

Uma Metodologia de Feixe e Benders Aplicada a um Problema  
Linear Inteiro de Grande Porte

PLÁCIDO ROGÉRIO PINHEIRO

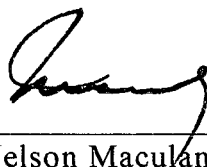
TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS  
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA  
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS  
REQUISITOS NECESSÁRIOS PARA OBTENÇÃO DO GRAU DE  
DOUTOR EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por:



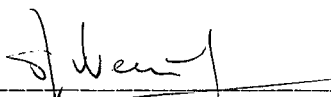
---

Prof. Dr. Ing. Paulo Roberto Oliveira  
(Presidente)



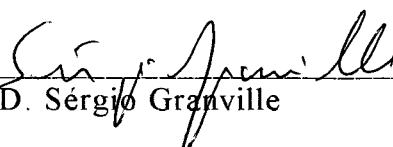
---

Prof. D.Sc. Nelson Maculan Filho



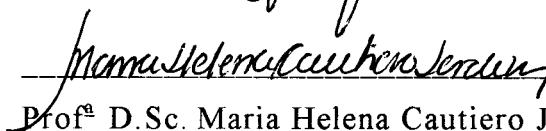
---

Prof<sup>a</sup> D.Sc. Susana Scheimberg de Makler



---

Prof. Ph.D. Sérgio Granville



---

Prof<sup>a</sup> D.Sc. Maria Helena Cautiero Jardim

Rio de Janeiro, RJ - Brasil

Março de 1998

PINHEIRO, PLÁCIDO ROGÉRIO

Uma Metodologia de Feixes e Benders Aplicado a um Problema Linear Inteiro de Grande Porte [Rio de Janeiro] 1998.

VIII, 82p., 29.7cm (COPPE/UFRJ. D.Sc., Engenharia de Sistemas e Computação, 1998)

Tese - Universidade Federal do Rio de Janeiro, COPPE.

1 - Programação Linear Inteira

2 - Relaxação Lagrangeana

3 - Método de Feixes

4 - "Cross Decomposition"

5- Decomposição de Benders

I. COPPE/UFRJ

II. Título (série)

**Dedicatória**

- À minha esposa Mirian pelo amor, carinho, companheirismo e compreensão. A quem muito devo, pela renúncia e paciência.
  
- Ao meu filho Pedro Gabriel que na turbulência de seus primeiros anos de vida, tanto se inquietou com minhas intermináveis horas de estudos.
  
- Aos meus pais Rodrigues e Ana Rosa, que me impulsionaram rumo a formação acadêmica.

## Agradecimentos

Agradeço à todos que durante este período, conviveram comigo e que de alguma forma contribuíram para realização deste ideal. Em especial:

Ao Prof. Paulo Roberto Oliveira, meu orientador, que soube com sua agradável convivência, sabedoria, competência, paciência e abnegação as causas científicas, encaminhar-me nesta jornada de lutas, renúncia e disciplina.

À Prof<sup>a</sup> Susana Scheimberg de Makler que em muito me ajudou com suas valorosas orientações.

Ao Prof. Nelson Maculan Filho com seu exemplo de pesquisador, um estímulo nesta caminhada.

Ao Prof. Antônio Clécio Fontelles Thomaz que prestou-me todo apoio e incentivo para cursar o doutorado.

Aos demais colegas, que em muito contribuíram a esta realização, cada um a sua maneira.

As secretárias Ana Paula, Cláudia e Rose sempre prestativos. À Maurício Nardone e Ricardo Arantes com a valiosa ajuda computacional.

A Universidade Estadual do Ceará(UECE), em especial ao Departamento de Matemática, e a Universidade de Fortaleza(UNIFOR) pelo apoio e suporte recebido.

A CAPES pela ajuda financeira.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para obtenção do grau de Doutor em Ciências (D.Sc.)

## Uma Metodologia de Feixes e Benders Aplicado a um Problema Linear Inteiro de Grande Porte

Plácido Rogério Pinheiro

Março, 1998

Orientador: Paulo Roberto Oliveira

Programa: Engenharia de Sistemas e Computação

Consideremos um problema linear inteiro de grande porte, caracterizado por possuir a seguinte estrutura: a matriz de restrições é esparsa, possuindo uma estrutura bloco-diagonal, juntamente com variáveis e restrições de acoplamento. Um dos grupos de restrições dificulta a aplicação do esquema de decomposição de Benders.

Propomos o seguinte esquema algoritmo. Uma relaxação Lagrangeana é feita sobre o citado conjunto de restrições. Apresentamos um processo heurístico para o cálculo do multiplicador através da resolução do problema dual, estruturado a partir do método de feixes. Em cada iteração do algoritmo, propomos uma decomposição de Benders onde são fornecidos cotas para o valor da função e um  $\epsilon$ -subgradiente.

Abstract of Thesis Presented to COPPE/UFRJ as a Partial Fulfillment of the Requirements for the Degree of Doutor of Science (D.Sc.)

A Methodology of Bundle and Benders Applied Large Mixed Linear Integer Problem

Plácido Rogério Pinheiro

March, 1998

Thesis Supervisor: Paulo Roberto Oliveira

Department: Computing and Systems Engineering

We consider a large mixed linear integer problem. The structure of the constraint matrix is sparse, with independent blocks, and coupling constraints and variables. One of the groups of constraints to make difficult the application of Benders scheme decomposition.

In this work we propose the following algorithm. A Lagrangian Relaxation is made on the mentioned set of constraints; we presented a process heuristic for the calculation of the multiplier through the resolution of the dual problem, structured starting from the method of bundle. For each iterations of that multiplier, we propose a Benders decomposition scheme where is supplied quotas for the value of the function and an  $\varepsilon$ -subgradient.

## SUMÁRIO

<b>INTRODUÇÃO</b> .....	1
<b>CAPÍTULO 1 - O Estado de Arte da “Cross Decomposition” e Principais Trabalhos em Métodos de Feixes</b> .....	4
<b>CAPÍTULO 2 - O Problema Primal</b> .....	14
2.1. Introdução.....	14
2.2. Notações.....	15
2.3. Os Problemas Linear, Inteiro e Misto.....	15
<b>CAPÍTULO 3 - Dualidade Lagrangeana e "Cross Decomposition"</b> .....	17
3.1. Introdução.....	17
3.2. Relaxação Lagrangeana.....	17
3.2.1. Relaxação Lagrangeana I.....	17
3.2.2. Relaxação Lagrangeana II.....	19
3.3. Condições de Otimalidade dos Problemas Duais.....	20
3.4. Técnicas de Decomposição.....	21
3.4.1. Decomposição de Dantzig-Wolfe.....	21
3.4.2 Decomposição de Benders.....	22
3.4.3 "Cross Decomposition".....	24
<b>CAPÍTULO 4 – Uma Aplicação do Método de Feixes</b> .....	26
4.1 Introdução.....	26
4.2 O Método de Feixes Aplicado à Relaxação II.....	26
4.3 Descrição do Algoritmo.....	30

4.3.1	Algoritmo de Feixes com Penalidade.....	32
4.4	Convergência do Método de Feixes.....	34
4.5	Recuperação das Variáveis Primais no Caso Linear Contínuo.....	37
4.6	Feixes x "Cross Decomposition".....	38
<b>CAPÍTULO 5</b>	<b>– Uma Metodologia de Regularização e Decomposição de     Benders.....</b>	<b>40</b>
5.1	Introdução.....	40
5.2	Decomposição de Benders para o Problema Relaxado I.....	40
5.3	Cotas.....	43
5.4	Regularização Quadrática do Problema Dual.....	43
5.5	O Algoritmo de Feixes Aproximado.....	50
5.5.1	O Algoritmo Parcial de Benders.....	50
5.5.2	Teste de Armijo Aproximado.....	51
5.5.3	Algoritmo de Regularização para Atualização dos Multiplicadores com a Relaxação I.....	53
5.5.4	Algoritmo Parcial de Benders para o Problema Linear Inteiro com a Relaxação I (Oráculo Aproximado .....	56
5.6	Sobre a Convergência.....	60
<b>CAPÍTULO 6</b>	<b>- Conclusões.....</b>	<b>62</b>
<b>CAPÍTULO 7</b>	<b>- Referências Bibliográficas.....</b>	<b>64</b>



## INTRODUÇÃO

O principal objetivo deste trabalho é o desenvolvimento de uma metodologia de decomposição que combine a decomposição de Benders e um processo heurístico para o cálculo dos multiplicadores aplicado a um problema relaxado para a resolução de um problema linear inteiro de grande porte do tipo

$$\text{mín } c^t x + d^t y$$

$$\text{s.a. } Ax + By \leq b$$

$$x \in X_R, y \in Y$$

onde  $X_R = \{x: Dx \leq d, x \geq 0\}$  e  $Y = \{y: Fy \leq f, y \geq 0, y \text{ inteiro}\}$

Relaxando-se parte das restrições tem-se o respectivo multiplicador atualizado por um processo heurístico, que resolve um modelo local do dual relaxado. A cada iteração, com o multiplicador obtido, aplicamos iterações da decomposição de Benders sobre o problema relaxado, obtendo-se um  $\epsilon$ -subgradiente e cotas inferior e superior da solução ótima. A título de comparação metodológica aplicaremos também a “Cross Decomposition” ([Van Roy. 83], [Holmberg. 90a] e [Holmberg. 92a]) ao problema linear inteiro.

A motivação deste trabalho surge de um problema linear inteiro de grande porte proveniente do planejamento de expansão da rede de transmissão digital de um sistema de telecomunicações de uma área

urbana equivalente a uma cidade com as dimensões do Rio de Janeiro [Moreno. 92].

O conteúdo desta tese seguirá a seguinte divisão:

Uma revisão bibliográfica dos principais trabalhos sobre o método de feixes e da "Cross Decomposition" é relatada no capítulo 1.

No capítulo 2 apresentamos os problemas nas versões de programação linear, linear inteira e linear inteira mista.

O capítulo 3 será dedicado à apresentação das relaxações lagrangeanas dos problemas. Abordamos as técnicas clássicas de decomposição de Dantzig-Wolfe, Benders e "Cross Decomposition" aplicada ao problema linear inteiro.

Através de uma proposta de relaxação e considerando o problema dual, descreveremos o método de feixes, conforme a abordagem de [Lemaréchal. 92], [Hiriart-Urruty and Lemaréchal. 93] e [Lemaréchal and Sagastizábal. 95]. Este é o tema do capítulo 4.

No capítulo 5, consideraremos uma outra relaxação do problema original. Apresentamos um processo heurístico para o cálculo do multiplicador, estruturado a partir do método de feixe. No lugar do oráculo, com o valor da função e um subgradiente, são fornecidos cotas para a função e um  $\varepsilon$ -subgradiente. Isto é feito através do uso de algumas iterações de Benders, relativas ao particionamento primal do problema de programação linear inteiro relaxado. Enfatizaremos também algumas propriedades de convergência.

Finalmente apresentaremos algumas conclusões no capítulo 6, visando à continuidade da pesquisa.

## CAPÍTULO 1

### ESTADO DE ARTE DA “CROSS DECOMPOSITION” E PRINCIPAIS TRABALHOS EM MÉTODO DE FEIXES

#### **Introdução: Relaxação Lagrangeana**

A programação linear inteira é uma eficiente ferramenta na modelagem e solução de um grande número de problemas. São eles provenientes das mais diversas áreas, tais como recolhimento de lixo, planejamento de sistemas de telecomunicações, cortes em uma, duas ou três dimensões, etc.

Os problemas de programação linear inteiro de grande porte são freqüentemente impossíveis de serem resolvidos diretamente através de softwares comerciais. Em tais casos a relaxação lagrangeana, combinada com a otimização por subgradientes é muitas vezes usada para achar cotas inferiores para o valor ótimo da função objetivo. Estas cotas podem ser usados, por exemplo, no método de Branch - Bound [Salkin and Mathur. 89], ou apenas para medir a qualidade das soluções viáveis. Tais propriedades encontram-se atualmente incorporadas em softwares comerciais como [CPLEX3.0 97], [OSL1.2 90], conforme descrito em MINTO [Nemhauser et alii. 94]. Outras estratégias também são consideradas: obtenção de cotas superiores, uma maior eficiência nas rotinas sobre a geração de cortes além do

uso do processamento paralelo, veja por exemplo [Nemhauser. 94] para o primeiro e [Lee. 94] para o último.

A relaxação lagrangeana foi utilizada por [Held and Karp. 70] e [Held and Karp. 71] com seus trabalhos sobre problemas do caixeiro viajante; os métodos de Branch-and-Bound e enumeração implícita tiveram considerável ganho em [Geoffrion 74] com a relaxação lagrangeana; em [Fisher. 81], encontramos várias questões norteadoras sobre a relaxação lagrangeana nos problemas lineares inteiros, dentre estas como calcular os multiplicadores de Lagrange, de que maneira escolher dentre as várias relaxações do problema e como obter soluções viáveis para o problema primal. Recentes técnicas para resolver o dual lagrangeano relaxado de problemas de otimização combinatória, em um tempo polinomial, utilizando como subrotina o algoritmo dos Elipsóides [Khachian. 79] ou o algoritmo de Vaidya [Vaidya. 90], tem sido apresentadas desde 1994, com [Bertsimas and Orlin. 94].

Outras metodologias de decomposição usam heurísticas lagrangeanas, combinando a solução do dual lagrangeano, através do método de subgradientes, juntamente com soluções primais viáveis heurísticas. Estas técnicas foram aplicadas a problemas de fluxo em redes “multicommodity” em [Holmberg. 96a], [Holmberg and Yuan.

96b], [Holmberg and Hellstrand. 96c], a problemas de localização capacitada em [Holmberg and Ling. 97b] e [Holmberg et alii. 97].

### **Decomposição de Benders**

Para problemas de programação linear inteiro, a decomposição de Benders[Benders. 62] é um método exato e finito. Ela é eficiente quando o número de variáveis inteiras é muito menor do que o número de variáveis contínuas pois neste caso o problema mestre tem dimensão bem menor do que o problema original. Entretanto para problemas de grande porte, o problema mestre de Benders pode ser de difícil solução, devido à elevada dimensão. Junta-se a isto a velocidade de convergência, em geral lenta, tornando este método, em muitos casos ineficiente [Wong. 78]. Além disso, experiências computacionais têm mostrado que um código geral de Branch-Bound aplicado para resolver o problema mestre de Benders produz freqüentemente uma árvore muito maior do que para resolver o problema original. Assim sendo, a desvantagem desta decomposição é muitas vezes a dificuldade de resolver o problema mestre, tornando-a ineficiente. Foram apresentados vários trabalhos com o objetivo de resolver o problema mestre aproximadamente com uma maior eficiência global. Dentre estes [Aardal and Larsson. 90], [Cote and Laughton. 84], [Fisher and Jaikumar. 78], [Hoc. 82], [Paula Jr. and Maculan. 88] e [Rana and Vickson. 88] que utilizam a decomposição de Benders com a relaxação

lagrangeana aplicada aos cortes no problema mestre. Este é então resolvido aproximadamente com o método de subgradientes [Poljak. 1967], [Poljak. 1969], [Held et alii. 74] e [Shor. 85]. Isto transfere a dificuldade do problema mestre para um muito mais fácil, ou seja, calcular iterativamente o máximo da função dual. Isto não soluciona entretanto todas as dificuldades. Em [Ferland and Florian. 79] e [Holmberg. 94b] este método é recusado devido à falta de controlabilidade (a solução ótima no problema mestre de Benders pode jamais alcançar o ótimo no problema mestre relaxado) na solução do problema mestre relaxado. Existem também sugestões de como obter um bom conjunto inicial de cortes para o problema mestre de Benders [Minoux. 84 e 86]. [McDaniel and Devine. 77] sugerem o uso da relaxação linear para o problema mestre de Benders em um número inicial de iterações.

