

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

PROBLEMAS DE LOCALIZAÇÃO:  
SOLUÇÃO POR DECOMPOSIÇÃO

PAULO MORELATO FRANÇA

Orientador: HERMANO DE M.F. TAVARES

10/79

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

MARÇO - 1979

UNICAMP  
BIBLIOTECA CENTRAL

## R E S U M O

O assunto deste trabalho é relativo aos problemas de localização. Numa rede onde circula um certo produto entre pontos que o produzem (ou estocam) e pontos que o consomem (ou demandam), o problema de localização procura encontrar locais para instalação de centros produtores (ou distribuidores) de modo a tornar mínimos os custos que intervêm nesse sistema de distribuição. Entre outros, destacam-se os custos de transportar o produto, o custo de construir e/ou operar o centro produtor e o custo de produzir o produto.

Distinguem-se três partes no conteúdo deste trabalho. Na primeira, constituída pelo capítulo I, é feita uma revisão geral da literatura que se ocupa do problema de localização onde são discutidos modelos e técnicas do problema. Devida à grande proliferação de formulações e métodos para esse problema, sua análise geral é feita sob uma classificação que procura agrupar problemas típicos. Numa segunda parte - capítulos II e III - é estudado um problema específico de localização, o problema de localização estocástico (PLE), onde as demandas são consideradas variáveis aleatórias. Esse problema - um problema de programação mista não-linear - sendo encarado como a fusão do problema de localização de armazéns e o problema de transporte estocástico, sugere a aplicação de um método de decomposição, onde cada um desses problemas pode ser separadamente resolvidos, fugindo-se assim da não-convexidade imposta pelo problema global. A técnica de decomposição utilizada é a decomposição de Benders generalizada (DBG) e no capítulo III são relatadas experiências computacionais que atestam o bom comportamento do método numa série de problemas. É enfatizada uma regra de escolha de multiplicadores que fornece um eficiente "corte de Benders". Por fim são apresentadas extensões ao PLE e discutido o modo de resolvê-las.

*À Jesus, Mafalda, Angela e Ana*

## A G R A D E C I M E N T O S

Quero expressar meus agradecimentos às pessoas que direta ou indiretamente me ajudaram na elaboração desta Tese, e em especial ao HERMÃO pelo incentivo e encorajamento nesses anos to dos de convivência; ao PACCA, pelo tempo que tra balhamos em assuntos análogos e aos colegas da F.E.C., SECUNDINO, JURA, RUI, AKEBÔ, RAUL, CHRIS TIANO, AUTHIÊ, BOTTURA e GEROMEL por terem torna do este trabalho mais agradável; à TAKINHA, pelo paciente trabalho na confecção dos originais.

Este trabalho de pesquisa contou com o suporte do Conselho de Desenvolvimento Científico e Tecnológico - CNPq e da Coordenação de Aperfei çamento do Pessoal de Ensino Superior - CAPES.

# Í N D I C E

	<u>Pág.</u>
<u>CAPÍTULO I - PROBLEMAS DE LOCALIZAÇÃO: MODELOS E TÉCNICAS DE SOLUÇÃO</u> .....	1
I.1 - INTRODUÇÃO .....	1
I.2 - PROBLEMAS DE LOCALIZAÇÃO CONTÍNUOS .....	3
I.3 - PROBLEMAS DE LOCALIZAÇÃO DISCRETOS .....	6
I.3.1 - PROBLEMAS MINIMAX DE LOCALIZAÇÃO .....	7
I.3.2 - PROBLEMAS MINISOMA DE LOCALIZAÇÃO .....	8
I.3.2.1 - Problema das p-medianas .....	9
I.3.2.2 - Problema de localização de armazéns (PLA) .....	11
I.4 - PRINCIPAIS MODELOS E TÉCNICAS DE SOLUÇÃO DOS(PLA)..	12
I.4.1 - MÉTODOS HEURÍSTICOS .....	12
I.4.2 - ALGORITMOS DE SEPARAÇÃO E AVALIAÇÃO (BRANCH-AND-BOUND) .....	13
I.4.3 - ENFOQUE DE PROGRAMAÇÃO MISTA E DECOMPOSIÇÃO	16
I.5 - OUTROS PROBLEMAS DE LOCALIZAÇÃO DE ARMAZENS .....	18
I.5.1 - PROBLEMA DE LOCALIZAÇÃO ESTOCÁSTICO .....	18
I.5.2 - MODELOS COM CENTROS INTERMEDIÁRIOS E VÁRIOS PRODUTOS .....	19
I.5.3 - PROBLEMAS DE LOCALIZAÇÃO DINÂMICO .....	23
I.5.4 - EXTENSÃO PARA FUNÇÕES CUSTO NÃO-LINEARES ..	24
I.6 - REFERÊNCIAS BIBLIOGRÁFICAS .....	27

<u>CAPÍTULO II - PROBLEMA DE LOCALIZAÇÃO ESTOCÁSTICO RESOLVIDO PELA DECOMPOSIÇÃO DE BENDERS GENERALIZADA.</u>	37
II.1 - INTRODUÇÃO .....	37
II.1.1 - PROGRAMAÇÃO MATEMÁTICA DE GRANDE PORTE .	37
II.2 - DECOMPOSIÇÃO DE BENDERS GENERALIZADA (DBG) .....	39
II.3 - PROGRAMAÇÃO LINEAR ESTOCÁSTICA: EQUIVALENTES DETERMINÍSTICOS .....	49
II.3.1 - INTRODUÇÃO .....	49
II.3.2 - O PROBLEMA DE DOIS ESTÁGIOS .....	50
II.3.3 - PROBLEMAS ESTOCÁSTICOS COM RESTRIÇÕES DE RISCO .....	55
II.4 - FORMULAÇÃO DO PROBLEMA .....	58
II.5 - ESPECIALIZAÇÃO DA DBG .....	64
II.5.1 - ALGORITMO .....	68
II.5.2 - UM ALGORITMO ALTERNATIVO .....	72
II.6 - REFERÊNCIAS BIBLIOGRÁFICAS .....	74
<u>CAPÍTULO III - SOLUÇÃO COMPUTACIONAL DO PROBLEMA DE LOCALIZAÇÃO ESTOCÁSTICO</u> .....	76
III.1 - INTRODUÇÃO .....	76
III.2 - RESOLUÇÃO DO SUBPROBLEMA .....	77
III.2.1 - O ALGORITMO DE FRANK-WOLFE .....	78
III.2.2 - APLICAÇÃO DO ALGORITMO DE FRANK-WOLFE PARA O PROBLEMA DE TRANSPORTE ESTOCÁSTICO .....	79

	<u>Pág.</u>
III.2.3 - ANÁLISE DA CONVERGÊNCIA DO MÉTODO DE FRANK-WOLFE .....	82
III.3 - O SUBPROBLEMA E OS MULTIPLICADORES .....	86
III.4 - RESOLUÇÃO DO PROBLEMA MESTRE .....	90
III.4.1 - INTRODUÇÃO: MÉTODOS DE ENUMERAÇÃO IM- PLÍCITA .....	90
III.4.2 - GERAÇÃO DE SOLUÇÕES PARCIAIS .....	92
III.4.3 - SONDAÇÃO DE S .....	94
III.4.4 - EXEMPLO .....	96
III.5 - IMPLEMENTAÇÃO COMPUTACIONAL E RESULTADOS .....	98
III.6 - CONCLUSÕES .....	101
III.7 - REFERÊNCIAS BIBLIOGRÁFICAS .....	102
 <u>CAPÍTULO IV - EXTENSÕES DO PROBLEMA DE LOCALIZAÇÃO ESTOCÁS- TICO</u> .....	 105
IV.1 - PROBLEMA DE LOCALIZAÇÃO ESTOCÁSTICO COM RESTRI- ÇÕES DE RISCO .....	105
IV.2 - PROBLEMA DE LOCALIZAÇÃO ESTOCÁSTICO CONSIDERANDO CENTROS DE DISTRIBUIÇÃO INTERMEDIÁRIOS .....	110
IV.3 - EXTENSÃO PARA FUNÇÕES CUSTO DESCONTÍNUAS .....	113
IV.4 - REFERÊNCIAS BIBLIOGRÁFICAS .....	117
IV.5 - APÊNDICE I .....	118

## CAPÍTULO I

### PROBLEMAS DE LOCALIZAÇÃO: MODELOS E TÉCNICAS DE SOLUÇÃO

#### I.1 - INTRODUÇÃO

A intenção deste capítulo é apresentar um panorama do problema de localização, através de uma revisão da literatura no assunto. O rápido e diversificado desenvolvimento de problemas, modelos e técnicas de solução envolvendo a análise de localização destrói qualquer intenção de ser exaustiva esta revisão e mesmo uma tentativa de mostrar o assunto dentro de um enfoque unificado e coerente. Procurar-se-á neste capítulo apresentar não só modelos mas também discutir técnicas de solução até agora aplicadas ao problema. Depois de uma breve discussão na área geral da análise de localização e na área dos problemas contínuos, a atenção será concentrada em modelos e técnicas para os problemas discretos, onde se desenvolve esta tese.

A preocupação matemática sobre decisões de localização é bem antiga. Este problema já aparecia no século XV com Cavalieri e posteriormente com Fagnano, Tedenat, Steiner e outros [31]. Fermat também é considerado o introdutor do problema e consta que Torricelli resolveu-o para o caso de três pontos (achar o ponto cuja soma das distâncias a 3 pontos dados é mínima). Mas, contemporaneamente foi Alfred Weber [92] quem formalizou o problema de localizar uma fábrica com o objetivo de minimizar os custos de transporte em relação a três pontos, dois fornecedores e um outro de consumo. Desde então muito se tem publicado no assunto, como pode ser atestado, por exemplo, pelo artigo de Francis e Goldstein [28] que relaciona 226 trabalhos numa tentativa de dar uma bibliografia relativa a assunto. Outros artigos tipo "survey" foram feitos, como o exaustivo (273 páginas) trabalho de Lea [58], ou Eilon et al. [20], cap. 2, Reville et al. [78], Scott [81], Galvão [31]. Na mesma linha há trabalhos que procuram discutir uma classe particular de problemas como em McGinnis [69] que analisa e compara os problemas de localização que usam a técnica de solução MIP (mixed-integer

programming) ou ainda El-Shaieb [24] que se ocupa só do problema de Weber simples.

Um problema de localização pode ser motivado de duas maneiras em relação ao objetivo a alcançar. A primeira quando quer-se tomar uma decisão de localização, procurando-se minimizar custos, por exemplo, de transportar mercadorias entre pontos de uma rede, custos de construir instalações na rede, ou ainda custos de armazenagem. Uma outra maneira se apresenta quando os custos ou benefícios são difíceis de quantificar e então procuram-se tomar decisões de localizar baseado em critérios de distância ou tempo mínimo. Os critérios de minimização de custos aparecem frequentemente em problemas que afetam o setor privado, como construção de fábricas, armazéns, entrepostos, onde os critérios de minimização de custos ou maximização de lucros são imperativos. Outros critérios que procuram retratar a maximização de alguma utilidade são mais comuns nos problemas que afetam o setor público [78], como localizar escolas, hospitais, serviços públicos, bombeiros, etc.

A motivação de um problema de localização pode ainda ser visualizada quando consideramos, por exemplo, o caso de localizar armazéns numa rede de distribuição, onde intervêm custos de armazenagem do produto, custos de transporte e custos de construção dos armazéns. É intuitivo que quanto maior o número de armazéns na rede, menores os custos de distribuição e maiores os custos de construção e armazenagem. Na procura de uma solução de compromisso, encontra-se o número ideal de armazéns e sua localização. A figura I.1 mostra isso.

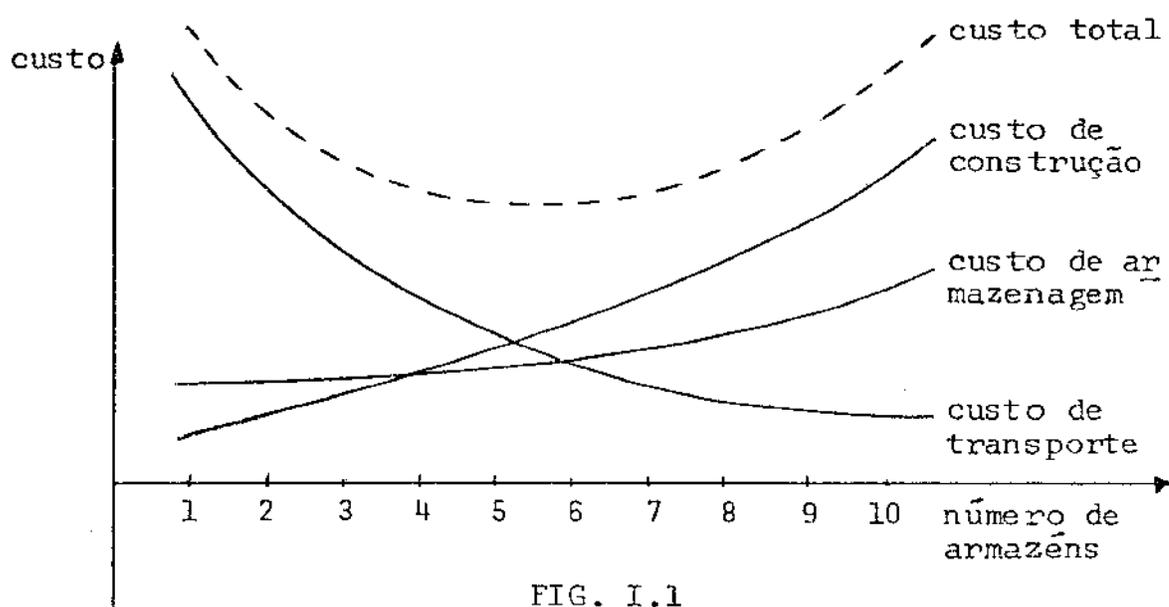


FIG. I.1

Alguns artigos discutem objetivos a serem alcançados e traçam orientações metodológicas para a escolha de modelos, funções objetivos e técnicas para diferentes problemas de localização. Entre outros destacam os trabalhos de Geoffrion [34], Rand [76], Revelle et al. [78] e Beattie [6].

Nas duas próximas seções será feito um esforço analítico no sentido de classificar os problemas de localização. Deve-se entretanto ressaltar que não se pode enquadrar todos os problemas e modelos nessa classificação ou qualquer outra, devido a existência de problemas combinados que escapam a qualquer rigidez que se queira impor à análise. Porém, num sentido amplo, há consenso sobre uma divisão dos problemas de localização em duas grandes categorias: os problemas contínuos e os discretos. Esta divisão apresenta-se também com outros nomes, guardando porém a sua essência. Reconhecem-se, por exemplo, os problemas de localização no plano e em redes [78] ou ainda como problemas de localização com espaço infinito de soluções e espaço finito de soluções [31].

## I.2 - PROBLEMAS DE LOCALIZAÇÃO CONTÍNUOS

Os problemas contínuos caracterizam-se:

- a. por procurar localizar pontos de serviços ("facilities") num espaço infinito de soluções (no plano, por exemplo) e não se restringe aos nós e arcos de uma rede.
- b. por usar uma medida de distância de acordo com uma métrica particular.

A utilização de uma certa métrica serve para subdividir os problemas contínuos. Assim reconhecem-se os problemas com distância Euclidiana e os problemas com distância retilínea. Usando métrica Euclidiana define-se distância como:

$$d_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 \quad (1)$$

onde

$d_{ij}$  = distância entre pontos  $i$  e  $j$

$(x_i, y_i)$  = coordenadas do ponto  $i$ , num sistema retangular

Nos problemas com distância retilínea, (métrica das cidades), define-se:

$$d_{ij} = |x_i - x_j| + |y_i - y_j| \quad (2)$$

Como já dito, foi A. Weber no começo do século quem examinou o problema de localizar uma fábrica para atender um mercado e dois fornecedores de matéria prima. Posteriormente o problema de Weber foi generalizado para achar um ponto cuja soma das distâncias ponderadas a  $m$  pontos dados no plano seja mínima.

Este difícil problema foi solucionado, com métrica Euclidiana, pela primeira vez por Weiszfeld [93] em 1937, que usou um procedimento iterativo ainda hoje considerado um excelente algoritmo. Os trabalhos mais recentes de Kuhn e Kuenne [56] e Cooper [13] redescobriram este procedimento ao estender o problema para a localização de  $n$  pontos no plano (fábricas) para atender a  $m$  pontos dados (consumidores). Quer-se, então, formalmente

$$\text{minimizar } f = \sum_{i=1}^n \sum_{j=1}^m z_{ij} \xi_j d_{ij},$$

onde

$\xi_j$  = fator que pondera o consumidor  $j$ .

$d_{ij}$  = distância entre fábrica  $i$  e consumidor  $j$ , definida por (1) ou (2).

$z_{ij} =$   
1 se consumidor  $j$  é serviço pela fábrica  $i$   
0 caso contrário

Na ausência de restrições, tomam-se derivadas parciais para a obtenção da solução ótima:

$$\frac{\partial f}{\partial x_i} = \sum_{j=1}^m \left[ z_{ij} \xi_j (x_i - x_j) / d_{ij} \right] = 0, \quad i=1,2,\dots,n$$

$$\frac{\partial f}{\partial y_i} = \sum_{j=1}^m \left[ z_{ij} \xi_j (y_i - y_j) / d_{ij} \right] = 0, \quad i=1,2,\dots,n.$$

A solução das equações acima fornece

$$x_i = \frac{\sum_{j=1}^m (z_{ij} \xi_j x_j / d_{ij})}{\sum_{j=1}^m z_{ij} \xi_j / d_{ij}}, \quad i=1,2,\dots,n$$

$$y_i = \frac{\sum_{j=1}^m (z_{ij} \xi_j y_j / d_{ij})}{\sum_{j=1}^m z_{ij} \xi_j / d_{ij}}, \quad i=1,2,\dots,n.$$

Essas equações podem ser resolvidas iterativamente. Seja  $k$  um índice que indica a iteração. Então:

$$x_i^{k+1} = \frac{\sum_{j=1}^m (z_{ij} \xi_j x_j^k / d_{ij})}{\sum_{j=1}^m (z_{ij} \xi_j / d_{ij})}, \quad i=1,2,\dots,n$$

$$y_i^{k+1} = \frac{\sum_{j=1}^m (z_{ij} \xi_j y_j^k / d_{ij})}{\sum_{j=1}^m (z_{ij} \xi_j / d_{ij})}, \quad i=1,2,\dots,n$$

O procedimento termina quando  $(x_i^{k+1} - x_i^k)$  e  $(y_i^{k+1} - y_i^k)$  forem desprezíveis.

É demonstrado que este procedimento converge para uma única solução no caso que  $n = 1$ , (veja, e.g., Kuhn e Kuenne [56] e El - Shaieb [24] ) pois a função  $f(\cdot)$  é convexa. No caso geral, ( $n > 1$ ), apenas é garantido um ótimo local. Ostresh [74] discute a convergência desses métodos iterativos para o problema de Weber. Em [24] é sugerido o uso de métodos de busca para solução do caso ( $n = 1$ ) e é relatada uma experiência usando os métodos de Fibonacci e aproximação quadrática. Bellman [7] aplicou a programação dinâmica neste caso também sem alcançar muito sucesso. Há um conjunto de trabalhos que procuram tratar o problema geral ( $n > 1$ ) sob o enfoque de programação

linear. Entre outros, são conhecidos os trabalhos de Juel e Love [51], Wesolowsky e Love [96], Cabot e al. [11], Pritsker e Ghare [75], Sherali e Shetty [83] que se utilizam de técnicas primais e duais. Os dois últimos levam em conta a existência de fluxos entre os pontos e satisfação de demandas dos consumidores.

Algumas extensões aos problemas contínuos de localização foram realizadas. Watson-Gandy e Eilon [91] estudaram o caso de funções custo descontínuas. Em Drezner e Wesolowsky [18] é considerada a generalização de se localizar um ponto numa esfera usando diversas normas. O interessante trabalho de McGinnis e White [70] funde sob o enfoque multicritério as formulações clássicas de minisoma e minimax, que são discutidas na seção I.3. Usando o conceito de restrição de risco, Seppala [82] desenvolveu um algoritmo para o caso em que as ponderações na distância são variáveis aleatórias. Cooper [14] também considera o caso estocástico.

### I.3 - PROBLEMAS DE LOCALIZAÇÃO DISCRETOS

Os problemas de localização discretos são caracterizados fundamentalmente por:

- a. um espaço de soluções que consiste dos pontos de uma rede (podendo ser nós ou pontos nos arcos).
- b. uma distância medida ao longo da rede,

por exemplo,  $d_{ij}$  = comprimento do caminho mínimo entre os nós  $i$  e  $j$ .

Como os problemas discretos supõem implicitamente a existência de uma rede para selecionar locais, eles também são chamados de problemas de localização em redes. Verifica-se também que são esses problemas que têm recebido maior atenção e desenvolvimento.

De acordo com a natureza da função objetivo, os problemas discretos se dividem em dois grupos:

\* quando os pontos de serviço a serem localizados

têm um caráter de atendimento de emergência, como hospitais ou bombeiros, etc., o objetivo mais racional é procurar localizá-los de modo a minimizar a maior distância que separa um ponto de demanda da rede do ponto de serviço mais próximo. Nesse caso tratamos os problemas minimax.

- \* quando procura-se localizar armazéns, fábricas, escolas, etc., uma função objetivo mais apropriada é minimizar a distância total a ser percorrida para a satisfação da demanda. São os problemas minisoma. Estudaremos resumidamente alguns resultados alcançados em minimax e mais extensivamente em minisoma.

### 1.3.1 - PROBLEMAS MINIMAX DE LOCALIZAÇÃO

Há uma grande variedade de problemas minimax. Handler [46] propôs uma classificação para distinguir os diferentes modelos; composta de 4 caracteres:

- conjunto dos locais possíveis para localizar: N ou P.
- conjunto dos pontos de demanda: N ou P
- número de centros a localizar ou máxima distância:  $p$  ou  $\delta^{-1}$  (respectivamente).
- tipo da rede: T ou G.

onde N denota só os nós da rede e P compreende os nós ou pontos nos arcos;  $p$  é o número de centros a localizar e  $\delta$  a distância crítica, ou seja, a máxima distância permitida entre um centro de demanda e o ponto de serviço mais próximo. O símbolo  $\delta^{-1}$  é usado para o problema inverso, ou seja, determinar o mínimo número de pontos de serviço de modo que todos os centros de demanda fiquem a uma distância de pelo menos  $\delta$ , de no mínimo um ponto de serviço. T significa estarmos lidando com árvores; G caso contrário. Um problema seria então definido como, por exemplo, P/N/P/G, aliás, o mais clássico dos problemas minimax. Deve-se ressaltar que o problema acima, assim como todos os demais onde os pontos de serviço podem ser localizados em locais quaisquer da rede (nós e arcos), o espaço de soluções não é discreto mais sim o contínuo de pontos do grafo. Porém, manteremos tais casos dentro da classificação dos problemas discretos, por simplicidade.

O problema  $N/N/1/G$  foi introduzido e resolvido por Hakimi [41] e sua generalização para  $p$  centros,  $N/N/p/G$ , foi tratada por Toregas et al. [89]. Em Hakimi [41] também é introduzido e solucionado o problema  $P/N/1/G$  e a generalização para  $P/P/1/G$  foi feita por Frank [30]. A variedade  $P/N/p/G$ , também chamado do problema dos  $p$ -centros, foi proposto por Hakimi [42] e solucionado por Minieka [71], Christofides e Viola [12], Handler [46] e Garfinkel et al. [32], entre outros. Em Handler [46] também são vistos os problemas  $P/P/p/G$  e  $N/P/p/G$ . Os algoritmos desenvolvidos nos trabalhos acima baseiam-se na solução sucessiva de uma série de problemas de recobrimento. Porém, no caso de árvores, algoritmos fundamentados da teoria de grafos mostram-se mais eficientes, como mostrado em Handler [47] para o caso de um único centro, em Handler [48] para o caso de  $p=2$  ( $P/P/2/G$  e  $P/N/2/G$ ) e Hakimi et al [44] quando  $p > 2$ . Outros resultados que tratam com localizações em árvores foram conseguidos por Dearing e Francis [16], Goldman [38] e Halfin [45].

Há uma variante aos problemas minimax que trabalham com distâncias Euclidianas ou retilíneas. Nesses casos quer-se instalar, por exemplo,  $N$  novas facilidades entre  $M$  já existentes com o objetivo de minimizar a máxima distância Euclidiana ponderada entre todos os pontos de serviço. Um exemplo de aplicação desses problemas é quando se deseja instalar um certo número de radares para cobrir o tráfego de determinados aeroportos. Cada radar precisa estar o mais próximo de algum aeroporto para controlar seu tráfego e perto o suficiente de outro radar para que se tenha uma rede integrada de radares. Dearing e Francis [17], Love et al [61], Wesolowsky [96], Morris [73], e Elzinga et al. [27] são alguns dos trabalhos nessa linha. Em geral os métodos usados são de programação não-linear (teoria de dualidade [27] e SUMT [61]) ou programação linear ([96],[73]).

### I.3.2 - PROBLEMAS MINISOMA DE LOCALIZAÇÃO

Dentro do contexto de fluxos em redes, Baumol e Wolfe [5] foram os primeiros a formularem um problema de localização onde o objetivo era minimizar a soma total dos custos envolvidos. Este trabalho criou o problema conhecido como o problema de localização de armazéns (PLA) - ("plant", "warehouse" ou "facility location problem") onde concentrou-se grande parte

do esforço da análise de localizações.

Dentre as múltiplas variações de formulação que esse problema sofreu, a introduzida por Hakimi [41,42] foi particularmente importante pela atenção que despertou e pelo desenvolvimento que apresentou posteriormente. Trata-se do problema das p-medianas ou p-centros<sup>1</sup>. Esta formulação não trabalha com os fluxos das mercadorias, e ao contrário do PLA, inclui como possíveis pontos de localização toda a rede considerada, ou seja, nós e arcos. Nos ocuparemos brevemente dos problemas das p-medianas e a seguir será feito um levantamento do estado atual de desenvolvimento dos PLA.

### I.3.2.1 - Problema das p-medianas

O problema das p-medianas é um problema de localização minisoma que procura localizar um certo número (p) de pontos de serviço numa rede, tal que a soma das mínimas distâncias de cada nó da rede ao mais próximo ponto de serviço seja minimizada.

Seja um grafo conexo, não ordenado,  $G=(N,A)$  onde  $N$  é o conjunto dos  $k$  nós e  $A$  o de arcos. Define-se um ponto do grafo  $G$ , a qualquer ponto situado sobre  $G$ , sendo um nó ou não. O comprimento do caminho mínimo entre dois pontos  $x$  e  $y$  é definido por  $d(x,y)$ , a distância entre  $x$  e  $y$ . Seja  $X_p = \{x_1, x_2, \dots, x_p\}$  um conjunto de  $p$  pontos que devem ser localizados em  $G$  e define-se para cada  $1 \leq i \leq k$

$$d(n_i, X_p) = \min \{d(n_i, x_1), d(n_i, x_2), \dots, d(n_i, x_p)\},$$

que é a mínima distância de  $n_i$  ao mais próximo ponto de  $X_p$ . Se  $w(n_i) \cdot d(n_i, X_p)$  é o custo de satisfazer a demanda no nó  $n_i$  então

---

<sup>1</sup> Em geral o problema dos p-centros é identificado como um problema minimax, ficando a denominação p-medianas para os problemas minisoma, mas esse rigor nem sempre é observado pelos autores.

$$f(X_p) = \sum_{i=1}^k w(n_i) d(n_i, X_p)$$

é o custo de satisfazer todas as demandas, assumindo que os  $p$  pontos de serviço estão localizados em  $G$ , correspondem a  $X_p$  e que não existem restrições nos seus tamanhos. Os locais ótimos de instalação correspondem aos pontos  $X_p^*$  em  $G$  tal que

$$f(X_p^*) \leq f(X_p) \text{ para todo } X_p \text{ em } G.$$

Esse é um problema muito complexo para resolver. A fim de mostrar os resultados teóricos que permitem solucioná-lo, vamos simplificar a formulação. Suponha o caso onde  $k=1$ . Deve-se achar, então um ponto  $y^*$  em  $G$  tal que  $f(y^*) \leq f(y)$ , para todo  $y$  em  $G$ . Goldstein provou que no caso de  $G$  ser uma árvore,  $y^*$  estará sempre sobre um nó. Hakimi [41] generalizou este resultado para  $G$ , ao provar o teorema abaixo.

