unesp **UNIVERSIDADE ESTADUAL PAULISTA**
Instituto de Biociências, Letras e Ciências Exatas
DEPARTAMENTO DE CIÊNCIAS DE COMPUTAÇÃO E ESTATÍSTICA

HEURÍSTICA *SURROGATE* PARA O PROBLEMA DE CARREGAMENTO DE PALETES DO PRODUTOR

Bruna de Lima Alcântara Kitamura

Dissertação de Mestrado
Pós-Graduação em Matemática

Rua Cristovão Colombo, 2265
15054-000 - São José do Rio Preto - SP - Brasil - 2009
Telefone: (017) 3221-2444
Fax: (017) 3221-2445

HEURÍSTICA *SURROGATE* PARA O PROBLEMA DE CARREGAMENTO DE PALETES DO PRODUTOR

Bruna de Lima Alcântara Kitamura [1]

Dissertação apresentada ao Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de São José do Rio Preto, São Paulo, como parte dos requisitos para a obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. Silvio Alexandre de Araujo
Agências Financiadoras: CNPq e FAPESP

São José do Rio Preto
2009

[1]Contato:brunakitamura@yahoo.com.br

Kitamura, Bruna de Lima Alcântara.
Heurística *Surrogate* para o problema de carregamento de paletes do produtor / Bruna de Lima Alcântara Kitamura. - São José do Rio Preto : [s.n.], 2009.
109 f. : il. ; 30 cm.

Orientador: Silvio Alexandre de Araujo
Dissertação (mestrado) - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas

1. Otimização matemática. 2. Pesquisa operacional. 3. Programação inteira. 4. Carregamento de paletes. 5. Paletes. 6. Heurístrica surrogate. I. Araujo, Silvio Alexandre de. II. Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU - 519.8

Ficha catalográfica elaborada pela Biblioteca do IBILCE
Campus de São José do Rio Preto - UNESP

Bruna de Lima Alcântara Kitamura

HEURÍSTICA *SURROGATE* PARA O PROBLEMA DE CARREGAMENTO DE PALETES DO PRODUTOR

Dissertação apresentada para a obtenção do título de Mestre em Matemática do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de São José do Rio Preto.

BANCA EXAMINADORA

Prof. Dr. Silvio Alexandre de Araujo
Professor Assistente Doutor
UNESP - São José do Rio Preto

Prof. Dr. Reinaldo Morabito
Professor Livre Docente
Universidade Federal de São Carlos - UFSCar

Prof. Dr. Geraldo Nunes Silva
Professor Livre Docente
UNESP - São José do Rio Preto

São José do Rio Preto, 2 de fevereiro de 2009.

Aos meus avós maternos, Irineu e Elza, e paternos, Valdir (*in memoriam*) e Alice.

Aos meus pais Valdir e Sandra.

Ao meu marido Rodrigo.

Dedico.

Agradecimentos

A Deus, sempre, pela vida, oportunidades e conquistas.

Ao meu professor Silvio Araujo pela orientação, dedicação e conselhos durante toda a elaboração deste trabalho. Obrigada pela compreensão e auxílio nos momentos mais difíceis.

Ao meu marido, Rodrigo, por apoiar e incentivar sempre meus hábitos ao estudo e crescimento profissional e intelectual. Por compartilhar comigo mais esta etapa de nossas vidas de forma tão paciente e amorosa. Obrigada por ser meu companheiro em todos os momentos e tão único em minha vida!

Aos meus pais pela compreensão e por, mesmo sem entender, me apoiarem em tudo que faço com tanto amor e carinho. Por serem pais tão presentes e por darem sempre educação e dedicação incondicionais. Por acreditarem em mim, quando eu mesma desacreditava. E por serem inesquecíveis sempre!

A todos os amigos, conquistados durante a vida, que tanto me auxiliaram no decorrer destes dois anos. Em especial gostaria de agradecer à Fernanda por sempre estar presente nas horas que mais precisei. Por, mesmo estando tão longe, ser minha confidente de todas as horas e auxiliar na análise deste trabalho. Em especial também, nunca poderei agradecer o suficiente à Marina, por me receber de braços abertos no início do curso e por facilitar tanto as noites em claro de estudo. Obrigada por ser uma amiga tão generosa e companheira.

A todos os professores que colaboraram para a minha formação.

À CNPq e à FAPESP pelo auxílio financeiro.

"A Matemática apresenta invenções tão sutis que poderão servir não só para satisfazer os curiosos como, também para auxiliar as artes e poupar trabalho aos homens."

Descartes

Resumo

O objetivo deste trabalho é estudar um caso particular dos problemas de corte e empacotamento, denominado Problema de Carregamento de Paletes do Produtor. Inicialmente, uma formulação proposta na literatura é avaliada com um pacote computacional. Posteriormente, as heurísticas lagrangiana e *surrogate* são estudadas e um método de atualização dos multiplicadores *surrogate* é adaptado para este problema.

A importância em se estudar o Problema de Carregamento de Paletes do Produtor é que, devido à escala e extensão de certos sistemas logísticos, um pequeno aumento do número de produtos a serem carregados sobre cada palete pode resultar em economias substanciais. A motivação em se estudar o método de atualização *surrogate* proposto é que, além da adaptação do presente trabalho não ter sido realizada na literatura, uma posterior aplicação desta heurística em conjunto com um procedimento *branch and bound* poderá render melhores resultados que outras heurísticas.

Palavras-chave: Problema de Carregamento de Paletes do Produtor; Heurística *Surrogate*; Atualização dos Multiplicadores *Surrogate*.

Abstract

The aim of this work is studying a particular case of cutting and packing problem, so-called the Manufacturer's Pallet Loading Problem. Initially, a formulation proposed in the literature is evaluated with a computer package. Subsequently, the lagrangian and *surrogate* heuristics are studied and a method to update the *surrogate* multiplier is adapted for this problem.

The importance of studying the manufacturer's pallet loading problem is that, due to the scale and scope of some logistics systems, a small increase in the number of products to be loaded on each pallet can result in substantial savings. The motivation of studying the proposed method of updating the *surrogate* multipliers is that, besides the adaptation of this work has not been carried out in the literature, further application of heuristics within a procedure *branch and bound* can yield better results than other heuristics.

Keywords: Manufacturer's Pallet Loading Problem; *Surrogate* Heuristic ; Update of the *Surrogate* Multipliers.

Sumário

Capítulo 1

Introdução

Algumas décadas depois da Revolução Industrial, a possibilidade de fabricar produtos em larga escala com preços inferiores trouxe a necessidade da capacidade de estocagem em maiores quantidades à logística empresarial. Com o aumento da demanda destes produtos e com o problema de serem fabricados em locais diferentes, foram exigidas soluções em como transportar estes produtos até os clientes em potencial de forma a minimizar o custo da empresa e maximizar a satisfação do cliente.

Algumas destas soluções foram, por exemplo, a de se estocar uma única vez um produto, por exemplo em um armazém de distribuição em local intermediário entre a empresa e o cliente, pois o custo de estocagem representa de um a dois terços dos custos logísticos totais da empresa; outra solução foi a unitização da carga, isto é, vários itens arranjados de forma a serem manuseados como um único volume; bem como a paletização da carga, que tem como característica reunir vários itens em uma plataforma, em geral de madeira, disposta horizontalmente, a fim de que a carga possa ser empilhada e estabilizada.

Na paletização da carga, há um melhor controle de estoque, a carga fica protegida contra roubos e danos físicos e pode ser transportada de uma só vez por meio de empilhadeiras, guindastes, caminhões, diminuindo o tempo de carga e descarga [1]. Dessa solução, em particular, para o problema em como transportar os produtos de uma indústria à um cliente em potencial, surge o problema de carregamento de paletes. Tal problema consiste em otimizar o aproveitamento do palete ao se carregar produtos embalados em caixas sobre um palete retangular.

O problema de carregamento de paletes é mais complexo do que aparenta, pois, dentre outras dificuldades existentes, o aumento cada vez maior do número de caixas a serem dispostas sobre um palete aumenta a dificuldade de sua resolução. Isto é, apesar da complexidade computacional destes problemas ainda ser desconhecida, a dificuldade na resolução aumenta proporcionalmente à razão entre as dimensões do palete e as dimensões da caixa. Os algoritmos exatos conhecidos, bem como os softwares utilizados na resolução do problema, apresentam complexidade exponencial sobre o tempo computacional, pois utilizam métodos de busca em árvore que podem chegar a 2^n resoluções de subproblemas, onde n é o número de caixas como solução ótima. Por isso, sua resolução é geralmente

não identificada por modelos exatos, principalmente para problemas que possuem como solução ótima um número elevado de caixas, tornando-se viável e de completo interesse o desenvolvimento e utilização de novas heurísticas para sua resolução.

Ao estudar o problema em 1982, Hodgson [2] dividiu-o em dois casos: o problema de carregamento de paletes do produtor, PCP do Produtor (*Manufacturer's Pallet Loading Problems*), onde os paletes são carregados somente com caixas iguais; e o problema de carregamento de paletes do distribuidor, PCP do distribuidor (*Distribuidor's Pallet Loading Problem*), onde existem caixas de diferentes dimensões.

No primeiro caso, que é o de interesse para este trabalho, o produtor fabrica um item que é embalado em caixas iguais de dimensões (l, w, h), onde l representa o comprimento, w, a largura e h, a altura da caixa. Estas caixas devem ser então carregadas em camadas horizontais sobre paletes de dimensões (L, W, H) onde L e W são o comprimento e largura do palete, respectivamente, e H é a altura máxima tolerável do carregamento. O problema pode ser decomposto em dois subproblemas: (i) para cada face das caixas, o problema bidimensional de arranjar ortogonalmente o máximo número de retângulos (l, w) (faces das caixas) dentro do retângulo (L, W) (superfície do palete), e (ii) o problema unidimensional de arranjar o máximo número de camadas ao longo da altura H do palete.

O principal objetivo deste trabalho é estudar o PCP do Produtor, em particular o subproblema (i). Fazer referência ao PCP do Produtor neste presente trabalho será tratar especificamente o subproblema (i) descrito acima. Modelos propostos e adaptados para a resolução do problema na literatura serão estudados, bem como procedimentos heurísticos que resolvam o problema de forma aproximada e que podem ser utilizados em métodos exatos de enumeração implícita.

Dentre as várias justificativas em se estudar estes problemas, uma delas se deve ao fato de que uma análise estatística baseada na solução do PCP do Produtor pode ser útil para escolher o palete mais adequado para uma empresa (dentre um conjunto de paletes padronizados disponíveis), assim como para projetar novos paletes [3]. Além disso, métodos mais efetivos para resolver o PCP do Produtor permitem encontrar melhores padrões de empacotamento para a armazenagem e distribuição de um produto, sem mencionar que com um programa computacional eficaz, a geração de padrões para cada produto, pode ser desempenhada com maior rapidez e flexibilidade em cada centro de distribuição, junto com o pessoal encarregado da montagem da carga paletizada [4].

A computação eficiente de limitantes bem apertados é o interesse primário de qualquer procedimento *branch and bound* aplicado à resolução de problemas de programação inteira. Muitas aproximações *branch and bound* bem sucedidas utilizam relaxação linear para limitar a solução ótima deste problema. Neste trabalho, interesse significante foi dado aos duais lagrangiano e *surrogate* como origem alternativa de limintes. A existência de técnicas eficientes, tais como o *método de otimização do subgradiente* para o caso lagrangiano, tem levado a ótimas aplicações da dualidade lagrangiana na solução de problemas especialmente estruturados. Enquanto o dual *surrogate* tem sido mostrado na

teoria fornecer limitantes mais fortes, a dificuldade na busca de multiplicadores duais *surrogate* tem desecorajado sua utilização na resolução de problemas inteiros. Baseados no desenvolvimento de novas relações entre a dualidade lagrangiana e *surrogate*, é sugerida uma nova estratégia para computar valores duais *surrogate*. A aproximação proposta permite utilizar diretamente métodos de busca lagrangiana pré-estabelecidos para explorar multiplicadores duais *surrogate* [5].

No Capítulo 2, primeiramente são fornecidas noções sobre o problema de corte e empacotamento e a identificação do PCP do Produtor segundo às classificações de Harald Dyckhoff [6]. Também é mostrada a equivalência deste problema aos problemas de corte e empacotamento, bem como considerações sobre o problema ser do tipo não-guilhotinado. Em seguida, são apresentadas algumas formulações matemáticas que foram adaptadas para o PCP do Produtor (não-guilhotinado) pelo trabalho de Oliveira [4]. Dentre eles, são descritos o modelo de Beasley [8], de Tsai *et al.* [9], de Hadjiconstantinou e Christofides [10] e de Chen *et al.* [11].

No Capítulo 3, o modelo de Beasley é utilizado para dar continuidade ao presente trabalho, sendo fornecidas justificativas para tal escolha. São apresentadas melhorias propostas na literatura, com relação à redução das variáveis e restrições desta formulação, bem como dois exemplos numéricos ilustrando tais melhorias e o modelo de Beasley. Exemplos da literatura são então analisados e comparados ao trabalho de Oliveira [4], verificando a dificuldade dos softwares matemáticos em resolver o PCP do Produtor para problemas que envolvem um número maior de caixas como solução ótima.

