RELAXAÇÃO LAGRANGEANA EM PROGRAMAÇÃO INTEIRA

Helenice de Oliveira Florentino

Orientador:
Prof. Dr. Marcos Nereu Arenales

Dissertação apresentada ao Instituto de Ciências Matemáticas de São Carlos - USP, como parte dos requisitos para obtenção do título de Mestre em Ciências de Computação e Matemática Computacional.

USP - São Carlos
SETEMBRO - 1990

Ao André e à minha família,
especialmente aos meus pais.

Agradeço

ao prof. Marcos Nereu Arenales pela orientação e amizade;

aos amigos de Ilha Solteira, pelo auxílio e compreensão;

aos amigos Solange, Edson e Naiara pelo estímulo e carinho;

aos amigos da sala 5, em especial a Ariovaldo e Margarida, que fizeram esta etapa mais agradável com suas companhias;

aos amigos da Pós, em especial à Marisa e Kasuo pelo apoio nos momentos difíceis;

à CAPES *pelo apoio financeiro;*

a todos que de um modo ou de outro contribuiram para a realização deste trabalho.

ÍNDICE

RESUMO i

ABSTRACT ii

INTRODUÇÃO 1

CAPÍTULO I

A programação Inteira e o Método Branch-and-Bound

1. Introdução 2
2. O Problema de Programação Linear Inteira 3
3. O Método Branch-and-Bound 3

CAPÍTULO II

Relaxação Lagrangeana

1. Introdução 7
2. Relaxação Lagrangeana e o Problema Dual 8
3. Outras Relaxações do Problema (P) e suas Relações 9
4. O Problema Perturbado e o Gap Dual 17

CAPÍTULO III

O Uso da Relaxação Lagrangeana no algoritmo Branch-and-Bound

1. A Relaxação Lagrangeana no Algoritmo Branch-and-Bound 32
2. Algoritmo Genérico 35
3. Considerações Finais..... 37

CAPÍTULO IV

Otimização do Problema Dual

1. Propriedades e Definições Básicas 38
2. Método Subgradiente 42

CAPÍTULO V

Aplicações

1. O Problema do Caixeiro Viajante 49
2. Problema de Localização de Armazéns 55
3. Problema de Designação Generalizada 60

CONCLUSÕES 64

APÊNDICE *A* 66

BIBLIOGRAFIA 69

RESUMO

Neste trabalho abordamos a teoria da relaxação lagrangeana para resolução de problemas de programação linear inteira, a qual tem sido extensivamente usada e apresentado resultados satisfatórios. Esta abordagem busca reformular um problema inteiro, fazendo deste um problema mais simples. Para tal, relaxa-se algumas restrições, colocando-as como um termo "penalidade" na função objetivo, criando assim o chamado "problema lagrangeano". É formulado o problema dual, o qual pode ser resolvido pelo método subgradiente ou variações deste.

A relaxação lagrangeana tem mostrado muita eficiência também quando usada para gerar limitantes para o algoritmo "Branch-and-Bound". Em muitos casos tais limitantes são melhores que os dado pela relaxação linear, gerando uma árvore de tamanho reduzido.

Esta técnica lagrangeana tem sido aplicada com sucesso a um grande número de problemas importantes de pesquisa operacional, por exemplo: rotas, localização, sequenciamento, designação, cobertura entre outros.

ABSTRACT

In this work we survey the lagrangean relaxation theory to solve integer linear programming problems, which has been extensively used and showed satisfactory results. This approach searches a new formulation for the original problem, in which some constraints are removed and replaced as a "penalty" term in the objective function. This new problem is called "lagrangean problem". So, the dual problem is formulated, which can be solved via the subgradient method or its variants.

The Lagrangean relaxation has proved to be efficient, when used to obtain bounds for the Branch-and-Bound algorithm. In many cases these bounds are better than those provided by the linear relaxation. In general, it yields a reduced tree.

This lagrangean technique has been successfully applied to number of important problems of operational research as, for example: routing, location, scheduling, assignment, set covering and others.

INTRODUÇÃO

A programação inteira é de muita importância, pois grande parte dos problemas da vida prática tem esta formulação e, infelizmente os algoritmos de resolução destes, em geral, são computacionalmente pouco eficientes. Assim sendo, idear tais algoritmos é uma área vigorosa de pesquisa entre especialistas e, evidências de progresso em resolver problemas reais de programação inteira estão aumentando significativamente.

Em meio a estas pesquisas está a relaxação lagrangeana, que a partir da década de 70 teve um enorme avanço e tem demonstrado ser uma excelente ferramenta para um grande número de aplicações.

Muitos problemas de programação inteira possuem um conjunto de restrições complexas que os tornam difíceis de resolver. A relaxação lagrangeana é fundamentada na observação de que tais problemas podem ser remodelados de forma que as restrições "complicadas" são relaxadas e postas como um objetivo a ser alcançado. Tal problema é chamado "problema lagrangeano" e está formalmente definido no capítulo II. As soluções ótimas dos problemas lagrangeanos fornecem limitantes para a função objetivo do problema original, os quais têm sido usados no algoritmo branch-and-bound com excelentes resultados.

Este trabalho está organizado da seguinte forma:

No capítulo I introduzimos algumas definições e terminologias que serão usadas no texto, bem como o algoritmo branch-and-bound. No capítulo II apresentamos em forma mais detalhada a teoria da relaxação lagrangeana. No capítulo III veremos o uso da relaxação lagrangeana no algoritmo branch-and-bound. No capítulo IV mostramos a resolução do problema dual através do algoritmo subgradiente e, finalizando, no capítulo V apresentamos alguns exemplos de aplicações.

CAPÍTULO I

A PROGRAMAÇÃO INTEIRA E O MÉTODO BRANCH-AND-BOUND

1. Introdução

Uma das dificuldades inerentes a problemas de programação inteira, é que não há maneira fácil de verificar se uma solução viável dada é realmente ótima. Em algumas aplicações particulares a melhor abordagem é enumerar todas as soluções e escolher a melhor, isto no caso de haver somente umas poucas soluções possíveis, tal procedimento pode ser bem mais fácil que qualquer dos algoritmos conhecidos. Mas o número de soluções viáveis não é sempre finito, e mesmo quando o seja, é comumente enorme, fazendo com que esta abordagem se torne exaustiva e na maioria das vezes impraticável. Portanto, para um algoritmo ser de uso geral, na solução de problemas de programação inteira, ele deve enumerar parcialmente um número tratável de possibilidades, e enumerar implicitamente todo o resto. Baseado nesta abordagem, algoritmos parcialmente enumerativos, é que tem-se buscado métodos com esta filosofia para resolver problemas de programação inteira. O processo mais adotado para resolver tais problemas, usa um método de busca em árvore chamado método "**Branch-and-Bound**".

O nosso objetivo neste capítulo, é definir o problema de programação linear inteira e apresentar o método Branch-and-Bound.

2. O Problema de Programação Linear Inteira

Considere o modelo

$$
\begin{array}{lll}
 & Z(P) = \min cx & (1) \\
 & \text{sujeito a} & \\
(P) & Ex \geq e & (2) \\
 & x \geq 0 & (3) \\
 & x_j \text{ inteiro, } j \in I. & (4)
\end{array}
$$

Onde $c^t \in R^n$, $x \in R^n$, $E \in R^{q \times n}$, $e \in R^q$ e I é o conjunto dos índices das variáveis restritas a assumirem valores inteiros.

Este tipo de modelo de otimização é conhecido como um **problema de programação linear inteira.**

Quando todas as variáveis são restritas a assumirem valores inteiros, o modelo é chamado de problema de programação inteira **pura**, caso contrário, é chamado de problema de programação inteira **mista**.

Dependendo da aplicação particular, a função objetivo (1) deve ser minimizada ou maximizada e as restrições (2) podem também incluir igualdades e desigualdades ($\leq$).

Na seção seguinte, veremos um método para resolução deste tipo de problema, o qual pode ser aplicado a problemas de programação inteira mista, bem como pura.

3. O Método Branch-And-Bound

Considere o problema (P), suponha que seu espaço de definição seja limitado. Assim, o número de pontos inteiros que podem

ser investigados é finito. Portanto, para resolução de tal problema, um caminho simples é enumerar todos os pontos, descartando infactibilidades e procurando a solução factível que dê o melhor valor para a função objetivo. Quando tal enumeração estiver completa, a solução ótima, se existir, será a associada com o melhor valor da função objetivo.

Para que a idéia acima se torne atrativa computacionalmente, teremos que refiná-la, fazendo com que alguns dos pontos inteiros não promissores sejam descartados sem ser testados. Este é o principal propósito do método Branch-and-Bound.

O procedimento Branch-and-Bound não trata diretamente com o problema de programação inteira (P). Ele considera um problema contínuo definido pela relaxação da integralidade do problema (P), tal relaxação é denotada por $(\bar{P})$.

A solução ótima de $(\bar{P})$, $\bar{x}$, pode não ser uma solução viável para o problema original, o que em geral acontece. Uma estratégia consiste em tomar um índice $j \in I$, cuja variável x_j tenha valor fracional na solução $\bar{x}$, e criar dois subproblemas, o que chamaremos de problemas candidatos (PC): (PC_1) é o problema (P) acrescentado a restrição ($x_j \leq [\bar{x}_j]$) e (PC_2) é (P) acrescentado a restrição ($x_j \geq [\bar{x}_j] + 1$), onde $[\bar{x}_j]$ é a parte inteira de $\bar{x}_j$. A partir daí, temos dois problemas a avaliar, (PC_1) e (PC_2). Suponha que escolhemos (PC_1), para avaliá-lo usaremos o mesmo processo visto para (P), determinamos a solução do problema contínuo $(\overline{PC_1})$, se x_j for inteiro para todo $j \in I$ temos uma solução factível para (P), guardamos tal solução, a qual será comparada com as demais soluções factíveis; caso contrário, deveremos outra vez, separar o problema em dois outros, (PC_3) e (PC_4), acrescentando as restrições $(x_j \leq [\bar{x}_j^1])$ e $(x_j \geq [\bar{x}_j^1] + 1)$ em (PC_1), com $\bar{x}^1$ solução de (PC_1). Agora temos a avaliar (PC_2), (PC_3) e (PC_4). Escolhemos um destes e repetimos todo o processo. A solução ótima para (P) será dada pelo melhor valor factível encontrado, depois de avaliados todos os problemas candidatos.

Este processo pode ser organizado através de uma estrura de árvore, que facilita e simplifica sua visualização, como podemos ver a seguir:

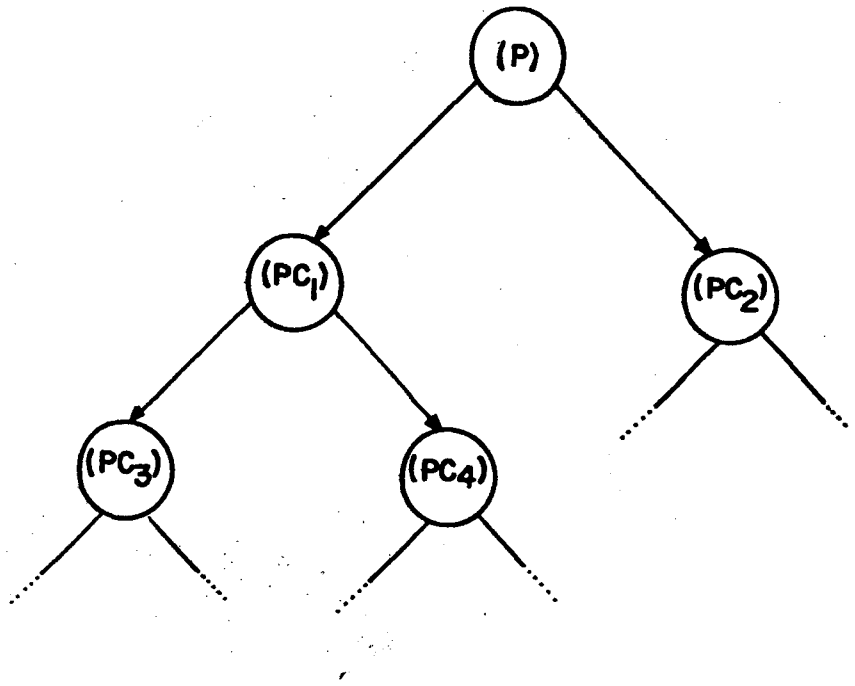

Figura 3.1 – Representação gráfica do algoritmo Branch-and-Bound.

O método, da maneira como foi descrito acima, pode se tornar bastante exaustivo. Para evitar isto, usaremos o chamado processo de sondagem, o qual descreveremos a seguir.

Suponha Z^* o melhor valor da função objetivo de (P), para as soluções factíveis conhecidas. $Z^* = \infty$ indica que não tem disponível uma solução factível.

Dizemos que um problema candidato (PC) é sondado, isto é, não há possibilidade de se encontrar soluções melhores, se ocorrer qualquer um dos ítens abaixo:

(i) (PC) é infactível (quando $(\overline{PC})$ for infactível), ou

(ii) $Z(PC) \geq Z^*$ (quando $Z(\overline{PC}) \geq Z^*$), ou

(iii) $Z(PC) < Z^*$ e a solução ótima de (PC) é conhecida (ou seja, a solução ótima de $(\overline{PC})$ satisfaz a condição de integralidade). Neste caso Z^* é atualizado.

Havendo sondagem, (PC) é retirado da lista de problemas candidatos a serem avaliados. Caso contrário ($Z(\overline{PC}) < Z^*$ e a solução ótima de $(\overline{PC})$ não sendo factível em (P)), (PC) não é sondado e sim separado em dois subproblemas, os quais serão adicionados à lista de problemas candidatos . Os problemas desta lista (nós da árvore sem ramificação e sem sondagem) continuarão a ser avaliados, com sondagem ou separação a cada passo, até que a lista se esvazie (que não haja mais nós a serem avaliados). A solução de (P), será dada pela última atualização de Z^*.

Observamos que em cada nó da árvore o valor ótimo da função objetivo de $(\overline{PC})$ limita inferiormente (no caso de (P) ser de minimização) o valor da função objetivo do problema candidato, (PC). Portanto para acelerar a sondagem devemos buscar melhores limitantes, ou seja, maiores limitantes inferiores. Tendo assim que percorrer um menor número de nós, ou melhor, avaliar um número reduzido de problemas.

Todo estudo que se tem feito com este método, é justamente com esta abordagem, aumentar o número de sondagens, tornando o processo mais rápido. Veremos no capítulo III como a relaxação lagrangeana contribui na produção de limitantes inferiores melhores.

CAPÍTULO II

RELAXAÇÃO LAGRANGEANA

1. Introdução

Neste capítulo apresentaremos uma outra técnica para resolução de um problema de programação inteira, a relaxação lagrangeana, a qual tem demonstrado muita eficiência em várias experiências práticas.

A relaxação lagrangeana consiste em "relaxar restrições complicadas", ou seja retirar as restrições que sem as quais, o problema apresenta uma estrutura especial (por exemplo uma estrutura particular em que conhecemos um bom método de resolução), colocando tais restrições como uma penalização na função objetivo.

Acredita-se que o trabalho de LORIE e SAVAGE(1955) seja a primeira aplicação publicada dos métodos de relaxação lagrangeana na solução de problemas de otimização combinatória. Um outro trabalho pioneiro é o de EVERETT(1963), mas foram os trabalhos de HELD e KARP(1970,1971) que impulsionaram, realmente, esta linha de pesquisa. O termo "relaxação lagrangeana" foi introduzido por GEOFFRION(1974).

Nas seções seguintes faremos um estudo mais detalhado da teoria da relaxação lagrangeana.

2. Relaxação Lagrangeana e o Problema Dual

Seja o problema de programação linear inteira

$$
(P) \qquad \begin{array}{l} Z(P) = \min \ cx \\ \text{sujeito a} \\ Ax \geq b \\ Bx \geq d \\ x \geq 0 \\ x_j \ \text{inteiro}, \quad j \in I, \end{array}
$$

onde $c^t \in R^n$, $x \in R^n$, $A \in R^{mxn}$, $b \in R^m$, $B \in R^{pxn}$, $d \in R^p$ e I denota o conjunto dos índices das variáveis restritas a assumir valores inteiros. Denotaremos por F(P) o conjunto das soluções factíveis de (P).

Suponhamos que o conjunto $Bx \geq d$ apresente uma estrutura especial, desta forma definimos a **Relaxação Lagrangeana** de (P) relativa ao conjunto de restrições $Ax \geq b$ e um vetor $\lambda \geq 0$, $\lambda^t \in R^m$, como :

$$
(RP_\lambda) \qquad \begin{array}{l} Z(RP_\lambda) = \min \ L(x,\lambda) \\ \text{sujeito a} \\ Bx \geq d \\ x \geq 0 \\ x_j \ \text{inteiro}, \quad j \in I, \end{array}
$$

onde $L(x,\lambda) = cx + \lambda(b - Ax)$ é denominada função Lagrangeana.