**Teorema 1** - Há pelo menos um nó  $n$  de  $G = (N, A)$  para o qual  $f(n) \leq f(y)$  para qualquer  $y$  em  $G$ .

Posteriormente, Hakimi [42] generalizou o teorema 1 para o caso de  $k=p \geq 1$ ,

**Teorema 2** - Há pelo menos um subconjunto  $X_p \subset N$  contendo  $p$  nós, tal que  $f(X_p) \leq f(Y_p)$  para qualquer conjunto  $Y_p$  de  $p$  pontos situados em nós ou arcos de  $G$ .

O importante resultado desses teoremas diz, em outras palavras, que a procura das soluções ótimas do problema dos  $p$ -medianas pode ficar restrita aos nós da rede.

Esses resultados foram ainda ampliados por Levy [60] que provou serem válidos os teoremas 1 e 2 no caso de custos côncavos e por Goldman [39] que mostrou também a aplicabilidade dos teoremas 1 e 2 no caso de um nó ser ponto de demanda e/ou de oferta. Hakimi e Maheshwari [43] e Wendell e Hurter [94] provaram a validade dos teoremas mesmo quando estão presentes restrições de capacidade nos arcos e mesmo restrições de tamanho dos centros, desde que se permita localizar mais de um centro num nó.

Quanto aos métodos de solução para o problema das p-medianas distinguem-se:

\*\* algoritmos heurísticos; Maranzana [65], Teitz e Bart [88].

\*\* algoritmos de separação e avaliação (branch-and-bound).

Entre outros, citam-se os trabalhos de El-Shaieb [25] e Järvinen et al. [49].

\*\* algoritmos baseados na programação linear.

Um enfoque é formular o problema como programação inteira e solucioná-lo por relaxação. Os problemas relaxados são de programação linear e algoritmos foram desenvolvidos por Reville e Swain [79], Garfinkel et al [33]. O enfoque de Marsten [68] usa dualidade e programação paramétrica para propor um interessante algoritmo. Para uma discussão aprofundada de modelos e métodos do problema das p-medianas veja o trabalho de Galvão [31].

#### I.3.2.2 - Problema de Localização de Armazéns -(PLA)

Foram os problemas de localização de armazéns que há vinte anos atrás deram início ao rápido desenvolvimento contemporâneo da análise de localizações. O problema, que parece ter sido proposto inicialmente por Baumol e Wolfe [5], teve sua formulação inspirada no clássico problema de transporte da programação linear. Apesar de na formulação original os autores tratarem especificamente do caso de localização de armazéns, o problema estendeu-se para localização de fábricas e depois, mais genericamente, para "facilidades".

O problema pode ser encarado como encontrar locais ótimos para instalação de armazéns (fábricas, etc.) dentre um número finito de locais candidatos e de modo a satisfazer a demanda de centros consumidores, procurando minimizar os custos envolvidos (transporte, estocagem, operação, construção, etc.).

O que caracteriza fundamentalmente os PLA são o tipo de sua função objetivo e a presença ou não de restrições de capacidade. Em geral os problemas supõem função objetivo linear (custos de transporte) com custo fixo de instalação.

Esses modelos serão analisados na próxima seção. O caso de funções não-lineares (custos côncavos ou lineares por partes) são estudados na seção I.5.4. Os PLA sofrem outras caracterizações, como a presença ou não de centros intermediários (seção I.5.2), se é considerado um só produto ou vários (seção I.5.2) e ainda se a demanda é sazonal (seção I.5.3) ou então se ela é estocástica (seção I.5.1).

#### I.4 - PRINCIPAIS MODELOS E TÉCNICAS DE SOLUÇÃO DOS PLA

Em relação ao modelamento, distinguem-se duas formulações gerais, os modelos com capacidade ilimitada e os com capacidade restrita. Em relação às técnicas usadas na solução dos problemas reconhecem-se três grupos:

- \* métodos heurísticos
- \* algoritmos de separação e avaliação
- \* enfoque de programação mista e decomposição.

##### I.4.1 - MÉTODOS HEURÍSTICOS

A preocupação de se desenvolver métodos heurísticos para resolver os PLA se deve à dificuldade computacional inerente a resolver de maneira ótima (exata) tais problemas. Os algoritmos heurísticos contentam-se com soluções boas (não-ótimas) e sua utilização precisa ser precedida de uma análise do que se quer retirar dos resultados da otimização do sistema. Em alguns casos práticos, a utilização de tais métodos impossibilita respostas às perguntas importantes, como análise de sensibilidade, prioridades de implementação, etc. [34]. Além disso o aparecimento de métodos exatos de otimização poderosos e eficientes, permite resolver mesmo problemas de grandes dimensões sem necessidade de lançar mão de heurísticas. Uma exposição da filosofia do enfoque heurístico é encontrada em Kuehn e Hamburger [55].

Um significativo esforço, porém, foi dedicado aos algoritmos heurísticos. Um dos primeiros algoritmos bem sucedidos foi o de Kuehn e Hamburger [55]. Eles dividem sua heurística em duas partes: Na primeira, um programa principal localiza

armazéns, um de cada vez, até que nenhum mais possa ser colocado na rede sem que aumente o custo total. Depois, uma outra rotina tenta modificar a solução do programa principal, trocando ou tirando armazéns de modo a melhorar a solução atual. Feldman et al. [28] propuseram uma heurística semelhante, e consideraram economias de escala, assim como o trabalho de Manne [64]. Outros métodos heurísticos foram propostos por Sá [80], Marks e Lieberman [67]. O já citado pioneiro trabalho de Baumol e Wolfe [5] sugere um algoritmo heurístico para encontrar boas soluções de um problema com função objeto côncava (custo de armazenagem côncavo) e restrições sobre a capacidade dos armazéns. Trabalhos aplicados ou estudo de casos sugerem heurísticas apropriadas (e.g. Burstall et al. [10], Lawrence e Pengilly [57], Monterosso [72], e Silva e Lins [84]).

#### I.4.2 - ALGORITMOS DE SEPARAÇÃO E AVALIAÇÃO - (BRANCH-AND-BOUND)

A técnica de separação e avaliação, em geral, se aplica bem aos modelos mais simples dos PLA, ou seja, modelos com um único produto, custos lineares de transporte e operação e um único estágio de distribuição (fábrica - armazém ou armazém - consumidor). Aqui só relataremos os principais algoritmos propostos em separação e avaliação especificamente endereçados a esses problemas mais simples, visto que os problemas com vários produtos e dois estágios são analisados em I.5.2 e os custos não-lineares em I.5.4.

O primeiro algoritmo exato proposto para resolver o PLA foi o de separação e avaliação de Efroymsen e Ray [19]. O modelo deles é sem restrição de capacidade e é a generalização de um modelo mais simples devido a Balinsky [3]. O problema formulado é

$$\text{minimizar } z = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} + \sum_{i=1}^m k_i z_i$$

sujeito a

$$\sum_{i \in N_j} x_{ij} = 1, \quad j = 1, 2, \dots, n,$$

$$0 \leq \sum_{j \in P_i} x_{ij} \leq n_i z_i, \quad i = 1, 2, \dots, m,$$

$$z_i = 0, 1$$

onde

$$c_{ij} = t_{ij} \cdot D_j$$

$t_{ij}$  = custo unitário de transporte do armazém  $i$  ao consumidor  $j$ ,

$D_j$  = demanda do consumidor  $j$ ,

$k_i$  = custo fixo de instalação do armazém  $i$ ,

$z_i$  = 1 se o armazém for instalado em  $i$ ; 0 caso contrário,

$x_{ij}$  = fração da demanda do consumidor  $j$  que é suprida pelo armazém  $i$ ,

$m$  = número de possíveis locais dos armazéns,

$n$  = número de centros consumidores,

$P_i$  = conjunto de consumidores que são supridos pelo armazém  $i$ ,

$n_i$  = número de elementos de  $P_i$ ,

$N_j$  = conjunto de armazéns que podem suprir o consumidor  $j$ .

A eficiência do enfoque de Efraymson e Ray está baseada na solução imediata do problema de programação linear que resulta da fixação do vetor de localizações  $z$ . Uma experiência computacional relatada com um problema de 50 armazéns (50 variáveis 0-1) e 200 pontos de consumo, consumiu cerca de 10 minutos de CPU (IBM 7094). O algoritmo pode ser estendido para funções custo lineares por partes. Spielberg [86] discute o mesmo problema e sugere melhorias para aumentar a eficiência computacional do algoritmo, além de incorporar ao modelo algumas restrições adicionais, como orçamento disponível. Muitos outros algoritmos foram desenvolvidos para o PLA sem restrições de capaci-

dade<sup>\*</sup>, mas o de Khumawala [53] parece ser o melhor do ponto de vista de eficiência e flexibilidade.

O chamado PLA capacitado considera que os armazéns têm uma capacidade limitada de armazenamento (ou as fábricas têm uma limitada capacidade de produção) e vários algoritmos de separação e avaliação foram propostos para resolvê-lo. Uma possível formulação é

$$\text{minimizar } \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{i \in I} k_i z_i$$

sujeito a

$$\sum_{j \in J} x_{ij} \leq S_i z_i, \quad i \in I = \{1, 2, \dots, m\},$$

$$\sum_{i \in I} x_{ij} = D_j, \quad j \in J = \{1, 2, \dots, n\}$$

$$x_{ij} > 0, \quad i \in I, \quad j \in J$$

$$z_i = 0, 1, \quad i \in I$$

onde

$x_{ij}$  = fluxo de produto entre  $i$  e  $j$

$z_i$  = 0 se o armazém  $i$  é instalado; 0 caso contrário

$c_{ij}$  = custo unitário de transporte

$D_j$  = demanda do consumidor  $j$

$S_i$  = capacidade do armazém  $i$

$k_i$  = custo fixo associado e instalação (operação) do armazém  $i$ .

Dentre os vários algoritmos propostos citam-se os de Davis e Ray [15], Ellewein e Gray [22], Ellewein [21], Sá [80]. Os dois últimos são comparados com o algoritmo desenvol

---

<sup>\*</sup> veja Balinsky e Spielberg [4]

vido por Akinc e Khumawala [1] que se mostrou bem mais eficiente numa série de 12 problemas de dimensões variadas. Nesta mesma classe de modelos, Geoffrion e McBride [35] desenvolveram um algoritmo capaz de tratar várias restrições adicionais, como limitantes inferiores e superiores na capacidade dos armazéns, restrições adicionais sobre as variáveis de fluxo e de localização e podem restringir a satisfação da demanda de um consumidor através de um único armazém. Eles relatam que um problema médio (25 armazéns e 100 consumidores) pode ser resolvido em aproximadamente 1 minuto de CPU (IBM 360/91).

#### I.4.3 - ENFOQUE DE PROGRAMAÇÃO MISTA E DECOMPOSIÇÃO

Um PLA pode ser encarado como um problema de programação mista -MIP- (contém variáveis reais e inteiras). O desenvolvimento de técnicas eficientes em programação inteira, em conjunto com o aparecimento da técnica de separação e avaliação, criou condições de se desenvolverem sistemas de programação mista de uso geral. Dentro então, de certas limitações, um PLA pode ser solucionado com esses pacotes comerciais de MIP. Os de uso geral mais conhecidos são o UMPIRE da UNIVAC, o MPSX-MIP para os 360 e 370 da IBM e Ophelie Mixed da CDC. Todos eles usam algoritmos de separação e avaliação. Esses sistemas permitem tratar de maneira ótima versões bastante sofisticados do PLA, como funções custo linear por partes, restrições sobre capacidade e outras restrições sobre variáveis reais e inteiras, permitindo muita flexibilidade no tratamento do problema. A mais séria limitação no uso desses pacotes se encontra no tamanho possível dos problemas tratados e na eficiência computacional [34]. No caso do tempo de máquina disponível ser um fator limitado e desde que os erros de cálculo acumulados não comprometam o resultado, esses pacotes se destinam mais a subotimizações. Além disso, só problemas de pequeno ou médio porte podem ser tratados.

Uma outra categoria de pacotes de otimização é constituída por sistemas especialmente desenvolvidos para os problemas de distribuição e localização. Os sistemas Poligami para o CDC 6600 e Capflo permitem que os modelos incorporem relativa flexibilidade, como, um único produto, dois estágios de distri-

buição, restrições na capacidade dos armazéns, funções custo com economias de escala (lineares por partes). Poligami trabalha dentro de um enfoque convencional de separação e avaliação que resolve um problema de fluxo em redes pelo método "out-of-kilter" para fazer a avaliação. Como relatado em Geoffrion [34] o sistema é apropriado a solução de problemas de no máximo 40 variáveis zero-um. Capflo guarda quase as mesmas características.

São raras as publicações que trazem experiência computacional com esse tipo de sistemas. Tentando mostrar a aplicabilidade do sistema Ophelie Mixed, Elson [26] relata que levou em torno de 8 minutos para resolver um problema com 15 fábricas, 3 produtos, 45 armazéns e 50 pontos de consumo.

Decomposição é um outro enfoque dividido para resolver problemas mistos de programação matemática. A idéia intuitiva da decomposição é dividir um problema em vários outros menores e de mais simples solução. Como decorrência óbvia, os métodos de decomposição se mostraram adequados aos problemas grandes, de muitas variáveis e por isso passaram a pertencer ao ferramental de solução dos sistemas de grande porte. No caso dos PLA o enfoque de decomposição se aplica procurando separar a parte inteira da parte real do problema e resolvê-las em separado, com a supervisão de um esquema coordenador. O algoritmo de decomposição de Benders [8] procura realizar isso. Ele não foi muito utilizado de início, para problemas de localização, mas quando os problemas práticos começaram a exigir modelos de grandes dimensões, de muitos produtos e boa flexibilidade, ela ganhou importância.

O algoritmo "linear" de Benders foi desenvolvido para problemas de programação mista onde o PLA é um caso particular, mas ele se aplica a problemas de forma geral.

$$\text{minimizar } cx + f(y)$$

$$\text{sujeito a } Ax + F(y) \leq b, x \geq 0, y \in Y$$

onde as variáveis  $y$  são "complicantes", no sentido que quando fixadas temporariamente num valor, o problema resultante é um programa linear. Daí o sentido de chamar algoritmo de Benders "linear". No caso do PLA,  $y$  são as variáveis zero-um.

Balinsky [3] foi o primeiro a usar a decomposição de Benders num PLA com capacidade ilimitada dos armazéns, mas foi o trabalho de Geoffrion e Graves [37] que consolidou-a como um poderoso método para solução dos PLA, principalmente no caso de vários produtos. Eles relatam extensa experiência computacional no projeto de uma rede de distribuição de alimentos com 14 fábricas, 17 produtos diferentes, 45 locais possíveis de instalação de centros de distribuição, e 121 zonas consumidoras. Em média, a solução do problema levou em torno de 1 minuto de CPU (IBM 360/91). Num trabalho posterior, Geoffrion et al. [35] ampliam o modelo anteriormente usado permitindo maior sofisticação do problema tratado e usando basicamente o mesmo enfoque da decomposição de Benders. Eles relatam que para problemas com 100 produtos, 100 centros de distribuição e 400 zonas de demanda, um computador de grande porte não leva mais que 15 minutos de CPU, com critérios de tolerância em torno de 1%.

## I.5 - OUTROS PROBLEMAS DE LOCALIZAÇÃO

### I.5.1 - PROBLEMA DE LOCALIZAÇÃO ESTOCÁSTICO

O problema de localização estocástico (PLE) é uma extensão natural dos PLA e ele aparece quando o conhecimento sobre algum parâmetro não é completo. Supondo porém que tais incertezas seguem uma lei conhecida (distribuição de probabilidades), a resolução do problema deve apelar para o campo da programação estocástica. A idéia fundamental é formular um problema equivalente determinístico e a partir daí olhar o PLE como um simples problema de localização. Como acontece geralmente, a inclusão de considerações estocásticas dificultam a solução dos problemas, e no caso considerado, o equivalente obtido cai no campo da programação mista não-linear.

O PLE pode ser descrito como saber definir locais ótimos para a instalação de "facilidades" (fábricas, armazéns, edifícios, etc.) dentre um conjunto finito de locais candidatos, de modo a satisfazer as demandas aleatórias de um conjunto de pontos consumidores procurando minimizar os custos de

de instalação (e produção) e de transporte envolvidos.

O problema acima, numa formulação específica é resolvido nos capítulos II e III através de um enfoque de decomposição.

Alguns trabalhos já foram publicados tendo o PLE como objetivo. Leblanc [59] apresenta um método heurístico de resolução do PLE dentro de uma formulação clássica considerando as demandas aleatórias. O ótimo do problema não é garantido nesse caso e a eficácia da heurística proposta não é possível de ser avaliada devida à carência de aplicações práticas, principalmente considerando problemas de grande porte. Usando o enfoque de separação e avaliação, Balachandran e Jain [2] formularam um PLE mais geral, considerando as demandas aleatórias e funções de custo descontínuas. O algoritmo proposto generaliza, para o caso estocástico, uma técnica de parametrização para problemas lineares ("operator theory"), mas não traz experiência computacional que valide o procedimento iterativo.

Fazendo um paralelo com o desenvolvido campo dos problemas de localização determinísticos, a solução do PLE apresentada nos capítulos II e III, através de um enfoque de programação mista tratada por decomposição, vem preencher uma lacuna nos modos de resolver os problemas de localização estocásticos e principalmente aqueles de grande porte.

Vale ainda ressaltar o trabalho de Jucker e Carlson [50], mas cuja formulação não usual, com considerações multicritério, não permite enquadrá-lo aqui.

#### I.5.2 - MODELOS COM CENTROS INTERMEDIÁRIOS (2 ESTÁGIOS) E VÁRIOS PRODUTOS

Até agora todos os modelos de PLA estudados eram de um único estágio (fábrica - consumidor ou armazém - consumidor) e um único produto. Porém, sistemas de distribuição mais complexos frequentemente tem necessidade de considerar dois estágios na distribuição e também trabalham com diferentes produtos. Esses modelos mais elaborados, geralmente não podem ser resolvidos pelos algoritmos simples relatados anteriormente. Ao nível das técnicas, as que encontram mais sucesso nestes casos

são a programação mista e decomposição.

Uma possível formulação de um problema de dois estágios e um único produto é

$$\text{minimizar } \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} \sum_{k \in K} c_{jk} x_{jk} + \sum_{j \in J} k_j z_j$$

sujeito a

$$\sum_{j \in J} x_{ij} \leq S_i \quad , \quad i \in I = \{1, 2, \dots, m\}$$

$$\sum_{i \in I} x_{ij} = \sum_{k \in K} x_{jk} \quad , \quad j \in J = \{1, 2, \dots, n\}$$

$$\sum_{i \in I} x_{ij} \leq C_j z_j \quad , \quad j \in J$$

$$\underline{D}_k \leq \sum_{j \in J} x_{jk} \leq \bar{D}_k \quad , \quad k \in K = \{1, 2, \dots, p\}$$

$$x_{ij} \quad , \quad x_{jk} \geq 0 \quad , \quad i \in I \quad , \quad j \in J \quad , \quad k \in K$$

$$y_j = 0, 1 \quad , \quad j \in J,$$

onde

$x_{ij}$  = fluxo de produto entre fábrica  $i$  e armazém  $j$ ,

$x_{jk}$  = fluxo de produto entre armazém  $j$  e ponto de consumo  $k$ ,

$z_j$  = 1 se o armazém  $j$  está aberto; 0 caso contrário,

$c_{ij} = c'_{ij} + T_i + V_j$  = custo unitário para transferir o produto de  $i$  para  $j$ ,

$c'_{ij}$  = custo unitário de transporte de  $i$  para  $j$ ,

$T_i$  = custo variável unitário de produção da fábrica  $i$ ,

$V_j$  = custo variável unitário por usar o armazém  $j$ ,

$c_{jk} = c'_{jk} + R_k$  = custo unitário para transferir o produto de  $j$  para  $k$ ,

- $c_{jk}$  = custo unitário de transporte entre j e k,
- $R_k$  = custo variável unitário por usar o ponto de consumo k,
- $k_j$  = custo fixo de operação do armazém j,
- $S_i$  = quantidade de produto fabricado em i,
- $\underline{D}_k$  = limitante inferior da quantidade demandada em k,
- $C_j$  = capacidade do armazém j,
- $\bar{D}_R$  = limitante superior da quantidade demandada em k.

Marks [66] aplicou um procedimento de separação e avaliação convencional para o modelo acima, onde a função objetivo com custo fixo (fixed charge) é substituída por uma função substituidora linear. A partir daí o esquema convencional de separação e avaliação é aplicado, com a avaliação sendo feita pela resolução de um problema de transbordo pelo método "out-of-kilter".

Estendendo os modelos de dois estágios para vários produtos, encontram-se os artigos de Elson [26] e Geoffrion e Graves [37]. O primeiro considera um modelo parecido com o descrito acima, mas incorpora além de vários produtos, a possibilidade de construir-se novos armazéns ou aumentar a capacidade dos já construídos. Ele usa o pacote MIP da CDC, o Ophelie Mixed, e relata tempos aceitáveis na solução de problemas de porte médio. O modelo usado por Geoffrion é diferente e lança mão de variáveis de índices quádruplos. O modelo é

$$\text{minimizar } \sum_{ijkl} c_{ijkl} x_{ijkl} + \sum_k \left[ f_k z_k + v_k \sum_{il} D_{il} y_{kl} \right]$$

$x \geq 0; y, z = 0, 1$

sujeito a

$$\sum_{kl} x_{ijkl} \leq S_{ij} \quad , \quad \text{para todo } ij$$

$$\sum_j x_{ijkl} = D_{il} y_{kl} \quad , \quad \text{para todo } ikl,$$

$$\sum_k y_{kl} = 1 \quad , \quad \text{para todo } k,$$

(1)

$$\underline{V}_k z_k \leq \sum_{il} D_{il} y_{kl} \leq \bar{V}_k z_k, \text{ para todo } k,$$

e restrições adicionais lineares sobre y e/ou z, onde

i = índice de produto,

j = índice de fábricas

k = índice de possível centro de distribuição (armazém),

l = índice de zona de consumo,

$S_{ij}$  = capacidade de produção do produto i na fábrica j,

$D_{il}$  = demanda pelo produto i na zona de consumo l,

$\underline{V}_k, \bar{V}_k$  = mínima e máxima capacidade do centro no local k,

$f_k$  = custo fixo de operação e/ou construção do centro k,

$v_k$  = custo variável unitário de operação do centro k,

$c_{ijkl}$  = custo unitário médio de produzir e transportar o produto i, da fábrica j, através do centro k, até a zona de consumo l.

$x_{ijkl}$  = fluxo do produto i, da fábrica j, através do centro k, até a zona de consumo l.

$y_{kl}$  = zero-um que será 1 se o centro de distribuição k servir a zona de consumo l, e 0 caso contrário.

$z_k$  = variável zero-um que será 1 se um centro de distribuição for colocado em k e zero caso contrário.

As possíveis restrições adicionais sobre z e y dão flexibilidade ao modelo no sentido de incorporarem, por exemplo, número máximo ou mínimo de centros funcionando, especificação de qual subconjunto de centros que devem permanecer abertos ou fechados, restrições adicionais de capacidade conjunta de vários centros que lidam com mesmo produto, etc.

O enfoque de decomposição de Benders aplicado por Geoffrion e Graves ao problema acima separa a sua resolução em duas partes, uma resolvendo um problema só em variáveis zero-um e outra resolvendo uma série de problemas simples de transporte, um para cada produto i. Como já dito na seção anterior, a aplicação de decomposição em problemas de grandes dimensões foi bem sucedida, com esforços computacionais bastante aceitáveis.

Além disso, uma vantagem desse enfoque é permitir um modelamento bastante complexo\*, além de permitir, devido ao método de decomposição empregado, interpretações econômicas que enriquecem a análise da solução do problema [62, 63].

### I.5.3 - PROBLEMA DE LOCALIZAÇÃO DINÂMICO

Nos problemas até agora estudados as decisões de localização eram feitas procurando satisfazer uma demanda agregada de produto, por exemplo, a demanda anual. Porém, quando a demanda (ou produção é altamente sazonal, o erro que se comete com a agregação pode ser grande. Nestes casos, as melhores decisões de localização devem variar ao longo do tempo, se alterando de maneira ótima para acompanhar as variações ocorridas na demanda e/ou produção.

Classicamente, o artifício matemático que estende os problemas de distribuição para o caso dinâmico consiste em repetir o número de fontes e destinos tantos quantos os períodos de tempos considerados relevantes. Os arcos de transporte só devem ligar uma fonte aos destinos situados depois no tempo. No caso geral, do PLA, esse artifício aumenta consideravelmente as dimensões do problema, podendo comprometer sua solução. Sweney e Thatan [87] desenvolveram, porém, um algoritmo dinâmico para o PLA com centros intermediários (demanda razoável), que evita a maneira clássica apresentada. Basicamente, eles calculam para cada instante de tempo uma série de configurações de localização com valores crescentes de custo. Como é imputado um custo para se mudar o estado de um armazém (fechar - abrir ou abrir - fechar), eles provam que há uma trajetória ótima de configurações no tempo. A configuração ótima e as sub-ótimas de um instante de tempo são calculadas pelo algoritmo de decomposição de Benders e a trajetória posterior calcula-se aplicando programação dinâmica.

Outro trabalho que apresenta uma solução do PLA para o caso dinâmico é o de Warszawski [90]. Ele sugere três algoritmos diferentes e aplica-os ao caso de localizar usinas produtoras de concreto numa obra civil.

---

\* em [35] é descrito um outro modelo mais complexo que o anterior

I.5.4 - EXTENSÃO PARA FUNÇÕES CUSTO NÃO-LINEARES

Até agora só foram considerados funções custo onde os custos de transporte e/ou de produção aumentam linearmente com a quantidade transportada ou produzida, como mostra a figura I.2. Em muitos casos reais estas suposições tornam o modelo pouco adequado ao problema que se deseja resolver; por exemplo, quando se consideram que o aumento no tamanho de uma unidade produtora ou da capacidade de um armazém proporcionam economias de escala. De modo geral os custos unitários de produção tendem a diminuir à medida que aumenta o nível da atividade econômica, como mostra a figura I.3.