Um desenvolvimento teórico é então inicializado pelo Capítulo 4 para problemas genéricos de maximização. A dualidade lagrangiana e a dualidade *surrogate* são apresentadas juntamente com suas propriedades de integralidade. Exemplos numéricos para o caso de maximização, pouco encontrados na literatura, são apresentados para ilustrar a necessidade da busca de multiplicadores duais ótimos para as relaxações lagrangiana e *surrogate*, respectivamente. Resultados encontrados em trabalhos da literatura para problemas de minimização são então adaptados e suas respectivas demonstrações são desenvolvidas independentemente de outros trabalhos ou, então, são adaptadas de demonstrações existentes para o caso de minimização. Algumas relações entre os problemas duais lagrangiano e *surrogate*, bem como suas demonstrações, são adaptadas de Karwan e Rardin [12], também para problemas de maximização. Finalmente, um contra-exemplo ilustra a dificuldade encontrada na obtenção de multiplicadores ótimos *surrogate* desde que o subgradiente não é necessariamente direção de decrescimento da relaxação *surrogate*.

Métodos de solução são apresentados no Capítulo 5 para resolução do dual lagrangiano e do dual *surrogate*: o método de otimização do subgradiente aplicado ao dual lagrangiano e aplicado ao dual *surrogate* e o método de Sarin *et al.* aplicado ao dual *surrogate*. São mostradas as convergências ou não convergências destes métodos e as justificativas para que tais propriedades aconteçam. O Método de Sarin *et al.* desenvolvido em [5] para problemas de minimização é adaptado com mais minúcias para problemas genéricos de

maximização, bem como as idéias implícitas e resultados deste mesmo trabalho.

No Capítulo 6, toda a teoria anterior desenvolvida é aplicada ao PCP do Produtor: as relaxações lagrangiana e *surrogate* e os respectivos problemas duais associados, a verificação da propriedade de integralidade lagrangiana e a dificuldade na demonstração da propriedade de integralidade *surrogate*, bem como um novo algoritmo de busca dual *surrogate* baseado nas idéias e algoritmo de Sarin *et al.* [5]. Um exemplo numérico finaliza o capítulo, ilustrando a aplicação deste algoritmo, ainda não desenvolvido na literatura, para o PCP do Produtor.

Finalmente no Capítulo 7 o trabalho é concluído e são apresentadas algumas propostas futuras.

Capítulo 2

Modelos Matemáticos para o PCP do Produtor

2.1 Introdução

No ambiente de produção de algumas indústrias, freqüentemente é necessário cortar painéis grandes (objetos) em partes menores (itens). Este problema é conhecido como *problema de corte* (*cutting problem*). É importante que os cortes sejam planejados, pois, assim, os efeitos negativos, tais como o desperdício de matéria-prima, podem ser minimizados, diminuindo os custos de produção. Também na indústria ocorre o *problema de empacotamento* (*packing problem*), que consiste em alocar, de uma forma econômica, uma coleção de itens dentro do objeto. Empacotar itens (por exemplo, caixas) em objetos maiores (por exemplo, paletes) e cortar objetos (por exemplo, couro) para produzir itens menores (por exemplo, carteiras) são problemas que apresentam uma estrutura lógica similar. Por isto, problemas de alocar ou cortar itens são referenciados na literatura simplesmente como problemas de corte e empacotamento [13].

Ainda segundo [13], dada a grande diversidade de tipos de problema que ocorrem na prática, Harald Dyckhoff sugeriu em 1990 [6] uma maneira de sistematizá-los. Apesar de serem completamente diferentes do ponto de vista prático, os problemas de corte e empacotamento são abordados matematicamente com as mesmas formulações e estratégias de resolução. Assim, as quatro principais características propostas no trabalho de Dyckhoff [6] para classificar problemas de corte são: dimensão, tipo de alocação, sortimento de objetos e sortimento de itens.

Com respeito à dimensão, o problema *bidimensional* é representado pelo algarismo (2) e é utilizado quando duas dimensões são relevantes no processo de corte, como é o caso do PCP do Produtor. Com respeito ao tipo de alocação (seleção dos objetos e itens), este problema possui itens e objetos selecionados de acordo com a seguinte combinação: não há seleção de objetos e há seleção de itens (B). Neste caso, a quantidade de objetos existentes em estoque não é suficiente para atender todos os itens demandados e, com

isto, alguns itens não são selecionados. Com respeito ao sortimento de objetos, que é na verdade o aspecto dos objetos, Dyckhoff classificou um único objeto (palete no PCP do Produtor) representando-o por (O); bem como classificou para o sortimento de itens (aspecto dos itens), a representação (C) como itens de tamanhos iguais (caixas no PCP do Produtor). Mais detalhes sobre as classificações podem ser encontradas em [6] e em [13]. Assim, de acordo com a tipologia de Dyckhoff [6], o PCP do Produtor é classificado em $2/B/0/C$, isto é, (2) bidimensional, (B) alocando apenas uma seleção de itens (O) em um único objeto grande (palete) e (C) de forma que os itens (caixas) possuam tamanhos iguais.

Em 2007, Wäscher *et al.* [7] aperfeiçoaram a tipologia de Dyckhoff para que outros problemas de corte fizessem parte da classificação proposta. Cinco critérios básicos são utilizados para a classificação dos problemas: dimensionalidade, seleção de itens e objetos, sortimento de objetos, sortimento de itens e forma dos itens. Adicionalmente são consideradas algumas variações para tratar casos específicos. Na dimensionalidade, o PCP do Produtor é considerado como bidimensional, enquanto que para a seleção de itens e objetos, é classificado sobre a maximização da produção de item (*output maximization*), isto é, em que um conjuno de itens pequenos (caixas) deve ser designado para um conjunto de objetos grandes (palete), mas este não é suficiente para acomodar todos os itens. Em relação ao sortimento de itens, o PCP do Produtor é classificado em itens idênticos e em relação ao sortimento de objetos, como um único tipo de objeto que terá dimensões fixas. A forma dos itens é considerada regular, desde que no PCP do Produtor, os itens considerados são as caixas.

Basicamente, o PCP do Produtor consiste em arranjar, sem sobreposição, o máximo número de dimensões (l, w) ou (w, l) (faces inferiores das caixas), sobre um retângulo maior (L, W) (superfície do palete). Admite-se que as caixas estão disponíveis em grandes quantidades e devam ser arranjadas ortogonalmente, isto é, com seus lados paralelos aos lados do palete. Estes arranjos de caixas formam camadas horizontais iguais de altura h que são então empilhadas uma sobre as outras, dentro do palete de altura H e de orientação vertical (geralmente) fixada.

De acordo com Oliveira [4], essas considerações práticas fazem com que o problema original de gerar um arranjo ótimo de três dimensões (para o problema de carregar caixas idênticas sobre o palete) seja reduzido a um problema bidimensional de arranjar caixas (retângulos menores) dentro do palete (retângulo maior), como descrito anteriormente. Mais ainda, com estas classificações, o PCP do Produtor equivale a um problema de corte, onde maximizar o número de caixas a serem dispostas sobre o palete é um problema idêntico a maximizar o número de peças a serem cortadas de uma peça maior. Neste contexto, as peças cortadas devem possuir dimensões iguais e os cortes realizados podem ser de forma não-guilhotinada, isto é, cortes produzidos por equipamentos que permitem interromper o corte antes de chegar à aresta oposta àquela de onde iniciou-se o corte dentro do palete.

No restante deste capítulo são revisadas algumas formulações matemáticas adaptadas

para o PCP do Produtor, com base no trabalho Oliveira [4], e que podem ser encontradas de forma mais geral na literatura em Beasley [8], Tsai *et al.* [9], Hadjiconstantinou e Christofides [10] e Chen *et al.* [11]. Em particular, é mostrado que o modelo adaptado de Hadjiconstantinou e Christofides [10] nem sempre é válido através de um contra-exemplo apresentado em Letchford e Amaral [14].

2.2 Formulação de Beasley

O trabalho de Beasley [8] formula matematicamente, como um problema de otimização inteira 0-1, o problema de cortar um número de peças retangulares de um único retângulo maior de forma a maximizar o valor das peças cortadas. De acordo com Farago e Morabito [15], o PCP do Produtor pode ser modelado como um caso particular do modelo de Beasley também como um programa inteiro 0-1, da forma descrita a seguir.

Considere caixas iguais com faces inferiores (l, w), de comprimento l e largura w, a serem colocadas sobre o palete de dimensões (L, W), análogas às da caixa. Admite-se, sem perda de generalidade, que $w \leq l \leq \min\{L, W\}$ e que l, w, L e W são números inteiros, pois é possível tomar tais dimensões proporcionais a uma unidade no sistema cartesiano. Defina $(l, w) = (l_1, w_1)$ e $(w, l) = (l_2, w_2)$ onde (l_i, w_i), $i = 1, 2$, são as dimensões de cada caixa com orientação i e, P e Q, como os números mínimo e máximo, respectivamente, de caixas por camada a serem carregadas sobre o palete. Note que a orientação das peças é fixada com $i = 1$ indicando posição horizontal de cada caixa e $i = 2$ representando a posição vertical (uma caixa de comprimento p e largura q não é a mesma que uma caixa de comprimento q e largura p).

Adota-se, então, um sistema de coordenadas cartesianas em que a origem representa o canto inferior esquerdo do palete e os pontos cartesianos (p, q) representam as possíveis posições em que cada canto inferior esquerdo de uma caixa pode ser colocada. Dessa forma, devem existir restrições que limitem os valores de p e q para garantir que estas possíveis posições estejam dentro do palete (retângulo maior de dimensões L e W) e restrições sobre a não sobreposição de caixas, isto é, se determinado ponto (p, q) for escolhido como posição para colocação de uma caixa dentro do palete, então devem ser eliminados todos os pontos (r, s) para os quais há sobreposição de caixas (veja Figura 2.1).

Para o ponto (p, q) fornecer uma posição em que a caixa (l, w) ou (w, l) permaneça dentro do palete de dimensões (L, W), deve-se ter, pelo menos, $p \leq L - w$ e $q \leq W - w$. Estas duas restrições garantem que as caixas, mesmo que não possam ser colocadas na posição horizontal $(i = 1)$ nos bordos do palete, ainda podem ser colocadas na posição vertical, uma vez que, por hipótese, $w \leq l$. Desta idéia, defina:

$$X = \{p \mid 0 \leq p \leq L - w, \text{ inteiro}\} \quad \text{e} \quad Y = \{q \mid 0 \leq q \leq W - w, \text{ inteiro}\} \tag{2.1}$$

como os conjuntos dos possíveis valores para p e q nos eixos x e y, respectivamente.

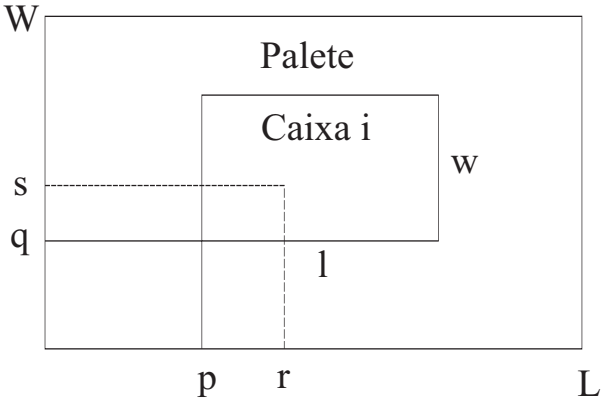


Figura 2.1: Posição (r, s) que deve ser eliminada em função da colocação de uma caixa na posição (p, q) com orientação i

Para as restrições de sobreposição, deseja-se obter um cálculo prévio de todos os pontos $(r, s) \in X \times Y$ que estão encobertos por uma determinada caixa que se encontra na posição $(p, q) \in X \times Y$ com orientação i. Assim, com o intuito de evitar tal sobreposição de caixas e garantir que suas posições possíveis estejam dentro dos limitantes do palete, defina a função a da seguinte forma:

$$a_{ipqrs} = \begin{cases} 1, & \text{se } 0 \leq p \leq r \leq p + l_i - 1 \leq L - 1 \text{ e } 0 \leq q \leq s \leq q + w_i - 1 \leq W - 1; \\ 0, & \text{caso contrário .} \end{cases} \tag{2.2}$$

Como pode ser observado pela Figura 2.1, quando $0 \leq p \leq r$ e $0 \leq q \leq s$, $a_{ipqrs} = 1$ indica que quando uma caixa de orientação i é colocada na posição (p, q), a posição (r, s) estará encoberta ao longo de todo comprimento e largura da caixa ($p + l_i$ e $q + w_i$, respectivamente) menos 1 unidade. A subtração de unidade deve-se ao fato de que uma nova caixa pode ser colocada na posição em que a caixa atual termina (nas posições dos bordos lateral e superior da caixa atual).