Dizemos que as restrições $Ax \geq b$, do problema (P), foram dualizadas e o problema resultante é chamado **Problema Lagrangeano**.

.....

*** Notação: $x,y \in R^n$

$x \leq y \Leftrightarrow x_i \leq y_i$, $i = 1,\ldots,n$.

Podemos observar que, para todo $x \in F(P)$ temos $Ax \geq b$ e, com $\lambda \geq 0$ então

$$L(x,\lambda) \leq cx \qquad \text{para todo } x \in F(P) \text{ e } \lambda \geq 0.$$

O termo "relaxação" geralmente é usado com a seguinte idéia: Um problema de minimização, (Q), é dito ser uma relaxação de um problema (P) se $F(Q) \supseteq F(P)$ e a função objetivo de (Q) for menor ou igual a função objetivo de (P) sobre F(P). Deste modo, claramente (RP_λ) é uma relaxação de (P) para todo $\lambda \geq 0$.

A importância de qualquer relaxação de (P) é determinada pela proximidade do valor objetivo desta e o valor objetivo de (P). Diante disto determinamos um critério pelo qual podemos medir a qualidade de uma escolha de λ, que seria feita no caso ideal com λ como solução ótima do problema:

$$\text{(D)} \qquad Z(D) = \max_{\lambda \geq 0} Z(RP_\lambda) .$$

Este problema é denominado "problema dual lagrangeano" ou também conhecido como "problema pseudo-dual". A solução deste nos fornece a mais apropriada escolha de λ, portanto sua resolução é de grande importância para nosso estudo e este é o motivo pelo qual reservamos o capítulo IV para apresentar um método com esta finalidade, o método subgradiente.

Eficientes aplicações da relaxação lagrangeana, em casos específicos, requer uma sensata partição do conjunto de restrições em dois grupos, $Ax \geq b$ e $Bx \geq d$, e uma apropriada escolha do valor $\lambda \geq 0$, como está ilustrado no capítulo V.

3. Outras Relaxações do Problema (P) e Suas Relações

Nesta seção veremos, através de alguns teoremas, as relações básicas existentes entre (P), (RP_λ) e outras relaxações de (P).

Em todo este texto denotaremos por $(\bar{.})$ o problema $(.)$ sem a condição de integralidade imposta sobre suas variáveis. Portanto $(\bar{P})$ será definido como :

$$
(\bar{P}) \qquad \begin{aligned} Z(\bar{P}) = \min\ & cx \\ \text{sujeito a}\ & \\ Ax &\geq b \\ Bx &\geq d \\ x &\geq 0. \end{aligned}
$$

Definimos $\bar{\lambda}$ como o vetor multiplicador ótimo correspondente às restrições $Ax \geq b$ em $(\bar{P})$, quando este tem solução ótima finita.

Definimos o problema (P^*) como sendo :

$$
(P^*) \qquad \begin{aligned} Z(P^*) = \min\ & cx \\ \text{sujeito a}\ & \\ Ax &\geq b \\ x &\in Co\{\ x \geq 0 : Bx \geq d \ \text{ e } \ x_j \text{ inteiro, } j \in I\ \}, \end{aligned}
$$

onde Co denota o casco convexo de um conjunto, isto é, o menor convexo que contém um dado conjunto . Embora seja muito difícil expressar este casco convexo como um conjunto explícito de restrições lineares, vale ressaltar que o problema (P^*) é um problema de programação linear.

Definimos λ^* como o vetor multiplicador ótimo correspondente às restrições $Ax \geq b$, quando (P^*) tem solução ótima finita.

Convencionalmente o valor ótimo de um problema de otimização infactível será $+\infty$ se este for de minimização e $-\infty$ no caso de maximização.

Notamos que o problema original (P) é equivalente a resolver um problema da forma

$$
\begin{aligned} \min\ & cx \\ \text{sujeito a}\ & \\ x &\in Co\{\ x \geq 0 : Ax \geq b,\ Bx \geq d \text{ e } x_j \text{ inteiro, } j \in I\}, \end{aligned}
$$

o qual é um problema linear onde sua solução ótima é um valor inteiro e coincide com a solução ótima de (P).

Para melhor visualização dos problemas antes definidos, assim como de suas regiões de factibilidade, exemplificamos com as figuras 3.1.

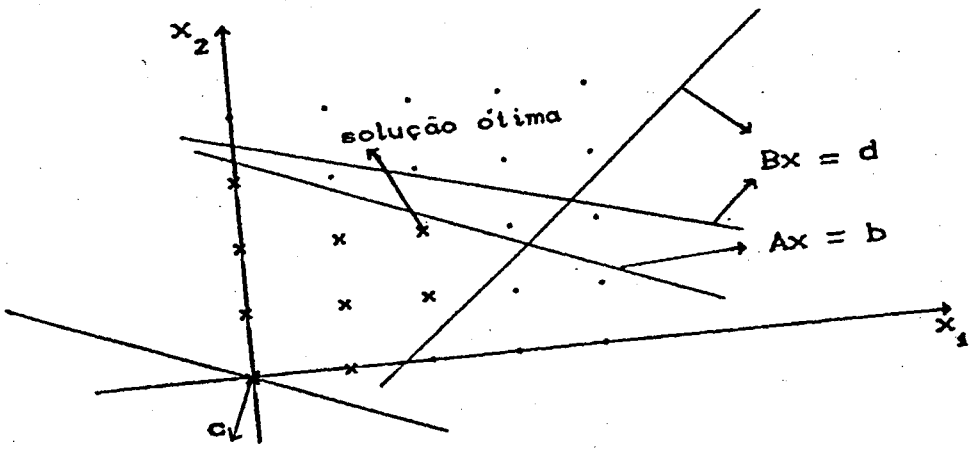

Figura 3.1.a - Os pontos marcados com x representam as soluções factíveis de um dado problema (P), ou seja, representam o conjunto F(P).

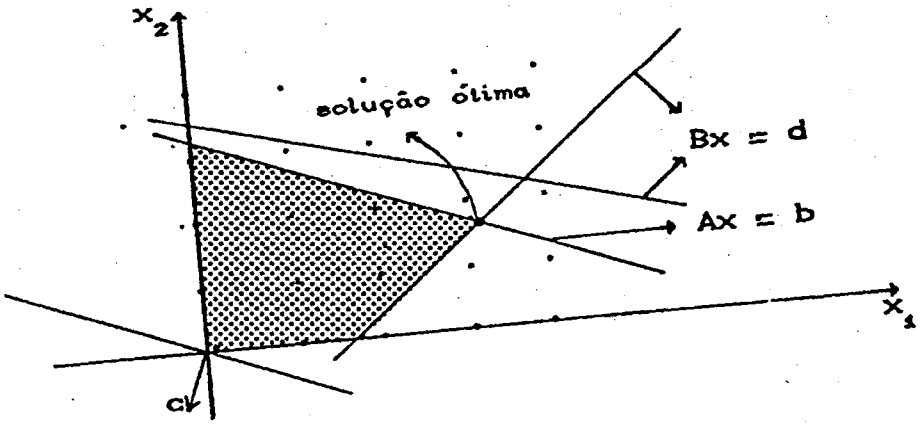

Figura 3.1.b - A parte sombreada da figura representa a região de factibilidade $F(\bar{P})$, do problema $(\bar{P})$.

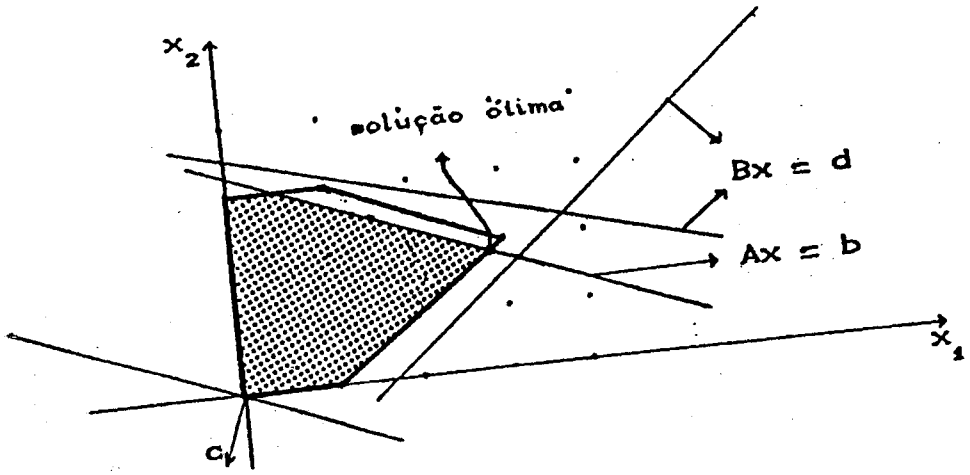

Figura 3.1.c – A região sombreada representa $F(P^*)$, e o polígono em traços mais fortes juntamente com seu interior reprenta o $Co\{x \geq 0 : Bx \geq 0,$ e x_j inteiro, $j \in I\}$.

Com as definições vistas, podemos enunciar o seguinte teorema , o qual descreve algumas das relações básicas entre (P), (RP_λ), (D), (P^*) e $(\bar{P})$.

Teorema 3.1 (GEOFFRION (1974))

(a) $F(\bar{P}) \geq F(P^*) \geq F(P)$,
$Z(\bar{P}) \leq Z(P^*) \leq Z(P)$,
$F(RP_\lambda) \geq F(P)$,
$Z(RP_\lambda) \leq Z(P) \qquad \forall\ \lambda \geq 0.$

(b) Se $(\bar{P})$ for factível, então

$$Z(\bar{P}) \leq Z(RP_{\bar{\lambda}}).$$

(c) Dado $\lambda^o \geq 0$, suponha que exista um vetor x^o satisfazendo as seguintes três condições:

(i) x^o é ótimo de (RP_{λ^o});

(ii) $Ax^o \geq b$;

(iii) $\lambda^o\ (\ b - Ax^o\ ) = 0$;

então x^o é solução ótima de (P). Se x^o satisfizer (i),

(ii) mas não (iii), então dizemos que x^o é uma solução ε-ótima de (P) com $\varepsilon = \lambda^o (Ax^o - b)$.

(d) Se (P^*) for factível então

$$Z(D) \equiv \max_{\lambda \geq 0} Z(RP_\lambda) = Z(RP_{\lambda^*}) = Z(P^*).$$

Prova:

(a) Segue direto das definições dos problemas.

(b)
$$Z(\bar{P}) \overset{*}{=} \max_{\lambda \geq 0} Z(\overline{RP_\lambda}) \qquad (* \text{ propriedade 1 do apêndice } A)$$

$$\overset{*}{=} Z(\overline{RP_{\bar{\lambda}}}) \qquad (* \text{ por definição de } \bar{\lambda})$$

$$\overset{*}{\leq} Z(RP_{\bar{\lambda}}) \qquad (* \text{ pois } F(\overline{RP_{\bar{\lambda}}}) \supseteq F(RP_{\bar{\lambda}})).$$

(c) As condições (i) e (ii) garantem a factibilidade de x^o em (P), pois de (i) x^o satisfaz $Bx^o \geq d$, $x^o \geq 0$ e x^o_j inteiro para $j \in I$ e de (ii) temos $Ax^o \geq b$.

Da condição (i) temos também que

$Z(RP_{\lambda^o}) = \min cx + \lambda^o(b - Ax) = cx^o + \lambda^o(b - Ax^o)$,

logo

$cx^o + \lambda^o(b - Ax^o) \leq cx + \lambda^o(b - Ax) \quad \forall\, x \in F(P)$.

De (iii) sabemos que $\lambda^o(b - Ax^o) = 0$; portanto a inequação anterior torna-se

$cx^o \leq cx + \lambda^o(b - Ax) \qquad \forall\, x \in F(P)$,

como $\lambda^o \geq 0$ e $(b - Ax) \leq 0 \qquad \forall\, x \in F(P)$, segue que $\lambda^o(b - Ax) \leq 0$.

Assim

$$cx^o \leq cx \qquad \forall\, x \in F(P). \qquad \text{(II)}$$

Da factibilidade de x^o e de (II) podemos concluir que x^o é ótimo para (P).

(d) $Z(D) \overset{d}{=} \max_{\lambda \geq 0} Z(RP_\lambda)$.

Podemos mostrar que (ver propriedade 2 no apêndice *A*)

$$Z(RP_\lambda) = [\min_x \; cx + \lambda(b - Ax)$$
$$\text{sujeito a}$$
$$x \in Co\{x \geq 0 : Bx \geq d \text{ e } x_j \text{ inteiro}, \; j \in I\}].$$

daí temos

$$\max_{\lambda \geq 0} Z(RP_\lambda) = \max_{\lambda \geq 0} [\min_x cx + \lambda(b - Ax)$$
$$\text{sujeito a}$$
$$x \in Co\{x \geq 0 : Bx \geq d \text{ e } x_j \text{ inteiro}, \; j \in I\}]$$

$\overset{*}{=} Z(P^*)$ (* ver propriedade 1, apêndice *A*).

Portanto

$$Z(D) \overset{d}{=} \max_{\lambda \geq 0} Z(RP_\lambda) \doteq Z(RP_{\lambda^*}) = Z(P^*).$$

■

A parte (a) do teorema acima mostra as mais óbvias relações entre (P) e suas relaxações $(\bar{P})$, (P^*) e (RP_λ).

A parte (b) mostra que a Relaxação Lagrangeana pode produzir resultados melhores ou no mínimo iguais à relaxação padrão $(\bar{P})$ em termos da proximidade de Z(P). Este fato sugere o uso da relaxação lagrangeana para determinação de limitantes no algoritmo Branch-and-Bound, que será explorado no próximo capítulo.

A parte (c) mostra as condições suficientes para que uma solução de (RP_λ) seja solução ótima ou quase-ótima de (P).

Com estas observações notamos que a Relaxação Lagrangeana é de interesse não somente para determinar limitantes inferiores para Z(P), mas também possibilita a determinação de solução ótima ou "quase-ótima" para o problema original (P). A qualidade de uma solução factível obtida pela Relaxação Lagrangeana é dada pelo "gap" dual entre Z(P) e Z(D), que é a diferença Z(P) - Z(D).

A parte (d) estabelece que o melhor valor de λ determinado através da otimização de (D), nestas condições é justamente λ^*, variável dual associada a $Ax \geq b$ no problema (P^*). Assim, a posição de $Z(P^*)$ no intervalo $[Z(\bar{P}), Z(P)]$ é uma referência para

a análise do potencial da relaxação lagrangeana aplicado a uma classe de problemas particulares.

A seguir apresentaremos um teorema, o qual mostrará relações adicionais entre $Z(\bar{P})$ e $Z(P^*)$ e entre $Z(P^*)$ e $Z(P)$. Mas, para enunciá-lo, precisaremos da seguinte propriedade:

Propriedade 3.1 (Propriedade de integralidade)

O problema (RP_λ) possui a propriedade de integralidade se $Z(RP_\lambda)$ não for alterada quando retirarmos a condição de integralidade sobre as variáveis de (RP_λ), ou seja, $Z(RP_\lambda) = Z(\overline{RP_\lambda})$ para todo $\lambda \geq 0$.

Teorema 3.2 (GEOFFRION (1974))

Suponha que $(\bar{P})$ seja factível e que (RP_λ) possua a propriedade de integralidade, então (P^*) é factível e

$$Z(\bar{P}) = Z(RP_{\bar{\lambda}}) = Z(D) = Z(RP_{\lambda^*}) = Z(P^*).$$

Prova :

Em vista dos itens (b) e (d) do Teorema 3.1, basta mostrar que $Z(\bar{P}) = Z(P^*)$.

Por definição

$$Z(\overline{RP_\lambda}) \stackrel{d}{=} \min cx + \lambda(b - Ax)$$
$$\text{sujeito a}$$
$$Bx \geq d$$
$$x \geq 0,$$

logo

$$\max_{\lambda \geq 0} Z(\overline{RP_\lambda}) = \max_{\lambda \geq 0} [\min cx + \lambda(b - Ax)$$
$$\text{sujeito a}$$
$$Bx \geq d$$
$$x \geq 0 \ ]$$

$$\overset{*}{=} \max_{\lambda \geq 0} \left[\max_{\substack{v \geq 0 \\ vB \leq c - \lambda A}} \lambda b + vd \right. \qquad (\ * \text{ dualidade em P.L})$$

$$= \max_{\substack{\lambda A + vB \leq c \\ \lambda \geq 0,\ v \geq 0}} \lambda b + vd$$

$$\overset{*}{=} \min_{\substack{\text{sujeito a} \\ Ax \geq b \\ Bx \geq d \\ x \geq 0,}} cx \quad \overset{d}{=} Z(\bar{P}) \qquad (\ * \text{ dualidade em P.L})$$

então $Z(\bar{P}) = \max_{\lambda \geq 0} Z(\overline{RP_\lambda})$.

