Decomposição em Programação

Inteira

: Kasuo Aoyanagi

orientador : Prof. Dr. Marcos Nereu Arenales

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

Ao meu pai e ao meu irmão (In Memorian) A minha mãe e aos meus irmãos

Ao Prof. Dr. Marcos Nereu Arenales, por sua orientação segura, dedicada e paciente durante todo o período de preparação deste trabalho; aos professores do "Departamento de Matemática Computacional", pelos ensinamentos recebidos; ao meu pai e ao meu irmão que sempre estiveram ao meu lado; a minha mãe e aos meus irmãos, , pela compreensão, pelos incentivos recebidos e pelos esforços feitos para que eu conseguisse concluir mais uma etapa de minha vida; aos professores e funcionários do ICMSC - USP, pela acolhida e atenção dispensada; aos professores da UFSCar - Departamento de Matemática, em especial aos professores Artur Darezzo Filho, Dorival Marcos Milani, Carlos Lazarini e principalmente à professora Selma Helena de Vasconcelos Arenales que mostraram que a matemática não é um "bicho de sete cabeças"; aos meus amigos, em especial a Helenice e a Marisa, cuja a amizade sempre honrarei; a Capes, que criou condições para a realização do curso de Mestrado.

Sempre me pareceu estranho que todos aqueles que estudam seriamente esta ciência acabam tomados de uma espécie de paixão pela mesma. Em verdade, o que proporciona o máximo prazer não é o conhecimento e sim a aprendizagem, não é a posse mas a aquisição, não é a presença mas o ato de atingir a meta.

"Carl Friedrich Gauss"

Resumo

Muitos métodos de resolução de problemas de programação inteira são baseados na decomposição primal ou dual. Nossa proposta é rever alguns algoritmos de decomposição primal, dual e um novo método proposto por Van Roy, o qual designamos por decomposição híbrida, onde procura-se utilizar ambas as decomposições em um único procedimento.

Além disso, fizemos uma implementação do algoritmo de decomposição híbrida para resolver uma aplicação prática.

Abstract

Some methods for solving integer programming problems are based on primal and dual decomposition. In this work, we propose to review some algorithms of primal and dual decomposition and a new method proposed by Van Roy, the cross decomposition method, which uses simultaneously both decomposition.

Futhermore, we implemented the cross decomposition algorithm to solve a pratical problem.

Prefácio

Muitos problemas de otimização, apresentam variáveis que assumem valores reais bem como variáveis que assumem valores inteiros e as relações entre as variáveis são lineares, assim como a função objetivo a ser otimizada é linear nas variáveis de decisão. Isto define um problema de programação linear inteira mista, que será objeto de especial atenção deste trabalho.

Um dos primeiros algoritmos estudados para a resolução desse tipo de problema foi apresentado por Gomory, em 1958, que baseia-se em adicionar sistematicamente restrições (planos de cortes) ao problema original, com a condição de integralidade relaxada, de forma que as "as soluções fracionárias" sejam eliminadas.

Muitos dos métodos conhecidos hoje, para a resolução de problemas de programação linear inteira mista, são baseados na decomposição primal ou dual do problema, que segue a idéia de planos de cortes de Gomory. Neste trabalho revemos alguns métodos clássicos que são baseados na decomposição primal ou dual e um método híbrido apresentado inicialmente por Van Roy, que utiliza ambas as decomposições em um único procedimento e fazemos um estudo de cada caso.

No capítulo 1 introduziremos a notação e a terminologia juntamente com alguns teoremas que serão empregados durante o trabalho.

No capítulo 2 revemos alguns métodos clássicos de decomposição primal e decomposição dual.

No capítulo 3 apresentaremos o método de decomposição Híbrida.

No capítulo 4 modelamos um problema de sequenciamento de uma fundição e apresentamos alguns resultados computacionais.

No capítulo 5 apresentamos as conclusões finais.

No apêndice A, encontram-se definições e conceitos básicos utilizados neste trabalho.

No apêndice B, encontra-se o método de decomposição de "Rosen", que é utilizado neste trabalho.

Ao final, apresentamos as referências bibliograficas usadas na elaboração deste trabalho.

Conteúdo

1	Def	inições	e Resultados Básicos	1
	1.1	Introd	ução	1
	1.2	Dualio	lade	3
		1.2.1	Propriedades do Problema Dual	6
		1.2.2	Dualidade em Programação Linear	10
2	Mé	todos o	de Decomposição	12
	2.1	Introd	lução	12
	2.2	Métod	lo de decomposição de Benders	12
		2.2.1	Algoritmo	19
		2.2.2	Considerações Finais	20
	2.3	Relax	ação Lagrangeana	21
		2.3.1	Introdução	21
		2.3.2	Método do Subgradiente	22
		2.3.3	Algoritmo	26
		2.3.4	Convergência	27
		2.3.5	Considerações Finais	29

3	Dec	omposição Hibrida	31
	3.1	Introdução	31
	3.2	Decomposição Híbrida	35
	3.3	Equivalência entre os Subproblemas	36
	3.4	Convergência do Algoritmo de Decomposição Híbrida	44
	3.5	Algoritmo de Decomposição Híbrida	45
	3.6	Considerações Finais	47
4	Est	udo do Caso: Planejamento de uma Fundição	48
	4.1	Definição do Problema	48
	4.2	Modelo Matemático	49
		4.2.1 Formulação	50
		4.2.2 Exemplo	51
		4.2.3 Comentários	54
	4.3	Método de Decomposição Híbrida aplicado ao Modelo	55
		4.3.1 Introdução	55
		4.3.2 Subproblema Primal	56
		4.3.3 Subproblema Dual	57
	4.4	Exemplo	59
		4.4.1 Problema Mestre	60
	4.5	Solução do Problema Teste:	61
	4.6	Implementação	62
	4.7	Considerações Finais	63
5	Coı	nclus ão	64

Apêndice A	66
Apêndice B	69
Bibliografia	72

. "

Capítulo 1

Definições e Resultados Básicos

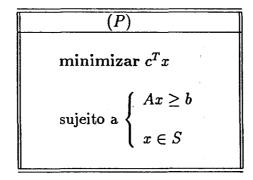
Neste capítulo inicial apresentaremos algumas definições e teoremas de dualidade, que serão referidos durante o trabalho. Omitiremos aqui algumas demonstrações dos teoremas, os quais podem ser encontrados em livros de programação matemática e de análise convexa.

1.1 Introdução

Um problema de programação linear inteira mista consiste em minimizar (ou maximizar) uma função linear (função objetivo) cujas variáveis devem satisfazer ao seguinte conjunto de restrições:

- 1. um sistema de equações ou inequações lineares.
- 2. restrições de integralidade de algumas variáveis.

Considere o seguinte problema de programação linear mista:



onde:

c, x : são vetores n-dimensionais.
b : é um vetor m-dimensional.

S: conjunto do Rⁿ, em geral indicando não negatividade e integralidade do algumes goordonades do um veter a C. S.

integralidade de algumas coordenadas de um vetor $x \in S$.

A: matriz de ordem $m \times n$.

Em função da integralidade de algumas variáveis e da estrutura especial do conjunto de restrições, podemos particionar a matriz A de tal forma que $n = n_1 + n_2$ e $m = m_1 + m_2$ e o problema (P) pode ser reescrito da seguinte forma:

(P) minimizar $c^1x_1 + c^2x_2$

sujeito a
$$\begin{cases} A_1^1 x_1 + A_1^2 x_2 \ge b_1 \\ A_2^1 x_1 + A_2^2 x_2 \ge b_2 \\ (x_1, x_2) \in S \end{cases}$$

onde:

$$\begin{split} c^T &= (c^1, c^2) \\ x^T &= (x_1, x_2) \\ S &= \{ \ x^T = (x_1, x_2) \ \text{tal que } x_1 \in \mathbf{R}_+^{\mathbf{n}_1} \ \text{e} \ x_2 \in \mathbf{Z}_+^{\mathbf{n}_2} \ \} \end{split}$$

Usaremos a seguinte notação:

- (.) problema a ser otimizado.
- v(.) valor da função objetivo avaliado numa solução ótima.
- F(.) região factível.
- (-) é o problema (.) com as restrições de integralidade relaxadas.

1.2 Dualidade

Considere o seguinte problema de programação matemática:

Primal (P)	_
minimizar $f(x)$	
$\int g(x) \leq 0$	
sujeito a $\left\{ x \in S \right\}$	

onde:

$$S \subseteq \mathbf{R}_+^{\mathbf{n}}$$
.
 $f: S \mapsto \mathbf{R}$.
 $g^T(x) = (g_1(x), \dots, g_m(x))$ tal que $g_i: S \mapsto \mathbf{R}$.

O que se desenvolverá a seguir é uma análise das condições que um vetor x^0 deve satisfazer, para ser solução ótima de (P)

Definição 1.2.1 Associado ao problema (P), definimos a "função lagrangeana" da seguinte forma:

$$L(x, u) = f(x) + u^{T}g(x)$$

onde $u \in \mathbf{R}^{\mathbf{m}}_{+}$ é chamado vetor multiplicador de Lagrange.

Definição 1.2.2 Seja $x^0 \in S$ e $u^0 \ge 0$, então (x^0, u^0) é um ponto de sela se satifizer:

1.
$$L(x^0, u^0) \le L(x, u^0)$$
, para todo $x \in S$.

2.
$$L(x^0, u^0) \ge L(x^0, u)$$
, para todo $u \ge 0$.

Isto é, x^0 minimiza $L(x,u^0)$ sobre S e u^0 maximiza $L(x^0,u)$ em todos os pontos $u \ge 0$.

Teorema 1.2.1 Seja $u^0 \ge 0$ e $x^0 \in S$ então (x^0, u^0) é um ponto de sela para L(x, u) se, e somente se,

- 1. x^0 minimiza $L(x, u^0)$ sobre S.
- 2. $g(x^0) \leq 0$.
- 3. $(u^0)^T g(x^0) = 0$.

A demonstração pode ser encontrada em Lasdon [15] página 84.

Teorema 1.2.2 (suficiência) Seja $u^0 \ge 0$ e $x^0 \in S$ se (x^0, u^0) é um ponto de sela para L(x, u), então x^0 resolve o problema (P).

A demonstração pode ser encontrada em Lasdon [15] página 85.

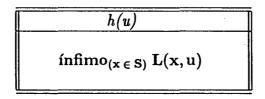
Teorema 1.2.3 Suponha que o problema (P) seja convexo, isto é, S é um conjunto convexo, f_i e g_i ($i=1,\ldots,m$) são funções convexas e que exista um ponto $x^0 \in S$ tal que $g(x^0) < 0$. Se x^0 resolve o problema (P), então existe um vetor de multiplicadores $u \geq 0$ tal que (x^0, u^0) é um ponto de sela de L(x, u).

A demonstração pode ser encontrada em Lasdon [15] página 88.

Iremos assumir as seguintes hipóteses para a análise do problema dual:

- ullet f e g são funções contínuas.
- o conjunto $S \subseteq \mathbf{R}^n$ é compacto (fechado e limitado).

Definição 1.2.3 Definimos a "função dual h(u)" como o ínfimo de $x \in S$ da função lagrangeana, para cada $u \ge 0$.



Lema 1.2.1 A função dual é côncava.

A demonstração pode ser encontrada em Lasdon [15] página 412.

Lema 1.2.2 A função dual é contínua.

A demonstração pode ser encontrada em Rockafellar [22] página 82.

Teorema 1.2.4 O valor da função dual é um limitante inferior do valor da função primal, isto é, $h(u) \leq f(x)$, para todo $u \geq 0$ e para todo $x \in S$ tal que $g(x) \leq 0$.

A demonstração pode ser encontrada em Lasdon [15] página 412.

O Teorema 1.2.4 sugere o problema de determinar o melhor limitante inferior dentre as infinitas famílias da relaxação $\{h(u)\}_{u\geq 0}$ é $Z_{DL}(u^*)$, onde u^* é solução ótima de:

$$\mathbf{Z_{DL}}\left(\mathbf{u^{*}}\right) = \mathbf{maximizar_{(u \ge 0)}} \ \mathbf{h(u)}$$

O problema (DL) é chamado de "Dual Lagrangeano" de (P) em relação ao conjunto de restrições $g(x) \leq 0$ ou problema pseudo-dual.

Corolário 1.2.1 (a) Se o valor ótimo de (DL) é ilimitado (∞), então (P) é infactível.

(b) Se o valor ótimo de (P) é ilimitado ($-\infty$), então (DL) é infactível.

A demonstração pode ser encontrada em Lasdon [15] página 412.

Teorema 1.2.5 Seja $u^0 \ge 0$ e $x^0 \in S$, (x^0, u^0) é um ponto de sela de L(x, u) se, e somente se, x^0 for primal factivel, u^0 for dual factivel e $f(x^0) = h(u^0)$.

A demonstração pode ser encontrada em Lasdon [15] página 413.

Definição 1.2.4 A diferença entre max $\{h(u)\}$ e min $\{f(x)\}$ é definido como sendo o "gap dual" ou seja:

$$gap \ dual = max\{ \ h(u) \ \} - min\{ \ f(x) \ \}$$

1.2.1 Propriedades do Problema Dual

Definição 1.2.5 O vetor m-dimensional γ é um subgradiente em $u \in \mathbb{R}_+^m$ da função Lagrangeana h em \mathbb{R}_+^{n+m} se, e somente se,

$$h(x) - h(u) \le \gamma^{T}(x - u) \text{ para } t \text{ odo } x \in Dom(h) \subseteq \mathbb{R}_{+}^{m}.$$

Lema 1.2.3 Seja $x \in S$ uma solução ótima de h(u), então g(x) é o subgradiente de h(u) no ponto $u = u^0$.

demonstração:

$$h(u^{0}) + (u - u^{0})^{T} g(x) =$$

$$= f(x) + (u^{0})^{T} g(x) + (u - u^{0})^{T} g(x)$$

$$= f(x) + u^{T} g(x) \ge h(u) \text{ para todo } u \ge 0.$$

$$Logo g(x) \text{ \'e um subgradiente de } h(u) \text{ no ponto } u = u^{0}.$$

Definição 1.2.6 O conjunto de todos os subgradientes, denotado por $\partial h(u)$, de h no ponto u é chamado de subdiferencial.