Também é definida uma variável binária que controla quando uma caixa é colocada ou não na posição (p, q), com orientação i:

$$x_{ipq} = \begin{cases} 1, & \text{se uma caixa de orientação } i \text{ é colocada na posição } (p, q) \text{ do palete;} \\ 0, & \text{caso contrário .} \end{cases}$$

Note que o elemento p deve ser um valor inteiro em X tal que se a caixa for colocada em algum ponto inteiro pertencente à reta $x = p$, então a caixa deve caber dentro do palete, independentemente de sua orientação i, isto é, $p \leq L - l_i$ para $i = 1, 2$. De maneira análoga, pode-se concluir que $q \leq W - w_i$ para $i = 1, 2$. Assim, a formulação de Beasley para o PCP do Produtor é dada da seguinte forma:

$$\text{maximizar} \quad \sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} x_{ipq} \tag{2.3}$$

sujeito a:

$$\sum_{i=1}^{2} \sum_{\{p\in X|p\leq L-l_i\}} \sum_{\{q\in Y|q\leq W-w_i\}} a_{ipqrs}x_{ipq} \leq 1, \quad \text{para todo } r \in X \text{ e } s \in Y \tag{2.4}$$

$$P \leq \sum_{i=1}^{2} \sum_{\{p\in X|p\leq L-l_i\}} \sum_{\{q\in Y|q\leq W-w_i\}} x_{ipq} \leq Q \tag{2.5}$$

$$x_{ipq} \in \{0,1\},\ p \in X \text{ tal que } p \leq L-l_i,\ q \in Y \text{ tal que } q \leq W-w_i. \tag{2.6}$$

A função objetivo (2.3) maximiza o número de caixas, enquanto as restrições (2.4) e (2.5) representam, respectivamente, a não sobreposição de caixas e a garantia de que a quantidade de caixas deve estar entre o número mínimo e máximo de caixas estipulados anteriormente. Para o caso onde o número máximo e mínimo de caixas por orientação é estipulado, P_i e Q_i, $i = 1, 2$, a restrição (2.5) é reescrita da seguinte forma:

$$P_i \leq \sum_{\{p\in X|p\leq L-l_i\}} \sum_{\{q\in Y|q\leq W-w_i\}} x_{ipq} \leq Q_i, \quad i = 1, 2.$$

O problema (2.3)-(2.6) envolve $O(|X||Y|)$ variáveis e restrições e torna-se difícil de ser resolvido otimamente na maioria dos casos práticos porque depende diretamente da cardinalidade dos conjuntos X e Y. Por isso, é importante investir em ferramentas que reduzam o tamanho de tais conjuntos para reduzir, conseqüentemente, o tempo computacional ao se utilizar a formulação de Beasley para resolver o PCP do Produtor (veja Capítulo 3).

2.3 Formulação de Tsai *et al.*

Tsai *et al.* [9] apresentam uma versão do problema de carregamento de paletes tridimensional com caixas de vários tamanhos. O método de carregamento permite que tais caixas sejam colocadas no mesmo palete com uma restrição imposta no número de caixas de cada tamanho a ser carregada. Tal problema é formulado como um problema de otimização inteira 0-1 em que, geralmente, é considerada uma proporção pré-determinada entre o número de caixas de cada tipo e o total de caixas carregadas no palete. Essa restrição de proporção não será considerada neste trabalho.

Desta forma, conforme realizado pelo trabalho de Oliveira [4], o modelo original de Tsai *et al.* [9] pode ser adaptado para modelar o PCP do Produtor como a seguir:

- S : conjunto de n caixas diferentes a serem consideradas;
- (l_i, w_i) : dimensões da caixa i no conjunto S (comprimento e largura) tais que $i = 1, \ldots, n$;
- (L, W) : dimensões do palete (comprimento e largura). Observe que $(l_i, w_i) = (l, w)$ ou ainda, $(l_i, w_i) = (w, l)$;
- (X°, Y°) : canto inferior esquerdo do palete pertencente ao plano cartesiano xy;

- (x_i, y_i) : variáveis de decisão (coordenadas x e y do canto inferior esquerdo da caixa i);
- P_k : variável de decisão binária associada à k-ésima caixa do conjunto S. A caixa k é colocada dentro do palete se $P_k = 1$ e é descartada do conjunto S se $P_k = 0$.

Assim, para $i = 1, 2, \ldots, n-1$, $j = i+1, i+2, \ldots, n$ e $k = 1, 2, \ldots, n$, a formulação matemática é dada por:

$$\text{maximizar} \quad \sum_{k=1}^{n} P_k \tag{2.7}$$

sujeito a:

$$x_i - x_j \geq l_j \qquad \text{para todo } i \text{ e } j \tag{2.8}$$

ou

$$x_j - x_i \geq l_i \qquad \text{para todo } i \text{ e } j \tag{2.9}$$

ou

$$y_i - y_j \geq w_j \qquad \text{para todo } i \text{ e } j \tag{2.10}$$

ou

$$y_j - y_i \geq w_i \qquad \text{para todo } i \text{ e } j \tag{2.11}$$

$$x_k \geq X^{\circ} P_k \qquad \text{para todo } k \tag{2.12}$$

$$y_k \geq Y^{\circ} P_k \qquad \text{para todo } k \tag{2.13}$$

$$x_k \leq (X^{\circ} + L) - l_k \qquad \text{para todo } k \tag{2.14}$$

$$y_k \leq (Y^{\circ} + W) - w_k \qquad \text{para todo } k \tag{2.15}$$

$$P_k \in \{0, 1\} \tag{2.16}$$

$$x_k, y_k \geq 0. \tag{2.17}$$

A função objetivo (2.7) maximiza o total de área coberta no palete pelas caixas, uma vez que maximiza o número de caixas. As restrições (2.8)-(2.11) garantem a não sobreposição de caixas dentro do palete. As restrições (2.12)-(2.15) delimitam a colocação de caixas dentro do palete (retângulo maior de dimensões $X^{\circ} + L$ e largura $Y^{\circ} + W$). E, por fim, as restrições (2.16)-(2.17) garantem a não-negatividade das variáveis contínuas x_i e y_i e a condição binária da variável P_k.

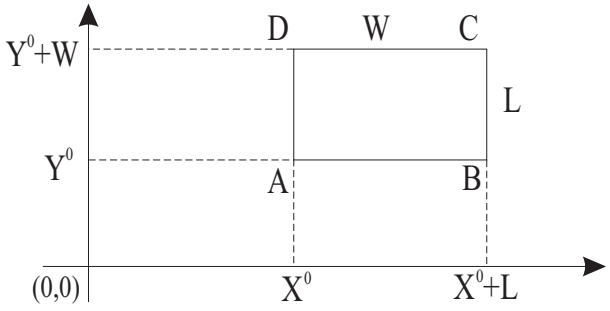


Figura 2.2: Palete ABCD

A posição (X°, Y°) é a coordenada do plano cartesiano tal que um subconjunto de caixas selecionadas de S são alocadas dentro do retângulo maior de comprimento $X^\circ + L$ e largura $Y^\circ + W$. Ao contrário da Formulação de Beasley, onde $(X^\circ, Y^\circ) = (0, 0)$, deve-se considerar o canto inferior esquerdo do palete, (X°, Y°), em qualquer posição suficientemente grande tal que as caixas que não são colocadas dentro do palete estejam em posições anteriores a ele com respeito ao plano xy. A Figura 2.2 representa a localização do palete ABCD no eixo xy.

As restrições (2.8)-(2.11) podem ser substituídas pelas seguintes restrições disjuntivas equivalentes:

$$x_j - x_i \leq -l_j + M(u_{ij}^1 + u_{ij}^2) \tag{2.18}$$

$$x_i - x_j \leq -l_i + M[1 - (u_{ij}^2 - u_{ij}^1)] \tag{2.19}$$

$$y_j - y_i \leq -w_j + M[1 - (u_{ij}^1 - u_{ij}^2)] \tag{2.20}$$

$$y_i - y_j \leq -w_i + M[2 - (u_{ij}^1 + u_{ij}^2)] \tag{2.21}$$

$$u_{ij}^1, u_{ij}^2 \in \{0, 1\} \tag{2.22}$$

onde M é um número suficientemente grande e u_{ij}^1 e u_{ij}^2 são variáveis binárias.

O modelo (2.7), (2.12)-(2.17) e (2.18)-(2.22) torna-se então um problema de otimização inteira $0-1$ que fornece o número de caixas necessárias $\sum_{k=1}^{n} P_k$ (e a localização de cada caixa (x_k, y_k) dentro do palete) para maximizar a área total coberta do palete. Nesta formulação, o número de variáveis envolvidas é da $O(3n + 2n(n-1)/2)$ e o número de restrições é da $O(3n + 2n(n-1)/2)$, ou seja, seu crescimento é dado pelo aumento do número n de caixas disponíveis.

2.4 Formulação de Hadjiconstantinou e Christofides

No trabalho de Hadjiconstantinou e Christofides [10] é apresentado um novo procedimento de busca em árvore para resolver problemas de corte bidimensional no qual um número de peças retangulares pequenas, cada uma com dado tamanho e valor, são cortadas de uma peça retangular maior. O objetivo é maximizar o valor de peças cortadas ou minimizar a sobra. Considera-se um número máximo de vezes que uma peça pode ser usada em um padrão de corte. O algoritmo limita o tamanho da busca em árvore utilizando um limitante derivado da relaxação lagrangiana em uma formulação de otimização inteira 0-1 deste problema. Otimização do subgradiente é usada para otimizar este limitante.

Considera-se uma peça retangular maior de dimensões (L, W), onde L representa o comprimento e W, a largura, a ser cortada em peças retangulares menores de dimensões $(l_1, w_1), (l_2, w_2), \ldots, (l_m, w_m)$ com valores (de rendimento lucrativo) correspondentes $v_1, v_2, \ldots, v_m$. O objetivo é construir um padrão de corte não-guilhotinado para (L, W) com valor total mais alto possível e utilizar não mais que Q_i réplicas de cada peça i, $i = 1, \ldots, m$. Também pode-se considerar um número mínimo de peças P_i para cada peça i a ser usada no padrão de corte [10]. No trabalho em questão, assume-se, sem perda de generalidade, que todas dimensões (l_i, w_i), $i = 1, \ldots, m$, são inteiras e que os cortes efetuados no retângulo maior são realizados em passos inteiros ao longo dos eixos x e y.

Para o caso particular do PCP do Produtor, existem apenas duas peças a serem cortadas do retângulo maior (L, W), ou seja, $m = 2$ cujas respectivas dimensões são dadas por $(l_1, w_1) = (l, w)$

e $(l_2, w_2) = (w, l)$ com $v_1 = v_2 = 1$, por exemplo. Note que os valores v_1 e v_2 podem possuir quaisquer valores desde que sejam necessariamente iguais, pois não haverá maior lucro em colocar uma caixa em particular na direção horizontal sobre o palete ou o mesmo com a posição vertical. Neste caso, será também considerado Q_1 e Q_2 suficientemente grandes. Defina:

$$x_{ijp} = \begin{cases} 1, & \text{se a } j\text{-ésima réplica da peça } i \text{ é cortada com seu canto inferior esquerdo na} \\ & \text{posição } p \text{ do eixo } x \text{ onde } j = 1, \ldots, Q_i \text{ e } p = 1, \ldots, L - l_i; \\ 0, & \text{caso contrário .} \end{cases}$$

$$y_{ijq} = \begin{cases} 1, & \text{se a } j\text{-ésima réplica da peça } i \text{ é cortada com seu canto inferior esquerdo na} \\ & \text{posição } q \text{ do eixo } y \text{ onde } j = 1, \ldots, Q_i \text{ e } q = 1, \ldots, W - w_i; \\ 0, & \text{caso contrário .} \end{cases}$$

$$z_{rs} = \begin{cases} 1, & \text{se o ponto } (r, s) \in (L, W), \quad r = 0, \ldots, l_i - 1, s = 0, \ldots, w_i - 1, \text{ não é cortado} \\ & \text{por nenhuma peça;} \\ 0, & \text{caso contrário .} \end{cases}$$

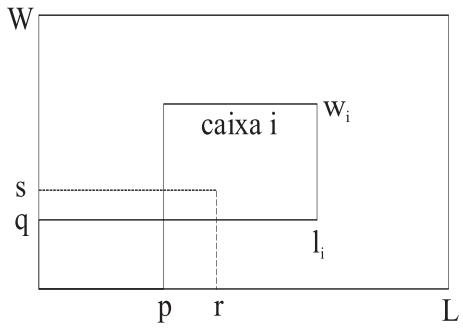


Figura 2.3: Localização de uma réplica da caixa i dentro de um palete

Note que uma réplica da peça i cortada de uma peça maior na posição (p, q) representa uma caixa de orientação i a ser colocada sobre o palete nesta mesma posição. A Figura 2.3 mostra tal réplica (caixa (l_i, w_i)), em um retângulo maior (palete) de dimensões (L, W), onde (r, s) é uma posição não permitida em função da colocação de uma caixa em (p, q). Assim, a formulação genérica dada por Hadjiconstantinou e Christofides [10], incluindo o caso particular do PCP do Produtor, é da forma:

$$\text{maximizar} \quad \sum_{i=1}^{m} v_i \sum_{j=1}^{Q_i} \sum_{p=0}^{L-l_i} x_{ijp} \tag{2.23}$$

sujeito a:

$$\sum_{s=q}^{q+w_i-1} \sum_{r=p}^{p+l_i-1} z_{rs} \leq (2 - x_{ijp} - y_{ijq}) l_i w_i \qquad i = 1, \ldots, m; j = 1, \ldots, Q_i;$$

$$p = 0, \ldots, L - l_i; q = 0, \ldots, W - w_i; \tag{2.24}$$

$$\sum_{p=0}^{L-l_i} x_{ijp} \leq 1 \qquad i = 1, \ldots, m; j = 1, \ldots, Q_i; \tag{2.25}$$