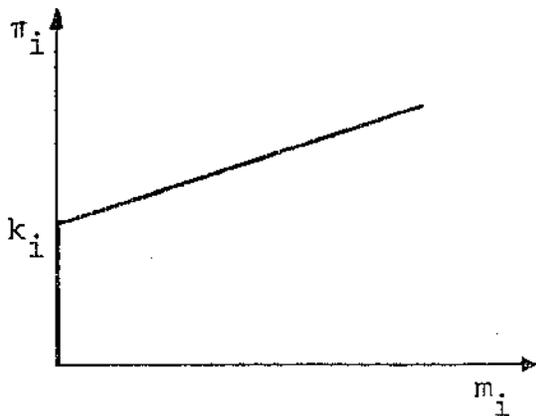


FIG. I.2

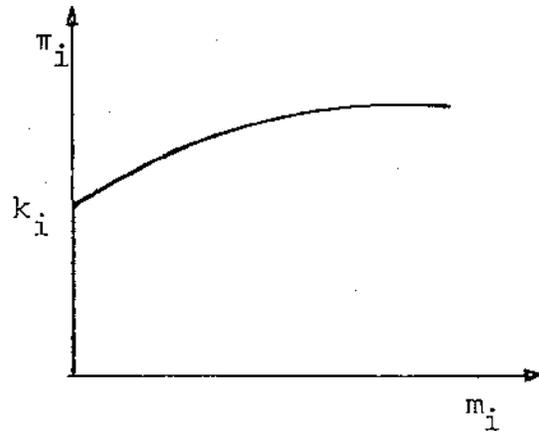


FIG. I.3

- $\pi_i$  - custo de produção na unidade  $i$
- $m_i$  - volume produzido na unidade  $i$
- $k_i$  - custo fixo da unidade  $i$

Por outro lado, em relação aos custos de transporte também podem ocorrer custos marginais decrescentes para maiores quantidades transportadas. Analogamente então, pode-se admitir os custos de transporte lineares (fig. I.4) ou côncavos (fig. I.5).

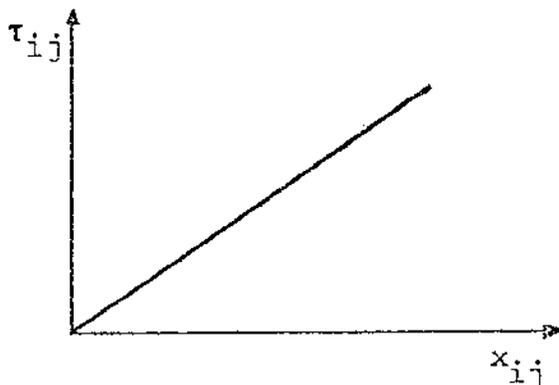


FIG. I.4

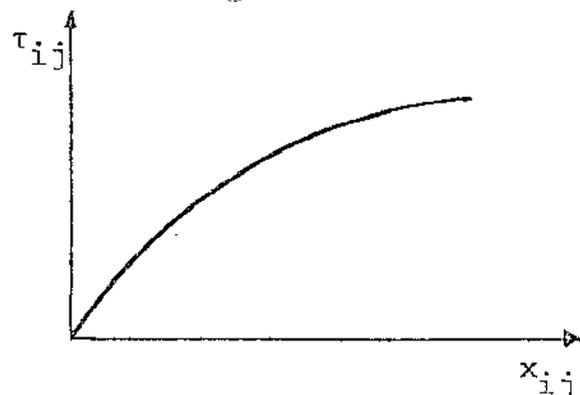


FIG. I.5

$\tau_{ij}$  - custo de transporte entre  $i$  e  $j$

$x_{ij}$  - fluxo entre  $i$  e  $j$

Genericamente a função objetiva de um problema de localização pode ser escrita como

$$\text{minimizar } \sum_{i \in I} \pi_i(m_i) + \sum_{i \in I} \sum_{j \in J} \tau_{ij}(x_{ij})$$

$$I = \{1, 2, \dots, m\}$$

$$J = \{1, 2, \dots, n\}$$

Os trabalhos de Efraymson e Ray [19], Spielberg [86], Khumawala [53], Davis e Ray [15], Ellwein [21], Gray [40], Akinc e Khumawala [1] e Marks [66], entre outros, assumem que  $\tau_{ij}(x_{ij})$  é linear e  $\pi_i(m_i)$  da forma abaixo - ("fixed-charge").

$$\pi_i(m_i) = \begin{cases} 0 & \text{se } m_i = 0 \\ k_i + \lambda_i m_i & \text{se } m_i > 0 \end{cases}$$

Porém, vários dos autores acima sugerem modificações em seus algoritmos de separação e avaliação de modo a comportarem funções custo lineares por partes. Entretanto, há trabalhos que desenvolvem dispositivos específicos para tratar funções que não sejam lineares. Dentro desta ótica pode-se citar o trabalho de Klein e Klimpel [54] que supõem os custos de produção da forma da fig. I.6 e sugerem um algoritmo de programação não-linear baseado no método de Rosen. Num estudo para localização de departamentos hospitalares, Elshafei [23] também considera custos do tipo da fig. I.6 e aplica um método de partição para chegar a solução do problema. Marks e Liebman [67] consideram custos da forma da fig. I.7 para problemas de localização com centros intermediários.

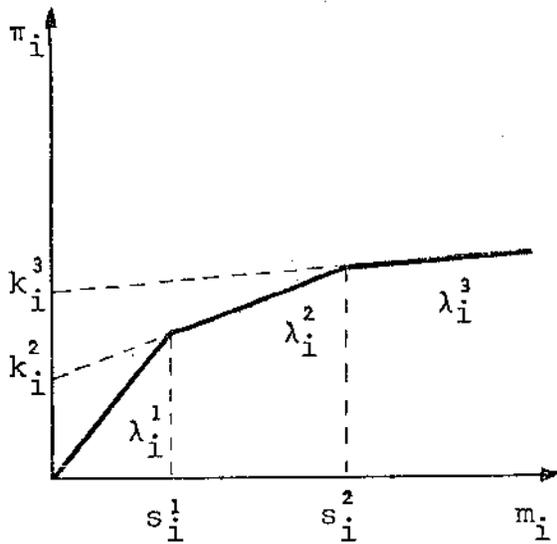


FIG. I.6

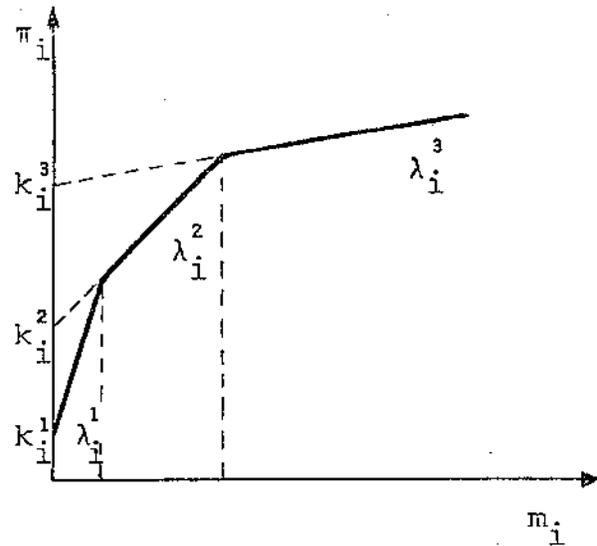


FIG. I.7

O artigo de Rech e Barton [77] pede somente que as funções  $\pi_i$  e  $\tau_{ij}$  sejam lineares por parte, não exigindo continuidade e utiliza envelopes convexos lineares por partes para substituir as funções. Daí, resolvem uma sequência finita de problemas de fluxo de custo mínimo pelo algoritmo "out-of-kilter" com a qual chegam a solução do problema original.

Sã [80] desenvolveu um algoritmo de separação e avaliação que permite que os custos de produção sejam côncavos e contínuos, como mostra a fig. I.3, sendo os custos de transporte lineares. Sua experiência computacional relata que o método se aplica a problemas até com 25 origens (variáveis inteiras), com aproximadamente 15 minutos de CPU para achar a solução ótima. O trabalho de Soland [84] admite que os custos de transporte e/ou produção sejam igualmente côncavos e seu algoritmo também de separação e avaliação caminha para a solução ótima resolvendo uma série de problemas de transporte comuns obtidos através de substituições lineares de  $\pi_i$  e  $\tau_{ij}$ . Supondo como Sã, custos côncavos para produção e lineares para transporte, Feldman, Lehrer e Ray [28] desenvolveram uma heurística para achar boas soluções.

REFERÊNCIAS BIBLIOGRÁFICAS - CAP. I

- [1] AKINC, U. e KHUMAWALA, B.M., "An Efficient Branch and Bound Algorithm for the Capacitated Warehouse Location Problem" Management Science, Vol. 23, nº 6, 585-594, (1977).
- [2] BALACHANDRAN, V. e JAIN, S., "Optimal Facility Location under Randon Demand with Geral Cost Structure", Naval Res. Log. Quartely 23, 421-435, (1976).
- [3] BALINSKY, M.L., "On Finding Integer Solutions to Linear Programs", Mathematica, Princeton, N.J., (1964).
- [4] \_\_\_\_\_ e SPIELBERG, K., "Methods for Integer Programming: Algebraic, Combinatorial and Enumerative", in J.S. Aronofsky (ed.), Progress in Operations Research, Vol. III, Wiley, N.Y., 1969.
- [5] BAUMOL, W.J. e WOLFE, P., "A Warehouse Location Problem", Operations Research 6, 252-263, (1958).
- [6] BEATTIE, D.W., "Improving the Structure of a Distribution System", Operational Research Quartely, Vol. 24, 3, 353-364, (1973).
- [7] BELLMAN, R., "An Application of Dynamic Programing to. Location - Allocation Problems", SIAM Review 7, 126-128, (1965).
- [8] BENDERS, J.F., "Partitioning Procedures for Solving Mixed-Variables Programming Problems", Numerische Mathematik, 4, 238-252, (1962).
- [9] BUZBY, B.R., "Nonlinear Distribution Problems", private communication, 1965.
- [10] BURSTALL, R.M., LEAVER, R.A. e SUSSANS, J.E., "Evaluation of transport costs for alternative factory sites - a case study", Operational Research Quarterly 13, 345-354, (1962).

- [11] CABOT, A.V., FRANCIS, R.L. e STARY, M.A., "A Network Flow Solution to a Rectilinear Distance Facility Location Problem", AIIE Transactions 2, 132-141, (1970).
- [12] CHRISTOFIDES, N. e VIOLA, P., "The Optimum Location of Multi-Centres on a Graph", Operational Research Quart. 22, 145-154, (1971).
- [13] COOPER, L., "Location - Allocation Problems", Operations Research 11, 331-343, (1963).
- [14] ———, "A Randon Locational Equilibrium Problem", Journal of Regional Science, 14, 47-54, (1974).
- [15] DAVIS, P.S. e RAY, T.L., "A Branch-Bound Algorithm for the Capacitated Facilities Location Problem", Naval Res. Log. Quarterly 16, 331-344, (1969).
- [16] DEARING, D.M. e FRANCIS, R.L., "A Minimax Location Problem on a Network", Transportation Science 8, 333-343, (1974)
- [17] ——— e ———., "A Network Flow Solution to a Multi-facility Location Problem Using Rectilinear Distances", Transportation Science 8, 126-141, (1974).
- [18] DREZNER, Z. e WESOLOWSKY, G.O., "Facility Location on a Sphere", J. Op. Res. Soc. 29, n° 10, 997-1004, (1978).
- [19] EFROYMSON, M.A. e RAY, T.L., "A Branch - Bound Algorithm for Plant Location", Operations Research 14, 361-368, (1966).
- [20] EILON, S., WATSON-GANDY, C.D.T. e CHRISTOFIDES, N., Distribution Management: Mathematical Modelling & Practical Analysis, Griffim, London, 1971.
- [21] ELLWEIN, L.B., "Fixed charge Location - Allocation Problems with Capacity and Configuration Constraints", Technical Rep. N° 70-2, Department of Ind. Engineering, Stanford University, (1970).

- [22] ——— e GRAY, P., "Solving Fixed Charge Location - Allocation Problems with Capacity and Configuration Constraints", AIIE Transactions, III, 4, 290-298, (1971).
- [23] ELSHAFEI, A.N., "An Approach to Locational Analysis", Operational Research Quarterly, Vol. 26, n° 1, 167-181, (1975).
- [24] EL-SHAIEB, A.M. "The single source Weber problem-survey and extensions", J. Opl. Res. Soc. Vol. 29, n° 5, 469-476, (1978).
- [25] ———, "A New Algorithm for Locating Sources Among Destinations", Management Science 20, 221-231, (1973).
- [26] ELSON, D.G., "Site Location via Mixed-Integer Programming", Operational Research Quarterly, 23, n° 1, 31-43, (1972)
- [27] ELZINGA, J., HEARN, D., e RANDOLPH, W.D., "Minimax Multi-facility Location with Euclidian Distances", Transportation Science 10, 321-336, (1976).
- [28] FELDMAN, E., LEHRER, F. e RAY, T., "Warehouse Location Under Continuous Economies of Scale", Management Science, Vol. 12, p. 670, (1966).
- [29] FRANCIS, R.L. e GOLDSTEIN, J.M., "Location Theory: A Selective Bibliography", Operations Research 22, 400-410, (1974).
- [30] FRANK, H., "A Note on a Graph Theoretic Game of Hakimi's", Operations Research 15, 567-570, (1967).
- [31] GALVÃO, R.D., "The Optimal Location of Facilities on a Network", Ph.D. Thesis, Imperial College, London, 1977.
- [32] GARFINKEL, R.S., NEEBE, A.W., RAO, M.R., "The m-Center Problem: Minimax Facility Location", Management Science 23, 1133-1142, (1977)

- [33] ———, NEEBE, A.W. e RAO, M.R., "An Algorithm for the m-Median Plant Location Problem", *Transportation Science* 8, 217-236, (1974).
- [34] GEOFFRION, A.M., "A Guide to Computer-Assisted Methods for Distribution System Planning", *Sloan Management Review*, winter 1975, 17-41.
- [35] ———, GRAVES, G.W. e LEE, S., "Strategic Distribution System Planning: A Status Report", in A. Hax (ed.), *Studies in Operations Management*, North Holland/Elsevier, 1978.
- [36] ———, e MCBRIDE, R.D., "The Capacited Plant Location Problem with Additional Constraints", *Joint National Meeting of AIIE, ORSA and TIMS*, Atlantic City, N.J., 1972.
- [37] ——— e GRAVES, G.W., "Multicommodity Distribution System Design by Benders Decomposition", *Management Science*, 20, 822-844, (1974).
- [38] GOLDMAN, A.J., "Minimax Location of a Facility in a Network", *Transportation Science* 6, 407-418, (1972).
- [39] ———, "Optimum Locations for Centers in Networks", *Transportation Science* 3, 352-360, (1969).
- [40] GRAY, P., "Exact Solution of the Site Selection Problem by Mixed Integer Programming", in *Applications of Mathematical Programming Techniques*, E.M. L. Beale (ed.) - Am. Elsevier Publ. Co. New York, 1970.
- [41] HAKIMI, S.L. "Optimal distribution of switching centers and medians of a graph", *Operations Research* 12, 450-459, (1964).
- [42] ———, "Optimum distribution of switching centers in a communication network and some related graph theoretic problems", *Operations Research* 13, 462-475, (1965).

- [43] ———, e MAHESWARI , S.N., "Optimum Locations of Centers in Networks", Operations Research 20, n° 5 967-973, (1972).
- [44] ———, SCHMEICHEL, E.F. e PIERCE, J.G., "On p-Centers in Networks", Transportation Science 12, 1-15, (1978).
- [45] HALFIN, S., "On Finding the Absolute and Vertex Centers of a Tree with Distances", Transportation Science 8, 75-77, (1974).
- [46] HANDLER, G.Y., "Minimax Network Location: Theory and Algorithms", Tech. Rep. 107, Op. Res. Center, M.I.T., (1974).
- [47] ———, "Minimax Location of a Facility in a Undirected Tree Graph", Transportation Science 7, 287-293, (1974)
- [48] ———, "Finding Two-Centers of a Tree: The Continuous Case", Transportation Science 12, n° 2, 93-106, (1978).
- [49] JÄRVINEN, P, RAJALA, J., e SINERVO, H., "A Branch-and-Bound Algorithm for Seeking the p-Median", Operations Research 20, 173-178, (1972).
- [50] JUCKER, J.V. e CARLSON , R.C., "The Simple Plant-Location Problem under Uncertainty", Operations Research, Vol. 24, n° 6, 1045-1055, (1976).
- [51] JUEL, H. e LOVE, R.F., "An Efficient Computational Procedure for Solving the Multifacility Rectilinear Facilities Location Problem", Operational Research Quarterly 27, 697-703, (1976).
- [52] KHUMAWALA, B.M., "An Efficient Branch and Bound Algorithm for the Warehouse Location Problem", Management Science 18, B718-B731, (1972).
- [53] ———, "An Efficient Branch-and-Bound Algorithm for the Warehouse Location Problem", Management Science 18, B718-B731, (1972).

- [54] KLEIN, M. e KLIMPEL, R.R., "Applications of Linearly Constrained Nonlinear Optimizations to Plant Location and Sizing", J. Indust. Engng. 18, 90, (1967).
- [55] KUEHN, A.A. e HAMBURGER, J.J., "A Heuristic Program for Locating Warehouses", Management Science 9, 643-666, (1963).
- [56] KUHN, H.W. e KUENNE, R.E., "An Efficient Algorithm for the Numerical Solution of the Generalized Weber Problem in Spatial Economics", J. Regional Science 4, 21-34, (1962)
- [57] LAWRENCE, R.M. e PENGILLY, P.J., "The Number and Location of Depots Required for Handling Products for Distribution to Retail Stores in South-East England" Operational Research Quarterly 20, 23-32, (1969).
- [58] LEA, A.C., "Location - Allocation Systems: An Annotated Bibliography," Discussion Paper n° 13, Department of Geography, University of Toronto, Canada, (1973).
- [59] LEBLANC, L.J., "A Heuristic Approach for Large Scale Discrete Stochastic Transportation - Location Problems" Comp. & Maths. with Appls., Vol. 3, 87-94, (1977).
- [60] LEVY, J., "An Extended Theorem for Location on a Network", Operational Research Quarterly 18, 433-442, (1967).
- [61] LOVE, R.F., WESOLOWSKY, G.O. e KRAEMER, S.A., "A Multi-facility Minimax Method for Euclidian Distances", Int. J. Prod. Res. 11, 37-45, (1973).
- [62] LUNA, H.P.L., "Les Techniques de Decomposition-Coordination dans les Modèles Economiques d'Optimisation", Thèse d'Etat, Un. P. Sabatier, 1978.
- [63] ———, "Economic Interpretation of the Benders Decomposition Technique Applied to Location Problems", 2<sup>nd</sup> Int. Conf. Dyn. Model Control of Nat. Economies, IFAC/IFORS/IIASA, Wien, 1977.

- [64] MANNE, A.S., "Plant Location under Economies - of - Scale - Decentralization and Computation", Management Science 11, 213-235, (1964).
- [65] MARANZANA, F.E., "On the Location of Supply Points to Minimize Transport Costs", Operational Research Quarterly 15, 261-270, (1964).
- [66] MARKS, D.H., "Facility Location and Routing Models in Solid Waste Collection Systems", Ph.D. Thesis, The Johns Hopkins Univ., 1969.
- [67] ——— and LIEBMAN, J.C., "Mathematical Analysis of Solid Waste Collection", Report (SW-5 rg), V.S. Dept. of H.E.W., Washington - DC, (1970).
- [68] MARSTEN, R.E., "An Algorithm for Finding Almost All of the Medians of a Network", Discussion Paper n° 23, Northwestern University, (1972).
- [69] MCGINNIS, L.F., "A survey of recent results for a class of facilities location problems", AIIE Transactions, Vol. 9, n° 1, 11,18, (1977).
- [70] ——— , e WHITE, J.A., "A Single Facility Rectilinear Location Problem with Multiple Criteria", Transportation Science 12, n° 3, 217-231, (1978).
- [71] MINIEKA, E., "The m-Center Problem", SIAM Review 12, 138-139, (1970).
- [72] MONTEROSSO, C.D.B., "Um Método Heurístico para a Localização e Dimensionamento de Armazéns em Sistemas de Grande Porte Considerando Economias de Escala", Tese de Mestrado - COPPE, UFRJ, 1977.
- [73] MORRIS, J.G., "A Linear Programming Approach of the Solution of Constrained Multi-Facility Minimax Location Problems where Distances are Rectangular", Operational Research Quarterly 24, 419-439, (1973)

- [74] OSTRESH, JR., L.M., "Convergence and Descent in the Fermat Location Problem", *Transportation Science* 12, nº 2, 153-164, (1978).
- [75] PRITSKER, A.B. e GHARE, P.M., "Locating New Facilities with Respect to Existing Facilities", *AIIE Transactions* 2, 290-297, (1970).
- [76] RAND, G.K., "Methodological Choices in Depot Location Studies", *Operational Research Quarterly*, Vol. 27, 1, 241-249, (1976).
- [77] RECH, P. e BARTON, L.G., "A Non-Convex Transportation Algorithm", in *Applications of Mathematical Programming Techniques*, E.M.L. Beale (ed.). Am. Elsevier Publ. Co., New York, 1970.
- [78] REVELLE, C., MARKS, D. e LIEBMAN, J.C., "An analysis of private and public sector location models", *Management Science* 16, nº 11, 692-707, (1970).
- [79] ———, e SWAIN, R.W., "Central Facilities Location", *Geogr. Anal.*, Vol. 2, 30-42, (1970).
- [80] SÁ, G., "Branch-and-Bound and Approximate Solutions to the Capacited Plant-Location Problem", *Operations Research* 17, 1005-1016, (1969).
- [81] SCOTT, A.J., "Location - Allocation Systems: A Review", *Geograph. Anal.* 2, 95-119, (1970).
- [82] SEPPALA, Y., "On a Stochastic Multi-Facility Location Problem" *AIIE Transactions*, Vol. 7, nº 1, 56-62 (1975).
- [83] SHERALI, H.D. e SHETTY, C.M., "A Primal Simplex Based Solution Procedure for the Rectilinear Distance Multifacility Location Problem", *J. Opl. Res. Soc.* 29 nº 4, 373-381, (1978).

- [84] SILVA, R.V. e LINS, M.L., "Localização de Depósitos entre Pontos de Produção e Consumo", IX Simpósio da SOBRAPO, Brasília, 1978.
- [85] SOLAND, R.M., "Optimal Facility Location with Concave Costs", Operations Research 22, 373-382, (1974).
- [86] SPIELBERG, K., "An Algorithm for the Simple Plant Location Problem with Some Side Conditions", Operations Research 17, 85-111, (1969).
- [87] SWENEY, D.J. e TATHAN, R.L., "An Improved Long-Run Model for Multiple Warehouse Location", Management Science 22, 748-758, (1976).
- [88] TEITZ, M.B. e BART, P., "Heuristic Method for Estimating the Generalized Vertex Median of a Graph", Operations Research 16, 955-961, (1968).
- [89] TOREGAS, C., SWAIN, R., REVELLE, C. e BERGMAN, L., "The Location of Emergency Service Facilities", Operations Research 19, 1363-1373, (1971).
- [90] WARSZAWSKY, A., "Multi-dimensional Location Problems", Operational Research Quarterly 24, nº 2, 165-179, (1973).
- [91] WATSON-GANDY, C.D. e EILON, S., "The Depot Siting Problem with Discontinuous Delinery Cost", Operational Research Quarterly 23, 277, (1972).
- [92] WEBER, A., "Über den Standort des Industrien", Tübingen, (1909), traduzido para o inglês como "Alfred Weber's Theory of the Location of Industries" por Friedrich, C.J., University of Chicago Press, (1929).
- [93] WEISZFELD, E., "Sur le Point Pour Lequel la Somme des Distances de n Points Donnés est Minimum", Tôhoku Math. J. 43, 355-386, (1937),

- [94] WENDELL, R.E. e HURTER JR., A.P., "Optimum Locations on a Network", Transportation Science 7, 19-33, (1973).
  
- [95] WESOLOWSKY, G.O., "Rectangular distance location under the minimax optimality criterion", Transportation Science 6, 103-113, (1972).
  
- [96] ————— , e LOVE R.F., "The Optimal Location of a New Facilities Using Rectangular Distances", Operations Research 19, 124-130, (1971).

## CAPÍTULO II

### PROBLEMA DE LOCALIZAÇÃO ESTOCÁSTICO RESOLVIDO PELA DECOM- POSIÇÃO DE BENDERS GENERALIZADA

#### II.1 - INTRODUÇÃO

O conteúdo principal deste capítulo é a resolução do problema de localização com demandas estocásticas. Como levantado no item I.5.1 as propostas de resolução de tal problema não se apresentam satisfatórias do ponto de vista da solução global e da dimensão do problema tratado. Aqui é sugerida a especialização da técnica de otimização de decomposição de Benders generalizada (DBG), que permite com um modelamento bastante geral, encontrar a solução ótima global do problema, além de abrir a possibilidade de tratamento de problemas de grande porte.

Neste capítulo é apresentado um resumo da programação matemática de grande porte e como as técnicas de decomposição - e em particular a de Benders - se enquadram nela. A seguir, há um resumo da DBG e levados pela necessidade de criar um problema equivalente determinístico, são apresentadas as duas maneiras clássicas de fazê-lo. Só a primeira maneira é usada na resolução apresentada, mas a outra maneira sugere uma interessante variação que é explorada no capítulo IV.

#### II.1.1 - PROGRAMAÇÃO MATEMÁTICA DE GRANDE PORTE

Paralelamente ao grande desenvolvimento experimentado nos últimos vinte anos no campo da otimização e programação matemática, verificou-se uma crescente preocupação com uma classe particular de problemas, que devido às suas dimensões não eram satisfatoriamente resolvidos pelas técnicas clássicas. Foi então, que no final da última década, técnicas já desenvolvidas foram sendo adaptadas e outras sendo criadas para tratar esses problemas de grande porte. O princípio de decomposição de

Dantzig-Wolfe dá início, em 1960, a uma série de algoritmos que explorando a estrutura particular de alguns problemas conseguem resolvê-los com grande vantagem em termos de tempo de processamento e/ou capacidade de memória. Dentre outras, algumas estruturas especiais que podem ser exploradas são, por exemplo, a forma bloco diagonal do conjunto de restrições, problemas não-lineares ou não-convexos que tornam-se lineares ou convexos quando algumas variáveis são fixadas, ou ainda quando estão presentes funções aditivamente separáveis.

Do ponto de vista computacional da programação matemática de grande porte pode-se dizer que as pesquisas tem se concentrado em duas categorias distintas: trabalhos que procuram melhorar a eficiência computacional de métodos já conhecidos, e trabalhos que desenvolvem novas técnicas. Na primeira categoria, as maiores contribuições se deram sobre o método Simplex de programação linear: algoritmos de fluxos em rede, forma produto da inversa, triangularização, técnicas de limitação superior ("upper bounding"), geração de colunas, etc. Na segunda categoria, tornaram-se clássicos os métodos de Dantzig-Wolfe, Rosen, Benders, Aproximação Tangencial, Uzawa, etc. Cabe notar que o desenvolvimento dessas técnicas foi possível graças ao avanço do suporte teórico necessário, principalmente da teoria da dualidade, em que as maiores contribuições partiram de Rockafellar, Lasdon e Geoffrion.

Porém, é fundamental repetir a importância da estrutura especial dos problemas a que se referem a programação matemática de grande porte. Somente o tamanho dos problemas não serve para caracterizar a aplicabilidade destas técnicas específicas. Por isso, talvez, o nome de "programação matemática estrutural" fizesse maior justiça para essa classe de problemas considerados.

A literatura classifica diversos tipos de estruturas: multidivisional, combinatorial, dinâmica e estocástica. Multidivisional diz respeito àqueles problemas naturalmente subdivididos (subproblemas) como reservatórios de sistemas hidráulicos, divisões de uma firma, setores de uma economia. Os problemas de estrutura combinatorial tem geralmente grande número de variáveis, devido ao grande número de possibilidades (alternativas) presentes, por exemplo nos problemas de fluxos. Problemas

dinâmicos são grandes quando as restrições têm que ser repetidas para os diversos instantes de tempo ou tornam uma estrutura particular que permite a sua divisão temporal. O caráter estocástico de alguns problemas também pode induzir estruturas especiais ou aumentar muito suas dimensões.