Como (RP_λ) possui a propriedade de integralidade

$$Z(\bar{P}) = \max_{\lambda \geq 0} Z(RP_\lambda) \overset{*}{=} \max_{\lambda \geq 0} \left[\min_{\substack{x \\ \text{sujeito a} \\ x \in co\{x \geq 0 \,:\, Bx \geq d \text{ e } x_j \text{ inteiro}, j \in I\}}} cx + \lambda(b - Ax) \right]$$

(* ver propriedade 2 no apêndice *A*)

$$\overset{**}{=} Z(P^*) \qquad (\ ** \text{ por dualidade}).$$

■

Pelo teorema acima notamos que, quando (RP_λ) possui a propriedade de integralidade, a melhor escolha de λ para (RP_λ) coincidirá com $\bar{\lambda}$, o multiplicador ótimo de $(\bar{P})$. Nestas circunstâncias a relaxação lagrangeana tem sido questionada, a menos que se determine uma solução quase-ótima para (D) mais rapidamente do que se faria resolvendo $(\bar{P})$ por métodos de programação linear. Mas, em geral é mais promissor usar relaxação lagrangeana quando (RP_λ) não possuir tal propriedade [GEOFFRION(1974) e MACULAN(1984)]. Em casos específicos, podemos construir a partição das restrições de (P), sem que a relaxação possua esta propriedade como veremos na seção 3 do capítulo V.

4. O Problema Perturbado e o Gap Dual

Até aqui vimos que, partindo de um problema linear inteiro de minimização (P), cujas restrições apresentam uma estrutura complexa, tornando-o difícil de resolver, podemos dualizar um subconjunto das restrições, $Ax \geq b$, colocando-as como termo penalidade na função objetivo, obtendo assim um problema com um conjunto de restrições, as restantes, com uma estrutura especial. Desta forma definimos a relaxação lagrangeana (RP_λ) de (P).

Mas, com isto surgem duas questões:

- Como determinar bons multiplicadores λ ?
- E em que condições temos garantido uma solução ótima para (P) a partir de (RP_λ) ?

Mostramos que $Z(RP_\lambda)$ é limitante inferior para Z(P), teorema 3.1 (a). Portanto um bom multiplicador seria aquele que otimizasse o problema (D), ou seja, que fizesse com que a função objetivo $Z(RP_\lambda)$ aproximasse o máximo possível de Z(P). Este assunto será tratado em mais detalhe no capítulo IV, o que nos responderá como determinar bons multiplicadores.

As condições em que temos uma solução ótima para (P), foram dadas pelo ítem (c) do teorema 3.1. As quais equivalem a dizer que este novo problema, (RP_λ), resolve o problema original se a sua solução ótima for factível para (P), ou seja, se satisfizer as restrições relaxadas $Ax \geq b$ e seu valor objetivo ótimo for igual a Z(P).

Voltaremos nossa atenção agora para as condições em que temos a igualdade das funções objetivos primal e dual, ou seja, em que temos o gap dual (diferença entre as funções objetivos primal e dual) nulo . Para isto usaremos alguns teoremas e definições que serão apresentados a seguir.

Definimos o problema perturbado (PPy), que é o problema (P) com uma perturbação y no vetor b do conjunto de restrições $Ax \geq b$, da seguinte forma

$$
(PPy) \qquad \begin{aligned} Z(PPy) = \min\ & cx \\ \text{sujeito a}\ & \\ & Ax \geq b - y \\ & Bx \geq d \\ & x \geq 0 \\ & x_j \text{ inteiro},\ j \in I, \end{aligned}
$$

onde $y \in R^m$.

Note que para $y = 0$ o problema (PPy) coincide com o problema (P), então $Z(PP_0) = Z(P)$.

Definição 4.1:

A função Z(PPy) que associa cada vetor de perturbação y o valor ótimo do problema perturbado, (PPy), é denominada **função b-perturbação** associada ao problema (P).

O domínio de Z(PPy) é dado pela condição de factibilidade do problema (PPy).

Dom Z(PPy) = { y / F(PPy) é não vazio }.

A função Z(PPy) é o objeto chave para se entender quando Z(P) = Z(D), então para melhor conhecê-la, ilustramos com a figura 4.1 as funções b-perturbação para os casos de problemas de programação inteira pura, mista e também para o problema contínuo, em seguida estudaremos seu comportamento através de alguns teoremas.

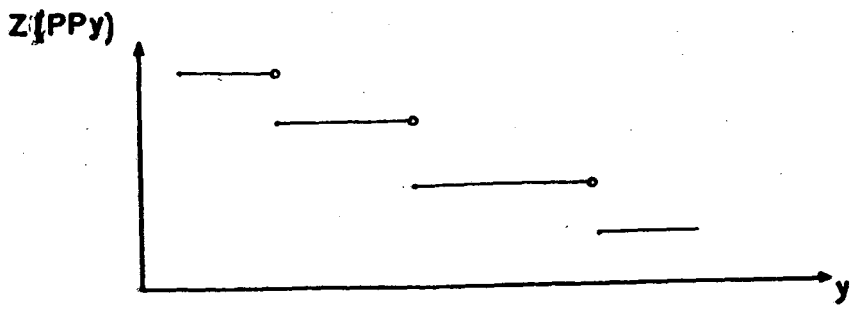

Figura 4.1.a - Ilustração da função b-perturbação associada a um problema de programação linear inteiro puro.

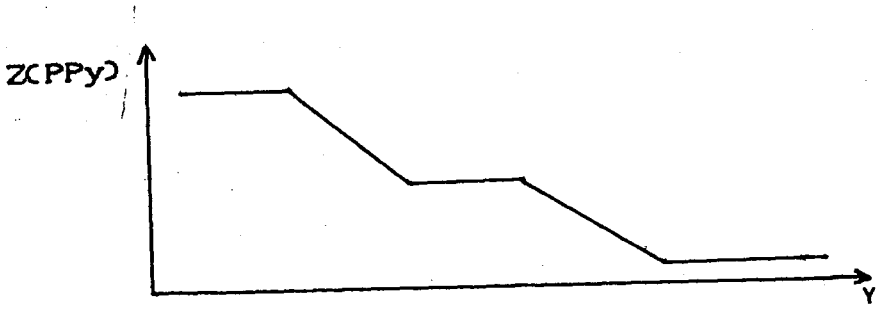

figura 4.1.b - Ilustração da função b-perturbação associada a um problema de programação linear inteiro misto.

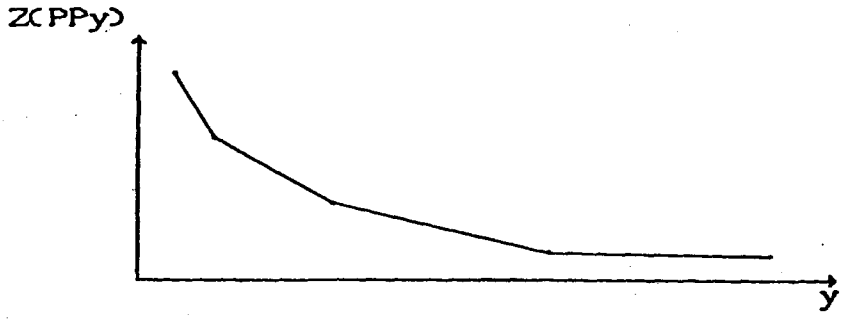

Figura 4.1.c - Ilustração da função b-perturbação associada a um problema de programação linear.

A descontinuidade pode ocorrer quando uma solução isolada (devido a integralidade), torna-se factível com a variação de y (pelo menos uma componente de y deve ter sido aumentada para "captar" esta solução) e passa a ser ótima de modo que o valor de Z(PPy) é dado por $c\bar{x}$ onde $\bar{x}$ é a solução incorporada à região de factibilidade. As idéias acima podem ser confirmadas pelos teoremas abaixo.

Teorema 4.1

A função Z(PPy) é semi contínua inferior.

Teorema 4.2

A função Z(PPy) é não crescente.

Prova:

Sejam $y_1, y_2 \in \text{Dom}(PPy)$ tal que $y_1 \leq y_2$ (I)

Temos que

$$Z(PPy_1) \stackrel{d}{=} \min[cx \ , \ x \in F(PPy_1) \ ] = cx_1 \ ,$$

o que implica em $b - Ax_1 \leq y_1$. (II)

De (I) e (II) temos $b - Ax_1 \leq y_2$ e portanto $x_1 \in F(PPy_2)$.

Por outro lado temos

$Z(PPy_2) \stackrel{d}{=} \min[\ cx \ , \ x \in F(PPy_2) \ ] \leq cx \qquad \forall\ x \in F(PPy_2)$.

Desde que $x_1 \in F(PPy_2)$ segue

$$Z(PPy_2) \leq cx_1 = Z(PPy_1). \ \blacksquare$$

Para um dado $\lambda^o \geq 0$ o conjunto solução do problema (RP_{λ^o}) é dado pelos valores de $x^o \in F(RP_{\lambda^o})$ tal que x^o minimiza a função $L(x,\lambda^o)$. EVERETT(1963) mostrou que cada uma das soluções do problema (RP_{λ^o}) está associada à solução de um problema perturbado, como veremos a seguir.

Teorema 4.3 (EVERETT(1963))

Suponha que x^o resolva (RP_{λ^o}), $\lambda^o \geq 0$; então x^o resolve o problema perturbado (PPy)

$$
(PPy) \qquad \begin{aligned} Z(PPy) = \min\ & cx \\ & Ax \geq b - y \\ & Bx \geq d \\ & x \geq 0 \\ & x_j \text{ inteiro, } j \in I, \end{aligned}
$$

onde $y_i = b_i - a_i x^o$ se $\lambda_i^o > 0$;

$y_i \geq b_i - a_i x^o$ se $\lambda_i^o = 0$, $i=1,\ldots,m$;

(a_i, $i = 1,\ldots,m$, sao linhas da matriz A).

Prova:

Como x^o resolve (RP_{λ^o}) temos

$$cx^o + \lambda^o(b - Ax^o) \leq cx + \lambda^o(b - Ax) \qquad \forall\ x \in (RP_{\lambda^o}),$$

ou seja

$$cx^o \leq cx + \lambda^o[(b - Ax)-(b - Ax^o)] \qquad \forall\ x \in (RP_{\lambda^o}).$$

Assim, desde que $\lambda^o \geq 0$, segue que

$$cx^o \leq cx, \qquad \text{(I)}$$

para todo x tal que: $x \in F(RP_{\lambda^o})$ e (II)

$$Ax \geq b - (b - Ax^o) \geq b - y, \qquad \text{(III)}$$

onde y é dado no enunciado do teorema. As condições (II) e (III)

descrevem a factibilidade do problema perturbado. Desde que x^o satisfaça (II) e (III) e como (I) se verifica, segue o teorema. ∎

Seja a função b-perturbação, Z(PP'y), associada com o seguinte problema convexo

$$(p') \qquad Z(P') = \min \; f(x) \quad \text{sujeito a} \quad g(x) \leq 0, \; x \in S,$$

onde $x \in R^n$, $S \subseteq R^n$, $f : S \longrightarrow R$ e $g : S \longrightarrow R^m$. Para o problema (PP'y)

$$(PP'y) \qquad Z(PP'y) = \min f(x) \quad \text{sujeito a} \quad g(x) \leq y, \; x \in S,$$

Dom Z(PP'y) = { y / ∃ x ∈ S; g(x) ≤ y }.

Temos o seguinte teorema:

Teorema 4.4

Se $g_i(x)$, $i = 1,\ldots,m$ forem funções convexas sobre S convexo, então Dom Z(PP'y) será convexo sobre S. Se além disso f(x) for convexa, então Z(PP'y) será convexa sobre domínio convexo.

Prova:

As hipóteses são:

(1) convexidade de f(x);

(2) convexidade de g(x);

(3) convexidade de S.

Mostraremos inicialmente a tese da convexidade do Dom Z(PP'y)

$$y_1 \in \text{Dom } Z(PP'y) \iff \exists\, x_1 \in S \,/\, g(x_1) \le y_1 \;;$$
$$y_2 \in \text{Dom } Z(PP'y) \iff \exists\, x_2 \in S \,/\, g(x_2) \le y_2 \;;$$

tomando uma combinação convexa desses pontos ($0 \le \tau \le 1$; $\tau + \bar{\tau} = 1$) temos:

$$y = \tau y_1 + \bar{\tau} y_2 \ge \tau g(x_1) + \bar{\tau} g(x_2) \ge g(\tau x_1 + \bar{\tau} x_2).$$

A primeira desigualdade segue do fato que τ e $\bar{\tau}$ são não negativos e a segunda devido a hipótese (2).

Desde que S é convexo (hipótese (3)), segue que $\tau x_1 + \bar{\tau} x_2 \in S$, logo $y \in$ Dom Z(PP'y), o que demonstra a primeira parte do teorema.

Para se mostrar a convexidade de Z(PP'y), lembrando-se que uma função é convexa se somente se tem epigrafo convexo, analisaremos a convexidade de:

$$H = \{ (t,y) \,/\, t \ge f(x),\ y \ge g(x),\ x \in S \},$$

que é o epigrafo de Z(PP'y).

Sejam $(t_1,y_1) \in H$ e $(t_2,y_2) \in H$, $0 \le \tau = 1 - \bar{\tau} \le 1$.

A combinação convexa $(t,y) = (\tau t_1 + \bar{\tau} t_2,\ \tau y_1 + \bar{\tau} y_2)$

é tal que

$\tau t_1 + \bar{\tau} t_2 \ge \tau f(x_1) + \bar{\tau} f(x_2) \ge f(\tau x_1 + \bar{\tau} x_2)$, devido à hipótese (1) e

$\tau y_1 + \bar{\tau} y_2 \ge \tau g(x_1) + \bar{\tau} g(x_2) \ge g(\tau x_1 + \bar{\tau} x_2)$, devido a (2) e

$\tau x_1 + \bar{\tau} x_2 \in S$, devido a (3).

Donde se conclue que $(t, y) \in H$. Assim H é convexo e consequentemente Z(PP'y) também é convexa. ∎

Pelo teorema acima concluimos que a função b-perturbação $Z(PP^*y)$ associada com o problema (P^*) é convexa.

Teorema 4.5

Dado $\lambda^o \geq 0$, seja x^o solução de $\min_{x \in F(RP_{\lambda^o})} L(x, \lambda^o)$,

Então

$$Z(PPy) \geq Z(PPy^o) - \lambda^o(y - y^o), \qquad \forall \ y \in \text{Dom } Z(PPy),$$

onde $y^o = b - Ax^o$.

Prova:

Como x^o minimiza $L(x,\lambda^o)$ para $\lambda^o \geq 0$, temos:

$$cx^o + \lambda^o(b - Ax^o) \leq cx + \lambda^o(b - Ax), \qquad \forall \ x \in F(RP_{\lambda^o})$$

ou

$$cx \geq cx^o - \lambda^o((b - Ax) - (b - Ax^o)), \qquad \forall \ x \in F(RP_{\lambda^o}).$$

Como $b - Ax^o = y^o$ e pelo teorema de EVERETT (teorema 4.3) $cx^o = Z(PPy^o)$ temos

$$cx \geq Z(PPy^o) - \lambda^o((b - Ax) - y^o) \quad \forall \ x \in F(RP_{\lambda^o}).$$

Mas como $\lambda^o \geq 0$ segue-se que para todo $y \in \text{Dom } Z(PPy)$, podemos escrever

$$cx \geq Z(PPy^o) - \lambda^o(y - y^o), \qquad \forall \ x \in F(RP_{\lambda^o}) \quad \text{tal que}$$

$b - Ax \leq y$.

Tomando a expressão acima para x que minimiza cx sujeito a $x \in F(RP_{\lambda^o})$, $b - Ax \leq y$, ou seja, x que resolve o problema perturbado, segue

$$Z(PPy) \geq Z(PPy^o) - \lambda^o(y - y^o), \qquad \forall \ y \in \text{Dom } Z(PPy). \blacksquare$$

Deste teorema podemos deduzir importante interpretação geométrica que relaciona $Z(PPy)$, $Z(RP_{\lambda^o})$, $L(x^o,\lambda^o)$, $Z(P)$ e $Z(D)$.