### **“Cross Decomposition”**

Motivados por estes insucessos, [Van Roy. 80 e 83] desenvolveu a “Cross Decomposition” explorando simultaneamente as estruturas dos problemas primal e dual, combinando as vantagens das decomposições de Dantzig-Wolfe [Dantzig and Wolfe. 60] e Benders [Benders. 62]. [Holmberg. 94b] realiza um estudo comparativo de diversas aproximações do problema mestre de Benders, apresentando como um método eficiente na solução de um problema linear inteiro a “Cross

Decomposition” [Van Roy. 80 e 83], [Holmberg. 90], [Holmberg. 92a] e [Holmberg. 94a] .

Aspectos teóricos da decomposição de Benders juntamente com a “Cross Decomposition” são também analisados em [Holmberg. 89 e 90a].

Modificações na “Cross Decomposition” para problemas de programação linear e inteira foram feitas por [Holmberg. 97a e 92b]. Estas modificações são efetuadas através da generalização do método de Kornai and Liptak [Kornai and Liptak. 65], em que se elimina a necessidade do uso dos problemas mestre primal e dual. A dinâmica desta decomposição está nos subproblemas, onde se itera os subproblemas primal e dual. Em vez de se utilizar a solução do último subproblema como entrada para o outro, se usa a média de todas as soluções anteriores. A prova da convergência desta metodologia encontra-se em [Holmberg. 94e].

A “Cross Decomposition” aplicada a problemas que tenham uma parte linear possui convergência finita. Aplicações para esta classe especial são: Problemas de Programação Linear, Programação Inteira Mista e Programação Não-Linear Convexa com Restrições Lineares. A decomposição generalizada de Benders [Geoffrion. 72] é uma estrutura na qual podemos também tratar problemas de Programação Não-Linear com Restrições Convexas e Problemas de Programação Inteira. Nela, o



subproblema primal não é de programação linear, entretanto ele necessita ser convexo para que não haja salto de dualidade. A mesma generalização também é possível para “Cross Decomposition” [Holmberg. 90a], onde se conclui que: “A “Cross Decomposition” generalizada tem convergência finita em problemas (Lipschitzianos) para os quais o algoritmo de decomposição de Dantzig-Wolfe ou o algoritmo da decomposição generalizada de Benders tem convergência finita.

[Van Roy. 86] aplicou a “Cross Decomposition” a problemas de localização capacitada com resultados computacionais dos mais eficientes. Problemas de localização capacitada com custos fixos também apresentaram bons resultados computacionais para problemas de pequeno porte, quando abordados pela metodologia da “Cross Decomposition” [Holmberg. 94c]. Para uma certa classe de problemas em localização de pequeno porte são apresentados métodos de solução exata estruturados a partir da “Cross Decomposition” em [Holmberg and Jörnsten. 91a]. Outros problemas de localização capacitada generalizada (multitipo e multiproduto), juntamente com resultados computacionais satisfatórios e encorajadores, foram abordados através da “Cross Decomposition” em [Lee Y. 93]. Em [Holmberg. 90] são apresentados resultados de experimentos computacionais com métodos de solução dual, primal e primal-dual para problemas de compra e venda, em problemas de programação linear estruturados de grande

porte. Nesta aplicação, Holmberg considerou a “Cross Decomposition” com Benders/ Dantzig-Wolfe e também a generalização de Kornai and Liptak. Ambas metodologias foram também aplicadas a problemas de planejamento organizacional [Holmberg. 94d]. Uma comparação das técnicas de Kornai and Liptak e Cross Decomposition para problemas lineares com estruturas bloco-angular, bem como resultados computacionais estão analisados em [Aadal and Ari. 90]. Um algoritmo simplificado da “Cross Decomposition” para múltipla escolha com o lado direito das restrições foi apresentado por [Kim et alii. 89]. Aplicações envolvendo problemas de transporte estocástico foram abordados em [Holmberg and Jörnsten. 84a], com estudo comparativo envolvendo outros métodos. Com a “Cross Decomposition” [Holmberg and Jörnsten. 93] solucionou Problemas Generalizados da Mochila com coeficientes variáveis.

### **Método de Feixes**

A atualização dos multiplicadores pode ser feita por vários métodos. Se formulado como um problema linear, o simplex é tradicionalmente utilizado. Por outro lado tem-se, em geral, um dual não-diferenciável e a abordagem clássica é o método de subgradientes ([Poljak. 67], [Poljak. 69] e [Shor. 85]), que se sabe não ser um método de descida. Uma alternativa aparentemente ausente na literatura ocidental é o uso das técnicas de dilatação de Shor [Shor. 70], onde se garante a monotonicidade dos valores iterados da função

objetivo. Embora mais complexas, as técnicas de feixes desenvolvidas originariamente por [Lemaréchal. 75] e [Wolfe. 75] vêm sendo cada vez mais utilizadas. Sobre elas nos deteremos a seguir.

O método de feixes explora os dados das iterações anteriores, vetores iterados, função objetivo, e subgradientes - o **feixe** de informações, para produzir a nova iteração. Teve como trabalho pioneiro o método de  $\varepsilon$ -descida [Lemaréchal. 76] que tem como modelo o método de subgradientes conjugados da programação diferenciável [Lemaréchal. 75] e [Wolfe. 75]. Kiwiel, em seu livro [Kiwiel. 85], apresenta uma nova visão sobre o método de feixes fundamentada nos métodos dos planos cortantes clássicos desenvolvidos por [Kelley. 60] e [Cheney and Goldstein. 59]. A idéia básica da generalização dos planos cortantes é de acrescentar uma regularização quadrática à aproximação linear convexa por partes para a função objetivo, esta usando as linearizações geradas por subgradientes. Para evitar um feixe de grande dimensão, é necessário limitá-lo. Kiwiel, por exemplo, apresentou uma estratégia de seleção de subgradientes baseada nos multiplicadores associados ao modelo local. Com isto o feixe se mantém em  $n + 2$  subgradientes,  $n$  sendo a dimensão da variável do problema.

Considera-se três abordagens para especificar o processo de estabilização quadrática, que são essencialmente equivalentes. A primeira utiliza a técnica de regiões de confiança, veja [Schramm and

Zowe. 87], [Schramm. 89]. A regularização de Moreau-Yosida gera o método proximal usado por [Kiwiel. 90]. O método de nível [Lemaréchal et alii. 91] e [Kiwiel. 92] é a terceira técnica de estabilização. Todos podem ser vistos no livro [Hiriart-Urruty and Lemaréchal. 93]. Uma moderna síntese das técnicas de feixes e uso de métrica variável é feita a partir do conceito de regularização de Moreau-Yosida em [Lemaréchal and Sagastizábal. 97] e [Qi and Chen. 97]. Aplicações em problemas de controle, envolvendo o método de feixes podem ser encontradas em [Mäkelä and Neittaanmäki. 92] e outras aplicações utilizando decomposição lagrangeana, redes e testes comparativos com outros algoritmos estão desenvolvidas em [Schramm and Zowe. 92]; decomposições de grande porte e otimização paralela em [Medhi. 91]; problemas de roteamento envolvendo janela de tempo em [Kohl et alii. 95]; otimização da produção diária em uma rede elétrica [Lemaréchal and Sagastizábal. 95].

Os bem conhecidos códigos de Fortran M1FC1 [Lemaréchal and Bancora. 85] e M2FC2 desenvolvidos por Lemaréchal empregam a metodologia da  $\varepsilon$ -descida. Mais recentemente, usando a regularização proximal, tem-se os códigos em Fortran BT, BTNC, BTNCBC, BTCLC, BTNCLC e NOA desta última metodologia encontram-se em [Oustrata. 91] e [Kiwiel and Stachurski. 88].

Segundo Lemaréchal em [SIAG/OPT. 94] “não é exagero afirmar que 90 por cento das aplicações da não-diferenciabilidade aparecem

nas decomposições, de uma forma ou de outra, enquanto que os 10 por cento restante mostra-se via o cálculo de autovalores”.

Citamos ainda [Lemaréchal. 96] quando afirma que “a otimização não-diferenciável tem como maior deficiência a velocidade de convergência”.

## CAPÍTULO 2

### O PROBLEMA PRIMAL

#### 2.1 Introdução

Consideremos o problema linear inteiro (P), motivado por uma aplicação em um sistema urbano de telecomunicações [Moreno et alii. 92].

$$(P)[1] \quad v_P = \text{mín} \sum_{k=1}^4 c_k^t x_k + \sum_{j=1}^2 e_j^t y_j$$

$$\text{s.a} \quad D_k x_k = d_k, \quad k = 1, \dots, 4$$

$$F_j y_j \leq f_j, \quad j = 1, 2$$

$$A_k x_k + B_k y_1 = K_k, \quad k = 1, \dots, 4$$

$$C_3 x_3 + C_4 x_4 + C_2 y_2 = K_5$$

$$x_k \geq 0 \text{ reais}, y_j \geq 0 \text{ inteiros}, k = 1, \dots, 4 \text{ e } j = 1, 2$$

onde as matrizes  $A_k$ ,  $B_k$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $D_k$ ,  $F_j$  têm dimensões adequadas com os vetores  $c_k$ ,  $d_k$ ,  $e_j$ ,  $f_j$ ,  $K_k$ ,  $K_5$ ,  $x_k$  e  $y_j$  envolvidos.

Para efeito de apresentação e generalização dos resultados, consideraremos inicialmente todas as variáveis inteiras.

Sejam  $X = \prod_{k=1}^4 X_k$  onde  $X_k = \{x_k; D_k x_k - d_k = 0 \wedge x_k \geq 0, x_k$

inteiros} e  $Y = \prod_{j=1}^2 Y_j$  onde  $Y_j = \{y_j; F_j y_j - f_j \leq 0 \wedge y_j \geq 0, y_j \text{ inteiros}\}$ ,

supostos não-vazios e limitados, isto é, finitos. Nós também assumimos que para quaisquer  $\hat{y}_j \in Y_j$   $j = 1, 2$  existem  $\hat{x}_k \in X_k$   $k = 1, \dots, 4$  tais que os blocos de restrições acopladas sejam satisfeitos.

## 2.2. Notações

Consideremos as seguintes notações: se  $(. .)$  é um problema de otimização então  $v(. .)$  é o valor ótimo,  $RV(. .)$  região viável do problema  $(. .)$ ,  $\text{Argmín}\{(. .)\}$  denota o conjunto de todos os pontos de mínimo da função objetivo sobre o conjunto de restrições do  $(. .)$ . Façamos  $x = (x_k)_{k=1, \dots, 4}$  e  $y = (y_j)_{j=1, 2}$ .

## 2.3. Os Problemas Linear, Inteiro e Misto

A formulação inteira de (P)[1] se escreve:

$$(PI)[2] \quad v_{PI} = \min \sum_{k=1}^4 c_k^1 x_k + \sum_{j=1}^2 e_j^1 y_j$$

$$\text{s.a.} \quad A_k x_k + B_k y_1 = K_k, \quad k = 1, \dots, 4$$

$$C_3 x_3 + C_4 x_4 + C_2 y_2 = K_5$$

$$x \in X, y \in Y$$

Uma relaxação contínua da variável  $x$  de (PI)[2] gera (P)[1] que se reescreve:

$$\begin{aligned}
 \text{(P)[1]} \quad v_P &= \min \sum_{k=1}^4 c_k^t x_k + \sum_{j=1}^2 e_j^t y_j \\
 \text{s.a.} \quad &A_k x_k + B_k y_1 = K_k, \quad k = 1, \dots, 4 \\
 &C_3 x_3 + C_4 x_4 + C_2 y_2 = K_5 \\
 &x \in X_R, \quad y \in Y
 \end{aligned}$$

sendo  $X_R = \prod_{k=1}^4 X_{(k)R}$  onde  $X_{(k)R} = \{x_k; D_k x_k - d_k = 0 \wedge x_k \geq 0\}$ .

Ao relaxarmos também a variável  $y$  de (P)[1] se obtém

$$\begin{aligned}
 \text{(PR)[3]} \quad v_{PR} &= \min \sum_{k=1}^4 c_k^t x_k + \sum_{j=1}^2 e_j^t y_j \\
 \text{s.a.} \quad &A_k x_k + B_k y_1 = K_k, \quad k = 1, \dots, 4 \\
 &C_3 x_3 + C_4 x_4 + C_2 y_2 = K_5 \\
 &x \in X_R, \quad y \in Y_R
 \end{aligned}$$

onde  $Y_R = \prod_{j=1}^2 Y_{(j)R}$ ,  $Y_{(j)R} = \{y_j; F_j y_j - f_j \leq 0 \wedge y_j \geq 0\}$  sendo  $X_R$  e  $Y_R$  não-vazios e limitados.

**Proposição 2.1:**  $RV(PR) \supseteq RV(P) \supseteq RV(PI)$

$$v_{PR} \leq v_P \leq v_{PI}$$



## CAPÍTULO 3

### DUALIDADE LAGRANGEANA E "CROSS DECOMPOSITION"

#### 3.1. Introdução

Com o objetivo de apresentar uma cota inferior para o custo ótimo, desenvolveremos um particionamento estruturado a partir da relaxação lagrangeana do problema.

#### 3.2. Relaxação Lagrangeana

Apresentamos uma relaxação lagrangeana referente ao último bloco de restrições de (P)[1] ou (PI)[2] e uma relaxação lagrangeana dos blocos de restrições de (P)[1] ou (PI)[2].

##### 3.2.1. Relaxação Lagrangeana I

Relaxando-se o último bloco de restrições de (P)[1], obtemos o que denominaremos de **relaxação I**; tem-se o dual

$$(D_I)[4] \quad v_{D_I} = \max_{\lambda} \varphi(\lambda)$$

onde,  $\forall \lambda$  se define a função dual

$$\begin{aligned}
 (\varphi)[5] \quad \varphi(\lambda) = \min_{(x, y) \in W_I} \mathcal{L}(x, y, \lambda) = \min \sum_{k=1}^4 c_k^t x_k + \sum_{j=1}^2 e_j^t y_j + \lambda^t (C_3 x_3 + C_4 x_4 \\
 + C_2 y_2 - K_5) \text{ s.a } (x, y) \in W_I \text{ onde } W_I = \{(x, y); x \in X_R, y \in Y, \\
 A_k x_k + B_k y_1 = K_k, k = 1, \dots, 4\}.
 \end{aligned}$$

O objetivo desta relaxação é garantir a separabilidade dos blocos  $x_3$  e  $x_4$  de variáveis, em relação a  $y_2$ , visando em seguida a aplicação da decomposição de Benders. Nesta serão separados os blocos  $x_k$ ,  $k = 1, \dots, 4$ , de  $y_1$ .

Similarmente podemos definir outras duas relaxações, a primeira, substituindo-se  $X_R$  por  $X$ , à qual corresponde o dual que denominamos  $DI_I$ , com função dual  $\varphi_I(\lambda)$ , e, em seguida  $Y$  por  $Y_R$ , em correspondência ao dual  $DR_I$ , cuja função dual é  $\varphi_R$ . Com estas notações respectivas obtemos a seguinte proposição da dualidade fraca.

**Proposição 3.1:**  $\varphi_R(\lambda) \leq \varphi(\lambda) \leq \varphi_I(\lambda), \quad \forall \lambda$

$$v_{DR_I} \leq v_{D_I} \leq v_{DI_I}$$

$$\varphi_I(\lambda) \leq v_{PI}, \quad \varphi_R(\lambda) \leq v_{PR} \text{ e } \varphi(\lambda) \leq v_P, \quad \forall \lambda$$

$$v_{DI_I} \leq v_{PI} \text{ e } v_{D_I} \leq v_P$$

### 3.2.2. Relaxação Lagrangeana II

Com o objetivo de separar todos os blocos de variáveis  $x \in X_R$  e  $y \in Y$  consideremos a relaxação dos blocos de restrições em (P)[1], (PI)[2] ou (PR)[3]. Denominaremos de **relaxação II**.