Os métodos da programação matemática de grande porte, explorando a estrutura particular de um problema, geralmente o decompõe em outros mais simples de se resolver. O mecanismo que faz essa decomposição pode ser visto como se primeiro se efetua-se uma manipulação que procura uma forma alternativa para o problema original e que seja mais simples de resolver. Manipulações geralmente dão origem aos problemas mestres. A seguir, uma estratégia de resolução visa gerar uma sequência de problemas simplificados para resolução do problema mestre. É onde aparecem os subproblemas de solução mais fácil e mais rápida.

As manipulações mais conhecidas são dualização, projeção, linearização externa e interna. Como estratégias, aparecem a restrição, relaxação, direções factíveis e otimização por partes. Um método geralmente caracteriza-se por uma combinação de uma (ou mais) manipulação acima com uma estratégia. Para um aprofundamento, ver o trabalho de Geoffrion [6] que é uma tentativa de dar uma visão unificada da programação matemática de grande porte.

## II.2 - DECOMPOSIÇÃO DE BENDERS GENERALIZADA (DBG)

A DBG desenvolvida por Geoffrion [7], é uma extensão para o caso não-linear do algoritmo de Benders "linear" (no sentido já explicado no Cap. I), e é uma decorrência natural do desenvolvimento recente da teoria da dualidade para problemas não-lineares. A exemplo de seu ancestral, a DBG trata problemas onde intervêm variáveis "complicantes", no sentido que quando elas são fixadas temporariamente o problema resultante se torna de solução mais simples. Casos típicos onde ela se aplica com vantagens são:

- quando o problema resultante se separa em vários subproblemas independentes.

- quando o problema resultante assume uma estrutura especial e conhecida (problema de transporte, por exemplo) para o qual há solução eficiente.
- quando o problema original é não-convexo, mas o resultante é, permitindo com isso que as não-convexidades sejam tratadas separadamente.

Apresenta-se aqui um resumo com interpretações do método, que trata com problemas da forma geral

$$\begin{array}{ll} \text{minimizar} & \phi(x,z) \\ & x,z \end{array} \quad (1)$$

sujeito a  $G(x,z) \leq 0$ ,  $x \in X$ ,  $z \in Z$ ,

onde  $z$  é um vetor do  $R^{n_2}$  que faz o papel das variáveis complicantes e  $x$  é um vetor do  $R^{n_1}$ .  $G$  é uma função vetorial de  $m$  componentes definida em  $X \times Z$ . O algoritmo é desenvolvido fazendo-se uma manipulação fundamental de projeção e a seguir duas outras manipulações de dualização.

A projeção de (1) sobre o espaço dos  $z$  significa

$$\begin{array}{ll} \text{minimizar} & v(z) \\ & z \end{array} \quad \text{sujeito a} \quad z \in Z \cap V \quad (2)$$

onde

$$v(z) \equiv \inf_x \phi(x,z) \quad \text{sujeito a} \quad G(x,z) \leq 0, \quad x \in X \quad (3)$$

e

$$V \equiv \left\{ z / G(x,z) \leq 0 \text{ para algum } x \in X \right\} . \quad (4)$$

Esta projeção é justificada pelo fato que resolver (2) é mais fácil, em princípio, que resolver (1). A introdução

de  $V$  é necessária para garantir a existência de  $v(z)$ , o que pode ser visto na figura abaixo.

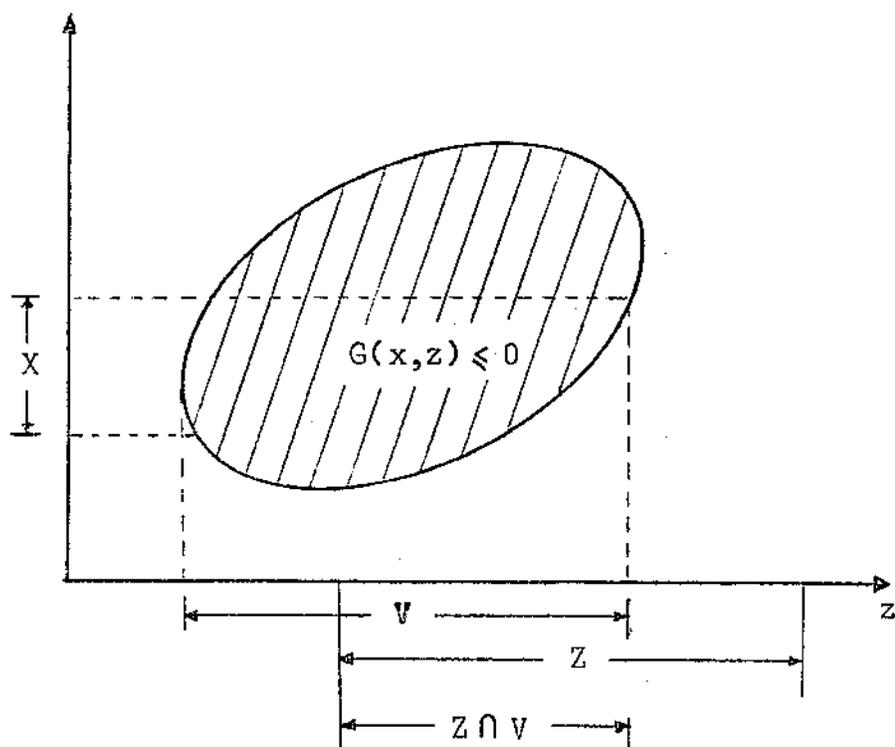


FIG. II.1

A expressão (3) pode ser obtida resolvendo-se um problema de otimização

$$\begin{aligned} &\text{minimizar } \phi(x,z) \text{ sujeito a } G(x,z) \leq 0 . & (5) \\ &x \in X \end{aligned}$$

O problema acima é o subproblema do algoritmo e posteriormente será explicado o papel que exercerá no procedimento iterativo.

Após a manipulação de projetar (1) sobre o espaço  $z$  (que visa separar o problema nos dois conjuntos de variáveis), a tarefa de resolver (2) é equivalente a resolver (1). Duas outras manipulações são feitas para deixar (2) numa forma mais tratável:

a) apelar para a representação dual de  $V$  de modo a facilitar a verificação da satisfação das restrições de (2).

b) apelar para a forma dual de  $v(z)$  como um máximo

ponto a ponto de uma coleção de funções que a englobe.

As três manipulações discutidas são precisadas respectivamente através dos seguintes teoremas:

Teorema 1 - Projeção

O problema (1) é infactível ou tem valor ótimo ilimitado se e somente se ocorre o mesmo para (2). Se  $(x^*, z^*)$  é solução ótima de (1), então  $z^*$  é solução ótima de (2). Se  $z^*$  é ótimo de (2) e  $x^*$  fornece o ínfimo de (3) com  $z = z^*$ , então  $(x^*, z^*)$  é ótimo de (1). Se  $\bar{z}$  é  $\epsilon_1$  - ótimo em (2) e  $\bar{x}$  é  $\epsilon_2$  - ótimo em (5), então  $(\bar{x}, \bar{z})$  é  $(\epsilon_1 + \epsilon_2)$  - ótimo em (1). A prova está em [7].

Teorema 2 - Representação de V

Assumindo que  $X$  é convexo e não-vazio e que  $G$  é convexa em  $X$  para cada  $z \in Z$  fixado; e assumindo ainda que o conjunto  $K_Z \equiv \{ k \in \mathbb{R}^m / G(x, z) \leq k \text{ para algum } x \in X \}$  seja fechado para cada  $z \in Z$ , então um ponto  $\bar{z} \in Z$  está também no conjunto  $V$  se e somente se  $\bar{z}$  satisfizer o sistema (infinito)

$$\left[ \inf_{x \in X} \lambda G(x, z) \right] \leq 0 \quad \text{para todo } \lambda \in \Lambda, \quad (6)$$

onde

$$\Lambda \equiv \left\{ \lambda \in \mathbb{R}^m / \lambda_i \geq 0 \quad \text{e} \quad \sum_{i=1}^m \lambda_i = 1 \right\}$$

A prova está em [7].

Teorema 3 - Forma dual de  $v(z)$

Assuma que  $X$  é um conjunto convexo não-vazio e que

$\phi$  e  $G$  sejam convexas em  $X$  para cada  $z \in Z$ . Assuma que para cada  $\bar{z} \in Z \cap V$  pelo menos uma das três seguintes condições ocorra:

a)  $v(\bar{z})$  é finito e (5) possui um vetor multiplicador ótimo\*.

b)  $v(\bar{z})$  é finito,  $G(x, \bar{z})$  e  $\phi(x, \bar{z})$  são contínuas em  $X$ ,  $X$  é fechado e o conjunto de soluções  $\epsilon$ -ótimas de (5) para  $z = \bar{z}$  é não-vazio e limitado para algum  $\epsilon \geq 0$ .

c)  $v(\bar{z}) = -\infty$

Então o valor ótimo de (5) é igual ao seu dual em  $Z \cap V$ , ou seja,

$$v(z) \equiv \sup_{u \geq 0} \left[ \inf_{x \in X} \phi(x, z) + uG(x, z) \right]$$

para todo  $z \in Z \cap V$ .

A prova também está em [7]. A hipótese (a) é satisfeita quando se verifica alguma condição de regularidade (qualificação de restrição) para o problema, que sendo imposta somente para evitar casos patológicos, faz de (a) uma hipótese usualmente satisfeita. A hipótese (c) considera o caso do dual ser inactível, onde toma-se o supremo como  $-\infty$ .

Verificadas as hipóteses dos teoremas 2 e 3, o problema (2) se escreve

$$\text{minimizar } \left[ \sup_{u \geq 0} \left[ \inf_{x \in X} \phi(x, z) + uG(x, z) \right] \right] \text{ sujeito a (6)} \\ z \in Z$$

---

\* Para perfeita compreensão desses teoremas e de suas provas é essencial o conhecimento da teoria da dualidade em programação não-linear apresentada em Lasdon [9], ou em Geoffrion [8] no seu aspecto estritamente convexo.

Usando ainda o fato que um supremo pode ser representado pelo menor limitante superior, chega-se após a sequência de manipulações, ao seguinte problema mestre:

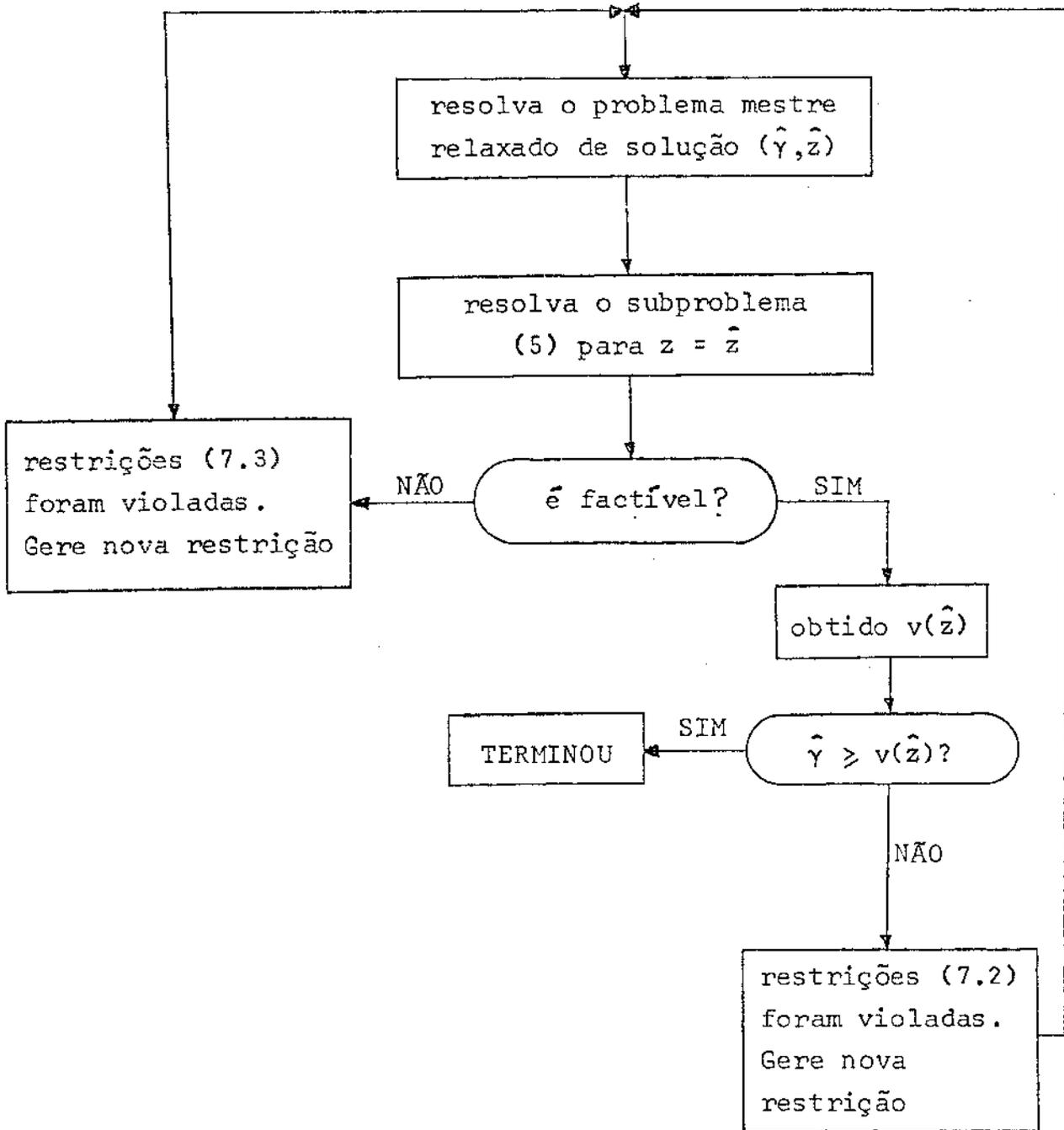
$$\begin{aligned} & \text{minimizar } \gamma && (7.1) \\ & \gamma, z \in Z \end{aligned}$$

sujeito a

$$\gamma \geq \underset{x \in X}{\text{ínfimo}} \left[ \phi(x, z) + uG(x, z) \right], \text{ para todo } u \geq 0 \quad (7.2)$$

$$\underset{x \in X}{\text{ínfimo}} \left[ \lambda G(x, z) \right] \leq 0, \text{ para todo } \lambda \in \Lambda \quad (7.3)$$

A forma a que o problema mestre acima foi levado a assumir (com infinitas restrições) sugere a aplicação da estratégia de resolução da relaxação, ou seja, ignoram-se quase todas as restrições de (7) e resolve-se o problema relaxado. Se a solução resultante não satisfizer todas as restrições ignoradas, gerar e adicionar uma ou mais restrições que foram violadas e resolver novamente, até que a solução relaxada satisfaça todas as restrições ignoradas. Suponha, então que  $(\hat{z}, \hat{\gamma})$  é um ótimo relaxado de (7). Como testar a factibilidade de tal solução com respeito às restrições de (7)? E se ocorrer infactibilidade como gerar uma restrição violada? A resposta a estas perguntas é dada pela resolução do subproblema (5). Ou seja, resolva (5) para  $z = \hat{z}$ ; se a solução achada for factível, então  $\hat{z} \in V$  e pelo teorema 2, satisfaz (7.3). A recíproca é válida. Suponha ainda que a solução de (5) para  $z = \hat{z}$  é factível e  $\hat{\gamma} \geq v(\hat{z})$ , então pelo teorema 3,  $(\hat{\gamma}, \hat{z})$  satisfaz (7.2). A recíproca também é válida. Até o momento, pode-se escrever o seguinte diagrama simplificado do algoritmo (sem considerações de como inicializar).



Vejamos como gerar uma restrição que foi violada. Um vetor  $\hat{u} \geq 0$  ou  $\lambda \in \Lambda$  gera uma restrição violada se verificarem respectivamente as relações

$$\gamma < \inf_{x \in X} \left[ \phi(x, \hat{z}) + \hat{u} G(x, \hat{z}) \right] \quad (8.1)$$

ou

$$\inf_{x \in X} \left[ \hat{\lambda} G(x, \bar{z}) \right] > 0 \quad (8.2)$$

Sabe-se da teoria da dualidade que se se tomar  $\hat{u}$  como o vetor multiplicador ótimo do subproblema (5) para  $z = \bar{z}$  (8.1) não só será satisfeita como seu lado direito alcançará máximo valor (devido à função dual ser côncava e o vetor multiplicador ótimo ser solução ótima do dual, cujo valor é igual a solução ótima do primal). Com isso conclui-se que quando o subproblema for factível e tiver solução ótima finita o seu vetor multiplicador ótimo gera a restrição mais violada dentre as restrições (7.2). No caso de (5) ser infactível, a utilização de um algoritmo de tipo dual facilita a obtenção de  $\hat{\lambda}$ .

O algoritmo da DBG pode então ser formalizado num procedimento iterativo. Antes é conveniente definir as seguintes funções auxiliares:

$$L^*(z, u) = \inf_{x \in X} \left[ \phi(x, z) + u G(x, z) \right], \quad z \in Z, \quad u \geq 0 \quad (9.1)$$

$$L_*(z, \lambda) = \inf_{x \in X} \left[ \lambda G(x, z) \right], \quad z \in Z, \quad \lambda \in \Lambda \quad (9.2)$$

Passo 1 - Seja um ponto  $\bar{z} \in Z \cap V$  conhecido. Resolva o subproblema (5) para  $z = \bar{z}$  e obtenha um vetor multiplicador ótimo  $\bar{u}$  e a função  $L^*(z, \bar{u})$ . Faça  $p = 1$   
 $q = 0$ ,  $u^1 = \bar{u}$ ,  $LS = v(\bar{z})$ .

Selecione um parâmetro de convergência  $\epsilon > 0$

Passo 2 - Resolva o problema mestre relaxado corrente

minimizar  $\gamma$  sujeito a  
 $z \in Z, \gamma$

$$\gamma \geq L^*(z, u^j) \quad , \quad j = 1, 2 \dots, p \quad (10)$$

$$L_*(z, \lambda^j) \leq 0 \quad , \quad j = 1, 2 \dots, q$$

Seja  $(\hat{z}, \hat{\gamma})$  uma solução ótima;  $\hat{\gamma}$  é um limitante inferior do valor ótimo de (1).

Se  $LS \leq \hat{\gamma} + \epsilon$ , TERMINOU.

Passo 3 - Resolva o subproblema (5) para  $z = \hat{z}$ . Um dos seguintes casos vai ocorrer:

Passo 3A - O valor de  $v(\hat{z})$  é finito. Se  $v(\hat{z}) \leq \hat{\gamma} + \epsilon$ , TERMINOU. Senão determine o vetor multiplicador ótimo  $\hat{u}$  e a função  $L^*(z, \hat{u})$ . Incremente  $p$  de 1 e faça  $u^p = \hat{u}$ . Se  $v(\hat{z}) < LS$ , faça  $LS = v(\hat{z})$ .  $LS$  é pois um limitante superior do valor ótimo de (1). Volte ao PASSO 2.

Passo 3B - O subproblema é infactível. Determine  $\hat{\lambda} \in \Lambda$  satisfazendo (8.2) e a função  $L_*(z, \hat{\lambda})$ . Aumente  $q$  de 1 e faça  $\lambda^q = \hat{\lambda}$ . Volte ao PASSO 2.

Algumas observações para efeito de implementação computacional fazem-se necessárias. O problema mestre relaxado (10) é sempre factível, mas pode ocorrer uma solução ilimitada,

que pode ser evitada tomando  $Z$  como um conjunto compacto (hipótese frequente na prática, por exemplo, nos problemas de localização onde  $Z$  é um conjunto discreto). A inicialização através do passo 1, escolhendo um  $\bar{z} \in Z \cap V$ , possibilita através de um bom conhecimento prático do problema, uma boa escolha inicial. Caso tal ponto factível seja desconhecido, é necessário lançar mão de um procedimento tipo fase-um capaz de obtê-lo. Uma outra inicialização, muito útil em problemas de grande dimensão, é resolver vários subproblemas para um subconjunto dos  $z \in Z \cap V$  que suspeita-se serem bons candidatos, e só depois executar o problema mestre relaxado.

No passo 2 resolve-se a cada iteração um problema mestre que é uma versão cada vez menos relaxada do problema (7) e cujo valor ótimo é  $\hat{\gamma}$ . Então  $\hat{\gamma}$  é um limitante inferior do problema (2), e sua sequência de valores é monótona não-decrescente.

Deve-se notar que quando se resolve um subproblema com  $z = \hat{z}$ , a solução ótima obtida -  $\hat{x}$  - fornece uma solução  $(\hat{x}, \hat{z})$  de (1) cujo valor é  $v(\hat{z})$ . Dentro de LS é sempre colocada a melhor avaliação já feita pelo subproblema das diferentes propostas  $\hat{z}$  e por isso fornece o melhor limitante superior até agora conhecido do problema (2). Note que a sequência de valores de LS não é necessariamente monótona não-crescente.

À luz desses limitantes, o critério de parada da DBG é quando eles se encontrarem, respeitando uma tolerância. Isto pode ocorrer tanto no passo 2 - avaliação do limitante inferior -, quanto no passo 3A - avaliação do limitante superior.

Uma outra observação importante é a não necessidade da função  $L_*$  e do passo 3B caso a factibilidade do subproblema seja sempre garantida, qualquer que seja  $z \in Z$  selecionado - o que, aliás, ocorre também nos problemas de localização -.

Um detalhe até agora não comentado é que a avaliação das funções  $L^*$  e  $L_*$  envolve um problema de otimização, que pode representar um problema crucial na solução de (1). Porém em muitos casos, tais funções podem ser obtidas sem qualquer esforço adicional. É o caso quando as funções  $\phi$  e  $G$  são linearmente superáveis, ou seja

$$\phi(x, z) = \phi_1(x) + \phi_2(z)$$

$$G(x,z) = G_1(x) + G_2(z)$$

Como será mostrado no item II.5, esta propriedade (que chamaremos de propriedade P) que está presente no problema de localização estocástico, será aplicada de tal modo que a obtenção da função  $L^*$  será feita usando simplesmente resultados da resolução do subproblema da iteração anterior.

A prova de convergência da DBG é imediata no caso em que  $Z$  é um conjunto discreto: como um  $\bar{z}$  não pode ser repetido a cada execução do passo 2 e  $Z$  tem um número finito de elementos, a convergência pode ser garantida num número finito de iterações. A prova para o caso geral está em [7].

## II.3 - PROGRAMAÇÃO LINEAR ESTOCÁSTICA: EQUIVALENTES DETERMINÍSTICOS

### II.3.1 - INTRODUÇÃO

No presente item, levados pela necessidade de fazer uso da programação estocástica para resolver o PLE, faremos um breve resumo das técnicas mais usadas e em especial desenvolvemos os fundamentos teóricos do problema de dois estágios e da restrição de risco. Esses dois enfoques da programação estocástica linear serão usados, o primeiro no presente capítulo, e o segundo no capítulo IV, para a solução do problema de localização estocástico.

Dentro do campo dos problemas matemáticos aplicados, os problemas de otimização tem recebido muita atenção nas últimas décadas, em grande parte devido ao aparecimento e rápido desenvolvimento do computador. Na tentativa de modelar um problema real e definir um objetivo a otimizar, verifica-se a existência de certos coeficientes ou parâmetros que intervêm no problema mas sobre os quais não há um pleno conhecimento, ou que são afetados de uma incerteza. Os problemas onde se verificam a presença de tais coeficientes podem ser de dois tipos:

- quando os possíveis valores dos coeficientes tomam a forma de uma função de um ou mais parâmetros com domínios conhecidos, estamos frente a um problema embutido dentro do campo da programação paramétrica.
- quando os coeficientes obedecem uma determinada distribuição de probabilidades, estamos tratando com a programação estocástica.

É esperado supor que a dificuldade em tratar os problemas de programação estocástica seja maior que no caso determinístico. Por isso, só a programação linear estocástica experimentou um desenvolvimento satisfatório até o momento. Apesar dos resultados de Rockafellar e Wets sobre a teoria da dualidade estocástica, o campo da programação não-linear estocástica não apresenta ainda resultados passíveis de aplicações.

Várias atitudes foram propostas para encarar a resolução de um problema de programação estocástica linear. O enfoque passivo, ou também chamado por Mandansky [10] como enfoque "esperar e olhar", consiste em levantar a função de distribuição do ótimo da função objetivo. A partir de Tintner [14] muitos trabalhos foram desenvolvidos nesta linha. Um outro enfoque (ativo) procura colocar o problema original numa forma determinística equivalente. Esse último enfoque coloca a resolução do problema como um problema de decisão, no sentido que uma atitude deve ser tomada "aqui e agora", antes do conhecimento do valor que a variável aleatória vai tomar.

O enfoque ativo por sua vez apresenta duas maneiras fundamentais de ser aplicado, ou seja, um equivalente determinístico pode ser achado através do problema de dois estágios, proposto por Dantzig [5] ou através da adoção de "restrições de risco", proposto por Charnes e Cooper [1].

### II.3.2 - O PROBLEMA DE DOIS ESTÁGIOS

Seja o problema linear de

$$\text{minimizar } cx \quad (11)$$

$$\text{sujeito a } Ax = b, \quad x \geq 0 \quad (12)$$

com

$$A(m \times n), \quad c(n \times 1), \quad b(m \times 1)$$

Vamos assumir que  $b$  é uma variável aleatória e que conhecemos sua função de distribuição de média  $\bar{b}$ . Nestas condições é impossível determinar um vetor  $x$  que satisfaça (12) para qualquer valor de  $b$ . Porém, podemos observar a discrepância entre  $Ax$  e  $b$ , que por sua vez será também uma variável aleatória e cuja função de distribuição dependerá de  $x$ . Se imaginarmos que se deve pagar uma penalidade por essas discrepâncias, podemos contornar o problema da aleatoriedade sobre  $b$ , procurando minimizar a soma de  $cx$  com o valor esperado das penalidades. É possível penalizar as discrepâncias de muitas maneiras. Um modo simples é supô-la proporcional ao valor absoluto da diferença (erro) observada na  $i^{\text{ésima}}$  restrição, porém punindo de modo diferente se a diferença for para mais ou para menos.

Formalmente temos as penalidades

$$p_i(b_i - A_i x) \quad \text{quando} \quad b_i \geq A_i x$$

$$h_i(A_i x - b_i) \quad \text{quando} \quad A_i x \geq b_i$$

sendo que  $p_i \geq 0$  e  $h_i \geq 0$  são os pesos considerados.

Vamos introduzir as variáveis

$$a) \quad y_i^+ = b_i - A_i x \quad \text{e} \quad \bar{y}_i = 0$$

se  $b_i \geq A_i x$

b)  $y_i^+ = 0$  e  $\bar{y}_i = A_i x - b_i$

se  $A_i x \geq b_i$

A adoção das variáveis  $y^+ = [y_1^+, y_2^+, \dots, y_m^+]$  e

$y^- = [y_1^-, y_2^-, \dots, y_m^-]$  pode ser vista como uma correção que deve ser feita quando, sendo escolhido "a priori" um vetor  $x \geq 0$ , ocorrer uma realização da variável aleatória  $b$ . Tendo como objetivo minimizar o custo de tal correção, podemos definir a escolha ótima dos  $y^+$  e  $y^-$  como um problema de otimização, conhecido como o problema do segundo estágio.

$$\text{minimizar } \left\{ \begin{array}{l} py^+ + hy^- \\ y^+, y^- \geq 0 \end{array} \right\}$$

sujeito a

(13)

$$y^+ - y^- = b - Ax$$

Assume-se no problema acima que os  $p_i$  e  $h_i$  são não negativos, além de não poderem ser simultaneamente iguais a zero para algum  $i$ . Com isso assegura-se que a solução de (13) terá pelo menos  $y^+$  ou  $y^-$  igual a zero para cada  $i$ . Contudo, para que tal ocorra é suficiente pedir  $p + h \geq 0$ , que é uma condição menos restritiva e como veremos adiante, acompanhada de um significado econômico.