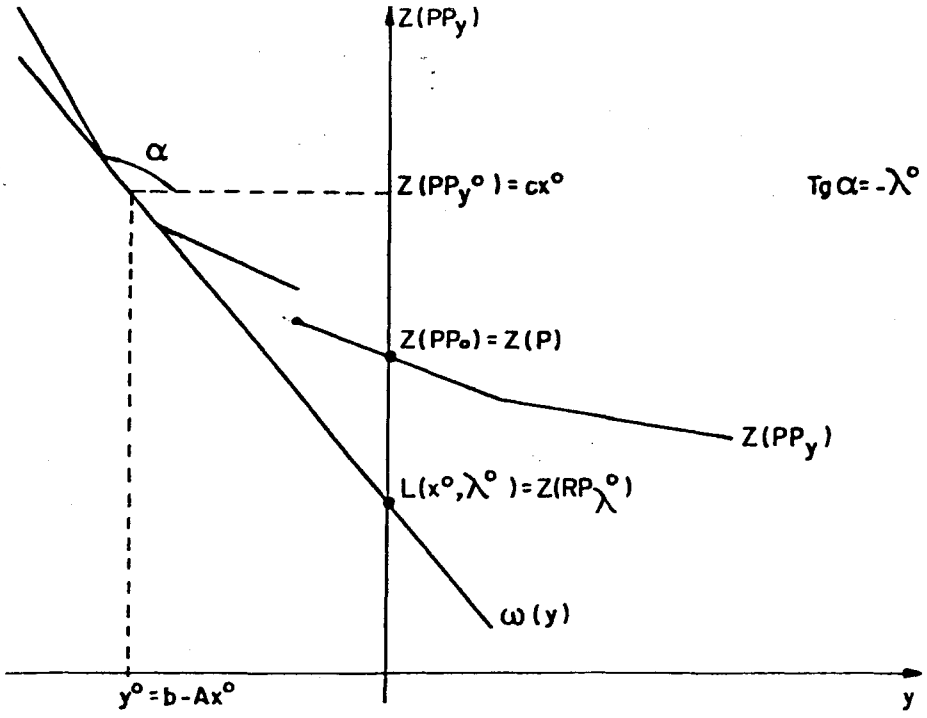

figura 4.2 - Interpretação geométrica do teorema 4.1 .

Pela figura 4.2 notamos que resolver o problema lagrangeano para $\lambda^o \geq 0$ significa geometricamente determinar o hiperplano suporte à função b-perturbação com inclinação $-\lambda^o$. O ponto suporte é dado por ($b-Ax^o$, cx^o) onde x^o é solução da relaxação lagrangeana (RP_{λ^o}) de (P). A equação do hiperplano é dada por

$$W(y) = Z(PPy^o) - \lambda^o(y - y^o) = cx^o - \lambda^o(y - (b - Ax^o)) = L(x^o, \lambda^o) - \lambda^o y .$$

A intersecção do hiperplano suporte com o eixo y = 0 fornece o valor de $Z(RP_{\lambda^o})$.

Assim, o problema dual pode ser visualizado geometricamente como achar a inclinação $-\lambda^o$ do hiperplano suporte à função b-perturbação que maximiza a intersecção com o eixo y = 0. Lembrando que Z(P) = Z(PPo), intuitivamente podemos esperar que, com a referida inclinação, a solução da relaxação lagrangeana é uma solução ótima ou quase ótima do problema primal, o que de fato é verdadeiro conforme o teorema 4.6, a seguir.

Definição 4.2:

Um vetor γ, tal que $\gamma^t \in R^m$, é um **subgradiente global** de $Z(PPy)$ em $\bar{y} \in \text{Dom } Z(PPy)$, se

$$Z(PPy) \geq Z(PP\bar{y}) + \gamma\,(y - \bar{y}) \qquad \forall\, y \in \text{Dom } Z(PPy).$$

O seguinte teorema apresenta um critério para a igualdade $Z(P) = Z(D)$ em termos da existência de um subgradiente global de $Z(PPy)$ na origem e a identificação deste subgradiente com a solução de (D).

Teorema 4.6 (GEOFFRION (1974))

Assuma (P) factível.

(a) As seguintes afirmações são equivalentes

(1) $Z(P) = Z(D)$.

(2) Existe um subgradiente global de $Z(PPy)$ em $\bar{y} = 0$.

(3) Existe um par (x^o, λ^o) satisfazendo $\lambda^o \geq 0$ e o teorema 3.1. (c):

(i) x^o é ótimo de (RP_{λ^o}),

(ii) $Ax^o \geq b$,

(iii) $\lambda^o\,(\,b - Ax^o\,) = 0$.

Prova:

(a) provaremos que (1) $\Rightarrow$ (3), ou seja,

se $Z(P) = Z(D)$ então existirá um par (x, λ) satisfazendo as condições (i),(ii) e (iii) do teorema 3.1, ítem (c).

Sejam x^o e λ^o respectivamente as soluções ótimas para $Z(P)$ e $Z(D)$.

Portanto

$$b - Ax^o \leq 0. \qquad \text{(I)}$$

Suponha por absurdo que x^o não seja solução ótima para (RP_{λ^o}); então, esta solução será dada por $\bar{x}$, logo

$$Z(D) = c\bar{x} + \lambda^o(b - A\bar{x}) < cx^o + \lambda^o(b-Ax^o).$$

Temos por hipótese que $Z(D) = Z(P)$,
substituindo na expressão acima temos

$$Z(P) < cx^o + \lambda^o(b - Ax^o),$$

que implica $\qquad cx^o < cx^o + \lambda^o(b - Ax^o).$

Portanto

$$\lambda^o(b - Ax^o) > 0.$$

Como $\lambda^o \geq 0$ temos um absurdo pois $(b - Ax^o) \leq 0$.

Portanto x^o é ótimo para (RP_{λ^o}). (II)

Sabemos por hipótese que $Z(P) = Z(D)$,
portanto

$cx^o = cx^o + \lambda^o\ (b - Ax^o)$ o que implica

$$\lambda^o\ (b - Ax^o) = 0. \qquad \text{(III)}$$

Portanto (I), (II) e (III) provam esta parte.

Provaremos que (3) $\Rightarrow$ (2)

Dado um par (x^o, λ^o) satisfazendo as condições (i), (ii) e (iii), provaremos que existe um subgradiente global de $Z(PPy)$ em $\bar{y} = 0$.

Desde que x^o é ótimo de (P) (veja teorema 3.1 deste capítulo) temos:

$Z(PPo) = cx^o.$ (I)

De (i) temos:

$$cx^o + \lambda^o(b - Ax^o) \leq cx + \lambda^o(b - Ax), \qquad \forall\ x \in F(RP_\lambda).$$

Usando a condição (iii) obtemos

$$cx^o \leq cx + \lambda^o(b - Ax), \qquad \forall\ x \in F(RP_\lambda).$$

Substituindo em (I) temos

$$Z(PPo) \leq cx + \lambda^o(b - Ax), \qquad \forall\ x \in F(RP_\lambda),$$

como $F(PPy) \subseteq F(RP_\lambda)$ logo

$$Z(PPo) \leq cx + \lambda^o(b - Ax), \qquad \forall \; x \in F(PPy).$$

Como $x \in F(PPy)$ então $b - Ax \leq y$ portanto

$$Z(PPo) \leq cx + \lambda^o y, \qquad \forall \; x \in F(PPy).$$

Logo

$$Z(PPo) - \lambda^o y \leq cx, \qquad \forall \; x \in F(PPy).$$

Tomando em particular x que otimiza (PPy), temos

$$Z(PPo) - \lambda^o y \leq Z(PPy).$$

Portanto $-\lambda^o$ é subgradiente global em $\bar{y} = 0$.

Provaremos que (2) $\Rightarrow$ (1).

Sabemos do teorema 3.1 que:

$$Z(P) \geq Z(RP_\lambda) \qquad \forall \; \lambda \geq 0. \qquad \text{(I)}$$

Por hipótese existe γ, um subgradiente global de Z(PPy) em $\bar{y} = 0$.

Logo $Z(PPy) \geq Z(PP_o) + \gamma \, y \qquad \forall \; y \in \text{Dom } Z(PPy).$ (II)

Escolhendo $y = \varepsilon \, e_i = \begin{pmatrix} 0 \\ \vdots \\ \varepsilon \\ \vdots \\ 0 \end{pmatrix}$ onde $\varepsilon > 0$ e $e_i = \begin{pmatrix} 0 \\ \vdots \\ i \\ \vdots \\ 0 \end{pmatrix}$.

Pelo teorema 4.2 deste capítulo temos

$$Z(PPy) \leq Z(PP_o). \qquad \text{(III)}$$

Por $y > 0$ e de (II) e (III) temos $\gamma \, y \leq Z(PPy) - Z(PP_o) \leq 0$

Portanto $\gamma \leq 0$.

Tomemos em particular, na equação (I), $\lambda = \lambda^o = -\gamma$ logo

$$Z(P) \geq Z(RP_{\lambda^o}) = cx^o + \lambda^o(b - Ax^o).$$

Da definição de subgradiente e da hipótese :

$Z(PPy) \geq Z(PPo) + \gamma \, y \qquad \forall \; y \in \text{Dom } Z(PPy)$ que implica

$Z(PPy) \geq Z(P) + \gamma \, y \qquad \forall \; y \in \text{Dom } Z(PPy)$ que implica

$$Z(P) \leq Z(PPy) - \gamma \, y \qquad \forall \; y \in \text{Dom } Z(PPy).$$

Tomando em particular $y = b - Ax^o$ temos pelo teorema 4.3 :

$$Z(P) \leq cx^o - \gamma(b - Ax^o) = cx^o + \lambda^o(b - Ax^o) = Z(RP_{\lambda^o}),$$

portanto

$$Z(P) \leq Z(RP_{\lambda^o}). \qquad \text{(IV)}$$

De (I) e (IV) temos que λ^o maximiza $Z(RP_\lambda)$ e

$$Z(D) = Z(RP_{\lambda^o}).$$

Como $Z(D) = \max Z(RP_\lambda) = Z(RP_{\lambda^o})$ segue o teorema. ∎

Teorema 4.7 (GEOFFRION (1974))

A função b-perturbação $Z(PP^*y)$, associada com (P^*), é precisamente o envoltório convexo inferior da função $Z(PPy)$, associada a (P).

Prova:

Para provar este teorema, basta mostrar que o epigrafo de $Z(PP^*y)$ é o casco convexo do epigrafo de $Z(PPy)$, ou seja,

$$\text{Epi } [Z(PP^*y)] = Co\{ \text{Epi } [Z(PPy)] \}.$$

Temos

$$\text{Epi } [Z(PPy)] \stackrel{d}{=} \{(\mu, y): \mu \geq Z(PPy)\}$$

$$= \{(\mu,y): \mu \geq cx \text{ e } b - Ax \leq y \text{ para algum } x \in X \},$$

onde

$$X \stackrel{d}{=} \{ x \geq 0 : Bx \geq d \text{ e } x_j \text{ inteiro para } j \in I \},$$

e para Epi $[PP^*y]$ substitui-se X por $Co\{X\}$. Suponha que $(\bar{\mu},\bar{y}) \in$ Epi $[PP^*y]$. Então $\bar{\mu} \geq c\bar{x}$ e $b - A\bar{x} \leq \bar{y}$ para algum $\bar{x} \in Co\{X\}$. Seja $\bar{x} = \sum_h \theta_h x^h$, onde $x^h \in X$, $\theta_h \geq 0$ para todo h e $\sum_h \theta_h = 1$. Definimos $\mu^h = cx^h$ e $y^h = b - Ax^h$ para todo h. Claramente $(\mu^h, y^h) \in$ Epi [PPy] para todo h. Portanto $\sum_h \theta_h(\mu^h, y^h) \in Co\{\text{Epi}[PPy]\}$. Mas

$$\sum_h \theta_h(\mu^h, y^h) = (\sum_h \theta_h\mu^h, \sum_h \theta_h y^h)$$

$$= (\sum_h \theta_h cx^h, \sum_h \theta_h(b - Ax^h)) = (c\bar{x}, b - A\bar{x}) \leq (\bar{\mu}, \bar{y})$$

e $(\bar{\mu}, \bar{y})$ também está em $Co\{Epi\ [PPy]\}$. Isto mostra que $Epi\ [PP^*y] \subseteq Co\{Epi\ [PPy]\}$. Agora suponha $(\bar{\mu}, \bar{y}) \in Co\{Epi\ [PPy]\}$. Seja $(\bar{\mu}, \bar{y}) = \sum_h \theta_h(\mu^h, y^h)$, onde $(\mu^h, y^h) \in Epi\ [PPy]$, $\theta_h \geq 0$ para todo h e $\sum_h \theta_h = 1$. Seja x^h um ponto qualquer em X satisfazendo $\mu^h \geq cx^h$ e $b - Ax^h \leq y^h$. Então

$$\sum_h \theta_h\mu^h \geq \sum_h \theta_h cx^h = c\bar{x} ,$$

$$\sum_h \theta_h y^h \geq \sum_h \theta_h(b - Ax^h) = (b - A\bar{x})$$

onde $\bar{x} = \sum_h \theta_h x^h$. Assim $(\bar{\mu}, \bar{y}) \geq (c\bar{x}, b - A\bar{x})$ com $\bar{x} \in Co\{X\}$, mostrando que $(\bar{\mu}, \bar{y}) \in Epi\ [PP^*y]$. Isto completa a prova. ■

O teorema acima mostra a relação entre os problemas (P) e (P^*), através das funções b-perturbação associadas a estes. Notamos também que o gap dual é precisamente igual à diferença entre a função b-perturbação de (P) e seu envoltório convexo inferior, ambos avaliados na origem.

Algumas destas idéias estão ilustradas na figura 4.3 para um caso hipotético de um problema de programação inteira misto, com uma única restrição $Ax \geq b$, portanto $y \in R$.

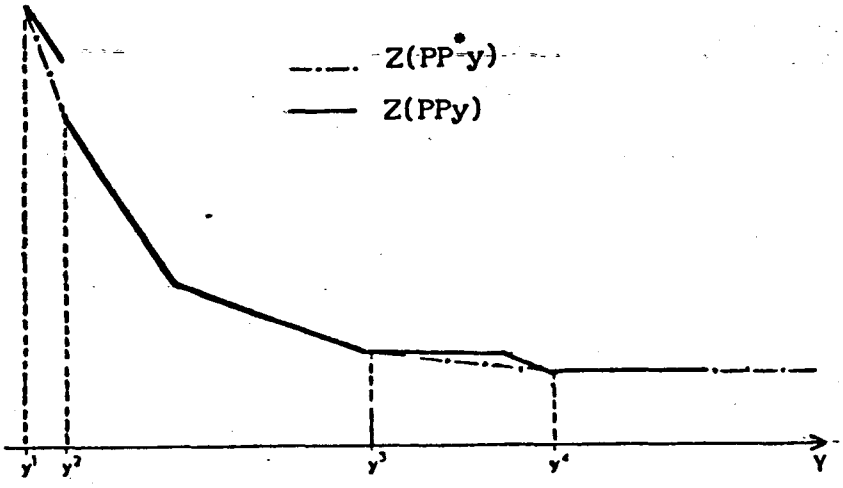

Figura 4.3 - Ilustração hipotética das funções b-perturbação

O envoltório convexo inferior $Z(PP^*y)$ de $Z(PPy)$, está representado por traços e pontos alternados. Observamos que não existe gap dual para $y^2 \leq y \leq y^3$ ou $y^4 \leq y$. Um subgradiente global de $Z(PPy)$ em $\bar{y} = 0$ existirá, se o eixo das ordenadas estiver em algum dos intervalos citados acima. Por outro lado, se a origem estiver entre y^1 e y^2, ou y^3 e y^4, existirá um gap dual e portanto não existirá um subgradiente global de $Z(PPy)$ em $\bar{y} = 0$.

CAPÍTULO III

O USO DA RELAXAÇÃO LAGRANGEANA NO ALGORITMO BRANCH-AND-BOUND

1. A Relaxação Lagrangeana no Algoritmo Branch-and-Bound

Neste capítulo discutiremos o uso da relaxação lagrangeana, como um dispositivo para possível melhoramento da eficiência do algoritmo Branch-and-Bound na resolução de problemas de programação linear inteira.

Vimos no capítulo I, que o método consiste de uma busca em árvore onde, cada nó representa um problema (PC) similar ao problema original (P), porém mais restrito. Tais nós são usualmente avaliados por $Z(\overline{PC})$ que fornece um limitante inferior (caso de minimização) para Z(PC). Este limitante é utilizado para indicar o quão promissor é o nó, podendo informar que não há qualquer solução promissora em F(PC) quando $Z(\overline{PC}) \geq Z^*$, onde Z^* é o melhor valor da função objetivo de (P) para as soluções factíveis conhecidas de (P), no instante em que (PC) está sendo avaliado.

Pelo que foi discutido no capítulo I, para que tal método não se torne bastante exaustivo, é necessário acelerar o processo de sondagem, e isto pode ser obtido com melhores limitantes para (PC). Pelo ítem (b) do teorema 3.1 no capítulo II, observamos que a relaxação lagrangeana (RPC_λ) de (PC) pode oferecer limitantes melhores que os dado por $(\overline{PC})$. Pelo mesmo teorema podemos observar que:

i) infactibilidade de $(RPC_\lambda) \Rightarrow$ infactibilidade de (PC);

ii) $Z(RPC_\lambda) \geq Z^* \Rightarrow Z(PC) \geq Z^*$;

temos então, dois casos de sondagem usando $Z(RP_\lambda)$. Por outro lado seja x^R uma solução de (RPC_λ), temos que:

iii) se $x^R \in F(PC)$ / $Z^* > cx^R \Rightarrow$ atualização de Z^*;

iv) se x^R for ótimo em (PC) $\Rightarrow$ sondagem de (PC).