Com  $\mu = (\mu_k)_{k=1, \dots, 5}$ , consideremos o dual do problema linear inteiro

(P)[1]

$$(D_{II})[6] \quad v_{D_{II}} = \max_{\mu} \theta(\mu)$$

onde,  $\forall \mu$  se define a função dual

$$(\theta)[7] \quad \theta(\mu) = \min_{(x, y) \in W_{II}} \mathcal{L}(x, y, \mu) = \min \sum_{k=1}^4 c_k^t x_k + \sum_{j=1}^2 e_j^t y_j + \sum_{k=1}^4 \mu_k^t (A_k x_k +$$

$$B_k y_1 - K_k) + \mu_5^t (C_3 x_3 + C_4 x_4 + C_2 y_2 - K_5) \text{ s.a. } (x, y) \in W_{II} \text{ onde}$$

$$W_{II} = \{(x, y); x \in X_R \text{ e } y \in Y\}.$$

Como precedentemente, relacionamos aos problemas (PI)[2] e (PR)[3], respectivamente a  $DI_{II}$  com função dual  $\theta_I(\mu)$  e  $DR_{II}$ , cuja função dual é  $\theta_R(\mu)$ ; tem-se

**Proposição 3.2:**  $\theta_R(\mu) \leq \theta(\mu) \leq \theta_I(\mu) \quad \forall \mu,$

$$v_{DR_{II}} \leq v_{D_{II}} \leq v_P \leq v_{PI},$$

$$v_{DR_{II}} \leq v_{D_{II}} \leq v_{DI_{II}} \leq v_{PI}.$$

Observação: Os problemas linear, inteiro e misto relaxados  $\theta_I$ ,  $\theta$  e  $\theta_R$  são separáveis nas variáveis  $x$  e  $y$ .

### 3.3 Condições de Otimalidade dos Problemas Duais

As condições de otimalidade a seguir são encontradas em [Rockafellar. 72], parágrafo 28, teorema 28.3, pág. 281.

A função  $(-\varphi(\lambda))$  é semi-contínua inferiormente, por ser o máximo de funções lineares. Admitindo-se que  $W_I$  é compacto, tem-se que o subdiferencial de  $\varphi(\lambda)$ [5] é dado por

$$[8] \quad \partial\varphi(\lambda) = \left\{ \sum_r \alpha_r g^r : \alpha_r \geq 0, \sum_r \alpha_r = 1, g^r = C_3 x_3^r + C_4 x_4^r + C_2 y_2^r - K_5, \right.$$

$\mathcal{L}(x^r, y^r, \lambda) = \varphi(\lambda) \}$  ver [Lemaréchal. 92].

Como a função objetivo do problema dual é côncava,  $\lambda^*$  é uma solução ótima para  $(D_I)$ [4] se e somente se

$$0 \in \partial\varphi(\lambda^*)$$

Analogamente, para a segunda relaxação se tem, no lugar de  $g^r$ , acima, o subgradiente:

$$[9] \quad h^i(\mu) = h^i = \begin{pmatrix} A_1 x_1^i + B_1 y_1^i - K_1 \\ \vdots \\ A_4 x_4^i + B_4 y_1^i - K_4 \\ C_3 x_3^i + C_4 x_4^i + C_2 y_2^i - K_5 \end{pmatrix}$$

sendo  $\mathcal{L}(x^i, y^i, \mu) = \theta(\mu)$

### 3.4 Técnicas de Decomposição

Objetivando resolver o problema linear inteiro consideremos as seguintes metodologias de decomposição, aplicadas a (P)[1].

#### 3.4.1. Decomposição de Dantzig-Wolfe

Inicialmente abordaremos a metodologia de Dantzig-Wolfe aplicada à solução do problema (D<sub>II</sub>)[6].

Para  $\mu$  fixo, explicitando  $(\theta)$ [7] tem-se o **subproblema dual**

$$\begin{aligned}
 (\theta)[10] \quad v_\theta = \min & \sum_{k=1}^4 c_k^t x_k + \sum_{j=1}^2 e_j^t y_j + \sum_{k=1}^4 \mu_k^t (A_k x_k + B_k y_1 - K_k) + \mu_5^t (C_3 x_3 + \\
 & C_4 x_4 + C_2 y_2 - K_5) \\
 \text{s.a.} \quad & D_k x_k = d_k, \quad k = 1, \dots, 4 \\
 & F_j y_j = f_j \quad j = 1, 2 \\
 & x \geq 0 \text{ real, } y \geq 0 \text{ inteiro}
 \end{aligned}$$

Seja  $X = \{(x, y); D_k x_k = d_k, F_j y_j \leq f_j, x_k \geq 0 \text{ reais, } y_j \geq 0 \text{ inteiros, } k = 1, \dots, 4 \text{ e } j = 1, 2 \text{ inteiros}\}$  um conjunto limitado e sejam  $\{(x^p, y^p)\}$   $\forall p \in P_X$  (finito) um conjunto dos pontos extremos de  $X$ . Assim sendo,  $(\theta)$ [7] é equivalente a

$$\min_{p \in P_X} \sum_{k=1}^4 c_k^t x_k^p + \sum_{j=1}^2 e_j^t y_j^p + \sum_{k=1}^4 \mu_k^t (A_k x_k^p + B_k y_1^p - K_k) + \mu_5^t (C_3 x_3^p + C_4 x_4^p + C_2 y_2^p - K_5)$$

O problema mestre de Dantzig-Wolfe(DM)[11] que incluirá um subconjunto  $P'_X \subseteq P_X$  é da forma

$$(DM)[11] \quad v_{DM} = \text{máx } \Delta$$

$$\text{s.a. } \Delta \leq \sum_{k=1}^4 c_k^t x_k^p + \sum_{j=1}^2 e_j^t y_j^p + \sum_{k=1}^4 \mu_k^t (A_k x_k^p + B_k y_1^p - K_k) +$$

$$\mu_5^t (C_3 x_3^p + C_4 x_4^p + C_2 y_2^p - K_5)$$

$$\Delta \in \mathbb{R}, \mu \text{ real}, \forall p \in P'_X$$

O algoritmo da Decomposição de Dantzig-Wolfe[Dantzig and Wolfe. 60] é um processo iterativo entre o subproblema dual  $(\theta)[10]$  e o problema mestre de Dantzig-Wolfe (DM)[11]. (DM)[11] fornece um  $\mu$  para  $(\theta)[10]$  e  $(\theta)[10]$  gera pontos extremos  $(x, y)$  e um novo corte para (DM)[11]. A cada iteração,  $P'_X$  cresce com  $v_\theta \leq v_{D_{\Pi}}$ , para  $\mu$  fixo e  $v_{DM} \geq v_{D_{\Pi}}$ . Após um número finito de iterações o algoritmo de Dantzig-Wolfe[Dantzig. 63] para, com  $v_\theta = v_{DM} = v_{D_{\Pi}}$ .

### 3.4.2. Decomposição de Benders

O problema (P)[1] pode ser reescrito como

$$\min_{y \in Q} \left\{ \sum_{j=1}^2 e_j^t y_j + \min_{k=1}^4 \left\{ \sum_{k=1}^4 c_k^t x_k; A_k x_k = K_k - B_k y_1, C_3 x_3 + C_4 x_4 = K_5 - C_2 y_2 \right. \right.$$

$k = 1, \dots, 4, x \in X_R \} \} \text{ onde}$

$$Q = \{y \in Y; \exists x \in X_R \text{ tal que } A_k x_k = K_k - B_k y_1, C_3 x_3 + C_4 x_4 = K_5 - C_2 y_2$$

$k = 1, \dots, 4\}$ , que supomos não vazio.

Para  $y \in Q$  fixo, tem-se o subproblema de minimização interna

(com  $X_R$  explicitado):

$$\begin{aligned}
 \text{(P1)[12]} \quad v_P &= \min \sum_{k=1}^4 c_k^t x_k \\
 \text{s.a.} \quad D_k x_k &= d_k \quad k = 1, \dots, 4 \\
 A_k x_k &= K_k - B_k y_1 \quad k = 1, \dots, 4 \\
 C_3 x_3 + C_4 x_4 &= K_5 - C_2 y_2 \\
 x_k &\geq 0 \quad k = 1, \dots, 4
 \end{aligned}$$

Seu dual é dado por

$$\begin{aligned}
 \text{(DB)[13]} \quad v_{DB} &= \max_{(v, u, w)} \sum_{i=1}^4 d_i^t v_i + \sum_{i=1}^4 (K_i - B_i y_1)^t \mu_i + (K_5 - C_2 y_2)^t \mu_5 \\
 \text{s.a.} \quad D_i^t v_i + A_i^t \mu_i &\leq c_i \quad i = 1, 2 \\
 D_i^t v_i + A_i^t \mu_i + C_i^t \mu_5 &\leq c_i \quad i = 3, 4
 \end{aligned}$$

onde  $\mu = (\mu_1, \dots, \mu_5)^t$ ,  $v = (v_1, \dots, v_4)^t$

Suponhamos que o poliedro

$$V = \{(v, \mu); D_i^t v_i + A_i^t \mu_i \leq c_i, \quad i = 1, 2 \text{ e } D_i^t v_i + A_i^t \mu_i + C_i^t \mu_5 \leq c_i, \quad i = 3, 4\}$$

é limitado. Assim sendo podemos definir o conjunto  $\{(v^p, \mu^p)\}$

$\forall p \in P_V$  (finito) de pontos extremos de  $V$ . Neste caso, (P)[1] é equivalente a

$$\text{(P)[1]} \quad \min_{y \in Q} \left\{ \sum_{j=1}^2 e_j^t y_j + \max_{p \in P_V} \left\{ \sum_{i=1}^4 d_i^t v_i^p + \sum_{i=1}^4 (K_i - B_i y_1)^t \mu_i^p + (K_5 - C_2 y_2)^t \mu_5^p \right\} \right\}$$

Para algum subconjunto  $P'_V \subseteq P_V$ , tem-se o chamado problema mestre de Benders.

$$(MB)[14] \quad v_{MB} = \text{mín } \Gamma$$

$$\text{s.a. } \Gamma \geq \sum_{j=1}^2 e_j^t y_j + \sum_{i=1}^4 d_i^t v_i^p + \sum_{i=1}^4 (K_i - B_i y_1)^t \mu_i^p + (K_5 - C_2 y_2)^t \mu_5^p$$

$$\Gamma \in \mathbb{R}, y \in Y \text{ e } \forall p \in P'_V$$

O algoritmo da Decomposição de Benders[Benders. 62] é um processo iterativo entre o dual do subproblema primal (DB)[13] e o problema mestre de Benders(MB)[14]. (MB)[14] produz um  $y$  para (DB)[13], este gera um ponto extremo  $(v, \mu)$  de  $V$  e um novo corte para (MB)[14]. A cada iteração,  $P'_V$  cresce com  $v_{DB} \geq v_p$  e  $v_{MB} \leq v_p$ . O algoritmo converge em um número finito de passos, com  $v_{DB} = v_{MB} = v_p$ .

### 3.4.3. “Cross Decomposition”

O algoritmo da “Cross Decomposition” [Van Roy. 83] e [Homberg. 85] consiste de uma fase de subproblemas, onde iteramos o subproblema dual  $(\theta)[10]$  e o dual do subproblema primal (DB)[13].  $(\theta)[10]$  gera um  $y$  para (DB)[13] e uma cota inferior  $v_\theta \leq v_p$ , (DB)[13] gera  $\mu$  para  $(\theta)[10]$  e uma cota superior  $v_{DB} \geq v_p$ . Os subproblemas  $(\theta)[10]$  e (DB)[13] também geram cortes para os problemas mestre de Dantzig-Wolfe(DM)[11] e Benders(MB)[14], respectivamente. Lembramos que as cotas não têm, em geral, um comportamento monótono entre iterações sucessivas. Assim, os problemas mestres serão utilizados quando não houver uma diminuição da cota superior



correspondente a iteração atual, ou, respectivamente, um aumento da cota inferior. Após isto, retornamos ao subproblemas.

Mais precisamente, temos os seguintes testes de convergência primal(TCP) e dual (TCD). Sejam  $v^+$  e  $v^-$  as melhores cotas superior e inferior sobre  $v_p$ . Por aperfeiçoar, significa aperfeiçoar as cotas superior e inferior ou gerar um novo e não determinado corte.

(TCP): Se  $\sum_{i=1}^4 d_i^t v_i^p + \sum_{i=1}^4 (K_i - B_i y_1)^t \mu_i^p + (K_5 - C_2 y_2)^t \mu_5^p < v^+$ , isto é, se

$v_{DB} < v^+$ , então  $y$  é uma solução aperfeiçoada em (DB)[13]. Se não, use o problema mestre de Benders, cuja solução  $y$  é levada a (DB)[13], retornando ao processo iterativo entre os subproblemas ( $\theta$ )[10] e (DB)[13].

(TCD): Se  $\sum_{k=1}^4 c_k^t x_k^p + \sum_{j=1}^2 e_j^t y_j^p + \sum_{k=1}^4 \mu_k^t (A_k x_k^p + B_k y_1^p - K_k) + \mu_5^t (C_3 x_3^p +$

$C_4 x_4^p + C_2 y_2^p - K_5) > v^-$ , isto é, se  $v_D > v^-$ , então  $\mu$  é uma solução aperfeiçoada em ( $\theta$ )[10]. Se não, use o problema mestre de Dantzig-Wolfe. Analogamente acima, a solução  $\mu$  é levada a ( $\theta$ )[10].

O algoritmo da “Cross Decomposition” resolve (P)[1] em um número finito de passos [Homberg. 90a].

## CAPÍTULO 4

### UMA APLICAÇÃO DO MÉTODO DE FEIXES

#### 4.1. Introdução

Como veremos, não é possível aplicar diretamente o método de feixes à metodologia que desenvolveremos junto à relaxação I, com a decomposição de Benders, porém ele é a motivação básica para o nosso método. Por isto, a título de ilustração (similar a [Lemaréchal. 92] e [Lemaréchal and Sagastizábal. 95]) aplicaremos o método de feixes na resolução do dual correspondente à relaxação II.

Procedimentos similares também são encontrados em [Schramm and Zowe. 92], veja também [Schramm and Zowe. 87], [Schramm. 89], [Oustrata et alii. 91] e [Kiwiel. 95].

#### 4.2. Método de Feixes Aplicado à Relaxação II

Passemos ao método de feixes aplicado à resolução de  $(D_{II})$ [6]. Suponha, na iteração  $n$ , conhecidos os multiplicadores (variáveis do problema dual)  $\mu^i$ ,  $i = 1, \dots, n$ , e respectivos subgradientes  $h^i \in \partial\theta(\mu^i)$ . Considere um conjunto  $\{1, \dots, \ell\}$  de índices correspondentes aos elementos incorporados ao feixe. Seja  $\{\pi^i\}$ ,  $i \in \{1, \dots, \ell\}$  uma coleção de pontos auxiliares. Estes pontos são os efetivamente gerados pelo algoritmo, e que serão ou não aceitos como iteração em função de um teste

de subida do tipo Armijo([Armijo. 66] e [Pchénitchny and Daniline. 65]). Acompanharemos a formulação usual da metodologia de feixes (veja [Hiriart-Urruty and Lemaréchal. 93] v. II) e definiremos o problema regularizado, que é uma aproximação local de  $(D_{II})$ [6], dado por

$$(F_{II})[15] \quad \max_{(z, \pi)} z - \frac{1}{2t_n} \|\pi - \mu^n\|^2$$

$$\text{s.a } z \leq (h^i)^t(\pi - \mu^n) + e_i + \theta(\mu^n), \quad i \in \{1, \dots, \ell\}$$

onde  $e_i$  é o erro de linearização dado por

$$e_i := e(\mu^n, \pi^i, h^i) := \theta(\pi^i) - \theta(\mu^n) + (h^i)^t(\mu^n - \pi^i),$$

e  $t_n > 0$ , o qual determina o tamanho da direção  $\pi - \mu^n$  e é escolhido através de um determinado critério. A direção  $\pi - \mu^n$ , de subida para  $\theta$ , a partir de  $\mu^n$ , gerará um novo ponto  $\pi$ , que será ou não aceito como novo iterado, respectivamente correspondendo a um passo sério ou nulo. No primeiro caso se terá  $\mu^{n+1}$ , caso contrário se terá gerado um novo ponto auxiliar  $\pi^{\ell+1}$ .