Como já dito, o valor ótimo de (13) depende de  $x$ . Então o problema (11) - (12) com  $b$  aleatório pode ser escrito na forma final equivalente e determinística

$$\text{minimizar } \left\{ \begin{array}{l} cx + E \min (py^+ + hy^-) \\ x \geq 0 \quad y \geq 0 \end{array} \right\}$$

(14)

sujeito a  $y^+ - y^- = b - Ax$

onde  $E$  denota a esperança matemática.

Teorema 4 - A função  $cx + E \min (py^+ + hy^-)$  é convexa em  $x$ .

Prova - Mostraremos primeiro que  $\min (py^+ + hy^-) = \pi(x)$  é uma função convexa em  $x$ . Queremos provar então que

$$\pi(\lambda x_1 + (1-\lambda)x_2) \leq \lambda\pi(x_1) + (1-\lambda)\pi(x_2); 0 \leq \lambda \leq 1 \quad (15)$$

Para uma determinada ocorrência aleatória  $\bar{b}$ , seja

$y_1 = y_1^+ - y_1^-$  e  $y_2 = y_2^+ - y_2^-$  as soluções ótimas do problema (13) para  $x_1$  e  $x_2$  respectivamente. Então

$$y_1 = \bar{b} - Ax_1 \quad (16)$$

$$y_2 = \bar{b} - Ax_2 \quad (17)$$

Multiplicando (16) por  $\lambda$  e (17) por  $(1 - \lambda)$  e somando

$$\lambda y_1 + (1-\lambda)y_2 = \bar{b} - A(\lambda x_1 + (1-\lambda)x_2) \quad (18)$$

Se (18) tiramos que  $\lambda y_1 + (1 - \lambda)y_2$  é uma solução factível de

(13) para  $x = \lambda x_1 + (1 - \lambda)x_2$ . Isto significa que

$$\pi(\lambda x_1 + (1-\lambda)x_2) \leq p(\lambda y_1^+ + (1-\lambda)y_2^+) + h(\lambda y_1^- + (1-\lambda)y_2^-) =$$

$$= \lambda(py_1^+ + hy_1^-) + (1-\lambda)(py_2^+ + hy_2^-)$$

$$= \lambda\pi(x_1) + (1-\lambda)\pi(x_2)$$

Como (15) é válida para qualquer ocorrência de  $b$

$$E\pi(\lambda x_1 + (1-\lambda)x_2) \leq E\left[\lambda\pi(x_1) + (1-\lambda)\pi(x_2)\right] =$$

$$= \lambda E\pi(x_1) + (1-\lambda)E\pi(x_2)$$

Como  $cx$  é uma função linear, o teorema está provado.

Em decorrência do teorema 4 podemos afirmar que o problema (11) - (12) com  $b$  estocástico é equivalente a (14), um problema de programação convexa com restrições lineares. Dantzig [5] enunciou o problema de dois estágios (que também leva o seu nome) numa forma mais geral.

$$\min_x \left\{ cx + E \min_y dy \right\}$$

$$Ax + By = b$$

$$x, y \geq 0$$

onde  $d(mx1)$  é um vetor de penalidades e  $B(rxm)$

### II.3.3 - PROBLEMAS ESTOCÁSTICOS COM RESTRIÇÕES DE RISCO

O enfoque de restrição de risco em programação estocástica recebeu muitas contribuições desde os trabalhos pioneiros de Charnes e Cooper [1, 2, 3]. Ele consiste basicamente em não requerer que as restrições com coeficientes aleatórios sejam satisfeitas completamente, ou seja, é suficiente que elas sejam satisfeitas numa dada proporção de casos. Para tal, arbitra-se uma determinada probabilidade em que uma restrição deve ser satisfeita. Formalmente, um dos problemas mais simples de serem tratados é:

minimizar  $cx$

sujeito a

(19)

$$\text{prob} \left[ \sum_{j=1}^n a_{ij} x_j \geq b_i \right] \geq \beta_i, \quad i \in I = \{1, 2 \dots m\}$$

$$x_j \geq 0, \quad 0 \leq \beta_i \leq 1$$

onde  $\beta_i, i \in I$  são probabilidades dadas e  $b_i, i \in I$  variáveis aleatórias independentes com funções distribuição de probabilidades  $F_i, i \in I$ .

O lado esquerdo da restrição acima pode ser visto como a expressão de uma função de distribuição acumulada

$$\text{prob} \left[ \sum_{j=1}^n a_{ij} x_j \geq b_i \right] = F_i \left( \sum_{j=1}^n a_{ij} x_j \right), \quad i \in I$$

O que se quer é o menor valor  $\hat{b}_i$  tal que

$$F_i(\hat{b}_i) = \beta_i, \quad i \in I,$$

ou seja

$$\hat{b}_i = F_i^{-1}(\beta_i), \quad i \in I$$

As restrições podem então assumir uma forma determinística equivalente

$$\sum_{j=1}^n a_{ij} x_j \geq \hat{b}_i$$

Sendo que a função  $F_i$  é caracterizada pela área sobre a curva abaixo (função densidade), a figura II.2 ilustra a equivalência pretendida.

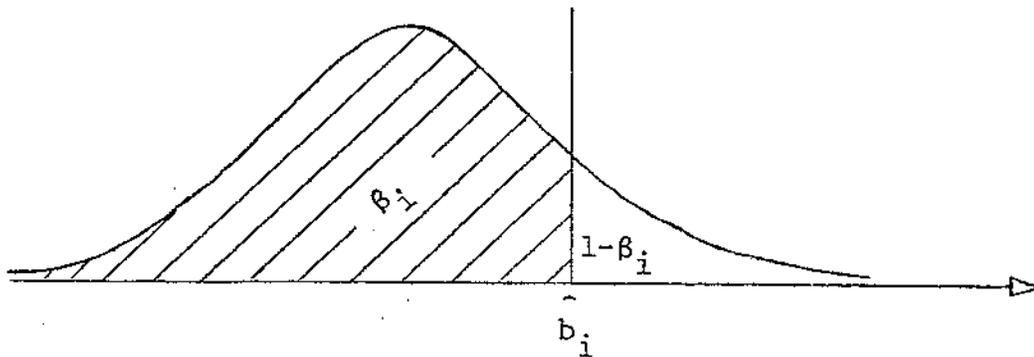


FIG. II.2

O problema (19) pode ser mais complicado. Vejamos algumas extensões possíveis.

O vetor  $c$  e a matriz  $[a_{ij}]$  podem ser variáveis aleatórias também (o caso de  $c$  aleatório é trivial, bastando substituí-lo pela sua média) e em Vajda [15] é mostrado o desenvolvimento de um equivalente determinístico para o caso das funções de distribuição chamadas estáveis.

Em Miller e Wagner [11] é visto o caso em que os  $b_i$  são variáveis aleatórias independentes, mas se deseja que a distribuição de probabilidades conjunta de todas inequações seja satisfeita a um nível  $\beta$ . Ou seja, quer-se

minimizar  $cx$

sujeito a  $\text{prob} \left[ Ax \geq b \right] \geq \beta$

$x \geq 0$  ,  $0 \leq \beta \leq 1$

Para o caso em que a suposição de independência entre os  $b_i$ ,  $i \in I$  é uma hipótese não razoável, Prékopa [12, 13] desenvolveu a teoria de função logaritmico-côncava que permite resolver problemas para uma classe de distribuições (inclusive a normal), através da aplicação de algoritmos da programação não-linear, como o de direções factíveis e o SUMT.

Até agora, nos problemas com restrição de risco, a solução ótima do problema era achada a priori, ou seja, antes do valor da variável aleatória tornar-se conhecido. Um outro método consiste em esperar que esses valores tornem-se conhecidos, mas decidir antecipadamente como esse conhecimento será utilizado, através de uma regra pré-fixada. Por exemplo, consideremos o problema (19). Charnes e Cooper [4] sugerem uma regra linear do tipo

$$x_j = \sum_{k=1}^m d_{jk} b_k \quad (20)$$

Substituindo (20) na função objetivo e nas restrições de (19) e desenvolvendo para o caso da distribuição normal, chega-se a um problema equivalente de programação convexa nas variáveis  $d_{jk}$ . Esse modelo encontra aplicações em problemas dinâmicos, onde existe dependência no tempo entre as variáveis aleatórias.

#### II.4 - FORMULAÇÃO DO PROBLEMA

O problema de localização estocástico consiste em encontrar locais ótimos para a instalação de plantas produtivas de modo a satisfazer uma demanda aleatória e procurando minimizar os custos fixos de instalação (construção) e os custos de transporte. Uma possível formulação apela para um problema determinístico equivalente através do enfoque de dois estágios, onde a não satisfação completa das demandas, para um dado plano de localização, produção e transporte é penalizada no critério. Seja a variável aleatória  $D_j$  a demanda por produto num destino  $j$ . Supondo que aí seja alocada uma quantidade de produto  $y_j$  e que após a observação da ocorrência da variável aleatória, ocorrer que  $y_j > D_j$ , então haverá um excesso de produto em  $j$  que deve ser armazenado e portanto incorre-se num custo  $h_j(y_j - D_j)$ . Por outro lado, se  $y_j < D_j$ , haverá falta de produto que deve

ser emergencialmente suprida com um custo  $p_j(D_j - y_j)$ . Os custos  $h_j$  e  $p_j$  quase sempre tem significação física cuja definição depende de cada modelamento feito e porisso pode ser avaliado sem dificuldade. Considerando ainda que o transporte de mercadorias de um centro de produção  $i$  a um centro de consumo  $j$  incorre num custo  $c_{ij}$  e a instalação de uma planta tem um custo fixo, uma possível formulação matemática baseada no equivalente determinístico de dois estágios é:

$$\text{Minimizar } \phi(x,z) = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + E \left[ \sum_{j \in J} \theta_j (D_j - y_j) \right] + \sum_{i \in I} k_i z_i$$

$x, z$

sujeito a

$$\sum_{j \in J} x_{ij} \leq S_i z_i, \quad i \in I$$

(21)

$$\sum_{i \in I} x_{ij} = y_j, \quad j \in J$$

$$z_i = 0, 1, \quad i \in I$$

$$x_{ij} \geq 0, \quad i \in I, \quad j \in J,$$

onde  $E$  significa esperança matemática,

$$\theta_j(D_j - y_j) = \begin{cases} h_j(D_j - y_j) & \text{se } y_j \geq D_j \\ p_j(D_j - y_j) & \text{se } y_j \leq D_j \end{cases}$$

e

$i$  índice de possível localização de planta,  $i \in I = \{1, 2, \dots, m\}$ ,

$j$  índice de ponto de demanda,  $j \in J = \{1, 2, \dots, n\}$ ,

$c_{ij}$  custo unitário de transporte da origem  $i$  ao destino  $j$ ,  
( $c_{ij} \geq 0$ ),

$k_i$  custo fixo de instalação de uma planta no local  $i$  ( $k_i \geq 0$ ),

$S_i$  capacidade de produção no local  $i$ ,

$p_j$  custo unitário de falta no destino  $j$ , ( $p_j \geq 0$ ),

$h_j$  custo unitário de excesso no destino  $j$ , ( $p_j - h_j \geq 0$ ),

$x_{ij}$  variável quantidade de produto transportado da planta  $i$   
para o ponto de demanda  $j$  ( $x$  é um vetor de componentes  
 $x_{ij}$ ,  $i \in I$ ,  $j \in J$ ),

$z_i$  variável zero-um que representa a decisão de localização  
de planta: será 1 se o local  $i$  for escolhido e 0 caso  
contrário ( $z$  é um vetor de componentes  $z_i$ ,  $i \in I$ ).

As hipóteses de  $p_j \geq 0$  e  $p_j \geq h_j$  são importantes do ponto de vista econômico. O caso  $p_j < h_j$  não tem nenhum interesse. Além disso são necessárias para a convexidade do problema, como já levantado quando se tratou o problema de dois estágios. A inversão de sinal sobre  $h_j$  se deve à definição da função penalidade  $\theta_j(\cdot)$ . Dependendo do problema considerado,  $h_j$  pode ser negativo ou positivo, sendo interpretado num caso como um custo (por exemplo, custo de armazenagem) e noutro como uma receita extra (por exemplo, receita por exportação).

A formulação (21) é muito semelhante ao problema de dois estágios apresentado em II.3.2 sendo que a restrição

$\sum_{i \in I} x_{ij} = y_j$  é apenas definicional.

Um fato muito importante que deve ser notado é que para uma dada estrutura fixa de localização ( $z_i, i \in I$  fixado), o problema resultante é de encontrar um fluxo ótimo que satisfaça uma demanda aleatória. Tal problema é um clássico problema de transporte estocástico (PTE), que como veremos mais tarde, será o subproblema a ser resolvido quando da aplicação da decomposição de Benders. Por ora é suficiente observar que desprezando o termo constante que aparece com a fixação de  $z$ , a função objetivo de (21) é análoga a de (14), garantindo através do teorema 4 que o problema de transporte estocástico é convexo.

Supondo que  $F(v_1, \dots, v_j, \dots, v_n)$  seja a função de distribuição acumulada conjunta das demandas satisfazendo as condições, (i) assume valor zero se qualquer de seus argumentos for negativo, (ii) cada uma de suas funções de distribuição marginal é contínua, a função objetivo pode ser desenvolvida aplicando-se a definição de esperança matemática.

$$\phi(x, z) = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \int_{E^{n+}} \sum_{j \in J} \theta_j(v_j - y_j) f(v_1, \dots, v_j, \dots, v_n)$$

$$dv_1 \dots dv_j \dots dv_n + \sum_{i \in I} k_i z_i,$$

onde  $f(v_1, \dots, v_j, \dots, v_n)$  é a função densidade de probabilidade. Tirando a somatória para fora da integral observa-se que o integrando só depende da variável  $v_j$ , o que permite integrar diretamente em relação a todas as outras.

$$\phi(x, z) = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} \int_0^{\infty} \theta_j(v_j - y_j) f_j(v_j) dv_j + \sum_{i \in I} k_i z_i,$$

onde  $f_j(v_j)$  é a função densidade marginal da demanda  $j$ .  
 Substituindo a função  $\theta_j$  pela sua expressão, tem-se

$$\phi(x,z) = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} \int_0^{y_j} h_j(v_j - y_j) f_j(v_j) dv_j +$$

$$\sum_{j \in J} \int_{y_j}^{\infty} p_j(v_j - y_j) f_j(v_j) dv_j + \sum_{i \in I} k_i z_i$$

Desenvolvendo a expressão acima, tem-se

$$\phi(x,z) = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} \left\{ \int_0^{y_j} h_j(v_j - y_j) f_j(v_j) dv_j + \right.$$

$$\left. p_j \left[ \int_0^{\infty} (v_j - y_j) f_j(v_j) dv_j - \int_0^{y_j} (v_j - y_j) f_j(v_j) dv_j \right] \right\} + \sum_{i \in I} k_i z_i$$

$$= \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} \left\{ h_j \int_0^{y_j} (v_j - y_j) f_j(v_j) dv_j + p_j \left[ \mu_j - y_j \right. \right.$$

$$\left. \left. - \int_0^{y_j} (v_j - y_j) f_j(v_j) dv_j \right] \right\} + \sum_{i \in I} k_i z_i$$

$$= \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} \left[ p_j \mu_j - p_j y_j + (h_j - p_j) \int_0^{y_j} (v_j - y_j) \right.$$

$$\left. f_j(v_j) dv_j \right] + \sum_{i \in I} k_i z_i,$$

onde  $\mu_j$  é o valor médio da demanda no destino  $j$ . Realizando uma integração por partes na expressão acima e omitindo o termo constante  $\sum_{j \in J} p_j \mu_j$ , a função objetivo fica

$$\phi(x,z) = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} \left[ -p_j y_j - (h_j - p_j) \int_0^{y_j} F_j(v_j) dv_j \right] + \sum_{i \in I} k_i z_i,$$

onde  $F_j(v_j)$  é a função de distribuição marginal da demanda em  $j$ . Substituindo o valor de  $y_j$ , chega-se à forma equivalente determinística final do PLE:

$$\begin{aligned} \text{Minimizar } \phi(x,z) = & \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} g_j \left( \sum_{i \in I} x_{ij} \right) + \\ x \geq 0, z = 0, 1 & \sum_{i \in I} k_i z_i \end{aligned} \quad (22)$$

$$\text{sujeito a } \sum_{j \in J} x_{ij} \leq S_i z_i, \quad i \in I \quad (23)$$

$$\begin{aligned} & \text{(restrições lineares adicionais sobre } z) \\ & \quad \quad \quad (24) \end{aligned}$$

onde

$$g_i \left( \sum_{i \in I} x_{ij} \right) = -p_j \sum_{i \in I} x_{ij} + (p_j - h_j) \int_0^{\sum_{i \in I} x_{ij}} F_j(v_j) dv_j$$

é uma função convexa e diferenciável

Algumas observações devem ser feitas. As restrições (24) dão flexibilidade para incorporar ao modelo restrições lineares que contenham somente variáveis  $z$ , por exemplo, uma limitação no número total de locais que devam existir ou acoplamentos lógicos entre diferentes locais, etc. A diferenciabilidade da função  $g_j$  é garantida pela suposição que  $F_j$ ,  $j \in J$  são contínuas e sua convexidade é imediata, podendo ainda ser verificada por analogia com o problema de dois estágios e usando o teorema 4. Vale ressaltar que a aplicação da DBG do problema (22-24) exige a convexidade da função objetivo, mas não a diferenciabilidade. Essa última exigência é no sentido de possibilitar a eficiente regra de recuperação do vetor multiplicador ótimo explicada no item III.3. Essa observação abre a possibilidade de se tratar funções de distribuição discretas com ligeiras mudanças no desenvolvimento apresentado. A formulação escolhida procura ser simples para facilitar a apresentação da solução do problema. Porém outras considerações podem ser incorporadas sem prejudicar o método de resolução apresentado neste capítulo. Por exemplo, se considerados os custos de produção devem ser incluídos nos coeficientes  $c_{ij}$ .

## II.5 - ESPECIALIZAÇÃO DA DBG

O problema (22-24) é um problema de programação mista não-linear (não-convexo) mas quando um vetor  $z$  é fixado ele se transforma em um problema clássico de programação não-linear convexa (PTE). Então, as variáveis  $z$  são obviamente identificadas como variáveis "complicantes".

Definindo

$$G_i(x,z) = \sum_{j \in J} x_{ij} - S_i z_i \quad e$$

$X = \{ x/x \geq 0 \}$  e  $Z = \{ z/z = (z_1, \dots, z_i, \dots, z_m), z_i=0,1 \}$  e  
 $z$  satisfazendo (24) },

a manipulação de projetar (22-24) no espaço dos  $z$  fornece

$$\underset{z}{\text{minimizar}} \quad v(z) \quad \text{sujeito a} \quad z \in Z \quad (25)$$

onde  $v(z) \equiv \underset{x \in X}{\text{ínfimo}} \phi(x, z) \quad \text{sujeito a} \quad G(x, z) \leq 0 . \quad (26)$

O problema de otimização embutido em (26) é identificado como o problema de transporte estocástico (PTE), que aparece quando temporariamente é fixado (projeção) um vetor de localização  $z = z^*$  no problema (22-24); é o subproblema do esquema de decomposição:

$$\text{Minimizar } \phi(x, z^*) = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} g_j \left( \sum_{i \in I} x_{ij} \right) + \sum_{i \in I} k_i z_i^* \quad (PTE)$$

sujeito a  $\sum_{j \in J} x_{ij} \leq S_i z_i^* , \quad i \in I .$

Note que o valor ótimo do subproblema acima é  $v(z^*)$ , que é uma avaliação dos custos totais (instalação e distribuição) da decisão de localização  $z^*$ , a menos do termo constante  $\sum_j \mu_j p_j$  que foi omitido.

Deve-se notar que para qualquer  $z \in Z$  sempre existe um  $x \in X$  que satisfaz  $G(x, z) \leq 0$ , por exemplo,  $x=0$  - não transportar nada - sempre satisfaz as restrições de produção,

qualquer que seja a configuração de localização escolhida. Isso significa que não é necessário considerar o conjunto V definido em II.2, e em consequência o teorema 2 não se aplica na presente especialização.

A dualização do problema projetado (25) através da aplicação do teorema 3 é possível, visto que suas hipóteses são satisfeitas. A hipótese de regularidade é satisfeita (restrições lineares) e as condições de Kuhn - Tucker são necessárias e suficientes para ótimo. Então

$$v(z) = \sup_{u \geq 0} \left[ \inf_{x \in X} [\phi(x,z) + uG(x,z)] \right], \text{ para todo } z \in Z. \quad (27)$$

E o problema (25) é equivalente a

$$\text{minimizar}_{z \in Z} \left\{ \sup_{u \geq 0} \left[ \inf_{x \in X} [\phi(x,z) + uG(x,z)] \right] \right\},$$

ou usando o fato de um supremo ser o menor limitante superior, o PLE (22-24) é equivalente ao problema mestre

$$\begin{aligned} &\text{minimizar } \gamma && (28) \\ &\gamma, z \in Z \end{aligned}$$

sujeito a

$$\gamma \geq \inf_{x \in X} [\phi(x,z) + uG(x,z)], \text{ para todo } u \geq 0 \quad (29)$$

O sucesso da aplicação da DBG ao PLE se deve em grande parte à separabilidade das funções  $\phi$  e  $G$  em  $x$  e  $z$ . Isto permite dizer que o PLE tem a propriedade P, ou seja, o problema de otimização que aparece do lado direito da restrição do problema mestre pode ser resolvido utilizando-se de informações da solução do subproblema da iteração anterior e com desprezível esforço computacional. Seja então um vetor dado de localização  $z^*$  e que para o qual a solução do subproblema (PTE) é  $v(z^*)$ . Se  $u^*$  é o vetor multiplicador ótimo associado, então, de (27) o valor ótimo de  $v(z^*)$  ocorre para  $u=u^*$  e é dado por

$$\begin{aligned}
 v(z^*) = \inf_{x \in X} & \left[ \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} g_j \left( \sum_{i \in I} x_{ij} \right) + \sum_{i \in I} k_i z_i^* \right. \\
 & \left. + \sum_{i \in I} \sum_{j \in J} u_i^* x_{ij} - \sum_{i \in I} u_i^* S_i z_i^* \right] = \sum_{i \in I} k_i z_i^* - \\
 & \sum_{i \in I} u_i^* S_i z_i^* + \inf_{x \geq 0} \left[ \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \right. \\
 & \left. \sum_{j \in J} g_j \left( \sum_{i \in I} x_{ij} \right) + \sum_{i \in I} \sum_{j \in J} u_i^* x_{ij} \right] \quad (30)
 \end{aligned}$$

De (29) segue que associado a  $u^*$  tem-se

$$\begin{aligned}
 \gamma \geq \inf_{x \in X} & \left[ \phi(x, z) + u^* G(x, z) \right] = \inf_{x \geq 0} \left[ \sum_{i \in I} \sum_{j \in J} \right. \\
 & c_{ij} x_{ij} + \sum_{j \in J} g_j \left( \sum_{i \in I} x_{ij} \right) + \sum_{i \in I} k_i z_i +
 \end{aligned}$$

$$\left[ \sum_{i \in I} \sum_{j \in J} u_i^* x_{ij} - \sum_{i \in I} u_i^* S_i z_i \right] = \sum_{i \in I} k_i z_i -$$

$$\sum_{i \in I} u_i^* S_i z_i + \text{infimo}_{x \in X} \left[ \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} +$$

$$\sum_{j \in J} g_j \left( \sum_{i \in I} x_{ij} \right) + \sum_{i \in I} \sum_{j \in J} u_i^* x_{ij} \right].$$

Substituindo acima a expressão do infimo tirado de (30), tem-se

$$\begin{aligned} \gamma &\geq \sum_{i \in I} k_i z_i - \sum_{i \in I} u_i^* S_i z_i + v(z^*) - \sum_{i \in I} k_i z_i^* - \\ &+ \sum_{i \in I} u_i^* S_i z_i^* \end{aligned}$$

Rearranjando os termos chega-se finalmente ao corte gerado por  $z^*$  e  $u^*$

$$\gamma \geq v(z^*) + \sum_{i \in I} (k_i - u_i^* S_i) (z_i - z_i^*) \quad (31)$$

### II.5.1 - ALGORITMO

A resolução do problema se dá por um procedimento iterativo a dois níveis, onde a cada vez é gerado no nível

superior do problema mestre um novo corte (31) através do vetor multiplicador ótimo obtido da resolução do PTE no nível inferior. O problema mestre é, a cada iteração, menos relaxado no sentido que são geradas novas restrições e ele produz um limitante inferior ( $\gamma^*$ ) do valor ótimo do PLE. A resolução do subproblema PTE fornece uma solução factível do PLE cujo valor  $v(z^*)$  pode ou não constituir-se no melhor limitante superior até agora conhecido, dependendo do valor da solução incumbente (define-se solução incumbente ao par  $(\bar{x}, \bar{z})$  onde  $\bar{z}$  é o melhor vetor de localizações até agora conhecido e  $\bar{x}$  é o vetor de fluxos ótimos associado).

Para um dado parâmetro de tolerância  $\epsilon > 0$ , um procedimento iterativo é:

Passo 1 - Inicialize com qualquer vetor de localizações convenientemente escolhido  $z^0$ . Resolva o subproblema PTE para  $z = z^0$  e obtenha o vetor multiplicador ótimo  $u^0$ . Faça  $p = 0$ , o limitante superior  $LS = v(z^0)$ , e o limitante inferior  $LI = 0$ .

Passo 2 - Resolva o problema mestre corrente

minimizar  $\gamma$  sujeito a

$$\gamma \geq v(z^t) + \sum_{i \in I} (k_i - u_i^t S_i)(z_i - z_i^t), \quad t=0,1,\dots,p,$$

através de uma rotina zero-um que leve em conta a variável real  $\gamma$ . Seja  $(\gamma^*, z^*)$  uma solução ótima;  $\gamma^*$  é o novo limitante inferior, i.é.,  $LI = \gamma^*$ . Se  $LS - LI \leq \epsilon$ , pare: a solução incumbente é uma solução ótima.

Passo 3 - Resolva o PTE para  $z = z^*$  e determine a sua solução ótima  $x^*$  e o valor da função objetivo  $v(z^*)$ . Se  $v(z^*) < LS$ , faça  $LS = v(z^*)$  e atualize a solução incumbente, i.é., faça  $(\bar{x}, \bar{z}) = (x^*, z^*)$ . Se  $LS - LI \leq \epsilon$ , pare: a solução ótima é  $(\bar{x}, \bar{z})$ . Senão, obtenha o vetor



Sendo  $z^0$  a solução inicial, o primeiro problema mestre a ser resolvido tem somente um corte, cujo "coeficiente angular" é  $k - u^0S$  e a solução ótima é  $z^1$ . Na segunda iteração vamos achar um  $\gamma^2$  mínimo, que é superior aos dois cortes. O terceiro corte, a exemplo dos outros não necessitaria passar tangente a curva em  $z^2$ ; qualquer corte passando tangente à curva entre  $(z^1, z^0)$  geraria uma restrição violada, no sentido que  $\gamma^2$  não conseguiria ser superior aos três cortes simultaneamente. É também possível verificar que o terceiro corte - construído com o vetor multiplicador ótimo - gera a restrição mais violada dentre todas. A observação de que os valores de LS não seguem uma sequência monótona não-crescente é verificada aqui (no caso de  $z$  discreto isso é mais evidente); e também se verifica que os valores de LI são monótonos não-decrescentes e que a cada iteração é resolvido um mestre cada vez menos relaxado, no sentido que a linearização externa que aproxima  $v(z)$  é cada vez mais exata. O critério de parada fica também evidente na figura.