Teorema 1.2.6 Se $\partial h(u) \neq \phi$, então $\partial h(u)$ é fechado e convexo.

demonstração:

$\frac{\partial h(u)}{\partial h(u)}$ é convexo

De fato, seja γ_1 e $\gamma_2 \in \partial h(u)$ então:

$$h(x) \le h(u) + (\gamma_1)^T (x - u) \text{ para todo } x \in S$$
 (1.1)

$$h(x) \le h(u) + (\gamma_2)^T (x - u) \text{ para todo } x \in S$$
 (1.2)

Multiplicando (1.1) por α e (1.2) por $(1-\alpha)$, $0<\alpha<1$ e somando o resultado obtemos:

$$h(x) \le h(u) + [\alpha \gamma_1 + (1 - \alpha)\gamma_2]^T(x - u)$$
 para todo $x \in S$
Assim $\alpha \gamma_1 + (1 - \alpha)\gamma_2 \in \partial h(u)$

$\partial h(u)$ é fechado

Suponha por absurdo que $\{\gamma_i^*\}$ seja uma seqüência convergente de elementos de $\partial h(u)$ com ponto limite igual a γ^* , tal que $\gamma^* \notin \partial h(u)$. Então existe um $x \in S$ e $\varepsilon > 0$ tal que:

$$h(x) - \varepsilon = h(u) + (\gamma^*)^T (x - u)$$
(1.3)

Para qualquer elemento γ_n^* da seqüência $\{\gamma_i^*\}$ temos:

$$h(x) \le h(u) + (\gamma_n^*)^T (x - u)$$
 (1.4)

Substituindo (1.3) em (1.4) e usando a desigualdade de Cauchy-Schwarz (ver apêndice A) obtemos:

$$\varepsilon \le (\gamma_n^* - \gamma^*)^T (x - u) \le ||\gamma_n^* - \gamma^*|| ||x - u||$$

$$\|\gamma_n^* - \gamma^*\| \ge \frac{\varepsilon}{\|x - u\|} > 0 \ para \ todo \ n$$

O que contradiz o fato que $\{\gamma_i\} \to \gamma^*$ logo $\gamma^* \in \partial h(u)$ e portanto $\partial h(u)$ é fechado.

Lema 1.2.4 Se γ é um subgradiente de $\partial h(u)$ no ponto \overline{u} , então o conjunto $\{u \geq 0 \text{ tal que } \gamma^T(u-\overline{u}) \geq 0\}$ contém todas as soluções ótimas do problema de maximização (DL).

demonstração:

Desde que γ é o subgradiente, temos que para todo u:

$$\gamma^T(u-\overline{u}) \ge h(u) - h(\overline{u})$$

Se u é solução ótima para o problema (DL)
$$\Rightarrow$$
 $h(u)$ - $h(\overline{u}) \geq 0$ logo $(u - \overline{u})^T \gamma \geq 0$.

O lema acima afirma, que qualquer subgradiente γ da função h em um ponto \overline{u} aponta para o semi-espaço que contém todas as soluções ótimas do problema (DL).

Teorema 1.2.7 A derivada direcional à direita de h em qualquer ponto \overline{u} e em qualquer direção d, existe e é dada por:

$$\nabla h(\overline{u}, d) = \min_{(\gamma \in \partial h(\overline{u}))} \mathbf{d}^{\mathbf{T}} \gamma$$

A demonstração pode ser encontrada em Grinald [11].

Se \overline{u} é um ponto de ótimo de (DL), então podemos concluir que $\nabla h(\overline{u},d) \leq 0$ em todas as direções, isto é, a função decresce em todas as direções e como a função dual é côncava, o seguinte teorema emerge:

Teorema 1.2.8 O ponto \overline{u} é ótimo em (DL) se, e somente se, $\nabla h(\overline{u},d) \leq 0$ em todas as direções no ponto ótimo.

Para algum $\overline{u} \ge 0$ considere os seguintes conjuntos:

$$I(\overline{u}) = \{ i \text{ tal que } \overline{u}_i = 0 \}$$

 $I^{c}(\overline{u})$ seu complementar

Teorema 1.2.9 O ponto $\overline{u} \geq 0$ é solução ótima de (DL) se, e somente se,

$$m\acute{a}ximo_{(v\in V)} \nabla h(\overline{u},v) = 0 \tag{1.5}$$

onde $V \subseteq \mathbb{R}^m$ é o conjunto que satisfaz:

$$0 \leq v_i \leq 1, i \in I(\overline{u})$$

$$-1 \leq v_i \leq 1, i \in I^c(\overline{u})$$

demonstração:

 (\Leftarrow) máximo $(\mathbf{v} \in \mathbf{V})$ $\nabla h(\overline{u}, v) = 0 \Rightarrow \overline{u} \geq 0$ é solução ótima.

Seja $u \geq 0$ uma outra solução ótima para o problema (DL) tal que $u \neq \overline{u}$.

Então $u - \overline{u} = \theta v$, ($\theta > 0$), $v \in V$.

Como máximo_($v \in V$) $\nabla h(\overline{u}, v) = 0 \Rightarrow \nabla h(\overline{u}, v) \leq 0, \forall v \in V.$

E para qualquer $\overline{\gamma} \in \partial h(\overline{u})$ e $v \in V$ obtemos:

$$h(u) \le h(\overline{u}) + \overline{\gamma}^T (u - \overline{u}) = h(\overline{u}) + \theta v^T \overline{\gamma}$$

Como $v^T \overline{\gamma} \leq 0$ e $\theta > 0$ (por hipótese) $\Rightarrow h(u) \leq h(\overline{u})$.

Desde que u é uma solução ótima arbitrária $\Rightarrow \overline{u}$ é solução ótima.

 (\Rightarrow) máximo $(v \in V)$ $\nabla h(\overline{u}, v) > 0 \Rightarrow \overline{u}$ não é solução ótima.

Note que o máximo do problema (1.5) é maior ou igual a zero, desde que v=0 é factivel.

Por hipótese existe um subgradiente $\gamma \in \partial h(\overline{u})$ e uma direção v tal que:

$$\nabla h(\overline{u},v) = \lim_{\theta \to 0^+} \frac{h(\overline{u} + \theta v) - h(\overline{u})}{\theta} > 0$$

Isto implica que $h(\overline{u}+\theta v)>h(\overline{u})$ (para algum $\theta>0$ e $\overline{u}+\theta v\in\mathbf{R}^m$), logo \overline{u} não é solução ótima.

Teorema 1.2.10 A solução \overline{u} é ótima para (DL) se, e somente se, $0 \in \partial h(\overline{u})$.

demonstração:

O vetor $0 \in \partial h(\overline{u})$ se, e somente se, $h(u) - h(\overline{u}) \leq 0(u - \overline{u})$ para todo $u \in \mathbf{R}_+^{\mathbf{m}}$. Logo $h(u) \leq h(\overline{u})$ é satisfeito se, e somente se, \overline{u} é ponto de ótimo de h.

1.2.2 Dualidade em Programação Linear

Considere o seguinte problema de programação linear:

minimizar
$$c^T x$$
 sujeito a $\left\{egin{array}{l} Ax \geq b \ x \geq 0 \end{array}
ight.$

onde:

 $c, x \in \mathbb{R}^n$. $b \in \mathbb{R}^m$. $A \in \mathbb{R}^{m \times n}$ tal que rank $(A) = m \in m \le n$.

Assim a função lagrangeana (cf. definição 1.2.1) é escrita da seguinte forma:

$$L(x,u)=c^Tx+u^T(b-Ax),\ u\geq 0$$

E a função dual:

$$h(\mathbf{u}) = \min_{\mathbf{z} \in \mathbf{u}} \mathbf{z} \mathbf{a}_{(\mathbf{z} > \mathbf{0})} (c - A^T u)^T x + u^T b$$

O mínimo existe se, e somente se, $(c - A^T u) \ge 0$, isto é, $A^T u \le c$, caso contrário, se $(c - A^T u)_i < 0$ para algum i, poderiamos tomar x_i arbitrariamente grande, assim $(c - A^T u)^T x$ seria arbitrariamente negativo. Para u satisfazendo $A^T u \le c$, o valor mínimo é conseguido quando $(c - A^T u)^T x = 0$ logo $h(u) = u^T b$.

Assim o dual do problema linear dado, fica da seguinte forma:

maximizar
$$h(u) = u^T b$$
 sujeito a $\left\{egin{array}{l} A^T u \leq c \\ u \geq 0 \end{array}
ight.$

Capítulo 2

Métodos de Decomposição

2.1 Introdução

O objetivo principal deste capítulo é apresentar alguns dos mais importantes métodos de otimização para problemas de programação linear inteira mista. Nos últimos anos, a complexidade cada vez maior, dos problemas tratados pela programação matemática, estimulou o aparecimento tanto de técnicas preocupadas com o aumento da eficiência computacional dos métodos tradicionais, bem como de novos métodos onde a tônica principal tem sido a decomposição.

Os métodos de decomposição (primal ou dual) podem ser vistos como uma manipulação que visa transformar o problema original em um problema equivalente (problema mestre), onde evidenciam-se simplificações advindas da estrutura original do problema e de uma estratégia de resolução, que tirando proveito destas simplificações, fornecem um procedimento iterativo para solução do problema mestre. A decomposição em subproblemas menores, dessa forma é passível de solução por técnicas conhecidas.

2.2 Método de decomposição de Benders

O algoritmo de decomposição de Benders procura explorar a estrutura primal do problema de programação linear inteira mista com "variáveis complicadas" (em nosso estudo variáveis inteiras). Para esta classe de problemas, estudada por Benders [5], fixamos temporariamente o valor destas variáveis

complicadas, reduzindo à um problema de programação linear (p.p.l.), que pode ser resolvido utilizando algum código Simplex. O problema mestre escrito nas "variáveis complicadas" pode ser resolvido adicionando iterativamente planos de cortes.

Considere o seguinte problema de programação linear inteira mis-

ta:

Primal (P)

minimizar_{((x₁,x₂) ∈ S)}
$$c^1x_1 + c^2x_2$$

sujeito a
$$\begin{cases} A^1x_1 + A^2x_2 \ge b \\ S = \{ x^T = (x_1, x_2) \ tal \ que \ x_1 \in \mathbf{R}_+^{\mathbf{n}_1}, x_2 \in \mathbf{Z}_+^{\mathbf{n}_2} \} \end{cases}$$

Fixando a variável x_2 , obtemos o seguinte p.p.l.:

Problema (
$$P_1$$
)

minimizar $c^1x_1+c^2x_2$

sujeito a $\left\{egin{array}{l} A^1x_1\geq b-A^2x_2 \ x_1\geq 0 \end{array}
ight.$

Mas para aplicar a técnica acima, a variável x_2 não pode ser escolhida arbitrariamente, uma vez que o problema (P_1) deve ser não vazio. Para expressar esta condição usaremos o Lema de Farkas-Minkowski (ver Lasdon [15] página 372).

Aplicando o Lema de Farkas-Minkowiski, o sistema linear

$$\begin{cases} A^1 x_1 \ge b - A^2 x_2 \\ x_1 \ge 0 \end{cases}$$

será factível se, e somente se,

$$(b - A^2 x_2)^T v \le 0 \tag{2.1}$$

para todo v tal que

$$\begin{cases} v^T A^1 \le 0 \\ v \ge 0 \end{cases}$$

Note que o cone $C = \{v \ tal \ que \ v^T A^1 \le 0, v \ge 0\}$ é um poliedro que não depende de x_2 e existe um número finito de direções extremais (raios extremos) que geram o cone, isto é, existem vetores v^j , $j = 1, \ldots, n_p$, tal que para qualquer vetor $u \in C$, pode ser escrito como uma combinação linear não negativa das direções extremas. Logo:

$$v = \sum_{j=1}^{n_p} \lambda_j v^j, \tag{2.2}$$

$$\lambda_j \ge 0, \tag{2.3}$$

Substituindo (2.2) e (2.3) em (2.1) obtemos:

$$\sum_{j=1}^{n_p} \lambda_j (b - A^2 x_2)^T v^j \leq 0,$$

O qual é verdadeiro para todo $\lambda_j \geq 0$ se, e somente se,

$$(b-A^2x_2)^Tv^j \leq 0, \quad j=1,\ldots,n_p.$$
 (2.4)

Seja J_{PA} o conjunto de vetores $x_2 \in \mathbf{Z}_+^{\mathbf{n}_2}$, que satisfaz (2.4) (necessariamente $J_{PA} \neq \phi$ se (P) tiver solução). O problema (P) é então equivalente a:

 $\text{minimizar}_{(\mathbf{x_2} \in \mathbf{J_{PA}})} \ \{ \ c^2x_2 + \ \text{minimo}_{(x_1 \ge 0)} \ c^1x_1 \ tal \ que \ A^1x_1 \ge b - A^2x_2 \ \}$

Isto corresponde à noção inicial de fixar a variável x_2 , resolver o problema linear (P_1) e escolher um melhor valor para variável x_2 .

Para x_2 fixo, a minimização interna acima é um p.p.l.

Subproblema Primal (SP x_2)

minimizar $(x_1 \ge 0)$ c^1x_1 sujeito a $A^1x_1 \ge b - A^2x_2$

Seja u o vetor das variáveis duais associadas à restrição de (SPx_2) , logo o dual é definido da seguinte forma:

Dual do Subproblema Primal (DSP x_2)

maximizar $(u \ge 0)$ $(b - A^2x_2)^Tu$ sujeito a $u^TA^1 \le c^1$

Considere agora o seguinte conjunto de soluções duais factíveis $F(DSPx_2) = \{ u \ tal \ que \ u^TA^1 \le c^1, u \ge 0 \}$, este conjunto não depende de x_2 e v^j ($j = 1, ..., n_p$), como definidos anteriormente, são seus raios extremos.

Se o conjunto $F(DSPx_2)$ for vazio, temos pelo teorema de dualidade (cf. capítulo 1, corolário 1.2.1) que:

(a) (SPx_2) não tem solução

ou

(b) (SPx2) é ilimitado

Mas por definição $x_2 \in J_{PA}$ então (SP x_2) tem solução e conseqüentemente, se $F(DSPx_2)$ for vazio então (SP x_2) seria ilimitado para $x_2 \in J_{PA}$. Portanto o problema (P) será ilimitado.

O máximo é tomado como $(-\infty)$ se o dual for infactível e usando o teorema de dualidade podemos escrever (P) como sendo:

$$\min \operatorname{minimizar}_{(\mathbf{x_2} \in \mathbf{J_{PA}})} \ \{ \ c^2x_2 + \operatorname{maximizar} \{ \ (b - A^2x_2)^Tu \ tal \ que \ u^TA^1 \leq c^1, \ u \geq 0 \ \} \ \}$$

O máximo de (DSP x_2) não pode ser ilimitado para $x_2 \in J_{PA}$ (de fato, se (DSP x_2) for ilimitado então (SP x_2) não terá solução, o que contradiz o fato que $x_2 \in J_{PA}$).