Observamos que em  $(F_{II})$ [15] está implícito o modelo local associado a  $\theta$ , dado pela função côncava seccionalmente afim:

$$\psi(\pi) := \theta(\mu^n) + \min_{i=1, \dots, \ell} \{(h^i)^t(\pi - \mu^n) + e_i\}$$

Assim,  $(F_{II})$ [15] é equivalente a

$$\max_{\pi} \left\{ \psi(\pi) - \frac{1}{2t_n} \|\pi - \mu^n\|^2 \right\}$$

Temos então o seguinte

**Lema 4.1.** O problema quadrático  $(F_{II})$ [15] possui uma solução única  $\hat{\pi}$  caracterizada por

$$\hat{\pi} = \mu^n + t_n \hat{h}, \quad \hat{h} \in \partial\psi(\hat{\pi})$$

Além disso,

$$\psi(\pi) \leq \theta(\mu^n) + \hat{h}^t(\pi - \mu^n) + \hat{e}, \quad \forall \pi$$

onde

$$\hat{e} := \psi(\hat{\pi}) - \theta(\mu^n) - t_n \|\hat{h}\|^2 \geq 0 \quad \blacklozenge$$

Sempre seguindo o texto citado, temos a função

$$\theta^a(\pi) = \psi(\hat{\pi}) + \hat{h}^t(\pi - \hat{\pi})$$

que é a chamada linearização agregada de  $\theta$ . Em particular, a solução  $\hat{\pi}$  de  $(F_{II})$ [15] também maximiza a função

$$\theta^a(\pi) - \frac{1}{2t_n} \|\pi - \mu^n\|^2 = \psi(\hat{\pi}) + \hat{h}^t(\pi - \hat{\pi}) - \frac{1}{2t_n} \|\pi - \mu^n\|^2$$

Este resultado, junto ao lema anterior levam ao novo modelo

$$\psi_+(\pi) = \min\{\psi(\pi), \theta(\hat{\pi}) + \hat{h}^t(\pi - \hat{\pi})\}.$$

O que se fez foi acrescentar ao feixe de informações o conjunto  $\{\hat{\pi}, \theta(\hat{\pi}), \hat{h}\}$ , onde  $\hat{h}$  é um subgradiente de  $\theta(\hat{\pi})$ . (veja detalhes no texto citado).

## Dualização do Modelo Local

Em lugar da resolução direta de (F<sub>II</sub>)[15](geração de linhas) se preferirá o dual(geração de colunas). Definiremos o lagrangeano e a partir dele obteremos as condições de otimalidade.

Para  $d := \pi - \mu^n$ ,

$$\mathcal{L}^*(z, d, \xi) = z - \frac{1}{2t_n} \|d\|^2 - \sum_{i=1}^{\ell} \xi^i (z - (\mathbf{h}^i)^t d - e_i - \theta(\mu^n)), \quad \xi \geq 0,$$

$$\mathcal{L}_z^* = 0 \quad \therefore \quad \sum_{i=1}^{\ell} \xi^i = 1,$$

$$\mathcal{L}_d^* = 0 \quad \therefore \quad -\frac{1}{t_n} d + \sum_{i=1}^{\ell} \xi^i \mathbf{h}^i = 0$$

Complementaridade:

$$\xi^i (z - (\mathbf{h}^i)^t d - e_i - \theta(\mu^n)) = 0, \quad i \in \{1, \dots, \ell\}$$

Substituindo-se estas equações em  $\mathcal{L}^*$ , chegamos ao dual de (F<sub>II</sub>)[15], que é o seguinte problema quadrático com restrições lineares

$$(DF_{II})[16] \quad \min \varphi(\xi)$$

$$\text{s.a.} \quad \sum_{i=1}^{\ell} \xi^i = 1$$

$$\xi^i \geq 0$$

$$\text{onde } \varphi(\xi) = \frac{t_n}{2} \left\| \sum_{i=1}^{\ell} \xi^i \mathbf{h}^i \right\|^2 + \sum_{i=1}^{\ell} \xi^i e_i + \theta(\mu^n)$$

Das condições de otimalidade tem-se também a atualização do multiplicador

$$\pi = \mu^n + t_n \left( \sum_{i=1}^{\ell} \xi^i h_1^i, \dots, \sum_{i=1}^{\ell} \xi^i h_5^i \right)$$

onde  $\xi$  é solução única do problema dual acima. [Bazaraa et alii. 93]

### 4.3. Descrição do Algoritmo

Descrevemos o algoritmo de feixes aplicado ao problema dual (D<sub>II</sub>)[6].

A aceitação da solução  $\pi^{\ell+1}$ , o chamado passo sério, é, essencialmente uma avaliação numérica de que “ $\theta$  cresceu suficientemente” e  $\pi^{\ell+1}$  é suficientemente afastado de  $\mu^n$ , dado pelo respectivo teste do tipo Armijo. Caso não haja aceitação do vetor  $\pi^{\ell+1}$  calculado, temos o passo nulo. Esta é uma característica específica dos métodos de feixes: em ambos os casos, novas informações serão adicionadas ao modelo local da função  $\theta$  (isto é, mais uma restrição em (F<sub>II</sub>)[15]) e se pode também modificar o valor de  $t_n$ . Um ponto crítico dos métodos de feixes é justamente a escolha, quando o passo é nulo, entre as duas seguintes opções:

a)  $t_n$  é muito pequeno (observe que valores diferentes de  $t_n$  causam diferentes direções de subida).

b) O modelo local para  $\theta$ (gerado pelo conjunto dos resíduos das restrições até então consideradas e a agregação) é insuficiente para o cálculo de uma adequada direção de subida.

Usualmente se enriquece o modelo com novos cortes, e, eventualmente se atualiza o passo. Observamos enfim que cuidados devem ser tomados de modo que a cardinalidade do feixe se mantenha razoável. Regra de descarte é portanto um ponto importante na implementação efetiva do método de feixes.

**Definição 4.2.** Um ponto  $\mu$  é uma  $\varepsilon$ -solução do problema  $(\theta)[7]$  se  $\theta(\mu) \leq \theta(\tilde{\mu}) + \varepsilon$  para todo ponto  $\tilde{\mu}$ .

### 4.3.1. Algoritmo de Feixes com Penalidade

Inicialização: São dados o ponto inicial  $\mu^1$ , juntamente com uma tolerância de parada  $\underline{\delta} \geq 0$  e  $\bar{\ell} > 0$  o tamanho máximo do feixe. Escolha um coeficiente de subida  $m_1 \in (0, 1)$ , inicialize o conjunto de subida  $N = \emptyset$ , o contador de iterações  $n = 1$  e o tamanho do feixe  $\ell = 1$ . Calcule  $\theta(\mu^1)$ , através da resolução de  $(\theta)$ [7] e  $h^1 = h(\mu^1)$  usando [9]. Faça  $e_1 = 0$ , correspondendo ao feixe inicial  $(h^1, e_1)$ , e o modelo inicial

$$\pi \rightarrow \psi_1(\pi) := \theta(\mu^1) + (h^1)^t(\pi - \mu^1)$$

#### Passo 1.0: Cálculo Principal e Teste de Parada

Escolha  $t_n > 0$  e seja  $\pi^{n+1}$  a solução única do problema quadrático  $(F_{II})$ [15] tal que

$$\pi^{n+1} = \mu^n + t_n \hat{h}^n \text{ com } \hat{h}^n \in \partial \psi_n(\pi^{n+1})$$

Faça

$$\hat{\epsilon}_n := \psi_n(\pi^{n+1}) - \theta(\mu^n) - t_n \|\hat{h}^n\|^2$$

$$\delta_n := \psi_n(\pi^{n+1}) - \theta(\mu^n) - \frac{t_n}{2} \|\hat{h}^n\|^2$$

Se  $\delta_n \leq \underline{\delta}$  pare.

#### Passo 2.0: Teste de Subida (Busca de Armijo):

Calcule  $\theta(\pi^{n+1})$  através de  $(\theta)$ [7] e  $h(\pi^{n+1})$  usando [9].

Se  $\theta(\pi^{n+1}) - \theta(\mu^n) \geq m_1 \delta_n$  não se verificar, trata-se de

“passo nulo”, vá para o passo 4.



**Passo 3.0: Passo Sério**

Faça  $\mu^{n+1} = \pi^{n+1}$ . Acrescente  $n$  ao conjunto  $N$ ; para  $i = 1, \dots, \ell$  permutar  $e_i$  e  $\hat{e}_n$  por, respectivamente

$$e_i + \theta(\mu^n) - \theta(\mu^{n+1}) + (h^i)^t(\mu^{n+1} - \mu^n)$$

$$\hat{e}_n + \theta(\mu^n) - \theta(\mu^{n+1}) + (\hat{h}^i)^t(\mu^{n+1} - \mu^n)$$

**Passo 4.0: Controle do Tamanho do Feixe**

Se  $\ell = \bar{\ell}$  então elimine no mínimo 2 elementos do feixe e insira o elemento  $(\hat{h}^n, \hat{e}_n)$ .

Denomine  $(h^i, e_i)_{i=1, \dots, \ell}$  o novo feixe obtido (com  $\ell < \bar{\ell}$ ).

**Passo 5.0: Insira  $(h^{\ell+1}, e_{\ell+1})$  ao feixe, onde  $e_{\ell+1} = 0$  no caso de passo sério, e no caso de passo nulo**

$$e_{\ell+1} = \theta(\pi^{n+1}) - \theta(\mu^n) + (h^{\ell+1})^t(\mu^n - \pi^{n+1})$$

**Substitua  $\ell$  por  $\ell + 1$  e defina o modelo**

$$\pi \rightarrow \psi_{n+1}(\pi) := \theta(\mu^{n+1}) + \min_{i=1, \dots, \ell} [e_i + (h^i)^t(\pi - \mu^{n+1})]$$

**Passo 6.0: Faça  $n = n + 1$  e volte para o passo 1.**

#### 4.4 Convergência do Método de Feixes

Para garantirmos a convergência do método de feixe, é suficiente fixarmos o parâmetro  $t_n = t > 0$  [Lemaréchal. 92]. Entretanto visando a uma melhor eficiência numérica é desejável adaptar-se  $t_n$  em cada iteração. Além disso, deve-se observar que pequenos valores de  $t_n$  são desastrosos para a convergência, enquanto que o teste de Armijo pode não ser válido para valor algum de  $t_n > 0$ . Para possíveis implementações computacionais, referimos a [Kiwiel. 90], [Outrata et alii. 91] e [Schramm and Zowe. 92].

A convergência do **algoritmo** é consequência direta dos teoremas a seguir, em conformidade com [Hiriart-Urruty and Lemaréchal. 93].

##### **Teorema 4.1: Número infinito de passos sérios.**

Seja o **algoritmo** aplicado ao problema de maximização  $(D_{II})[6]$ , com uma tolerância de parada  $\underline{\delta} = 0$ . Assuma que  $N$  é um conjunto infinito.

i) Se

$$\sum_{n \in N} t_n = +\infty$$

então  $\{\mu^n\}$  é uma seqüência maximizante.

ii) Se, além disso,  $\{t_n\}$  tem uma cota superior em  $N$ , e se  $(D_{II})[6]$  tem um conjunto não-vazio de soluções, então a seqüência  $\{\mu^n\}$  converge para a solução.

#### **Teorema 4.2: Número infinito de passos nulos**

Considere o **algoritmo** com uma tolerância de parada  $\underline{\delta} = 0$ . Assuma que  $N$  é finito: para algum  $n$ , cada iteração  $n \geq n_0$  produz um passo nulo. Se

$$t_n \leq t_{n-1}, \text{ e para todo } n > n_0$$

$$\sum_{n>n_0} \frac{t_n^2}{t_{n-1}} = +\infty$$

então  $\mu^{n_0}$  maximiza  $\theta$ .

#### **Convergência Finita**

Uma variante do algoritmo acima descrito foi proposta por [Kiwiel. 87b] onde não há agregação (só entra o subgradiente do oráculo no passo 2.0) e é definido um processo denominado de seleção. Neste se determina uma cardinalidade máxima para o feixe, igual à dimensão da variável  $\mu$  do problema mais 2. A seleção dos subgradientes está associada aos multiplicadores não nulos do subproblema quadrático, que serão em quantidade máxima dada por aquela dimensão mais 1. (veja o artigo citado).

Neste trabalho, é demonstrado que sob condição de regularidade do tipo Haar (essencialmente, independência linear dos subgradientes do feixe) a convergência, para problemas seccionalmente afins - é o caso da função  $\theta$ , é obtida em um número finito de passos.

Obs: É razoável supor que este resultado seja extensível para o **algoritmo** acima descrito.

#### 4.5. Recuperação das Variáveis Primais no Caso Linear Contínuo

Para o problema linear contínuo  $(\theta_R)$  podemos estimar as variáveis primais do problema (PR)[3] como uma combinação convexa das variáveis primais obtidas das iterações anteriores, conforme [Lemaréchal. 92] e [Hiriart-Urruty and Lemaréchal. 93] fazem para o método de planos secantes.

Suponha, no passo 2 do algoritmo que as variáveis primais obtidas no cálculo de  $\theta(\pi^{n+1})$  sejam armazenadas em um conjunto  $\{x^k, y^k\}$ ,  $k = 1, \dots, \ell$ . Assuma também que  $(F_{II})$ [15] é resolvido através do dual  $(DF_{II})$ [16]. Neste caso, podemos considerar a fórmula

$$(x(\xi), y(\xi)) = \sum_{k=1}^{\ell} \xi^k (x^k, y^k),$$

da combinação convexa gerada pela solução de  $(DF_{II})$ [16]. Temos então que

**Teorema 4.3.** As variáveis primais podem ser obtidas através da fórmula

$(x(\xi), y(\xi)) = \sum_{\ell=1}^n \xi_{\ell} (x^{\ell}, y^{\ell})$ , onde  $\xi$  é o multiplicador solução do

problema dual.

**Demonstração:** De fato, observamos que o problema primal (PR)[3] e seu dual  $(D_{II})$ [6] não apresentam salto de dualidade, assim como  $(F_{II})$ [15] e  $(DF_{II})$ [16]. Observe que  $(x(\xi), y(\xi))$  é viável (os resíduos serão nulos devido à otimalidade). Assim devido também à linearidade

da função custo, e chamando de  $\mu = \mu^n$  o multiplicador ótimo associado a  $\xi$ , temos

$$\psi(\mu) = \sum_{\ell} \xi^{\ell} \left[ \sum_{k=1}^4 c_k^t x_k^{\ell} + \sum_{j=1}^2 e_j^t y_j^{\ell} \right] = \sum_{k=1}^4 c_k^t x_k(\xi) + \sum_{j=1}^2 e_j^t y_j(\xi) \quad \blacklozenge$$

Observação: É evidente que a simples substituição do multiplicador ótimo no problema primal relaxado deverá também resultar em uma solução primal ótima.