Em (iii) temos que, se x^R for factível em (PC), mas não satisfizer às folgas complementares (ítem (c) do teorema 3.1 do capítulo II), ela poderá melhorar Z^*, isto é, seria usada para atualizar Z^* ainda que (PC) não fosse sondado. Em (iv) além do caso de sondagem do (PC) podemos ter o de atualização de Z^*. Lembramos que x^R será solução ótima para (PC) se satisfizer às folgas complementares, e $x^R \in F(PC)$.

Se x^R não for factível em (PC), métodos heurísticos podem ser desenvolvidos, de tal forma que explorem estruturas específicas do problema (PC) de modo a ganhar a factibilidade na esperança de melhorar o valor de Z^*. Um método assim é dado por FISHER(1981, 1985), o qual é chamado "heurística lagrangeana" e está ilustrado na seção 3 do capítulo V.

No caso de não haver sondagem, o problema será separado em dois ou mais subproblemas, (PC_1), ..., (PC_L), de tal forma que $F(PC) = F(PC_1) \cup \ldots \cup F(PC_L)$ e de preferência que $F(PC_i) \cap F(PC_j) = \emptyset$, $i \neq j$. Usualmente as restrições de separação são simplesmente limites sobre variáveis de índice j tal que $j \in I$.

Veremos agora, como a relaxação lagrangeana pode ser usada como complemento para a relaxação linear.

Suponha que uma variável de índice $j \in I$ tenha um valor fracional na solução $\bar{x}$ de $(\overline{PC})$. Estamos interessados em limites inferiores sobre $Z(PC / x_j \leq [\bar{x}_j])$ e $Z(PC / x_j \geq [\bar{x}_j] + 1)$, onde "/" significa que as restrições seguintes são acrescentadas ao problema e $[\bar{x}_j]$ é a parte inteira de $\bar{x}_j$. Tais limites podem ser dados por:

$$Z_s(j) \stackrel{d}{=} Z(\ RPC_\lambda\ /\ x_j \le [\bar{x}_j]\ ),$$
$$Z_i(j) \stackrel{d}{=} Z(\ RPC_\lambda\ /\ x_j \ge [\bar{x}_j] + 1\ ) \qquad (1)$$

ou $Z_s(j)$ é o valor da função objetivo da relaxação lagrangeana do problema (PC), com a restrição adicional $x_j \le [\bar{x}_j]$.

$Z_i(j)$ é o valor da função objetivo da relaxação lagrangeana do problema (PC), com a restrição adicional $x_j \ge [\bar{x}_j] + 1$.

Se $Z_s(j) \ge Z^*$ e $Z_i(j) \ge Z^*$ então o problema (PC) é sondado. Os limites em (1), podem também ser usados para guiar a separação, quando (PC) não for sondado. Sejam $Z_s(j)$ e $Z_i(j)$ calculados para todo $j \in I$ tal que $\bar{x}_j$ seja fracional. A escolha da variável a separar poderá ser feita por aquela de índice j tal que dê o maior valor para o maior entre $Z_s(j)$ e $Z_i(j)$ para todo $j \in I$ tal que $\bar{x}_j$ seja fracional. Uma vez selecionada a variável de separação j_o , $Z_s(j_o)$ e $Z_i(j_o)$ produzem limites inferiores para futuras referências sobre os novos subproblemas candidatos criados. Segundo GEOFFRION(1974), vários problemas de programação inteira foram resolvidos com sucesso usando critérios análogos, ou seja, usando $(\overline{PC})$ antes de (RPC_λ).

Assim a relaxação lagrangeana pode ser usada no algoritmo Branch-and-Bound padrão com a tarefa de sondar, gerar melhores soluções factíveis e guiar as separações.

No uso de (RPC_λ) como complemento para $(\overline{PC})$, a questão preocupante é se (RPC_λ) deve ser usado antes, depois ou no lugar de $(\overline{PC})$. Esta questão não pode ser respondida de forma geral, mas uma importante observação a fazer é se o problema relaxado possui ou não a propriedade de integralidade, definida no capítulo II, para uma partição particular das restrições.

Suponha que o problema possua a propriedade de integralidade. Então (RPC_λ) será infactível somente quando $(\overline{PC})$ for infactível e se $(\overline{PC})$ for factível, pelo teorema 3.2, do capítulo II, o melhor valor de λ para (RPC_λ) será $\bar{\lambda}$, e

$Z(RPC_{\bar{\lambda}})=Z(PC)$. Assim (RPC_{λ}) não sondaria (PC) por infactibilidade ou por limitantes sem que $(\overline{PC})$ o fizesse.

Estes fatos discutem a oposição ao uso da relaxação lagrangeana quando (RPC_{λ}) possui a propriedade de integralidade, pois neste caso esta relaxação tem pouca vantagem sobre $(\overline{PC})$, mas pode ser muito promissor usá-la como recurso para melhorar soluções factíveis (GEOFFRION(1974)). Estas conclusões negativas apoiam-se no fato de que $(\overline{PC})$ seja manejável com técnicas de programação linear, se este não for o caso, então (RPC_{λ}) será uma atrativa alternativa computacional.

Agora suponha que a relaxação lagrangeana de (PC) não possua a propriedade de integralidade. Então o problema $(\overline{PC})$ não necessariamente possui a melhor escolha de λ e (RPC_{λ}) pode com sucesso sondar onde a relaxação $(\overline{PC})$ falhar. Diante disto, faz sentido usar (RPC_{λ}) antes, depois ou até mesmo no lugar de $(\overline{PC})$, dependendo do gasto computacional destas duas relaxações. A estratégia mais efetiva depende do desempenho de $(\overline{PC})$ na geração de λ a ser usado em (RPC_{λ}). $(\overline{PC})$ pode ser usado para gerar o valor de λ inicial e então melhorado por um método independente. Um possível método para determinar um valor apropriado para λ será visto no capítulo seguinte.

2. Algoritmo Genérico

Baseado no trabalho de SHAPIRO(1979), apresentamos nesta seção um diagrama que resume os passos para a determinação de uma solução ótima de (P) através do método Branch-and-Bound com auxílio da relaxação Lagrangeana.

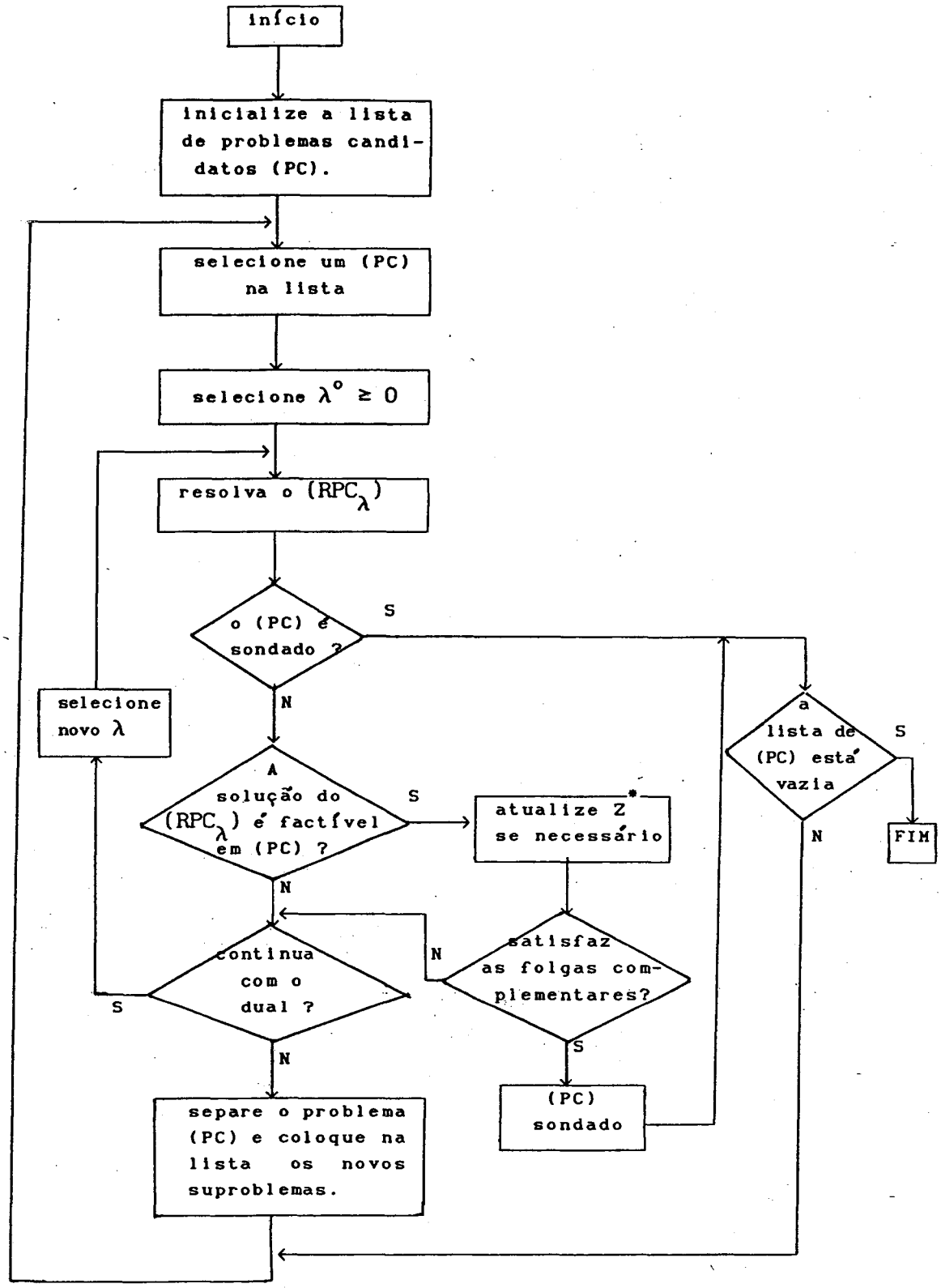

Figura 2.1 - Diagrama de blocos mostrando como a relaxação lagrangeana pode auxiliar o algoritmo Branch-and-Bound.

3. Considerações Finais

A figura 2.1 mostra um algoritmo genérico para o uso da relaxação lagrangeana no método Branch-and-Bound. Pelo que podemos observar, em publicações de experiências computacionais, este procedimento geralmente é usado para completar a solução de (P) dada por (RP_λ), ou seja, a solução ótima $\tilde{x} \in F(P)$ de ($RP_{\tilde{\lambda}}$) não é ótima em (P), o que implica na existência de um gap dual (onde $\tilde{\lambda} \geq 0$ é a melhor escolha para o valor de λ, a qual pode ser calculada partindo de $\bar{\lambda}$, solução ótima dual associada à $Ax \geq b$ em ($\bar{P}$), pelo algoritmo subgradiente que veremos no capítulo seguinte), então inicia-se o Branch-and-Bound tomando $Z^* = c\tilde{x}$. Inicialmente a lista de problemas candidatos contêm apenas o problema (P), o qual será particionado em outros subproblemas que serão adicionados a lista. Selecionado o novo problema candidato a ser avaliado, um bom valor de λ^o inicial é usualmente dado pelo valor de λ usado no cálculo do limitante do nó anterior (SHAPIRO(1979), FISHER(1981)).

Uma comparação do tempo computacional, para resolução de problemas usando o método Branch-and-Bound com limites dados pela relaxação lagrangeana ou apenas pela relaxação linear, é dada em FISHER(1981) onde relata que para problemas pequenos o tempo computacional usando limitantes dados pela relaxação linear é em média um quarto do tempo que se faria usando a relaxação lagrangeana, mas no caso de problemas grandes, este tempo é em média duas vezes e meia maior que se faria usando a relaxação lagrangeana.

OTIMIZAÇÃO DO PROBLEMA DUAL (D)

1. Propriedades e Definições Básicas

Neste capítulo apresentaremos um método para otimização do problema dual (D), o tradicional método subgradiente o qual tem sido muito usado e demonstra ter bastante eficiência computacional. Para seu desenvolvimento precisaremos dos resultados e definições a seguir.

O problema dual definido no capítulo anterior será agora objeto de atenção:

$$(D) \qquad Z(D) = \max_{\lambda \geq 0} Z(RP_\lambda)$$

onde

$$Z(RP_\lambda) = \min_{x \in X} [cx + \lambda(b - Ax)] \quad , \quad X = \{x \ / \ Bx \geq d \ , \ x \geq 0 \text{ e } x_j \text{ inteiro } j \in I \},$$

Propriedade 1.1

A função $Z(RP_\lambda)$ é côncava em λ e linear por partes.

Prova:

Sejam $\lambda^1 \geq 0$ e $\lambda^2 \geq 0$ arbitrários e para $\theta \in [0,1]$ seja $\lambda = \theta\,\lambda^1 + (1 - \theta)\,\lambda^2$.

Pela definição de $Z(RP_\lambda)$, existe $\bar{x} \in F(RP_\lambda)$ tal que

$$Z(RP_\lambda) = c\bar{x} + \lambda\,(b - A\bar{x}),$$

por definição de $Z(RP_\lambda 1)$ e $Z(RP_\lambda 2)$ temos

$$Z(RP_\lambda 1) \leq c\bar{x} + \lambda^1\,(b - A\bar{x}),$$

$$Z(RP_\lambda 2) \leq c\bar{x} + \lambda^2\,(b - A\bar{x}).$$

Multiplicando a primeira desigualdade por $\theta \geq 0$, e a segunda por $(1 - \theta) \geq 0$. Somando-as obtemos:

$$\theta\,Z(RP_\lambda 1) + (1 - \theta)\,Z(RP_\lambda 2) \leq c\bar{x} + \lambda\,(b - A\bar{x}) = Z(RP_\lambda).$$

O que prova a concavidade de $Z(RP_\lambda)$.

Para mostrarmos a linearidade por partes, é suficiente observarmos que $Z(RP_\lambda)$ é o envoltório inferior de uma família de hiperplanos com equações

$$\begin{aligned} Z &= cx^1 + \lambda\,(b - Ax^1); \\ Z &= cx^2 + \lambda\,(b - Ax^2); \\ &\vdots \\ Z &= cx^K + \lambda\,(b - Ax^K). \end{aligned}$$

onde $x^1, x^2, \ldots, x^k$ são os pontos que otimizam (RP_λ) a cada valor de λ (ver figura 1.1).

■

A figura 1.1, a seguir, esquematiza a função $Z(RP_\lambda)$ para o caso em que a matriz A possui apenas uma linha ($\lambda \in R$) :

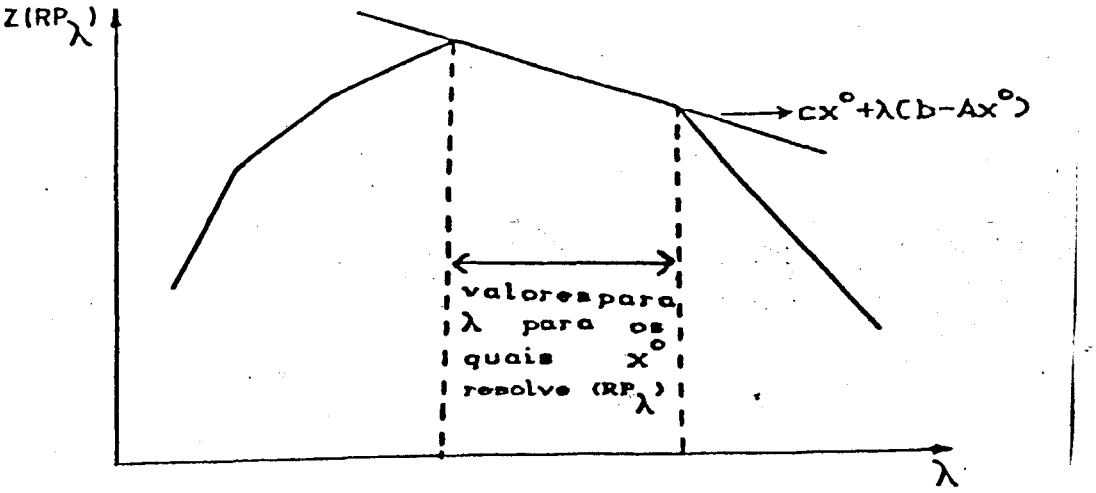

Figura 1.1 - Função $Z(RP_\lambda)$.

Definição 1.1

Um vetor $g(\lambda^o)$, tal que $g(\lambda^o) \in R^m$ é um subgradiente da função côncava $Z(RP_\lambda)$ no ponto λ^o se

$$Z(RP_\lambda) \leq Z(RP_{\lambda^o}) + (\lambda - \lambda^o)\, g(\lambda^o) \quad \text{para todo } \lambda \geq 0.$$

Observamos que, se $Z(RP_\lambda)$ for diferenciável em um ponto λ^o então $g(\lambda^o) = \nabla Z(RP_{\lambda^o})$, onde $\nabla Z(RP_{\lambda^o})$ é o gradiente de $Z(RP_\lambda)$ no ponto λ^o.

Definição 1.2

O subdiferencial de $Z(RP_\lambda)$ no ponto λ^o, denotado por $\partial Z(RP_{\lambda^o})$, é o conjunto de todos os subgradientes de $Z(RP_\lambda)$ em λ^o.