(2.26)

$$\sum_{p=0}^{L-l_i} x_{ijp} = \sum_{q=0}^{W-w_i} y_{ijp} \qquad i = 1, \ldots, m; j = 1, \ldots, Q_i; \tag{2.27}$$

$$\sum_{i=1}^{m} w_i \sum_{j=1}^{Q_i} \sum_{p=\max\{0, r-l_i+1\}}^{\min\{r, L-l_i\}} x_{ijp} + \sum_{s=0}^{W-1} z_{rs} = W \qquad r = 0, \ldots, L-1; \tag{2.28}$$

$$\sum_{i=1}^{m} l_i \sum_{j=1}^{Q_i} \sum_{q=\max\{0, s-w_i+1\}}^{\min\{s, W-w_i\}} y_{ijp} + \sum_{r=0}^{L-1} z_{rs} = L \qquad s = 0, \ldots, W-1; \tag{2.29}$$

$$x_{ijp} \in \{0,1\} \qquad i = 1, \ldots, m; j = 1, \ldots, Q_i;\ p = 0, \ldots, L - l_i; \tag{2.30}$$

$$y_{ijq} \in \{0,1\} \qquad i = 1, \ldots, m; j = 1, \ldots, Q_i; q = 0, \ldots, W - w_i; \tag{2.31}$$

$$z_{rs} \in \{0,1\} \qquad r = 0, \ldots, L-1; s = 0, \ldots, W-1. \tag{2.32}$$

Para o caso particular do PCP do Produtor, em que $m = 2$, a função objetivo (2.23) maximiza o número de réplicas de cada peça i, a serem cortadas do retângulo maior (L, W); equivale dizer, respectivamente, que a função objetivo maximiza o número de réplicas de cada caixa i, $i = 1, 2$, a serem dispostas no palete de dimensões (L, W). Seguindo com este raciocínio e modificando para o PCP do Produtor o problema de corte genérico (2.23)-(2.32), a restrição (2.24) garante que qualquer ponto em (L, W) será no máximo uma vez o local em que uma réplica de alguma das caixas i será colocada. O conjunto de restrições (2.25) e (2.27) expressam que qualquer caixa é colocada no máximo uma vez em (L, W). A restrição (2.28) impede que a soma das larguras das caixas a serem colocadas em (L, W) com seus cantos inferiores esquerdos no mesmo comprimento excedam a largura do palete, W. Similarmente, a restrição (2.29) garante que se a soma dos comprimentos de algumas caixas excedem L, então nem todas estas caixas podem ser colocadas em (L, W) com seus cantos inferiores esquerdos na mesma largura.

Observe que $x_{ijp} = 1$ e $y_{ijq} = 1$ significa que a j-ésima caixa do tipo i é colocada no palete com seu canto inferior esquerdo em (p, q). Se tal j-ésima caixa não é colocada no palete, então x_{ijp} e y_{ijq} devem possuir valor zero para todo p e para todo q, respectivamente [14]. Analogamente à formulação de Beasley, os índices p e q são considerados tais que $p \in X$ e $q \in Y$. Como citado anteriormente, a vantagem em diminuir os elementos de tais conjuntos através de relações como as de dominância encontradas no Capítulo 3 é reduzir de forma significativa o tempo computacional para obter uma solução ou aproximação para a solução do PCP do Produtor com esta modelagem. Tal vantagem acontece porque a formulação de Hadjiconstantinou e Christofides, (2.23)-(2.32), envolve N_c restrições e N_v variáveis inteiras, dadas por:

$$N_c = 2\sum_{i=1}^{m} Q_i + \sum_{i=1}^{m} Q_i |X||Y| + L + W$$

e

$$N_v = \sum_{i=1}^{m} Q_i (|X| + |Y|) + LW$$

que dependem dos conjuntos X e Y. O tamanho desta formulação também depende do número total de caixas dado pelo número máximo Q_i de réplicas de cada caixa. Como $m = 2$, Q_1 e

Q_2 representam o número máximo de caixas com orientação $i = 1, 2$, respectivamente, a serem carregadas sobre o palete.

De acordo com Letchford e Amaral [14], esta formulação nem sempre é válida, pois existem soluções inteiras que não representam padrões de corte factíveis. Isto é facilmente verificado pelo contra-exemplo a seguir. Sejam $m = 2$, $(L, W) = (3, 3)$, $(l_1, w_1, Q_1, v_1) = (2, 1, 2, 1)$ e $(l_2, w_2, Q_2, v_2) = (1, 2, 2, 1)$. Ao considerar x_{110}, x_{121}, x_{210}, x_{222}, y_{110}, y_{122}, y_{210}, y_{221} e z_{11} iguais a um e todas as outras variáveis iguais a zero, é obtida uma solução inteira factível para (2.23)-(2.32). No entanto, como é possível observar pela Figura 2.4, a primeira peça do tipo 1 e a primeira do tipo 2 (caixas 1 e 3) sobrepõem-se, assim como a segunda peça do tipo 1 e a segunda do tipo 2 (caixas 2 e 4). Isto mostra que este padrão de corte não é factível, tornando a formulação (2.23)-(2.32) inválida para este exemplo.

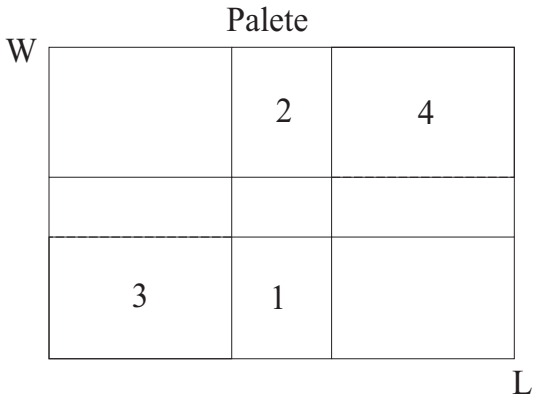


Figura 2.4: Caixas 1 e 2 sobrepostas pelas Caixas 3 e 4

Para validar este modelo, Letchford e Amaral [14] propõem a formulação abaixo, eliminando as variáveis z e substituindo as restrições (2.24), (2.28), (2.29) e (2.31) por restrições da forma $x_{ijp} + y_{ijq} + x_{klr} + y_{kls} \leq 3$ para todo i, p, q, k, r e s, tal que a peça do tipo i, colocada com seu canto inferior esquerdo em (p, q), sobrepõe a peça do tipo k com seu canto inferior esquerdo em (r, s). Além disso, para $j = 1, \ldots, Q_i$ e $l = 1, \ldots, Q_k$ tais restrições fornecem ou $i \neq k$ ou $j \neq l$. Apesar de válida, a formulação de Letchford e Amaral [14], envolve um número de variáveis e restrições computacionalmente intratáveis nos problemas práticos, não sendo viável sua utilização. Tal modelo é dado da seguinte forma:

$$\text{maximizar} \quad \sum_{i=1}^{m} v_i \sum_{j=1}^{Q_i} \sum_{p=0}^{L-l_i} x_{ipq} \tag{2.33}$$

sujeito a:

$$\sum_{p=0}^{L-l_i} x_{ijp} \leq 1 \qquad i = 1, \ldots, m;\ j = 1, \ldots, Q_i \tag{2.34}$$

$$\sum_{p=0}^{L-l_i} x_{ijp} = \sum_{q=0}^{W-w_i} y_{ijp} \qquad i = 1, \ldots, m;\ j = 1, \ldots, Q_i \tag{2.35}$$

$$
\begin{aligned}
x_{ijp} + y_{ijq} + x_{klr} + y_{kls} \leq 3 \quad & i = 1, \ldots, m \\
& p = 1, \ldots, L - l_i;\ q = 0, \ldots, W - w_i \\
& r = 0, \ldots, L - 1;\ s = 0, \ldots, W - 1 \\
& j = 1, \ldots, Q_i;\ l = 1, \ldots, Q_k \\
& k = 1, \ldots, m,\ i \neq k \ \text{ ou } \ j \neq l. \qquad (2.36)
\end{aligned}
$$

2.5 Formulação de Chen *et al.*

No trabalho de Chen *et. al* [11] é considerado o problema de carregamento de contêineres com caixas de tamanho não uniforme. Os autores apresentam um modelo de otimização inteira 0-1 que considera múltiplos contêineres e vários tamanhos para as caixas colocadas sobre o palete. De acordo com Oliveira [4], casos especiais do problema de carregamento de contêineres são discutidos baseados em um modelo geral. As soluções não tem garantia de serem carregamentos estáveis. Um caso especial do problema de carregamento de contêineres é selecionar um contêiner para um dado conjunto de caixas a fim de minimizar a perda de espaço. Baseado neste caso especial, Oliveira [4] apresenta a seguinte formulação adaptada para o PCP do Produtor.

Sejam os seguintes parâmetros e variáveis para o modelo:

- M : número suficientemente grande;
- m : número total de paletes disponíveis;
- N : número total de caixas a serem empacotadas;
- (l_i, w_i) : parâmetros que indicam o comprimento e largura da caixa i;
- (L, W) : parâmetros que indicam o comprimento e largura do palete;
- n_j : variável binária igual a 1 se o palete j é usado e igual a zero, caso contrário;
- (x_i, y_i) : variáveis contínuas (para localização) que indicam as coordenadas do canto inferior frontal esquerdo da caixa i;
- (lx_i, ly_i) : variáveis binárias que indicam se o comprimento da caixa i é paralelo ao eixo x ou y. Por exemplo, o valor de lx_i é igual a 1 se o comprimento da caixa i é paralelo ao eixo x; caso contrário, ele é igual a zero;
- (wx_i, wy_i) : variáveis binárias que indicam se a largura da caixa i é paralela ao eixo x ou y. Por exemplo, o valor de wx_i é igual a 1 se a largura da caixa i é paralela ao eixo x; caso contrário, ele é igual a zero;
- a_{ik}, b_{ik}, c_{ik}, d_{ik} : variáveis binárias definidas para indicar a localização das caixas em relação umas às outras. A variável a_{ik} é igual a 1 se a caixa i está no lado esquerdo da caixa k. Similarmente, as variáveis b_{ik}, c_{ik} e d_{ik} definem se a caixa i está à direita, em cima ou abaixo da caixa k, respectivamente. Essas variáveis são definidas para $i < k$.

A Formulação de Chen *et. al* é dada por:

$$\text{minimizar} \quad \sum_{j=1}^{m} LWn_j - \sum_{i=1}^{N} l_i w_i \tag{2.37}$$

sujeito a:

$$x_i + l_i lx_i + w_i wx_i \leq x_k + M(1 - a_{ik}) \qquad \text{para todo } i,\ k,\ i < k \tag{2.38}$$

$$x_k + l_k lx_k + w_k wx_k \leq x_i + M(1 - b_{ik}) \qquad \text{para todo } i,\ k,\ i < k \tag{2.39}$$

$$y_i + w_i wy_i + l_i ly_i \leq y_k + M(1 - c_{ik}) \qquad \text{para todo } i,\ k,\ i < k \tag{2.40}$$

$$y_k + w_k wy_k + l_k ly_k \leq y_i + M(1 - d_{ik}) \qquad \text{para todo } i,\ k,\ i < k \tag{2.41}$$

$$a_{ik} + b_{ik} + c_{ik} + d_{ik} \geq 1 \qquad \text{para todo } i,\ k,\ i < k \tag{2.42}$$

$$x_i + l_i lx_i + w_i wx_i \leq \sum_{j=1}^{m} Ln_j \qquad \text{para todo } i \tag{2.43}$$

$$y_i + w_i wy_i + l_i ly_i \leq \sum_{j=1}^{m} Wn_j \qquad \text{para todo } i \tag{2.44}$$

$$\sum_{j=1}^{m} n_j = 1 \tag{2.45}$$

$$lx_i, ly_i, wx_i, wy_i \in \{0, 1\} \tag{2.46}$$

$$a_{ik}, b_{ik}, c_{ik}, d_{ik} \in \{0, 1\} \tag{2.47}$$

$$x_i, y_i \geq 0. \tag{2.48}$$

A função objetivo (2.37) minimiza o espaço total não utilizado nos paletes. As restrições (2.38)-(2.41) garantem que não haja sobreposição de caixas, o que é verificado pela restrição (2.42). As restrições (2.43) e (2.44) garantem que apenas um palete é selecionado para um dado conjunto de caixas.