#### 4.6. Feixes x "Cross Decomposition"

A comparação efetiva entre as duas metodologias aplicadas à mesma relaxação exigiria uma bateria de testes computacionais. Gostaríamos no entanto de observar que o esforço, por iteração, leva ao seguinte:

i) "Cross Decomposition":

i.1. Iteração com aperfeiçoamento de cotas (poderíamos denominar de passo sério): resolve-se um problema linear inteiro ( $\theta$ )[10] e um linear (DB)[13].

i.2. Iteração sem aperfeiçoamento de cotas (passo nulo): além de  $\theta$  [10] e (DB)[13], resolve-se o problema mestre linear de Dantzig-Wolfe (DM)[11], e/ou o problema mestre inteiro de Benders (MB)[14].

ii) Feixes

ii.1. Passo sério: Resolve-se um problema quadrático linear (F<sub>II</sub>)[15] e um problema linear inteiro (θ)[7]( oráculo).

ii.2. Passo nulo: o mesmo acima.

Assim, se se supuser que em ambos os algoritmos se tenha passo sério, K passos nulos, obtemos a seguinte comparação, por passo sério:

i) " Cross Decomposition": de  $3K + 2$  a  $4K + 2$  subproblemas devem ser resolvidos;

ii) Feixes:  $2K + 2$  subproblemas devem ser resolvidos.

Lembramos que a complexidade do problema quadrático linear é equivalente à da programação linear, sendo que a utilização da forma dual (DF<sub>II</sub>)[16] permite o uso de algoritmos particularmente eficientes, veja [Kiwiel. 86]. Acrescentamos que o problema mestre de Benders (MB)[14] é inteiro, o que pode acarretar maior complexidade aritmética.

Deste modo, é, no mínimo, desejável que se realize comparações numéricas entre os dois algoritmos.

## CAPÍTULO 5

### UMA METODOLOGIA DE REGULARIZAÇÃO E DECOMPOSIÇÃO DE BENDERS

#### 5.1 Introdução

A conhecida convergência lenta dos algoritmos baseados na decomposição de Benders em problemas lineares inteiros de grande porte motivou o desenvolvimento da metodologia que apresentaremos a seguir, e que pretende acelerar o método clássico. Em particular [Cabral. 93] obteve resultados extremamente lentos para o problema original citado [Moreno et alii. 92], apesar de ter considerado uma instância de pequena dimensão.

Aplicaremos a decomposição de Benders ao problema linear inteiro com a relaxação I (§ 3.2.1), com atualização dos multiplicadores por um método do tipo feixes.

#### 5.2 Decomposição de Benders para o Problema Relaxado I.

O método da decomposição de Benders aplicado ao problema relaxado  $(\varphi)$  [5] consiste em reformular este em um problema equivalente contendo somente  $y$ -variáveis inteiras e uma variável contínua. Suporemos que aquele problema tem solução ótima finita para todo  $\lambda$ .



Para cada  $\lambda$ ,  $(\varphi)[5]$  pode ser reescrito como:

$$(\varphi)[5] \min_{y \in Q} \left\{ \sum_{j=1}^2 e_j^t y_j + \lambda^t C_2 y_2 + \min_{x \in X_R} \left\{ \sum_{k=1}^4 c_k^t x_k + \lambda^t (C_3 x_3 + C_4 x_4); A_k x_k = K_k - B_k y_1, k = 1, \dots, 4, x \in X_R \right\} \right\}$$

onde

$Q = \{y \in Y; \exists x \in X_R \text{ tal que } A_k x_k = K_k - B_k y_1, k = 1, \dots, 4\}$ , que supomos não vazio.

Para  $y \in Q$  com  $y_1$  fixo, o subproblema de minimização interna (com  $X_R$  explicitado),

$$\begin{aligned} (L)[17] v_L = \min_x \sum_{k=1}^4 c_k^t x_k + \lambda^t (C_3 x_3 + C_4 x_4) \\ \text{s.a.} \quad D_k x_k = d_k \quad k = 1, \dots, 4 \\ A_k x_k = K_k - B_k y_1 \quad k = 1, \dots, 4 \\ x_k \geq 0 \quad k = 1, \dots, 4 \end{aligned}$$

tem seu dual dado por

$$\begin{aligned} (D)[18] v_D = \max_{(v,u)} \sum_{i=1}^4 d_i^t v_i + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i \\ \text{s.a.} \quad D_i^t v_i + A_i^t u_i \leq \tilde{c}_i \quad \text{onde} \quad \tilde{c}_i = \begin{cases} c_i & i = 1, 2 \\ c_i + C_i^t \lambda & i = 3, 4, \end{cases} \end{aligned}$$

onde  $u = (u_1, \dots, u_4)^t$  e  $v = (v_1, \dots, v_4)^t$ .

Assumiremos que os poliedros

$$U(\lambda) = \{(v, u); D_1^t v_i + A_i^t u_i \leq \tilde{c}_i \text{ onde } \tilde{c}_i = \begin{cases} c_i, & i=1, 2 \\ c_i + C_i^t \lambda, & i=3, 4 \end{cases}\}$$

são uniformemente limitados, se necessário acrescentando-se cotas às variáveis  $(v, u)$  e a  $\lambda$ . Deste modo podemos definir o conjunto  $\{(v^q, u^q)_\lambda\} \forall q \in P_{U(\lambda)}$  (finito) de pontos extremos de  $U(\lambda)$ . Neste caso,  $(\varphi)[5]$  é equivalente a

$$(\varphi)[5] \min_{y \in Q} \left\{ \sum_{j=1}^2 e_j^t y_j + \lambda^t C_2 y_2 + \max_{q \in P_{U(\lambda)}} \left\{ \sum_{i=1}^4 d_i^t v_i^q + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^q \right\} \right\}$$

Chamando de  $z_L(\lambda)$  o argumento do mínimo, tem-se para qualquer subconjunto  $P'_{U(\lambda)} \subseteq P_{U(\lambda)}$ , o chamado problema mestre relaxado de Benders:

$$(\text{MB})[19] \quad z_L(\lambda) = \min_{(z, y)} z$$

$$\text{s.a } z \geq \sum_{j=1}^2 e_j^t y_j + \lambda^t C_2 y_2 + \sum_{i=1}^4 d_i^t v_i^q + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^q$$

$$z \in \mathbb{R}, y \in Y \quad \forall q \in P'_{U(\lambda)}$$

### 5.3 Cotas

Para  $\lambda$  fixo, sejam  $z_U(\lambda) = \sum_{j=1}^2 e_j^t y_j + \lambda^t C_2 y_2 + v_D$  o limite superior,

onde  $v_D$  foi obtido no dual do subproblema primal relaxado (D)[18] e  $z_L(\lambda)$  o limite inferior proveniente do problema mestre de Benders (MB)[19].

Então

$$z_L(\lambda) \leq \varphi(\lambda) = \min_{(x,y) \in W_I} \mathcal{L}(x, y, \lambda) \leq z_U(\lambda)$$

Para  $\lambda$  variável, se supusermos, como será feito no algoritmo, que as restrições do problema mestre relaxado de Benders serão mantidas, de uma a outra iteração em  $\lambda$ , então  $z_L(\lambda^{p+1}) \geq z_L(\lambda^p)$ .

### 5.4 Regularização Quadrática do Problema Dual

A resolução iterativa do problema dual de maximização em  $\varphi$  (D<sub>I</sub>)[4], que atualiza o multiplicador  $\lambda$ , é feita através de um modelo regularizado local, como na metodologia de feixes. Entretanto, diferentemente do capítulo anterior, não conhecemos, para cada  $\lambda$ , o valor de  $\varphi(\lambda)$ , apenas dispomos das cotas inferior( $z_L(\lambda)$ ) e superior( $z_U(\lambda)$ ), dadas acima.

Suponhamos que estamos na  $p$ -ésima iteração  $\lambda^p$ . Definiremos o

modelo abaixo, motivados pelo capítulo anterior, veja  $(F_{II})[15]$ :

$$(F_I)[20] \quad w(\lambda^p) = \max_{(w, \rho)} w - \frac{1}{2t_p} \|\rho - \lambda^p\|^2$$

$$\text{s.a. } w \leq (g^r)^t (\rho - \lambda^p) + z_U(\rho^r), \quad r \geq 1$$

sendo  $t_p > 0$ , o qual determina o tamanho da direção  $\rho - \lambda^p$ .

Comparando-se com a aplicação da metodologia de feixes do capítulo anterior, a formulação acima adapta aquela estrutura, de acordo com o seguinte:

a) o valor de  $g^r := C_3 x_3^r + C_4 x_4^r + C_2 y_2^r - K_5$  não corresponde a um subgradiente, mas a algum  $\varepsilon_r$ -subgradiente de  $\varphi$  em  $\lambda^p$ .

De fato, para  $(x^r, y^r) \in W_1$  e  $\lambda$  qualquer

$$\varphi(\lambda) = \min_{(x, y) \in W_1} \mathcal{L}(x, y, \lambda)$$

$$\leq \mathcal{L}(x^r, y^r, \lambda) \text{ para algum } \varepsilon_r \geq 0$$

Da definição de  $\mathcal{L}$ , temos

$$\varphi(\lambda) \leq (\lambda - \lambda^p)^t g^r + \mathcal{L}(x^r, y^r, \lambda^p) + \varepsilon_r, \quad \text{para algum } \varepsilon_r \geq 0$$

$$= (\lambda - \lambda^p)^t g^r + \varphi(\lambda) + \varepsilon_r,$$

isto é,  $g^r \in \partial_{\varepsilon_r} \varphi(\lambda^p)$

b) os cortes lineares, correspondentes ao modelo poliedral local, não são exatos, devido ao descrito no item anterior, e também ao desconhecimento de  $\varphi(\lambda^r)$ . Este valor é substituído pela cota superior  $z_U(\rho^r)$ , fornecida pelo dual do subproblema primal relaxado(D)[18]. A obtenção de  $g^r$ ,  $z_U(\rho^r)$  (e de  $z_L(\rho^r)$ ) pode exigir algumas iterações do algoritmo de Benders. De fato, consideraremos aceitável  $(g^r, z_U(\rho^r), z_L(\rho^r))$  se o seguinte **teste de qualidade da aproximação de  $\varphi(\rho^r)$**  é verificado:

$$[21] \quad z_U(\rho^r) - z_L(\rho^r) \leq \alpha(z_U(\lambda^{p-1}) - z_L(\lambda^{p-1}))$$

para algum  $0 < \alpha < 1$ .

Observamos que a convergência do método de Benders garante que o teste será verificado em um número finito de iterações [Benders. 62]. Com este teste se garante que o erro máximo no cálculo de  $\varphi$ , de uma iteração para outra em  $\lambda$ , decresça. Indiretamente podemos também esperar que  $\varepsilon_r \rightarrow 0$ , onde  $\varepsilon_r$  foi definido no item anterior.

Com este conjunto de informações, temos o modelo "aproximado"

$$[22] \quad \Omega_p(\rho) := \min_r \{ (g^r)^t (\rho - \lambda^p) + z_U(\rho^r) \}$$

Assim, equivalentemente, a (F<sub>I</sub>)[20] temos:

$$[23] \quad \max_p \{ \Omega_p(\rho) - \frac{1}{2t_p} \|\rho - \lambda^p\|^2 \}$$

É interessante observar que o modelo regularizado  $(F_I)$ [20] tem nele embutido o processo(decomposto) dos planos secantes, e pretende determinar uma direção de subida através dos resíduos acumulados, com o cálculo aproximado da função dual  $\varphi(\lambda)$  em  $(D_I)$ [4], através da decomposição de Benders.

O lema e a proposição que se seguem visam justificar, no algoritmo 5.5.3, a existência e unicidade da solução do subproblema quadrático, assim como seu subproduto que é o subgradiente agregado. Para maiores detalhes sobre a técnica de agregação, veja o livro [Hiriart-Urruty and Lemaréchal. 93], seção XV.3.1.

**Lema 5.1**(veja o lema XV.3.1.1, do livro [Hiriart-Urruty and Lemaréchal. 93])

O problema [23] tem uma solução única  $\rho^{p+1}$  caracterizada pela seguinte fórmula

$$[24] \quad \rho^{p+1} = \lambda^p + t_p \hat{g}_p, \quad \hat{g}_p \in \partial\Omega_p(\rho^{p+1})$$

Além disso

$$\Omega(\lambda) \leq z_L(\lambda_p) + (\hat{g}_p)^t(\lambda - \lambda_p) + \hat{e}_p, \quad \forall \lambda,$$

onde

$$[25] \quad \hat{e}_p = \Omega(\rho^{p+1}) - z_L(\lambda_p) - t_p \|\hat{g}_p\|^2.$$

Demonstração:

Suponha não vazio o conjunto gerado pelas restrições lineares, a existência e unicidade da solução  $\rho^{p+1}$  seguem da definição positiva da quadrática. A condição de otimalidade para esta solução é

$$0 \in \partial\Omega_p(\rho^{p+1}) - \frac{1}{t_p}(\rho^{p+1} - \lambda^p),$$

que é [24].

Temos então

$$\Omega(\lambda) \leq \Omega(\rho^{p+1}) + (\hat{g}_p)^t(\lambda - \rho^{p+1})$$

que é equivalente a

$$\Omega(\lambda) \leq z_L(\lambda_p) + (\hat{g}_p)^t(\lambda - \lambda_p) - z_L(\lambda_p) + \Omega(\rho^{p+1}) + (\hat{g}_p)^t(\lambda_p - \rho^{p+1})$$

Em vista de [24], reconhece-se a expressão [25] de  $\hat{e}_p$ . ♦

**Proposição 5.2** (veja a proposição XV.3.1.2, do livro [Hiriart-Urruty and Lemaréchal. 93])

Com a notação do **lema 5.1**, considere uma função quadrática

$\Psi: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$  satisfazendo

$$[26] \quad \Psi(\lambda) \leq z_L(\lambda_p) + (\hat{g}_p)^t(\lambda - \lambda_p) + \hat{e}_p =: \tilde{\varphi}(\lambda), \quad \forall \lambda,$$

com igualdade em  $\lambda = \rho^{p+1}$ . Então  $\rho^{p+1}$  maximiza a função

$$\tilde{\Psi}(\lambda) := \Psi(\lambda) - \frac{1}{2t_p} \|\lambda - \lambda^p\|^2$$

Demonstração:

Usando [24] e [25], as relações definindo  $\Psi$  podem ser escritas, sucessivamente:

$$\Psi(\lambda) \leq \Omega(\rho^{p+1}) + (\hat{g}_p)^t(\lambda - \rho^{p+1}) - \Omega(\rho^{p+1}) + (\hat{g}_p)^t(\rho^{p+1} - \lambda_p) + z_L(\lambda_p) + \hat{e}_p \therefore$$

$$\Psi(\lambda) \leq \Omega(\rho^{p+1}) + (\hat{g}_p)^t(\lambda - \rho^{p+1})$$

com igualdade em  $\lambda = \rho^{p+1}$ . Subtraindo o termo  $\frac{1}{2t_p} \|\lambda - \lambda_p\|^2$  de ambos

os lados,

$$\tilde{\Psi}(\lambda) \leq \Omega(\rho^{p+1}) + (\hat{g}_p)^t(\lambda - \rho^{p+1}) - \frac{1}{2t_p} \|\lambda - \lambda_p\|^2,$$

ainda com igualdade em  $\lambda = \rho^{p+1}$ . Agora observe que a função do lado direito é maximizada quando

$$\hat{g}_p - \frac{1}{t_p}(\lambda - \lambda_p) = 0,$$

que corresponde a  $\rho^{p+1}$ , dado por [24]. ♦

A função  $\tilde{\varphi}(\lambda)$  é denominada linearização agregada à aproximação de  $\varphi$ . Ele majora o modelo  $\Omega$ , conforme descrito no **lema 5.1**.

