y_{ij} ε -ativos é simplesmente introduzir restrições no problema diretor, que garantem a proximidade da fronteira do politopo na obtenção da direção viável.

Frisamos que esse tipo de ε -atividade não é

necessária para a convergência do algoritmo considerado, já que basicamente operamos sobre um politopo (vide Zoutendijk [38]). Conduz, no entanto, à obtenção de direções que propiciam convergência (se a mesma ocorrer) mais rápida. Por essa razão, adotaremos (vide capítulo V) forma simplificada de se lidar com a ϵ -atividade dos vínculos g_{ij} .

Vínculos g_{ij} ϵ -ativos.

O vínculo g_{ij} , $i=1\dots k$ e $j=1\dots m$, é dito ϵ -ativo em \bar{x}_i se,

$$\epsilon > \bar{y}_{ij} - g_{ij}(\bar{x}_i) \quad \epsilon > 0.$$

Considerar esse tipo de ϵ -atividade no problema de obtenção da direção viável para o problema mestre, significa incluir as restrições

$$\nabla g_{ij}^t(\bar{x}_i) \omega_i - z_{ij} \leq 0 \quad \text{para } j \text{ tais que } \epsilon > \bar{y}_{ij} - g_{ij}(\bar{x}_i)$$

Esse tipo de ϵ -atividade é fundamental na obtenção da convergência de algoritmos de direções viáveis.

Hogan [16] a partir da consideração de ϵ -atividade para tais vínculos, consegue provar a convergência de um algoritmo de Frank-Wolfe no caso em que a função objetivo do problema é côncava e não diferenciável.

Intuitivamente, a consideração de vínculos ϵ -ativos para os subproblemas, i.e., para os vínculos g_{ij} , deteta a proximidade de "quinas" no gráfico da função

perturbação, que é côncava e não diferenciável. Evita-se assim a ocorrência de fenômenos de ziguezague devidos a não diferenciabilidade da função critério.

A convergência de algoritmos para funções não diferenciáveis é ainda quase totalmente desconhecida, conhecendo-se alguns resultados para problemas de min-max, referenciados em Hogan [16].

É interessante observar que o problema de minimizar uma função perturbação, pode ser em geral colocado no formato de um problema de min-max, após alguma manipulação de vínculos, o que sugere um interessante caminho para futuras pesquisas.

C A P I T U L O V

IMPLEMENTAÇÃOIntrodução.

No presente capítulo trataremos da implementação parcial do algoritmo de direções viáveis estudado, i.e., descreveremos um conjunto de subrotinas utilizadas na obtenção da direção viável para o problema mestre II.2.9 .

Apresentaremos em linhas gerais na secção 1, descrição suscinta do algoritmo, bem como destacaremos a área coberta pelo programa para computadores, elaborado em linguagem FORTRAN. Nessa secção, discutiremos o volume informacional necessário ao funcionamento do programa, o qual se detalhará na secção 2.

A secção 2 descreverá as subrotinas programadas em FORTRAN-IV GH, visando obter-se a direção viável para o problema mestre II.2.9 . Introduziremos aí algumas das modificações propostas no capítulo IV, guardando-se a secção seguinte para um exemplo.

SECÇÃO 1 - O ALGORITMO PARA IMPLEMENTAÇÃO.

Computacionalmente, como já discutimos, será con-

veniente resolver-se o problema mestre e daí derivar a solução do problema original via subproblemas.

1 Problema Original

$$\begin{aligned} \text{Maximizar}_{\mathbf{x}} \quad & \sum_{i=1}^k f_i(x_i) \text{ s.a } (\forall i \in I_k)(x_i \in X_i) \wedge \\ & \wedge \sum_{i=1}^k g_i(x_i) \leq b. \end{aligned}$$

2 Problema Mestre

$$\begin{aligned} \text{Maximizar}_{\mathbf{y}} \quad & \sum_{i=1}^k v_i(y_i) \text{ s.a } (\forall i \in I_k)(y_i \in Y_i) \\ & \wedge \sum_{i=1}^k y_i \leq b. \end{aligned}$$

3 Função Perturbação

$$v_i(y_i) = \sup \{ f_i(x_i) \text{ s.a } x_i \in X_i \wedge g_i(x_i) \leq y_i \}$$

4 Subproblemas $P_{y_i}^i$

$$\text{Maximizar}_{x_i} f_i(x_i) \text{ s.a } x_i \in X_i \wedge g_i(x_i) \leq y_i.$$

A seguir esquematizaremos o algoritmo primal de direções viáveis utilizado para resolução do problema mestre (2), o qual é baseado no algoritmo expos-
to no capítulo III, considerando-se algumas das modifica-
ções discutidas no capítulo IV:

- ξ -atividade para os vínculos y_i ;
- procedimento da redefinição .

Não serão, portanto, consideradas na parte com-
putacional, a ξ -atividade para os vínculos g_i e a dife-
renciabilidade, quando ocorre, das funções v_i . Não nos

preocuparemos também em manusear os valores de ϵ , que se rão supostos dados. Note-se que as ϵ -atividades dos vínculos servem somente para aumentar a rapidez do algoritmo (ver capítulo IV).

O conjunto das subrotinas que descreveremos na próxima secção, cobrirá o passo 2 do algoritmo que segue.

5 Algoritmo para Implementação.

PASSO 1. É dado $\bar{y} = (\bar{y}_1, \dots, \bar{y}_k)$ viável em (2). Resolver os subproblemas $P_{\bar{y}_i}$, obtendo-se as soluções ótimas \bar{x}_i . Vá para o passo 2.

Observações. Suponhamos que os subproblemas sejam resolvidos por um algoritmo de direções viáveis de Zoutendijk para funções diferenciáveis continuamente. As informações que fluem do passo 1 para o passo 2 são :

- a) ponto $\bar{y} = (\bar{y}_1, \dots, \bar{y}_k)$ viável em (2);
- b) as soluções ótimas \bar{x}_i dos subproblemas $P_{\bar{y}_i}$;
- c) os vínculos g_{ij} e h_{ij} ativos nos pontos \bar{x}_i (essa informação é um subproduto da resolução do problema de busca da direção viável para cada subproblema no ponto \bar{x}_i).

PASSO 2. Obter uma direção viável pelo programa III.3.44 substituindo-se os vínculos y_{ij} ativos por ϵ -ativos, determinados no programa principal. Para os vínculos g_{ij} e h_{ij} consideraremos somente os ativos, sendo, no entanto, imediata a inclusão da ϵ -atividade como mostraremos na descrição do programa principal (subrotina ODIVAL). Teste a optimalidade da solução viável \bar{y} no problema mestre; se o valor ótimo do programa linear III.3.44 for nulo (não ha

verá direções de crescimento local) faça $y^{\text{ótimo}} = \bar{y}$ e $x^{\text{ótimo}} = \bar{x} = (\bar{x}_1, \dots, \bar{x}_k)$, pare; caso contrário, complete a obtenção da direção viável, usando o procedimento da redefinição. Vá para o passo 3.

Observação. Na descrição do programa principal apresentaremos forma simplificada de se lidar com ϵ -atividade dos vínculos y_{ij} , conveniente quando a região viável é um polítopo.

PASSO 3. Resolva o problema mestre restrito III.3.47.

Seja $\bar{y} + \bar{\theta}\bar{z}$ o novo ponto obtido. Trocar no passo 1 \bar{y} por $\bar{y} + \bar{\theta}\bar{z}$.

Observações. O manuseio das perturbações possivelmente incluídas na otimização unidirecional, deverá ser considerado para cada caso particular. Usando-se técnicas paramétricas ou busca por Fibonacci, teremos sem trabalho extra, a solução ótima dos subproblemas $P_{\bar{y} + \bar{\theta}\bar{z}_i}^i$, razão pela qual não será necessária, após a primeira iteração, a resolução dos subproblemas $P_{\bar{y}}^i$ no passo 1.

Subrotinas e Volume Informacional.

Das subrotinas utilizadas para obtenção da direção viável, destacamos:

- a- subrotina principal;
- b- subrotinas satélites.

Na classe (a) incluímos a subrotina ODIVAL que obterá a direção viável para (2), utilizando para tal as subrotinas incluídas em (b) :

- PROLIN;
- DETMAX;
- NOVIN;
- GRAGH;
- GRAF;

O volume informacional necessário para resolução do problema computacional de obtenção da direção viável pela subrotina principal, pode ser classificado em:

6 - dados de entrada;
 7 - dados provenientes do passo 1.

Em (6) incluiremos dados inerentes à estrutura do problema original (1):

- coeficientes determinativos das funções h_i ,
 g_i e f_i ;
- número de subsistemas;
- dimensão do vetor atividade associado a cada subsistema;
- dimensões dos espaços dos contradomínios das funções h_i ;
- vetor-recursos;
- recursos não utilizados pelos subsistemas;
- ξ -valores.

Em (7) serão incluídos os dados necessários à subrotina principal provenientes do passo 1 do algoritmo:

- ponto $\bar{y}=(\bar{y}_1, \dots, \bar{y}_k)$;
- soluções ótimas, \bar{x}_i , dos subproblemas $P_{\bar{y}_i}^i$;
- vínculos g_{ij} e h_{ij} ativos em \bar{x}_i ;

A seguir discutiremos o significado de cada subrotina na obtenção da direção viável para o problema mestre (2).

SEÇÃO 2 - DESCRIÇÃO DAS SUBROTINAS.

Antes de passarmos à descrição das subrotinas satélites e principal, apresentaremos os tipos de funções consideradas nesse trabalho. Queremos, no entanto, ressaltar que particularidades aqui assumidas com respeito às funções, não restringe a aplicabilidade da subrotina principal. As funções como também seus gradientes são resultados de subrotinas, sendo, portanto, fácil a mudança dessas subrotinas por outras, conforme a conveniência do usuário. Por essa razão, preferimos o programa sublocado, i.e., o programa constituído por série de subrotinas, o qual permite facilmente a mudança de qualquer particularidade assumida no trabalho computacional.

8 Funções Vínculo.

Tanto as funções g_{ij} como as funções h_{ij} serão, como já mencionamos, definidas em \mathbb{R}^n com valores em \mathbb{R} . O número das funções g_{ij} , para cada subsistema, será no máximo m , dimensão do vetor-recursos. Poderá, no entanto, ser menor que m para um subsistema, se o mesmo não utilizar determinadas componentes do vetor-recursos b . As funções h_{ij} são, para cada subsistemas, em número qualquer.

Assim sendo, sejam os parâmetros presentes na subrotina principal :

S - número de referência do subsistema;

MR - dimensão do vetor-recursos b;

KS - número de subsistemas;

L - lista de tamanho KS; L(S) representa o número de funções g_{ij} não identicamente nulas associadas ao subsistema S;

M - lista de tamanho KS; M(S) representa o número de funções h_{ij} associadas ao subsistema S;

N - lista de tamanho KS; N(S) é a dimensão do espaço do vetor atividade associado ao subsistema S;

IGPER - lista que contém os índices das componentes do vetor-recursos não utilizadas pelos subsistemas; essa lista é formada de KS partições cada uma de tamanho MR-L(S), $S=1 \dots KS$; em cada partição os índices aparecem em ordem crescente (mais a frente exemplificaremos o uso dessa lista).

Tanto as funções g_{ij} quanto as funções h_{ij} serão da forma:

$$\sum_{t=1}^{n_t} \alpha_t x_t^{\beta_t}$$

onde $x_t \in \mathbb{R}$, $t=1 \dots n_t$, e as constantes reais α_t e β_t são dadas pelas listas AA (CC) e BB (DD) respectivamente, para os vínculos g_{ij} (h_{ij}); n_t é a dimensão do vetor atividade, sendo dada na lista N para cada subsistema.

Observação. Antes de apresentarmos as listas AA e BB, notemos que uma matriz qualquer pode ser convenientemente

transformada numa lista, a qual se associam duas quantidades que a caracterizam:

PC - passo da coluna da matriz;

PL - passo da linha da matriz.

Seja a matriz M abaixo:

$$M = \begin{bmatrix} 1 & 5 \\ 2 & 2 \\ 3 & 1 \end{bmatrix}$$

\uparrow $M(3,2)$

A matriz M pode ser transformada na lista LM , por exemplo, listando os elementos de M coluna após coluna, i.e.:

$$LM = (1, 2, 3, 5, 2, 1)$$

\uparrow $LM(6)$

Assim sendo, dado o elemento $M(I,J)$ de M , obtemos o correspondente elemento da lista LM , usando as quantidades PCM , passo da coluna de M , e PLM , passo da linha de M . Optando-se por listar-se colunas após colunas teremos:

$$PLM = 1$$

$$PCM = 3$$

Logo,

$$M(I,J) = LM((I-1)PLM + (J-1)PCM + 1)$$

Por exemplo,

$$M(3,2) = LM((3-1)1 + (2-1)3 + 1) = LM(6)$$

Destarte, da observação anterior qualquer matriz pode ser transformada numa lista através de uma relação

biunívoca. Essa forma de proceder utilizaremos durante toda parte computacional, i.e., toda matriz a que nos referirmos será uma lista coluna após coluna.