Adaptando as observações do trabalho de Chen *et. al* [11], as variáveis binárias, lx_i, ly_i, wx_i e wy_i são dependentes e satisfazem as seguintes relações:

$$lx_i + ly_i = 1$$
$$wx_i + wy_i = 1$$
$$lx_i + wx_i = 1$$
$$ly_i + wy_i = 1$$

Desde que a última relação pode ser escrita como combinação linear das demais, então pode-se reescrever $wx_i = 1 - wy_i$, $lx_i = wy_i$ e $ly_i = 1 - wy_i$. A Formulação de Chen *et al.* (2.37)-(2.48) é então reduzida ao seguinte modelo:

$$\text{minimizar} \quad \sum_{j=1}^{m} LWn_j - \sum_{i=1}^{N} l_i w_i \tag{2.49}$$

sujeito a:

$$x_i + l_i wy_i + w_i(1 - wy_i) \leq x_k + M(1 - a_{ik}) \quad \text{para todo } i,\ k,\ i < k \tag{2.50}$$

$$x_k + l_k wy_k + w_k(1 - wy_k) \leq x_i + M(1 - b_{ik}) \quad \text{para todo } i,\ k,\ i < k \tag{2.51}$$

$$y_i + w_i wy_i + l_i(1 - wy_i) \leq y_k + M(1 - c_{ik}) \quad \text{para todo } i,\ k,\ i < k \tag{2.52}$$

$$y_k + w_k wy_k + l_k(1 - wy_k) \leq y_i + M(1 - d_{ik}) \quad \text{para todo } i,\ k,\ i < k \tag{2.53}$$

$$a_{ik} + b_{ik} + c_{ik} + d_{ik} \geq 1 \quad \text{para todo } i,\ k,\ i < k \tag{2.54}$$

$$x_i + l_i wy_i + w_i(1 - wy_i) \leq \sum_{j=1}^{m} Ln_j \quad \text{para todo } i \tag{2.55}$$

$$y_i + w_i wy_i + l_i(1 - wy_i) \leq \sum_{j=1}^{m} Wn_j \quad \text{para todo } i \tag{2.56}$$

$$\sum_{j=1}^{m} n_j = 1 \tag{2.57}$$

$$wy_i \in \{0, 1\} \tag{2.58}$$

$$a_{ik}, b_{ik}, c_{ik}, d_{ik} \in \{0, 1\} \tag{2.59}$$

$$x_i, y_i \geq 0. \tag{2.60}$$

Para adaptar ao PCP do Produtor, Oliveira [4] apresenta modificações no modelo inicial (2.37)-(2.48), pois não é necessária a seleção de um palete dentre alguns paletes existentes. Deseja-se que um palete já esteja estabelecido, com suas dimensões previamente fornecidas, e deve-se maximizar o número de caixas nele dispostas ao invés de minimizar a sobra. Dessa forma, sejam:

- m : limitante superior para o número de caixas
- (X^o, Y^o) : canto inferior esquerdo do palete no plano cartesiano ao longo dos eixos x e y;
- P_i : variável de decisão binária onde $P_i = 1$ se a caixa i é colocada no palete e $P_i = 0$, caso contrário.

O novo problema de otimização inteira 0-1 pode ser formulado como a seguir:

$$\text{maximizar} \sum_{i=1}^{m} P_i \tag{2.61}$$

sujeito a:

$$x_i + l_i lx_i + w_i wx_i \leq x_k + M(1 - a_{ik}) \quad \text{para todo } i,\ k,\ i < k \tag{2.62}$$

$$x_k + l_k lx_k + w_k wx_k \leq x_i + M(1 - b_{ik}) \quad \text{para todo } i,\ k,\ i < k \tag{2.63}$$

$$y_i + w_i wy_i + l_i ly_i \leq y_k + M(1 - c_{ik}) \quad \text{para todo } i,\ k,\ i < k \tag{2.64}$$

$$y_k + w_k wy_k + l_k ly_k \leq y_i + M(1 - d_{ik}) \quad \text{para todo } i,\ k,\ i < k \tag{2.65}$$

$$a_{ik} + b_{ik} + c_{ik} + d_{ik} \geq 1 \quad \text{para todo } i,\ k,\ i < k \tag{2.66}$$

$$x_i \geq X^o P_i \quad \text{para todo } i \tag{2.67}$$

$$y_i \geq Y^o P_i \quad \text{para todo } i \tag{2.68}$$

$$x_i + l_i lx_i + w_i wx_i \leq (X^o + L) \quad \text{para todo } i \tag{2.69}$$

$$y_i + w_i wy_i + l_i ly_i \leq (Y^o + W) \qquad \text{para todo } i \tag{2.70}$$

$$lx_i, ly_i, wx_i, wy_i \in \{0,1\} \tag{2.71}$$

$$P_i, a_{ik}, b_{ik}, c_{ik}, d_{ik} \in \{0,1\} \tag{2.72}$$

$$x_i, y_i \geq 0. \tag{2.73}$$

A função objetivo (2.61) maximiza o número total de caixas dispostas sobre o palete. As restrições (2.62)-(2.65) garantem que não haja sobreposição de caixas, o que é verificado pela restrição (2.66). As restrições (2.67)-(2.70) garantem que todas as caixas colocadas dentro do palete estejam dentro de suas dimensões limitantes.

Novamente, é possível diminuir o número re variáveis envolvidas no modelo, desde elas são dependentes entre si. Assim, a formulação (2.61)-(2.73) pode ser reescrita por:

$$\text{maximizar} \sum_{i=1}^{m} P_i \tag{2.74}$$

sujeito a:

$$x_i + l_i wy_i + w_i(1 - wy_i) \leq x_k + M(1 - a_{ik}) \qquad \text{para todo } i,\ k,\ i < k \tag{2.75}$$

$$x_k + l_k wy_k + w_k(1 - wy_k) \leq x_i + M(1 - b_{ik}) \qquad \text{para todo } i,\ k,\ i < k \tag{2.76}$$

$$y_i + w_i wy_i + l_i(1 - wy_i) \leq y_k + M(1 - c_{ik}) \qquad \text{para todo } i,\ k,\ i < k \tag{2.77}$$

$$y_k + w_k wy_k + l_k(1 - wy_k) \leq y_i + M(1 - d_{ik}) \qquad \text{para todo } i,\ k,\ i < k \tag{2.78}$$

$$a_{ik} + b_{ik} + c_{ik} + d_{ik} \geq 1 \qquad \text{para todo } i,\ k,\ i < k \tag{2.79}$$

$$x_i \geq X^o P_i \qquad \text{para todo } i \tag{2.80}$$

$$y_i \geq Y^o P_i \qquad \text{para todo } i \tag{2.81}$$

$$x_i + l_i wy_i + w_i(1 - wy_i) \leq (X^o + L) \qquad \text{para todo } i \tag{2.82}$$

$$y_i + w_i wy_i + l_i(1 - wy_i) \leq (Y^o + W) \qquad \text{para todo } i \tag{2.83}$$

$$wy_i \in \{0,1\} \tag{2.84}$$

$$P_i, a_{ik}, b_{ik}, c_{ik}, d_{ik} \in \{0,1\} \tag{2.85}$$

$$x_i, y_i \geq 0. \tag{2.86}$$

Capítulo 3

Estudo Computacional do Modelo de Beasley

3.1 Introdução

Neste capítulo, a Formulação de Beasley (2.3)-(2.6) adaptada ao PCP do Produtor é explorada com mais detalhes, pois este modelo será utilizado por capítulos posteriores em que se estudam as relaxações lagrangiana e *surrogate* para esta formulação. A escolha deste modelo foi baseada na experiência e resultados adquiridos por Oliveira [4], e na prova dada por Letchford e Amaral [14] e Morabito e Farago [16] mostrando que os limitantes superiores obtidos com relaxação linear e relaxação lagrangiana e *surrogate* com esta formulação para o PCP do Produtor são bem apertados.

Como discutido previamente, investir em melhorias para os conjuntos X e Y deste modelo torna-se um interesse essencial para encontrar ferramentas que auxiliem na diminuição do tempo computacional para a maioria dos exemplos encontrados na literatura, salvos casos de pequeno porte onde os conjuntos melhorados mantém-se iguais. Tais melhorias são encontradas na literatura no sentido de que a redução na cardinalidade destes conjuntos implicam diretamente na diminuição do número de variáveis e de restrições deste modelo, respectivamente dados pelas aproximações $2|X||Y|$ e $|X||Y|$.

Inicialmente, serão apresentados neste capítulo os "Conjuntos Normais", desenvolvidos nos trabalhos de Herz [17] e Christofides e Whitlock [18], e os "Conjuntos *Raster Points*", encontrados pela primeira vez no trabalho de Scheithauer e Terno [19]. Ambas melhorias dos conjuntos X e Y também são apresentadas para o PCP do Produtor nos artigos de Beasley [8], Alvarez-Valdez *et al.* [20], Oliveira [4] e Oliveira e Morabito [21]. Em seguida, um exemplo trivial é desenvolvido detalhadamente para ilustrar a formulação de Beasley e a redução dos conjuntos X e Y. Por fim, outros exemplos são utilizados e alguns resultados computacionais são obtidos com a implementação do modelo de Beasley no sistema de resolução CPLEX acoplado à linguagem AMPL. A implementação para o problema trivial está anexada ao fim deste trabalho e as demais implementações são idênticas a esta, com exceção dos dados fornecidos a cada exemplo teste da Seção 3.4.1.

3.2 Melhorias nos Conjuntos X e Y

Herz [17] e Christofides e Whitlock [10], a partir da observação de que uma caixa colocada na posição (p, q) pode ser movida para a esquerda ou para baixo até que seu bordo esquerdo e seu bordo inferior sejam ambos adjacentes a outras caixas (ou adjacentes ao próprio palete (L, W)), restringiram os conjuntos X e Y aos *conjuntos normais*:

$$X = \{p \mid p = \sum_{i=1}^{2} l_i b_i, \quad 0 \leq p \leq L - w, \quad b_i \geq 0 \text{ e inteiro}, i = 1, 2\} \tag{3.1}$$

e

$$Y = \{q \mid q = \sum_{i=1}^{2} w_i b_i, \quad 0 \leq q \leq W - w, \quad b_i \geq 0 \text{ e inteiro}, i = 1, 2\} \tag{3.2}$$

Os elementos dos conjuntos normais são as combinações lineares entre os comprimentos l_1 e l_2 e larguras w_1 e w_2 respeitando os limitantes do palete, isto é, $0 \leq p \leq L-w$ e $0 \leq q \leq W-w$. Note que para fins de notação, a partir de agora serão denotados por X e Y os conjuntos normais. Posteriormente, utilizando considerações de dominância, Scheithauer e Terno [19] reduziram os conjuntos normais X e Y aos *conjuntos raster points*:

$$X' = \{(L - p)_X | p \in X\} \cup \{0\} \tag{3.3}$$

e

$$Y' = \{(W - q)_Y | q \in Y\} \cup \{0\} \tag{3.4}$$

onde $(L-p)_X = \max\{p|p \in X, p \leq L-p\}$ e $(W-q)_Y = \max\{q|q \in Y, q \leq W-q\}$. Observe que o conjunto *raster point* X' é composto dos elementos do *conjunto normal* X que são menores ou iguais a $(L-p)$. O elemento $(L-p)$ só pertence a X' quando também pertencer a X, o que nem sempre é válido. Analogamente, é possível compreender a formação dos elementos do conjunto Y'.

3.3 Aplicação da Teoria

3.3.1 Construção das Melhorias

Nesta seção, ilustra-se o funcionamento das melhorias anteriormente propostas com um exemplo particular do PCP do Produtor modelado de acordo com a Formulação de Beasley. Considere, portanto, $(L, W, l, w) = (11, 10, 4, 3)$. Primeiramente, apresenta-se os conjuntos normais, pois os conjuntos *raster points* são construídos a partir deles. Por (2.1), os conjuntos X e Y são dados por:

$$X = \{p \mid 0 \leq p \leq 8, \text{ inteiro}\} = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$$

e

$$Y = \{q \mid 0 \leq q \leq 7, \text{ inteiro}\} = \{0, 1, 2, 3, 4, 5, 6, 7\}$$

Construindo os Conjuntos Normais

A partir de (3.1), tem-se:

$$X = \{p \mid p = 4b_1 + 3b_2, \quad 0 \leq p \leq 8, \quad b_i \geq 0 \text{ e inteiro}, i = 1, 2\} \tag{3.5}$$

Logo:

$b_1 = 0$: $b_2 = 0 \Rightarrow p = 0 \leq 8 \Rightarrow 0 \in X$
$b_2 = 1 \Rightarrow p = 3 \leq 8 \Rightarrow 3 \in X$
$b_2 = 2 \Rightarrow p = 6 \leq 8 \Rightarrow 6 \in X$
$b_2 = 3 \Rightarrow p = 9 > 8 \Rightarrow$ A partir daqui, não existem mais elementos em X.

$b_1 = 1$: $b_2 = 0 \Rightarrow p = 4 \leq 8 \Rightarrow 4 \in X$
$b_2 = 1 \Rightarrow p = 7 \leq 8 \Rightarrow 7 \in X$
$b_2 = 2 \Rightarrow p = 10 > 8 \Rightarrow$ A partir daqui, não existem mais elementos em X.

$b_1 = 2$: $b_2 = 0 \Rightarrow p = 8 \leq 8 \Rightarrow 8 \in X$
$b_2 = 1 \Rightarrow p = 12 > 8 \Rightarrow$ A partir daqui, não existem mais elementos em X.

$b_1 = 3$: $b_2 = 0 \Rightarrow p = 12 > 8 \Rightarrow$ A partir daqui, não existem mais elementos em X.

Portanto, dos valores para p que são menores que $L - w = 8$, compõe-se o conjunto normal $X = \{0, 3, 4, 6, 7, 8\}$. Da mesma forma, se

$$Y = \{q \mid q = 3b_1 + 4b_2, \quad 0 \leq q \leq 7, \quad b_i \geq 0 \text{ e inteiro}, i = 1, 2\} \tag{3.6}$$

então:

$b_1 = 0$: $b_2 = 0 \Rightarrow q = 0 \leq 7 \Rightarrow 0 \in Y$
$b_2 = 1 \Rightarrow q = 4 \leq 7 \Rightarrow 4 \in Y$
$b_2 = 2 \Rightarrow q = 8 > 7 \Rightarrow$ A partir daqui, não existem mais elementos em Y.