Se (P) for não vazio, a maximização interna pode ser tomada em um ponto extremo (vértice) de (DSP x_2), como existe um número finito de vértices u^k , $k = 1, \ldots, n_q$. Podemos escrever o problema (P) da seguinte forma:

$$\min \mathbf{zar}_{(\mathbf{x_2} \in \mathbf{J_{PA}})} \{ \ c^2x_2 + \mathbf{maximizar}_{(\mathbf{1} \le \mathbf{k} \le \mathbf{n_q})} \{ \ (b - A^2x_2)^T u^k, \ u^k \ge 0 \} \ \}$$

O qual é equivalente a:

minimizar
$$w$$
sujeito a $\begin{cases} w \geq c^2x_2 + (b - A^2x_2)^Tu^k, \ u^k \in F(DSPx_2) \\ x_2 \in J_{PA} \end{cases}$

Substituindo J_{PA} , pela restrições que o definem, obtemos o seguinte problema:

Problema Mestre (PM)

minimizar
$$_{(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})} w$$

sujeito a
$$\begin{cases} w \geq c^2 x_2 + (b - A^2 x_2)^T u^k, & u^k \in F(DSPx_2) \\ (b - A^2 x_2)^T v^j \leq 0, & v^j \in J_{PA} \end{cases}$$

O número de restrições do problema (PM) é igual ao número de vértices e de raios extremos de $(DSPx_2)$, que em geral é enorme mesmo para um

problema de dimensão moderada. Entretanto somente uma pequena parte das restrições de (PM) serão ativas na solução ótima, mais precisamente, o número de restrições ativas de um problema nunca excede a $n_2 + 1$, onde n_2 é o número de variáveis de x_2 , se a solução ótima relaxada, $x_2 \ge 0$, for não degenerada.

Isto sugere um problema mestre relaxado (PMR), formado através de subconjuntos $T_P \subseteq F(DSPx_2)$ e $J_P \subseteq J_{PA}$ de restrições de (PM).

Problema Mestre Relaxado (PMR)

minimizar
$$(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})^w$$

sujeito a $\left\{ \begin{array}{l} w \geq c^2 x_2 + (b - A^2 x_2)^T u^k, & (para\ todo\ u^k \in T_P) \\ (b - A^2 x_2)^T v^j \leq 0, & (para\ todo\ v^j \in J_P) \end{array} \right.$

Seja $(\overline{w}, \overline{x}_2)$ solução de (PMR) (se (PMR) não tem solução factível, então (P) também não tem solução factível), desde que (PMR) é uma relaxação de (PM), \overline{w} é um limitante inferior de w^* (o valor ótimo de (PM) e consequentemente de (P)).

Uma condição necessária e suficiente para que $(\overline{w}, \overline{x}_2)$ seja uma solução ótima de (PM) e consequentemente de (P) é que o ponto $(\overline{w}, \overline{x}_2)$ satisfaça todas as restrições de (PM) ou seja:

$$\begin{cases} w \ge c^2 x_2 + (b - A^2 x_2)^T u^k, & (u^k \in F(DSPx_2)) \\ (b - A^2 x_2)^T v^j \le 0, & (v^j \in J_{PA}) \end{cases}$$

Caso contrário, devemos encontrar uma restrição que seja "violada" (de preferência a mais "violada").

O ideal seria calcular:

$$\text{máximo}_{(1 \le \mathbf{j} \le \mathbf{n}_{\mathbf{p}})} (b - A^2 x_2)^T v^j$$

Embora não tenhamos todos os vértices disponíveis podemos resolver o problema acima através de $(DSPx_2)$.

Se excluirmos o caso onde $(DSP\overline{x}_2)$ não tem solução (caso contrário o problema (P) não teria solução ou é ilimitado), o problema pode ser dividido em três casos:

1. O valor de (DSP \overline{x}_2) é ilimitado. O método Simplex aplicado à (DSP \overline{x}_2) fornece um raio extremo $\overline{v} \in J_{PA}$ tal que:

$$\begin{cases} (b - A^2 \overline{x}_2)^T \overline{v} > 0 \\ (A^1)^T \overline{v} \le c^1 \end{cases}$$

Assim a restrição $(b - A^2\overline{x}_2)^T\overline{v} \leq 0$ não é satisfeita por \overline{x}_2 logo $(\overline{w}, \overline{x}_2)$ não é solução de (PM). Então a restrição $(b - A^2\overline{x}_2)^T\overline{v} \leq 0$ é adicionada à (PMR).

2. O ótimo de $(DSP\overline{x}_2)$ possui um valor finito (consequentemente é obtido um vértice $\overline{u} \in F(DSPx_2)$) tal que:

$$c^2\overline{x}_2 + (b - A^2\overline{x}_2)^T\overline{u} = \overline{w}$$

Desde que \overline{u} é um vértice de $F(DSPx_2)$ que maximiza $(b-A^2\overline{x}_2)^Tu$, obtemos:

$$c^2\overline{x}_2 + (b - A^2\overline{x}_2)^T u^k \leq \overline{w}$$
 para todo $u^k \in F(DSPx_2)$.

Além disso temos que $(b - A^2 \overline{x}_2)^T v^k \leq 0$ para todos os raios extremos v^k de J_{PA} . De fato, se tivessemos $(b - A^2 \overline{x}_2)^T v^k > 0$ para algum v^k , então a solução ótima de $(DSP\overline{x}_2)$ seria ilimitada e o caso 1 seria identificado.

Segue que $(\overline{w}, \overline{x}_2)$ é solução ótima de (PM) e o algoritmo termina.

3. A solução ótima de $(DSP\overline{x}_2)$ é limitada e obtemos um vértice $\overline{u} \in F(DSPx_2)$, mas ao contrário do caso 2, obtemos:

$$c^2\overline{x}_2 + (b - A^2\overline{x}_2)^T\overline{u} > \overline{w}$$

Logo a restrição $c^2x_2 + (b - A^2x_2)^T \overline{u} \le w$ é acrescentada em (PMR).

Enquanto o teste de otimalidade não for satisfeito (caso 2) podemos adicionar ao programa mestre relaxado, as restrições que não satisfazem a solução $(\overline{w}, \overline{x}_2)$. Assim um novo programa relaxado é resolvido para (w, x_2) .

Se em algum estágio, a solução ótima do problema mestre relaxado não existir, então o problema (P) será infactível e o algoritmo termina.

Se
$$F(PMR) = \phi$$
 então $F(PM) = \phi$.

Se $v(PMR) = \infty$ nada podemos concluir sobre v(PM).

Finalmente, se $F(DSPx_2)$ for vazio e (P) tiver uma solução factível concluimos que (P) será ilimitado e o algoritmo termina.

2.2.1 Algoritmo

- 1. Inicialização
 - 1.1. $w := (-\infty);$
 - 1.2. $x_2 := \text{qualquer elemento de S};$
 - 1.3. $\varepsilon \geq 0$; (tolerância);
- 2. Resolva o problema (DSP x_2);
 - 2.1. Se (DSPx2) é Ilimitado Então
 - 2.1.1. Gerar corte de Benders
 - 2.1.1.1. Determinar um raio extremo \overline{v} ;
 - 2.1.1.2. Acrescentar a restrição em (PMR) 2.1.1.2.1. $(b - A^2x_2)^T \overline{v} \leq 0$;
 - 2.2. Se $(DSPx_2)$ é Infactível Então
 - 2.2.1. Fim. O problema (P) é ilimitado.
 - 2.3. Se $|w c^2x_2 (b A^2x_2)^Tu| \le \varepsilon$ Então
 - 2.3.1. Fim. Solução ótima;

Senão Gerar corte de Benders

2.3.2. Acrescentar a restrição em (PMR)
$$3.3.2.1. c^2x_2 + (b - A^2x_2)^Tu \le w;$$

- 3. Resolva o problema (PMR);
 - 3.1. <u>Se</u> (PMR) é Infactível <u>Então</u>

3.1.1. Fim. O problema (P) é infactível;

3.2. Se (PMR) é Ilimitado Então

3.2.1.
$$w := (-\infty);$$

3.2.2. $x_2 := \text{qualquer elemento de S}$

<u>Senão</u>

3.2.3. Obtemos uma solução ótima finita (w, x_2) ;

4. Volte a passo 2;

Teorema 2.2.1 O algoritmo converge em um número finito de iterações.

A demonstração pode ser encontrada em Salkin [24] página 264.

Para problemas de grande porte, a convergência pode ser muito lenta, contudo o erro de arredondamento resultante da regra de parada (caso 2) pode não ser satisfeita. Mas em cada iteração, a solução ótima do problema mestre relaxado (PMR) é um limitante inferior de w^* , além disso, se em um dado estágio o subproblema (SP x_2) tiver uma solução ótima, então $(\overline{x}_1, \overline{x}_2)$ é solução de (P) e $\underline{w} = c^2 \overline{x}_2 + c^1 \overline{x}_1$ é um limitante superior de w^* . Logo o algoritmo pode ser interrompido quando $|\overline{w} - \underline{w}| \le \varepsilon$ (tolerância).

2.2.2 Considerações Finais

- i- A velocidade de convergência do algoritmo está intimamente ligada com o código Simplex utilizado para resolver o problema $(DSPx_2)$ e principalmente com o método utilizado para a resolução do problema linear inteiro (PMR). No problema de otimização em redes, Magnanti et al. [16] observaram que o algoritmo de Benders converge muito lentamente, pois requer a resolução de um número muito grande de problemas inteiros.
- ii- Assim a grande dificuldade do algoritmo está justamente na resolução do problema inteiro, pois a cada iteração novos cortes (restrições) são adicionadas ao problema mestre relaxado. Côté et al. [6] propõem algumas heurísticas, para a resolução aproximada do problema mestre relaxado, com objetivo de acelerar a busca, contudo os métodos propostos são heurísticos, assim não existe uma garantia que uma solução factível possa ser determinada, ainda que ela exista.

2.3 Relaxação Lagrangeana

2.3.1 Introdução

Considere o seguinte problema de programação linear mista:

$$\begin{array}{c} \text{Primal (P)} \\ \\ \text{minimizar } c^Tx \\ \\ \text{sujeito a} \left\{ \begin{array}{l} A_2x \geq b_2 \\ \\ x \in X = \{ \ x = (x_1, x_2) \ \text{tal que } A_1x \geq b_1, x_1 \in \mathbf{R}^{\mathbf{n}_1}_+, x_2 \in \mathbf{Z}^{\mathbf{n}_2}_+ \} \end{array} \right. \end{array}$$

A idéia básica da Relaxação Lagrangeana é de relaxar um conjunto de restrições "complicadas", obtendo-se assim um problema que seja mais "fácil" de se resolver.

Para cada restrição "complicada" associamos uma variável u_i ($u_i \ge 0$) chamada de multiplicador de Lagrange ou variável dual. Denotemos por u o vetor formado pelos multiplicadores de Lagrange.

Assim podemos definir a função L(x,u) (cf. capítulo 1, definição 1.2.1) da seguinte forma:

$$L(x, u) = c^T x + u^T (b_2 - A_2 x) = u^T b_2 + (c^T - u^T A_2) x$$

E a função dual é dada por:

$$h(\mathbf{u}) = \min_{\mathbf{x} \in \mathbf{X}} \{ L(x, u) \} = u^T b_2 + \min_{\mathbf{x} \in \mathbf{X}} \{ (c^T - u^T A_2) x \}$$

O problema h(u) é chamado de relaxação Lagrangeana de (P) em relação ao conjunto de restrições $A_2x \geq b_2$. Podemos observar que h(u) não contém as restrições "complicadas" $A_2x \geq b_2$ uma vez que estas foram adicionadas na função objetivo com um termo de penalização, logo obtemos um problema "mais fácil" de se resolver.

Desse modo, definimos o seguinte problema (cf. capítulo 1 seção 1.2):

$$\mathbf{Z_{DL}} \; (\mathbf{u^*}) = \mathbf{maximizar_{(\mathbf{u} \geq \mathbf{0})}} \; h(u)$$

A solução de (DL) nos fornece uma escolha mais apropriada da "variável dual u", consequentemente sua resolução é de grande importância para o nosso estudo.

2.3.2 Método do Subgradiente

Apresentaremos uma aproximação bem simples, do algoritmo de passo ascendente, para a resolução de (DL) chamado de "algoritmo do subgradiente", que foi proposto inicialmente por Shor (1962) que é uma generalização do método do gradiente para funções diferenciáveis. Posteriormente foi estudada por Poljak [20], mas a idéia central é baseado nos trabalhos de Agmon [1], Motzkin e Schoenberg [18].

Basicamente o algoritmo procura gerar uma sequência de soluções $\{u^k\}$ através da seguinte regra:

$$u^{k+1} = Proj_{U} \left\{ u^{k} + \frac{\theta_{k} \gamma^{k}}{\|\gamma^{k}\|} \right\}$$
 (2.5)

onde:

 $\begin{array}{lll} \mathbf{U} & = & \{ u \text{ tal que } u \in \mathbf{R}_+^{\mathbf{m}} \}. \\ Proj_{\mathbf{U}} & = & \text{\'e o operador proje\'ção de } \mathbf{R}^{\mathbf{m}} \text{ em U.} \\ \theta_k & = & \text{escalar maior que zero.} \\ \gamma^k & = & (b_2 - A_2 x^k) \text{\'e um subgradiente de h em (} u^k \text{).} \end{array}$

Observe que não temos nenhuma garantia que γ^k seja uma direção de subida, dessa forma a convergência do algoritmo depende essencialmente da

escolha do tamanho do passo θ_k em cada iteração. Existem muitos resultados teóricos que determinam o tamanho do passo a ser usado, mas um resultado mais geral é dado pelo seguinte teorema:

Teorema 2.3.1 (Poljak [20]) Seja h uma função côncava e contínua definida em um conjunto compacto $X \subseteq \mathbf{R^{n+m}}$. Se θ_k ($\theta_k > 0$) é escolhido de tal modo que:

$$\lim_{k\to\infty}\theta_k=0$$

$$\sum_{k=1}^{\infty} \theta_k = \infty$$

então a seqüência $\{u^k\}$ com $\gamma^k \in \partial h(u^k)$ é tal que:

$$\lim_{k\to\infty} \sup h(u^k) = h(u^*) = \operatorname{maximizar}_{(\mathbf{u}\in\mathbb{R}^{\mathbf{m}}_+)} \{ h(u) \}$$

demonstração:

É suficiente mostrar que dado $\varepsilon > 0$, $\exists k \text{ tal que } u^k \in \mathcal{V}(\epsilon) = \{ u \geq 0 \text{ tal que } h(u) > h(u^*) - \varepsilon \}.$

Suponhamos por absurdo que $u^k \notin \mathcal{V}(\varepsilon)$, para todo k.

Em outras palavras, $h(u^k) \leq h(u^*) - \varepsilon$, para todo u^k .

Como h é contínua em $X \Rightarrow \exists \delta > 0$ tal que:

$$||u - u^*|| \le \delta \implies h(u) > h(u^*) - \varepsilon$$

Temos $\gamma^k \neq 0$, para qualquer k (caso contrário u^k seria um ponto de ótimo e o algoritmo termina).