A seguinte propriedade nos fornece uma maneira de determinar um subgradiente de $Z(RP_\lambda)$ no ponto λ^o.

Seja $Z(RP_{\lambda^o}) = cx^o + \lambda^o(b - Ax^o)$, isto é, x^o é uma solução

ótima do problema lagrangeano:

$$Z(RP_{\lambda^o}) = \min cx + \lambda^o(b - Ax)$$

sujeito a

$$Bx \geq d$$
$$x \geq 0,$$
$$x_j \text{ inteiro}, \ j \in I.$$

Propriedade 1.2

O vetor $(b - Ax^o)$ é um subgradiente de $Z(RP_\lambda)$ no ponto $\lambda = \lambda^o$.

Prova:

$$Z(RP_{\lambda^o}) + (\lambda - \lambda^o)(b - Ax^o) = cx^o + \lambda^o(b-Ax^o) + (\lambda - \lambda^o)(b-Ax^o)$$
$$= cx^o + \lambda(b - Ax^o) \geq Z(RP_\lambda) \quad \text{para todo } \lambda \geq 0.$$

Portanto $(b - Ax^o)$ é um subgradiente de $Z(RP_\lambda)$ em λ^o. ∎

Propriedade 1.3

$\partial Z(RP_\lambda)$ é convexo e fechado.

Prova:

Se $\partial Z(RP_\lambda) = \emptyset$ não há nada a mostrar. Suponha $\partial Z(RP_\lambda) \neq \emptyset$.

Para provar a convexidade, sejam g_1 e $g_2 \in \partial Z(RP_{\bar{\lambda}})$;

então $Z(RP_{\bar{\lambda}}) \geq Z(RP_\lambda) + (\bar{\lambda} - \lambda)g_1$ para todo $\bar{\lambda} \geq 0$, (I)

$Z(RP_{\bar{\lambda}}) \geq Z(RP_\lambda) + (\bar{\lambda} - \lambda)g_2$ para todo $\bar{\lambda} \geq 0$, (II)

multiplicando (I) por α e (II) por $(1 - \alpha)$, $0 < \alpha < 1$, somando

temos

$$Z(RP_{\bar{\lambda}}) \geq Z(RP_{\lambda}) + (\bar{\lambda} - \lambda)(\alpha g_1 + (1 - \alpha)g_2) \quad \text{para todo } \bar{\lambda} \geq 0.$$

Portanto

$$(\alpha g_1 + (1 - \alpha)g_2) \in \partial Z(RP_{\lambda}).$$

Para mostrar que $\partial Z(RP_{\lambda})$ é fechado, seja $\{g_i^*\}$ uma sequência convergente de elementos de $\partial Z(RP_{\lambda})$ com limite g^* e assuma $g^* \notin \partial Z(RP_{\lambda})$. Então existe um $\hat{\lambda} \geq 0$ e um $\varepsilon > 0$ tal que

$$Z(RP_{\hat{\lambda}}) + \varepsilon = Z(RP_{\lambda}) + (\hat{\lambda} - \lambda)g^*. \qquad \text{(III)}$$

Para qualquer elemento g_n^* da sequência $\{g_i^*\}$

$$Z(RP_{\hat{\lambda}}) \geq Z(RP_{\lambda}) + (\hat{\lambda} - \lambda)g_n^* \quad , \qquad \text{(IV)}$$

subtraindo (III) de (IV) temos

$$\varepsilon \leq (\hat{\lambda} - \lambda)(g^* - g_n^*) \leq || g^* - g_n^* || \; || \hat{\lambda} - \lambda || \; ,$$

dando a desigualdade de Schwartz. Assim

$$|| g^* - g_n^* || \geq \frac{\varepsilon}{|| \hat{\lambda} - \lambda ||} > 0 \quad \text{para todo } n \; ,$$

o que contradiz a convergência de $\{g_i^*\}$ para g^*. Assim, $g^* \in \partial Z(RP_{\lambda})$ e $\partial Z(RP_{\lambda})$ é fechado. ∎

2. Método Subgradiente

Aparentemente o método subgradiente foi originado por SHOR(1968) e talvez tenha sido melhor explorado por POLYAK(1969). Mas a idéia central baseia-se nos trabalhos de AGMON(1954) e MOTZKIN e SCHOENBERG(1954) - métodos de relaxação iterativos para resolver sistemas de desigualdades lineares - que são denominados processos de Féjer para os quais a distância ao domínio desejado decresce monotonicamente a cada passo. Uma

primeira extensão foi proposta por EREMIN(1962,1965) que explorou os processos de aproximação de Féjer com mais detalhes e elaborou um algoritmo para a solução de sistemas de inequações. SHOR(1968) fez uma generalização para torná-los aplicáveis ao problema de minimização de uma função convexa, obtendo os também chamados métodos de SHOR (ou KIEV).

Estes métodos consistem basicamente no deslocamento, a cada iteração k, na direção de um subgradiente da função considerada, ou seja

$$\lambda^{k+1} = \lambda^{k} + t_k \frac{g^t(\lambda^k)}{\| g(\lambda^k) \|}, \qquad (2.1)$$

onde $g(\lambda^k)$ é um subgradiente da função considerada.

Dentro desta classe de métodos temos algoritmos importantes propostos por POLYAK(1969) e por HELD, WOLFE e CROWDER(1974).

A dificuldade dos métodos que utilizam (2.1) decorre do fato de $g(\lambda^k)$ não ser geralmente uma direção de subida, estando então a convergência de tais algoritmos estritamente ligada à determinação do passo t_k na direção $g(\lambda^k)$.

Existem alguns resultados teóricos que relatam uma apropriada escolha para o tamanho do passo t_k, mas o resultado mais geral é dado como veremos a seguir, pela regra das séries divergentes. Para tal, consideraremos dados λ^o e a sequência $\{t_k\}$ de escalares positivos. Definiremos a sequência $\{\lambda^k\}$ por:

$$\lambda^{k+1} = \mathrm{proj}_S \left(\lambda^k + t_k \frac{g^t(\lambda^k)}{\| g(\lambda^k) \|}\right) \quad \text{para } k = 0,1,2.... .$$

Onde $S = \{ \lambda \ / \ \lambda \geq 0 \}$ e proj_S é o operador projeção de R^m sobre S e $g(\lambda^k) \in \partial Z(RP_\lambda)$.

Teorema 2.1 (ERMOLIEV(1966), POLYAK(1967))

Seja f uma função côncava e contínua de R^m em R, S um conjunto convexo (assumimos que ou $f(\lambda) \longrightarrow -\infty$ quando $\|\lambda\| \longrightarrow \infty$

ou S é limitado), se t_k dado em

$$\lambda^{k+1} = \text{proj}_S \left(\lambda^k + t_k \frac{g^t(\lambda^k)}{||g(\lambda^k)||}\right) \quad \text{para } k = 0,1,2.... \quad ,$$

é tal que satisfaz (Regra das séries divergentes)

i) $t_k \longrightarrow 0$ quando $k \longrightarrow +\infty$;

ii) $\sum_{k=0}^{\infty} t_k = +\infty$;

então esta sequência $\{\lambda^k\}$, $k \in N$ com $g(\lambda^k) \in \partial f$, é tal que

$$\limsup \quad f(\lambda^k) = f^* \stackrel{d}{=} \max_{\lambda \in S} \{ f(\lambda) \},$$

onde $f^* = f(\lambda^*)$.

Prova:

É suficiente mostrar que $\forall\, \varepsilon > 0$, $\exists\, N$ tal que $\forall\, i \geq N$

$$\lambda^i \in V(\varepsilon) = \{\lambda \,/\, f(\lambda) \geq f^* - \varepsilon \;;\; \lambda \in S \}.$$

Supomos por absurdo que $\forall\, k$, $\lambda^k \notin V(\varepsilon)$.

Em outras palavras $\forall\, k$: $f(\lambda^k) < f^* - \varepsilon$. (I)

Como f é contínua em R^m então existe $\alpha > 0$ tal que

$$|| \lambda - \lambda^* || \leq \alpha \Rightarrow f(\lambda) \geq f^* - \varepsilon. \qquad \text{(II)}$$

Podemos assumir que $\forall\, k$, $g(\lambda^k) \neq 0$ (caso contrário, λ^k seria solução ótima e o algoritmo terminaria)

Considere $\bar{\lambda}^k$ definido por

$$\bar{\lambda}^k = \lambda^* - \alpha \frac{g^t(\lambda^k)}{||g(\lambda^k)||} \quad , \qquad \text{(III)}$$

temos

$|| \bar{\lambda}^k - \lambda^* || = \alpha$, portanto, de (II) temos,

$$f(\bar{\lambda}^k) \geq f^* - \varepsilon \quad . \qquad \text{(IV)}$$

Pela definição de subgradiente e por (I) e (IV) temos, $\forall$ k :

$$(\lambda^k - \bar{\lambda}^k)g(\lambda^k) \leq f(\lambda^k) - f(\bar{\lambda}^k) < 0. \qquad \text{(V)}$$

Sabendo-se que o operador projeção é uma contração (MINOUX(1983), página 449) podemos escrever

$$||\lambda^{k+1} - \lambda^*|| = \left|\left|\text{proj}_S \left[\lambda^k + t_k \frac{g^t(\lambda^k)}{||g(\lambda^k)||}\right] - \text{proj}_S\,[\lambda^*]\right|\right|$$

$$\leq \left|\left| \lambda^k + t_k \frac{g^t(\lambda^k)}{||g(\lambda^k)||} - \lambda^*\right|\right| .$$

Disto segue

$$||\lambda^{k+1} - \lambda^*||^2 \leq ||\lambda^k - \lambda^*||^2 + t_k^2 + 2t_k \frac{(\lambda^k-\lambda^*)g(\lambda^k)}{||g(\lambda^k)||} . \qquad \text{(VI)}$$

Isolando λ^* em (III) e substituindo em $(\lambda^k - \lambda^*)\, g(\lambda^k)$ temos

$$(\lambda^k - \lambda^*)\, g(\lambda^k) = \left[\lambda^k - \bar{\lambda}^k - \alpha \frac{g^t(\lambda^k)}{||g(\lambda^k)||}\right] g(\lambda^k) =$$

$(\lambda^k - \bar{\lambda}^k)\, g(\lambda^k) - \alpha\, ||g(\lambda^k)||$ substituindo em (VI) obtemos

$$||\lambda^{k+1} - \lambda^*||^2 \leq ||\lambda^k - \lambda^*||^2 + t_k^2 + 2t_k \frac{(\lambda^k - \bar{\lambda}^k)\, g(\lambda^k)}{||g(\lambda^k)||} - 2\,\alpha\, t_k,$$

deste resultado e (V) temos

$$||\lambda^{k+1} - \lambda^*||^2 < ||\lambda^k - \lambda^*||^2 + t_k(\, t_k - 2\alpha\,),$$

como $t_k \longrightarrow 0$, existe K tal que $k \geq K \Rightarrow t_k \leq \alpha$.

Temos então $\forall\, k \geq K$:

$$||\lambda^{k+1} - \lambda^*||^2 < ||\lambda^k - \lambda^*||^2 - \alpha\, t_k ,$$

por adição sobre k , k = K, K+1, ..., K + P ;

$$\alpha \sum_{k=K}^{K+P} t_k < ||\, \lambda^K - \lambda^*||^2 - ||\, \lambda^{K+P+1} - \lambda^*||^2 < ||\lambda^K - \lambda^*||^2 .$$

Então, escolhendo P suficientemente grande, obtemos uma contradição com o fato que $\sum_{k=0}^{\infty} t_k = \infty$ e o teorema está provado.

■

Este resultado nada diz a respeito da velocidade de convergência, mas AGMON(1954) mostrou convergência linear deste procedimento quando usado para resolver inequações.

Várias condições sobre a sequência $\{t_k\}$ foram apresentadas usando uma variedade de hipóteses técnicas. POLYAK(1969) mostrou que a sequência de valores $Z(RP_{\lambda^k})$ tem convergência geométrica quando t_k for definido da seguinte forma

$$t_k = \sigma_k \frac{Z(D) - Z(RP_{\lambda^k})}{\| g(\lambda^k) \|^2}, \text{ onde } 0 < \sigma_k \leq 2, \ \forall k.$$

A escolha do coeficiente de relaxação σ_k desta forma, $0 < \sigma_k \leq 2$, pode ser justificada pelo que segue:

Sejam $\Omega = \{\lambda^t \in R^m / 0 \leq \lambda \ ; \ Z(RP_\lambda) = Z(D)\}$, $S(\lambda)$ a projeção de λ sobre Ω e $d(\lambda) = \|S(\lambda) - \lambda\|$ a distância de λ ao conjunto dos pontos ótimos.

Suponha $\lambda^{k+1} = \lambda^k + t_k g(\lambda^k)$.

Podemos escrever

$$d(\lambda^{k+1}) \stackrel{d}{=} \|\lambda^{k+1} - S(\lambda^{k+1})\| \leq \|\lambda^{k+1} - S(\lambda^k)\| ,$$

e

$$\|\lambda^{k+1} - S(\lambda^k)\|^2 = \|\lambda^k - S(\lambda^k) + t_k \ g(\lambda^k)\|^2 =$$

$$= \|\lambda^k - S(\lambda^k)\|^2 + 2t_k(\lambda^k - S(\lambda^k))g(\lambda^k) + t_k^2 \ \|g(\lambda^k)\|^2 ,$$

portanto

$$d^2(\lambda^{k+1}) \leq d^2(\lambda^k) + t_k \ [t_k \|g(\lambda^k)\|^2 + 2 \ [\lambda^k - S(\lambda^k)] \ g(\lambda^k)].$$

Como a função $Z(RP_\lambda)$ é côncava temos

$$[S(\lambda^k) \ - \lambda^k] g(\lambda^k) \geq Z(D) - Z(RP_{\lambda^k}) ,$$

logo

$$d^2(\lambda^{k+1}) \leq d^2(\lambda^k) + t_k \ [t_k \|g(\lambda^k)\|^2 - 2 \ [Z(D) - Z(RP_{\lambda^k})] .$$

Entretanto, se t_k for escolhido entre 0 e $2 \dfrac{(Z(D) - Z(RP_{\lambda^k})}{\|g(\lambda^k)\|^2}$,

então $d(\lambda^{k+1}) < d(\lambda^k)$, ou seja, a distância entre λ^k e o conjunto Ω de soluções ótimas decresce estritamente para todo k, o que justifica esta escolha para t_k.

Na prática o desconhecimento de Z(D) leva a trabalhar com estimativas $\hat{Z}$ de Z(D), portanto é comum o passo ser escolhido pela fórmula

$$t_k = \sigma_k \frac{\hat{Z} - Z(RP_{\lambda^k})}{\| g(\lambda^k) \|^2}, \qquad (2.2)$$

onde $0 < \sigma \leq 2$ e $\hat{Z}$ é uma estimativa para Z(D).

POLYAK(1969) mostrou que usando uma estimativa $\hat{Z}_- < Z(D)$, ou a sequência de $Z(RP_{\lambda^k})$ converge para $\hat{Z}_-$ ou obtem-se, em um número finito de passos, um ponto λ^j para o qual $\hat{Z}_- \leq Z(RP_{\lambda^j}) \leq Z(D)$, usando $\sigma = 2$, $\forall$ k. Uma estimativa $\hat{Z}_- < Z(D)$ pode ser calculada se uma solução factível para o problema dual for acessível.

HELD, WOLFE e CROWDER(1974) usaram em suas experiências práticas a estimativa $\hat{Z}^- \geq Z(D)$ sem afetar a convergência do algoritmo. Neste caso a condição $t_k \longrightarrow 0$ ($k \longrightarrow \infty$) exige uma escolha de $\sigma_k \longrightarrow 0$ ($k \longrightarrow \infty$). Sugeriram também uma estratégia empírica para a determinação de σ_k, que é a seguinte: $\sigma_k = 2$ durante 2n iterações, após isto faz-se $\sigma_k \leftarrow \frac{\sigma_k}{2}$, após n iterações, depois repita para $\frac{n}{2}$ iterações, $\frac{n}{4}$, $\frac{n}{8}$, ... até um número, r, pré-fixado de iterações, resultando t_k pequeno. Usando tais condições, eles obtiveram bons resultados no tratamento dos problemas de alocação, caixeiro viajante e multi-fluxos.

Finalmente apresentamos, em uma forma resumida, o algoritmo do método:

ALGORITMO SUBGRADIENTE

Inicialização -

Escolha $\lambda^o \geq 0$ e $g(\lambda^o) \in \partial Z(RP_\lambda)$;
$k = 0$.

Passo principal -

Enquanto $g(\lambda^k) \neq 0$ e $k \leq$ número máximo de iterações, faça:

início

$$\lambda^{k+1} = proj_S \left(\lambda^k + t_k \frac{g^t(\lambda^k)}{||g(\lambda^k)||}\right) ;$$

escolha $g(\lambda^{k+1}) \in \partial Z(RP_{\lambda^{k+1}})$;
$k \longleftarrow k + 1$;

fim.