### **Dualização do modelo local**

Aqui também, é mais conveniente que a resolução de  $(F_1)$ [20] seja feita através do dual. Definiremos o lagrangeano e respectivas condições de otimalidade.



Para  $\mathbf{d} := \rho - \lambda^p$ ,

$$\mathcal{L}^*(\mathbf{w}, \mathbf{d}, \eta) = \mathbf{w} - \frac{1}{2t_p} \|\mathbf{d}\|^2 - \sum_{r \in \{1, \dots, \ell\}} \eta^r (\mathbf{w} - (\mathbf{g}^r)^t \mathbf{d} - z_U(\rho^r)), \quad \eta \geq 0$$

$$\mathcal{L}_w = 0 \quad \therefore \quad \sum_{r \in \{1, \dots, \ell\}} \eta^r = 1,$$

$$\mathcal{L}_d^* = 0 \quad \therefore \quad -\frac{1}{t_p} \mathbf{d} + \sum_{r \in \{1, \dots, \ell\}} \eta^r \mathbf{g}^r = 0,$$

Complementaridade:

$$\eta_r [\mathbf{w} - (\mathbf{g}^r)^t \mathbf{d} - z_U(\rho^r)] = 0, \quad r \geq 1$$

Substituindo-se estas equações em  $\mathcal{L}^*$ , chegamos ao seguinte problema quadrático linear

(DF<sub>1</sub>)[26] mín  $\wp(\eta)$

$$\text{s. a.} \quad \sum_{r \in \{1, \dots, \ell\}} \eta^r = 1$$

$$\eta^r \geq 0$$

$$\text{onde } \wp(\eta) = \frac{t_p}{2} \left\| \sum_{r \in \{1, \dots, \ell\}} \eta^r \mathbf{g}^r \right\|^2 + \sum_{r \in \{1, \dots, \ell\}} \eta^r z_U(\rho^r)$$

Das condições de otimalidade tem-se também a atualização do multiplicador

$$\rho = \lambda^p + t_p \left( \sum_{r \in \{1, \dots, \ell\}} \eta^r \mathbf{g}^r \right),$$

onde  $\eta$  é solução única de (DF<sub>1</sub>)[26].

## 5.5 O Algoritmo de Feixes Aproximado

### 5.5.1. O Algoritmo Parcial de Benders

A cada iteração, o multiplicador  $\lambda$  é usado no subproblema (D)[18], que, resolvido, fornece um limite superior  $z_U$  e gera um novo corte de Benders para ser incluído no problema mestre relaxado (MB)[19]. A resolução deste fornece um limite inferior  $z_L$  e uma variável  $y$  para o subproblema (L)[17], que por sua vez, é resolvido em  $x$ . Com  $\lambda$  fixo, este processo é repetido, acumulando-se todos os cortes no problema mestre de Benders MB[19], até que o teste [21] seja satisfeito. Ao final deste processo os valores de  $x$ ,  $y$ ,  $z_U$  e  $z_L$  são levados ao modelo regularizado, para uma nova atualização do multiplicador  $\lambda$ .

Observação: Optamos por incluir no modelo quadrático apenas o corte que corresponde à realização do teste [21]. Poder-se-ia, no entanto, incluir todos os cortes, deixando para a política de seleção a eliminação adequada. (veja passo 4.0, "Controle do Tamanho do Feixe", algoritmo 5.5.3).

### 5.5.2. Teste de Armijo Aproximado

Sempre motivado pelo capítulo anterior, uma aproximação do teste de Armijo (passo 2) do algoritmo 4.3.1, determinará aqui a direção de subida suficiente da aproximação de  $\varphi$ . Assim, seja

$$\delta_p := \Omega(\rho^{p+1}) - z_L(\rho^p) - \frac{1}{2t_p} \|\rho^{p+1} - \lambda^p\|^2 \text{ onde } \rho^{p+1} = \lambda^p + t_p \hat{g}$$

onde  $\Omega$  é dado por [22].

Aproximando-se os valores da  $\varphi$  pelas cotas inferior ( $z_L(\lambda)$ ) e superior ( $z_U(\rho)$ ) tem-se

$$\varphi(\rho^{p+1}) - \varphi(\lambda^p) \leq z_U(\rho^{p+1}) - z_L(\lambda^p)$$

Para  $0 < m_1 < 1$  fornecido, uma aproximação do teste de Armijo será satisfeita em  $\rho^{p+1}$  se:

$$z_U(\rho^{p+1}) - z_L(\lambda^p) \geq m_1 \delta_p$$

onde o lado esquerdo é positivo porque

$$z_U(\rho^{p+1}) - z_L(\lambda^p) \geq z_U(\rho^{p+1}) - z_L(\rho^{p+1}) \geq 0$$

Se compararmos este teste ao que correspondia ao cálculo exato da função  $\varphi$ , observamos que a diferença entre os valores atual e o candidato foi substituída por uma majoração, tanto quanto  $\delta_p$  é uma majoração do valor exato. Isto permite esperar que o teste de

parada do método aproximado de feixes não será antecipado, desde que também se assegure uma boa aproximação para a função  $\varphi$ .

Nos dois próximos tópicos apresentaremos, em separado, o algoritmo externo de atualização de  $\lambda$ , baseado em um processo regularizado, seguindo a decomposição de Benders que é o fornecedor do oráculo aproximado.

### 5.5.3. Algoritmo de Regularização para Atualização dos Multiplificadores com a Relaxação I.

Antes de apresentarmos o algoritmo e com a finalidade de mantermos a notação do capítulo anterior, substituiremos o modelo  $F_1[20]$ . Levando em conta a observação feita na seção 4.2, ele é equivalente a

$$[27] \quad w(\lambda^p) = \max_{(w, \rho)} w - \frac{1}{2t_p} \|\rho - \lambda^p\|^2$$

$$\text{s.a. } w \leq (g^r)^t (\rho - \rho^r) + e_r + z_U(\lambda^p), \quad r \geq 1$$

onde  $e_r := e(\lambda^p, \rho^r, g^r) := z_U(\rho^r) - z_U(\lambda^p) + (g^r)^t (\rho^r - \lambda^p)$ .

Usamos indistintamente  $g(\rho^r)$  e  $g^r$ .

#### Algoritmo

Inicialização: São dadas tolerância de parada  $\underline{\delta} \geq 0$  e  $\theta > 0$ . Seja  $\bar{\ell} > 0$  o tamanho máximo do feixe,  $t_1 > 0$ . Obtenha uma solução viável dual inicial  $\lambda^1, y^0 \in Y$  e  $x^0$  solução viável inicial de (L)[17], isto é, para  $y = y^0$ ,  $x^0$  é solução de

$$A_k x_k = K_k - B_k y^0 \quad k = 1, \dots, 4$$

$$x_k \in X_k, \quad k = 1, \dots, 4$$

Calcule  $g^1 = g(\lambda^1)$ . Faça  $z_U(\lambda^1) := \mathcal{L}(x^0, y^0, \lambda^1)$ . Estime  $z_L(\lambda^1)$ , por exemplo, através de uma iteração do algoritmo de Benders. Escolha  $m_1 \in (0, 1)$  a redução do teste de Armijo,  $\alpha \in (0, 1)$  é a redução do teste

de qualidade da aproximação de  $\varphi$ . Inicialize o conjunto de subida  $P = \emptyset$ , o contador de iterações  $p = 1$  e o tamanho do feixe  $\ell = 1$ . Faça  $e_1 = 0$ , correspondendo ao feixe inicial  $(g^1, e_1)$ , e o modelo inicial

$$\rho \rightarrow \Omega_1(\rho) := z_U(\lambda^1) + (g^1)^t(\rho - \lambda^1)$$

### Passo 1.0: Cálculo Principal e Teste de Parada

Seja  $\rho^{p+1}$  a solução única do problema quadrático [27],

tal que,

$$\rho^{p+1} = \lambda^p + t_p \hat{g}^p \text{ com } \hat{g}^p \in \partial\Omega_p(\rho^{p+1})$$

Faça

$$\hat{e}_p := \Omega_p(\rho^{p+1}) - z_L(\lambda^p) - t_p \|\hat{g}^p\|^2$$

$$\delta_p := \Omega_p(\rho^{p+1}) - z_L(\lambda^p) - \frac{t_p}{2} \|\hat{g}^p\|^2$$

Calcule através do algoritmo (5.5.4)(oráculo aproximado)

$$z_U(\rho^{p+1}), z_L(\rho^{p+1}) \text{ e } g(\rho^{p+1}).$$

Se  $\delta_p \leq \underline{\delta}$  e  $z_U(\rho^{p+1}) - z_L(\rho^{p+1}) < \theta$  pare.

### Passo 2.0: Aproximação do Teste de Armijo

Se  $z_U(\rho^{p+1}) - z_L(\lambda^p) \geq m_1 \delta_p$ ,  $m_1 \in (0, 1)$  “passo sério”; caso contrário, trata-se de “passo nulo”, vá para o passo 4.

**Passo 3.0: Passo S3rio**

Faça  $\lambda^{p+1} = \rho^{p+1}$ . Acrescente  $p$  ao conjunto  $P$ ; para  $r = 1, \dots, \ell$

Permutar  $e_r$  e  $\hat{e}_p$  por, respectivamente,

$$e_r + z_U(\lambda^p) - z_U(\lambda^{p+1}) + (g^r)^t(\lambda^p - \lambda^{p+1})$$

$$\hat{e}_p + z_U(\lambda^p) - z_U(\lambda^{p+1}) + (\hat{g}^r)^t(\lambda^p - \lambda^{p+1})$$

**Passo 4.0: Controle do Tamanho do Feixe**

Se  $\ell = \bar{\ell}$  ent3o elimine no m3nimo 2 elementos do feixe e

insira o elemento  $(\hat{g}^p, \hat{e}_p)$ .

Denomine  $(g^t, e_t)_{t=1, \dots, \ell}$  o novo feixe obtido (com  $\ell < \bar{\ell}$ ).

**Passo 5.0:** Insira  $(g^{\ell+1}, e_{\ell+1})$  ao feixe, onde  $e_{\ell+1} = 0$  no caso de passo

s3rio, e no caso de passo nulo

$$e_{\ell+1} = z_U(\rho^{p+1}) - z_U(\lambda^p) + (g^{\ell+1})^t(\rho^{p+1} - \lambda^p)$$

Substitua  $\ell$  por  $\ell + 1$  e atualize o modelo

$$\rho \rightarrow \Omega_{p+1}(\rho) := \min_r \{ (g^r)^t(\rho - \lambda^{p+1}) + z_U(\rho^r) \}$$

**Passo 6.0:** Faça  $p = p + 1$  e volte para o passo 1.

### 5.5.4 Algoritmo Parcial de Benders para o Problema Linear Inteiro com a Relaxação I (Oráculo Aproximado)

Inicialização: Faça  $q = 1$

Passo 1: Resolver

$$v_D = \max_{(v_i, u_i)} \sum_{i=1}^4 d_i^t v_i + \sum_{i=1}^4 (K_i - B_i y_1^p)^t u_i$$

$$\text{s.a. } D_i^t v_i + A_i^t u_i \leq \tilde{c}_i \text{ onde } \tilde{c}_i = \begin{cases} c_i & i = 1, 2 \\ c_i + C_i^t \rho^{p+1} & i = 3, 4 \end{cases}$$

Se não existe uma solução, pare:  $(\varphi)[5]$  não tem solução viável. Caso contrário, seja  $(v^{p,q}, u^{p,q})$  uma solução, e faça

$$z_U(\rho^{p+1}) = \sum_{j=1}^2 e_j^t y_j + \rho^t C_2 y_2 + v_D$$

Gere um nova restrição(corte) a partir de  $(\rho^{p+1}, v^{p,q}, u^{p,q})$ .

Vá para o passo 2.

Passo 2: Resolver

min  $z$

$$\text{s.a. } z \geq \sum_{j=1}^2 e_j^t y_j + (\rho^{p+1})^t C_2 y_2 + \sum_{i=1}^4 d_i^t v_i^q + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^q$$

$$z \in \mathbb{R}, y \in Y \text{ e } \forall q$$

Seja  $(z_L(\rho^{p+1}), y^{p+1})$  a solução ótima; vá para o passo 3.



**Passo 3: Resolva**

$$\min \sum_{k=1}^4 c_k^t x_k + (\rho^{p+1})^t (C_3 x_3 + C_4 x_4).$$

$$\text{s.a. } A_k x_k = K_k - B_k y_l^{p+1} \quad k = 1, \dots, 4$$

$$x_k \in X_R \quad k = 1, \dots, 4$$

Seja  $x^{p+1}$  a solução; vá para o passo 4.

**Passo 4: Teste de Qualidade da Aproximação de  $\varphi$**

Se

$$z_U(\rho^{p+1}) - z_L(\rho^{p+1}) \leq \alpha(z_U(\lambda^p) - z_L(\lambda^p))$$

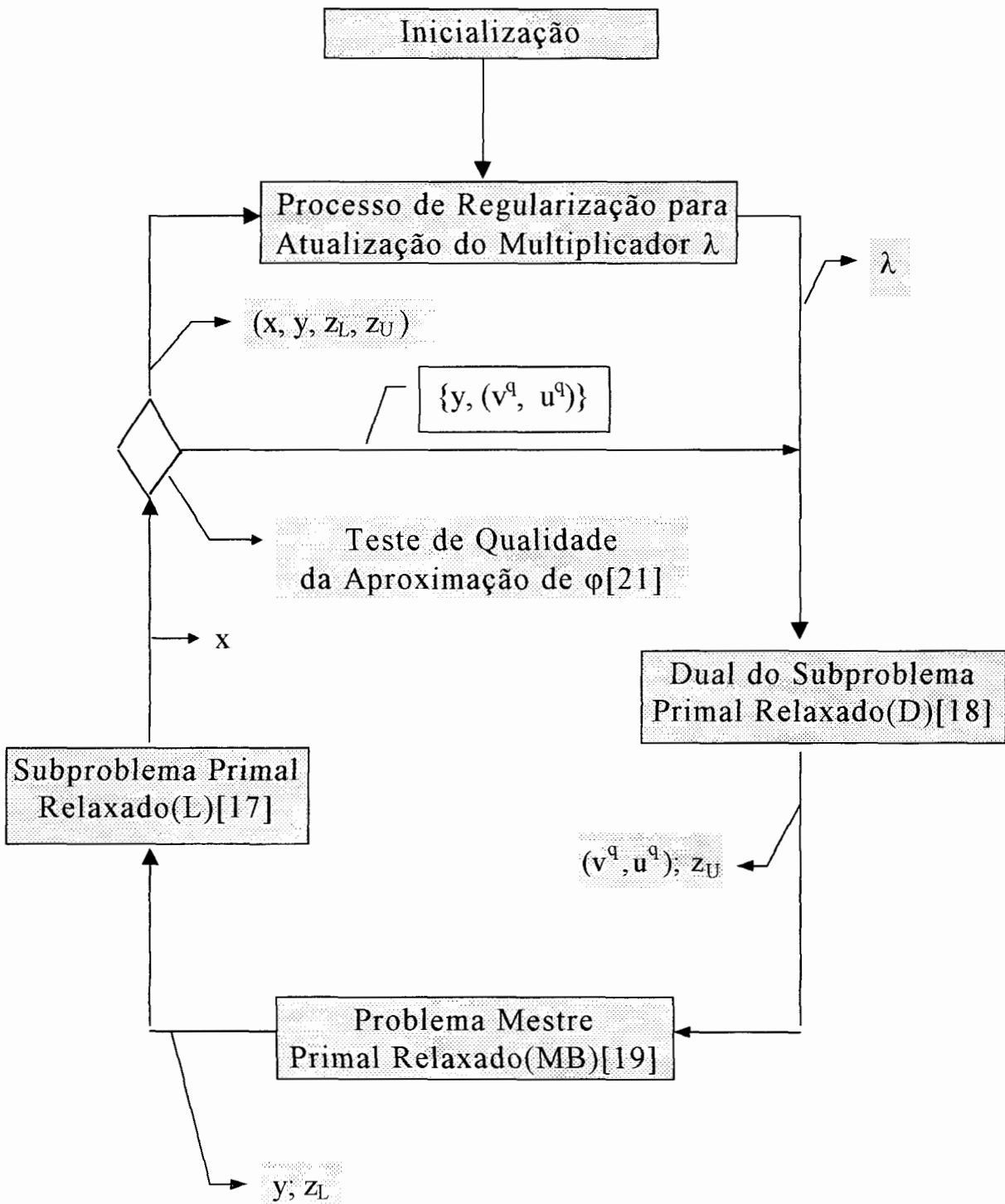
fim

Caso contrário, faça  $p + 1 = p$ ,  $q = q + 1$  e retorne ao passo 1.