Seja \underline{u}^k definido da seguinte forma:

$$\underline{u}^k = u^* - \delta \frac{\gamma^k}{\|\gamma^k\|} \tag{2.6}$$

Temos que $\|\underline{u}^k - u^*\| = \delta$ e portanto $h(\underline{u}^k) > h(u^*) - \varepsilon$. Usando o fato que $h(u^k) \le h(u^*) - \varepsilon \Rightarrow h(u^k) - h(\underline{u}^k) < 0$ Como γ^k é o subgradiente da função h no ponto u^k temos:

$$h(\underline{u}^k) - h(u^k) \le (\gamma^k)^T (\underline{u}^k - u^k)$$
 para todo k

Logo:

$$(\gamma^k)^T(u^k - \underline{u}^k) \le h(u^k) - h(\underline{u}^k) < 0$$
 (2.7)

$$Como \|u^{k+1} - u^*\|^2 = \|Proj_U\{ u^k + \frac{\theta_k \gamma^k}{\|\gamma^k\|} \} - Proj_U\{ u^* \}\|^2$$

e usando o fato que o operador projeção é uma contração (ver Minoux [17]) obtemos:

$$||u^{k+1} - u^*||^2 \le ||u^k + \frac{\theta_k \gamma^k}{||\gamma^k||} - u^*||^2$$

$$||u^{k+1} - u^*||^2 \le ||u^k - u^*||^2 + \theta_k^2 + \frac{2\theta_k(\gamma^k)^T (u^k - u^*)}{||\gamma^k||}$$
(2.8)

De (2.6) temos que:

$$(\gamma^k)^T(u^k-u^*)=(\gamma^k)^T(u^k-\underline{u}^k-\frac{\delta\gamma^k}{\|\gamma^k\|})$$

$$(\gamma^{k})^{T}(u^{k} - u^{*}) = (\gamma^{k})^{T}(u^{k} - \underline{u}^{k}) - \delta \|\gamma^{k}\|$$
(2.9)

Substituindo (2.9) em (2.8) obtemos:

$$||u^{k+1} - u^*||^2 \le ||u^k - u^*||^2 + \theta_k^2 + \frac{2\theta_k(\gamma^k)^T(u^k - \underline{u}^k)}{||\gamma^k||} - 2\delta\theta_k$$

De (2.7) temos que $(\gamma^k)^T(u^k-\underline{u}^k)<0$ e como $\theta_k>0$ (por hipótese) então:

$$||u^{k+1} - u^*||^2 \le ||u^k - u^*||^2 + \theta_k(\theta_k - 2\delta)$$

Como $\{\theta_k\}$ é uma seqüência que tende para zero quando k tende para infinito então $\exists K \in N$ tal que $k \geq K \Rightarrow \theta_k \leq \delta$.

Assim para todo $k \geq K$

$$||u^{k+1} - u^*||^2 \le ||u^k - u^*||^2 - \theta_k \delta$$

Somando sobre k, k = K, ..., K + p obtemos:

$$\delta \sum_{k=K}^{K+p} \theta_k \le \|u^K - u^*\|^2 - \|u^{K+p+1} - u^*\|^2 \le \|u^K - u^*\|^2$$

Tomando p suficientemente grande, obtemos uma contradição do fato que $\sum_{k=1}^{\infty} \theta_k = \infty$.

Evidentemente o algoritmo baseado no procedimento acima para a determinação do passo, não possui nenhum interesse prático, pois a convergência torna-se excessivamente lenta. Uma outra estratégia para a escolha do passo foi proposta por Poljak [21] onde θ_k é definido da seguinte forma:

$$\theta_k = \frac{\lambda_k \{ h(u^*) - h(u^k) \}}{\|b_2 - A_2 x^k\|}$$
 (2.10)

onde:

$$0 < \varepsilon_1 \le \lambda_k \le 2 - \varepsilon_2$$
 ($\varepsilon_1 > 0$, $\varepsilon_2 > 0$). $h(u^*)$ é o valor ótimo do problema (DL).

Usando o resultado acima, obtemos o seguinte algoritmo do subgradiente:

2.3.3 Algoritmo

- 1. Inicialização
 - 1.1. $\varepsilon \geq 0$; (tolerância)
 - 1.2. k := 1;
 - 1.3. $u^k := 0$;
 - 1.4 max > 0 (número máximo de iterações)
- 2. Selecione algum subgradiente $\gamma^k \in \partial h(u^k)$
 - 2.1. Se $\|\gamma^k\| \le \varepsilon$ Então

2.1.1. Fim.

<u>Senão</u>

2.1.2.
$$u^{k+1} := Proj_U \{ u^k + \theta_k \gamma^k \};$$

- 3. Teste de Parada
 - 3.1. Se $\max = k$ Então

3.1.1. Fim. (Solução aproximada)

Senão

$$3.1.2. k := k + 1;$$

4. Retorne ao passo 2

onde:

h(u*) é o valor do ponto ótimo no problema (DL).

$$\theta_k = \frac{\lambda_k \{ h(u^*) - h(u^k) \}}{\|\gamma^k\|^2}$$

 $\gamma^k = (b_2 - A_2 x^k)$ é o subgradiente de $h(u^k)$.

$$0 < \varepsilon_1 \le \lambda_k \le 2 - \varepsilon_2 \ (\ \varepsilon_1 > 0 \ , \ \varepsilon_2 > 0 \)$$

2.3.4 Convergência

A convergência do algoritmo (2.3.3) é garantida pelo seguinte teorema:

Teorema 2.3.2 Suponha que a solução ótima u^* do problema (DL) exista e ao aplicarmos o algoritmo do subgradiente com a hipótese adicional de que existe um M>0 tal que $\|\gamma^k\|^2 \leq M$, para todo $\gamma^k \in \partial h(u)$ e para todo u pertencente ao conjunto $\{u \text{ tal que } \|u-u^*\| \leq \|u^0-u^*\| \}$, então o $\lim_{k\to\infty} h(u^k) = h(u^*)$ e algum ponto limite da seqüência $\{u^k\}$ é solução ótima.

demonstração:

Inicialmente iremos mostrar que a seqüência $\{\|u-u^*\|^2\}$ é monotonicamente decrescente.

Usando o fato que $u^{k+1} = Proj_U\{u^k + \theta_k \gamma^k\}$ e que o operador projeção é uma contração temos que:

$$||u^{k+1} - u^*||^2 = ||Proj_U\{u^k + \theta_k \gamma^k\} - Proj_U\{u^*\}||^2$$

$$\leq ||u^k + \theta_k \gamma^k - u^*||^2$$

$$= ||u^k - u^*||^2 - \theta_k^2 ||\gamma^k||^2 + 2\theta_k (u^k - u^*)^T \gamma^k$$

Como $h(u^*) - h(u^k) \le (\gamma^k)^T (u^* - u^k)$ temos que:

$$||u^{k+1} - u^*||^2 \le ||u^k - u^*||^2 - \theta_k^2 ||\gamma^k||^2 - 2\theta_k [h(u^*) - h(u^k)]$$

Desde que:

$$\theta_k = \frac{\lambda_k \{ h(u^*) - h(u^k) \}}{\|\gamma^k\|^2}.$$

Temos que:

$$||u^{k+1} - u^*||^2 \le ||u^k - u^*||^2 + \frac{(\lambda_k^2 - 2\lambda_k)[|h(u^*) - h(u^0)|]^2}{||\gamma^k||^2}$$
(2.11)

$$Seja \Delta = \frac{[h(u^*) - h(u^k)]^2}{\|\gamma^k\|^2} \ge 0$$

Logo:

$$(\lambda_k^2 - 2\lambda_k)\Delta = \lambda_k(\lambda_k - 2)\Delta \le -\varepsilon_1\varepsilon_2\Delta \tag{2.12}$$

Desde que $0 < \varepsilon_1 \le \lambda_k \le 2 - \varepsilon_2$ ($\varepsilon_1 > 0$, $\varepsilon_2 > 0$) (por hipótese). Temos por (2.11) e (2.12) que:

$$||u^{k+1} - u^*||^2 \le ||u^k - u^*||^2 - \frac{(\varepsilon_1 \varepsilon_2)[|h(u^*) - h(u^k)|]^2}{||\gamma^k||^2}$$
(2.13)

Um resultado imediato de (2.13) é que a seqüência $\{\|u^k - u^*\|^2\}$ é monotonicamente decrescente, isto implica que:

$$\lim_{k\to\infty} ||u^k - u^*||^2$$
 existe

A existência do limite implica de (2.13) que:

$$\lim_{k\to\infty} \frac{[h(u^*) - h(u^k)]^2}{\|\gamma^k\|^2} = 0$$

Caso contrário teriamos uma subseqüência infinita de $\{\|u^{k+1} - u^*\|^2\}$ decrescendo em cada passo para algum $\varepsilon_3 \geq 0$, o que é um absurdo.

Desde que $\|\gamma^k\|^2 < M$ para todo k podemos concluir que:

$$\lim_{k\to\infty}h(u^k)=h(u^*).$$

Finalmente a seqüência $\{u^k\}$ possui no mínimo um subseqüência convergente, pois u^k é limitado pelo conjunto $\{u \text{ tal que } ||u-u^*|| \le ||u^0-u^*|| \}$.

Seja $\{u^{k_i}\}$ uma subseqüência que converge para u^{**} e desde que a função h é continua (cf. capítulo 1 lema 1.2.2) temos que:

$$\lim_{i\to\infty}h(u^{k_i})=h(\lim_{i\to\infty}u^{k_i})=h(u^{**})=h(u^*)$$

E assim u** é ótimo.

Como desconhecemos o valor ótimo do problema, utiliza-se uma estimativa \overline{w} de $h(u^*)$ e o passo é escolhido da seguinte forma:

$$\theta_k = \frac{\lambda_k \{ \overline{w} - h(u^k) \}}{\|\gamma^k\|^2} \tag{2.14}$$

$$0 < \varepsilon_1 \le \lambda_k \le 2 - \varepsilon_2 \tag{2.15}$$

Usando (2.14) e (2.15) Held et al. [12] sugerem a seguinte regra heurística.

- 1. Faça $\lambda_0 = 2$ durante $n_0 = n$ iterações (onde n é o número de variáveis do problema).
- 2. Para i = 1 até r (número máximo de iterações) faça $\lambda_i = \lambda_{i-1}/2$ durante $n_i = n_{i-1}/2$ iterações.

Um outro resultado pode ser encontrado em Allen et al. [2] onde a restrição $0 < \varepsilon_1 \le \lambda_k \le 2 - \varepsilon_2$ ($\varepsilon_1 > 0$, $\varepsilon_2 > 0$) é relaxada permitindo que a seqüência $\{\lambda_k\}$ aproxime-se de zero.

2.3.5 Considerações Finais

- i- A grande vantagem do algoritmo do subgradiente, reside na sua simplicidade de implementação e no conhecimento de somente um gradiente em cada ponto, mas com o inconveniente de que a taxa de convergência ser geralmente lenta.
- ii- A escolha do tamanho do passo é uma decisão crucial. Este problema foi estudado inicialmente por Poljak [20,21] e posteriormente por Held et al. [12] e Bazaraa e Sherali [2] que fornecem algumas regras heurísticas para a solução deste problema. Contudo estas escolhas sobre o tamanho do passo são de natureza puramente heurísticas. Assim diferentes variações do tamanho do passo tem diferentes variações na convergência do algoritmo.
- iii- Os criterios de parada são completamente arbitrários, dado que não possuimos critérios que assegurem que o ponto encontrado seja de ótimo. Por exemplo quando a função h(u) é "suave" e $\{u^k\}$ converge para o ponto de máximo u^* de h então $\{\gamma^k\} \to \gamma^* = 0$, isto é, o gradiente converge mas isto

pode não ocorrer em algumas funções não diferenciáveis, assim critérios do tipo $\|\gamma\| \le \varepsilon$ podem se tornar inuteis. Desse modo um dos critérios mais utilizados é o de delimitar o número de iterações.

iv- O algoritmo do subgradiente tem sido considerado apenas em um esquema heurístico dado que seus limitantes são de natureza especulativa, pelo fato do algoritmo não dar indicação sobre o "gap" de otimalidade. Sen e Sherali [25] propõem uma nova estratégia de resolução usando o algoritmo do subgradiente primal - dual. Desse modo conhecendo o valor do "gap dual", podemos estabelecer um critério de parada mais confiável.

v- Podemos observar que o conhecimento "a priori" do valor ótimo (ou uma estimativa) é de grande importância para o desenvolvimento do algoritmo, Kim et al. [16] propõem uma nova estratégia onde este dado não é conhecido.

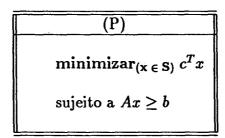
Capítulo 3

Decomposição Híbrida

3.1 Introdução

Os princípios de decomposição apresentados neste trabalho, tem fornecido muitos algoritmos eficientes para a resolução de problemas de programação linear inteira mista. Os algoritmos de decomposição de Benders e o método relaxação Lagrangeana procuram explorar da estrutura do problema (seja nas variáveis ou restrições), para resolver uma seqüência de subproblemas "fáceis". Contudo muitos problemas apresentam um estrutura especial tanto para as variáveis como para as restrições. O problema de planejamento de multi-item [27] é um exemplo, dado que o subproblema resultante da relaxação Lagrangeana é reduzido à resolução de vários problemas planejamento de um único item, bem como o subproblema de Benders resulta na resolução de um problema de fluxos em redes a custo mínimo. Assim a idéia básica da decomposição híbrida [28] é usar ambos os principios de decomposição em um único procedimento.