Essa forma opcional de manipular matrizes, leva à generalidade das subrotinas, não havendo preocupações com comandos COMMON.

As listas AA (ou CC) e BB (ou DD) terão a forma abaixo:

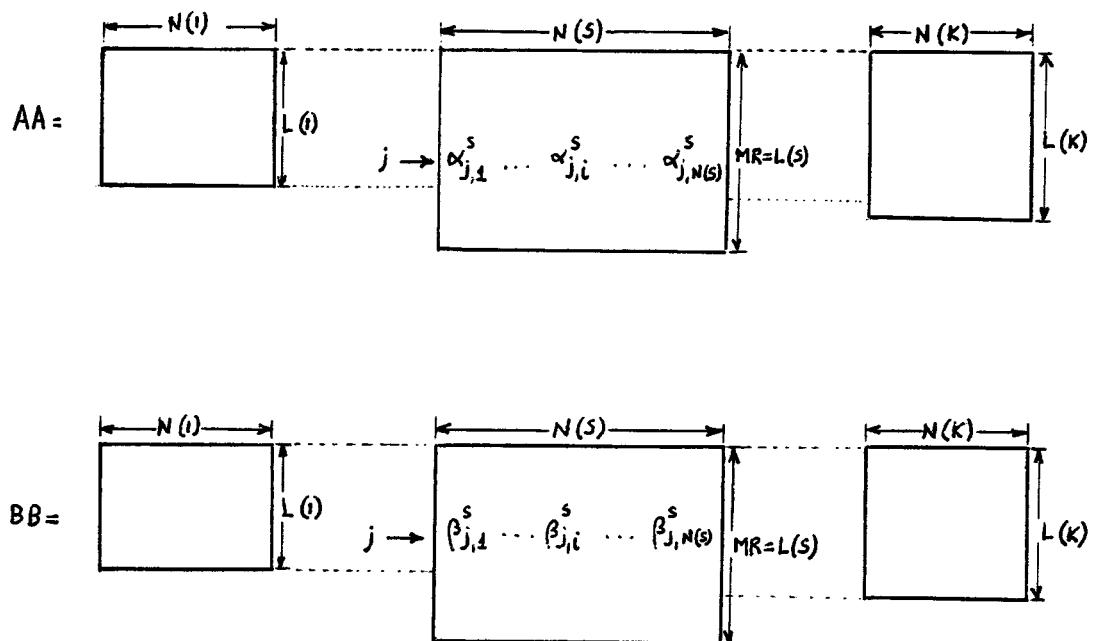


fig.1

A lista AA (fig.1) é constituída de várias submatrizes. A primeira partição da lista AA, refere-se à primeira submatriz, estando associada ao primeiro subsistema. Portanto a lista AA terá KS partições, sendo o pri

meiro elemento da primeira partição $AA(1)$, e o último elemento da KS -ésima partição $AA(\sum_{s=1}^{KS} N(s) \cdot L(s))$, i.e., após listarmos a primeira submatriz coluna após coluna, passamos a listar a segunda submatriz, sempre aumentando o índice. Mais a frente exemplificaremos a lista AA e BB.

Na construção do vínculo g_{sj} em $x^s \in \mathbb{R}^{N(s)}$, a linha j da S -ésima submatriz da lista AA (vide fig.1), representa os coeficientes das componentes do vetor atividade x^s , sendo os correspondentes coeficientes da lista BB, os expoentes dessas componentes, i.e.,

$$g_{sj}(x^s) = \alpha_{j,1}^s x_1^{\beta_{j,1}^s} + \dots + \alpha_{j,i}^s x_i^{\beta_{j,i}^s} + \dots + \alpha_{j,N(s)}^s x_{N(s)}^{\beta_{j,N(s)}^s}$$

onde

$$x^s = (x_1, \dots, x_{N(s)}).$$

$$10 \quad \alpha_{j,i}^s = AA \left(\sum_{t=1}^{s-1} N(t) \cdot L(t) + (I-1) \cdot PCC + (J-1) \cdot PLC + 1 \right) \quad i=1 \dots N(s)$$

sendo, PCC passo da coluna da S -ésima submatriz da lista AA e PLC o passo da linha da referida submatriz.

Vemos, portanto, que tendo as listas L e N, anteriormente definidas, para cada subsistema poderemos obter qualquer coeficiente (ou expoente) associado a qualquer componente do vetor atividade através de (10), e portanto, qualquer função vínculo.

Na construção das funções g_{ij} requer-se um dado. Ele provém do fato de que certas componentes do vetor-vetor recursos, para determinados subsistemas, pode não ter utilidade. Como dissemos no capítulo II, $g_{ij}(x_i)$ representa o nível do recurso de disponibilidade b_j , uti-

lizado pelo subsistema i . Se esse subsistema não utiliza tal recurso, teremos sempre nível zero, ou seja, $g_{ij}(.) \geq 0$. Nessas condições, os coeficientes α_{ij} em (9), correspondentes a esse vínculo g_{ij} serão todos nulos, não sendo incluídos na correspondente submatriz da lista AA. A indicação de quais componentes do vetor-recursos não são utilizadas, é como já dissemos, dada pela lista IGPER.

Seja o exemplo,

11
$$AA = \begin{array}{|c|c|c|c|c|c|} \hline & 1 & 2 & 3 & 7 & 1 & -1 & -4 & 1 & 2 & 3 \\ \hline 1 & 4 & 5 & 6 & 8 & 1 & 0 & 3 & & & \\ \hline & 9 & 2 & 3 & 2 & & & & & & \\ \hline \end{array}$$

fig.2

ou, na forma de lista,

$$AA = (1, 4, 2, 5, 3, 6, 7, 8, 9, \dots, 1, 2, 3)$$

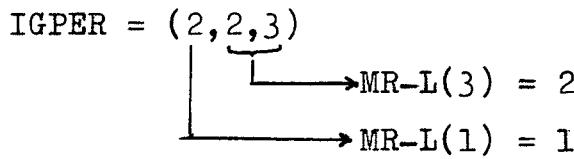
Da figura 2 conforme a figura 1 tiramos:

$$\begin{array}{ll} L(1) = 2 & N(1) = 3 \\ L(2) = 3 & N(2) = 4 \\ L(3) = 1 & N(3) = 3 \end{array}$$

Suponhamos que a dimensão do vetor-recursos seja 3, i.e., $MR=3$.

Pela figura 2 notamos que, o subsistema 1 não utiliza alguma componente de recurso e o subsistema 3 não duas delas. A lista IGPER dará para cada subsistema as componentes não utilizadas. Seja IGPER dada abaixo:

12



Observação. A partição referente ao subsistema 2 em IGPER não existe, pois esse subsistemas utiliza todas componentes do vetor-recursos b.

Assim de (11) e (12) notamos que a primeira submatriz da lista AA, representa os coeficientes de g_{11} , porém a segunda linha os coeficientes de g_{13} e não g_{12} (dado que o índice 2 está em IGPER, e portanto, o subsistema 1 não utiliza essa componente do vetor-recursos). Para o subsistema 3, os coeficientes em AA correspondem a g_{44} por mera casualidade.

Esse fato, i.e., não utilização de determinadas componentes do vetor-recursos b, leva-nos, mais adiante, a definir a lista IALG.

Como vimos, portanto, a descrição das funções vínculo g_{ij} (h_{ij}) são feitas pelas listas AA (CC) e BB (DD), que serão lidas no programa principal.

13 Funções Objetivo.

As funções objetivo aqui consideradas serão da forma:

14

$$\sum_{t=2}^{n_t+1} \alpha_t x_t^{\beta_{t-1}} + \alpha_1 e^{\sum_{t=n_t+2}^{n_t+1} \alpha_t x_{t-(n_t+1)}^{\beta_{t-1}}}$$

onde $x_t \in \mathbb{R}$, $t=1 \dots n_t$, α_t e β_t são constantes reais dadas respectivamente pelas listas EE e GG, e n_t é dado pela

lista N. As listas EE e GG terão a forma:

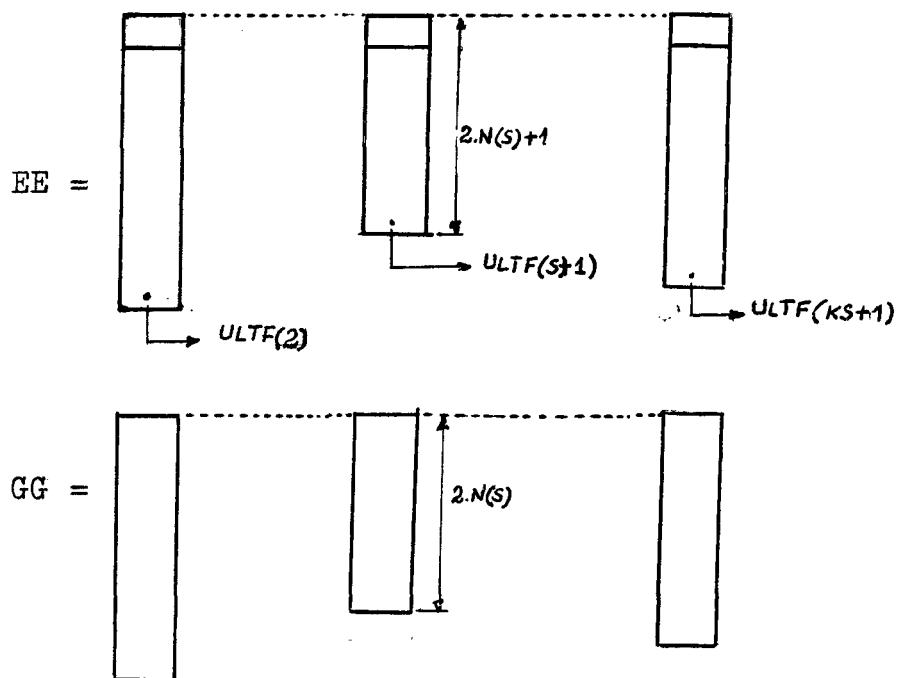


fig.3

Cada submatriz coluna da lista EE possui os co
ficientes α_t referentes à função objetivo de um determi-
nado subsistema, e cada submatriz correspondente na lista
GG os expoentes β_t . As submatrizes colunas tanto em EE co-
mo em GG são listadas colunas após colunas, como em AA.

O primeiro elemento de cada submatriz em EE é
um indicador da existência ou não da parte exponencial,
i.e., se $\alpha_1 = 0$, observando-se (14), não existirá a parte
exponencial, sendo que os coeficientes nas submatrizes
em EE correspondem ao termo polinomial. São possíveis
ainda duas outras variantes:

16

- existência da parte exponencial somente.

Observando-se (14) e a figura 3, depreendemos que a função objetivo do subsistema S, enquadra-se em (15).

No caso (16) $\alpha_i \neq 0$, mas o número de termos na submatriz em EE será menor que $2N(S) + 1$ (vide fig.3), e o número de termos na correspondente submatriz em GG será menor que $2N(S)$.

Para sabermos, baseado no número de termos dessas submatrizes, quando $\alpha_i \neq 0$, se a função objetivo possui ou não termos polinomiais, definimos a lista ULTF de tamanho $KS + 1$, a ser lida no programa principal. Sendo:

$ULTF(1) = 0;$

$ULTF(2) =$ índice do último elemento da primeira submatriz em GG (vide fig.3);

$ULTF(S) =$ índice do último elemento da $(S-1)$ -ésima submatriz em GG;

$ULTF(KS + 1) =$ índice do último elemento da KS -ésima submatriz de GG.

Assim sendo caso α_i , para o subsistema considereado seja não nulo, verificamos o valor de:

$$Q = ULTF(S+1) - ULTF(S) - 2N(S) - 1$$

Se Q for nulo então a função objetivo do subsistema S pos

sui as partes polinomial e exponencial. Caso contrário, $Q=0$, i.e., a parte polinomial não existe.