Pensando o problema de localização, o problema mestre pode ser interpretado ainda, aproveitando-se das observações acima. Cada lado direito de suas restrições constitui-se numa função (linear) que sub-estima qualquer  $v(z)$ ,  $z \in Z$ . Porém o conjunto das restrições fornece a melhor função (mínimo ponto a ponto) sub-estimadora de  $v(z)$  e o problema mestre procura o vetor de localização que minimiza tal função. A solução do mestre sendo  $(\gamma^*, z^*)$ , significa que  $\gamma^*$  é o sub-estimador do custo esperado relativo a nova proposta de localização  $z^*$  e o correspondente subproblema de transporte estocástico fornece um novo multiplicador ótimo  $u^*$  que gera novo corte e etc.

II.5.2 - UM ALGORITMO ALTERNATIVO

A observação de que a solução do problema mestre nas primeiras iterações lida com aproximações ainda pobres, faz sugerir um algoritmo alternativo onde não necessariamente se busca otimalidade estrita quando da resolução do mestre. A idéia é procurar somente factibilidade, principalmente quando  $p$  é pequeno, ou seja, procurar uma solução cujo valor seja inferior a  $LS - \epsilon$ . O problema mestre pode então ser formulado como achar  $z \in Z$  tal que

$$\gamma \geq v(z^t) + \sum_{i \in I} (k_i - u_i^t S_i)(z_i - z_i^t), \quad t = 0, 1 \dots p$$

$$\gamma \leq LS - \epsilon$$

Eliminando  $\gamma$ , temos que

Achar  $z \in Z$ , tal que

(32)

$$v(z^t) + \sum_{i \in I} (k_i - u_i^t S_i)(z_i - z_i^t) \leq LS - \epsilon, \quad t = 0, 1 \dots p$$

O problema (32) busca simplesmente uma solução  $z \in Z$  que seja factível e qualquer função objetivo linear pode ser considerada para efeito de resolvê-lo através de um algoritmo de programação zero-um. Experiências computacionais mostraram bons resultados se se considerar o lado direito do último corte gerado como a própria função objetivo. Uma certa vantagem desse algoritmo alternativo é que o problema mestre tornou-se um problema de programação zero-um puro com a eliminação da variável real  $\gamma$  e pode ser resolvido por rotinas

padrões.

Com a subotimização do mestre, sua solução ótima não mais fornece necessariamente um limitante inferior da solução ótima de PLE. A variável LI pode ser eliminada e o algoritmo tem a seguinte forma:

Passo 1 - Escolhido um parâmetro de tolerância  $\epsilon > 0$ , inicialize com qualquer vetor de localizações convenientemente escolhido  $z^0$ . Resolva o subproblema PTE para  $z = z^0$  e obtenha o vetor multiplicador ótimo  $u^0$ . Faça  $p = 0$ , o limitante superior  $LS = v(z^0)$ .

Passo 2 - Resolva o problema mestre modificado corrente

$$\min v(z^p) + \sum_{i \in I} (k_i - u_i^p S_i)(z_i - z_i^p)$$

sujeito a

$$v(z^t) + \sum_{i \in I} (k_i - u_i^t S_i)(z_i - z_i^t) \leq LS - \epsilon, \quad t = 0, 1, \dots, p$$

por uma rotina zero-um. Se o problema não tem solução factível, terminou: a solução incumbente é uma solução ótima. Se uma solução ótima é  $z^*$ , vá ao passo 3.

Passo 3 - Resolva o PTE para  $z = z^*$ . Seja  $x^*$  uma solução ótima e  $v(z^*)$  o valor ótimo da função objetivo. Se  $v(z^*) < LS$ , faça  $LS = v(z^*)$  e atualize a solução incumbente, isto é, faça  $(\bar{x}, \bar{z}) = (x^*, z^*)$ . Faça  $p = p + 1$ ,  $u^p = u^*$ , (vetor multiplicador ótimo) e  $z^p = z^*$ . Volte ao passo 2.

REFERÊNCIAS BIBLIOGRÁFICAS - CAP. II

- [1] CHARNES, A. e COOPER, W.W., "Chance Constrained Programming", Management Science 6, 73-79, (1959).
- [2] \_\_\_\_\_ e \_\_\_\_\_, "Chance Constrained Programs with Normal Deviates and Linear Decision Rules", Naval Res. Log. Quarterly 7, 533-544, (1960).
- [3] \_\_\_\_\_ e \_\_\_\_\_, "Chance Constrains and Normal Deviates", J. Amer. Statist. Assoc. 57, 134-148, (1962)
- [4] \_\_\_\_\_ e \_\_\_\_\_, "Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints". Operations Research 11, 18-39, (1963).
- [5] DANTZIG, G.B., "Linear Programming under Uncertainty", Management Science 1, 197-206, (1955).
- [6] GEOFFRION, A.M. "Elements of Large - Scale Mathematical Programming" in Perspectives on Optimization, (ed. A. M. Geoffrion). Addison - Wesley, 1972
- [7] \_\_\_\_\_, "Generalized Benders Decomposition", J. Optimization Theory and Applic. 10, 237-260, (1972).
- [8] \_\_\_\_\_, "Duality in Nonlinear Programming: A Simplified Application - Oriented Development". SIAM Review 13 (1971)
- [9] LASDON, L.S., Optimization Theory for Large Systems, MacMillan, London 1970, Cap. 8.
- [10] MADANSKY, A., "Inequalities for Stochastic Linear Programming Problems", Management Science 6, 197-204, (1960).
- [11] MILLER, B.L., and WAGNER, H.M. "Chance Constrained Programming with Joint Constraints". Operations Research 13, 930-945, (1965).

- [12] PRÉKOPA, A., "Logarithmic concave measures with application to stochastic programming", Acta Sci. Math. 32, 301-316, Hungria, (1971).
  
- [13] \_\_\_\_\_, "Contributions to the theory of stochastic programming", Mathematical Programming, 4, 202-221, 1973.
  
- [14] TINTNER, G., "Stochastic Linear Programming with Applications to Agricultural Economics", in Proc. 2nd Symp. Lin. Progr. (A. Antosiewicz ed.), Washington, 197-228; (1955)
  
- [15] VAJDA, S., Probabilistic Programming, Academic Press, New York, 1972.

CAPÍTULO III

SOLUÇÃO COMPUTACIONAL DO PROBLEMA DE LOCALIZAÇÃO ESTOCÁSTICO

III.1 - INTRODUÇÃO

O método de Benders quando aplicado a um problema, o decompõe através da manipulação matemática de projetar tal problema sobre o espaço das variáveis "complicantes" -  $z$  - como já visto no item II.2. Essa decomposição pode ser vista dentro de uma estrutura de dois níveis, onde no inferior resolve-se um subproblema decorrente da fixação das variáveis  $z$  num nível superior. Estabelece-se uma troca de informações entre os 2 níveis, onde, no caso do PLE, o nível superior envia o vetor zero-um de localizações calculado -  $z^*$  - e o nível inferior retorna o vetor de fluxos ótimos calculado -  $x^*$  - para aquelas localizações, assim como o vetor multiplicador ótimo associado -  $u^*$  -.

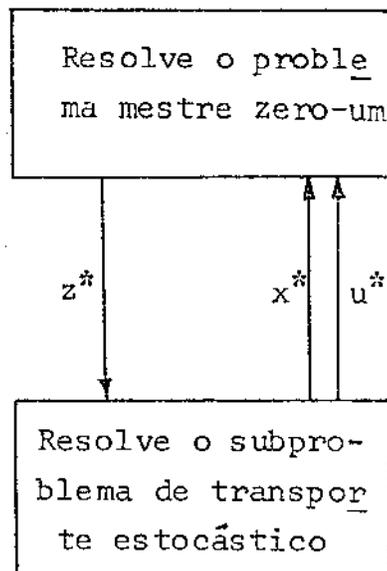


FIG. III.1

O presente capítulo ocupa-se de resolver o problema mestre e o subproblema. Inicialmente discutem-se os principais métodos adequados a resolver o subproblema de transporte estocástico e faz-se a aplicação de um deles. É dada ênfase à obtenção e ao papel que os multiplicadores ótimos desempenham no algoritmo de Benders. A seguir desenvolve-se um método de enumeração implícita para resolver o problema mestre de programação zero-um. É relatada, por final, as experimentações computacionais realizadas e as conclusões que foram daí tiradas.

### III.2 - RESOLUÇÃO DO SUBPROBLEMA

Vários métodos já foram propostos para resolver o problema de transporte estocástico. Tal problema foi colocado originalmente por Ferguson e Dantzig [8] para demandas com distribuição discreta. Para o caso contínuo, Elmaghraby [7] propôs um procedimento iterativo que, variando fluxos, busca satisfazer as condições de otimalidade de Kuhn - Tucker. Posteriormente Szwarc [18] estendeu o caso tratado em [8] para aproximações contínuas da demanda. Hadley [13] e Willians [19] apresentam técnicas semelhantes entre si, ambas apelando para o uso do princípio de decomposição de Dantzig - Wolfe generalizado, mas o último estende a solução para funções densidade de demanda mais gerais. Usando um equivalente determinístico alternativo, Balachandran e Jain [1] propuseram uma técnica paramétrica que trabalha com problemas clássicos de transporte mas não apresentam um exemplo numérico para mostrar a aplicabilidade do método. O trabalho de Ermolev e Mirzoakhmedov [6] resolve o PTE usando conceitos de quasi - gradientes estocásticos e o método de direções factíveis sem requerer que a função densidade exista, bastando conhecer a sequência de realizações da demanda aleatória. Mais recentemente, Cooper e Leblanc [5] propuseram um algoritmo baseado no método de Frank-Wolfe para programação não-linear. Devido a simplicidade do algoritmo e da possibilidade relatada de tratar eficientemente problemas de grande dimensão, este foi o método selecionado para a resolução do subproblema.

III.2.1 - O ALGORITMO DE FRANK-WOLFE

O método proposto por Frank e Wolfe [9] originalmente destinava-se para resolver problemas com funções objetivas convexas (quadráticas) e restrições lineares. Sejam então problemas da forma

$$\text{minimizar } \phi(x) \quad (1)$$

$$\text{sujeito a } Ax = b, \quad x \geq 0, \quad (2)$$

onde  $\phi(x)$  é uma função convexa diferenciável. A aproximação de primeira ordem de Taylor de  $\phi(\cdot)$  sobre o ponto factível  $x^k$  fornece o seguinte problema de programação linear:

$$\text{minimizar } \phi(x^k) + \nabla\phi(x^k) (y^k - x^k) \quad (3)$$

$$\text{sujeito a } Ay^k = b, \quad y^k \geq 0, \quad (4)$$

onde  $\nabla\phi(x^k)$  é o vetor gradiente avaliado no ponto  $x^k$ . Os termos  $\phi(x^k)$  e  $\nabla\phi(x^k) \cdot x^k$  são constantes e podem ser omitidos de (3).

Seja agora  $\hat{y}^k$  uma solução ótima de (3)-(4). Podemos definir então uma direção de busca

$$d^k \equiv \hat{y}^k - x^k.$$

Um novo ponto factível pode ser calculado com

$$x^{k+1} = x^k + \alpha d^k ,$$

sendo o parâmetro  $\alpha$  determinado através da resolução do problema de busca unidimensional

$$\begin{aligned} &\text{minimizar } \phi(x^k + \alpha d^k) \\ &0 \leq \alpha \leq 1 \end{aligned} \quad (5)$$

Prova-se (e.g. [20] e [14]) que a sequência de pontos gerados  $\{x^0, x^1, \dots, x^k, \dots\}$  converge para  $x^*$ , solução ótima de (1)-(2).

### III.2.2 - APLICAÇÃO DO ALGORITMO DE FRANK-WOLFE PARA O PROBLEMA DE TRANSPORTE ESTOCÁSTICO

Trabalharemos com o PTE em uma forma reduzida, ou seja, eliminando as fontes produtoras que estão fechadas temporariamente pela solução  $z^*$ , fornecida pelo problema mestre numa certa iteração. Pode-se então definir para tal vetor de localizações  $z^*$ , um subconjunto de índices  $I^* \subseteq I$ , tal que  $z_i^* = 1$ . Para cada  $z_i^* = 0$  (ou seja,  $i \in I - I^*$ ), o subproblema tem uma restrição  $\sum_{j \in J} x_{ij} \leq 0$ , resultando em soluções degeneradas  $x_{ij} = 0$ , para todo  $i \in I - I^*$ ,  $j \in J$ . Eliminando-se tais restrições e variáveis definimos o PTE reduzido (PTER):

$$\begin{aligned} &\text{minimizar } \phi(x) = \sum_{i \in I^*} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} g_j \left( \sum_{i \in I^*} x_{ij} \right) + \sum_{i \in I^*} k_i \\ &x \geq 0 \end{aligned} \quad (PTER)$$

$$\text{sujeito a } \sum_{j \in J} x_{ij} \leq S_i , \quad i \in I^* ,$$

onde  $g_j(\sum_{i \in I^*} x_{ij}) = -p_j \sum_{i \in I^*} x_{ij} + (p_j - h_j) \int_{\sum_{i \in I^*} x_{ij}}^{\infty} F_j(v_j) dv_j$

e  $x$  é o vetor de fluxos de componentes  $x_{ij}$ .

É óbvia a vantagem computacional de se trabalhar com o PTER, um problema de menor dimensão.

O vetor gradiente de  $\phi(x)$  é definido como  $\nabla\phi(x) = \frac{\delta\phi(x)}{\delta x}$

com componentes

$$\frac{\delta\phi(x)}{\delta x_{ij}} = c_{ij} + \frac{\delta g_j(\sum_{i \in I^*} x_{ij})}{\delta x_{ij}} = c_{ij} - p_j + (p_j - h_j) F_j(\sum_{i \in I^*} x_{ij}),$$

$$i \in I^*, j \in J.$$

Se  $x^k$  é uma solução factível do PTER podemos definir

$$\bar{c}_{ij} = \left. \frac{\delta\phi(x)}{\delta x_{ij}} \right|_{x=x^k} \quad (6)$$

Substituindo (6) em (3) e omitindo os termos constantes podemos construir o PL

$$\text{minimizar } \sum_{i \in I^*} \sum_{j \in J} \bar{c}_{ij} y_{ij}$$

$$\text{sujeito a } \sum_{j \in J} y_{ij} \leq S_i, \quad i \in I^* \quad (7)$$

$$y_{ij} \geq 0, \quad i \in I^*, j \in J.$$

O problema (7) tem uma estrutura angular particular que não tem presente qualquer restrição de acoplamento entre variáveis, o que o deixa completamente decomponível em cada origem. É esta particularidade que torna o método de Frank-Wolfe atrativo para a resolução do PTER. Se  $I^*$  tem  $m^*$  componentes, então teremos  $m^*$  subproblemas independentes - um para cada origem  $i \in I^*$  - do tipo

$$\left. \begin{array}{l} \text{minimizar} \quad \sum_{j \in J} \bar{c}_{ij} y_{ij} \\ \text{sujeito a} \quad \sum_{j \in J} y_{ij} \leq S_i \\ \quad \quad \quad y_{ij} \geq 0 \quad , \quad j \in J \end{array} \right\} \text{subproblema } i$$

A solução dos subproblemas acima é por simples inspeção. Inicialmente acha-se o índice  $s$  de modo que

$$\bar{c}_{is} = \min_{j \in J} \bar{c}_{ij}$$

A solução ótima será

$$\left. \begin{array}{l} \hat{y}_{is} = S_i \\ \hat{y}_{ij} = 0 \quad , \quad j \in J - \{s\} \end{array} \right\} \text{ se } \bar{c}_{is} < 0$$
$$\hat{y}_{ij} = 0 \quad , \quad j \in J \quad \text{ se } \quad \bar{c}_{is} \geq 0$$

A interpretação desta solução à luz do problema de transporte é evidente: se o menor estimador de primeira ordem dos custos esperados totais (de transporte, deficit e excesso) -  $\bar{c}_{is}$  - for negativo, então deveremos transportar todo o produto disponível na origem  $i$  para o destino  $s$ ; caso contrário, não transportaremos nada. Tal solução simples é possível porque não há restrição impondo satisfação de demanda nos destinos. O problema (5) de minimização unidimensional pode ser resolvido por qualquer rotina de busca que leve em conta a restrição de  $0 \leq \alpha \leq 1$ .

### III.2.3 - ANÁLISE DA CONVERGÊNCIA DO MÉTODO DE FRANK-WOLFE

Procuraremos aqui estabelecer condições para uma boa convergência do algoritmo de Frank-Wolfe proposto para resolver o PTER. Inicialmente calculemos o ótimo irrestrito do PTER, ou seja, o ótimo de sua função objetivo assumindo que a capacidade das fontes é ilimitada. Lembrando que a função objetivo do PTER é convexa e diferencial, as condições de Kuhn - Tucker são condições necessárias e suficientes para otimalidade. Um par  $(x,u)$ ,  $x \geq 0$ ,  $u \geq 0$  satisfaz tais condições se e somente se (ver Apêndice I):

$$c_{ij} - p_j + (p_j - h_j) F_j \left( \sum_{i \in I^*} x_{ij} \right) - \mu_{ij} \geq 0, \quad i \in I^*, \quad j \in J \quad (8)$$

$$x_{ij} \left[ c_{ij} - p_j + (p_j - h_j) F_j \left( \sum_{i \in I^*} x_{ij} \right) - \mu_{ij} \right] = 0, \quad i \in I^*, \quad j \in J \quad (9)$$

$$\mu_{ij} x_{ij} = 0, \quad i \in I^*, \quad j \in J \quad (10)$$

Supondo um determinado destino  $\bar{j} \in J$  e os fluxos convergentes

$x_{i\bar{j}} > 0$ ,  $i \in I^*$ , conclui-se de (10) que  $\mu_{i\bar{j}} = 0$ , para todo  $i \in I^*$ ,  
De (9) pode-se escrever

$$c_{i\bar{j}} = p_{\bar{j}} - (p_{\bar{j}} - h_{\bar{j}}) F_{\bar{j}} \left( \sum_{i \in I^*} x_{i\bar{j}} \right), \quad \forall i \in I^* . \quad (11)$$

Como o segundo membro é constante,  $\forall i \in I^*$  conclui-se que para  $\bar{j} \in J$  somente um fluxo convergente  $x_{i\bar{j}}$  deve ser não nulo, supondo que os  $c_{i\bar{j}}$ ,  $i \in I^*$  sejam diferentes entre si.  
O valor de  $x_{i\bar{j}}$  pode ser calculado de (11)

$$x_{i\bar{j}} = F_{\bar{j}}^{-1} \left( \frac{p_{\bar{j}} - c_{i\bar{j}}}{p_{\bar{j}} - h_{\bar{j}}} \right)$$

Como  $F_{\bar{j}}$  é uma função distribuição de probabilidades, deve satisfazer  $0 \leq F_{\bar{j}} \leq 1$ , implicando que

$$p_{\bar{j}} \geq c_{i\bar{j}} \geq h_{\bar{j}} .$$

Caso  $p_{\bar{j}} < c_{i\bar{j}}$ , temos que  $x_{i\bar{j}} = 0$  .

Resta saber como escolher  $\bar{i}$  entre os  $i \in I^*$ . Para isso tenta-se satisfazer (8). De (11) e para  $i = \bar{i}$ , vê-se que (8) é satisfeita na igualdade. Para os demais  $i \in I^*$  tem-se

$$c_{i\bar{j}} - \mu_{i\bar{j}} \geq p_{\bar{j}} - (p_{\bar{j}} - h_{\bar{j}}) F_{\bar{j}} \left( \sum_{i \in I^*} x_{i\bar{j}} \right), \quad \mu_{i\bar{j}} > 0 .$$

Como o segundo membro é constante  $\forall i \in I^*$ , só se encontrará um  $\mu_{i\bar{j}} > 0$  que satisfaça (12) se escolher-se  $c_{\bar{i}\bar{j}}$  como

$$c_{\bar{i}\bar{j}} = \min_{i \in I^*} c_{i\bar{j}} . \quad (12)$$

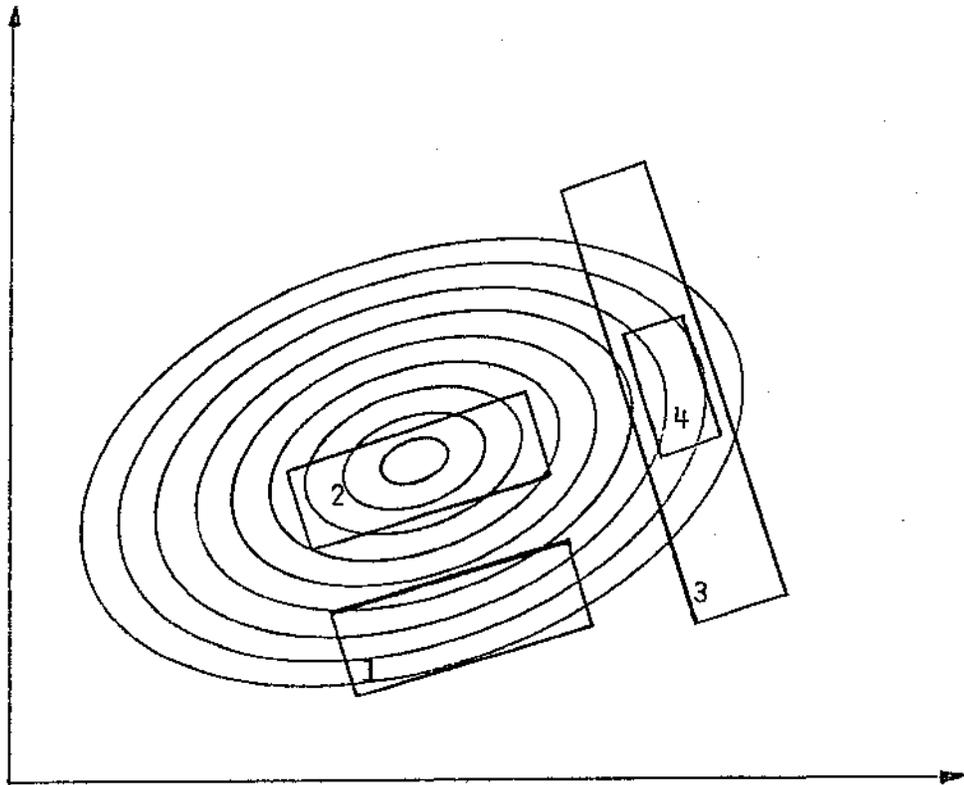
Por fim, vê-se que a solução  $x_{i\bar{j}} = 0$ ,  $i \in I^*$  não satisfaz a condição (8), supondo-se que no caso geral  $c_{i\bar{j}} < p_{\bar{j}}$ ,  $\forall i \in I^*$ .

Extendendo-se agora o mesmo raciocínio para todos os destinos  $j$ , conclui-se que o mínimo irrestrito do PTER pode ser calculado como

$$\left. \begin{aligned} x_{\bar{i}j} &= F_j^{-1} \left( \frac{p_j - c_{\bar{i}j}}{p_j - h_j} \right) \\ x_{ij} &= 0, \quad i \neq \bar{i} \\ \text{sendo } c_{\bar{i}j} &= \min_{i \in I^*} c_{ij} \end{aligned} \right\} \text{ para todo } j \in J \quad (13)$$

Deve-se salientar ainda que quando  $c_{\bar{i}j} < h_j$  para pelo menos um  $j \in J$ , a função objetiva não é limitada inferiormente.

A análise da convergência do algoritmo proposto no item anterior é baseada em experimentação computacional e explicações de alguns comportamentos são reforçados analiticamente, tentando estabelecer condições para que o algoritmo tenha um bom desempenho. O algoritmo de Frank-Wolfe trabalha iterativamente fazendo aproximações lineares da função objetivo não-linear. É fundamental e intuitivo num método desse tipo que a qualidade de tais aproximações influencie diretamente na sua eficiência. Ou seja, a convergência do algoritmo deverá ser melhor tanto quanto a função objetivo seja próxima da linear. (veja ref. [15], pag. 244). Porém, como o objetivo a otimizar está em presença de restrições, essa linearidade pretendida é função da posição relativa que ocupa o subconjunto de soluções factíveis definido pelo conjunto de restrições, assim como é função também do "tamanho" desse subconjunto. O exemplo bidimensional abaixo onde estão representados quatro conjuntos de restrições e as curvas de nível de uma função convexa explica estes aspectos levantados.



É evidente que uma aproximação linear da função objetivo é muito melhor sobre o conjunto 1 do que sobre o 2, mostrando que a proximidade do centro piora a aproximação. O mesmo ocorre em relação a 3 e 4. O último, por ser mais restrito, dá uma aproximação mais realista do que o primeiro. Conjugado a esses fatores, há também de se evidenciar a influência do mau condicionamento da função que pode, em certos casos, agravar a qualidade da aproximação.

Vários testes computacionais com o PTER foram realizados e pôde-se concluir que as condições para se obter convergência em intervalos de tempo satisfatórios são:

- 1 - O ponto de ótimo da função objetivo - calculado em (13) - deve estar fora do conjunto de soluções factíveis, e daí as restrições do PTER serão sempre saturadas. Isso ocorrerá sempre que  $\sum_{i \in I^*} S_i$  for significativamente menor que a soma das demandas médias.
- 2 - A condição acima não é suficiente para garantir

boa convergência. É ainda necessário garantir que o conjunto de soluções situe-se sobre uma região suave da função objetivo. Essa suavidade acentua-se tanto mais quanto o ponto de mínimo tende para o infinito. Para todos os dados do problema fixados e em particular  $p_j > h_j$ ,  $\forall j \in J$  verifica-se que a diminuição de todos os custos de transporte tende, por (13), a fazer aumentar os fluxos ótimos e melhorar a convergência. Essa constatação pode levar à errônea exigência de custos de transporte relativamente baixos como condição fundamental de convergência do método, como afirmado em Cooper e Leblanc [5]. Por exemplo, consegue-se boa convergência no caso de  $h_j > c_{ij}$ , para algum  $j \in J$  (nesse caso o mínimo é ilimitado inferiormente) e custos relativamente altos de transporte.

Concluindo, o algoritmo de Frank-Wolfe se mostrou eficiente para uma classe de problemas de transporte estocástico e abre possibilidades de, a partir dele procurar-se outros métodos que aproveitem o caráter específico do conjunto de restrições, como por exemplo levar em conta a informação obtida com linearizações já feitas em iterações anteriores.

### III.3 - O SUBPROBLEMA E OS MULTIPLICADORES

Como visto em II.5.1 a geração de um novo corte de Benders exige o conhecimento do vetor multiplicador ótimo -  $u^*$  - do subproblema. A obtenção de tal vetor esbarra em dois problemas: a técnica escolhida para resolver o PTER é primal e não fornece o vetor multiplicador ótimo como subproduto; a outra dificuldade é que resolve-se o PTER no nível inferior, ou seja, são consideradas as restrições sobre  $i \in I^*$ , mas necessitam-se também as variáveis duais das restrições  $i \in I - I^*$  para a construção do corte. Sabe-se também que dentre os vários possíveis multiplicadores, deve-se escolher aquele que fornece o mais efi

ciente corte de Benders.