Considere o seguinte problema de programação linear mista (cf. definição do capítulo 1):



Explicitando as variáveis "complicadas", aqui representadas por $x_2 \in \mathbf{Z}_+^{n_2}$ e utilizando o método de decomposição de Benders (ou primal), o problema (P) é reescrito na seguinte forma:

$$(P)$$
minimizar $(\mathbf{x} \in \mathbf{S})$ $c^1x_1 + c^2x_2$
sujeito a $A^1x_1 + A^2x_2 \geq b$

Assumiremos que o conjunto:

 $\{ x_1 \in \mathbf{R}^n_+; tal \ que \ A^1x_1 + A^2x_2 \geq b \}$ é não vazio para todo $x_2 \in \mathbf{Z}^{\mathbf{n}_2}_+.$

Fixando a variável x_2 obtemos o seguinte p.p.l.:

Subproblema Primal (SP
$$x_2$$
)

minimizar $_{(\mathbf{x_1} \geq \mathbf{0})} c^1x_1 + c^2x_2$

sujeito a $A^1x_1 \geq b - A^2x_2$

Dessa forma o problema (P) é equivalente a:

$$\begin{aligned} & \text{Primal } (P_1) = \text{minimizar}_{(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})} \ v(SPx_2) \\ \\ & \text{minimizar}_{(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})} \left\{ \begin{array}{l} \text{minimizar}_{(\mathbf{x_1} \geq \mathbf{0})} \ c^1x_1 + c^2x_2 \\ \\ sujeito \ a \ A^1x_1 + A^2x_2 \geq b \end{array} \right\} \end{aligned}$$

ou

ou

$$\operatorname{minimizar}_{(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})} \operatorname{máximo}_{(\mathbf{k})} \{ (b - A^2 x_2)^T u^k + c^2 x_2, \ tal \ que \ u^k \in T_{PA} \}$$

ou

$$(PM)$$
minimizar
 $(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})^{w}$
sujeito a
$$\begin{cases} b^T u^k + (c^2 - u^k A^2) x_2 \leq w \\ u^k \in T_{PA} \end{cases}$$

Onde T_{PA} é um conjunto de todos os pontos extremos de F(DSP), e (DSP) é o dual do subproblema (SPx_2) . Por simplicidade retiramos x_2 de F(DSP), desde que F(DSP) não depende desta variável. As restrições decorrentes dos elementos de T_{PA} são chamadas de cortes de Benders e geralmente em um número muito grande, desse modo procuramos resolver um problema relaxado (PMR), onde apenas alguns cortes são considerados.

$$(PMR)$$
minimizar $_{(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})} w$
sujeito a
$$\begin{cases} b^T u^k + (c^2 - (u^k)^T A^2) x_2 \leq w \\ u^k \in T_P \subseteq T_{PA} \end{cases}$$

Explicitando as restrições "complicadas", aqui representadas por $A_2x \geq b_2$, o procedimento para a decomposição dual é feito reescrevendo o problema (P) na seguinte forma:

$$(P)$$
 minimizar $_{(\mathbf{x} \in \mathbf{S})} c^T x$ sujeito a $\left\{egin{array}{l} A_1 x \geq b_1 \ A_2 x \geq b_2 \end{array}
ight.$

Dualizando em relação à restrição $A_2x \geq b_2$ obtemos o seguinte problema:

Subproblema Dual (SD u_2)

minimizar $(\mathbf{x} \in \mathbf{S})$ $c^Tx + (b_2 - A_2x)^Tu_2$ sujeito a $A_1x \geq b_1$

Assumiremos que o conjunto:

 $\{ x \in S \ tal \ que \ A_1 x \ge b_1 \ \}$ é não vazio e limitado.

O dual Lagrangeano do problema (P) é dado por:

ou

 $\text{maximizar}_{(u_2 \geq 0)} \{ \text{minimizar}_{(k)} c^T x^k + (b_2 - A_2 x^k)^T u_2, \text{ tal que } x^k \in T_{DA} \}$

ou

$$ext{Dual Mestre }(DM)$$
 $ext{maximizar}_{(\mathbf{u_2} \geq \mathbf{0})} z$ $ext{sujeito a} \left\{egin{array}{l} c^T x^k + (b_2 - A_2 x^k)^T u_2 \geq z \ & x^k \in T_{DA} \end{array}
ight.$

Onde T_{DA} é um conjunto de todos os pontos extremos do casco convexo C_{ℓ} { F(SD) }. Por simplicidade retiramos u_2 de F(SD), desde que F(SD) não depende desta variável. Assim, como no problema de Benders, não resolvemos o problema mestre completo, mas uma relaxação (DMR).

$$egin{aligned} ext{Dual Mestre Relaxado} & (DMR) \ & ext{maximizar}_{(\mathbf{u_2} \geq \mathbf{0})} \ z \ & ext{sujeito a} & \left\{ egin{aligned} c^T x^k + (b_2 - A_2 x^k)^T u_2 \geq z \ & ext{x}^k \in T_D \subseteq T_{DA} \end{aligned}
ight.$$

3.2 Decomposição Híbrida

Uma iteração híbrida para a resolução do problema (P) seria:

- Escolher o valor da variável x_2^k .
- Iteração:
 - 1. Resolver o subproblema primal (SPx_2^k) , e seja u_2^{k+1} as coordenadas da variável dual ótima associada à restrição $A_2x^k \leq b_2$.
 - 2. Resolver o subproblema dual (SDu_2^{k+1}) , e seja x^{k+1} a solução ótima primal.

O valor ótimo do subproblema primal (SPx_2) fornece um limitante superior (cf. capítulo 2 seção 1) e o valor ótimo do subproblema dual (SDu_2)

fornece um limitante inferior (cf. capítulo 2 seção 2) do valor da solução ótima. Desse modo obtemos um procedimento que gera limitantes (inferior e superior) para o valor ótimo do problema (P).

Porém, o algoritmo composto apenas pelos dois subproblemas não possui uma convergência garantida, para assegurarmos a convergência do método, o "ping-pong" (troca de informações entre os dois subproblemas) é imerso em um dos esquemas padrões de decomposição, onde a convergência é comprovada. Quando o "ping-pong" cessa, as soluções previamente obtidas são utilizadas na geração de cortes para o problema mestre primal (ou dual), que então é resolvido, gerando assim uma nova solução que servirá como ponto de partida na continuação do "ping-pong" entre os subproblemas.

3.3 Equivalência entre os Subproblemas

Neste parágrafo vamos mostrar que o "subproblema Lagrangeano" (SDu_2) pode ser utilizado como "problema mestre de Benders" (PMR) e analogamente o "subproblema de Benders" (SPx_2) pode ser utilizado como problema mestre da decomposição dual (DMR). Necessitamos da seguinte definição:

Definição 3.3.1 Sejam (P) e (Q) problemas de programação linear inteira mista. Dizemos que (P) e (Q) são problemas equivalentes em relação a um subconjunto de variáveis (P), neste subconjunto, são soluções ótimas de (P), neste subconjunto, são soluções ótimas de (Q) e vice - versa.

Por exemplo, o problema (P) e o problema mestre (PM) são equivalentes em relação à variável x_2 (devido a construção). Similarmente (DL) e (DM) são equivalentes em relação à variável u_2 . Note que a noção de equivalência envolve também as variáveis primal e dual.

Considere a seguinte partição nas restrições:

$$A_1x \geq b_1, A_2x \geq b_2$$

$$A_1 = \left[\begin{array}{c} A_1^1 \\ A_2^1 \end{array} \right]$$

$$A_2 = \left[\begin{array}{c} A_1^2 \\ A_2^2 \end{array} \right]$$

Logo podemos escrever (DSP) da seguinte forma:

(DSP) maximizar
$$u_1^T(b_1 - A_2^1x_1) + u_2^T(b_2 - A_2^2) + c^2x_2$$

sujeito a
$$\left\{egin{array}{l} u_1^TA_1^1+u_2^TA_1^2 \leq c^1 \ & \ u_1 \geq 0 \ & \ u_2 \geq 0 \end{array}
ight.$$

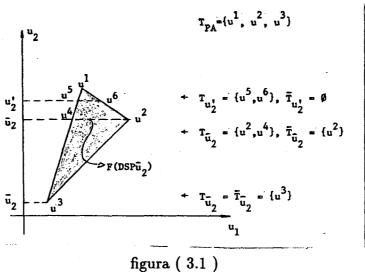
Necessitamos ainda da seguinte notação:

- $F(DSPu_2)$ é a região factível do conjunto de restrições de (DSP), com u_2 fixo.
- Tu_2 é o conjunto de todos os $u^k = (u_1^k, u_2)^T$, tal que u^k é ponto extremo de $F(DSPu_2)$.
- $\overline{T}u_2$ é o conjunto de todos os $u^k = (u_1^k, u_2)^T$, tal que u^k é ponto extremo de F(DSP).
- $(PMRu_2) \equiv (PMR : u^k \in Tu_2).$

$$\overline{T}u_2 \subseteq Tu_2 \; ; \; \overline{T}u_2 \subseteq T_{PA}$$
 (3.1)

$$Tu_2 \subseteq T_{PA}$$
 se, e somente se, $\overline{T}u_2 = Tu_2$ (3.2)

As inclusões (3.1) são ilustradas na figura (3.1).



E a seguinte desigualdade é verdadeira:

$$v(PMR \ tal \ que \ u^k \in \overline{T}u_2) \le v(PMRu_2) \equiv v(PMR \ tal \ que \ u^k \in Tu_2) \le v(PM)$$
 (3.3)

Teorema 3.3.1 Os cortes pertencentes à diferença $Tu_2 - \overline{T}u_2$ não pertencem à (PM).

demonstração:

De fato, seja $u \in Tu_2 - \overline{T}u_2$ e $u \in T_{PA} \Longrightarrow u \in Tu_2 \land u \notin \overline{T}u_2 \Longrightarrow u$ é um ponto não extremo de F(DSP). Desse modo, u pode ser escrito como uma combinação convexa dos pontos extremos de F(DSP).

Logo para cada $u^k \in T_{PA}$, existem $\delta_k \geq 0$ tal que:

$$u = \sum_{k} \delta_k u^k \tag{3.4}$$

$$\sum_{k} \delta_{k} = 1 \tag{3.5}$$

Um corte gerado por este ponto não extremo é dado por:

$$u^T b + (c^2 - u^T A^2) x_2 \le w (3.6)$$

onde:

w é a função objetivo de (PM) a ser minimizada.

Substituindo (3.4) em (3.6) obtemos:

$$\left(\sum_{k} \delta_{k} u^{k}\right)^{T} b + \left[c^{2} - \left(\sum_{k} \delta_{k} u^{k}\right)^{T} A^{2}\right] x_{2} \leq w$$
(3.7)

Logo:

$$c^{2}x_{2} + \sum_{k} \delta_{k} \{ (u^{k})^{T}b - (u^{k})^{T}A^{2}x_{2} \} \leq w$$
 (3.8)

Usando (3.5) obtemos:

$$c^2 \sum_k \delta_k = c^2 \tag{3.9}$$

$$w\sum_{k}\delta_{k}=w\tag{3.10}$$

| Substituindo (3.9) e (3.10) em (3.8) obtemos:

$$\sum_{k} \delta_{k} \{ (u^{k})^{T} b + [c^{2} - (u^{k})^{T} A^{2}] x_{2} \} \leq w \sum_{k} \delta_{k}$$
 (3.11)

Uma solução que satisfaz (3.11), também satisfaz (3.6), que é uma combinação positiva das restrições de (PM). Logo o conjunto de restrições de (PM) implica em (3.6) e portanto redundante.

O próximo teorema demonstra que o subproblema Lagrangeano é equivalente ao problema mestre relaxado de Benders (PMR u_2) em relação à variável x_2 .

Teorema 3.3.2 O problema mestre primal $(PMRu_2) \equiv \{PMR \ tal \ que \ u^k \in Tu_2\}$ é equivalente ao subproblema Lagrangeano (SDu_2) em relação a variável primal x_2 , e $v(PMRu_2) = v(SDu_2)$.

demonstração:

O problema (PMRu₂) é escrito da seguinte forma:

$$\begin{cases} & \text{minimizar}_{(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})} \ w \\ & sujeito \ a \ b^T u^k + [c^2 - (u^k)^T A^2] x_2 \leq w, \ u^k \in T u_2 \\ & ou \\ & \text{minimizar}_{(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})} \ \text{máximo} \{ \ (b - A^2 x_2)^T u^k + c^2 x_2, \ u^k \in T u_2 \ \} \\ & ou \end{cases}$$

$$\underset{(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})}{\operatorname{minimizar}_{(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})}} \left\{ \begin{array}{l} \operatorname{maximizar}_{(\mathbf{u} \geq \mathbf{0})} \ u^T b + (c^2 - u^T A^2) x_2 \\ \\ \operatorname{sujeito} \ a \ (u_1)^T A_1^1 \leq c^1 - u_2^T A_2^1 \end{array} \right\}$$

Desde que a solução ótima do problema de maximização interna é uma solução básica (ponto extremo do poliedro). Além disso o problema interno possui u₂ e x₂ fixos. Logo tomando o dual do problema interno obtemos:

$$\begin{aligned} & \text{minimizar}_{(\mathbf{x_2} \in \mathbf{Z_+^{n_2}})} \left\{ \begin{array}{l} & \text{minimizar}_{(\mathbf{x_1} \geq \mathbf{0})}(c^1 - u_2^T A_2^1) x_1 + (c^2 - u_2^T A_2^2) x_2 + u_2^T b_2 \\ & \\ & sujeito \ a \ A_1^1 x_1 \geq b_1 - A_1^2 x_2 \end{array} \right\} \end{aligned}$$

ou

 (SDu_2)

Desde que todos estes problemas são formulados na variável x_2 , a prova está completa. \Box

Um inconveniente do problema mestre de Benders, é a ilimitação da iteração inicial. A decomposição híbrida, contorna este problema utilizando

o valor da variável x_2 obtido da solução ótima do último problema Lagrangeano (SDu_2) .

A equivalência entre (SDu_2) e $(PMRu_2)$ não implica que ambos os subproblemas de programação matemática possam ser facilmente resolvidos. A "dificuldade" de um problema linear inteiro misto (Q) pode ser definida através do "gap relativo" do problema linear relaxado (\overline{Q}) e é dado por:

$$G(Q) = rac{v(Q) - v(\overline{Q})}{v(Q)}$$

Corolário 3.3.1 $G(SDu_2) = G(PMRu_2)$, isto é, (SDu_2) e $(PMRu_2)$ são problemas de igual dificuldade.

demonstração:

Pelo teorema 3.3.2, temos que $v(SDu_2) = v(PMRu_2)$, reescrevendo a prova do teorema 3.3.2 para $\overline{PMRu_2}$, isto é, $x_2 \in \mathbf{Z}_+^{n_2}$ é substituida por $x_2 \geq 0$, obtemos que $v(\overline{PMRu_2}) = v(\overline{SDu_2})$ e a prova está completa.

Consequentemente resolver o problema (SDu_2) em vez de resolver o problema $(PMRu_2)$ é mais atrativo, desde que a estrutura especial do subproblema dual seja explorada.

Em paralelo com o teorema 3.3.2 o próximo teorema diz respeito ao uso do subproblema de Benders (SPx_2) e o problema mestre da decomposição dual.