A seguir exemplificamos essas variantes. Seja o número de subsistemas $KS=3$, e as listas EE e GG dadas abaixo :

$EE =$	<table border="1"> <tr><td>1</td><td>2</td><td>0</td></tr> <tr><td>3</td><td>3</td><td>2</td></tr> <tr><td>4</td><td>4</td><td>5</td></tr> <tr><td>3</td><td>5</td><td>6</td></tr> <tr><td>2</td><td>6</td><td></td></tr> <tr><td>1</td><td>7</td><td></td></tr> <tr><td>8</td><td></td><td></td></tr> </table>	1	2	0	3	3	2	4	4	5	3	5	6	2	6		1	7		8			$GG =$	<table border="1"> <tr><td>4</td><td>4</td><td>1</td></tr> <tr><td>5</td><td>5</td><td>3</td></tr> <tr><td>6</td><td>6</td><td>2</td></tr> <tr><td>7</td><td>2</td><td></td></tr> <tr><td>8</td><td>8</td><td></td></tr> <tr><td>9</td><td></td><td></td></tr> </table>	4	4	1	5	5	3	6	6	2	7	2		8	8		9		
1	2	0																																								
3	3	2																																								
4	4	5																																								
3	5	6																																								
2	6																																									
1	7																																									
8																																										
4	4	1																																								
5	5	3																																								
6	6	2																																								
7	2																																									
8	8																																									
9																																										

ou na forma de listas teremos,

$$EE = (1, 3, 4, 3, 2, 1, 2, \dots, 2, 5, 6)$$

$$GG = (4, 5, 6, 7, 8, 4, 5, \dots, 1, 3, 2)$$

Suponhamos ainda a lista N dada a seguir:

$$N(1) = 5 \quad N(2) = 3 \quad N(3) = 3$$

Observando-se as listas EE e GG dadas acima, concluimos :

$$ULTF(1) = 0$$

$$ULTF(2) = 6$$

$$ULTF(3) = 13$$

$$ULTF(4) = 17$$

Assim sendo, vem :

$$f_1(x^1) = 1.e$$

$$f_2(x^2) = 3x_1^4 + 4x_2^5 + 5x_3^6 + 2e^{6x_1^2 + 7x_2^3 + 8x_3^9}$$

$$f_3(x^3) = 2x_1^1 + 5x_2^3 + 6x_3^2.$$

As funções objetivo ficam, consequentemente, determinadas dando-se as listas EE, GG, ULTF e N, que serão lidas no programa principal.

A seguir trataremos da descrição das subrotinas satélites e principal.

Morfologia das Subrotinas Satélites.

As subrotinas satélites, como já mencionamos, são : PROLIN, DETMAX, NOVIN, GRAGH, GRAF.

PROLIN.

Utilizada na resolução de um problema de programação linear da forma :

17 Maximizar $\underset{x}{c^t x}$ s.a $Mx \leq b$

onde $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ (sendo, portanto, M matriz $m \times n$), através do método Simplex Revisado [9; pg131]. Consideraremos $b \geq 0$.

A consideração $b \geq 0$ para nós não particulariza em nada o problema que propusemos resolver, i.e., obter uma direcção viável para o problema mestre (2), pois no

problema linear III.3.44 de obtenção da direção viável sempre ocorre $b \geq 0$. Como em III.3.44, não se incluem em (17) restrições de não-negatividade. Nessas condições, consideremos o problema modificado de (17) pela mudança de variável a seguir :

$$18 \quad x = x_2 - x_1 \quad x_1 \geq 0 \quad x_2 \geq 0$$

De (18) em (17), vem :

$$\underset{x_1, x_2}{\text{Maximizar}} \quad c^t (x_2 - x_1) \quad \text{s.a.} \quad M(x_2 - x_1) \leq b \quad x_1 \geq 0 \quad x_2 \geq 0$$

ou, equivalentemente,

$$19 \quad \underset{x_1, x_2}{\text{Maximizar}} \quad [c^t \quad -c^t] \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \quad \text{s.a.} \quad [M \quad -M] \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \leq b$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \geq 0$$

Consideremos a variável σ ,

$$20 \quad \sigma = (x_2, x_1)$$

obtendo-se de (19)

$$21 \quad \underset{\sigma}{\text{Maximizar}} \quad [c^t \quad -c^t] \sigma \quad \text{s.a.} \quad \sigma \geq 0 \quad [M \quad -M] \sigma \leq b$$

O problema (21) equivalente a (17) (pois as mudanças de variáveis obedecem a relações lineares), tem suas soluções ligadas pelas transformações (18) e (20). O problema (21) pode então ser resolvido via Simplex Re-

visado, obtendo-se daí a solução ótima de seu equivalente, (17).

Solução do problema (15).

A subrotina PROLIN efetua a partir de (17) as mudanças de variáveis recaindo no problema (21), o qual é resolvido pelo método Simplex Revisado. Obtida a solução de (21), a própria subrotina PROLIN deriva a solução ótima de (17).

Parâmetros da Subrotina PROLIN.

Descrevamos inicialmente os parâmetros que são os dados a serem lidos no programa principal.

Dados Lidos. Suponhamos ter o problema (17). A matriz A que serve de dados para subrotina PROLIN é dada abaixo :

22

$A =$	<table border="1"> <tr> <td>0</td><td>c_1</td><td>\dots</td><td>c_n</td><td>←</td><td>coeficientes de custo</td></tr> <tr> <td>b_1</td><td colspan="3" style="text-align: center;">- M</td><td></td><td></td></tr> <tr> <td>\vdots</td><td colspan="3"></td><td></td><td></td></tr> <tr> <td>b_m</td><td colspan="3"></td><td></td><td></td></tr> </table>	0	c_1	\dots	c_n	←	coeficientes de custo	b_1	- M					\vdots						b_m					
0	c_1	\dots	c_n	←	coeficientes de custo																				
b_1	- M																								
\vdots																									
b_m																									
			→ componentes do vetor b.																						

Lembramos que A deve ser lida na forma de lista coluna após coluna. A primeira posição de A, A(1), terá inicialmente o valor nulo.

Além de A serão lidos os parâmetros :

N - número de colunas da matriz M;

M - número de linhas da matriz M;
 PLA - passo da linha da matriz A;
 PCA - passo da coluna da matriz A;
 EPL - zero computacional.

Como para A usamos a forma de lista coluna após coluna, vem:

$$\begin{aligned} \text{PLA} &= 1; \\ \text{PCA} &= M+1. \end{aligned}$$

Parâmetros. Dado que haveremos de obter a solução de (17) via (21), a subrotina PROLIN transforma a matriz A, obtendo uma nova matriz A dada abaixo :

$$23 \quad A = \left[\begin{array}{c|cccc|cccc} 0 & c_1 & \cdots & c_n & -c_1 & \cdots & -c_n \\ \hline b_1 & & & & & & \\ \vdots & & -M & & & & -M \\ b_m & & & & & & \end{array} \right]$$

correspondente ao problema (21). A subrotina PROLIN opera sobre a matriz A em (23), gerando no lugar dos coeficientes de custo, os coeficientes relativos de custo, no lugar do vetor b a solução viável, e na posição A(1), o valor da função objetivo para essa solução viável. Abaixo descrevemos os parâmetros referindo-os ao problema (21) :

J1 - dimensão da lista A;
 B - matriz $(M+1) \times M$ gerada pela subrotina PROLIN correspondentes às variáveis de folga:

$$B = \begin{bmatrix} 0 & \dots & 0 \\ & -I \end{bmatrix}$$

coeficientes de custo das variáveis de folga.

onde I é a matriz identidade $M \times M$;

JB - dimensão da lista B ;

C - lista auxiliar;

JC - dimensão da lista C ;

$IVNB$ - lista dos índices das variáveis não básicas;

$JIVNB$ - dimensão da lista $IVNB$;

IVB - lista dos índices das variáveis básicas;

$JIVB$ - dimensão da lista IVB ;

$TIPO$ - indicador do tipo de solução, i.e., $TIPO=1$ a solução é infinita; $TIPO=0$ a solução é finita.

$L1$ - lista dos índices das colunas de A não básicas;

$JL1$ - dimensão da lista $L1$;

$L10$ - tamanho da lista $L1$ numa determinada iteração;

$L2$ - lista dos índices das colunas de B não básicas;

$JL2$ - dimensão da lista $L2$;

$L20$ - tamanho da lista $L2$ numa determinada iteração;

$L3$ - lista das colunas básicas;

$JL3$ - dimensão da lista $L3$;

\emptyset - lista cujos elementos são as componentes do vetor \emptyset em (21).

$J\emptyset$ - dimensão da lista \emptyset .

Observação. Para aumentar o campo de aplicação da subrotina PROLIN, independentemente do problema que objetivamos resolver, introduzimos a possibilidade de TIPO ser igual a 1, embora com as hipóteses que impusemos nos capítulos anteriores tal caso nunca ocorra.

Dimensões. Esses parâmetros servem para determinar o espaço de memória suficiente para a resolução do problema (21). Assim sendo, para dimensionar as listas haverá de se tomar cuidado, pois as listas anteriormente descritas (exceto B), referem-se não à matriz A construída para (17), mas à matriz A construída para (21). Este fato, deve-se à maneira como opera a subrotina PROLIN: dada a matriz A para (17), a subrotina transforma-a, obtendo a matriz A para (21). As dimensões podem ser calculadas pelas expressões abaixo:

$$\begin{aligned} J1 &= (2N+1)(M+1); \\ JB &= (M+1)(M+1); \\ JC &= (M+1); \\ JIVNB &= 2N; \\ JIVB &= M; \\ JL1 &= 2N; \\ JL2 &= M; \\ JL3 &= M; \\ J\emptyset &= 2N. \end{aligned}$$

Observação. Lembramos que essas dimensões correspondem às dimensões máximas das listas cujo tamanho varie durante o emprego da subrotina (ex.: IVNB, Ll, etc).

Saídas da Subrotina PROLIN .

Os parâmetros TIPO, a lista \emptyset e a lista A, contém os valores de saída da subrotina PROLIN que nos interessam.

TIPO - se o problema linear (17) tiver valor ótimo $+\infty$ então em TIPO ficará o valor 1; caso contrário, TIPO = 0;

\emptyset - a solução do problema (21) se existir (i.e., TIPO=0), é transferida da lista A (primeira coluna) para a lista \emptyset . Essa lista é então transformada segundo (18) e (20), obtendo-se a solução de (17). Assim sendo, cada elemento da lista terá o valor correspondente das componentes da solução ótima para (17). Por essa razão, somente os primeiros N elementos de \emptyset terão interesse. Lembremos que essa lista fora dimensionada para 2N elementos devido à dimensão do problema (21);

VALOR ÓTIMO - se a solução for finita esse valor ficará disponível em A(1).

NOVIN.

Essa subrotina atua como auxiliar da PROLIN, determinando em cada iteração a inversa da matriz da base atual, a solução viável atual e o valor da função objetivo para essa solução.

Parâmetros da Subrotina NOVIN.

Para cada uma das funções mencionadas acima, damos a significação dos parâmetros da subrotina NOVIN.

24 1. Na transformação da inversa da matriz da base :

B - inversa da matriz da base;
JB - dimensão da lista B;
C - lista auxiliar mencionada em PROLIN;
JC - dimensão da lista C;
M - parâmetro mencionado em PROLIN;
IP - índice da linha pivô de A para (21);
KS - parâmetro auxiliar; nesse caso, vale M;
PLB - passo da linha da matriz B;
PCB - passo da coluna de B.

Saída da Subrotina NOVIN para (24).

A matriz inversa da base atual fica disponível em B.

25 2. Na determinação da solução viável atual e do valor da função objetivo :

B - matriz para (21);
JB - dimensão de A;
KS - neste caso, vale 1;
PCB - por conveniência feito igual a 0.

Os demais parâmetros, C, JC, M, PLB, IP, tem o significado exposto em (24)

Saída da Subrotina NOVIN para 25.

Nesse caso, a subrotina transforma diretamente a primeira coluna de A, gerando em A(1), o valor da função objetivo para a solução viável atual, a qual ocupará as demais posições da primeira coluna.

DETMAX.

Utilizada na determinação do maior elemento da linha de índice igual a LCR da matriz A.

Essa subrotina é utilizada na PROLIN para determinar em cada iteração o maior dos coeficientes relativos de custo, que se encontram colocados na primeira linha de A, ou seja determinar a coluna pivô numa determinada iteração.

Parâmetros da Subrotina DETMAX.

Possuem a mesma significação dos parâmetros apresentados em PROLIN, adicionando-se :

LCR - linha dos coeficiente relativos de custo (no presente caso LCR=0);

KP - índice da coluna de A que possui o maior coeficiente relativo de custo (coluna pivô);

MAX - valor do maior coeficiente relativo de custo.

Saídas da Subrotina DETMAX.

As saídas constam dos valores em MAX e KP.

Antes de descrevermos as duas últimas subrotinas satélites, consideremos algumas listas e parâmetros que entram na composição da subrotina principal, e que serão utilizados pelas subrotinas que posteriormente descreveremos.

26 X - lista com KS (número de subsistemas) partições; cada partição terá N(S) elementos correspondentes às componentes do vetor-atividade associado ao subsistema S.

De posse da lista N, poderemos para qualquer subsistema determinar, através da lista X, os valores das componentes do vetor atividade associado a esse subsistema.

27 Y - matriz cuja S-ésima coluna contém os valores das componentes do vetor variável do problema mestre associado ao subsistema S. Essa matriz terá, então, MR linhas e KS colunas, ou seja, será uma lista tendo KS partições do mesmo tamanho;

28 IATG - lista que contém KS partições. A S-ésima partição possui, em ordem crescente, os índices dos vínculos g_{ij} ativos para o subsistema. Os vínculos $g_{ij} \geq 0$, por hipótese, nunca serão ativos;

29 IALG - lista com KS partições; a S-ésima partição possui, em ordem crescente, os índices das linhas de S-ésima submatriz em AA, correspondentes aos vínculos g_{ij} cujos índices a

parecem na S-ésima partição de IATG; essa lista, é necessária por causa da exclusão dos coeficientes das $g_{ij} \geq 0$ das submatrizes em AA;

30 KIATG - lista de tamanho KS; KIATG(S) é o tamanho da S-ésima partição da lista IATG ou IALG;

31 IATH - lista com KS partições; a S-ésima partição possui, em ordem crescente, os índices dos vínculos h_{ij} ativos, para o subsistema S;

32 KIATH - lista de tamanho KS; KIATH(S) é o tamanho da S-ésima partição da lista IATH.

GRAGH.