λ^k é a solução encontrada pelo método subgradiente.

O cálculo de $g(\lambda^k)$ pode ser feito pela propriedade (1.2) e o tamanho do passo pode ser calculado com auxílio da fórmula (2.2), usando estimativas inferiores ou superiores.

Os critérios de parada são bastante variados, dado que não temos um critério para terminar o processo quando na iteração corrente tem-se um ponto suficientemente próximo da solução desejada. No entanto, tem-se usado terminar o algoritmo por um número limite de iterações (FISHER(1985), MINOUX(1989)).

APLICAÇÕES

Neste capítulo apresentamos alguns problemas clássicos de programação inteira, sem a pretenção de esgotar o assunto, para os quais a relaxação lagrangeana foi usada com sucesso e os resultados publicados.

1. O Problema do Caixeiro Viajante

Nesta seção mostraremos o uso da relaxação lagrangeana na resolução do problema do caixeiro viajante, baseado nos trabalhos de BAZARAA e GOODE (1977) e HELD e KARP (1970, 1971).

O problema do caixeiro viajante consiste em determinar uma rota de custo mínimo, na qual deve-se visitar N cidades, passando apenas uma vez por cada uma dessas e então retornar ao ponto de partida. Tal problema pode ser tratado como veremos a seguir.

Seja $G = [X,U]$ um grafo direcionado, onde X denota o conjunto de vértices e $|X| = N$: número de vértices no grafo; U denota o conjunto de arcos e $|U| = M$: número de arcos no grafo. Cada arco $u \in U$ é um par ordenado de vértices, $u = (i,j)$, no qual i é chamado vértice inicial e j é chamado vértice final.

Um **caminho** é uma sequência de arcos $\{ u_1, u_2, \ldots , u_p \}$, tais que para todo $k = 1, 2, \ldots , p-1$, o vértice terminal de u_k

coincide com o vértice inicial de u_{k+1}.

Um **circuito** é um caminho $\{u_1, u_2, \ldots, u_p\}$, no qual o vértice inicial do arco u_1, coincide com o terminal de u_p.

Um **circuito elementar** é um circuito tal que existe no máximo um arco divergindo e um convergindo em cada vértice.

Um **circuito Hamiltoniano** é um circuito elementar, que contém todos os vértices do grafo.

A cada arco $u = (i,j) \in U$ de G, associaremos o número l_{ij}, o qual representa o custo de passar do vértice i para j.

O **problema do caixeiro viajante** equivale então a determinar um circuito Hamiltoniano de custo total (soma dos custos de todos os arcos no circuito) mínimo no grafo G.

A versão não direcionada do caixeiro viajante, ou caixeiro viajante simétrico, corresponde ao caso especial onde a existência do arco (i,j) entre os vértices i e j implica na existência do arco (j,i) e $l_{ij} = l_{ji}$. Neste caso a orientação dos arcos pode ser esquecida e o problema do caixeiro viajante busca um circuito Hamiltoniano de custo total mínimo em um grafo não orientado.

Para a formulação do problema, associaremos a cada arco $u = (i,j)$ de G uma variável inteira x_{ij} que assume 0 ou 1, como mostraremos a seguir

$$x_{ij} = \begin{cases} 1 & \text{se o arco } (i,j) \text{ está contido no circuito Hamiltoniano;} \\ 0 & \text{caso contrário.} \end{cases}$$

Seja $\bar{U}(x) = \{ u \,/\, u \in U,\ x_u = 1 \}$

Para um circuito Hamiltoniano temos as seguintes condições a serem satisfeitas

$$\begin{cases} \forall\, i \in X : \sum_{\substack{j \\ (i,j) \in U}} x_{ij} = 1 & C_1 \\ \\ \forall\, j \in X : \sum_{\substack{i \\ (i,j) \in U}} x_{ij} = 1 & C_2 \\ \\ \text{o grafo parcial } \bar{G}(x) = [\ X,\ \bar{U}(x)\ ] \text{ é conexo} & C_3 \end{cases}$$

As condições C_1 e C_2, são as condições necessárias para o circuito ser elementar.

Podemos escrever a condição C_3 como

(C'3) O grafo parcial $\bar{G}(x) = [X, \bar{U}(x)]$ é conexo e contém N arcos, 2 dos quais são incidentes no vértice 1.

O grafo parcial satisfazendo (C'3) é chamado de 1-árvore. Uma 1-árvore é um grafo consistindo de uma árvore sobre os vértices 2, 3,...,N , ligada ao vértice 1 por dois arcos distintos (ver figura 1.1). Assim uma 1-árvore contém um único ciclo, este ciclo contém o vértice 1 e tal vértice sempre tem grau dois.

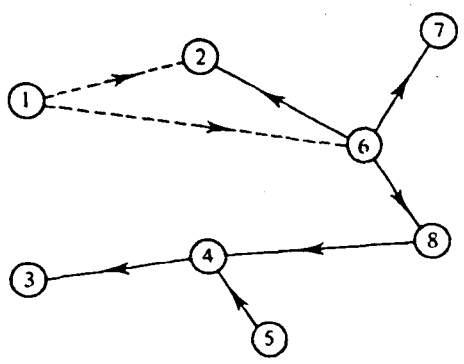

Figura 1.1 - 1-árvore consistindo de uma árvore sobre os vértices 2,3,...,8 e ligada por dois arcos distintos ao vértice 1.

Seja $\varphi_1 = \{ (x_{11}, \ldots, x_{1N}, \ldots, x_{NN}) : x_{ij} = 0$ ou $x_{ij} = 1$ $(i \neq j)$ e os arcos (i,j) com $x_{ij} = 1$ formam uma 1-árvore $\}$, com esta notação o problema do caixeiro viajante pode ser formulado como segue.

$$
\begin{cases}
\min \sum\limits_{(i,j) \in U} l_{ij} x_{ij} \\
\quad \text{sujeito a} \\
\sum\limits_{\substack{j \\ (i,j) \in U}} x_{ij} = 1 \qquad (\forall\, i \in X) \qquad (1) \\
\sum\limits_{\substack{i \\ (i,j) \in U}} x_{ij} = 1 \qquad (\forall\, j \in X) \qquad (2) \\
x \in \varphi_1 \qquad (3)
\end{cases}
$$

Note que calcular uma 1-árvore de custo mínimo na presença das restrições (1) e (2), é um problema muito difícil. No entanto podemos dualizar tais restrições usando, como já vimos, a relaxação lagrangeana.

Para definir a relaxação lagrangeana deste problema tomemos λ_i como a variável dual associada às restrições (1) e μ_j às restrições (2). Então a função lagrangeana torna-se

$$L(x,\lambda,\mu) = \sum_{(i,j)} l_{ij}\, x_{ij} + \sum_i \lambda_i \left[\sum_j x_{ij} - 1 \right] + \sum_j \mu_j \left[\sum_i x_{ij} - 1 \right]$$

$$= - \sum_i \lambda_i - \sum_j \mu_j + \sum_{(i,j)} (l_{ij} + \lambda_i + \mu_j)\, x_{ij}$$

Para λ e μ fixos, o valor da função dual $Z(RP_{\lambda\mu})$ é então obtido por

$$Z(RP_{\lambda\mu}) = - \sum_i \lambda_i - \sum_j \mu_j + \min_{x \in \varphi_1} \left\{ \sum_{(i,j)} (l_{ij} + \lambda_i + \mu_j)\, x_{ij} \right\}$$

e reduz portanto, a busca de uma 1-árvore de custo total mínimo, onde cada arco $u = (i,j)$ de G tem um custo $\tilde{l}_{ij} = l_{ij} + \lambda_i + \mu_j$.

Notamos aqui que uma 1-árvore de custo mínimo pode ser encontrada, determinando-se uma árvore de custo mínimo, a qual possui todos os vértices de $X - \{1\}$, e posteriormente adicionando a esta árvore dois arcos distintos de menor custo, incidentes no vértice 1. Então o cálculo da função dual reduz-se essencialmente a determinar uma árvore mínima geradora (HELD e KARP(1971)).

No caso simétrico, onde temos $l_{ij} = l_{ji}$ para qualquer $(i,j) \in U$ e a orientação dos arcos irrelevante, em vez de impor as duas condições C1 e C2 separadamente, é suficiente garantir a existência de precisamente dois arcos incidentes. Portanto, as restrições (1) e (2) podem ser substituidas por uma única condição

$$\sum_{\substack{j \,/ \\ \forall\, (i,j) \in U}} x_{ij} + \sum_{\substack{j \,/ \\ \forall\, (i,j) \in U}} x_{ji} = 2 \qquad (\forall\, i \in X)$$

Temos então não mais que N restrições de grau sobre os vértices e determinamos a formulação do caixeiro viajante simétrico sugerida por HELD e KARP (1971), como veremos a seguir.

$$
\begin{cases}
\min \sum_{(i,j) \in U} l_{ij} x_{ij} \\
\text{sujeito a} \\
\sum_{\substack{j / \\ (i,j) \in U}} x_{ij} + \sum_{\substack{j / \\ (i,j) \in U}} x_{ji} = 2 \qquad (\forall i \in X) \\
x \in \varphi_1
\end{cases}
$$

O problema dual é dado abaixo

$$
Z(D) = \max_{\lambda \text{ irrestrito}} Z(RP_\lambda)
$$

onde $\quad Z(RP_\lambda) = -2 \sum_{i=1} \lambda_i + \min_{x \in \varphi_1} \sum_{(i,j) \in U} (l_{ij} + \lambda_i + \lambda_j) x_{ij}$.

A otimização do problema dual, tanto no caso não simétrico quanto no caso simétrico, pode ser feita pelo método subgradiente, visto no capítulo anterior.

A experiência computacional apresentada no trabalho de HELD e KARP (1971), a respeito do problema simétrico para grafos contendo mais de 64 vértices, mostra que o gap dual é geralmente muito pequeno, raramente excedendo 1 ou 2 por cento, e também que a otimização do problema dual pelo método subgradiente, tem fornecido bons limitantes inferiores, os quais são usados no método branch-and-bound, com árvores de tamanho reduzido, para determinar soluções ótimas exatas.

As experiências de BAZARAA e GOODE (1977) mostram que no caso não simétrico o gap dual também é de tamanho reduzido permitindo encontrar bons limitantes inferiores, o que possibilita

resolver problemas de tamanho relativamente grande.

Recentemente, progressos importantes têm sido feito na solução do problema do caixeiro viajante, combinando relaxação lagrangeana e planos de corte. A idéia consiste em adicionar à formulação descrita, restrições extras correspondentes as faces do casco convexo das soluções inteiras. O uso dessas restrições, chamadas cortes, exclui soluções não inteiras e resulta em reduzir apreciavelmente o gap dual. Entretanto, uma completa caracterização do casco convexo de pontos inteiros não é conhecido no presente para o problema do caixeiro viajante, mas um grande número de tipos de cortes podem ser identificados. Nos casos simétrico e não simétrico problemas com 318 e 325 vértices respectivamente, foram resolvidos por esta abordagem (MINOUX (1989)), o que demonstra o grande sucesso e o avanço na solução de problemas inteiros com auxílio da relaxação lagrangeana.

2. Problema de Localização de Armazéns

O problema de *localização de armazéns com capacidade* consiste em alocar um certo número de armazéns os quais têm que servir um conjunto de consumidores a um custo mínimo, onde a cada consumidor tem associada uma demanda e existem restrições sobre a quantidade total de mercadorias demandadas de cada armazém. Os armazéns são possíveis de serem localizados apenas nos pontos de consumo.

Podemos formular este problema como um problema de programação linear inteira mista. Para isto considere os seguintes dados e variáveis:

Dados:

q_j é a demanda mínima do consumidor j

d_{ij} é o custo de atender toda a demanda do consumidor j por um armazém instalado na localidade i, com $d_{ij} \geq 0$

Q_i é a capacidade do armazém em i

l_i é o limite inferior sobre a demanda atendida por um armazém instalado na localidade i

p_l é o limite inferior sobre o número de armazéns a serem instalados

p_u é o limite superior sobre o número de armazéns a serem instalados

F_i é o custo fixo de instalação do armazém na localidade i

Variáveis:

x_{ij} é a fração de demanda de um consumidor j atendida por um armazém em i.

$$y_i = \begin{cases} 1 & \text{se um armazém for instalado na localidade i,} \\ 0 & \text{caso contrário.} \end{cases}$$

Então com n consumidores o problema de localização de armazéns com capacidade pode ser formulado por:

$$Z(P) = \min \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_{ij} + \sum_{i=1}^{n} F_i Y_i , \qquad (1)$$

sujeito a

$$\sum_{i=1}^{n} x_{ij} \geq 1, \qquad j = 1,\ldots,n \qquad (2)$$

$$l_i Y_i \leq \sum_{j=1}^{n} q_j x_{ij} \leq Q_i Y_i , \qquad i = 1,\ldots,n \qquad (3)$$

(P)

$$p_l \leq \sum_{i=1}^{n} Y_i \leq p_u \qquad (4)$$

$$0 \leq x_{ij} \leq 1, \qquad i = 1,\ldots,n \quad j = 1,\ldots,n \qquad (5)$$

$$Y_i \in \{0,1\} \qquad i = 1,\ldots,n \qquad (6)$$

As inequações (2) garantem que as demandas mínimas de todos os consumidores serão atendidas. As inequações (3) garantem que armazéns que não foram instalados (y_i=0) não ofertem mercadorias e que o total de mercadorias ofertado pelo armazém i (y_i=1) está dentro dos limites impostos sobre ele. As inequações (4) garantem que o número de armazéns está entre p_l e p_u. As inequações (5) definem x_{ij} como uma fração e finalmente as equações (6) são as restrições lógicas, sobre a instalação ou não de um armazém.

Podemos ter casos onde os armazéns têm capacidade suficiente para atender qualquer consumidor, portanto as restrições sobre a capacidade podem ser esquecidas, este problema é chamado de *localização simples* e pode ser formulado como veremos a seguir:

$$Z(P) = \min \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_{ij} + \sum_{i=1}^{n} F_i Y_i , \qquad (1)$$

sujeito a

$$\sum_{i=1}^{n} x_{ij} \geq 1, \qquad j = 1,\ldots,n \qquad (2)$$

$$x_{ij} \leq Y_i , \qquad i = 1,\ldots,n \quad j = 1,\ldots,n \qquad (3)$$

$$0 \leq x_{ij} \leq 1, \qquad i = 1,\ldots,n \quad j = 1,\ldots,n \qquad (5)$$

$$Y_i \in \{0,1\} \qquad i = 1,\ldots,n \qquad (6)$$

É sabido que o problema de localização é difícil de resolver e tem sido extensivamente estudado. ERLENKOTTER (1978) apresenta um método para resolver o problema de localização simples. Para isto obtém um problema linear contínuo, relaxando as restrições de integralidade sobre as variáveis y_i, resolvendo-o por um método dual. Os gaps duais são muito pequenos e em vários casos a solução do problema dual é uma solução primal inteira, a qual é ótima. Se as soluções obtidas não forem inteiras, então será aplicado o método branch-and-bound que, segundo o autor, apresenta um tempo computacional reduzido.

Voltando ao problema de localização com capacidade, vários métodos com soluções aproximadas para (P) estão sendo desenvolvidos, muitos dos quais utilizando a relaxação lagrangeana.

Em alguns casos as restrições sobre os consumidores (2) são relaxadas na tentativa de encontrar uma solução factível através do problema lagrangeano. Tal relaxação pode ser vista a seguir.

$$\min \sum_{i=1}^{n} \sum_{j=1}^{n} (d_{ij} - \lambda_j) x_{ij} + \sum_{i=1}^{n} F_i Y_i + \sum_{j=1}^{n} \lambda_j \qquad (7)$$

sujeito a

$$l_i Y_i \leq \sum_{j=1}^{n} q_j x_{ij} \leq Q_i Y_i , \qquad i = 1,\ldots,n \qquad (3)$$

$$p_l \leq \sum_{i=1}^{n} Y_i \leq p_u \qquad (4)$$

$$0 \leq x_{ij} \leq 1, \qquad i = 1,\ldots,n \quad j = 1,\ldots,n \qquad (5)$$

$$Y_i \in \{0,1\} \qquad i = 1,\ldots,n \qquad (6)$$

CRHISTOFIDES e BEASLEY(1983) apresentam um método de resolução para o problema (P), utilizando o procedimento branch-and-bound com limitantes dados por esta relaxação lagrangeana. Esta relaxação pode ser facilmente resolvida por um algoritmo apresentado por tais autores, no qual as variáveis duais são determinadas pelo método subgradiente. Apresentam também resultados computacionais (para problemas com 50 armazéns e 150 consumidores) onde mostram que o gap dual é muito pequeno.