$b_1 = 1$: $b_2 = 0 \Rightarrow q = 3 \leq 7 \Rightarrow 3 \in Y$
$b_2 = 1 \Rightarrow q = 7 \leq 7 \Rightarrow 7 \in Y$
$b_2 = 2 \Rightarrow q = 11 > 7 \Rightarrow$ A partir daqui, não existem mais elementos em Y.

$b_1 = 2$: $b_2 = 0 \Rightarrow q = 6 \leq 7 \Rightarrow 6 \in Y$
$b_2 = 1 \Rightarrow q = 10 > 7 \Rightarrow$ A partir daqui, não existem mais elementos em Y.

$b_1 = 3$: $b_2 = 0 \Rightarrow q = 9 > 7 \Rightarrow$ A partir daqui, não existem mais elementos em Y.

Daí, obtém-se o conjunto normal $Y = \{0, 3, 4, 6, 7\}$.

Construindo os Conjuntos *Raster Points*

Procedendo conforme (3.3) e (3.4), pode-se reduzir os conjuntos normais X e Y para os conjuntos *rasterpoints*:

$$X' = \{(11 - p)_X | p \in X\} \cup \{0\} \tag{3.7}$$

e

$$Y' = \{(10 - q)_Y | q \in Y\} \cup \{0\} \tag{3.8}$$

Para $p = 0$: $L - p = 11 - 0 = 11 \Rightarrow \max\{p | p \in X, p \leq 11\} = \max\{0, 3, 4, 6, 7, 8\} = 8 \Rightarrow 8 \in X'$
Para $p = 3$: $L - p = 11 - 3 = 8 \Rightarrow \max\{p | p \in X, p \leq 8\} = \max\{0, 3, 4, 6, 7, 8\} = 8 \Rightarrow 8 \in X'$
Para $p = 4$: $L - p = 11 - 4 = 7 \Rightarrow \max\{p | p \in X, p \leq 7\} = \max\{0, 3, 4, 6, 7\} = 7 \Rightarrow 7 \in X'$
Para $p = 6$: $L - p = 11 - 6 = 5 \Rightarrow \max\{p | p \in X, p \leq 5\} = \max\{0, 3, 4\} = 4 \Rightarrow 4 \in X'$
Para $p = 7$: $L - p = 11 - 7 = 4 \Rightarrow \max\{p | p \in X, p \leq 4\} = \max\{0, 3, 4\} = 4 \Rightarrow 4 \in X'$
Para $p = 8$: $L - p = 11 - 8 = 3 \Rightarrow \max\{p | p \in X, p \leq 3\} = \max\{0, 3\} = 3 \Rightarrow 3 \in X'$

Para $q = 0$: $W - q = 10 - 0 = 10 \Rightarrow \max\{q | q \in Y, q \leq 10\} = \max\{0, 3, 4, 6, 7\} = 7 \Rightarrow 7 \in Y'$
Para $q = 3$: $W - q = 10 - 3 = 7 \Rightarrow \max\{q | q \in Y, q \leq 7\} = \max\{0, 3, 4, 6, 7\} = 7 \Rightarrow 7 \in Y'$
Para $q = 4$: $W - q = 10 - 4 = 6 \Rightarrow \max\{q | q \in Y, q \leq 6\} = \max\{0, 3, 4, 6\} = 6 \Rightarrow 6 \in Y'$
Para $q = 6$: $W - q = 10 - 6 = 4 \Rightarrow \max\{q | q \in Y, q \leq 4\} = \max\{0, 3, 4\} = 4 \Rightarrow 4 \in Y'$
Para $q = 7$: $W - q = 10 - 7 = 3 \Rightarrow \max\{q | q \in Y, q \leq 3\} = \max\{0, 3\} = 3 \Rightarrow 3 \in Y'$

Logo, são construídos $X' = \{0, 3, 4, 7, 8\}$ e $Y' = \{0, 3, 4, 6, 7\}$. Note que por se tratar de um exemplo de pequeno porte, $Y' = Y$; mas, quando forem considerados exemplos cada vez maiores na Seção 3.4.1, maior será a redução de elementos dos conjuntos normais até os *raster points*. Essa diferença é explicitada através das tabelas da seção em questão.

3.3.2 Exemplo Numérico

Nesta seção apresenta-se um exemplo numérico simples retirado do trabalho de Oliveira para ilustrar a formulação de Beasley para o PCP do Produtor, bem como as melhorias propostas na Seção 3.2 para os conjuntos X e Y. Sejam um palete de dimensões $(L, W) = (5, 4)$ e caixas de dimensões iguais $(l, w) = (3, 2)$. Como cada caixa pode assumir as posições vertical ou horizontal, então é considerado respectivamente, $(l_1, w_1) = (3, 2)$ e $(l_2, w_2) = (2, 3)$.

As possíveis posições (p, q) para colocação de cada caixa dentro do palete pertencem à combinação de pontos entre os conjuntos:

$$X = \{p \mid 0 \leq p \leq L - w = 5 - 2 = 3, \text{ inteiro}\} = \{0, 1, 2, 3\} \tag{3.9}$$

e

$$Y = \{q \mid 0 \leq q \leq W - w = 4 - 2 = 2, \text{ inteiro}\} = \{0, 1, 2\} \tag{3.10}$$

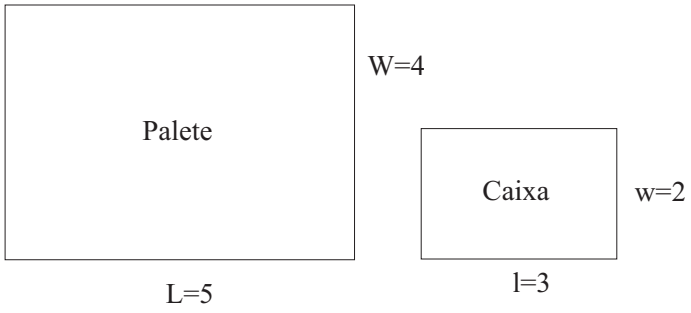


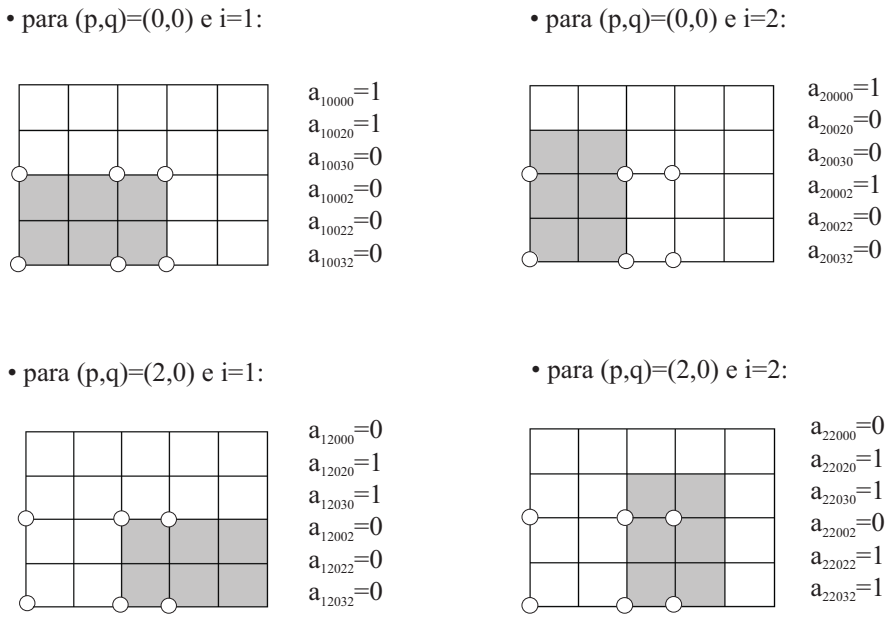
Figura 3.1: Palete de dimensões $(L, W) = (5, 4)$ e caixas de dimensões $(l, w) = (3, 2)$

Para obter os conjuntos normais deve-se fixar os valores de l_1 e de l_2 e, desde que p seja uma posição que esteja dentro das dimensões do palete, deve-se variar os inteiros b_1 e b_2 de todas as formas possíveis para construir o conjunto normal X. Procedendo de forma análoga, o conjunto normal Y também é obtido resultando, portanto, nos conjuntos normais:

$$X = \{0, 2, 3\} \quad \text{e} \quad Y = \{0, 2\} \tag{3.11}$$

Os conjuntos *raster points* X' e Y' são exatamente os conjuntos normais acima, ou seja, não há redução para este exemplo. Assim, para que a modelagem do PCP do Produtor segundo a formulação de Beasley seja exemplificada, basta encontrar os valores da função a, para todo $p \in X'$ e $q \in Y'$. Isto é, deve-se obter as informações de quais pontos (r, s) estarão encobertos, para todo $r \in X'$ e $s \in Y'$, quando há colocação de uma caixa de orientação i na posição (p, q).

De acordo com (5.1), a função a para este exemplo é descrita pela figura a seguir:



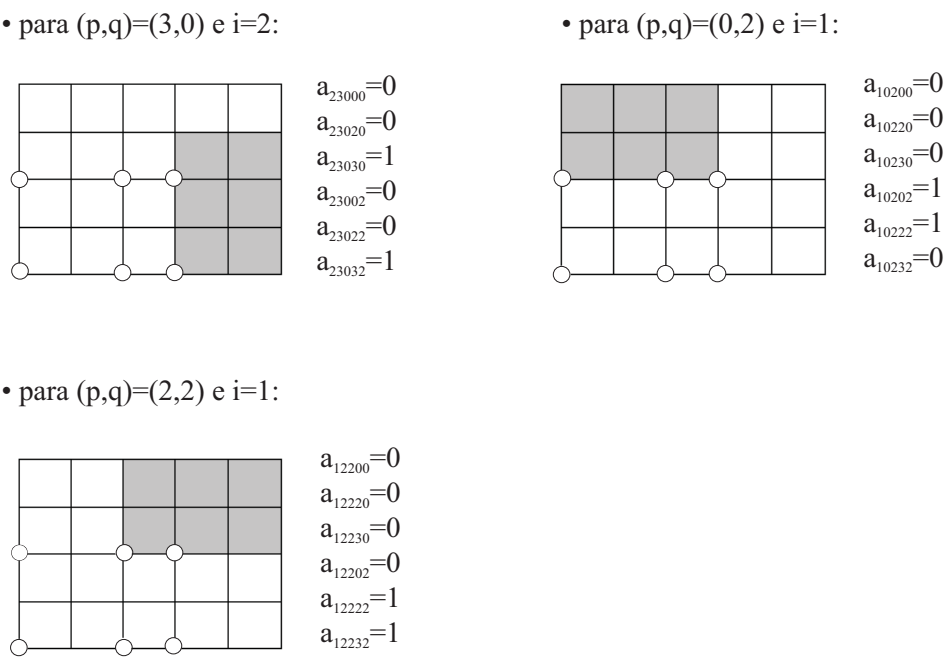


Figura 3.2: Detalhamento da função a

Pela formulação de Beasley, obtém-se de (2.4) as seguintes inequações:

- para $(r,s) = (0,0) : x_{100} + x_{200} \leq 1$;
- para $(r,s) = (0,2) : x_{102} + x_{200} \leq 1$;
- para $(r,s) = (2,0) : x_{100} + x_{120} + x_{220} \leq 1$;
- para $(r,s) = (2,2) : x_{102} + x_{122} + x_{220} \leq 1$;
- para $(r,s) = (3,0) : x_{120} + x_{220} + x_{230} \leq 1$;
- para $(r,s) = (3,2) : x_{122} + x_{220} + x_{230} \leq 1$;

Ao considerar P e Q dados por: $P = 0$ e $Q = \lfloor \frac{L*W}{l*w} \rfloor = \lfloor \frac{5*4}{3*2} \rfloor = \lfloor \frac{20}{6} \rfloor = 3$, então o modelo (2.3)-(2.6) aplicado ao exemplo é dado da seguinte forma:

$$\text{maximizar} \quad x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \tag{3.12}$$

sujeito a:

$$x_{100} + x_{200} \leq 1 \tag{3.13}$$
$$x_{102} + x_{200} \leq 1 \tag{3.14}$$
$$x_{100} + x_{120} + x_{220} \leq 1 \tag{3.15}$$
$$x_{102} + x_{122} + x_{220} \leq 1 \tag{3.16}$$
$$x_{120} + x_{220} + x_{230} \leq 1 \tag{3.17}$$
$$x_{122} + x_{220} + x_{230} \leq 1 \tag{3.18}$$
$$x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \leq 3 \tag{3.19}$$
$$x_{ipq} \in \{0,1\} \tag{3.20}$$

O número de restrições e variáveis envolvidas neste modelo é de aproximadamente $|X||Y|$ e $2|X||Y|$, respectivamente. As variáveis x_{ipq} implícitas (que "não aparecem") no modelo estão sendo multiplicadas por $a_{ipqrs} = 0$. Este problema foi resolvido pelo pacote AMPL/CPLEX obtendo solução ótima igual a três caixas: $x_{120} = x_{122} = x_{200} = 1$ e demais variáveis iguais a zero. A Figura 3.3 a seguir representa a solução obtida e a implementação do modelo encontra-se em anexo no Apêndice A.