Utilizada para o cálculo dos gradientes das funções g_{ij} e h_{ij} ativas, para cada subsistema.

Parâmetros.

S - número de referência do subsistema;

ALFA - lista dos coeficientes das componentes do vetor atividade dos vínculos g_{ij} (ou h_{ij});

JALFA - dimensão da lista ALFA;

BETA - lista dos expoentes das componentes do vetor atividade dos vínculos g_{ij} (ou h_{ij});

JBETA - dimensão da lista BETA ;

33 GR - matriz cujas linhas são constituídas pelas componentes dos gradientes dos vínculos g_{ij} (ou h_{ij}) ativos; para cada subsistema essa matriz será KIATG(S) \times N(S) (ou KIATH(S) \times N(S));

JGR - dimensão da lista GR;

PLGR - passo da linha da matriz GR;

34

PCGR - passo da coluna da matriz GR;
 X - lista com significação já mencionada;
 JX - dimensão da lista X;
 N - idem X;
 JN - dimensão da lista N;
 IND - S-ésima partição da lista IALG (IATH para h_{ij});
 KIAT - dimensão da lista IND, i.e., número de vínculos g_{ij} (ou h_{ij}) ativos;
 PCC - passo da coluna da S-ésima submatriz da lista AA (ou CC para h_{ij});
 ULT - índice do último elemento da (S-1)-ésima submatriz da lista AA (CC para h_{ij}).

Dimensões das Listas da Subrotina GRAGH.

O dimensionamento das listas para o cálculo dos gradientes dos vínculos g_{ij} ativos, é facilmente obtido, resultando,

$$\begin{aligned}
 JALFA &= \sum_{S=1}^s N(S) \cdot L(S); \\
 JBETA &= JALFA; \\
 JN &= \sum_{K=1}^s N(K); \\
 JGR &= KIATG(S) \cdot N(S); \\
 JX &= \sum_{K=1}^s N(K); \\
 KIAT &= KIATG(S).
 \end{aligned}$$

Notamos, portanto, que exceto para a lista IND, que se obtém transferindo os índices da S-ésima partição de IALG, dimensionamos as demais listas o suficiente para incluir os dados necessários ao cálculo dos gradientes

para o subsistema S.

Saída da Subrotina GRAGH.

Quando utilizada para o subsistema S, a saída consta da matriz GR (na forma de lista) com KIATG(S) linhas e N(S) colunas, sendo cada linha constituída pelas componentes dos gradientes dos vínculos g_{ij} (ou h_{ij}) ativos, cujos índices j estão na S-ésima partição da lista IALG.

GRAF.

Utilizada no cálculo das funções objetivo e das repectivos gradientes, para cada subsistema.

Parâmetros.

S - número de referência do subsistema;

ULTF - lista cuja significação demos em (12);

JULTF - dimensão da lista ULTF;

ALFA - lista dos coeficientes das componentes do vetor atividade associado ao subsistema S;

JALFA - dimensão da lista ALFA;

BETA - lista dos expoentes das componentes do vetor atividade associado ao subsistema S;

JBETA - dimensão da lista BETA;

N - significação já aludida;

JN - idem N;

X - idem N;

JX - idem N;

GRF - lista cujos elementos são as componentes do gradiente da função objetivo do subsistema

36

S;

JGRF - dimensão da lista GRF;

F - valor da função objetivo do subsistema S;

R - lista auxiliar;

JR - dimensão da lista R;

MF - indica o tipo de função, i.e., MF=1 a função objetivo tem termos polinomiais e exponenciais; MF=-1 a função objetivo tem somente a parte polinomial e MF=0 a função objetivo tem somente a parte exponencial.

Dimensões das listas da Subrotina GRF.

JULTF = S + 1;

JALFA = ULTF(S + 1);

JBETA = ULTF(S + 1) - S;

JN = KS;

JX = $\sum_{k=1}^s N(k);$

JGRF = N(S);

JR = N(S).

Portanto, para as listas JULTF, JALFA, JBETA e JX, dimensionamos o suficiente para incluir os dados necessários ao cálculo do gradiente e da função objetivo do subsistema S.

Saídas para Subrotina GRAF.

Três são os parâmetros de saída da subrotina GRAF quando empregada para o subsistema S:

F - valor da função objetivo;
 MF - tipo da função objetivo no sentido anteriormente exposto;
 GRF - lista cujos elementos são as componentes do gradiente da função objetivo.

Passemos, finalmente, à descrição da subrotina ODIVAL que obterá a direção viável para o problema mestre no ponto Y, se a mesma existir.

ODIVAL.

A subrotina ODIVAL como já dissemos, tratará de obter a direção viável de maior crescimento local para a função objetivo do problema mestre (2), através do programa linear III.3.44, além de incluir duas das modificações propostas no capítulo IV :

- ϵ -atividade para os vínculos y_{ij} ;
- procedimento da redefinição.

A inclusão da ϵ -atividade para os vínculos g_{ij} e h_{ij} é fácil e será discutida.

Iniciemos apresentando as etapas abrangidas pela subrotina principal :

- 1- Teste de Optimalidade Prévio;
- 2- Geração da Matriz D;
- 3- Resolução do Problema Linear;
- 4- Teste de Optimalidade;

5- Geração da Matriz das Direções Viáveis.

Teste de Optimalidade Prévio.

O presente teste verifica se o ponto $\bar{x} = (\bar{x}_1, \dots, \bar{x}_k)$ onde \bar{x}_i é solução ótima de $P_{\bar{y}}$ e $\bar{y} = (\bar{y}_1, \dots, \bar{y}_k)$ fora obtido da otimização unidirecional, é ótimo, baseado na não existência de vínculos g_{ij} , $i=1\dots k$ e $j=1\dots m$, ativos em \bar{x}_i , $i=1\dots k$. Se para todos subsistemas, não houver nenhum vínculo g_{ij} ativo em \bar{x}_i , então a cada restrição $g_i(x_i) \leq \bar{y}_i$ estará associado um só vetor multiplicador ótimo, o qual será nulo. Logo :

$$(\forall i \in I_k)(\forall z_i \in \mathbb{R}^m)(v_i'(\bar{y}_i; z_i) = 0) \implies (\forall z \in \mathbb{R}^{km})(\sum_{i=1}^k v_i'(\bar{y}_i; z_i) = 0).$$

deduz-se daí, a optimalidade de \bar{y} para o problema (2), e, portanto, a optimalidade de \bar{x} para (1).

Esse teste é possível com base nas informações oriundas do passo 1 do algoritmo (5), mais especificamente a lista KIATG. No caso de não existir nenhum vínculo g_{ij} ativo teremos :

$$\text{KIATG}(S) = 0 \quad S = 1 \dots KS.$$

Caso haja ao menos um vínculo g_{ij} , para algum subsistema, ativo em \bar{x}_i , passamos à geração da matriz D.

Antes porém, se desejarmos considerar ϵ -atividade para os vínculos g_{ij} e h_{ij} , deveremos dispor de uma subrotina que redefina as listas IATG, IALG, KIATG, KIATH, e, IATH. Assim, por exemplo, IATG deverá conter os índices dos vínculos g_{ij} ϵ -ativos. No presente trabalho não conside

raremos a existência de tal subrotina, e portanto, não consideraremos ϵ -atividade para os vínculos g_{ij} e h_{ij} . Como vemos a introdução de vínculos g_{ij} e h_{ij} ϵ -ativos não apresenta qualquer dificuldade.

37 Geração da Matriz D.

Essa matriz (parâmetro da subrotina ODIVAL) servirá de dado para a subrotina PROLIN. Assim sendo, deverá ser construída na forma da matriz A em (22), onde M representa a matriz do programa linear III.3.44 considerando-se vínculos y_{ij} ϵ -ativos e o procedimento de redefinição.

A própria subrotina principal encarrega-se de gerar a lista IATY, que contém, em ordem crescente, os índices j dos vínculos y_{ij} EPM-ativos, ou seja, os índices j tais que :

$$EPM \geq b_j - \sum_{i=1}^k \bar{y}_{ij} \geq 0$$

onde EPM (parâmetro da subrotina ODIVAL) deverá ser convenientemente fixado. Para tais valores de j, i.e., $j \in IATY$, serão consideradas as restrições $\sum_{i=1}^k z_{ij} \leq 0$, fixando-se os z_{ij} dessas restrições no valor -1, caso o correspondente vínculo g_{ij} seja não ativo (ou ϵ -ativo se considerarmos vínculos g_{ij} ϵ -ativos) em \bar{x}_i . Nesse caso, após fixados os devidos z_{ij} em -1, se existirem, a restrição $\sum_{i=1}^k z_{ij} \leq 0$ passará a ter o segundo membro não nulo. A lista CONTA (parâmetro da subrotina ODIVAL) terá para cada índice j em IATY, o valor desse segundo membro.

Portanto, considerando-se essas variações nas restrições de III.3.44, geramos a matriz D, dado para a

subrotina PROLIN que resolverá o problema linear.

Resolução do Problema Linear.

O problema linear resolvido para a matriz D anteriormente construída, tem seu valor ótimo testado conforme a etapa seguinte.

38 Teste de Optimalidade para o Programa Mestre.

Se o valor ótimo do programa linear for menor que uma quantidade EPO (parâmetro da subrotina ODIVAL) convenientemente definida, então testamos o valor da quantidade EPM ; se $EPM = 0$ temos \bar{y} é ótimo em (2) e \bar{x} é ótimo em (1); caso contrário, fazemos $EPM=0$ e geramos novamente a matriz D . Por outro lado, se o valor ótimo for maior que EPO , passamos à geração da matriz das direções viáveis.

Essa forma de manusear a ϵ -atividade para os vínculos y_{ij} é simplificada, podendo-se optar por outros procedimentos (vide [29]).

39 Geração da Matriz das Direções Viáveis.

A matriz Z das direções viáveis (parâmetro da subrotina ODIVAL) é gerada em várias etapas.

Essa matriz terá KS colunas, sendo cada coluna constituída pelas componentes da direção viável associada ao correspondente subsistema.

Inicialmente colocamos em Z os valores z_{ij} calculados pela subrotina PROLIN.

A seguir, recalculamos os z_{ij} fixados inicialmente em -1 pelo procedimento de redefinição (capítulo IV), se os

mesmos existirem. Esse cálculo é baseado nos valores existentes na lista CONTA (vide 37). Ou seja, para cada elemento não nulo na lista CONTA, geramos um elemento correspondente na lista BR (parâmetro da subrotina ODIVAL), tal que

$$BR(IATY(j)) = -1. + \frac{(CONTA(j) - E1 - W)}{CONTA(j)},$$

onde E1 (parâmetro da subrotina ODIVAL) é uma quantidade convenientemente fixada e W a soma dos valores z_{ij} calculados pela PROLIN para $j \in IATY$. Elementos nulos em CONTA indicam a não existência de z_{ij} fixados no valor -1 para o correspondente j em IATY. Nesse caso, todos esses z_{ij} serão calculados pela PROLIN ou serão nulos caso $g_{ij} \neq 0$ para $j \in IATY$.

Finalmente, colocamos os z_{ij} totalmente livres no valor 1, obtendo-se a matriz das direções viáveis.

Parâmetros da Subrotina ODIVAL.

Inicialmente descreveremos o significado de cada parâmetro, calculando-se posteriormente as dimensões das listas. O usuário de posse das expressões das dimensões das listas presentes na subrotina, e, do significado de cada parâmetro, poderá utilizá-la sem maiores dificuldades.

Entre os parâmetros destacaremos os dados a serem lidos no programa principal, as saídas e as dimensões das listas.

Para os parâmetros que possuem significação já mencionada em outras subrotinas, daremos simplesmente o número da referência onde poderão ser revistos.

Dados Lidos.

AA - (8)
BB - (8)
CC - (8)
DD - (8)
EE - (13)
GG - (13)
X - (26)
Y - (27)
IALG - (29)
KIATG - (30)
IATG - (28)
IATH - (31)
KIATH - (32)
IGPER - (8)
N - (8)
L - (8)
M - (8)
ULTF - (13)
KS - (8)
MR - (8)
EPL - (vide PROLIN)
EPM - (37)
EPO - (38)
El - (39)
RC - lista cujos elementos são as componentes do
vetor-recursos b.

Parâmetros. Os parâmetro abaixo entram no cálculo da direção viável para o problema mestre (2), sendo gerados durante o

processamento :

CONTA - (37)
 BR - (39)
 D - (37)
 GR - (33)
 GRF - (35)
 R - (36)
 B - vide PROLIN
 C - vide PROLIN
 Ø - vide PROLIN
 L1 - vide PROLIN
 L2 - vide PROLIN
 L3 - vide PROLIN
 IVNB - vide PROLIN
 IVB - vide PROLIN
 IATY - (37)
 IND - (34)

Saídas da Subrotina ODIVAL.