Consideremos agora a partição nas variáveis: $x^T = (x_1, x_2)$ e sejam $A^1 = [A_1^1 | A_1^2]$ e $A^2 = [A_2^1 | A_2^2]$. Dessa forma obtemos:

(SD) maximizar
$$c^1x_1 + c^2x_2 + u_1^T(b_1 - A_1^2x_2) + u_2^T(b_2 - A_2^2x_2)$$

sujeito a:
$$\left\{ \begin{array}{l} A_1^1x_1+A_1^2x_2\geq b_1\\ \\ x_1\in \mathbf{R}_+^{n_1}\\ \\ x_2\in \mathbf{Z}_+^{n_2} \end{array} \right.$$

Necessitamos ainda da seguinte notação:

- $F(SDx_2)$ é a região factível do conjunto de restrições de (SD), com x_2 fixo.
- Tx_2 é o conjunto de todos os $x^k = (x_1^k, x_2)^T$, tal que x^k é ponto extremo de $F(SDx_2)$.
- $\overline{T}x_2$ é o conjunto de todos $x^k = (x_1^k, x_2)^T$, tal que x^k é ponto extremo de F(SD).
- $\bullet (DMRx_2) \equiv (DMR : x^k \in Tx_2).$

Uma figura representativa, destes conjuntos é similar à (3.1) com (x_1, x_2) substituindo (u_1, u_2) .

Pela definição temos que:

$$\overline{T}x_2 \subseteq Tx_2$$
; $\overline{T}x_2 \subseteq T_{DA}$

$$Tx_2 \subseteq T_{DA} \text{ se, e somente se, } \overline{T}x_2 = Tx_2$$

$$\text{v(DMR tal que } x^k \in \overline{T}x_2) \geq \text{v(DMR}x_2) \geq \text{v(DM)}$$

Teorema 3.3.3 O problema mestre dual ($DMRx_2$) \equiv (DMR tal que $x^k \in Tx_2$) e o subproblema de Benders (SPx_2) são equivalentes em relação à variável dual u_2 e $v(DMRx_2) = v(SPx_2)$.

demonstração:

O problema (DMRx2) é escrito da seguinte forma:

$$\begin{cases} & \text{minimizar}_{(\mathbf{u_2} \geq \mathbf{0})} \ z \\ & sujeito \ a \ c^T x^k + (b^2 - A^2 x^k)^T u_2 \geq z, \ x^k \in T x_2 \end{cases}$$

$$ou$$

$$\text{maximizar}_{(\mathbf{u_2} \geq \mathbf{0})} & \text{mínimo} \{ \ c^T x^k + u_2^T (b_2 - A_2 x^k) \ , \ x^k \in T x_2 \ \}$$

$$ou$$

$$\text{maximixar}_{(\mathbf{u_2} \geq \mathbf{0})} \begin{cases} & \text{minimizar}_{(\mathbf{x} \in \mathbf{S})} \ c^T x^k + u_2^T (b_2 - A_2 x^k) \ \\ & sujeito \ a \ A_1^1 x_1 \geq b_1 - A_1^2 x_2 \end{cases}$$

Desde que a solução ótima do problema de minimização interna é uma solução básica (ponto extremo da região factível) e todos os seus pontos extremos são dados pelo conjunto Tx_2 , logo tomando o dual do p.p.l. interno com u_2 e x_2 fixos, podemos reescrever (DMR x_2) da seguinte forma:

$$\begin{cases} \text{maximizar}_{(\mathbf{u} \geq \mathbf{0})} \ c^2 x_2 + u_2 b_2 - u_2^T A_2^2 x_2 + u_1^T (b_1 - A_1^2 x_2) \\ sujeito \ a \ u_1^T A^1 \geq c^1 - u_2^T A_2^1 \\ ou \\ \begin{cases} \text{maximizar}_{(\mathbf{u} \geq \mathbf{0})} \ c^2 x_2 + u^T b - u^T A^2 x_2 \\ sujeito \ a \ u^T A^1 \leq c^1 \end{cases}$$

Que é justamente o dual de:

$$\begin{cases} \min_{(\mathbf{x_1} \ge \mathbf{0})} c^1 x_1 + c^2 x_2 \\ sujeito \ a \ A^1 x_1 \ge b - A^2 x_2 \end{cases}$$
ou
$$(SPx_2)$$

Desde que todos estes problemas são formulados na variável x_2 , a prova está completa.

O teorema 3.3.4, mostra como podemos verificar se uma solução é ótima.

Teorema 3.3.4 Seja x^* uma solução ótima de (P) e u^* uma solução ótima de (D). Então as seguintes afirmações são equivalentes:

- (a) A relaxacação Lagrangeana relativa a restrição $A_2x \geq b_2$ não possui "gap dual", isto é v(P) = v(D).
- (b) Existe uma solução ótima dual u^+ de (SPx_2^*) com $v(SDu_2^+) = v(SPx_2^*)$.
- (c) Existe uma solução ótima primal x^+ de (SDu_2^*) com $v(SPx_2^+)$ = $v(SDu_2^*)$.

3.4 Convergência do Algoritmo de Decomposição Híbrida

Nas seções anteriores foi discutido a relação entre a decomposição dual (via relaxação Lagrangeana) e a decomposição primal (Benders). Embora os subproblemas (SDu_2) e (SPx_2) sejam subproblemas mestre de um e de outro (teorema 3.3.2 e 3.3.4), isto não implica que o problema (P) podem ser resolvido usando o "ping-pong" entre os dois subproblemas. Por exemplo, nós nunca iremos terminar o algoritmo se a relaxação Lagrangeana não tiver gap dual zero, além disso, pode ocorrer que o algoritmo fique imerso em um processo de ciclagem (ainda que a relaxação Lagrangena não tenha gap dual). Portanto necessitamos ainda de mais alguns resultados a fim de auxiliar nos testes de convergência.

Lema 3.4.1 (a) seja u_0 a solução ótima dual de (SPx_2^0) e x^+ a solução ótima de (SDu_2^0) então $x_2^+ \neq x_2^0$ a menos que $v(SPx_2^0) = v(P)$.

(b) seja x^0 a solução ótima primal de (SDu_2^0) e u^+ a solução ótima dual de (SPx_2^0) então $u_2^+ \neq u_2^0$ a menos que $v(SDu_2^0) = v(P)$.

demonstração:

(a)

Suponha por absurdo que $v(SPx_2^0) > v(P)$ e $x_2^0 = x_2^+$. Como u^0 é solução ótima dual de (SPx_2^0) , pelo teorema 3.3.3, temos que u_2^0 é ótimo em $(DMRx_2^0)$ e:

$$c^T x^+ + u_2^0 (b_2 - A_2 x^+) \ge v(DMRx_2^0) = v(SPx_2^0)$$

Que contradiz o fato que x^+ é solução ótima de (SDu_2^0) desde que:

$$c^T x^+ + u_2^0 (b_2 - A_2 x^+) = v(SDu_2^0) \le v(P) < v(SPx_2^0).$$

O item (b) segue a mesma linha de raciocínio do item (a).

Suponha que o método de decomposição híbrida, resolva sucessivamente o subproblema primal (SPx_2^k) e o subproblema dual (SDu_2^k) , ou seja:

$$(SPx_2^1) \longrightarrow (SDu_2^2) \longrightarrow (SPx_2^3) \longrightarrow (SDu_2^4) \longrightarrow (SPx_2^5) \longrightarrow \dots$$
 (3.12)

Dessa forma não podemos reaplicar o método em um número de iterações menor que quatro. De fato, pelo lema 3.4.1 temos que $v(SPx_2^1) \neq v(SPx_2^2)$ desde que $x_2^3 = x_2^2 \neq x_2^1$ (onde x_2^2 é solução ótima primal associado ao subproblema dual (SDu_2^2)). Assim é possivel que $x_2^4 = x_2^1$, pois o lema 3.4.1 só se aplica a soluções consecutivas, um argumento análogo é usado no subproblema dual. Van Roy [28] sugere que um teste de convergência seja a contagem do número de iterações, sem que haja uma melhora da solução, desse modo se o contador chegar a quatro devemos resolver um dos problemas mestre e com uma nova solução reíniciamos o "ping-pong".

3.5 Algoritmo de Decomposição Híbrida

1. Inicialização

```
1.1. k := 0;

1.2. v_P := (+\infty);

1.3. v_D := (-\infty);

1.4. \varepsilon \ge 0; (tolerância)

1.5. \xi := 1;

1.6. Selecione u_2^1 \ge 0;
```

2. Subproblema Dual

```
2.1. k := k + 1;
2.2. Resolva o Subproblema Dual (SDu<sub>2</sub><sup>k</sup>)
2.3. Seja x<sup>k</sup> solução ótima de (SDu<sub>2</sub><sup>k</sup>);
2.4. Se v<sub>D</sub> < v(SDu<sub>2</sub><sup>k</sup>) Então

2.4.1. v<sub>D</sub> := v(SDu<sub>2</sub><sup>k</sup>);

2.4.2. ξ := 1

Senão

2.4.3. ξ := ξ + 1;
```

- 2.5. Se $|v_D v_P| \leq \varepsilon$ Então
 - 2.5.1. Fim. (A solução ótima do problema é a última solução do subproblema primal)
- 2.6. Se $\xi = 4$ Então
 - 2.6.1. Resolva o problema mestre Primal (PMR);
 - 2.6.2. Seja (w, x_2^{k+1}) solução ótima de (PMR);
 - $2.6.3. \ \xi := 1;$
 - 2.6.4. Se $|v_P w| \le \varepsilon$ Então
 - 2.6.4.1. Fim. (A solução ótima do problema é a última solução do subproblema primal)

3. Subproblema Primal

- 3.1. k := k + 1;
- 3.2. Resolva o Subproblema Primal (SPx_2^k)
- 3.3. Seja x^k e u_2^k solução ótima primal e dual de (SP x_2^k);
- 3.4. Se $v_P > v(SPx_2^k)$ Então
 - 3.4.1. $v_P := v(SPx_2^k);$
 - 3.4.2. $\xi := 1$

Senão

3.4.3.
$$\xi := \xi + 1$$
;

- 3.5. Se $|v_P v_D| \le \varepsilon$ ou $|v_P w| \le \varepsilon$ Então
 - 3.5.1. Fim. (A solução ótima do problema é a última solução do subproblema primal)
- 3.6. Se $\xi = 4$ Então
 - 3.6.1. Resolva o problema mestre Dual (DMR);
 - 3.6.2. Seja (u_0, u_2^{k+1}) solução ótima; de (DMR);
 - 3.6.3. $\xi := 1$;
 - 3.6.4. Se $|v_D u_0| \le \varepsilon$ Então
 - 3.6.4.1. Fim. (A solução ótima do problema é a última solução do subproblema primal)

4. Volte ao passo 2

Teorema 3.5.1 O algoritmo de decomposição híbrida é finito e resolve (P). demonstração:

Basta usar o lema 3.4.1 e o fato que os algoritmos de decomposição de Benders e dual são finitos e resolvem (P).

O lema (3.4.1) assume que o subproblema (SPx_2^k) possui uma solução ótima finita. Se (SPx_2) for ilimitado o teste de convergência não pode ser utilizado. Holmberg [13], sugere um outro teste de convergência, onde as soluções ilimitadas de (SPx_2^k) são consideradas.

3.6 Considerações Finais

i- O algoritmo de decomposição híbrida apesar de sua dificuldade de implementação, tem se mostrado bastante eficiente para a resolução de alguns problemas de programação linear inteira mista. Van Roy [29] mostra que o problema de alocação utilizando o algoritmo de decomposição híbrida converge mais rapidamente que o método de decomposição de Benders e a relaxação Lagrangeana via algoritmo do subgradiente.

ii- Podemos observar que o problema dos limitantes (inferior e superior) da solução ótima é controlado no algoritmo de decomposição híbrida, desse modo temos um controle maior sobre a solução do problema.

Capítulo 4

Estudo do Caso: Planejamento de uma Fundição

4.1 Definição do Problema

Uma fundição possui uma encomenda para a fabricação de várias peças (cada peça deve ser fabricada com um tipo específico de liga metálica), em um determinado horizonte de tempo. Para esta empreitada, a fundição possui um forno, onde somente um tipo de liga pode ser utilizada por período e máquinas, onde são fabricadas os moldes, que produzirão as peças requisitadas pelos clientes. O problema consiste em determinar qual liga deve ser utilizada em cada período e como deve ser programada cada máquina para a fabricação de cada peça, tal que o custo de fabricação seja mínimo.

São fornecidos os seguintes dados:

L = Número de ligas.

M = Número de máquinas.

T = Número de períodos.

P = Número de tipos de peças.

 a_{im} = Quantidade (tonelada) de peças do tipo "i"

que podem ser produzidas na máquina "m" por hora.

 F_t = Capacidade do forno (toneladas/hora) no período "t".

Demanda (toneladas) de peças do tipo "i".

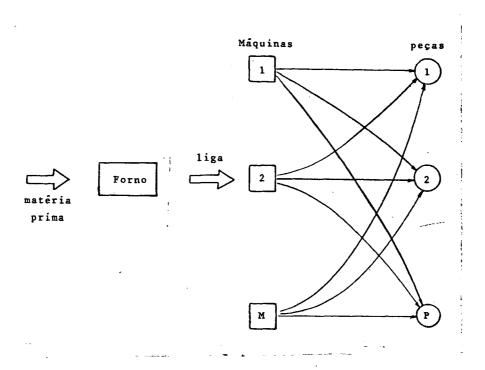
= Número de horas no período "t".

 $L_j = \{ \text{ peça "i" tal que utiliza a liga "j" } \}.$ $\overline{L}_j = \{ \text{ peça "i" tal que não utiliza a liga "j" } \}.$

= "Custo" de produzir a peça do tipo "i" no período "t".

O custo poderia refletir a urgência de se produzir uma determinada peça; por exemplo: $c_{jt} = c_j t^2$, onde c_j é um custo associado à peça "j" decorrente de clientes preferênciais.

O problema pode ser representado esquematicamente da seguinte forma:



Modelo Matemático 4.2

A programação das máquinas pode ser formulada pela fração de tempo que será dedicada para produzir uma determinada peça em cada período.