MACULAN e REINOSO (1989) apresentam um método exato para resolver o problema de localização capacitado. Para tal, usam o algoritmo branch-and-bound com limitantes obtidos pelo método de ajustamento de multiplicadores, que são heurísticas desenvolvidas em aplicações específicas que exploram alguma estrutura especial do problema dual. Descrevem também um modelo de decomposição lagrangeana, o qual consiste da formulação de um problema equivalente ao problema original (P), trabalhando com as variáveis de decisão, utilizando-as em cada conjunto de restrições de (P), e definindo uma relaxação lagrangeana para a nova formulação, a qual se decompõe aditivamente em vários subproblemas, conservando um dos conjuntos de restrição de (P). Os autores apresentam resultados de experiências computacionais, as quais mostram que a técnica referida é razoavelmente eficiente para problemas de médio porte.

Outros trabalhos com este tipo de método têm sido

desenvolvidos, e com isto têm-se definido novas e diferentes formas para uso da relaxação lagrangeana.

3. Problema de Designação Generalizada

O problema de designação generalizada é o problema de designar n tarefas a m homens a um custo mínimo, onde cada tarefa é atribuída a apenas um homem, ao qual está associda uma capacidade de trabalho.

Notação:

c_{ij} é o custo de designar a tarefa j ao homem i

a_{ij} é o tempo que o homem i requer para executar a tarefa j

b_i é o tempo disponível para o homem i

$$x_{ij} = \begin{cases} 1 & \text{se o homem i executa a tarefa j} \\ 0 & \text{caso contrário} \end{cases}$$

Formulação:

$$Z(P) = \text{mínimo} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}\, x_{ij} \qquad (1)$$

sujeito a

$$(P) \qquad \sum_{i=1}^{m} x_{ij} = 1 \quad , \quad j = 1,\ldots,n \qquad (2)$$

$$\sum_{j=1}^{n} a_{ij}\, x_{ij} \le b_i \quad , \quad i = 1,\ldots,m \qquad (3)$$

$$x_{ij} \in \{ 0, 1 \}, \quad i = 1,\ldots,m \ , \ e \ j = 1,\ldots,n \qquad (4)$$

As restrições (2) asseguram que cada tarefa é designada a apenas um homem, (3) respeita o tempo disponível de cada homem.

Apresentamos duas relaxações lagrangeanas para o problema. A primeira é por dualização das restrições (2):

$$Z_1(RP_u) = \text{mínimo} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}\, x_{ij} + \sum_{j=1}^{n} u_j \left(1 - \sum_{i=1}^{m} x_{ij} \right)$$

sujeito a

$$(RP_u) \qquad \sum_{j=1}^{n} a_{ij}\, x_{ij} \le b_i \quad , \quad i = 1,\ldots,m$$

$$x_{ij} \in \{ 0, 1 \}, \quad i = 1,\ldots,m \ , \ e \ j = 1,\ldots,n$$

onde $u = (u_1\ u_2\ \ldots\ u_n)$

Este problema reduz-se a m problemas da mochila 0-1 independentes e podem, segundo FISHER(1981), ser resolvidos em tempo proporcional a $n \sum_{i=1}^{m} b_i$.

A segunda relaxação é obtida dualizando as restrições (3) com $v \ge 0$

$$Z_2(RP_v) = \text{mínimo} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}\, x_{ij} + \sum_{i=1}^{m} v_i \left(\sum_{j=1}^{n} a_{ij}\, x_{ij} - b_i \right)$$

sujeito a

$$(RP_v) \qquad \sum_{i=1}^{m} x_{ij} = 1 \quad , \quad j = 1,\ldots,n \qquad (1)$$

$$x_{ij} \in \{ 0, 1 \}, \quad i = 1,\ldots,m \ , \ e \ j = 1,\ldots,n \qquad (2)$$

onde $v = (v_1\ v_2\ \ldots\ v_m) \geq 0$

Esta relaxação é definida para $v \geq 0$ pois é necessário assegurarmos que $Z_2(RP_v) \leq Z(P)$. Tal problema é fácil de resolver (FISHER (1981)) em tempo proporcional a m.n, determinando-se o $(c_{ij} + v_i a_{ij}) = \min_k (c_{kj} + v_k\, a_{kj})$ para cada j, fazendo-se $x_{ij} = 1$ e os x_{kl} restantes iguais a zero.

Podemos verificar que $Z_2(RP_v)$ não mudará de valor caso substituirmos (2) por $0 \leq x_{ij} \leq 1$, $i = 1,\ldots,$ n e $j = 1, \ldots,$ n. Isto indica que a relaxação possui a propriedade de integralidade. Portanto, esta não seria uma boa escolha de restrições a relaxar do ponto de vista de limitantes, comparando-o com $(\bar{P})$.

Temos visto, neste capítulo, exemplos de aplicações da relaxação lagrangeana. Vimos também, que tal método tem proporcionado boas soluções para o problema (P). Mas, as soluções do problema lagrangeano são sempre factíveis para o problema original ?

A solução da relaxação lagrangeana pode ser infactível para o problema (P), porém fornece subsídios para encontrar uma boa solução factível de (P) com uso de heurísticas. Exemplificaremos tal idéia , baseando-nos no trabalho de FISHER(1981), para um problema de designação generalizada. Para isto, imagine o problema de designação generalizada como o problema de colocar n ítens dentro de m mochilas, usando cada ítem apenas uma vez.

Portanto, no problema (RP_u) as restrições relaxadas $\sum_{i=1}^{m} x_{ij} = 1$, $j = 1,\ldots,n$ (que obrigam que cada ítem seja usado apenas uma vez) podem ser violadas. Suponha $\tilde{x}$ a solução

ótima da relaxação (RP_u) e defina :

$$S_1 = \left\{ j \ / \ \sum_{i=1}^{m} \tilde{x}_{ij} = 0 \right\}$$

$$S_2 = \left\{ j \ / \ \sum_{i=1}^{m} \tilde{x}_{ij} = 1 \right\}$$

$$S_3 = \left\{ j \ / \ \sum_{i=1}^{m} \tilde{x}_{ij} > 1 \right\}$$

As restrições de (P) são violadas por $\tilde{x}$ correspondente ao índice $j \in (S_1 \cup S_3)$.

Seja $j \in S_3$. Para satisfazer a restrição violada basta retirar o ítem j de todas as mochilas, com exceção de uma. Para determinar em qual mochila deixar o ítem j, FISHER(1981) usa a mochila que $\max_i \{(c_{ij} - u_i) / a_{ij}.\}$.

Agora tomemos $j \in S_1$. Para completar a construção da solução factível, basta designar ítens de S_1 para as mochilas. O autor, sugere ordenar os ítens em S_1 por valor decrescente de $\sum_{i=1}^{m} a_{ij}$ e colocar cada ítem, por sua vez, dentro de uma mochila com capacidade suficiente, a qual maximiza $(c_{ij} - u_i) / a_{ij}$.

CONCLUSÕES

Através de observações de experiências publicadas, concluimos que a relaxação lagrangeana fornece uma das técnicas mais eficientes para produzir bons limitantes para problemas de programação linear inteira, os quais reduzem substancialmente processos de busca em árvore. Geralmente, na escolha das restrições a relaxar, explora-se alguma estrutura especial apresentada pelo problema original, o que facilita a resolução do problema relaxado. Vários trabalhos (em particular MINOUX(1989)) mostram que um número razoavelmente grande de problemas, tem sido resolvido com sucesso usando a relaxação lagrangeana. Como exemplo temos os problemas de localização, cobertura, mochila quadrático, designação, caixeiro viajante, multi-fluxos, sequenciamento e muitos outros problemas isolados (FISHER(1985)), o que torna evidente a potencialidade da relaxação lagrangeana.

Observamos também que boas aproximações para o valor ótimo do problema dual podem ser conseguidas com o uso do método subgradiente, sendo recomendado por alguns autores, como MINOUX(1989), em vista, principalmente de sua simplicidade de implementação.

Experiências mostram que os gaps duais, diferença entre os valores ótimos dos objetivos dual e do problema original, são em geral, pequenos. Assim, o interesse da relaxação lagrangeana não está restrito apenas à determinação de bons limitantes, podendo também ser usada para obter boas aproximações para a solução ótima do problema original ou até mesmo soluções exatas.

A técnica de relaxação lagrangeana tem sido aplicada à problemas do tipo combinatoriais com procedimentos enumerativos em conjunção com outros métodos, como métodos de planos de cortes, restrição de substituição e outros, melhorando assim os limitantes

inferiores ou em alguns casos, obtendo uma solução ótima inteira.

Pelo que podemos observar nos trabalhos publicados, a relaxação lagrangeana tem sido muito usada com resultados satisfatórios, embora na prática exija o uso de algumas heurísticas, como para tornar a solução do problema lagrangeano factível para o problema original. Segundo FISHER(1981), tais soluções, quase sempre, são infactíveis, porém, são "quase" factíveis, podendo torná-las factíveis por heurísticas específicas ao problema, como por exemplo aquela descrita na seção 3 do capítulo V, para problemas de designação generalizada. Talvez, este fato seja uma das maiores dificuldades apresentadas na utilização deste método, embora esta dificuldade não tenha sido salientada nos trabalhos estudados.

O interesse pelo tema deste trabalho foi dispertado por seu extenso uso e praticidade. Nesta etapa, procuramos dar um enfoque mais teórico para esta abordagem, buscando os principais conceitos e relações que se fazem necessárias para entendimento e uso do método.

Portanto, é de nosso interesse em uma segunda fase de trabalho, implementar computacionalmente este método de relaxação para um problema específico, comparando o método subgradiente com outras formas de resolução do problema dual. MINOUX(1989) apresenta alguns problemas combinatoriais que foram resolvidos com sucesso, usando relaxação lagrangeana e planos de cortes, os quais foram sugeridos por GEOFFRION(1974). Portanto, também faz parte de nossas perspectivas futuras, estudar tal abordagem.

Como vimos, esta é uma área de extrema importância, por ser bastante aplicada a problemas práticos. Portanto todo esforço que se fizer na busca de melhores métodos ou aperfeiçoamento dos já existentes, será imprescindível no processo de desenvolvimento.

Propriedade 1

Sejam

$$(1)\quad Z(P) = \min cx \quad Ax \geq b \quad x \geq 0 \qquad (2)\quad Z(D) = \max_{\lambda \geq 0} \left(\min_{x \geq 0} cx + \lambda(b - Ax) \right)$$

Como (1) é um problema linear, então (2) coincide com seu dual padrão.

Prova:

$\min (cx + \lambda(b - Ax))$ implica

$\min ((c - \lambda A)x + \lambda b)$.

Este mínimo existe se e somente se $c - \lambda A \geq 0$, ou $\lambda A \leq c$.

Fazendo $D = \{ \lambda \ / \ \lambda A \leq c, \ \lambda \geq 0 \}$.

Para qualquer $\lambda \in D$, o mínimo para cx é alcançado quando $(c - A\lambda)x = 0$.

Assim, $\forall \lambda / \lambda A \leq c$,

$$\min cx + \lambda(b - Ax) = \lambda b. \quad \blacksquare$$

Definição 1

Um conjunto $S \subset R^n$ é convexo se e somente se

$$\left.\begin{array}{l} \forall x \in S, \\ \forall y \in S, \\ \forall \alpha \in [0 , 1]. \end{array}\right\} \Rightarrow \alpha x + (1-\alpha)y \in S.$$

Definição 2

O casco convexo de um conjuto A, $A \subset R^n$, denotado por $Co\{A\}$ é o menor conjunto convexo que contém A.

Exemplo:

Propriedade 2

Seja $A \subseteq R^n$ conjunto compacto, $f: A \longrightarrow R$ linear, e considere **a** o ponto de máximo da função f.

Se a função f acima for definida como $f: Co\{A\} \longrightarrow R$ então o ponto de máximo de f será **a**.

Prova:

Suponha por absurdo que **a** não seja o ponto de máximo da função f definida no casco convexo de A. Então $\exists\ b \in Co(A)$ / $K = f(b) > f(a)$.

Seja $\varepsilon > 0$ / $K - \varepsilon > f(a)$.

Definimos $S_{k-\varepsilon} = \{ x \in R^n / f(x) \leq K - \varepsilon \}$ que é fechado e convexo.

Desde que $A \subset S_{k-\varepsilon}$

segue $Co\{A\} \subset S_{k-\varepsilon}$. Absurdo pois $b \in Co\{A\}$ e

$K = f(b) < K - \varepsilon = f(b) - \varepsilon$. ■

Definição 3

Uma função $f: R^n \longrightarrow R$ definida sobre um conjunto convexo S é convexa se satisfizer

$$\forall\, x \in S,\ \forall\, y \in S,\ \forall\, \alpha \in [0,\ 1]:$$
$$f(\ \alpha x + (1 - \alpha)y\ ) \leq \alpha\, f(x) + (1 - \alpha)\, f(y),$$

e f será côncava se a função $-f(x)$, $x \in S$, for convexa.

Definição 4

O epigrafo de uma função $f : R^n \longrightarrow R$, denotado por epi [f], é o conjunto:

$$\{\ (\mu,\ x):\ f(x) \leq \mu,\ x \in R^n,\ \mu \in R\ \} \subset R^{n+1}\ .$$

Propriedade 3

Uma função f é convexa se e somente se epi [f] é um conjunto convexo.

Definição 5

Dizemos que um problema de programação matemática é convexo se consistir de uma minimização de uma função convexa sobre um domínio convexo.

Definição 6

Uma função $f: X \longrightarrow R$, $X \subset R^n$, diz-se semi contínua inferior no ponto $\mathbf{a} \in X$, quando para cada $\varepsilon \in R$, $\varepsilon > 0$, existe uma vizinhança V de **a** tal que $x \in V \Rightarrow f(\mathbf{a}) - \varepsilon \leq f(x)$.

BIBLIOGRAFIA

AGMON, S. (1954) - "The Relaxation Method for Linear Inequalities" *Canadian Journal of Mathematics*, Vol. 6., p. 382-392.

BAZARAA, M. S. & GOODE, J. J. (1977) - "The Travelling Salesman Problem: A Dual Approach", *Mathematical Programming*, 13, 221-237.

CHISTOFIDES, N. & BEASLEY, J. E. (1983) - "Extension to a Lagrangean Relaxation Approach for the Capacitated Warehouse Location Problem", *European Journal of Operation Research*, Vol. 13, p. 19-28.

EREMIN, I. J. (1962) - "An Iterative Method for Cebysev Approximations of Incompatible Systems of Linear Inequalities", *Soviet Mathematics Doklady*, Vol. 3, p. 570-672.

EREMIN, I. J. (1965) - "A Generalization of the Motzkin-Agmon Relaxation Method", *Soviet Mathematics Doklady*, Vol. 3, p. 570-672.

ERLENKOTTER, D. (1978) - " A Dual-Based Procedure for Uncapacited Facility Location", *Operations Research*, Vol. 26, No. 6, p. 992-1009.

ERMOLIEV, Yu M. (1966) - " Methods of Solution of Nonlinear Extremal Problems", *Kibernetika*, Vol. 2, No. 4, p. 1-17.

EVERETT, H. (1963) - "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources", *Operations Research*, vol. 11, p. 339-417.

FISHER, M. L., (1981) - "The Lagrangian Relaxation Method for Solving Integer Programming Problems", *Management Science*,

Vol. 27, p. 1-18.

FISHER, M. L., (1985) - "An Applications Oriented Guide to Lagrangian Relaxation", *Interfaces* 15, p. 10-21.

GEOFFRION, A. M. (1974) - "Lagrangian Relaxation of Integer Programming", *Mathematical Programming Study* 2, p. 82-114.

GONDRAN, M. & MINOUX, M. (1979) - "Graphes et algorithmes", Collection de la Direction des Éstudes et Recherches EDF, Eyrolles, Paris.

GUIGNARD, M. & ROSENWEIN, M. B. (1989) - "An Application-Oriented Guide for Designing Lagrangean Dual Ascendent Algorithms", *European Journal of Operational Research*, Vol. 43, No. 2, p. 197-205.

HELD, M. & KARP, R. M. (1970) - "The Traveling Salesman Problem and Minimum Spanning Trees", *Operations Research*, Vol. 18, p. 1138-1162.

HELD, M. & KARP, R. M. (1971) - "The Traveling Salesman Problem and Minimum Spanning Trees: Part II", *Mathematical Programming* 1, p. 6-25.

HELD, M., WOLFE, P. & CROWDER, H. D. (1974) - "Validation of Subgradient Optimization", *Mathematical Programming* 6, p. 62-88.

LASDON, L. S. (1970) - " *Optimization Theory for Large Systems*", Macmillan Series in Operations Research.

LORIE, J. & SAVAGE, L. J. (1955) - "Three Problems in Capital Rationing", *Journal of Business*, p. 229-239.

MACULAN, N.; CAMPELLO, R. E. & LOPES, L. B. R. (1984) - "Relaxação Lagrangeana para Programação Inteira" , VII *Congresso Nacional de Matematica Aplicada e Computacional*, Campinas, p. 293-315.

TAHA, H. A. (1975) - "*Integer Programming, Theory, Applications, Computations*", Academic Press, New York.

ICMSC-USP
PÓS-GRADUAÇÃO

INSTITUTO DE CIÊNCIAS MATEMÁTICAS E CUSP-USP-SÃO CARLOS