Na subrotina ODIVAL existem três pontos de saída :

40 - o teste de optimalidade prévio é positivo, ou
 seja, \bar{y} é ótimo em (2) e \bar{x} é ótimo em (1);
 41 - o teste de optimalidade é positivo, i.e., \bar{y} é
 ótimo em (2) e \bar{x} é ótimo em (1);
 42 - não ocorrem (40) e (41).

No caso (42) gera-se a matriz das direções viá-
veis que fica disponível em Z .

Dimensões das Listas da Subrotina ODIVAL.

As expressões abaixo indicam como se poderá calcular as dimensões das listas presentes na subrotina ODIVAL. Lembramos que essas dimensões correspondem aos valores máximos quando o número de elementos das listas, depender dos resultados do processamento.

Antes de apresentarmos as expressões sejam as definições de NLD_{MAX} e NCD_{MAX} :

$$NLD_{MAX} = \sum_{S=1}^{KS} M(S) + 3 \cdot \sum_{S=1}^{KS} L(S) + MR + 1$$

$$NCD_{MAX} = \sum_{S=1}^{KS} N(S) + \sum_{S=1}^{KS} L(S) + 1$$

Temos :

$$JAA = \sum_{S=1}^{KS} L(S) \cdot N(S) .$$

$$JBB = JAA$$

$$JCC = \sum_{S=1}^{KS} M(S) \cdot N(S)$$

$$JDD = JCC$$

$$JEE = \sum_{S \in I} (2N(S) + 1) + \sum_{S \notin I} (N(S) + 1)$$

$$JGG = JEE - KS$$

$$JX = \sum_{S=1}^{KS} N(S)$$

JY = KS.MR

JRC = MR

JCONTA = MR

JZ = JY

JBR = MR

JD = $\lceil 2(NCD_{MAX} - 1) + 1 \rceil \cdot NLD_{MAX}$

JGR = $\max_{S=1 \dots KS} \{ N(S) \cdot \max \{ L(S), M(S) \} \}$

JGRF = $\max_{S=1 \dots KS} \{ N(S) \}$

JR = JGRF

JB = $2NLD_{MAX}$

JC = NLD_{MAX}

JØ = $2(NCD_{MAX} - 1)$

JIATG = $\sum_{S=1}^{KS} L(S)$

JIALG = JIATG

JKIATG = KS

$$JIATH = \sum_{S=1}^{KS} M(S)$$

$$JKIATH = KS$$

$$JIATY = MR$$

$$JIGPER = MR \cdot KS - \sum_{S=1}^{KS} I(S)$$

$$IND = \max_{S=1 \dots KS} \{L(S), M(S)\}$$

$$JN = KS$$

$$JM = KS$$

$$JL = KS$$

$$JULTF = KS + 1$$

$$JIVNB = 2(NCD_{MAX} - 1)$$

$$JIVB = (NLD_{MAX} - 1)$$

$$JL1 = 2(NCD_{MAX} - 1)$$

$$JL2 = (NLD_{MAX} - 1)$$

$$JL3 = (NLD_{MAX} - 1).$$

SEÇÃO - EXEMPLO.

A presente secção terá por intuito testar o funcionamento da subrotina ODIVAL, cuja listagem é apresentada no apêndice B.

No exemplo que abaixo apresentaremos, as funções dadas- funções objetivo e funções vínculo -, possivelmente só cumprem com as hipóteses de diferenciabilidade.

O exemplo proposto, cujas funções possivelmente não são côncavas ou convexas e nem tão pouco os pontos $\bar{x}_i \in \mathbb{R}^n$ soluções ótimas para os subproblemas $P_{\bar{y}_i}$, tem por finalidade testar o funcionamento da subrotina ODIVAL. Para isso, é suficiente verificar se cada etapa é perfeitamente executada, independentemente, repetimos, das hipóteses de convexidade-concavidade e optimalidade dos pontos \bar{x}_i .

Destarte, dados pontos $\bar{x}_i \in \mathbb{R}^n$, as funções e seus gradientes devem ser perfeitamente calculados. A matriz D corretamente gerada. O problema linear corretamente resolvido, possibilitará obter-se a direção viável supondo-se todas as modificações objetivadas pelo procedimento de redefinição.

A correta execução de todas as etapas, poderá ser verificada na impressão dos resultados constantes da listagem do programa para o exemplo testado, cujos dados a seguir apresentamos.

MR = 3

KS = 3

EPM = 0,100

EPO = 0,001

EPL = 0,001

E1 = 0,010

$$RC = (7,0 ; 5,0 ; 4,0)$$

$$Y = (y ; y ; y) = (3,0 ; 3,0 ; 1,5 ; 2,5 ; 1,0 ; 1,5 ; 1,5 ; 0,0 ; 1,0)$$

Dados para o subsistema 1. Seja $x_1 = (a, b)$.

$$f_1(x_1) = \frac{a^2}{2} + \frac{b^2}{2}$$

$$g_{11}(x_1) = a^2 + \frac{b^2}{2}$$

$$g_{12}(x_1) \equiv 0$$

$$g_{13}(x_1) = a^2 + b^2$$

$$h_{11}(x_1) = \frac{a^2}{2} + \frac{b^2}{2}$$

$$\bar{x}_1 = (1,0 ; 1,0)$$

Dados para o subsistema 2. Seja $x_2 = (a, b)$.

$$f_2(x_2) = \frac{a^2}{2} + b$$

$$g_{21}(x_2) = a + b$$

$$g_{22}(x_2) = a^2$$

$$g_{23}(x_2) = b^2$$

$$\bar{x}_2 = (1,0 ; 1,0)$$

Dados para o subsistema 3. Seja $x_3 = a$.

$$f_3(x_3) = a$$

$$g_{31}(x_3) = a^4$$

$$g_{32}(x_3) = 1,5 a^2$$

$$g_{33}(x_3) \equiv 0$$

Listas.

```

L = (2 ; 3 ; 2)
M = (1 ; 0 ; 0)
N = (2 ; 2 ; 1)
IGPER = (2 ; 3)
IATG = (1 ; 1 ; 3 ; 2)
IALG = (1 ; 1 ; 3 ; 2)
KIATG = (1 ; 2 ; 1)
IATH = (1).
KIATH = (1 ; 0 ; 0)
ULTF = (0 ; 3 ; 6 ; 8)

```

Dos dados para os subsistemas decorre :

```

AA = (1,0 ; 1,0 ; 0,5 ; 1,0 ; 1,0 ; 1,0 ; 0,0 ;
      1,0 ; 0,0 ; 1,0 ; 1,0 ; 1,5)
BB = (2,0 ; 2,0 ; 2,0 ; 2,0 ; 1,0 ; 2,0 ; 1,0 ;
      1,0 ; 1,0 ; 2,0 ; 4,0 ; 2,0)
CC = (0,5 ; 0,5)
DD = (2,0 ; 2,0)
EE = (0,0 ; 0,5 ; 0,5 ; 0,0 ; 0,5 ; 1,0 ; 0,0 ;
      1,0)
GG = (2,0 ; 2,0 ; 2,0 ; 1,0 ; 1,0)
X = (x1 ; x2 ; x3) = (1,0 ; 1,0 ; 1,0 ; 1,0 ; 1,0)

```

Cálculo da dimensões das listas.

Esse cálculo é facilmente feito a partir das expressões em V.2 .

$$NLD_{max} = (1+0+0) + 3(2+3+2) + 3 + 1 = 26$$

NCD ~~max~~ = (2+2+1) + (2+3+2) + 1 = 13
JAA = 2.2 3.2 2.1 = 12
JBB = 12
JCC = 1.2+ 0.2+ 0.1 = 2
JDD = 2
JEE = (2+ 1) + (2+1) + (1+1) = 8
JGG = 8-3 = 5
JX = 2+ 2+ 1 = 5
JY = 3.3 = 9
JRC = 3
JCONTA = 3
JZ = 9
JBR = 3
JD = (2(13-1)+1).26 = 650
JGR = 6
JGRF = 2
JR = 2
JB = 26.26 = 676 (fixado em 256)
JC = 26
JO = 2(13-1) = 24
JIATG = 7 (fixado em 4)
JIALG = 7 (fixado em 4)
JIATH = 1
JKIATH = 3
IND = 3
JN = 3
JM = 3
JL = 3
JULTF = 4
JIVNB = 2(13-1) = 24

JIVB = (26-1) = 25

JL1 = 24

JL2 = 25

JL3 = 25

De posse desses resultados a subrotina ODIVAL pode ser utilizada na obtenção da direção viável, a qual para o exemplo proposto poderá ser observada na listagem dos resultados constantes do apêndice B .

C A P I T U L O VI

CONCLUSÕES

O algoritmo estudado apresenta particular importância quando o número de restrições que definem os conjuntos $X_i \subset \mathbb{R}^{n_i}$, é grande em relação à dimensão m do vetor-recurso. Tal fato, liga-se a aplicação do método de direções viáveis ao problema mestre, com consequente decentralização do trabalho computacional. Para que se possa assim proceder, importante resultado fora demonstrado, o qual deduz a equivalência entre os problemas mestre e original, exceto pela diferenciabilidade de suas funções objetivo. Além disso, tendo-se a solução ótima do problema mestre, pode-se imediatamente dispor da solução do problema original.

Na otimização de sistemas não-lineares de grande porte decomponíveis, o procedimento iterativo proposto pela coordenação primal, pode ser o caminho decisivo para resolução de tais problemas, já que reduz o trabalho computacional à resolução de uma sequencia de subproblemas menores.

No curso de toda dissertação, papel de importância básica é dado às funções perturbação, cuja não diferenciabilidade em \mathbb{R}^m seu domínio Y_i , leva à obtenção de um algoritmo cujas propriedades de convergência ainda são problemas em aberto. Notemos dessa forma que a decentralização

computacional conduz a uma interessante pesquisa, qual seja, a determinação de condições que levam à convergência do algoritmo estudado.

Mostramos que a diferenciabilidade local de algumas funções perturbação, pode ser utilizada com vantagem computacional, diminuindo não somente o trabalho computacional na resolução do problema linear de obtenção da direção viável, mas também o espaço de memória requerido.

Visando ainda tornar viável o emprego da diferenciabilidade local quando ela ocorre, desenvolveu-se, quando os subproblemas $P_{\bar{y}_i}$ são resolvidos por um método de direções viáveis, um critério de detecção dessa propriedade, baseado somente nas informações oriundas da resolução do max-min no ponto de ótimo \bar{x}_i para o subproblema $P_{\bar{y}_i}$.

Como resultado computacional, elaborou-se uma subrotina em FORTRAN IV a qual obtém a direção viável para cada iteração do algoritmo.

Queremos finalmente ressaltar que dois estudos merecem atenção, quando o intuito é a prova da convergência do algoritmo :

- introdução e estudo da ϵ -atividade para os vínculos g_{ij} constantes no problema diretor;
- estudo do comportamento das funções perturbação na fronteira de seu domínio.

A P E N D I C E A

DEFINIÇÕES E UM TEOREMA PARA FUNÇÕES CONCAVAS

Apresentamos a seguir definições largamente utilizadas no presente trabalho, as quais podem ser encontradas em Rockafellar [32] ou Geoffrion [12], onde são detalhadamente analisadas.

Um teorema para obtenção da derivada direcional de uma função côncava é enunciado, cuja demonstração pode ser encontrada em Rockafellar [32].

Definições.

Seja $f : X \rightarrow \mathbb{R}$, onde f é uma função côncava em $X \subset \mathbb{R}^n$ convexo e $g : X \rightarrow \mathbb{R}^m$ é convexa em X .

Definição- Seja o problema primal na forma canônica :

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

O problema (1) é definido como problema não perturbado e denotado por P_0 .

Definição- Definimos problema primal perturbado ao problema :

$$2 \quad \text{Maximizar} \quad f(x) \quad \text{s.a.} \quad g(x) \leq y \\ x \in X$$

onde y é qualquer vetor no \mathbb{R}^m , sendo definido como vetor-perturbação associado ao vínculo g de P_0 . O problema (2) é denotado por P_y .

Definição- Dado o problema (1), a função perturbação v associada ao vínculo g de P_0 , é definida por:

$$3 \quad v(y) \triangleq \sup \{ f(x) \mid x \in X \wedge g(x) \leq y \}$$

onde $v : Y \rightarrow \mathbb{R}$, sendo

$$Y \triangleq \{ y \in \mathbb{R}^m \mid (\exists x \in X) (g(x) \leq y) \}$$

Observação : se considerarmos que o supremo de uma função num conjunto vazio é $-\infty$, poderemos definir v no \mathbb{R}^m , estendendo o espaço do contradomínio para $\mathbb{R} \cup \{-\infty, +\infty\}$.

Definição- O valor ótimo do problema (1) é definido como

$$\sup \{ f(x) \mid x \in X \wedge g(x) \leq 0 \}$$

Admitindo que $\sup \{ f(x) \mid x \in \emptyset \} \triangleq -\infty$, temos que o problema

(1) sempre terá valor ótimo sendo possivelmente $-\infty$ ou $+\infty$.