As variáveis de decisão são:

Variáveis:

 T_{imt} = Fração de tempo da máquina "m", utilizado para produzir a peça do tipo "i" no período "t".

$$x_{jt} = \begin{cases} 1 & \text{se a liga do tipo "j" for utilizada no período "t".} \\ 0 & \text{caso contrário.} \end{cases}$$

4.2.1 Formulação

Função Objetivo

$$minimizar \sum_{t=1}^{T} \sum_{i=1}^{P} \sum_{m=1}^{M} h_t c_{it} a_{im} T_{imt}$$

Restrições

• Apenas uma liga pode ser utilizada por período

$$\sum_{j=1}^{L} x_{jt} = 1 \quad t = 1, \dots, T$$
 (4.1)

• Se a liga "j" for utilizada no período "t", apenas peças de L, podem ser fabricadas

$$\sum_{(i \in L_j)} T_{imt} \leq (1 - x_{jt}) \Lambda + 1 \quad t = 1, \dots, T$$

$$j = 1, \dots, L$$

$$m = 1, \dots, M$$

$$(4.2)$$

$$\sum_{(i \in \overline{L_j})} T_{imt} \leq (1 - x_{jt}) \Lambda \qquad t = 1, \dots, T \qquad (4.3)$$

$$j = 1, \dots, L$$

$$m = 1, \dots, M$$

 $\Lambda > 0 \ \acute{e} \ um \ n\'umero \ grande$

• Capacidade de Produção do Forno

$$\sum_{j=1}^{L} \sum_{m=1}^{M} \sum_{(i \in L_j)} a_{im} T_{imt} \le F_t \quad t = 1, \dots, T$$
 (4.4)

• Demanda das Peças

$$\sum_{t=1}^{T} \sum_{m=1}^{M} a_{im} h_t T_{imt} \ge b_i \quad i = 1, \dots, P$$
 (4.5)

4.2.2 Exemplo

Para ilustrar o problema, considere o seguinte conjunto de dados:

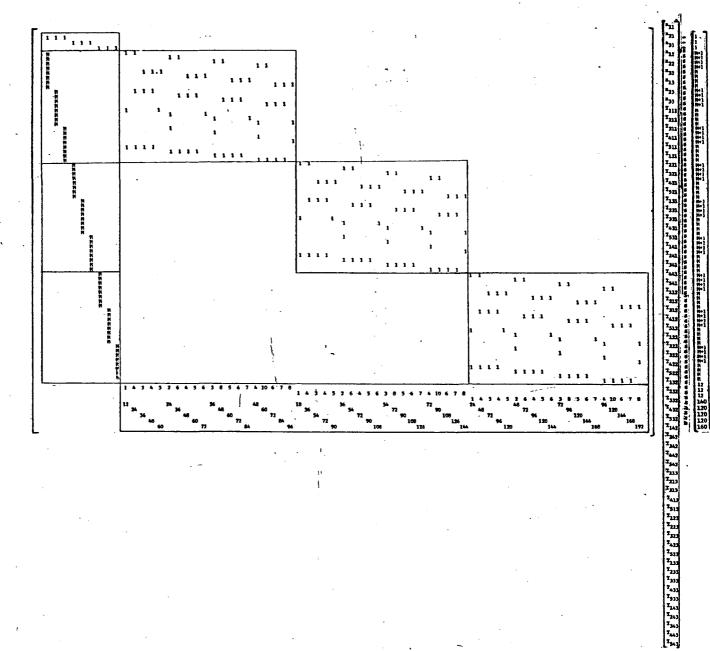
- Número de Ligas = 3.
- Número de Máquinas = 4.
- Número de Periodos = 3.
- Número de Peças = 5.
- Capacidade do Forno:
 - $F_t = 12$ toneladas/hora (i = 1,...,3).
- Número de horas por período:
 - $h_1 = 12 \text{ horas.}$
 - $-h_2 = 18 \text{ horas.}$
 - $-h_3 = 24$ horas.
- Peças por ligas.
 - $\text{ Liga } 1 = \{1, 2\}.$
 - Liga $2 = \{2, 3, 4\}$ (a peça 2 pode ser feita usando a liga 1 ou 2).
 - $\text{ Liga } 3 = \{5\}.$
- Demanda de Peças:
 - $-b_1 = 140$ toneladas.
 - $-b_2 = 120$ toneladas.

- $-b_3 = 120$ toneladas.
- $b_4 = 120$ toneladas.
- $-b_5 = 160$ toneladas.
- Quantidade de peças (toneladas) "i" produzidas na máquina "m":

	Máquina			
Peça	1	2	3	4
1	1	2	3	4
2	2	3	4	5
3	3	4	5	6
4	4	5	6	7
5	5	6	7	8

• Custo de produção das peças:

	Período		
Peça	1	2	3
1	2	3	4
2	3	4	5
3	4	5	6
4	5	6	7
5	6	7	8



4.2.3 Comentários

i- O problema apresenta em seu conjunto de restrições a seguinte estrutura:

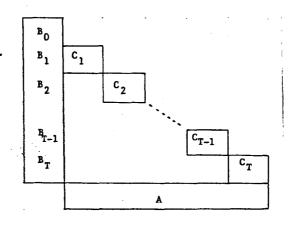


figura (4.1)

Que possuem as seguintes características:

- Matrizes C_1, \ldots, C_T :
 - São identicas.
 - Os elementos das matrizes são formados por zeros ou um.
- Matriz [$B_0 | \dots | B_T$] :
 - as colunas são linearmente independentes.

ii- O conjunto de restrições (4.2) e (4.3), são necessários para representar a factibilidade do problema (P). O seguinte equacionamento representa o problema, se $L_i \cap L_j = \phi$, $\forall i, j$. Desse modo as restrições (4.2) e (4.3) podem ser substituidas por:

$$\sum_{i \in L_j} T_{imt} \le x_{jt} \qquad t = 1, \dots, T$$

$$j = 1, \dots, L$$

$$m = 1, \dots, M$$

$$(4.6)$$

Exemplificando os dados:

- L=M=P=2.
- T=1.
- $L_1 = \{1, 2\}$.
- $L_2 = \{2\}.$

E suponha que:

- $x_{11} = 1$ (a liga 1 está sendo preparada no período 1).
- $x_{21} = 0$ (a liga $\frac{1}{2}$ não está sendo preparada no período 1).

Usando as esquações (4.6) obtemos o seguinte conjunto de restrições:

$$T_{111}+$$
 T_{211} = 1 (4.7)
 $T_{211}+$ T_{221} = 1 (4.8)
 T_{211} = 0 (4.9)
 T_{221} = 0 (4.10)

As restrições deveriam ser descritas apenas por (4.7) e (4.8).

As restrições (4.9) e (4.10) obriga $T_{2mt} = 0$ pois $x_{2t} = 0$ e $2 \in L_2$. Istojé falso, pois a peça 2 pode também ser produzida pela liga 1.

4.3 Método de Decomposição Híbrida aplicado ao Modelo

4.3.1 Introdução

Como vimos no capítulo 3, o método de decomposição híbrida procura utilizar as estruturas especiais do problema (P), para a resolução de 2 (dois) subproblemas, que são o subproblema primal (obtido pela fixação das variáveis inteiras em zero ou um) e subproblema dual (obtidos através da dualização das restrições complicadas).

4.3.2 Subproblema Primal

Suponha, por um momento, que os tipos de ligas a serem utilizadas em cada período já estejam selecionados, ou seja, a variável x_{jt} é fixada em 0 ou 1. Obtemos desse modo o seguinte p.p.l.:

$$(SPx_{jt})$$
 minimizar $\sum_{t=1}^{T}\sum_{i=1}^{P}\sum_{m=1}^{M}h_{t}c_{it}a_{im}T_{imt}$

sujeito a

$$\sum_{(i \in L_j)} T_{imt} \le (1 - x_{jt}) \Lambda + 1 \qquad t = 1, \dots, T \qquad (4.11)$$

$$j = 1, \dots, L$$

$$m = 1, \dots, M$$

$$\sum_{(i \in \overline{L_j})} T_{imt} \leq (1 - x_{jt}) \Lambda \qquad t = 1, \dots, T \qquad (4.12)$$

$$j = 1, \dots, L$$

$$m = 1, \dots, M$$

$$\sum_{j=1}^{L} \sum_{m=1}^{M} \sum_{(i \in L_j)} a_{im} T_{imt} \le F_t \quad t = 1, \dots, T$$
 (4.13)

$$\sum_{t=1}^{T} \sum_{m=1}^{M} a_{im} h_t T_{imt} \ge b_i \qquad i = 1, \dots, P \qquad (4.14)$$

$$T_{imt} \ge 0$$

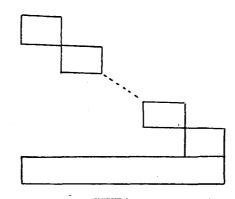
$$i = 1, \dots, P$$

$$m = 1, \dots, M$$

$$t = 1, \dots, T$$

 $\Lambda > 0$ é um número grande

Cuja a matriz de restrições apresenta a seguinte estrutura:



Que pode ser explorada utilizando o algoritmo de decomposição de "Dantzig-Wolfe" [7] ou o método de decomposição de "Rosen" [23].

Em nosso estudo, decidimos optar pelo método de decomposição de Rosen [23], devido à algumas características do problema que construiremos a seguir.

Comentários

Ao aplicarmos o método de decomposição de Rosen (ver apêndice B) no problema em questão, podemos observar que aparecem algumas simplificações que são:

- A matriz B_{i1} inicial é formada pelas colunas da matriz de restrição das variáveis de folgas.
- A atualização das matrizes particionadas, não se faz necessária, devido ao fato de ser composta apenas de zeros e uns.
- Devido a particularidade do problema, as matrizes B_{i2} são geradas de modo implicíto, ou seja, são gerados somente quando solicitado.

4.3.3 Subproblema Dual

Se dualizarmos o conjunto de restrições (4.4) e (4.5) do problema (P), obtemos o seguinte subproblema dual:

$$(SD\mu, \theta) \ minimizar \qquad \sum_{t=1}^{T} h_{t} \{ \sum_{i=1}^{P} c_{it} [\sum_{m=1}^{M} a_{im} T_{imt}] \} +$$

$$\sum_{i=1}^{P} \mu_{i} (\sum_{t=1}^{T} \sum_{m=1}^{M} b_{i} - a_{im} T_{imt}) +$$

$$\sum_{t=1}^{T} \theta_{t} (\sum_{j=1}^{L} \sum_{m=1}^{M} \sum_{i \in L_{j}} a_{im} T_{imt} - F_{t})$$

sujeito a

$$\sum_{j=1}^{L} x_{jt} = 1 \qquad t = 1, \dots, T \qquad (4.15)$$

$$\sum_{(i \in L_j)} T_{imt} \leq (1 - x_{jt}) \Lambda + 1 \quad t = 1, \dots, T \qquad (4.16)$$

$$j = 1, \dots, L \qquad m = 1, \dots, M$$

$$\sum_{(i \in \overline{L_j})} T_{imt} \leq (1 - x_{jt}) \Lambda \qquad t = 1, \dots, T \qquad (4.17)$$

$$j = 1, \dots, L \qquad m = 1, \dots, M$$

$$T_{imt} \geq 0 \qquad i = 1, \dots, M$$

$$T_{imt} \geq 0 \qquad i = 1, \dots, M$$

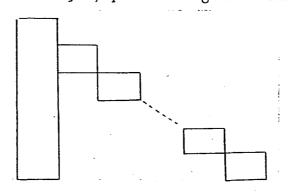
$$t = 1, \dots, M$$

$$t = 1, \dots, T$$

$$x_{jt} \in \{0, 1\} \qquad j = 1, \dots, L$$

 $\Lambda > 0$ é um número grande

Cuja a matriz de restrições, apresenta a seguinte estrutura:



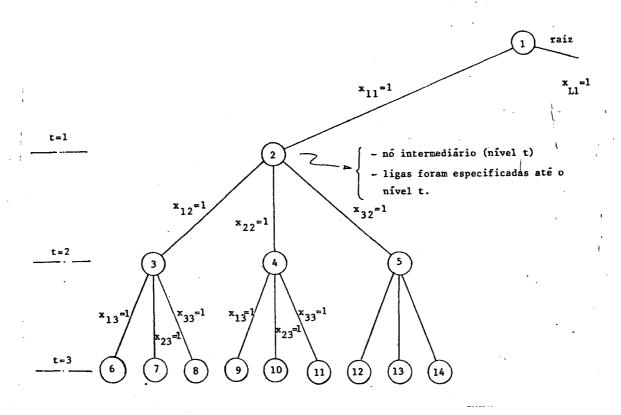
Desse modo, para a resolução do subproblema dual, optamos por

um algoritmo do tipo "enumeração", onde procuramos verificar todas as soluções factíveis inteiras possíveis, ou seja, ao fixarmos a variável inteira x_{ji} (tal que a restrição (4.11) seja satisfeita), procuramos verificar se as ligas escolhidas nos períodos podem fabricar as peças requeridas pelo cliente.

Dessa forma, não há necessidade de verificarmos "a priori" todas as combinações possíveis, que seria de L^T, mas um subconjunto que pode ser armazenado em uma lista e a solução ótima do problema pode ser determinada utilizando uma simples inspeção.

4.4 Exemplo

Utilizando o problema teste obtemos:



Considere os seguintes caminhos, formados pelos nós da árvore

acima:

- $C_1 = \{1,2,3,6\}$, que fornece a seguinte escolha de ligas nos períodos:
 - período 1 :- Liga 1.
 - período 2 :- Liga 1.
 - período 3 :- Liga 1.

Observe que $L_1 \cup L_1 \cup L_1 = \{1,2\}$, ou seja, só seram fabricadas as peças 1 e 2. Logo este caminho não deve ser considerado, pois esta escolha não atende às especificações do cliente, desse modo retiramos este caminho da árvore.

- $C_2 = \{1,2,4,11\}$, que fornece a seguinte escolha de ligas nos períodos:
 - período 1 :- Liga 1.
 - período 2 :- Liga 2.
 - período 3 :- Liga 3.

Obeserve que $L_1 \cup L_2 \cup L_3 = \{1,2,3,4,5\}$, ou seja, seram fabricas as peças 1,2,3,4 e 5. Logo se escolhermos este caminho a solução obtida será factível, pois as especificações do cliente são atendidas.

Comentários

Podemos determinar o valor das variáveis duais (μ, θ) , utilizando o algoritmo do subgradiente.

4.4.1 Problema Mestre

O método de decomposição híbrida está baseado no sucesso do "ping-pong" entre os subproblemas primal e dual, mas somente a iteração destes dois subproblemas não garante a convergência do algoritmo. Necessitamos desse modo, resolver a uma certa altura do algoritmo, um problema mestre (primal ou dual) que irá gerar uma nova solução que servirá como ponto de partida na continuação do problema.

O algoritmo de decomposição híbrida possibilita a escolha de apenas um problema mestre (primal ou dual). No nosso caso, escolhemos resolver o problema mestre primal, devido ao fato de conhecermos todas as soluções inteiras factíveis. Desse modo é necessário calcular o valor mínimo que a função do problema mestre assume com as soluções obtidas.

4.5 Solução do Problema Teste:

- * Solução Ótima do Problema Teste *
 - Ligas a serem utilizadas:
 - Período 1: Liga do tipo 2.
 - Período 2: Liga do tipo 3.
 - Período 3: Liga do tipo 1.
 - Valor Ótimo: 3.356,6667

Tabela de Distribuição de Seviço

• Período 1:

		MAQUINA			
PEÇA	1	2	3	4	
1					
2				*	
3			*****	*****	
4	*****	******	*		
5					

• Período 2:

	MĀQUINA			
PEÇA	1	2	3	4
1				
2				
3				
4				
5				*****

• Período 3:

	MÁQUINA				
PEÇA	1	2	3	4	
1		\ -	*****	****	
2	**	*****		***	
3		1	!		
4					
5					

4.6 Implementação

A implementação do algoritmo de decomposição híbrida foi feita em linguagem "Pascal", utilizando o compilador Turbo-Pascal versão 3.0 da Borland em um micro-computador da linha IBM Pc-AT com co-processador numérico 80287 e clock de 12 MHz com sistema operaconal MS-DOS versão 3.30 da Microsoft.