Estas dificuldades são contornadas através da análise das condições de otimalidade de Kuhn-Tucker para o PTE definido em II.5. Um par  $(x,u)$ ,  $x \geq 0$ ,  $u \geq 0$  satisfaz essas condições de otimalidade se e somente se (ver Apêndice I):

$$c_{ij} - p_j + (p_j - h_j) F_j \left( \sum_{i \in I} x_{ij} \right) + u_i \geq 0, \quad i \in I, \quad j \in J \quad (14.1)$$

$$\sum_{j \in J} x_{ij} - S_i \leq 0, \quad i \in I^* \quad ; \quad \sum_{j \in J} x_{ij} \leq 0, \quad i \in I - I^* \quad (14.2)$$

$$x_{ij} \left[ c_{ij} - p_j + (p_j - h_j) F_j \left( \sum_{i \in I} x_{ij} \right) + u_i \right] = 0, \quad i \in I, \quad j \in J \quad (14.3)$$

$$u_i \left( \sum_{j \in J} x_{ij} - S_i \right) = 0, \quad i \in I^* \quad ; \quad u_i \left( \sum_{j \in J} x_{ij} \right) = 0, \quad i \in I - I^* \quad (14.4)$$

A partir de (14) os  $u_i^*$ ,  $i \in I^*$  podem ser calculados como

$$u_i^* = \max_{j \in J} \left[ -c_{ij} + p_j - (p_j - h_j) F_j \left( \sum_{i \in I^*} x_{ij} \right) \right] \quad (15.1)$$

$$\text{se} \quad \sum_{j \in J} x_{ij} = S_i, \quad i \in I^* \quad (15.2)$$

$$\text{ou} \quad u_i^* = 0 \quad (16.1)$$

$$\text{se} \quad \sum_{j \in J} x_{ij} \leq S_i, \quad i \in I^* \quad (16.2)$$

Se para um determinado  $i \in I^*$ , vale (15.2), então (15.1) é deduzido primeiro analisando (14.3) e considerando as equações onde  $x_{ij} \neq 0$ ,  $j \in J$ , de onde se tiram diferentes valores possíveis de  $u_j$ . Para se satisfazer (14.1) para todo  $j \in J$ , deve-se tomar  $u_j^*$  definido por (15.1). Se, caso contrário verifica-se (16.2), a adoção de (16.1) é evidente.

Analisemos agora o caso da obtenção dos  $u_i^*$ ,  $i \in I - I^*$ . Para isso apelaremos para o conceito de função de perturbação<sup>1</sup> do PTE, quando  $z = z^*$  que nos guiará para uma boa escolha de tais multiplicadores. Primeiro definimos a função

$$w(z^*) = \left[ \sum_{j \in J} p_j \mu_j - \sum_{i \in I} k_i z_i^* \right] + v(z^*)$$

que representa o custo mínimo esperado de distribuição ( que inclui custos de transporte e custos esperados de deficit e excesso) associado com o vetor de localizações  $z^*$ . É óbvio que uma solução ótima  $x^*$  (e multiplicador ótimo  $u^*$ ) de PTE é também ótima num problema similar onde se elimina  $\sum_{i \in I} k_i z_i$  e se inclui  $\sum_{j \in J} p_j \mu_j$  no critério. Pode-se então dizer que  $-u^*$  é um subgradiente da função de perturbação<sup>2</sup> associada com tal problema, e

$$w(z) \geq w(z^*) - \sum_{i \in I} u_i^* S_i(z_i - z_i^*),$$

---

1 - Da teoria de dualidade em programação não-linear. Para maior compreensão ver o clássico Geoffrion [10] ou o trabalho compacto de Soares Filho [17].

2 - Note que, a fim de simplificar a notação, e sem perda de generalidade, usamos  $z$  como argumento da função de perturbação  $w(\cdot)$ , em vez do lado direito  $(S_1 z_1, \dots, S_m z_m)$ .

ou, o termo  $-\sum_{i \in I} u_i^* S_i(z_i - z_i^*)$  é um sub-estimador da variação dos mínimos custos esperados de distribuição quando seleciona-se a localização  $z$  em vez de  $z^*$ . Por outro lado, a expressão  $\sum_{i \in I} k_i(z_i - z_i^*)$  avalia a efetiva variação dos custos de instalação quando seleciona-se a estrutura de localização  $z$  em vez de  $z^*$ . Então, o lado direito de qualquer corte (31), definido em II.5, gerado por algum vetor multiplicador ótimo  $u^*$  associado ao vetor  $z^*$  e somado a  $\sum_{j \in J} p_j \mu_j$ , fornece um sub-estimador do custo esperado total (distribuição e instalação), relacionado com qualquer outra localização  $z \in Z$ . Ou ainda, o custo esperado  $v(z)$  pode ser estimado com base no custo esperado conhecido  $v(z^*)$  por

$$v(z^*) + \sum_{i \in I} (k_i - u_i^* S_i) (z_i - z_i^*) .$$

Como estamos interessados em escolher  $u_i^*$ ,  $i \in I - I^*$  que satisfazendo (14) ainda forneça o melhor sub-estimador de  $v(z)$ ,  $z \in Z$ , devemos então tomar  $u_i^*$  o menor possível, ou seja

$$u_i^* = \max \left[ -c_{ij} + p_j - (p_j - h_j) F_j \left( \sum_{i \in I} x_{ij} \right) \right],$$

$$\text{para } i \in I - I^* \quad (17)$$

Observamos que qualquer  $u_i^* \geq 0$  satisfazendo (14) pode ser usado para gerar os cortes. A adoção de (17) fornece porém uma melhor escolha, com desprezível esforço computacional adicional. Em III.5 é relatada a melhoria obtida com o uso da regra (17).

### III.4 - RESOLUÇÃO DO PROBLEMA MESTRE

Foge dos objetivos desse estudo uma análise dos inúmeros métodos e técnicas existentes para a resolução de um problema de programação inteira ou zero-um. Nem mesmo nos permitimos uma incursão exploratória a tais métodos, como foi feito no caso da resolução do subproblema devido ao seu caráter mais específico. Dos quatro enfoques clássicos de programação inteira - enumerativo, decomposição de Benders, plano-de-corte e teoria de grupo - escolhemos o primeiro para solução do problema mestre com base em experiências bem sucedidas, para problemas de médio porte, relatadas em Geoffrion e Marsten [12]. Mais informação pode ser obtida em artigos da mesma linha, e.g. Beale [4], Balinski [3] e Salkin [16].

#### III.4.1 - INTRODUÇÃO: MÉTODOS DE ENUMERAÇÃO IMPLÍCITA

Dentro do enfoque enumerativo, os métodos de enumeração implícita se destacam pela simplicidade algorítmica. Uma maneira óbvia de resolver um problema zero-um de  $n$  variáveis seria enumerar completamente as  $2^n$  possíveis soluções e avaliar aquela de custo mínimo. A idéia de enumeração implícita traz consigo a de exclusão de soluções e nisso reside a "inteligência" do método.

O algoritmo aqui apresentado é baseado no trabalho original de Balas [2] e na versão melhorada de Geoffrion [11].

Os problemas resolvidos são da forma:

$$\begin{aligned} & \text{minimizar} && \pi = cz \\ & \text{sujeito a} && Az + b \leq 0 \\ & && z = 0,1 \end{aligned}$$

$$A(m \times n) \quad , \quad c(1 \times n) \quad , \quad b(m \times 1)$$

São necessárias algumas definições:

O número total de soluções possíveis é  $2^n$ . Todas essas soluções constituem uma enumeração completa das soluções.

Seja S uma solução parcial, subconjunto de variáveis binárias. Diz-se que quem não está em S está livre.

Seja j um elemento de S

Considere-se que se 
$$\begin{cases} j > 0 \rightarrow z_j = 1 \\ j < 0 \rightarrow z_j = 0 \end{cases}$$

Exemplo: (n = 5)

Se  $S = \{3, 5, -2\} \rightarrow z_3 = 1, z_5 = 1, z_2 = 0$  .

Uma solução complementar de S é uma solução definida por S e mais uma especificação binária das variáveis livres. Para o S acima existem as seguintes soluções complementares

- (0, 0, 1, 0, 1)
- (0, 0, 1, 1, 1)
- (1, 0, 1, 0, 1)
- (1, 0, 1, 1, 1)

Uma solução parcial S com s elementos determina  $2^{n-s}$  soluções complementares (inclui S).

A melhor solução complementar factível é uma solução complementar que minimiza cz sobre todas as soluções complementares factíveis de S.

Sondar S significa:

- 1) Achar a melhor solução complementar factível de S

ou

- 2) Determinar que nenhuma solução complementar factível de S acarreta um valor na função objetivo menor que o da solução incumbente (a melhor até agora conhecida).

Quando uma solução parcial  $S$  foi sondada significa que todas as suas soluções complementares foram implicitamente enumeradas, ou seja, podem ser excluídas de outras considerações, exceto a melhor solução complementar factível, caso ela substitua a incumbente.

O procedimento geral do método consiste em gerar uma sequência de soluções parciais e ir tentando sondá-las. Veremos então como gerar uma sequência inteligente de soluções parciais e depois veremos como sondá-las.

### III.4.2 - GERAÇÃO DE SOLUÇÕES PARCIAIS

Uma sequência  $\langle S^v \rangle$  de soluções parciais é não-redundante se nenhuma solução complementar de uma solução parcial da sequência repete alguma solução complementar de uma solução parcial já sondada. A geração de uma sequência não-redundante de soluções parciais pode ser feita da seguinte forma: Seja inicialmente  $S^0 = \emptyset$

Se  $S^0$  puder ser sondado, então não há solução factível, ou há a melhor solução complementar factível é achada, e fim.

Se  $S^0$  não puder ser sondado, aumentar  $S^0$  dando um valor binário a uma variável livre de cada vez, até que na tentativa  $k_1$ ,  $S^{k_1}$  seja sondável.

É fácil de verificar que para garantir que a sequência  $\langle S^v \rangle$ ,  $v = k_1 + i$ ,  $i = 1, 2 \dots$  seja não-redundante é necessário e suficiente que se tenham todos os futuros  $S^v$  com pelo menos um elemento complementar a  $S^{k_1}$ . Então, para gerar  $S^{k_1+1}$  basta tomar  $S^{k_1}$  como seu último elemento multiplicado por  $-1$  e marcá-lo com um traço. A marcação traduz o fato que  $S^{k_1}$  foi sondado. Não é necessário guardar  $S^{k_1}$ .

Seja o exemplo:

$$k_1 = 3$$

$$S^3 = \{3, 5, -2\}$$

$S^3$  foi sondado  $\rightarrow S^4 = \{3, 5, \underline{2}\}$

A partir daí dois casos podem ocorrer:

a)  $S^4$  pode ser sondado

Basta então considerar que todos os complementos de  $S^3$  sem o último elemento foram implicitamente enumerados, pois sondar  $S^3$  e  $S^4$  é equivalente a sondar  $\{3, 5\}$ .

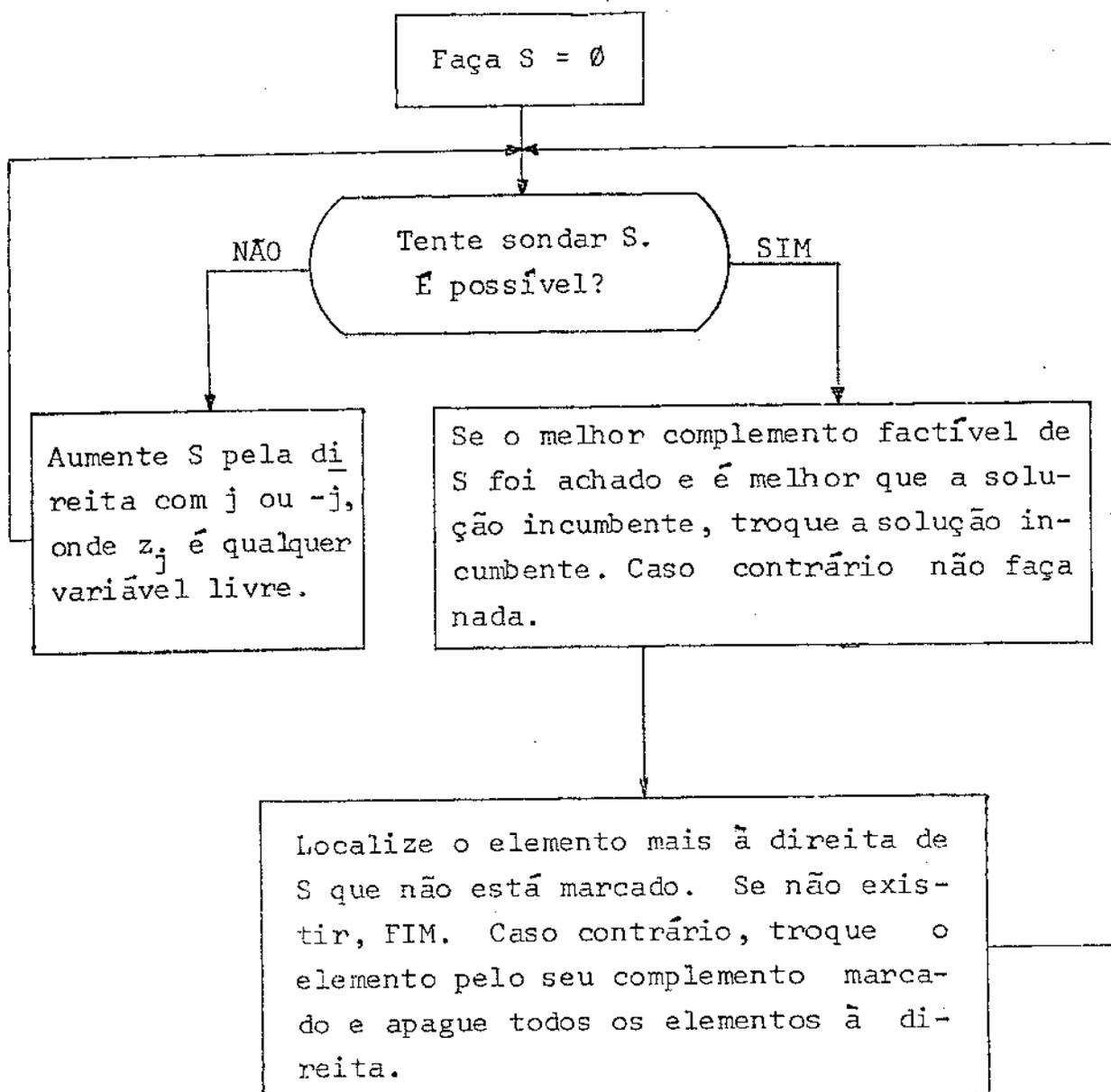
Note que para continuação da sequência,  $S^5$  já está determinado:  $S^5 = \{3, \underline{-5}\}$ .

b)  $S^4$  não pode ser sondado

Como no caso de  $S^0$ , aumentar  $S^4$  de uma variável livre binária de cada vez, tentando sondar a solução parcial resultante, até que na tentativa  $k_2$ ,  $S^{k_2}$  seja sondado.

Note que a sequência  $S^{k_1+1} \dots S^{k_2}$  é não redundante, como mostra o exemplo:  $S^4, S^5 \dots S^{k_2}$  conserva sempre  $z_2 = 1$  enquanto  $S^3$  tem  $z_2 = 0$ .  $S^{k_2}$  precisa ser armazenado e todos os  $S^v$  sucessores não só precisam ser não-redundantes em relação a  $S^{k_2}$ , como também em relação a  $S^{k_1}$ . Estes dois objetivos podem ser conseguidos fazendo  $S^{k_2+1}$  igual a  $S^{k_2}$  com o último elemento complementado e marcado. No exemplo,  $S^4$  não pôde ser sondado, então  $S^5 = \{3, 5, \underline{2}, 1\}$ , digamos. Supondo que  $S^5$  pôde ser sondado,  $S^6$  será  $\{3, 5, \underline{2}, \underline{-1}\}$ .

Até o momento, o método pode ser esquematizado da seguinte maneira:



Resta saber como sondar S.

### III.4.3 - SONDAGEM DE S

Inicialmente tentamos sondar S procurando achar a melhor solução complementar factível de S, que chamaremos de  $z^S$ .

Se  $c_j \geq 0$ , então  $z_j^S = 0$ ; se  $c_j < 0$ , então  $z_j^S = 1$ , sendo j uma variável livre. Supondo, sem perda de generalidade, que todos os custos são não-negativos (basta uma mudança de variáveis,

caso eles não sejam) basta fazermos  $z_j^S = 0$  para todas as variáveis livres e verificarmos se  $z^S$  é factível. Caso seja, S foi sondado. Caso contrário, procuramos verificar se não há complemento factível de S que melhore a solução incumbente. Verificar isso é equivalente a constatar que é impossível tirar a infactibilidade de  $z^S$  e ainda melhorar a solução incumbente. Para demonstrar a última afirmação arbitremos valor 1 para as variáveis do conjunto

$$T^S \equiv \{ j \text{ livre: } cz^S + c_j < \bar{\pi} \text{ e } a_{ij} < 0 \text{ para} \\ \text{algum } i \text{ tal que } y_i^S > 0 \},$$

onde  $y^S = Az^S + b$ ,  $\bar{\pi}$  é o valor da solução incumbente e  $a_{ij}$  é elemento da matriz A; arbitrando valor 1 para qualquer outra variável não de  $T^S$ , obteríamos uma solução de valor maior que  $\bar{\pi}$  ou não contribuiríamos para a diminuição da infactibilidade de  $z^S$ . Então, se  $T^S$  for vazio não há complemento factível de S que melhore a solução incumbente e S foi sondado. A mesma conclusão é válida se

$$y_i^S + \sum_{j \in T^S} \min \{0, a_{ij}\} > 0$$

para algum  $i$  tal que  $y_i^S > 0$ .

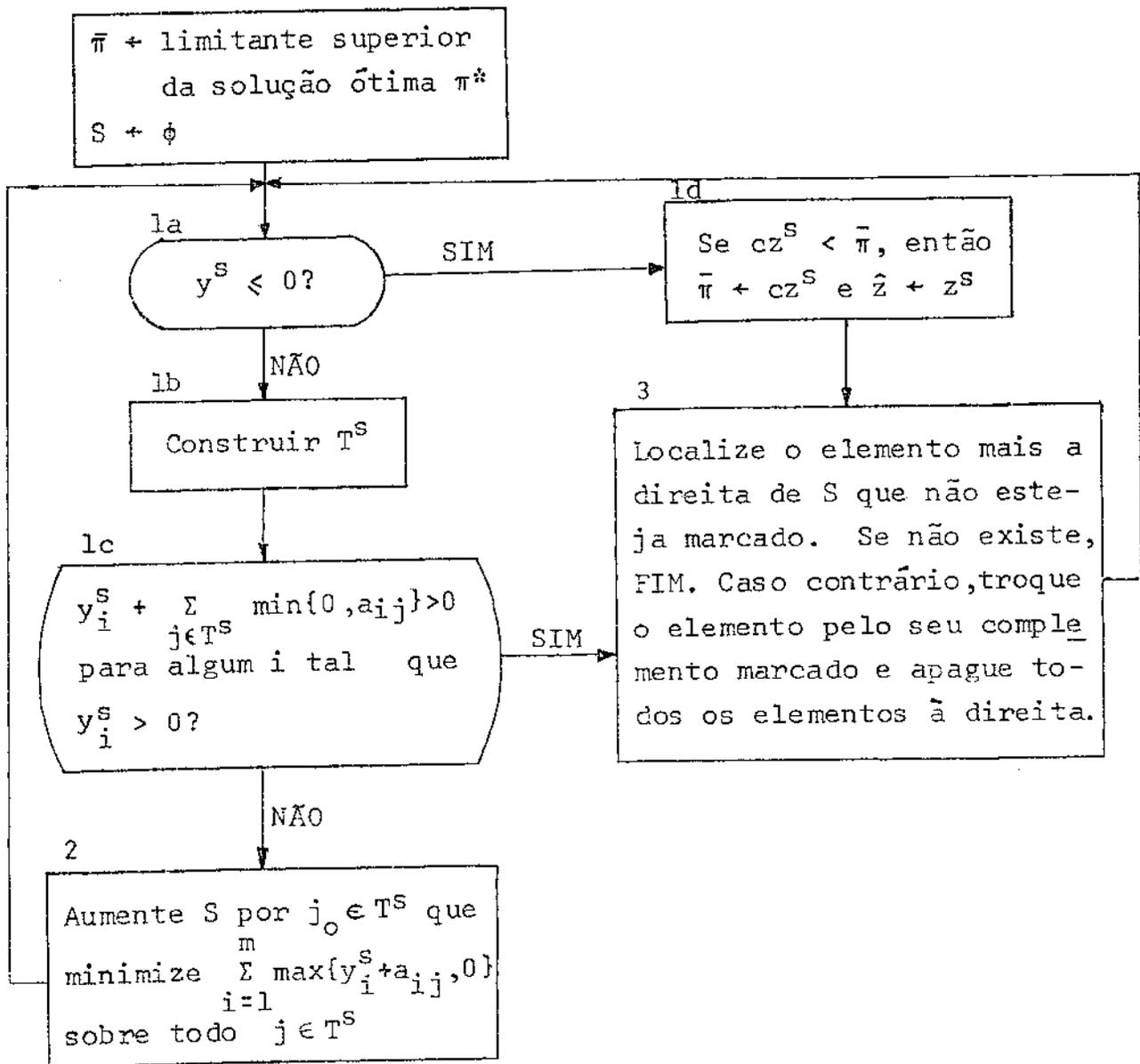
Caso S não seja sondado devemos aumentá-lo de uma variável livre. Tal variável pode ser  $j_0 \in T^S$  tal que

$$\text{minimize}_{j \in T^S} \left\{ \sum_{i=1}^m \text{máximo} \{y_i^S + a_{ij}, 0\} \right\}.$$

Com isso procuramos selecionar uma variável livre de  $T^S$  que

minimiza a infactibilidade "total" da nova solução  $z^{s+1}$ . Há outros critérios para o mecanismo de aumentar  $S$ , e.g., seleccionar o  $j_0 \in T^S$  tal que ele factibilize ao máximo a restrição mais violada, ou que  $j_0$  factibilize ao máximo o maior número de restrições infactíveis.

Em uma versão detalhada o algoritmo pode ser, sabendo que  $\pi^*$  é o valor da solução ótima e  $\hat{z}$  é a solução incumbente:



III.4.4 - EXEMPLO

Aplicamos o algoritmo acima ao problema

minimizar  $5z_1 + 7z_2 + 10z_3 + 3z_4 + z_5$

sujeito a

$$\begin{aligned}
 -z_1 + 3z_2 - 5z_3 - z_4 + 4z_5 + 2 &\leq 0 \\
 2z_1 - 6z_2 + 3z_3 + 2z_4 - 2z_5 &\leq 0 \\
 z_2 - 2z_3 + z_4 + z_5 + 1 &\leq 0 \\
 z_1, \dots, z_5 &\text{ sendo variáveis zero-um}
 \end{aligned}$$

inicialização

$$\bar{\pi} = \infty$$

$$S = \phi$$

1a. iteração

1a  $y = (2, 0, 1)$

1b  $T = \{1, 3, 4\}$

1c  $i = 1 : 2 - 7 \leq 0$

$i = 3 : 1 - 2 \leq 0$

2  $j = 1 : 1 + 2 + 1 = 4$

$j = 3 : 3 = \boxed{3}$

$j = 4 : 1 + 2 + 2 = 5$

$S^1 = \{3\}$

2a. iteração

1a  $y = \{-3, 3, -1\}$

1b  $T = \{2, 5\}$

1c  $i = 2 : -5 \leq 0$

2  $j = 2 : \boxed{0}$

$j = 5 : 1 + 1 = 2$

$S^2 = \{3, 2\}$

3a. iteração

1a  $y = \{0, -3, 0\} \leq 0 \rightarrow S^2$  sondado

1d  $cz^S = 17 < \infty; \bar{\pi} + 17; \hat{z} = (0, 1, 1, 0, 0)$

3  $S^3 = \{3, -2\}$

4a. iteração

1a  $y = \{-3, 3, -1\}$

1b  $T = \{5\}$

1c  $i = 2 : 3 - 2 > 0 \rightarrow S^3$  sondado

3  $S^4 = \{-3\}$

5a. iteração

1a  $y = (2, 0, 1)$

1b  $T = \{1, 4\}$

1c  $i = 1 : 2 - 1 - 1 = 0$

$i = 3 : 1 > 0 \rightarrow S^4$  sondado

3 FIM. A solução ótima é

$(0, 1, 1, 0, 0)$  e o valor ótimo

da função objetivo é 17.

### III.5 - IMPLEMENTAÇÃO COMPUTACIONAL E RESULTADOS

O algoritmo que resolve o PLE escolhido para ser implementado em computador foi o algoritmo alternativo descrito em II.5.2. A estrutura do programa compreende num nível, um programa principal da coordenação cuja função fundamental é ler os dados, escrever os resultados, monitorar a chamada de subrotinas e gerar novas restrições do problema mestre. Em outro nível está a subrotina que resolve o problema mestre e a subrotina que resolve o subproblema. Ainda há outras subrotinas que desempenham funções auxiliares, como a subrotina de avaliação da função objetivo, a subrotina de busca unidimensional, etc. Para esta última, usamos um método de interpolação quadrática, modificado para aproveitar a particularidade da busca se efetuar sobre um domínio restrito.

Para mostrar a exequibilidade e eficiência do enfoque de decomposição de Benders resolvendo o PLE, foram processados vários exemplos numéricos. Foram também feitas alterações nos dados e parâmetros de convergência para se analisar a sensibilidade do método a esses parâmetros.

Inicialmente foi processado um pequeno exemplo ( $m = n = 3$ ) para efeito de acerto do programa e ajuste de parâmetros. Este problema convergiu em 6 iterações. Em seguida, foram rodados 10 exemplos de dimensões maiores ( $m = 10, n = 50$ ), cujos resultados estão na tabela 2. Os dados dos problemas foram tomados dentro dos seguintes limites:

custos de transporte	-	[6,10]
custos de excesso	-	[0,9]
custos de deficit	-	[9,33]
ofertas nas origens	-	[600,900]
custos fixos de instalação	-	[6000,9000]
demandas médias	-	[50,500]

As funções de distribuição das demandas foram supostas exponenciais negativas pela facilidade de integração. O problema 1 é o fundamental e todos os outros são derivações dele. A tabela 1 mostra como a decomposição de Benders se com-

porta, iteração por iteração, neste problema. Os valores ótimos encontrados foram normalizados para efeito de comparação.

iteração	valor subproblema (LS)	valor problema mestre
1	109.62	96.56
2	101.62	82.57
3	104.03	93.85
4	101.15	86.95
5	103.68	96.07
6	102.78	86.71
7	100.54	90.06
8	103.59	92.70
9	100.00	convergência

TABELA 1

A segunda coluna registra os valores ótimos do subproblema que são limitantes superiores (LS) do problema global e a terceira coluna os valores ótimos do problema mestre resolvido em seguida. Pelo fato de ter sido utilizado o algoritmo alternativo descrito em II.5.2 esses últimos valores não significam um limitante inferior e explicam porque no ótimo não há coincidência de valores. Notar que LS segue uma sequência monótona não-crescente.

problema	$\epsilon$	$\delta$	nº de iterações	tempo de CPV		regra dual?
	%	%		total (seg.)	mestre (%)	
1	0.1	0.1	9	17.71	3	sim
2	0.1	0.001	3	20.02	1	sim
3	1	0.1	6	11.87	2	sim
4	0.1	0.1	15	30.14	4	não
5	0.1	0.1	10	29.10	1	sim
6	0.1	0.1	14	40.34	2	sim
7	0.1	0.1	11	31.07	1	sim
8	0.1	0.1	8	24.52	1	sim
9	0.1	0.001	6	18.52	1	sim
10	0.1	0.001	16	43.64	4	não

TABELA 2

Na tabela acima  $\epsilon$  é o parâmetro de convergência do

problema global e  $\delta$  é o parâmetro de convergência do subproblema de transporte estocástico, ambos em porcentagem de custo total ótimo respectivo. Em seguida é dado o número de iterações de Benders para convergência, o tempo total de CPU dispendido na execução e a porcentagem do tempo total gasto ocupada pelo problema mestre. Toda experimentação computacional foi realizada no computador PDP-10 da UNICAMP. A última coluna indica se é ou não usada a regra do item III.3 para recuperação dos multiplicadores ótimos do subproblema. Todas as rodadas foram inicializadas do mesmo modo ( $z_i = 1, \forall i \in I$ ) à exceção do problema 7. Os parâmetros de convergência internos das outras subrotinas foram tomados coerentes com os exigidos para os seus problemas fontes.