**Observações:**

1. O teste de parada do algoritmo 5.5.3 acrescenta à usual tolerância  $\underline{\delta}$  dos feixes, a exigência de que a aproximação da função seja razoável. De fato, para aproximações grosseiras de  $\varphi$ , é possível ter-se falsos passos sérios com o erro  $\delta$  falsamente pequeno, daí a necessidade da  $\theta$ -aproximação.
2. É também consequência deste teste que, se  $\theta$  for muito pequeno e o processo for convergente, nas iterações finais do algoritmo, poderá ocorrer a necessidade de um grande número de iterações do método de Benders em cada chamada do oráculo aproximado. Este fato delimita portanto nosso algoritmo a problemas em que a metodologia do método de Benders seja adequada(rápida).
3. O problema mestre relaxado de Benders MB[19] deveria ter alguma heurística para seleção de cortes, tendo em vista que a acumulação de todas as desigualdades explodiria o subproblema.

Apresentaremos a seguir o fluxograma do algoritmo para o problema linear inteiro com a relaxação I.



## 5.6. Sobre a Convergência

### Algoritmo Principal

Os resultados da convergência de [Hiriart-Urruty and Lemaréchal. 93], dados pelos lemas 3.2.1 e 3.2.3, teoremas 3.2.2 e 3.2.4, capítulo XV, vol. II, poderiam ser parcialmente adaptados para o algoritmo de feixes aproximado. Optamos no entanto por apenas observar que para  $\theta$  suficientemente pequeno os resultados citados correspondem à garantia da estabilidade do algoritmo de feixes. Isto pode ser observado através do acréscimo de um parâmetro positivo  $\theta \rightarrow 0$ , na expressão dos erros de linearização, assim como nos ganhos previstos pelo modelo (veja, em particular, o lema 3.2.1, capítulo XV, vol II em [Hiriart-Urruty and Lemaréchal. 93]). Desta forma se garantia apenas a convergência local. Por outro lado, o teste de qualidade da aproximação de  $\varphi$ [21] deve ser suficiente para a da obtenção da convergência global, pois força o processo iterativo a chegar à formulação usual dos feixes, com  $\theta = 0$ . Sem dúvida, com o risco de ser um algoritmo de elevado custo computacional, conforme já foi observado.

## Oráculo Aproximado

A seguir, apresentaremos o conhecido resultado que garante a não ciclagem do algoritmo de Benders

**Teorema 5.1:** Os vetores compostos pelos vértices e respectivos multiplicadores  $(v^p, u^p, \lambda^p)$  gerados a cada iteração pelo algoritmo 5.5.4 são diferentes entre si.

Demonstração: (Similar a [Salkin and Mathur. 89], pág. 360)

Suponhamos os primeiros  $(p \geq 1)$  pontos extremos, digamos  $(v^1, u^1), (v^2, u^2), \dots, (v^p, u^p)$  gerados do problema  $(D)(\lambda)$ [18] e  $\lambda^1, \dots, \lambda^p$  obtidos do problema regularizado  $(F_1)$ [20].

Então do passo 2, tem-se

min  $z$

$$\text{s. a. } z \geq \sum_{j=1}^2 e_j^t y_j + (\lambda^s)^t C_2 y_2 + \sum_{i=1}^4 b_i^t v_i^p + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^p$$

$$z \in \mathbb{R}, y \in Y \text{ e } s \leq p$$

Seja a solução ótima deste problema  $z, y$ , isto é para algum  $k$   $(1 \leq k \leq p)$

$$\begin{aligned} \text{[28] } z &= \sum_{j=1}^2 e_j^t y_j + (\lambda^k)^t C_2 y_2 + \sum_{i=1}^4 b_i^t v_i^k + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^k \geq \sum_{j=1}^2 e_j^t y_j + (\lambda^s)^t C_2 y_2 \\ &+ \sum_{i=1}^4 b_i^t v_i^s + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^s \quad s = 1, \dots, p \end{aligned}$$

Como  $z$  é um limite inferior sobre o custo ótimo primal relaxado

$\varphi$ ,  $\varphi \geq z$ , e de [28]

$$[29] \sum_{i=1}^4 b_i^t v_i^k + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^k \leq \varphi - \left( \sum_{j=1}^2 e_j^t y_j + (\lambda^k)^t C_2 y_2 \right)$$

Por outro lado, na iteração seguinte de (D)( $\lambda^{p+1}$ )[18] a solução ( $v^{p+1}$ ,  $u^{p+1}$ ) é um vértice de  $U(\lambda)$ . Então

$$[30] \sum_{i=1}^4 b_i^t v_i^{p+1} + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^{p+1} = \sum_{k=1}^4 c_k^t x_k + (\lambda^{p+1})^t (C_3 x_3 + C_4 x_4)$$

onde  $x$  é uma solução de (L)( $\lambda$ )[17].

Como  $(x, y)$  é uma solução viável de  $\varphi$  tem-se

$$\sum_{j=1}^2 e_j^t y_j + (\lambda^{p+1})^t C_2 y_2 + \sum_{k=1}^4 c_k^t x_k + (\lambda^{p+1})^t (C_3 x_3 + C_4 x_4) \geq \varphi$$

equivalentemente

$$\varphi - \sum_{j=1}^2 e_j^t y_j \leq \sum_{k=1}^4 c_k^t x_k + (\lambda^{p+1})^t (C_3 x_3 + C_4 x_4) + (\lambda^{p+1})^t C_2 y_2$$

Combinando-se [29] e [30] tem-se

$$\sum_{i=1}^4 b_i^t v_i^k + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^k + (\lambda^k)^t C_2 y_2 \leq \varphi - \sum_{j=1}^2 e_j^t y_j \leq \sum_{k=1}^4 c_k^t x_k +$$

$$(\lambda^{p+1})^t (C_3 x_3 + C_4 x_4) + (\lambda^{p+1})^t C_2 y_2 = \sum_{i=1}^4 b_i^t v_i^{p+1} + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^{p+1} +$$

$$(\lambda^{p+1})^t C_2 y_2.$$

Se  $\varphi - \left( \sum_{k=1}^4 c_k^t x_k + (\lambda^{p+1})^t (C_3 x_3 + C_4 x_4) \right) = \sum_{j=1}^2 e_j^t y_j + (\lambda^{p+1})^t C_2 y_2$  então

$(x, y)$  resolve o problema linear inteiro relaxado  $(\varphi)$ [5].

Caso contrário,

$$[31] \sum_{i=1}^4 b_i^t v_i^k + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^k + (\lambda^k)^t C_2 y_2 < \sum_{i=1}^4 b_i^t v_i^{p+1} + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^{p+1} + (\lambda^{p+1})^t C_2 y_2$$

e neste caso  $(v^k, u^k, \lambda^k) \neq (v^{p+1}, u^{p+1}, \lambda^{p+1})$ . Porém da desigualdade [28]

$$\sum_{i=1}^4 b_i^t v_i^k + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^k + (\lambda^k)^t C_2 y_2 \geq \sum_{i=1}^4 b_i^t v_i^s + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^s + (\lambda^s)^t C_2 y_2 \quad s = 1, \dots, p$$

Por outro lado, de [31] tem-se,

$$\sum_{i=1}^4 b_i^t v_i^s + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^s + (\lambda^s)^t C_2 y_2 < \sum_{i=1}^4 b_i^t v_i^{p+1} + \sum_{i=1}^4 (K_i - B_i y_1)^t u_i^{p+1} + (\lambda^{p+1})^t C_2 y_2 \quad s = 1, \dots, p$$

e assim sendo  $(v^{p+1}, u^{p+1}, \lambda^{p+1}) \neq (v^s, u^s, \lambda^s) \quad s = 1, \dots, p \quad \blacklozenge$

**Corolário 5.2.** Se  $q > 1$  na  $m$  - ésima iteração interna, então  $y^{m+1} \neq y^m$ .

Demonstração: (Similar a [Salkin and Mathur. 89], pág. 362)

Suponhamos o contrário, que ao resolver (MB)[19] com  $m$  cortes, a solução encontrada  $y$  seja repetida. Neste caso, ao resolver do problema (D)[18], obteríamos um vetor  $(v^{m+1}, u^{m+1})$  satisfazendo

$$\sum_{i=1}^4 d_i^t v_i^\ell + \sum_{i=1}^4 (K_i - B_i y_i)^t u_i^\ell = \sum_{i=1}^4 d_i^t v_i^{m+1} + \sum_{i=1}^4 (K_i - B_i y_i)^t u_i^{m+1}$$

para algum  $\ell = \{1, \dots, m\}$ . No entanto, isto só ocorre quando o critério de otimalidade é atingido. ♦



## CAPÍTULO 6

### CONCLUSÕES

Nosso principal objetivo foi apresentar uma técnica alternativa com uso da relaxação lagrangeana na solução de um problema em programação linear inteira.

O trabalho desenvolvido apresentou um novo algoritmo estruturado a partir da relaxação de um bloco de restrições que apresenta dificuldades ao problema, quando abordado por técnicas tradicionais de Benders. Esperamos poder tirar vantagem computacional do processo heurístico regularizador sobre outros algoritmos (Dantzig-Wolfe, subgradientes) pois sua direção de busca é determinada por processos similares ao método de feixes, que, comprovadamente vem apresentando resultados superiores àqueles em diversos problemas de grande porte [Schramm and Zowe. 92]. Parece-nos também que dificilmente a técnica de "Cross Decomposition" seria adaptável, por alguma forma de aproximação, à relaxação I, como o fizemos com o método de feixes.

Como extensões possíveis, consideremos:

1) Adaptar os resultados de convergência capítulo XV, vol. II de [Hiriart-Urruty and Lemaréchal. 93], objetivando obter uma convergência para o algoritmo de feixes aproximado § 5.5.3.

- 2) Implementar os algoritmos § 4.3.1 e § 5.5.3 comparar entre si e com outras relaxações, com o propósito de obter um menor salto de dualidade.
- 3) Implementar a "Cross Decomposition", objetivando avaliar seu desempenho computacional e comparar com os resultados do algoritmo § 4.3.1.
- 4) Investigar outras aplicações com o propósito de verificar a eficiência dos métodos em problemas estruturados.
- 5) Estender a decomposição para problemas não lineares e não lineares inteiros, utilizando-se da relaxação lagrangeana juntamente com o processo heurístico regularizador.

**CAPÍTULO 7****REFERÊNCIAS BIBLIOGRÁFICAS**

[Aardal and Larson. 90] Karen Aardal and T. Larsson, “A Benders Decomposition Based Heuristic for the Hierarchical Production Planning Problem”, *European Journal of Operational Research* 45, pág. 4 - 14, 1990.

[Aardal and Ari. 90] K. Aardal and Aysen Ari, “On the Resemblance Between the Kornai-Liptak and Cross Decomposition Techniques for Block-Angular Linear Programs”, *European Journal of Operational Research* 46, pág.. 393 - 398, 1990.

[Armijo. 66] Armijo, L., Minimization of Function having Continuous Partial Derivatives, *Pacific Journal of Mathematics*, 16, pág. 1-3, 1966.

[Bazaraa et alii. 93] M. S. Bazaraa, H. D. Sherali and C. M. Sherali, “Nonlinear Programming”, Second Edition, John Wiley & Sons, Inc, 1993.

[Benders. 62] J. F. Benders, “Partitioning Procedures for Solving Mixed-Variables Programming Problems”, *Numerische Mathematik* 4, pág. 238 - 252, 1962.

[Bertsimas and Orlin. 94] D. Bertsimas and J. B. Orlin, "A Technique for Speeding up the Solution of the Lagrangean Dual", *Mathematical Programming* 63(1), pág.. 23-45, 1994.

[Cabral. 93] L. A. F. Cabral, "Uma Variante do Método de Decomposição de Benders", Tese de Mestrado, COPPE/UFRJ - Universidade Federal do Rio de Janeiro, PESC, 1993.

[Cheney and Goldstein. 59] E. W. Cheney and A. A. Goldstein, "Newton's Methods for Convex Programming and Tchebycheff Approximation", *Numerische Mathematik*, 1, pág.. 253 - 268, 1959.

[CPLEX3.0 97] "Using the CPLEX<sup>TM</sup> Linear Optimizer", CPLEX Optimization Inc., 1997.

[Coté and Laughton. 84] G. Coté and M. A. Laughton, "Large-scale Mixed Integer Programming: Benders-type Heuristics", *European Journal of Operational Research* 16, pág.. 327 - 333, 1984.

[Dantzig. 63] G. B. Dantzig, "Linear Programming and Extensions". Princeton University Press, Princeton, NJ. 1963.

[Dantzig and Wolfe. 60] G. B. Dantzig and P. Wolfe, "Decomposition Principle for Linear Programs", *Operations Research* 8, pág. 101-111, 1960.

[Ferland and Florian. 79] J. A. Ferland and M. Florian, "A Sub-optimal Algorithm to Solve a Large Scale 0-1 Programming Problem", in: A. Prekopa(ed.), Survey of Mathematical Programming, North-Holland, Amsterdam, pág.. 461- 469, 1979.

[Fisher and Jaikumar. 78] M. L. Fisher and R. A. Jaikumar, "Decomposition Algorithm for Large Scale Vehicle Routing", Working Paper 78 -11- 05, Dept. of Decision Sciences, University of Pennsylvania.

[Fisher. 81] M. L. Fisher, "The Lagrangian Relaxation Method for Solving Integer Programming Problems", Management Science 27(1), January 1981.

[Geoffrion. 72] A. M. Geoffrion, "Generalized Benders Decomposition", Journal of Optimization Theory and Applications 10, pág.. 237 - 260, 1972.

[Geoffrion. 74] A. M. Geoffrion, "Lagrangian Relaxation for Integer Programming", Mathematical Programming Study 2, pág.. 82 - 114, 1974.

[Held and Karp. 70] M. Held and R. M. Karp, "The Traveling Salesman Problem and Minimum Spanning Trees", Operations Research 18, pág. 1138 - 1162, 1970.

[Held and Karp. 71] M. Held and R. M. Karp, "The Traveling Salesman Problem and Minimum Spanning Trees: Part II", *Mathematical Programming* 1, pág. 6 - 25, 1971.

[Held et alii. 74] M. Held, P. Wolfe and H. P. Crowder, "Validation of Subgradient Optimization", *Mathematical Programming* 6, pág. 62 - 88, 1974.

[Hiriart - Urruty and Lemaréchal. 93] J. B. Hiriart - Urruty, and C. Lemaréchal, "Convex Analysis and Minimization Algorithms" Vol I e II, Springer - Verlag, 1993.