Definição- Suponhamos que $f(\cdot)$ seja finita para $x = \bar{x}$. Diremos que um vetor $\gamma \in \mathbb{R}^n$ é supergradiente de f no ponto \bar{x} se :

$$(\forall x \in X) (f(x) \leq f(\bar{x}) + \gamma^t (x - \bar{x})).$$

Definição- Seja $\bar{x} \in \text{Int}X$. Definimos derivada direcional de f em \bar{x} e na direção $h \in \mathbb{R}^n$, ao valor do limite :

$$\lim_{\theta \rightarrow 0} \frac{f(\bar{x} + \theta h) - f(\bar{x})}{\theta}$$

se o mesmo existir. Caso o limite exista denotamos seu valor por $f'(\bar{x}; h)$.

Teorema.

Seja $\bar{x} \in \text{Int}X$, sendo $f(\bar{x})$ finito. Então

$$(\forall h \in \mathbb{R}^n) (f'(\bar{x}; h) = \min \{ p^t h \mid p \text{ é supergradiente de } f \text{ em } \bar{x} \}).$$

A P E N D I C E B

LISTAGEM DOS RESULTADOS COMPUTACIONAIS.

C
C
C
C
C

PROGRAMA PRINCIPAL

C

DIMENSION AA(12),BR(12),CC(2),DD(2),EE(8),GG(5),
1X(5),Y(9),RC(3),CONTA(3),Z(9),BR(3),D(650),GR(6),
2GRF(2),GH(2),B(256),C(26),D(24),R(2)

C
C

LETTURA DE DADOS

READ (8,1001) (IATG(I),I=1,4),(IALG(I),I=1,4),
1(KIATG(I),I=1,3),IATH(1),(KIATH(I),I=1,3),(L(I),I=1,3),
2(M(I),I=1,3),(N(I),I=1,3),(IGPER(I),I=1,2),KS,MR
READ (8,1002) (ULTF(I),I=1,4)

1002 FORMAT (4I2)

READ (8,1003) (RC(I),I=1,3),(Y(I),I=1,9),(X(I),I=1,5),
1(AA(I),I=1,12),(RB(I),I=1,12),(CC(I),I=1,2),(DD(I),I=1,2),
2(EE(I),I=1,8),(GG(I),I=1,5)
READ (8,1005) EPM,EPO,E1,EPL

1001 FORMAT (8D11)

1003 FORMAT (26F3.1)

1005 FORMAT (4F5.3)

C
C

IMPRESSAO DOS DADOS

WRITE (5,1004)

1004 FORMAT ('1',28X,' IMPRESSAO DOS DADOS ')
WRITE (5,93) (IATG(I),I=1,4),(IALG(I),I=1,4),
1(KIATG(I),I=1,3),IATH(1),(KIATH(I),I=1,3),(L(I),I=1,3),
2(M(I),I=1,3),(N(I),I=1,3),(IGPER(I),I=1,2),KS,MR
93 FORMAT (////15X,16I2))

WRITE (5,97) (ULTF(I),I=1,4)

97 FORMAT (///5X,4I3)

WRITE (5,94) EPM,EPO,E1,EPL

94 FORMAT (///4(5X,F5.3))

WRITE (5,95) (RC(I),I=1,3),(Y(I),I=1,9),(X(I),I=1,5),

```
1(AA(I),I=1,12),(BB(I),I=1,12),(CC(I),I=1,2),(DD(I),I=1,2),  
2(EE(I),I=1,8),(GG(I),I=1,5)  
95 FORMAT (///(5X,15F4.1))
```

```
C  
1 CALL DDIVAL (AA,12,BR,12,CC,2,DD,2,EE,8,GG,5,  
1 X,5,Y,9,RC,3,CCNTA,3,Z,9,BR,3,  
2 D,650,GR,6,GRF,2,R,2,B,256,C,26,0,  
3 24,IALG,4,KIATG,3,IATG,4,IATH,1,  
4 KIATH,3,IATY,3,IGPER,2,IND,3,N,3,  
5 L,3,M,3,ULTF,4,IVNB,24,IVB,25,L1,  
6 24,L2,25,L3,25,KS,MR,EPM,EPO,E1,EPL)
```

```
C  
CALL EXIT  
END
```

卷之三

* * *

* A SUBROTINA ODIVAL *

— 1 —

A decorative border consisting of a repeating pattern of small black stars on a white background.

*

* OBTENCAO DA DIRECAO VIAVEL PARA O PROBLEMA

* GUARANÁ, FORTÉ, DE ESTROGOFIA, BACON, CHIPIPE E DE
* FUNGOS, PRATICAMENTE NÃO DIFERENTIAVAM.

• FIGURAS GEOMÉTRICAS

* ETAPAS DA SUBROTINA

*

*-TESTE-DE-OTIMALIDADE

* GERACAO DA MATRIZ D

* RESOLUCAO DO PROBLEMA

* TESTE DE OTIMALIDADE

* GERACAO DA MATRIZ DA

```
1      SUBROUTINE ODIVAL(AA,JAA,BB,JBB,CC,JCC,DD,JDD,EE,JEE,
2                           GG,JGG,X,JX,Y,JY,RC,JRC,CONTA,JCONTA,
3                           Z,JZ,BR,JBR,D,JD,GR,JGR,GRF,JGRF,R,JR,
4                           B,JB,C,JC,O,JD,IALG,JIALG,KIATG,JKIATG,
5                           IATG,JIATG,IATH,JIATH,KIATH,JKIATH,IATY,
6                           JIATY,IGPER,JIGPER,IND,JIND,N,JN,L,JL,
7                           M,JM,ULTF,JULTF,IVNB,JIVNB,IVB,JIVB,
8                           L1,JL1,L2,JL2,L3,JL3,KS,MR,EPM,EPO,E1,
9                           EPI)
```

```

DIMENSION AA(JAA),BB(JBB),CC(JCC),DD(JDD),EE(JEE),
1   GG(JGG),X(JX),Y(JY),RC(JRC),CONTA(JCONTA),
2   Z(JZ),BR(JBR),D(JD),GR(JGR),GRF(JGRF),
3   R(JR),B(JB),C(JC),C(JC)

```

```

C   INTEGER IALG(JIALG),KIATG(JKIATG),IATG(JIATG),IATH(JIATH),
1   KIATH(JKIATH),IATY(JIATY),IGPER(JIGPER),IND(JIND),
2   N(JN),L(JL),M(JM),ULTF(JULTF),IVNB(JIVNB),
3   IVB(JIVB),L1(JL1),L2(JL2),L3(JL3),S,PLD,PCD,PLZ,
4   PCZ,PLGR,PCGR,PLY,PCY,TIPO,UTIAG,UTIAH,ULT

```

C

C

C

C

```

C   . . . . .
C   . TESTE DE OTIMALIDADE PREVIO .
C   .

```

```

C   . VERIFICAR SE TODOS KIATG(S) SAO NULOS .
C   .

```

C

C

```

DO 1 S=1,KS

```

```
IF(KIATG(S).EQ.0) GO TO 1

```

```
GO TO 9

```

```
1 CONTINUE

```

```
RETURN

```

C

C

C

```

C   . . . . .
C   . GERACAO DA MATRIZ D .
C   .

```

```

C   . DIMENSIONAMENTO DA MATRIZ D EXCETO ACOPLAGEMTO .

```

```
   . CALCULO DO NUMERO DE LINHAS/NLD/

```

```
   . CALCULO DO NUMERO DE COLUNAS/NCD/

```

```
   . CALCULO DO NUMERO DE ELEMENTOS/ND/

```

C

C

```

C   . . . . .
C   9 NLD=1

```

```
   NCD=1

```

```
   DO 10 S=1,KS

```

```
   IF(KIATG(S).EQ.0) GO TO 10

```

```
   NLD=NLD+3*KIATG(S)+KIATH(S)

```

```
   NCD=NCD+KIATG(S)+N(S)

```

```
10 CONTINUE

```

```
83 ND=NLD*NCD

```



```
I2=0
J2=0
UTIAG=0
UTIAH=0
C
C      PASSO DA LINHA DA MATRIZ D
C
C      PLD=1
C
C      PASSO DA COLUNA DA MATRIZ D
C
C      PCD=NLD+KIATY
C
C      GERACAO DOS ELEMENTOS NAO NULOS DA MATRIZ D
C
86 FORMAT (1H1//5X,15)
DO 12 S=1,KS
J01=KIATG(S)
J02=N(S)
J03=KIATH(S)
IF(J01.EQ.0) GO TO 6
DO 13 K=1,J01
J=J2+K
I=I2+K
KH1=I*PLD+1
KH2=KH1+2*J01+J03
KH3=I*PLD+J*PCD+1
KH4=KH3+J01
KH5=KH4+J01+J03
C
C      CALCULO DA PRIMEIRA COLUNA DA MATRIZ D
C      CALCULO DAS SUBMATRIZES IDENTIDADE E MENOS IDENTIDADE
C      CONTIDAS NA MATRIZ D
C
D(KH1)=1.
D(KH2)=1.
D(KH3)=1.
D(KH4)=1.
13 D(KH5)=-1.
C
C      TRANSFERENCIA DOS INDICES DAS LINHAS DAS MATRIZES DOS
C      COEFICIENTES DOS VINCULOS G ATIVOS DA LISTA IATG PARA IND
C
DO 14 K=1,J01
KH6=UTIAG+K
14 IND(K)=IALG(KH6)
C
C      CALCULO DAS DIMENSÕES DAS LISTAS E MATRIZES PARA OBTENCAO
```

```
C      DOS GRADIENTES DOS VINCULOS G ATIVOS PELA SUBROTINA GRAGH
C
C      JAA=0
C      JX=0
C      DO 8 K=1,S
C      JAA=JAA+N(K)*L(K)
C      8 JX=JX+N(K)
C
C      CALCULO DOS GRADIENTES DOS VINCULOS G
C      ATIVOS PELA SUBROTINA GRAGH
C
C      CALL GRAGH (S,AA,JAA,BB,JAA,JAA-J02*L(S),PLGR,PCGR,N,
C      1           KS,L(S),IND,J01,X,JX,GR,J01*J02)
C
C      COLOCACAO DOS GRADIENTES DAS G ATIVAS NA MATRIZ D
C
C      DO 15 K=1,J01
C      I=I+1
C      DO 16 J3=1,J02
C      KH7=(K-1)*PLGR+(J3-1)*PCGR+1
C      J=J+1
C      KH8=I*PLD+J*PCD+1
C      16 D(KH8)=-GR(KH7)
C      15 J=J-J02
C      IF (J03.EQ.0) GO TO 99
C
C      TRANSFERENCIA DOS INDICES DOS VINCULOS
C      H ATIVOS DA LISTA IATH PARA LISTA IND
C
C      DO 17 K=1,J03
C      KH6=K+UTIAH
C      17 IND(K)=IATH(KH6)
C
C      CALCULO DAS DIMENSÕES DAS LISTAS E MATRIZES PARA OBTENCAO
C      DOS GRADIENTES DOS VINCULOS H ATIVOS PELA SUBROTINA GRAGH
C
C      JCC=0
C      DO 7 K=1,S
C      7 JCC=JCC+M(K)*N(K)
C
C      CALCULO DOS GRADIENTES DOS VINCULOS H
C      ATIVOS PELA SUBROTINA GRAGH
C
C      CALL GRAGH (S,CC,JCC,DD,JCC,JCC-J02*M(S),PLGR,PCGR,N,
C      1           KS,M(S),IND,J03,X,JX,GR,J01*J02)
C
C      COLOCACAO DOS GRADIENTES DAS H ATIVAS NA MATRIZ D
```

26 IMA=1

C

C VERIFICAR SE O INDICE IATY(K) ESTA EM IATG. CASO AFIRMATIVO

C GERE O INDICE DA COLUNA/NVATA/DA SUBMATRIZ

C CORRESPONDENTE AO SUBSISTEMA S

C

NVATA=C

DO 28 J=1,J01

I=I3+J

IF (IATG(I)-IATY(K))30,29,31

29 NVATA=NVATA+1

KH21=(NLD+K-1)*PLD+(KH20+NVATA)*PCD+1

D(KH21)=-1.

CS=0.