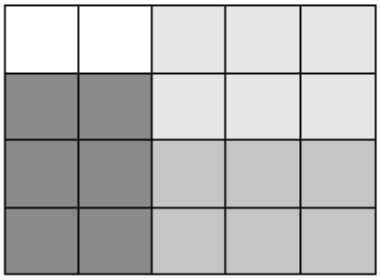


Figura 3.3: Padrão de carregamento ótimo para o exemplo

3.4 Testes Computacionais - AMPL/CPLEX

3.4.1 Exemplos Utilizados

O conjunto de exemplos utilizados (baseados em Oliveira [4]) foram retirados, em sua maioria, dos conjuntos Cover I e Cover II dispostos na literatura. Os exemplos foram divididos em três grupos e cada um deles, subdivididos em dois, totalizando seis grupos.

O Grupo 1 disposto nas Tabelas 3.1 e 3.2 possui 20 exemplos (L1-L20) retirados aleatoriamente do conjunto Cover I. O primeiro subconjunto, o Grupo 1A, possui os primeiros 10 exemplos tais que, considerando o modelo (2.3)-(2.6), não possuem *gap* entre o valor da solução ótima e o valor da solução da relaxação linear de (2.3)-(2.6). O Grupo 1B, possui os 10 exemplos restantes que, neste caso, possuem valores diferentes ao se comparar o valor da solução ótima e o valor da solução da relaxação linear.

De forma análoga, o segundo grupo foi dividido em dois grupos (Tabelas 3.3 e 3.4). O Grupo 2A possui nove exemplos retirados do conjunto Cover I, que foram utilizados em diversos artigos como, por exemplo, Farago e Morabito [15] e Scheithauer e Terno [19]; enquanto que o Grupo 2B é composto de 10 exemplos retirados de Letchford e Amaral [14]. Todos os exemplos encontrados nos Grupos 1 e 2 possuem solução ótima não-guilhotinada de primeira ordem em até 50 caixas e, portanto, são considerados fáceis de se resolver.

Os 32 exemplos da literatura dispostos no Grupo 3 possuem solução ótima contida no intervalo de 50 a 100 caixas. Eles foram retirados do conjunto Cover II e subdivididos em dois grupos (Tabelas 3.5 e 3.6). O Grupo 3A é composto de 16 exemplos que possuem solução ótima não-guilhotinada, enquanto que o Grupo 3B é composto dos outros 16 exemplos que, por não possuírem este tipo de solução, são considerados difíceis de se resolver.

Nas Tabelas 3.1 a 3.6, as colunas "$2|X||Y|$"e "$2|X'||Y'|$" representam os valores aproximados do número de variáveis 0-1 envolvidas na Formulação de Beasley para o PCP do Produtor

respectivamente para os conjuntos *normais* e *raster points*. É possível perceber em todas as tabelas a diminuição considerável deste número ao se utilizar os conjuntos *raster points* ao invés dos conjuntos *normais*.

GRUPO1A								
Exemplo	**(L,W)**	**(l,w)**	**|X|**	**|Y|**	**2(|X||Y|)**	**|X'|**	**|Y'|**	**2(|X'||Y'|)**
L1	(11,5)	(4,1)	11	5	110	11	5	110
L2	(16,11)	(4,3)	11	6	132	10	5	100
L3	(24,18)	(8,3)	15	9	270	10	6	120
L4	(28,27)	(8,3)	19	18	874	14	13	364
L5	(14,14)	(3,2)	12	12	288	12	12	288
L6	(20,16)	(5,2)	17	13	442	16	12	384
L7	(48,39)	(11,4)	30	21	1260	18	15	540
L8	(31,20)	(7,2)	27	16	864	25	14	700
L9	(11,8)	(2,1)	11	8	176	11	8	176
L10	(33,30)	(7,4)	24	21	1008	21	18	756
Média			**17,70**	**12,90**	**523,40**	**14,80**	**10,80**	**353,80**

Tabela 3.1: Grupo 1A de Exemplos

GRUPO1B								
Exemplo	**(L,W)**	**(l,w)**	**|X|**	**|Y|**	**2(|X||Y|)**	**|X'|**	**|Y'|**	**2(|X'||Y'|)**
L11	(36,28)	(8,3)	27	19	1026	22	14	616
L12	(36,36)	(10,3)	25	25	1250	18	18	648
L13	(60,33)	(9,5)	40	14	3840	28	11	616
L14	(37,25)	(7,3)	29	17	986	25	13	650
L15	(30,30)	(5,4)	21	21	882	18	18	648
L16	(50,36)	(8,5)	32	18	1152	22	14	616
L17	(39,39)	(11,3)	27	27	1458	19	19	722
L18	(33,30)	(7,3)	25	22	1100	21	18	756
L19	(37,26)	(5,4)	28	17	952	25	14	700
L20	(45,32)	(10,3)	34	21	1428	27	14	756
Média			**28,80**	**20,10**	**1135,40**	**22,50**	**15,30**	**672,80**

Tabela 3.2: Grupo 1B de Exemplos

GRUPO2A								
Exemplo	**(L,W)**	**(l,w)**	**|X|**	**|Y|**	**2(|X||Y|)**	**|X'|**	**|Y'|**	**2(|X'||Y'|)**
L21	(22,16)	(5,3)	16	10	320	14	8	224
L22	(86,82)	(15,11)	24	22	1056	17	16	544
L23	(57,44)	(12,5)	31	19	1178	19	14	532
L24	(42,39)	(9,4)	27	24	1296	20	15	600
L25	(124,81)	(21,10)	41	18	1476	24	13	624
L26	(40,25)	(7,3)	32	17	1088	28	13	728
L27	(52,33)	(9,4)	37	18	1332	28	13	728
L28	(56,52)	(12,5)	30	26	1560	21	17	714
L29	(87,47)	(7,6)	67	27	3618	57	17	1938
Média			**33,89**	**20,11**	**1436,00**	**25,33**	**14,00**	**736,89**

Tabela 3.3: Grupo 2A de Exemplos

GRUPO2B								
Exemplo	**(L,W)**	**(l,w)**	**|X|**	**|Y|**	**2(|X||Y|)**	**|X'|**	**|Y'|**	**2(|X'||Y'|)**
L30	(32,22)	(5,4)	23	13	598	20	11	440
L31	(40,26)	(7,4)	28	14	784	22	11	484
L32	(100,64)	(17,10)	32	14	896	22	11	484
L33	(32,27)	(5,4)	23	18	828	20	15	600
L34	(37,30)	(8,3)	28	21	1176	23	16	736
L35	(100,82)	(22,8)	32	23	1472	23	16	736
L36	(100,83)	(22,8)	32	23	1472	23	16	736
L37	(40,33)	(7,4)	28	21	1176	22	16	704
L38	(53,26)	(7,4)	41	14	1148	35	11	770
L39	(81,39)	(9,7)	51	13	1326	33	11	726
Média			**31,80**	**17,40**	**1087,60**	**24,30**	**13,40**	**641,60**

Tabela 3.4: Grupo 2B de Exemplos

GRUPO3A								
Exemplo	**(L,W)**	**(l,w)**	**|X|**	**|Y|**	**2(|X||Y|)**	**|X'|**	**|Y'|**	**2(|X'||Y'|)**
L40	(77,47)	(10,7)	44	17	1496	28	13	728
L41	(41,36)	(7,4)	29	24	1392	23	18	828
L42	(55,55)	(11,5)	31	31	1922	18	18	648
L43	(44,29)	(5,4)	35	20	1400	32	17	1088
L44	(85,46)	(12,5)	59	21	2478	42	15	1260
L45	(51,47)	(7,5)	35	31	2170	27	23	1702
L46	(70,43)	(7,6)	50	23	2300	40	16	1280
L47	(68,56)	(13,4)	47	35	3290	32	22	1408
L48	(52,52)	(9,4)	37	37	2738	28	28	1568
L49	(52,36)	(8,3)	43	27	2322	38	22	1672
L50	(78,40)	(8,5)	60	22	2640	50	16	1600
L51	(70,63)	(11,5)	46	39	3588	30	27	1620
L52	(120,64)	(19,5)	80	27	4320	48	19	1824
L53	(60,56)	(9,4)	45	41	3690	36	32	2304
L54	(100,82)	(17,5)	64	46	5888	42	27	2268
L55	(113,57)	(11,6)	83	27	4482	63	19	2394
Média			**49,25**	**29,25**	**2882,25**	**36,06**	**20,75**	**1483,25**

Tabela 3.5: Grupo 3A de Exemplos

GRUPO3B								
Exemplo	**(L,W)**	**(l,w)**	**|X|**	**|Y|**	**2(|X||Y|)**	**|X'|**	**|Y'|**	**2(|X'||Y'|)**
L56	(43,26)	(7,3)	35	18	1260	31	14	868
L57	(49,28)	(8,3)	40	19	1520	35	14	980
L58	(57,34)	(7,4)	45	22	1980	39	17	1326
L59	(63,44)	(8,5)	45	26	2340	35	19	1330
L60	(61,35)	(10,3)	50	24	2400	43	17	1462
L61	(67,37)	(11,3)	55	25	2750	47	18	1692
L62	(61,38)	(10,3)	50	27	2700	43	20	1720
L63	(61,38)	(6,5)	47	24	2256	41	19	1558
L64	(67,40)	(11,3)	55	28	3080	47	20	1880
L65	(74,49)	(11,4)	56	31	3472	44	22	1936
L66	(93,46)	(13,4)	72	25	3600	57	17	1938
L67	(106,59)	(13,5)	78	31	4836	58	19	2204
L68	(141,71)	(13,8)	92	26	4784	61	19	2318
L69	(74,46)	(7,5)	58	30	3480	50	23	2300
L70	(86,52)	(9,5)	66	32	4224	54	23	2484
L71	(108,65)	(10,7)	75	32	4800	54	23	2484
Média			**57,44**	**26,25**	**3092,63**	**46,19**	**19,00**	**1780,00**

Tabela 3.6: Grupo 3B de Exemplos

3.4.2 Resultados Computacionais

Todos os exemplos L1-L71 foram testados para ambos os conjuntos *normais* e *raster points* com a implementação do modelo de Beasley no sistema de resolução CPLEX (modo default - versão 10) acoplado à linguagem AMPL sob plataforma Windows XP Professional, versão 2002. Foi utilizado um computador AMD Athlon(tm) 64X2 Dual Core Processor 5600+ (3.00 GHz) com 2Gb de memória RAM.

Os resultados obtidos com os conjuntos *raster points* foram comparados àqueles obtidos pelos conjuntos *normais* e também aos resultados obtidos pelo trabalho de Oliveira [4] ao implementar o mesmo modelo (formulação de Beasley) em linguagem Pascal (compilador Delphi 5) com o pacote GAMS/CPLEX (versão 7) para a resolução dos mesmos conjuntos de exemplos em um computador Pentium III (650 MHz).

Nas Tabelas 3.7 e 3.8 são apresentadas a média das soluções ótimas de cada um dos seis grupos, bem como a média das soluções obtidas em um tempo máximo de três horas de execução, o número de nós e, finalmente, o tempo (em segundos) gasto para resolver cada grupo de exemplos (computado apenas para os exemplos resolvidos em menos de três horas).

		Grupo 1A		Grupo 1B		Grupo 2A	
		Média	Oliveira	Média	Oliveira	Média	Oliveira
Solução Ótima		31,90		44,20		48,56	
Conj. Normais	Sol. Obtida	31,90	-	44,20	-	48,56	-
	nº nós	0	-	0	-	0	-
	Tempo (s)	0,06	-	0,88	-	6,32	-
Conj. Raster Points	Sol. Obtida	31,90	31,90	44,20	44,20	48,56	48,56
	nº nós	0	0	0	0	0	0
	Tempo (s)	0,06	11,4	0,20	48,6	1,43	243

Tabela 3.7: Resultados - Grupos 1A, 1B e 2A

		Grupo 2B		Grupo 3A		Grupo 3B	
		Média	Oliveira	Média	Oliveira	Média	Oliveira
Solução Ótima		42,60		73,31		80,06	
Conj. Normais	Sol. Obtida	42,60	-	73,31	-	79,94	-
	nº nós	16,6	-	0,38	-	38,50	-
	Tempo (s)	8,11	-	79,60	-	200,04	-
Conj. Raster Points	Sol. Obtida	42,60	42,60	73,31	73,31	80,00	79,94
	nº nós	16	25,7	0	0	3,1	1,1
	Tempo (s)	2,91	292,2	3,03	553,8	23,9	1053

Tabela 3.8: Resultados - Grupos 2B, 3A e 3B

Ao observar as tabelas acima, em todos os grupos de exemplos testados percebe-se que a utilização dos conjuntos *raster points* faz com que o número menor de variáveis envolvidas reflita diretamente em elevado ganho de tempo computacional. Devido a diferença de ambiente computacional (máquina e versão CPLEX), resultados computacionais obtidos pelo presente

trabalho foram muito superiores em termos de tempo que àqueles obtidos no trabalho de Oliveira [4].