[Hoc. 82] Hoang Hai Hoc, "Topological Optimization of Networks: A Non-linear Mixed Integer Model Employing Generalized Benders Decomposition", *IEEE Transactions on Automatic Control* 27, pág. 164 - 169, 1982.

[Holmberg and Jörnsten. 84a] Kaj Holmberg and K. O. Jörnsten, "Cross Decomposition Applied to the Stochastic Transportation Problem", *European Journal of Operational Research* 17(3), pág. 361 - 368, 1984a.

[Holmberg. 85] Kaj Homberg, "Decompositon in Large Scale Mathematical Programming" PhD Dissertation, Linköping, Institute of Technology (Linköping, 1985).

[Holmberg. 89] Kaj Homberg, “On the Use of Valid Inequalities in Benders and Cross Decomposition”, Technical-Report LiTh-MAT-R-89-21, Linköping Institute of Technology, Sweden.

[Holmberg. 90] Kaj. Holmberg, “Computational Experiments with Primal-Dual Solution Methods for the Produce-or-Purchase Problem” Technical-Report LiTH-MAT-R-90-20, Linköping Institute of Technology, Sweden.

[Holmberg. 90a] Kaj. Holmberg, “On the Convergence of Cross Decomposition”, *Mathematical Programming* 47, pág. 269 - 296, 1990a.

[Holmberg and Jörnsten. 91a] Kaj Holmberg and Kurt Jörnsten, “Decomposition Methods for the Exact Formulation of the Simple Location Problem with Spatial Interaction”, Technical-Report LiTH-MAT/OPT-91-21, Linköping Institute of Technology, Sweden.

[Holmberg. 92a] Kaj. Holmberg, “Generalized Cross Decomposition applied to Nonlinear Integer Programming Problems: Duality Gaps and Convexification in Parts”, *Optimization* 23, pág. 341 - 356, 1992.

[Holmberg. 92b] Kaj. Holmberg, “Linear Mean Value Cross Decomposition: A Generalization of the Kornai-Liptak Method”, *European Journal of Operational Research* 62, pág. 55 - 73, 1992.

[Holmberg and Jörnsten. 93] Kaj Holmberg and Kurt Jörnsten, "Solving the Generalized Knapsack Problem with Variable Coefficients", Technical-Report LiTH-MAT-R-93-19, Linköping Institute of Technology, Sweden.

[Holmberg. 94a] Kaj. Holmberg, "Cross Decomposition Applied to Integer Programming Problems: Duality Gaps and Convexification in Parts", Operations Research 42(4), pág. 657 - 668, 1994.

[Holmberg. 94b] Kaj. Holmberg, "On Using Approximations of the Benders Master Problem", European Journal of Operational Research 77, pág. 657 - 668, 1994.

[Holmberg. 94c] Kaj. Holmberg, "Solving the Staircase Cost Facility Location Problem with Decomposition and Piecewise Linearization", European Journal of Operational Research 77, pág. 41 - 61, 1994.

[Holmberg. 94d] Kaj. Holmberg, "Primal and Dual Decomposition as Organizational Design: Price and/or Resource Directive Decomposition" Technical-Report LiTH-MAT-R-94-04, Linköping Institute of Technology, Sweden.

[Holmberg. 94e] Kaj. Holmberg, "A Convergence Proof for Linear Mean Value Cross Decomposition", Zeitschrift für Operational Research 39, pág. 157-186, 1994.



[Holmberg. 96a] Kaj Homberg, “Lagrangean Heuristics for Linear Cost Multicommodity Network Flow Problems”, Technical-Report LiTH-MAT/OPT-WP-95-01, Linköping Institute of Technology, Sweden.

[Holmberg and Yuan. 96b] Kaj Homberg and Di Yuan, “A Lagrangean Heuristics Based Branch-and-Bound Approach for the Capacitated Network Design Problem”, Technical-Report LiTH-MAT-R-1996-23, Linköping Institute of Technology, Sweden.

[Holmberg and Hellstrand. 96c] Kaj Homberg and Johan Hellstrand, “Solving the Uncapacitated Network Design Problem by a Lagrangean Heuristic and Branch-and-Bound”, Technical-Report LiTH-MAT-R-94-11, Linköping Institute of Technology, Sweden.

[Holmberg. 97a] Kaj Holmberg, “Mean Value Cross Decomposition Applied to Integer Programming Problems”, *European Journal of Operational Research* 97: pág. 124-138, 1997.

[Holmberg and Ling. 97b] Kaj Homberg and Jonas Ling, “A Lagrangean Heuristics for the Facility Location Problem with Staircase Costs”, *European Journal of Operational Research* 97: pág. 63-74, 1997.

[Holmberg et alii. 97] Kaj Homberg, Mikael Ronnqvist and Di Yuan, "Solving Capacitated Facility Location Problems with Single Sourcing using a Lagrangean Heuristic, Repeated matching and Branch-and-Bound", Technical-Report LiTH-MAT-R-1997-05, Linköping Institute of Technology, Sweden.

[Khachian. 79] L. G. Khachian, "A Polynomial Algorithm for Linear Programming", Soviet Mathematics Doklady 20, pág. 191-194, 1979.

[Kelley. 60] J. E. Kelley, "The Cutting Plane Method for Solving Convex Problems", SIAM J., 8, pág. 703 - 712, 1960.

[Kim et alii. 89] Sehun Kim, Seong Cheol Cho and Bong Sik Um, "A Simplified Cross Decomposition Algorithm for Multiple Right Hand Choice Linear Programming", Journal of the Operations Research Society of Japan 32(4), pág. 441- 449, 1989.

[Kiwiel. 85] K. C. Kiwiel, "Methods of Descent for Nondifferentiable Optimization", Lecture Notes in Mathematics 1133, Springer-Verlag, Berlin, 1985.

[Kiwiel. 86] K. C. Kiwiel, " A Method for Solving Certain Quadratic Programming Problems Arising in Nonsmooth Optimization, IMA Journal. Numerical Analysis (6), pág. 137-152, 1986.

[Kiwiel and Stachurski. 88] K. C. Kiwiel and A. Stachurski, "NOA - A Fortran Package of Nondifferentiable Optimization Algorithms", Systems Research Institute, Polish Academy of Sciences, Newelska 6, pág.. 01- 447, Warsaw, Poland, 1988.

[Kiwiel. 90] K. C. Kiwiel, "Proximity Control in Bundle Methods for Convex Nondifferentiable Minimization", *Mathematical Programming* 46, pág. 105 - 122, 1990.

[Kiwiel. 92] K. C. Kiwiel, "Proximal Level Bundle Methods for Convex Nondifferentiable Optimization, Saddle-Point Problems and Variational Inequalities", Research Report N<sup>o</sup> 1742, INRIA, Rocquencourt, 1992.

[Kiwiel. 95] K. C. Kiwiel, "Approximations in Proximal Bundle Methods and Decomposition of Convex Programs" *Journal of Optimization Theory and Applications* 84(3), pág. 529 - 548, march 1995.

[Kohl et alii. 95] Niklas Kohl and Oli B. G. Madsen, "An Optimization Algorithm for the Vehicle Routing Problem with Time Windows based on Lagrangean Relaxation" Technical-Report, Institute of Mathematical Modelling, Lyngby, Denmark, 1995.

[Kornai and Liptak. 65] J. Kornai and T. Liptak, "Two-level Planning", *Econometrica* 33, pág. 141- 191, 1965.

[Lee. 94] Eva Kwork-Yin Lee, “Parallelism in Mixed Integer Programming” present at the International Symposium on Mathematical Programming, Michigan, USA, 1994.

[Lee Y. 93] Choong Y. Lee, “A Cross Decomposition Algorithm for Multiproduct-Multitype Facility Location Problem”, *Computers Operations Research* 20(5), pág. 527-540, 1993.

[Lemaréchal. 75] C. Lemaréchal, “An Extension of Davidon Methods to Non Differentiable Problems”, in *Nondifferentiable Optimization, Mathematical Programming Study 3*”, (Balinski M. L. and Wolfe P.), pág. 95 - 109, 1975.

[Lemaréchal. 76] C. Lemaréchal, “Nondifferentiable Optimization, Subgradient and  $\varepsilon$  - subgradient Methods”, in *Optimization and Operations Research, Lecture Notes in Economics and Mathematical Systems 117*, Springer-Verlag, Berlin, pág. 191 - 199, 1976.

[Lemaréchal and Bancora. 85] C. Lemaréchal and M. Bancora Imbert, “Le Module M1FC1”, Tech. Report, Institut de Recherche d’Informatique et d’Automatique, Le Chesnay, 1985.

[Lemaréchal et alii. 91] C. Lemaréchal, A. Nemirovskii e Y. Nesterov., “New Variants of Bundle Methods”, Research Report N<sup>o</sup> 1508, INRIA, Rocquencourt, 1991.

[Lemaréchal. 92] C. Lemaréchal, “Lagrangian Decomposition and Nonsmooth Optimization: Bundle Algorithm, Prox Iteration, Augmented Lagrangian, Nonsmooth Optimization: Methods and Applications,” Edited by F. Giannessi, Gordon and Breach, Philadelphia, Pennsylvania, pág. 201-216, 1992.

[Lemaréchal and Sagastizábal. 95] C. Lemaréchal and C. Sagastizábal, “Application of Bundle Methods to the Unit-Commitment Problem”, Rapport Technique n<sup>o</sup> 0184, Institut de Recherche d’Informatique et d’Automatique, Le Chesnay, Novembre 1995.

[Lemaréchal 96] C. Lemaréchal, Interview with Claude Lemaréchal, Optima, 1996

[Lemaréchal and Sagastizábal. 97] C. Lemaréchal and C. Sagastizábal, “Variable Metric Bundle Methods: from Conceptual to Implementable Form”, Mathematical Programming 76, pág. 393-410, 1997.

[Mäkelä and Neittaanmäki. 92] M. M. Mäkelä and P. Neittaanmäki, “Nonsmooth Optimization”, World Scientific, 1992.

[McDaniel and Devine. 77] D. McDaniel and M. Devine, “A Modified Benders’ Partitioning Algorithm for Mixed Integer Programming”, Management Science 24, pág. 312 - 319, 1977.

[Medhi. 91] Deepankar Medhi, "Decomposition of Structured Large scale Optimization Problems and Parallel Optimization", PhD Thesis, University of Wisconsin, (1987).

[Minoux. 84] M. Minoux, "Subgradient Optimization and Benders Decomposition for Large Scale Programming", *Mathematical Programming*, R. W. Cottle, M. L. Kelmanson and B. Korte (Editors), 1984.

[Minoux. 86] M. Minoux, "Mathematical Programming", *Theory and Algorithms*, John Wiley and Sons, 1986.

[Moreno et alii]. 92] A. O. Moreno, E. Vidal Jr., C. A. Santos, P. R. Oliveira e S.I.M. Rodrigues, "Planejamento da Configuração Otimizada da Rede de Transmissão Digital", *Revista TELEBRÁS*, pág. 20 - 34, 1992.

[Nemhauser et alii. 94] G. L. Nemhauser, M. W. P. Savelsbergh and G. C. Sigismondi. "MINTO, a Mixed INTeger Optimizer", *Operations Research Letters* 15, pág. 47 - 58, 1994.

[Nemhauser. 94] G. L. Nemhauser, "Branch-and-Cut: Column Generation for Solving Huge Integer Programs" present at the International Symposium on Mathematical Programming, Michigan, USA, 1994.

[OSL 1.2. 90] "Optimization Subroutine Library, Guide and Reference". IBM Corporation 1990.

[Oustrata et alii. 91] Jiri Oustrata, Helga Schramm and Jochem Zowe, "Bundle Trust Methods: Fortran Codes for Nondifferentiable Optimization User's Guide", Report N<sup>o</sup> 269, Universität Bayreuth, 1991.

[Paula Jr. and Maculan. 88] G. G. Paula Jr. and Nelson Maculan, "A p-Median Location Algorithm based on the Convex Lagrangean Relation of the Benders master problem", present at the International Symposium on Mathematical Programming, Tokyo, Japan, 1988.

[Pchénitchny and Daniline. 65]. B. Pchénitchny and Y. Daniline, "Methodes Numériques dans les Problèmes d'Extrémum", Mir, Moscou, 1977.

[Poljak. 67] B. T. Poljak, "A General Method of Solving Extremum Problems", Soviet Mathematics Doklady 8, pág. 593 - 597, 1967.

[Poljak. 69] B. T. Poljak, "Minimization of Unsmooth Functionals", USSR Computational and Mathematical Physics 9, pág. 14 - 29, 1969.

[Qi and Chen. 97] Liqun Qi and Xiaojun Chen, "A Preconditioning Proximal Newton Method form Nondifferentiable Convex Optimization", Mathematical Programming 76, pág. 411-429, 1997.

[Rana and Vickson. 88] K. Rana and R. G. Vickson, "A Model and Solution Algorithm for Optimal Routing of a Time - Chartered Containership", Transportation Science 22, pág. 83 - 95, 1988.

[Rockafellar. 72] R. Tiriell Rockafellar, "Convex Analysis", Princeton University Press, 1972.

[Salkin and Mathur. 89] H. M. Salkin and K. Mathur, "Foundations of Integer Programming", North - Holland, 1989.

[Schramm. 89] H. Schramm, "Eine Kombination von Bundle-und Trust-Region-Verfahren zur Lösung Nichtdifferenzierbarer Optimierungsprobleme", Bayreuther Mathematische Schriften, Helf 30, Bayreuth, 1989.

[Schramm and Zowe. 87] H. Schramm and J. Zowe, "A Combination of the Bundle Approach and the Trust Region Concept", Report N<sup>o</sup> 20, Universität Bayreuth, 1987.

[Schramm and Zowe. 92] H. Schramm and J. Zowe, "A Version of the Bundle Idea for Minimizing a Nonsmooth Function: Conceptual Idea, Convergence Analysis, Numerical Results", Siam J. Optimization, Vol. 2, N<sup>o</sup> 1, pág.. 121 - 152, February, 1992.

[SIAG/OPT. 94] SIAG/OPT Views-and-News - A Forum for the SIAM Activity Group on Optimization, Larry Nazareth Editor, N<sup>o</sup> 4, Spring - Verlag, 1994.



[Shor. 70] N. Z. Shor, "Convergence Rate of the Gradient Descent Method with Dilatation of the Space", *Cybernetics* 6(2), pág.. 102-108, 1970.

[Shor. 85] N. Z. Shor, "Minimization Methods for Non - Differentiable Functions", Springer-Verlag, Berlin, 1985.

[Vaidya. 90] P. Vaidya, "A New Algorithm for Minimizing Convex Functions Over Convex Sets", AT&T Bell Laboratories, Murray Hill, NJ. 1990.

[Van Roy. 80] Tony J. Van Roy, "Cross Decomposition for Large - Scale Mixed Integer Linear Programming with Applications to Facility Location on Distribution Networks", Doctoral Dissertation, Applied Sciences, Katholieke Universiteit Leuven, 1980.

[Van Roy. 83] T. J. Van Roy, "Cross Decomposition for Mixed Integer Programming", *Mathematical Programming* 25, pág. 46 - 63, 1983.

[Van Roy. 86] T. J. Van Roy, "A Cross Decomposition Algorithm for Capacited Facility Location", *Operations Research* 34(1), pág. 145-163, 1986.

[Wolfe. 75] P. Wolfe, "A Method of Conjugate Subgradient for Minimizing Nondifferentiable Functions", in *Nondifferentiable Optimization, Mathematical Programming Study 3*, (Eds. Balinski M. L. and Wolfe P.), pág. 145 - 173, 1975.

[Wong. 78] R. T. Wong, "Acceletating Benders Decomposition for Network Design", Ph.D. thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, (January), 1978.