I2=MR-L(S)

GO TO 34

30 NVATA=NVATA+1

28 CONTINUE

C

C VERIFICAR SE O INDICE IATY(K) ESTA EM IGPER

C

31 I2=MR-L(S)

IF(I2) 33,32,33

C

C EXISTE IGPER PARA SUBSISTEMA S

C

33 DO 35 J=1,I2

I=I1+J

IF(IGPER(I)-IATY(K)) 35,36,32

36 CS=0.

GO TO 34

35 CONTINUE

C

C NAO EXISTE PARA O SUBSISTEMA S A LISTA IGPER

C

32 CS=1.

34 KH20=KH20+J01+N(S)*IMA

I1=I1+I2

I3=J01+I3

25 VONTA=VONTA+CS

C

C VALOR REDEFINIDO A SER COLOCADO NA PRIMEIRA

C COLUNA DO BLOCO DE ACOPLAMENTO

C

CONTA(K)=VONTA

IF(VONTA.EQ.0.) GO TO 24

KH22=(NLD+K-1)*PLD+1

D(KH22)=VONTA

```
DO 18 K=1,J03
I=J+1
DO 19 J3=1,J02
KH7=(K-1)*PLGR+(J3-1)*PCGR+1
J=J+1
KH8=I*PLD+J*PCD+1
19 D(KH8)=-GR(KH7)
18 J=J-J02
C
C      CALCULO DO GRADIENTE DA FUNCAO OBJETIVO DO SUBSISTEMA S
C
99 CALL GRAF (S,ULTF,S+1,EE,25,GG,21,
1           N,KS,X,JX,GRF,J02,F,MF,P,N(S))
C
C      COLOCACAO DO GRADIENTE DA FUNCAO F EM D
C
DO 20 K=1,J02
J=J+1
KH9=J*PCD+1
20 D(KH9)=GRF(K)
C
C      REDEFINICAO DOS INDICADORES/I2,J2,UTIAG,UTIAH/
C
I2=I2+3*J01+J03
J2=J2+J02+J01
UTIAG=UTIAG+J01
6  UTIAH=UTIAH+J03
12 CONTINUE
C
C      TESTE DE EXISTENCIA DO BLOCO DE ACOPLAGEMTO DA MATRIZ D
C
C      IF(KIATY.EQ.0)GO TO 23
C
C      GERACAO DO BLOCO DE ACOPLAGEMTO
C
DO 24 K=1,KIATY
KH20=0
J1=0
I3=0
VONTA=0.
DO 25 S=1,KS
J01=KIATG(S)
IF (J01)26,27,26
C
C      O SUBSISTEMA S NAO POSSUI SUBMATRIZES NO BLOCO DE ACOPLAGEMTO
C
27 IMA=0
GO TO 31
```

```

24 CONTINUE
  WRITE (5,96) (D(I),I=1,ND)
96 FORMAT (1H1///15X,'MATRIZ D (COLUNA APOS COLUNA)',1///(5X,4E16.7))

.
.
.
. RESOLUCAO DO PROGRAMA LINEAR .
.
.

23 CALL PROLTIN (D,ND+(NCD-1)*PCD,B,PCD*PCD,C,PCD,IVNB,
1           2*(NCD-1),IVB,PCD-1,NCD-1,PCD-1,TIPO,PLD,PCD,
2           L1,2*(NCD-1),L2,PCD-1,L3,PCD-1,O,2*(NCD-1),EPL)

.
.
.
. TESTE DE OTIMALIDADE PARA O PROGRAMA MESTRE .
.
.

IF(D(1).LT.0..OR.D(1).GT.EPC) GO TO 5
IF (EPM.EQ.0..) RETURN
EPM=0.
GO TO 9

.
.
.
. GERACAO DA MATRIZ /Z/ DAS DIRECOES VIAVEIS .
.
.

ZERACAO DA MATRIZ Z

5 NEZ=KS*MR
  DO 37 J=1,NEZ
37 Z(J)=0.

PASSOS DA LINHA E COLUNA DA MATRIZ Z

```

```

PLZ=1
PCZ=MR
C
JA=0
JB=0
C
C   COLOCACAO DOS VALORES ZIJ CALCULADOS
C   PELA SUBROTINA PROLIN NA MATRIZ Z
C
DO 38 S=1,KS
J01=KIATG(S)
IF(J01.EQ.0) GO TO 38
DO 39 J=1,J01
J1=J+JA
J2=J+JB
KH23=PLZ*(IATG(J2)-1)+PCZ*(S-1)+1
39 Z(KH23)=0(J1)
JA=J01+N(S)+JA
JB=J01+JB
38 CONTINUE
C
C   TESTE DE EXISTENCIA DA LISTA BR DOS VALORES USADOS
C   NA REDEFINICAO DOS CORRESPONDENTES ZIJ
C
IF(KIATY.EQ.0) GO TO 46
C
C   GERACAO DA LISTA BR
C
--DO 40 J=1,KIATY
IF(CONTA(J1)42,41,42
41 BR(IATY(J))=-2.
GO TO 40
C
C   CALCULO DE BR(J)
C
42 W=0.
DO 43 I=1,KS
KH25=PLZ*(IATY(J)-1)+PCZ*(I-1)+1
43 W=W+Z(KH25)
BR(IATY(J))=-1.+(CONT(I)-E1-W)/CONT(I)
IF(BR(IATY(J)).LE.1..AND.BR(IATY(J)).GE.-1..) GO TO 40
IF(1-BR(IATY(J))) 44,40,45
44 BR(IATY(J))=1.
GO TO 40
45 BR(IATY(J))=-1.
40 CONTINUE
IN=NCD-1
WRITE (5,99) (CONT(I), I=1, KIATY), (BR(I), I=1, 3, 2),

```

```

1D(1),(C(I),I=1,IN)
98 FORMAT (1H1////5X,'LISTA CONTA = (',E16.7,',',
1E16.7,',')'////5X,'LISTA BR = (',E16.7,',',E16.7,
2')'////5X,'VALOR OTIMO DO PROBLEMA DIRETOR = ',
3E16.7////5X,'LISTA O = (',3F16.7//(16X,3E16.7/))

```

```

C      CALCULC DOS ZIJ NAO OBTIDOS PELA SUBROTTINA PROLIN
C

```

```

46 JC=0
    JB=0
    DO 47 S=1,KS
        JD=MR-L(S)
        J01=KIATG(S)
        DO 48 J=1,MR
            1F(JD.EQ.0) GO TO 50

```

```

C      VERIFITCAR SE O INDICE J ESTA EM IGP
C

```

```

DO 49 J=1,JD
    I1=I+JC
    IF(IGPER(I1)-J)49,48,50
49 CONTINUE
50 IF(J01.EQ.0) GO TO 51

```

```

C      VERIFICAR SE O INDICE J ESTA NA LISTA IATG
C

```

```

DO 52 J=1,J01
    I1=I+JB
    IF(IATG(I1)-J)52,48,51
52 CONTINUE
51 IF(KIATY.EQ.0) GO TO 54

```

```

C      VERIFICAR SE O INDICE J ESTA NA LISTA IATY
C

```

```

DO 53 I=1,KIATY
    IF(IATY(I)-J) 53,56,54
56 KH24=PLZ*(J-1)+PCZ*(S-1)+1
    Z(KH24)=BR(J)
    GO TO 48
53 CONTINUE
54 KH24=PLZ*(J-1)+PCZ*(S-1)+1
    Z(KH24)=1.
48 CONTINUE
55 JC=JC+JD
    JB=JB+J01
47 CONTINUE
    WRITE (5,90) (Z(I),I=1,9)
90 FORMAT (1H1////20X,'DIRECAO VIAVEL-MATRIZ Z')

```

1////(5X,3E16.7))

RETURN

END

*

* A SUBROTINA PROLIN *

*

*

* RESOLUCAO DE UM PROBLEMA DE MAXIMIZACAO *

* EM PROGRAMACAO LINEAR SEM RESTRICOES DE *

* NAO-NEGATIVIDADE *

*

*

* PARAMETRO TIPO-INDICADOR DO TIPO DE SOLUCAO *

* TIPO=1 SOLUCAO INFINITA *

* TIPO=0 SOLUCAO FINITA *

*

*

* ETAPAS DA SUBROTINA PROLIN *

*

*

* ADAPTACAO DA MATRIZ DOS COEFICIENTES AS RESTRICOES *

* DE NAO NEGATIVIDADE *

* INTRODUCAO DAS VARIAVEIS DE FOLGA *

* DETERMINACAO DA COLUNA PIVOT *

* TESTE DE OTIMALIDADE *

* DETERMINACAO DA LINHA PIVOT *

* DETERMINACAO DA NOVA SOLUCAO VIAVEL E DO VALOR *

* DA FUNCAO OBJETIVO *

* DETERMINACAO DOS NOVOS COEFICIENTES RELATIVOS DE CUSTO *

* DETERMINACAO DA NOVA MATRIZ INVERSA *

* REORGANIZACAO DE LISTAS *

*

C
C SUBROUTINE PROLIN (A,J1,B,JB,C,JC,IVNR,IVNB,IVB,JIVB,N,M,
1 TIPO,PLA,PCA,L1,JL1,L2,JL2,L3,JL3,O,JO,EPL)

C
C DIMENSION A(J1),B(JB),C(JC),O(JO)

C
C DIMENSION L1(JL1),L2(JL2),L3(JL3)
C INTEGER IVNB(JIVNB),IVB(JIVB),TIPO,PLA,PCA
C REAL MAX

C
C
C
C . ADAPTACAO DA MATRIZ DOS COEFICIENTES .
C . AS RESTRIÇÕES DE NAO-NEGATIVIDADE .
C .
C

C
C DO 1 J=1,N
KT=M+1
DO 1 II=1,KT
I=II-1
N1=PLA*I+PCA*j+1
N2=N1+N*(M+1)
1 A(N2)=-A(N1)
N=2*N
DO 40 K=1,N
L1(K)=K
40 IVNB(K)=K
DO 2 I=1,M
2 IVB(I)=N+I
L10=N
L20=0

C
C
C
C . INTRODUÇAO DAS VARIAVEIS DE FOLGA .
C .
C

C
C M1=M+1
DO 3 I=1,M1
DO 3 K=1,M
KH=(I-1)*PLA+K*PCA+1
B(KH)=0.

```
3 IF(K.EQ.T-1)B(KH)=-1.  
C  
C  
C  
C . DETERMINACAO DA COLUNA PIVO E GERACAO DO VETOR C .  
C  
C  
C  
C  
4 MAX=0.  
Q1=0.  
IF(L10.EQ.0) GO TO 5  
CALL DETMAX (A,J1,0,PLA,PCA,KP,L1,L10,JL1,MAX)  
5 IF(L20.EQ.0) GO TO 6  
DO 7 I=1,L20  
KH=L2(I)*PCA+1  
IF(B(KH).LE.Q1) GO TO 7  
KS=I  
Q1=B(KH)  
7 CONTINUE  
C  
C  
C . TESTE DE OTIMALIDADE .  
C  
C  
C  
C  
6 IF(MAX.GT.EPL.CR.Q1.GT.EPL) GO TO 8  
C  
- - - TIP0=0 - - -  
DO 9 I=1,N  
9 O(I)=0.  
DO 10 I=1,M  
IF(IVB(I).GT.N) GO TO 10  
J=IVB(I)  
O(J)=A(PLA*I+1)  
10 CONTINUE  
K1=N/2  
DO 11 J1=1,K1  
J2=J1+K1  
11 O(J1)=C(J1)-C(J2)  
N=N/2  
RETURN  
8 K=0  
IF(MAX.GE.Q1) GO TO 12  
K=1  
KP=L2(KS)  
DO 13 II=1,M1  
KH=(II-1)*PLA+KS*PCA+1
```

```
13 C(II)=-B(KH)
GO TO 16
12 KH=KP*PCA+1
C(1)=-A(KH)
DO 14 I=1,M
Q1=0.
DO 15 J=1,M
KH1=KH+J*PLA
KH2=I*PLA+PCA*j+1
15 IF(A(KH1).NE.0..AND.B(KH2).NE.0..) Q1=Q1+A(KH1)*B(KH2)
14 C(I+1)=Q1
```

C
C
C
C
C
C
C
C

```
.....  
.  
DETERMINACAO DA LINHA PIVO  
TRANSFORMACAO DO VETOR C  
CALCULO DE TATAO  
.  
.....
```

```
16 IP=0
DO 17 I=1,M
IF(C(I+1).GT.0..) GO TO 18
17 CONTINUE
TIPO=1
RETURN
18 KH=I*PLA+1
Q1=A(KH)/C(I+1)
IP=I
20 J=IP+1
DO 21 I=J,M
IF(C(I+1).LE.0..) GO TO 21
KH=I*PLA+1
IF(A(KH)/C(I+1).GE.Q1) GO TO 21
Q1=A(KH)/C(I+1)
IP=I
21 CONTINUE
C(IP+1)=1./C(IP+1)
DO 22 II=1,M1
22 IF(II-1.NE.IP)C(II)=-C(II)*C(IP+1)
```

C
C
C
C
C
C
C
C

```
.....  
.  
TRANSFORMACAO DA PRIMEIRA COLUNA DE A  
NOVA SOLUCAO VIAVEL E NOVO VALOR DA FUNCAO OBJETIVO  
.  
.....
```

```
CALL NCVIN (A,J1,C,JC,M,IP,1,PLA,0)
```

□□□□□

- TRANSFORMACAO DA PRIMEIRA LINHA DE A
- NOVOS COEFICIENTES RELATIVOS DE CUSTO

```

DO 23 J=1,N
IF(J.EQ.KP) GO TO 23
Q1=0.
DO 24 I=1,M
KH=IP*PLA+I*PCA+1
KHI=I*PLA+PCA*N+1
24 IF(B(KH).NE.0.) Q1=Q1+B(KH)*A(KHI)
KH=J*PCA+1
A(KH)=A(KH)-Q1*C(1)
23 CONTINUE
KH=KP*PCA+1
A(KH)=-C(1)