Os problemas 2, 3 e 4 usam os mesmos dados gerais que 1. A experiência com o problema 2 mostra que exigindo maior significância para a solução do subproblema (e para suas variáveis duais) a convergência é mais rápida, devido a melhor qualidade dos cortes gerados mas o tempo total fica comprometido pelo maior esforço de cálculo, ao nível do subproblema. O problema 3 mostra que o tempo de resolução é fortemente afetado pela tolerância exigida para convergir. O número de iterações e o tempo necessário para resolver o problema 4 mostra a oportunidade e eficiência do uso da regra de recuperação proposta.

O problema 5 é semelhante ao 1 à exceção dos custos de transporte tomados pela metade.

No problema 6 os custos de deficit foram aumentados em 20%. O aumento no número de iterações faz aumentar o tempo gasto com o mestre pelo fato de ele estar resolvendo problemas com crescente número de restrições. O problema 7 é o mesmo que 6, apenas inicializado num melhor limitante superior. Em 8 é tentada uma experiência, executando-se o problema 6 com  $\epsilon$  variável, ou seja, a tolerância é suposta grande no início e vai sendo restringida para o valor 0.1% no curso da otimização. Os resultados são excelentes e explicam-se devido ao menor esforço computacional requerido nas primeiras iterações. Os problemas 9 e 10 são os mesmos que 6 e comprovam o que já havia sido observado em relação ao 2 e 4.

Existem outras alternativas para tentar melhorar a performance do enfoque de Benders. Em alguns casos é vantajoso

gerar vários cortes antes de perfazer uma iteração do mestre - (útil em problemas com grande número de variáveis binárias, onde o esforço maior concentra-se no mestre) - , ou ainda estudar outras funções objetivo para o mestre da versão alternativa.

O tempo total de processamento de um problema pode ser melhorado substancialmente se a resolução de um subproblema partir da solução conseguida na iteração anterior, visto que de iteração para iteração a distribuição ótima dos fluxos não varia muito.

### III.6 - CONCLUSÕES

A única experiência computacional com a aplicação da DBG que foi levantada, não traz informações sobre tempo de execução e outros detalhes, mas mostra para problemas de variados tamanhos, aquilo que é fundamental e típico no uso da técnica de decomposição de Benders na resolução de problemas: 0 pequeno número de iterações requeridas para a convergência. O uso do Benders "linear" (subproblemas lineares) é bastante difundido principalmente na solução de problemas de programação mista e em particular nos problemas de localização - distribuição confirmando a alta eficiência na exploração das alternativas possíveis. Esta eficiência pode ser vislumbrada se encarmos a técnica de Benders como um processo automático (e inteligente) de estudo de cada alternativa de localização. Para o caso de termos 10 variáveis zero-um as alternativas possíveis são  $2^{10} = 1024$ . Como mostrado, em média, basta analisar "inteligentemente" apenas 1% dessas possibilidades.

O enfoque da DBG para solução do PLE mostrou a mesma tendência geral de convergência rápida verificada em outras aplicações do algoritmo de Benders e abre, no campo estocástico dos problemas de localização, boas perspectivas para aplicação a problemas de grande porte e em particular para problemas com multi-produtos e para modelos mais sofisticados de sistemas de distribuição.

REFERÊNCIAS BIBLIOGRÁFICAS - CAP. III

- [1] BALACHANDRAN, V. e JAIN, S., "Optimal Facility Location under Random Demand with General Cost Structure", Naval Res. Log. Quarterly 23, 421-435, (1976).
- [2] BALAS, E., "An Additive Algorithm for Solving Linear Programs with Zero-One Variables", Operations Research 13, 517-546, (1965).
- [3] BALINSKI, M.L., "Integer Programming: Methods, Uses, Computation", Management Science, Vol. 12, nº 3, 253-313, (1965).
- [4] BEALE, E.M.L., "Survey of Integer Programming", Operational Research Quarterly, Vol. 16, nº 2, 219-238, (1965).
- [5] COOPER, L., e LEBLANC, L.J., "Stochastic Transportation Problems and other Network Related Convex Problems", Naval Res. Log. Quarterly, Vol. 24, nº 2, 327-337, (1978).
- [6] ERMOLEV, YU. M. e MIRZOAKHMEDOV, F., "Direct Methods of Stochastic Programming in Inventory Control Problems", Cibernetics 6, 887-895, (1977).
- [7] ELMAGHRABY, S.E., "Allocation under Uncertainty when the Demands has continuous D.F.", Management Science, Vol. 6, nº 3, 270-294, (1960).
- [8] FERGUSON, A.R. e DANTZIG, G.B., "The Allocation of Aircrafts to Routes, An Exemple of Linear Programming under Uncertain Demand", Management Science, Vol. 3, Nº 1, 45-73, (1956).

- [9] FRANK, M. e WOLFE, P., "An Algorithm for Quadratic Programming", Naval Res. Log. Quarterly 3, 95-110, (1956).
- [10] GEOFFRION, A.M., "Duality in Nonlinear Programming: A Simplified Application Oriented Development", SIAM Review 13, 1-36, (1971).
- [11] \_\_\_\_\_, "Integer Programming by Implicit Enumeration and Balas' Method", SIAM Review n° 2, 178-190, (1967).
- [12] \_\_\_\_\_ e MARSTEN, R.E., "Integer Programming Algorithms: A Survey", in Perspectives on Optimization: A Collection of Expository Articles, Ed. A.M. Geoffrion, Addison-Wesley, 1972.
- [13] HADLEY, G., Nonlinear and Dynamic Programming, Addison Wesley, Reading, 1964.
- [14] KARMANOV, V., Programmation Mathematique, Editions Mir, Moscou, 1975.
- [15] REKLAITIS, G.V. and PHILLIPS, D.T., "A Survey of Nonlinear Programming", AIIE Transactions, Vol. 7, n° 3, 235-256, (1975).
- [16] SALKIN, H.N., Integer Programming, Addison Wesley Publ. Co., 1975.
- [17] SOARES FILHO, S., "Programação Matemática de Grande Porte: Uma Aplicação à Sistemas Hidrotérmicos de Potência", Tese de Doutorado, FEC-UNICAMP, (1978).
- [18] SZWARC, W., "The Transportation Problem with Stochastic Demand", Management Science, Vol. 11, n° 1, 33-50, (1964).

- [19] WILLIAMS, A.C., "A Stochastic Transportation Problem",  
Operations Research 11, 759-770, (1963).
- [20] ZANGWILL, W., Nonlinear Programming: A Unified Approach,  
Prentice - Hall, Englewood Cliffs, (1969).

CAPÍTULO IV

EXTENSÕES DO PROBLEMA DE LOCALIZAÇÃO ESTOCÁSTICO

IV.1 - PROBLEMA DE LOCALIZAÇÃO ESTOCÁSTICO COM RESTRIÇÕES DE RISCO

No item II.3 discutimos e apresentamos duas maneiras de se formular um problema determinístico equivalente a um problema estocástico linear. A resolução do problema de localização estocástico apresentada até agora fez uso da primeira maneira - através do problema de dois estágios de Dantzig -. Apresentaremos em seguida uma maneira alternativa de formular o PLE - através de restrições de risco - que envolve uma visão diferente de tratar a estocasticidade presente. Como já dito, neste caso otimiza-se um critério procurando-se satisfazer as restrições que contêm variáveis aleatórias, dentro de um nível pré-determinado de satisfação.

Consideremos o problema de localização já enunciado em I.5.1 porém com a demanda tendo de ser satisfeita acima de uma porcentagem requerida,  $0 \leq \beta_j \leq 1$  e geralmente elevada, Matematicamente o problema pode ser formulado como

$$\text{minimizar } \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{i \in I} k_i z_i \quad (1)$$

$x, z$

$$\text{sujeito a } \sum_{j \in J} x_{ij} \leq S_i z_i, \quad i \in I \quad (2)$$

$$\text{prob } \left[ \sum_{i \in I} x_{ij} \geq D_j \right] \geq \beta_j, \quad j \in J \quad (3)$$

$$x_{ij} \geq 0, \quad i \in I, \quad j \in J$$

$$z_i = 0, 1, \quad i \in I \quad (4)$$

$$0 \leq \beta_j \leq 1, \quad j \in J.$$

Supondo que as variáveis aleatórias  $D_j$  são independentes, com função distribuição marginal acumulada  $F_j(b_j) = \text{prob} [D_j \leq b_j]$ , as restrições (3) podem ser escritas numa forma não-probabilística. Para isso basta procurar o menor valor  $\hat{D}_j$  tal que  $F_j(\hat{D}_j) = \beta_j$ , ou seja

$$\hat{D}_j = F_j^{-1}(\beta_j).$$

As restrições (3) são então equivalentes a

$$\sum_{i \in I} x_{ij} \geq \hat{D}_j, \quad j \in J. \quad (3-a)$$

O problema (1), (2), (3-a), (4) é um problema de programação linear mista e pode ser resolvido por qualquer das técnicas de otimização apresentadas no CAP. I, e em particular pela decomposição de Benders, cuja aplicação resulta em subproblemas lineares de transporte. Deve-se notar que as funções densidade das demandas podem ser quaisquer desde que sejam independentemente distribuídas e truncadas de modo a satisfazerem  $F(\xi) = 0$  para  $\xi < 0$ .

Além disso as inversas das suas funções acumuladas precisam ser calculadas. O esforço computacional para obter as inversas das funções de probabilidade é reconhecido no campo da programação estocástica como sendo relativamente grande. Mas a competência do enfoque de restrição de risco deve ser

analisada à luz da experiência computacional e de comparações. Não é intenção nossa aqui traçar paralelos de eficiência entre os dois enfoques de tratar um problema de programação matemática estocástica, e sim apresentar duas maneiras distintas, duas filosofias diferentes de encarar tais problemas.

Porém, quando as funções densidade são particulares - por exemplo, a uniforme - os PLE com restrição de risco tornam-se muito simples e fáceis de serem resolvidos, como será mostrado em seguida. Em Reese e Stedry [4] um problema de distribuição é tratado de maneira semelhante.

Supondo então, cada demanda  $D_j$ ,  $j \in J$  como variável aleatória independente e uniformemente distribuída no intervalo  $[D_j, \bar{D}_j]$  e  $f_j(b_j)$  as suas funções densidade de probabilidade, temos

$$f_j(b_j) = \frac{1}{\bar{D}_j - D_j} \Rightarrow F_j(b_j) = \frac{b_j - D_j}{\bar{D}_j - D_j} .$$

Sendo (3) da forma  $F_j(\sum_{i \in I} x_{ij}) \geq \beta_j$ , podemos substituir o resultado anterior

$$\sum_{i \in I} x_{ij} \geq \beta_j (\bar{D}_j - D_j) + D_j, \quad j \in J.$$

Definindo

$$\tilde{D}_j = \beta_j (\bar{D}_j - D_j) + D_j$$

(3) pode ser escrita na forma determinística

$$\sum_{i \in I} x_{ij} \geq \tilde{D}_j .$$

Obedecendo o limitante superior das demandas, temos finalmente

$$\bar{D}_j \leq \sum_{i \in I} x_{ij} \leq \bar{D}_j \quad . \quad (3-b)$$

A decomposição de Benders aplicada a (1), (2), (3-b), (4) resulta em subproblemas lineares de transporte com demandas limitadas inferior e superiormente. Charnes e Klingman [2] apresentam uma maneira de resolver tal problema formulando um equivalente com variáveis limitadas onde acrescenta-se somente mais uma linha em relação à resolução do problema clássico de transporte. Seja então o subproblema de transporte resultante da fixação do vetor de localizações em  $z_i = z_i^*$ ,  $i \in I$  e considere  $I^* \in I$  tal que  $z_i^* = 1$

$$\text{minimizar} \quad \sum_{i \in I^*} \sum_{j \in J} c_{ij} x_{ij}$$

$$\text{sujeito a} \quad \sum_{j \in J} x_{ij} \leq S_i \quad , \quad i \in I^*$$

(5)

$$\bar{D}_j \leq \sum_{i \in I^*} x_{ij} \leq \bar{D}_j \quad , \quad j \in J$$

$$x_{ij} \geq 0 \quad , \quad i \in I^* \quad , \quad j \in J \quad .$$

Definindo  $I_1^* = I^* + \{m + 1\}$  e

exigindo que para cada  $j \in J$

$$\sum_{i \in I^*} x_{ij} + x_{m+1,j} = \bar{D}_j$$

podemos escrever

$$x_{m+1,j} = \bar{D}_j - D_j \quad (6)$$

Como  $\underline{D}_j \leq D_j \leq \bar{D}_j$ ,

$$0 \leq x_{m+1,j} \leq \bar{D}_j - \underline{D}_j \quad (7)$$

De (6) tiramos

$$\sum_{j \in J} x_{m+1,j} \geq \sum_{j \in J} \bar{D}_j - \sum_{i \in I^*} S_i \quad (8)$$

Finalmente, o problema (5) é equivalente a

$$\text{minimizar } \sum_{i \in I^*} \sum_{j \in J} c_{ij} x_{ij}$$

$$\text{sujeito a } \sum_{j \in J} x_{ij} \leq S_i, \quad i \in I^*$$

$$\sum_{i \in I_1^*} x_{ij} = \bar{D}_j, \quad j \in J$$

$$\sum_{j \in J} x_{m+1,j} \geq \sum_{j \in J} \bar{D}_j - \sum_{i \in I^*} S_i$$

$$0 \leq x_{m+1,j} \leq \bar{D}_j - \underline{D}_j, \quad j \in J$$

$$x_{ij} \geq 0, \quad i \in I^*, \quad j \in J$$

onde  $c_{m+1,j} = 0$ ,  $\forall j \in J$ .

O problema acima é um clássico problema de transporte com algumas variáveis limitadas e pode ser resolvido por qualquer algoritmo de fluxos ("out-of-kilter", por exemplo).

#### IV.2 - PROBLEMA DE LOCALIZAÇÃO ESTOCÁSTICO CONSIDERANDO CENTROS DE DISTRIBUIÇÃO INTERMEDIÁRIOS

No item I.5.2 já foi relevada a importância de se considerar centros intermediários num problema de distribuição. Neste item mostraremos como tratar a estocasticidade quando se consideram esses centros. Uma outra formulação clássica diferente da já apresentada é:

$$\text{minimizar } \sum_{i \in I} \sum_{k \in K} c_{ik} x_{ik} + \sum_{k \in K} \sum_{j \in J} c_{kj} x_{kj} + \sum_{k \in K} z_k t_k \quad (9.1)$$

$$\text{sujeito a } \sum_{k \in K} x_{ik} \leq S_i, \quad i \in I = \{1, 2, \dots, l\} \quad (9.2)$$

$$\sum_{k \in K} x_{kj} = D_j \quad , \quad j \in J = \{1, 2, \dots, n\} \quad (9.3)$$

$$\sum_{i \in I} x_{ik} - \sum_{j \in J} x_{kj} = 0 \quad , \quad k \in K = \{1, 2, \dots, m\} \quad (9.4)$$

$$\sum_{i \in I} x_{ik} \leq z_k C_k \quad , \quad k \in K \quad (9.5)$$

$$x_{ik} \geq 0; x_{kj} \geq 0 \quad , \quad i \in I \quad , \quad k \in K \quad , \quad j \in J$$

$$z_k = 0, 1 \quad , \quad k \in K$$

onde

$t_k$  custo fixo para operar o centro de distribuição  $k$

$C_k$  capacidade do centro de distribuição  $k$  .

e as demais variáveis e parâmetros conservam as significações já definidas. Considerando agora que as demandas  $D_j$ ,  $j \in J$  sejam variáveis aleatórias com uma dada distribuição de probabilidades, a manipulação de penalizar excessos e deficits nos destinos através da eliminação da restrição (9.3) e inclusão na função objetivo de um termo penalizador, pode ser realizada sem dificuldades. A decomposição de Benders generalizada é uma técnica que pode ser empregada para resolver (9), lembrando que para uma determinada estrutura binária de localização, o subproblema resultante é um clássico problema de entreposto ("transshipment problem") com função objetivo convexa e cujas restrições podem ser transformadas para um problema de transporte (veja, e.g., Orden [3], Wagner [5])

Podem ser encontrados sistemas de distribuição onde não se tem controle sobre a variável de quanto produzir. Até agora todos os sistemas imaginados supõem que uma firma, tendo o controle sobre o volume de produção e transporte, procura encontrar locais ótimos para instalação de plantas. Porém, se um setor governamental preocupa-se com a absorção<sup>1</sup> do excedente comercializável de uma safra agrícola, dispondo de uma rede de armazenagem e tendo de satisfazer as demandas por tais produtos, configura-se uma situação onde o controle só pode ser efetuado sobre o volume a armazenar e sobre a maneira de transportar o produto. Além disso, a quantidade a ser ofertada pelos produtores é afetada de muita incerteza, assim como as demandas nos centros consumidores. Nesses casos é necessário tratar as ofertas também como variáveis aleatórias. Escolhendo-se pesos que penalizem as discrepâncias na satisfação das restrições de oferta (os pesos por "falta" de produto nas origens deve ser muito grande), as restrições (9.2) podem ser adicionadas à função objetivo de modo semelhante às (9.3).

O enfoque de restrição de risco pode ser usado também sobre a formulação do problema (1) do CAP. I, sendo porém,  $i$  índice de fábricas,  $k$  de armazém e  $j$  de demandas.

$$\min \sum_{i \in I} \sum_{k \in K} \sum_{j \in J} c_{ikj} x_{ikj} + \sum_{k \in K} f_k z_k + \sum_{k \in K} (v_k \sum_{j \in J} D_j y_{kj})$$

$$\text{s.a. } \text{prob} \left( \sum_{k \in K} \sum_{j \in J} x_{ikj} \leq S_i \right) \geq a_i, \quad i \in I$$

(10)

$$\text{prob} \left( \sum_{i \in I} x_{ikj} \geq D_j y_{kj} \right) \geq b_j, \quad k \in K, \quad j \in J$$

$$\sum_{k \in K} y_{kj} = 1, \quad k \in K$$

---

1 O governo compromete-se a comprar o excedente através da fixação de uma política de preços mínimos, por exemplo.

$$\bar{v}_k z_k \leq \sum_{j \in J} D_j y_{kj} \leq \bar{v}_k z_k, \quad k \in K$$

$$x_{ikj} \geq 0, \quad i \in I, k \in K, j \in J.$$

#### IV.3 - EXTENSÃO PARA FUNÇÕES CUSTO DESCONTÍNUAS

No item I.5.4 foram mostradas várias maneiras de se levar em conta ganhos por economias de escala no planejamento de instalação de estabelecimentos e da distribuição ótima de seus produtos. Neste item fazemos o mesmo tipo de extensão para os problemas de localização estocástico, considerando porém uma forma mais generalizada, ou seja, permitindo que os custos de produção sejam lineares por parte, mas não-côncavos (descontínuos). Suporemos, então, que os custos de produção de uma determinada planta  $i$  sejam do tipo da figura IV.1.

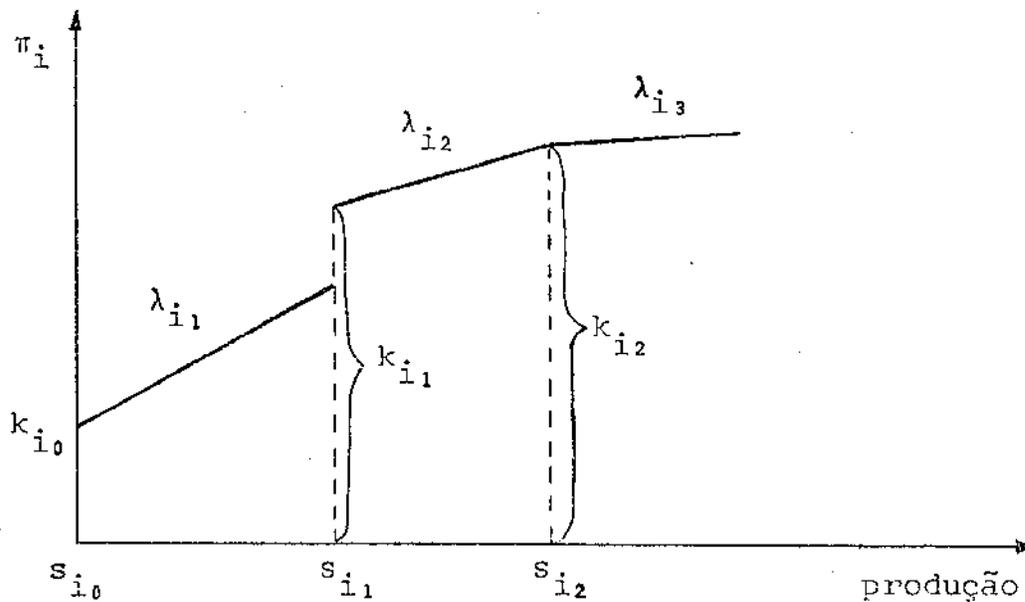


FIG. IV.1

Usando tal estrutura de custos, pode-se chegar a uma formulação do problema que permite a aplicação da DGB. Considere

- $D_j$  - demanda aleatória no destino  $j$ ,  $j \in J = \{1, 2, \dots, n\}$
- $c_{ij}$  - custo de transporte entre origem  $i$ ,  $i \in I = \{1, 2, \dots, m\}$   
e destino  $j \in J$
- $x_{ij}$  - fluxo entre  $i$  e  $j$
- $m_{ir}$  - quantidade produzida em  $i$  a um custo marginal  
 $\lambda_{ir}$ ;  $i \in I$ ,  $r \in R = \{1, 2, \dots, \ell\}$
- $z_{ir}$  - variáveis zero-um .

As variáveis  $m_{ir}$  indicam a quantidade produzida, devendo situar-se entre os limites  $s_{ir-1}$  e  $s_{ir}$  e somente uma deve ser positiva para cada  $i \in I$ . Para isso são definidas as restrições

$$z_{ir} s_{ir-1} \leq m_{ir} \leq z_{ir} s_{ir}, \quad i \in I \quad (11.1)$$

$$\sum_{r \in R} z_{ir} = 1, \quad z_{ir} = 0,1, \quad i \in I, \quad r \in R. \quad (11.2)$$

Deve ser garantido que só será transportado aquilo que for produzido, logo

$$\sum_{j \in J} x_{ij} = \sum_{r \in R} m_{ir}, \quad i \in I \quad (11.3)$$

$$x_{ij} \geq 0, \quad i \in I, \quad j \in J.$$

Usando o equivalente determinístico de dois estágios de Dantzig, apresentado II.3.2, as restrições de demanda desaparecem e a função objetivo de minimização dos custos de transporte, produção e instalação fica

$$\begin{aligned} \text{minimizar } & \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} g_j \left( \sum_{i \in I} x_{ij} \right) + \sum_{i \in I} \sum_{r \in R} \lambda_{ir} m_{ir} + \\ & \sum_{i \in I} \sum_{r \in R} z_{ir} (k_{ir-1} - \lambda_{ir} s_{ir-1}) \quad , \end{aligned} \quad (11.4)$$

onde  $g_j \left( \sum_{i \in I} x_{ij} \right)$  é a mesma função já definida em II.5 e os dois últimos termos saem de relações sobre a fig. IV.1. Usando um artifício semelhante, Busby [1] aplicou a decomposição de Benders no caso das demandas determinísticas.

No caso estocástico a aplicação da DBG é imediata. Seja então um conjunto de variáveis binárias determinado  $z_{ir}^*$ ,  $i \in I$ ,  $r \in R$  satisfazendo as restrições (11.2). É equivalente escolher um único  $r^*$  para cada  $i \in I$  tal que  $z_{ir^*} = 1$  e  $z_{ir} = 0$  para os demais. Com isso a restrição (11.3) torna-se

$$\sum_{j \in J} x_{ij} = m_{ir^*} \quad , \quad i \in I$$

e pode ser substituída em (11.1) gerando o seguinte subproblema

$$\begin{aligned} \min & \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} g_j \left( \sum_{i \in I} x_{ij} \right) + \sum_{i \in I} \sum_{j \in J} \lambda_{ir^*} x_{ij} + \\ & \sum_{i \in I} z_{ir^*} (k_{ir^*-1} - \lambda_{ir^*} s_{ir^*-1}) \end{aligned} \quad (12.1)$$

sujeito a

$$z_{ir}^* s_{ir}^* - 1 \leq \sum_{j \in J} x_{ij} \leq z_{ir}^* s_{ir}^* , \quad i \in I \quad (12.2)$$

$$x_{ij} \geq 0 , \quad i \in I , \quad j \in J$$

As restrições (12.2) podem ser transformadas em restrições simples de transporte através do mesmo artifício usado em IV.1.

A obtenção do problema mestre é imediata a partir do conhecimento das variáveis duais associadas a (12.2). Como no caso da aplicação da DBG ao problema com custos lineares, o teorema 2 do CAP. II não se aplica pois (12) será sempre factível qualquer que seja o  $z_{ir}^*$  escolhido. Aqui, a propriedade P também é válida devido a separabilidade entre as variáveis  $x$  e  $z$  na função objetivo e nas restrições. Da mesma maneira anterior os valores ótimos dos subproblemas podem ser usados para gerar os cortes do problema mestre sem esforço computacional adicional. A versão com custos de produção descontínuos introduz mais variáveis binárias ( $m \times l$  variáveis zero-um no total) e faz prever um esforço computacional concentrado na resolução do problema mestre.

REFERÊNCIAS BIBLIOGRÁFICAS - CAP. IV

- [1] BUZBY, B.R., "Nonlinear Distribution Problems", private communication, (1965).
- [2] CHARNES, A. e KLINGMAN, D., "The Distribution Problem with Upper and Lower Bounds on Node Requeriments", Management Science, 16, 638-642, (1970).
- [3] ORDEN, A., "The Transshipment Problem", Management Science, 2, 276-285, (1956).
- [4] REESE, R.M. e STEDRY, A.C., "A Chance - Constrained Distribution Problem", Naval Res. Log. Quarterly, Vol. 24, nº 1, 35-45, (1977).
- [5] WAGNER, H.M., Principles of Operations Research, Englewood Cliffs: Prentice - Hall, 1969.

APÊNDICE I

CONDIÇÕES DE ESTACIONALIDADE PARA MÍNIMO

(Condições de Kuhn - Tucker)

Seja um problema de programação não-linear com restrições

$$\begin{array}{l} \text{minimizar } f(x) \\ x \end{array}$$

sujeito a  $x \in S_1 \cap S_2 \cap S_3$

onde  $S_1 = (E^n)^+$

$$S_2 = \{x/w_i(x) = 0, \quad i = 1, 2, \dots, n_w\}$$

$$S_3 = \{x/g_i(x) \leq 0, \quad i = 1, 2, \dots, n_g\}$$

$x = n$ -vetor de variáveis reais

$f(x)$ ,  $w_i(x)$ ,  $g_i(x)$  = funções reais definidas em  $S_1$  e diferenciáveis.

Define-se a função lagrangeana

$$L(x, \lambda, u) = f(x) + \sum_{i=1}^{n_w} \lambda_i w_i(x) + \sum_{i=1}^{n_g} u_i g_i(x)$$

Dizemos que um ponto  $(x^0, \lambda^0, u^0)$  satisfaz as condições de estacionalidade\* se

$$1. \quad \nabla_x L(x^0, \lambda^0, u^0) \geq 0$$

$$x^0 \cdot \nabla_x L(x^0, \lambda^0, u^0) = 0$$

$$x^0 \geq 0$$

$$2. \quad w_i(x^0) = 0 \quad , \quad i = 1, 2, \dots, n_w$$

$$3. \quad g_i(x^0) \leq 0$$

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

$$u_i \geq 0 \quad , \quad i = 1, 2, \dots, n_g.$$

---

\* Desde que o problema original satisfaça alguma condição de regularidade.