Para a verificação da implementação do algoritmo de decomposição híbrida, foi feita uma rotina do tipo enumeração, onde todas as soluções factíveis são sondadas.

4.7 Considerações Finais

- i- Não foi possivel notar a influência do problema mestre primal na eficiência computacional do algoritmo, uma vez que nos testes realizados a sua resolução não foi necessária. Acreditamos que a influência no tempo computacional seja bastante grande, devido ao fato do problema mestre primal ser de difícil resolução. Desse modo o desempenho do algoritmo está intimamente ligado ao "ping-pong" entre os subproblemas.
- ii- Apesar da grande dificuldade de implementação, o algoritmo se mostra bastante promissor.
- iii- Vários exemplos, de dimensões semelhantes ao exemplo dado, foram resolvidos e em nenhum caso foi necessário a resolução do problema mestre primal, isto é, a solução ótima foi obtida em poucas iterações pelo "ping-pong".

Capítulo 5

Conclusão

Neste trabalho fizemos uma revisão de alguns procedimentos de decomposição de programação linear inteira mista, "primal" (Benders), "dual" (relaxação lagrangeana) e primal-dual" (híbrido).

Implementamos o algoritmo de decomposição híbrida para a resolução de um problema real, onde foi explorada particularidades da matriz de restrição. O algoritmo híbrido, embora permita explorar muitas particularidades do problema, é de difícil implementação, pois há a necessidade de algoritmos eficientes para a resolução dos subproblemas e de um problema mestre.

Devido a limitação (memória) do compilador "Turbo Pascal" não foi possível, nesta etapa, apresentar resultados computacionais significativos. Os casos estudados foram muito pequenos, porém apontam para uma direção promissora, se observarmos o tempo computacional extremamente baixo para a execução do problema, decorrente da facilidade de resolução dos subproblemas.

Desse modo, podemos conduzir pesquisas futuras nos seguintes tópicos:

- Métodos de enumeração eficiente e heurísticas na resolução do subproblema dual.
- A utilização de heurísticas para a solução do problema mestre dual e primal.
- Comparação computacional decorrente da escolha do problema mestre primal ou dual.

- A utilização do algoritmo em outros tipos de problemas lineares inteiros mistos, ou mesmo a resolução de outros tipos de problemas de programação matemática (Holmberg [13]).
- Verificar o desempenho computacional em problemas de grande porte e o efeito do problema mestre.

Apêndice A

Definição 1 Um conjunto $S \subseteq \mathbb{R}^n$ (não vazio) é convexo se, e somente se, $\alpha x_1 + (1-\alpha)x_2 \in S$, $\alpha \in [0,1]$ e para todo $x_1, x_2 \in S$.

Definição 2 Um politopo é um conjunto convexo formado por um número finito de subespaços convexos.

Definição 3 Um poliedro é um politopo limitado.

Definição 4 Seja $S \subseteq \mathbb{R}^n$ um conjunto convexo não vazio, o ponto $x \in S$ é um ponto extremo de S se não existem $x_1, x_2 \in S$, tal que $x = \alpha x_1 + (1 - \alpha)x_2$, $\alpha \in (0,1)$ e $x_1 \neq x_2$.

Definição 5 Um vetor n-dimensional d é uma raio extremo de um conjunto $S \subseteq \mathbb{R}^n$ se $x + \alpha d \in S$ para todo $x \in S$ e para todo $\alpha \geq 0$.

Corolário 1 Um poliedro possui um número finito de pontos extremos e raios extremos.

A demonstração pode ser encontrada em Lasdon [15].

Definição 6 Seja $S \subseteq \mathbb{R}^n$ um conjunto convexo, não vazio e limitado, então o ponto $x \in S$ é uma combinação convexa dos pontos $x_1, \ldots, x_r \in S$ se existirem escalares $\alpha_1, \ldots, \alpha_r$ tal que:

$$\begin{cases} x = \sum_{i=1}^{r} \alpha_i x_i \\ \sum_{i=1}^{r} \alpha_i = 1 \\ \alpha_i \ge 0 \ i = 1, \dots, r \end{cases}$$

Teorema 1 Seja $S \subseteq \mathbb{R}^n$ um conjunto convexo não vazio e $x^m \in S$ é uma combinação convexa dos pontos extremos x_1, \ldots, x_r de S mais uma combinação linear não negativa do raios extremos d_1, \ldots, d_s de S (desde que o conjunto não seja limitado), ou seja:

$$\begin{cases} x = \sum_{i=1}^{r} \alpha_i x_i + \sum_{j=1}^{s} \alpha_i d_j \\ \\ \sum_{i=1}^{r} \alpha_i = 1 \end{cases}$$

$$\alpha_i \ge 0 \ i = 1, \dots, r$$

$$\mu_j \ge 0 \ j = 1, \dots, s$$

A demonstração pode ser encontrada em Lasdon [17].

Definição 7 O casco convexo de um conjunto $S \subseteq \mathbf{R}^n$ (não vazio) é a intersecção de todos os conjuntos convexos que contém S e é denotado por C(S).

Definição 8 Um conjunto $S \subseteq \mathbf{R^n}$ (não vazio) é chamado de cone com vértice no ponto zero, se $x \in S$, então $\alpha x \in S$ para todo $\alpha \geq 0$. Se S for convexo então o cone é convexo.

Definição 9 Uma função $f: S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ é convexa se, e somente se, $f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha)f(x_2)$, para todo $x_1, x_2 \in S$ e $\alpha \in [0,1]$.

Definição 10 Uma função $f: S \subseteq \mathbb{R}^n \longmapsto \mathbb{R}$ é côncava se -f for convexa.

Definição 11 Uma função $f: S \subseteq \mathbf{R}^n \longmapsto \mathbf{R}$ é semi-contínua superior no ponto a, se para qualquer $\varepsilon > 0$ \exists uma vizinhança $\mathcal V$ tal que se $x \in S$, então $f(x) - f(a) < \varepsilon$.

Teorema 2 (Desigualdade de Cauchy-Schwarz) Seja $S \subseteq \mathbb{R}^n$ um conjunto convexo não vazio. Então para todo $x_1, x_2 \in S$ temos que:

$$|x_1^T x_2| \le ||x_1|| \; ||x_2||$$

A demonstração pode ser encontrada em Bazaraa [4]

Teorema 3 Seja $S \subseteq \mathbf{R^n}$ um conjunto convexo não vazio e $\|.\|$ a norma Euclidiana. Então para todo $x_1, x_2 \in S$ temos que:

$$||x_1 + x_2||^2 = ||x_1||^2 + ||x_2||^2 + 2x_1^T x_2$$

A demonstração pode ser encontrada em Bazaraa [4].

Apêndice B

Método de Decomposição de Rosen

Seja o seguinte problema de programação linear:

(P)
$$minimizar w = \sum_{i \in I} c_i^T x_i$$

sujeito a

$$\begin{array}{cccc} \sum_{i \in I} A_i x_i = & b_0 \\ B_i x_i = & b_i & i \in I \\ & x_i \geq & 0 & i \in I \end{array}$$

onde:

- as matrizes A_i são de ordem $m_0 \times n_i$ e possuem rank completo.
- as matrizes B_i são de ordem $m_i \times n_i$ e possuem rank completo.

Considere a seguinte partição nas matrizes $B_i = [B_{i1}|B_{i2}]$, onde o rank (B_i) = m_i e $B_{i1}^{-1}b_i \ge 0$ ($i \in I$), $A_i = [A_{i1}|A_{i2}]$ e nos vetores $x_i = [x_{i1}|x_{i2}]$.

Desse modo o problema (P) pode ser escrito da seguinte forma:

$$(P_1) \qquad minimizar \ w = \sum_{i \in I} (c_{i1}^T x_{i1} + c_{i2}^T x_{i2})$$

sujeito a

$$\sum_{i \in I} (A_{i1}x_{i1} + A_{i2}x_{i2}) = b_0$$

$$B_{i1}x_{i1} + B_{i2}x_{i2} = b_i \quad i \in I$$

$$x_{i1} \ge 0 \quad i \in I$$

$$x_{i2} \ge 0 \quad i \in I$$

Relaxando a restrição $x_{i1} \ge 0$, podemos escrever x_{i2} da seguinte forma:

$$x_{i2} = B_{i1}^{-1}b_1 - B_{i1}^{-1}B_{i2}x_{i2}$$

Desse o modo o problema (P_1) é equivalente a:

$$minimizar \sum_{i \in I} d_i x_{i2}$$

sujeito a

$$\sum_{i \in I} M_i x_{i2} = \widehat{b}_0 \quad i \in I$$

$$x_{i2} \geq 0 \quad i \in I$$

onde:

$$M_{i} = A_{i2} - A_{i1}B_{i1}^{-1}B_{i2}$$

$$d_{i} = c_{i2} - c_{i1}B_{i1}^{-1}B_{i2}$$

$$\hat{b}_{0} = b_{0} - \sum_{i \in I} A_{i1}B_{i1}^{-1}b_{i}$$

Obtemos desse modo um p.p.l. com um numero menor de colunas e variáveis, que agora pode ser resolvido utilizando algum código Simplex.

Teste de Otimalidade

O vetor $(x_{i1}^0,x_{i2}^0)^T$ resolve o problema bloco angular se, e somente se, $x_{i1}^0 \geq 0$.

Se o teste de otimalidade não for satisfeito, podemos escolher uma variável x_{i2} , tal que:

$$\theta_s = \min_{j \in J} \frac{x_{j1}}{x_{j1} - x_{j1}^0}$$

onde:

$$J = \{j \text{ tal que } \mathbf{x}_{j1}^0 < 0\}$$

Para atualizar a base devemos resolver o seguinte p.p.l.:

minimizar x_{1s}

sujeito a

$$\sum_{k=1}^{m} B_{i1}^{k} x_{k1} + \sum_{k \in K} x_{k2} = b$$

$$x_{k1} \ge 0$$

$$x_{k2} \ge 0$$

onde:

$$K=\{k \text{ tal que } x_{k2}^0>0\}$$

Bibliografia

- [1] Agmon, S., The Relaxation Methods for Linear Inequalites. Canadian Journal of Mathematics, 6, (1954), 382 392.
- [2] Allen, E., Helgason, R., Kennington, J., Shetty, B., A Generalization of Poljak's Convergence Result for Subgradient Optimization. Mathematical Programming, 37, (1987), 309 317.
- [3] Bazaraa, M.S., Sherali, H.D., On the Choise of Step Size in Subgradient Optimization. European Journal of Operation Research, 7, (1981), 380-388.
- [4] Bazaraa, M.S., Javis, J.J., Linear Programming and Network Flows. John-Wiley and Sons (1977).
- [5] Benders, J.F., Partitioning Procedures for Solving Mixed Variables Programming Problems. Numerische Mathematik, 4, (1962), 238 252.
- [6] Côté, G., Laughton, M. A., Large Mixed Integer Programming Benders Type Heuristics. European Journal of Operation Research, 16, (1984), 327 - 333.
- [7] Dantzig, G.B.; Wolfe, P., Decomposition Principle for Linear Programming. Operation Research, 8, vol 1, (1960), 101 111.
- [8] Fisher, M.L., The Lagragian Relaxation Methods for Solving Integer Programming Problems. Management Science, 1, vol 27, (1981), 1 18.
- [9] Geoffrion, A.M., Lagrangian Relaxation for Integer Programming. Mathematical Programming Study 2, (1974), 82 114.
- [10] Geoffrion, A.M., Graves, G.W., Multicommodity Distribution System Design by Benders Decomposition. Management Science, 5, vol 20, (1974), 822 -844.

- [11] Grinold, R.C., Stespest Ascent for Large Scale Linear Programming. SIAM Review, 3, vol 14, (1972), 447 464.
- [12] Held, M., Wolfe, P., Crowder, H.P., Validation of Subgradient Optimization. Mathematical Programming, 6, (1974), 62 - 88.
- [13] Holmberg, K., On the Convergence of Cross Decomposition. Mathematical Programming, 47, (1990), 269 296.
- [14] Kim, S., Ahn, H., Cho, S., Variable Target Value Subgradient Method. Mathematical Programming, 49 (1991), 359 369.
- [15] Lasdon, L.S., Optimization Theory for Large System. Mac.Millan (1970).
- [16] Magnanti, T.L., Wong, R.T., Accelerating Benders Decomposition Algorithmic Enhancement and Model Selection Criteria. Operation Research, 3, vol 29, (1981), 464 484.
- [17] Minoux, M., Mathematical Programming Theory and Algorithms. John-Wiley and Sons (1986).
- [18] Motzink, T.S., Schoenberg, I.J., The Relaxation Methods for Linear Inequalites. Canadian Journal of Mathematics, 6, (1954), 392 404.
- [19] Nemhauser, G.L., Wolsey, L.A., Integer and Combinatorial Optimizatiom. John-Wiley and Sons (1988).
- [20] Poljak, B.T., A General Method of Solving Extremum Problems. Soviet Mathematical Doklady, 3, vol 8, (1967), 593 597.
- [21] Poljak, B.T., Minimization of Unsmooth Functionals. USSR Computational Mathematics and Mathematics Physics, vol 9, (1969), 14 29.
- [22] Rockafellar, R.T., Convex Analysis. Princeton University Press (1972).
- [23] Rosen, J.B., Primal Partition Programming for Block Diagonal Matrices. Numerische Mathematik, vol 6, (1964), 250 - 260.
- [24] Salkin, H.M., Integer Programming. Addison Wesley Publishing Company (1975).
- [25] Sen, S., Sherali, H.D., A Class of Convergence Result of Subgradient Optimization. Mathematical Programming, 37, (1986), 279 297.
- [26] Shapiro, J.F., Mathematical Programming Structures and Algorithms. John-Wiley and Sons (1979).

- [27] Souza, K.X.S., Multi-item Lot Sizing by the Cross Decomposition Method. Alio-euro, Workshop: Pratical Combinatorial Optimization, (1989), Rio de Janeiro Brasil.
- [28] Van Roy, T.J., Cross Decomposition for Mixed Integer Programming. Mathematical Programming, 25, (1983), 46 63.
- [29] Van Roy, T.J., A Cross Decomposition Algorithm for Capacitated Facility Location. Operation Research, 1, vol 34, (1986), 145 - 163.