```

- TRANSFORMACAO DA INVERSA DA BASE

CALL-NEWIN (B,JB,C,JC,M,IP,M,PLA,PCA)

C C C C C C

```

IF(K.NE.0) GO TO 31
IF(L20.LT.1) GO TO 25
DO 26 J=1,L20
IF(IP.EQ.L2(J)) GO TO 27
CONTINUE
GO TO 25
IF(L3(J).EQ.KP) GO TO 38
IF(L10.LT.1) GO TO 29
DO 28 I=1,L10
IF(L1(I).EQ.KP) L1(I)=L3(I)
L3(J)=KP

```

```
GO TO 38
25 L20=L20+1
    L2(L20)=IP
    L3(L20)=KP
30 IF(L10.LT.1) GO TO 38
    DO 32 I=1,L10
        IF(L1(I).EQ.KP) GO TO 34
32 CONTINUE
    GO TO 38
34 L10=L10-1
    DO 33 J=I,L10
        L1(J)=L1(J+1)
    GO TO 38
31 IF(L20.LT.1) GO TO 37
    DO 35 J=1,L20
        IF(IP.EQ.L2(J)) GO TO 38
35 CONTINUE
37 L10=L10+1
    L1(L10)=L3(KS)
    L20=L20-1
    DO 36 J=KS,L20
        L2(J)=L2(J+1)
36 L3(J)=L3(J+1)
38 K=IVNB(KP)
    IVNB(KP)=IVB(IP)
    IVB(IP)=K
    GO TO 4
    END
```



```
C
C
C
C ****
C *
C * A SUBROTIINA NOVIN *
C *
C ****
C
C ****
C *
C * DETERMINACAO DA INVERSA DA BASE *
C * DO PROGRAMA LINEAR *
C * DETERMINACAO DA NOVA SOLUCAO VIABEL *
C *
C ****
```

```
SUBROUTINE NOVIN (B,JB,C,JC,M,IP,KS,PLB,PCB)
```

```
DIMENSION B(JB),C(JC)
INTEGER PLB,PCB
```

```
DO 2 K=1,KS
KH=IP*PLB+PCB*K+1
IF(B(KH).EQ.0.) GO TO 2
IM=M+1
DO 3 II=1,IM
I=II-1
IF (I.EQ.IP) GO TO 3
IF(C(II).EQ.0.) GO TO 3
KH=K*PCB+1
KHO=KH+I*PLB
KH1=KH+IP*PLB
B(KHO)=B(KHO)+B(KH1)*C(II)
3 CONTINUE
2 CONTINUE
DO 4 K=1,KS
KH=IP*PLB+PCB*K+1
4 B(KH)=B(KH)*C(IP+1)
RETURN
END
```



```
• CALCULO DA COMPONENTE J DO GRADIENTE G/H •  
•  
•.....  
2 GR(KH3)=ALFA(KH1)*BETA(KH1)*X(KH2)**(BETA(KH1)-1.)  
RETURN  
END
```


C
C
C F TEM PELÔ MENOS UM TERMO PCLINOMIAL
C

3 MF=-1
6 A=0.
 IR=N(S)
 DO 7 I=1,IR
 I1=I+IA
 I2=I1-S
 J=I+KH
 A=A+ALFA(I1)*X(J)**BETA(I2)
 GRF(I)=ALFA(I1)*BETA(I2)*X(J)**(BETA(I2)-1.)
7 CONTINUE
 IF(MF) 8,9,10
8 F=A
 RETURN
9 B=A
 A=0.
 P=ALFA(IA)*EXP(B)
 DO 11 I=1,IB
11 GRF(I)=GRF(I)*P
 GO TO 14
10 B=0.
 DO 12 I=1,TB
 I1=I1+I
 J=KH+I
 I2=I2+I
 B=B+ALFA(I1)*X(J)**BETA(I2)
 R(I)=ALFA(I1)*BETA(I2)*X(J)**(BETA(I2)-1.)
12 CONTINUE
 P=ALFA(IA)*EXP(B)
 DO 13 I=1,IB
13 GRF(I)=GRF(I)+P*R(I)
14 F=P+A
 RETURN
 END

IMPRESSAO DOS DADOS

1 1 3 2 1 1 3 2 1 2 1 1 1 0 0 2
3 2 1 0 0 2 2 1 2 3 3 3

0 3 6 8

0.100 0.001 0.010 0.001

7.0 5.0 4.0 3.0 3.0 1.5 2.5 1.0 1.5 1.5 0.0 1.0 1.0 1.0 1.0
1.0 1.0 1.0 1.0 0.5 1.0 1.0 1.0 0.0 1.0 0.0 1.0 1.0 1.0 1.5 2.0
2.0 2.0 2.0 1.0 2.0 1.0 1.0 1.0 2.0 4.0 2.0 0.5 0.5 2.0 2.0
0.0 0.5 0.5 0.0 0.5 1.0 0.0 1.0 2.0 2.0 2.0 1.0 1.0

IMPRESSAO DOS RESULTADOS

KIATY= 2

LISTA IATY =(1,3)

ND = NUMERO DE ELEMENTOS DA MATRIZ D = 160

NCD = NUMERO DE COLUNAS DA MATRIZ D = 10

MATRIZ D (COLUNA A PCS COLUNA)

0.0	0.100000E 01	0.0	0.0
0.100000E 01	0.100000E 01	0.100000E 01	0.0
0.0	0.100000E 01	0.100000E 01	0.100000E 01
0.0	0.100000E 01	0.100000E 01	0.100000E 01
0.0	0.100000E 01	0.100000E 01	0.0
-0.100000E 01	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	-0.100000E 01	0.0
0.100000E 01	0.0	-0.200000E 01	-0.100000E 01
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.100000E 01	0.0	-0.100000E 01	-0.100000E 01
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.100000E 01	0.0	0.100000E 01
0.0	-0.100000E 01	0.0	0.0
0.0	0.0	-0.100000E 01	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.100000E 01	0.0
0.100000E 01	0.0	-0.100000E 01	0.0
0.0	0.0	0.0	-0.100000E 01
0.0	0.0	0.0	0.0
0.0	0.0	0.0	-0.100000E 01
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
-0.200000E 01	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.100000E 01
0.100000E 01	-0.100000E 01	0.0	0.0
0.100000E 01	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
-0.300000E 01	0.0	0.0	0.0

LISTA CONTA = (0.1000000E 01, 0.1000000E 01)

LISTA BR = (-0.1000000E 01, -0.9999990E-02)

VALOR OTIMO DO PROBLEMA DIRETOR = 0.1333333E 01

LISTA 0 = (0.0 , 0.0 , 0.0
0.1000000E 01, 0.0 , 0.1000000E 01
0.0 , 0.1000000E 01, 0.3333333E 00)

DIRECAO VIAVEL-MATRIZ Z

0.0 0.0 -0.9999990E-02
0.1000000E 01 0.1000000E 01 0.0
-0.1000000E 01 0.1000000E 01 0.0

BIBLIOGRAFIA.

1. Arrow, K. J. and Hurwicz, L.: "Decentralization and Computation in Resource-Allocation". In Pfouts, R. (editor) : Essays in Economics and Econometrics. University of North Carolina Press, Chapel Hill, N.C. , 1960.
2. Balas, E.: "An Infeasibility-Pricing Decomposition Method for Linear Programs. Operations Research, Vol. 14, No. 5, Sept.-Oct. 1966.
3. Balinski, M. L. and Baumol, W. J.: "The Dual in Nonlinear Programming and Its Economic Interpretation. The Review of Economic Studies, Vol. 35, No. 3, July 1968.
4. Baumol, W.J. and Fabian, T.: "Decomposition, Pricing for Decentralization and External Economies. Management Science, Vol. 11, No. 1, Sept. 1964.
5. Dantzig, G. B. and Wolfe, P.: "Decomposition Principles for Linear Programs. Operations Research, Vol. 8, No. 1, Jan.-Feb. 1960.
6. Dantzig, G. B.: "Linear Programming and Extensions. Princeton University Press, Princeton, 1963.
7. Dantzig, G. B. and Wolfe , P.: "The Decomposition Algorithm

for Linear Programs. *Econometrica*, Vol. 29, No. 4, Oct. 1961.

8. Gale, D.: "The theory of Linear Economic Models." McGraw-Hill, New York, 1960.

9. Gass, S.I.: "Programacion Lineal Metodos y aplicaciones." Compañía Editorial Continental, S. A., México, 1969.

10. Geoffrion, A. M.: "Elements of Large-Scale Mathematical Programming." Working Paper No. 144, Western Management Science Institute, UCLA, 1969.

11. Geoffrion, A. M.: "Primal Resource-Directive Approaches for Optimizing Nonlinear Decomposable Systems." RAND Report No. RM-5829-PR, 1968.

12. Geoffrion, A. M.: "Duality in Nonlinear Programming : A Simplified Applications-Oriented Development." Working Paper No. 150, Western Management Science Institute
UCLA, 1969.

13. Geoffrion, A. M.: "Strictly Concave Parametric Programming, Part II : Additional Theory and Computational Considerations." Management Science, Vol. 13, No. 3, Nov. 1966.

14. Hirshleifer, J.: "On the Economics of Transfer Pricing." Journal of Business, Vol. 29, No. 3, July 1956.

15. Hirshleifer, J.: "Economics og the Divisionalized Firm." Journal of Business, Vol. 30, No. 2, April 1957.
16. Hogan, W. W. Jr.: "Optimization and Convergence for Extremal Value Function Arising from Structured Non-linear Programs." Western Management Science Institute, University of California, Los Angeles, Sept. 1971.
17. Jennegren, L.P.: "Studies in the Mathematical Theory of Decentralized Resource-Allocation." Ph.D., Stanford University, 1971.
18. Koopmans, T. C.: "Allocation of Resource and Price System." In Koopmans, T. C.: Three Essays on the State of Economic Science. McGraw-Hill, New York, 1957.
19. Kornai, J. and Liptak, Th.: "Two-Level Planning." Econometrica, Vol. 33, No. 1, January 1965.
20. Lasdon, L. S.: "Duality and Decomposition in Mathematical Programming." IEEE Transactions, Vol. SSC-4, No. 2, July 1968.
21. Lasdon, L. S.: "Optimization Theory for Large Systems." Macmillan, New York, 1970.
22. Malinvaud, E.: "Decentralized Procedures for Planning." In Bacharach, M. O. L. and Malinvaud, E. (editors): Activity Analysis in the Theory og Growth and Plan-

ning. Macmillan, London, 1967.

23. Mangasarian, O.: "Nonlinear Programming". McGraw-Hill, New York, 1969.

24. March, J. and Simon, H.: "Organizations". Wiley, New York, 1958.

25. Marschak, J.: "Elements for a Theory of Teams". Management Science, Vol. 1, No. 2, Jan. 1955.

26. Marschak, T.: "Centralization and Decentralization in Economic Organizations". Econometrica, Vol. 27, No. 3 July 1959.

27. Merbach, H. M.: "On Round-Off Errors in Linear Programming". Lectures Notes in Operations Research and Mathematical Systems. Spring-Verlag, Berlin, 1970.

28. Mesarovic, M. D., Macko, D. and Takahara, Y.: "Theory of Hierarchical, Multilevel, Systems". Academic Press, New York, 1970.

29. Polak E.: "Computational Methods in Optimization". Academic Press 1971.

30. Radner, R.: "The Application of Linear Programming to Team Decision Problems". Management Science, Vol. 5, No. 2, Jan. 1959.

31. Radner, R.: "Teams". Working Paper No. 255, Center of Research in Management Science, University of California at Berkeley, 1968.
32. Rockafellar, R. T.: "Convex Analysis", Princeton University Press, Princeton, New York, 1969.
33. Simonnard, M.: "Linear Programming". Prentice-Hall, Englewood Cliffs, New Jersey, 1966.
34. Stoer J. and Witzgall C.: "Convexity and Optimization in Finite Dimension I". Springer-Verlag, Berlin, Heidelberg, New York, 1970.
35. Varaiya, P.: "Trends in the Theory of Decision-Making in Large-Systems". Department of Electrical Engineering and Computer Sciences, University of California, Berkeley.
36. Weitzman, M.: "Iterative Multilevel Planning with Production Targets". *Econometrica*, Vol. 38, No. 1, Jan. 1970.
37. Zangwill, W.I.: "Nonlinear Programming : A Unified Approach", Prentice-Hall, Englewood Cliffs, New Jersey, 1969.
38. Zoutendijk, G.: "Methods of Feasible Directions". Elsevier Publishing Company, Amsterdam, 1960.
39. Zschau, E. V. W.: "A Primal Decomposition Algorithm for Linear Programming". Working Paper No. 91, Graduate

School of Business, Stanford University, 1967.

40. Zukhovitskii, S. I., Polyak, R. A. and Primak, M. E.:
"An Algorithm for the Solution of Convex Programming
Problems", DAN USSR 153(5), 991-1000, 1963.