Tomando o caso particular do Grupo 3B, observe que enquanto Oliveira gasta 1053 segundos para encontrar solução ótima para o Grupo 3B, o pacote AMPL/CLEX gasta apenas 23,9 segundos para o mesmo feito. Mais ainda, note que a qualidade da solução encontrada pelo trabalho de Oliveira é igual apenas àquele encontrado pelos conjuntos *normais* deste trabalho e inferior àquele encontrado ao se utilizar os mesmos conjuntos *raster points*. Não foi possível encontrar solução ótima no prazo de 3 horas apenas para o exemplo L70, enquanto que Oliveira não conseguiu encontrar solução ótima neste mesmo tempo para os exemplos L69 e L70.

A partir dos resultados obtidos para os problemas do Grupo 3B, que possuem maior número de caixas como solução ótima (entre 50 e 100 caixas), é possível observar a dificuldade em se resolver de maneira exata o PCP do Produtor com pacotes computacionais à medida que se aumentam tais números de caixas. Na literatura, pode-se encontrar métodos exatos e heurísticos para sua resolução. De acordo com Ribeiro e Lorena [22], os métodos exatos baseiam-se em estruturas de busca em árvore e podem ser encontrados em [23], [24], [20] e [21]. Heurísticas podem ser construtivas, dividindo o palete em blocos [25], através de métodos recursivos [26] e com técnicas baseadas em estruturas conhecidas como $G4$ [19] e L [27, 28]. Alguns outros trabalhos também podem ser aplicados à meta-heurísticas, tais como o *método tabu* [29, 30], e a algoritmos genéticos [31, 32].

Oliveira e Morabito [21] desenvolveram um método *branch and bound* para o PCP do Produtor em que, a cada nó, é resolvida uma heurística lagrangiana. Tal heurística foi escolhida, pois se mostrou superior às heurísticas *surrogate* e lagrangiana/*surrogate* comparadas em testes computacionais. Entretanto, como será discutido no decorrer deste trabalho, a heurística *surrogate* não possui garantias de convergência quando seus multiplicadores são atualizados pelo método de otimização do subgradiente. No presente trabalho, será apresentada o algoritmo de Sarin *et al.* aplicado ao PCP do Produtor para uma nova atualização de tais multiplicadores, pois é provado na literatura que as heurísticas *surrogate* fornecem limitantes melhores, ou pelo menos iguais, às heurísticas lagrangianas.

Capítulo 4

Os Problemas Lagrangiano e *Surrogate*

4.1 Introdução

Problemas de otimização inteira são aqueles nos quais maximizam ou minimizam uma função de muitas variáveis que está sujeita a: (i) Restrições de desigualdade e/ou igualdade e (ii) Restrições de integralidade sobre algumas ou todas as variáveis utilizadas [33]. Em alguns trabalhos, quando nem todas as variáveis possuem restrições de integralidade, tais problemas ainda são denotados por *problemas de otimização inteira mista*. Em geral, problemas de otimização inteira possuem maior dificuldade em sua resolução que problemas de programação linear e podem ser formulados da seguinte forma:

$$\begin{array}{lll} (P) & \text{maximizar} & c^T x \\ & \text{sujeito a:} & Ax \leq b,\ x \in S \end{array}$$

onde $S = \{x \geq 0 : Gx \leq h,\ x$ satisfaz algumas ou todas restrições discretas$\}$, A uma matriz $m \times n$, G, $k \times n$, b e c vetores de dimensões apropriadas e h, vetor $k \times 1$.

Para estes problemas de otimização inteira serem resolvidos, sub-problemas associados possuindo maior facilidade de resolução, chamados relaxações, podem ser utilizados para fornecerem limitantes (chamados duais) para a solução ótima de P. Nemhauser e Wolsey [33], definem uma relaxação de P como um problema de otimização que satisfaz as seguintes propriedades: (i) O conjunto de soluções factíveis de P é um subconjunto das soluções factíveis da relaxação e; (ii) O valor da função objetivo do problema P não é melhor que o correspondente valor da relaxação. Ou seja, relaxar o problema P é "expandir" sua região factível e obter um sub-problema (relaxação) a partir de P que seja mais fácil de ser resolvido. O fato da relaxação possuir uma região factível maior implica diretamente no fato dela possuir solução ótima melhor ou igual à de P. A solução ótima obtida para uma relaxação é chamada limitante dual para P e para o caso de maximização, foco deste presente trabalho, também pode ser chamada de limitante superior para P.

Dentre as relaxações mais utilizadas, destacam-se as de programação linear (em que as condições de integralidade das variáveis de decisão são relaxadas), lagrangiana, *surrogate* e combinada lagrangiana-*surrogate*. Em 1970, Held e Karp [34] tornaram-se os precursores para o que Geoffrion [35] chamou posteriormente de relaxação lagrangiana ao utilizarem um método

lagrangiano para resolver o problema do caixeiro viajante. A relaxação *surrogate*, apesar de ter sido introduzida anteriormente por Glover [36], em 1965, possui poucas variações de aplicações, se comparadas às desenvolvidas para a relaxação lagrangiana. A dificuldade em se resolver os problemas resultantes da relaxação *surrogate*, que em muitos casos se reduzem a problemas da mochila, pode ser uma das causas.

O problema do caixeiro viajante, o problema de atribuição generalizado, o problema de recobrimento e particionamento de conjuntos e o problema de roteamento são alguns exemplos de aplicações que utilizam a relaxação lagrangiana e que podem ser encontrados em [37], [38], [39], [40], [41] e [42]. Dentre as aplicações que utilizam a relaxação *surrogate* existem poucos trabalhos se comparadas às da lagrangiana. Pode-se citar a utilização *surrogate* e lagrangiana na decomposição de Benders encontrado em [43] e [44] e os problemas da mochila, de recobrimento de conjuntos e de atribuição generalizado, ambos com abordagem *surrogate* e contidos nos trabalhos [45], [46] e [47], respectivamente. Para mais referências sobre relaxações, veja [48].

Neste capítulo, a teoria dual lagrangiana e *surrogate* é apresentada de forma a introduzir premissas que permitam chegar à compreensão de métodos heurísticos, dispostos no Capítulo 5, a serem aplicados posteriormente à resolução do PCP do Produtor, no Capítulo 6. Algumas relações entre os duais lagrangiano e *surrogate*, incluindo condições para a existência de *gap* entre eles, são verificadas para facilitar o manuseio com o problema *surrogate* que, em geral, é muito mais difícil de ser resolvido que o problema lagrangiano. Conceitos primários e importantes demonstrações são encontrados no Apêndice B para facilitar a introdução deste capítulo.

4.2 Dualidade Lagrangiana

A *relaxação lagrangiana* é uma técnica que pode ser utilizada para obtenção de limitantes para problemas de programação inteira quando as restrições têm uma estrutura que se pode tirar partido. Se no problema primal P existir um conjunto de restrições que tornam o problema difícil de ser resolvido, a estratégia lagrangiana baseia-se em relaxar tais restrições "complicadas", integrado-as à função objetivo, e fazendo das restrições restantes uma estrutura mais simples que a do conjunto inicial. Considere:

$$\begin{array}{lll} (P) & \text{maximizar} & c^T x \\ & \text{sujeito a:} & Ax \leq b,\ x \in S \end{array}$$

o problema de otimização inteira descrito na Seção 4.1. Considere ainda que $Ax \leq b$ seja um conjunto de restrições complicadas, no sentido de que ao retirar tais restrições torna o problema mais fácil de ser resolvido.

Relaxar estas restrições de forma lagrangiana é não impor explicitamente que $Ax \leq b$ seja satisfeito na matriz de restrições, mas estabelecer um tipo de penalização na função objetivo, caso $Ax > b$. Define-se o *problema lagrangiano* (*relaxação lagrangiana*) da seguinte forma:

$$\begin{array}{lll} (P_\lambda) & \text{maximizar} & c^T x + \lambda^T (b - Ax) \\ & \text{sujeito a:} & x \in S \end{array}$$

onde λ é um vetor composto de elementos chamados *multiplicadores de Lagrange* que toma valores não-negativos, $\lambda \geq 0$. Caso $(b - Ax)$ seja negativo, isto é, $Ax > b$, então $\lambda(b - Ax) < 0$ penaliza a função objetivo, pois é uma parcela negativa em um problema cujo objetivo é a maximização. Para problemas de minimização tem-se o caso análogo para λ não positivo; para ambos problemas de minimização e maximização, caso existam restrições a serem relaxadas do tipo $Ax = b$, o sinal de λ é irrestrito.

Para exclusão de casos específicos, considere P factível e S limitado. Considere, ainda, as seguintes notações:

- $v(.)$: valor de uma solução ótima para o problema (.);
- $[T]$: politopo de um conjunto de restrições qualquer T;
- $\bar{T}$: conjunto de restrições qualquer, T, com as restrições de integralidade relaxadas. Por exemplo, $P(\bar{S})$ = relaxação linear de P. Note que $P(.)$ é o problema P previamente definido e tal que x pertence ao conjunto (.);
- $\Omega(.)$: conjunto de soluções ótimas para o problema (.).

Observe que agora é possível escrever $v(P) = \sup_{x \in S}\{c^T x : Ax \leq b\}$ e $v(P_\lambda) = \sup_{x \in S}\{c^T x + \lambda^T(b - Ax)\}$.

Implicitamente, ao relaxar $Ax \leq b$ no problema primal, a região de factibilidade de P estará sendo relaxada, uma vez que tais restrições são retiradas do conjunto de restrições de P (menos restrições implica em maior região de factibilidade). Assim, a função objetivo do problema lagrangiano fornece um valor melhor ou igual à solução ótima desejada, isto é, para o caso de maximização, P_λ fornece um limitante superior $v(P_\lambda)$ para o problema primal P (Teorema 4.1). Isto também pode ser concluído pela observação de que $v(P_\lambda) = c^T x + \lambda(b - Ax)$ enquanto que $v(P)$ é apenas $c^T x$ para algum x factível a P. O problema lagrangiano é um problema em x para um determinado λ fixo e a idéia é determinar o melhor valor para λ que forneça o limitante mais próximo possível para P, isto é, o menor limitante superior possível. Observe, ainda, que com as considerações anteriores, segundo Nemhauser e Wolsey [33], o problema lagrangiano é, de fato, uma relaxação para P (Seção 4.1).

Teorema 4.1. *(Dualidade Lagrangiana Fraca) Para qualquer escolha do vetor de multiplicadores de Lagrange, $\lambda \geq 0$, temos que $v(P) \leq v(P_\lambda)$.*

Demonstração: Qualquer ponto válido para P é válido para P_λ porque obedece ao conjunto de restrições de ambos problemas. Como $\lambda^T(b - Ax) \geq 0$ para todo x factível a P, então $c^T x \leq c^T x + \lambda^T(b - Ax)$, isto é, $v(P) \leq v(P_\lambda)$ para o caso particular do ponto x ótimo para P. ■

Geralmente, fixado um valor para λ, torna-se fácil resolver P_λ; por exemplo, por algum método enumerativo (o PCP do Produtor é resolvido por inspeção; veja Seção 6.2.2 e Oliveira [4]). O interesse principal é obter um valor para λ específico. Deseja-se encontrar um λ ótimo tal que a solução da relaxação lagrangiana, que é um limitante superior para P, seja o mais próximo possível (ou igual) do valor da solução ótima de P. Para isto, basta minimizar os limitantes superiores obtidos por P_λ, uma vez que, pelo Teorema 4.1, nunca será menor que a

solução ótima $v(P)$. Assim, o foco principal da relaxação lagrangiana está em resolver o seguinte *problema dual lagrangiano*:

$$(D_L) \qquad \text{minimizar}_{\ \lambda \geq 0} \ \{v(P_\lambda)\}$$

Resolver o dual lagrangiano é determinar o valor de λ para o qual a função lagrangiana $c^T x + \lambda^T(b - Ax)$ toma o valor mínimo. Assim, deve existir pelo menos um valor λ^* para o qual o problema lagrangiano toma o menor valor possível e que deve ser igual ao valor ótimo do problema dual lagrangiano, isto é, $v(D_L) = v(P_{\lambda^*}) = \inf_{\lambda \geq 0}\{v(P_\lambda)\}$. As seguintes observações ainda podem ser consideradas: D_L fornece o valor de λ^* ótimo, bem como a melhor solução lagrangiana $v(P_{\lambda^*})$; λ^* pode assumir valores diferentes para o(s) mesmo(s) valor(es) ótimo(s), $v(D_L)$, que é o limitante superior mais próximo (ou igual) a $v(P)$.

Corolário 4.1. *Para qualquer escolha do vetor de multiplicadores de Lagrange, $\lambda \geq 0$, temos que $v(P) \leq v(D_L)$.*

Demonstração: Pela estrutura de D_L e pelo Teorema 4.1, $v(P) \leq v(P_\lambda)$ acontece para todo λ factível a D_L, em particular, para $\inf\limits_{\lambda \geq 0} v(P_\lambda)$. O que implica diretamente em $v(D_L) \stackrel{\triangle}{=} \inf\limits_{\lambda \geq 0} v(P_\lambda) \geq v(P)$. ■

Como pode ser visto no Apêndice B.2, em otimização linear é correto afirmar que a solução ótima dual é sempre igual à solução ótima de seu próprio problema primal, $v(D) = v(PL)$. Esta afirmação não é sempre válida para problemas de otimização inteira; em particular, aplicar a relaxação lagrangiana ao problema primal implicará na função objetivo de D_L, que possui o incremento $\lambda^T(b - Ax)$, não necessariamente igual à função objetivo de P. Torna-se então natural o questionamento da existência de condições para que a solução ótima do dual lagrangiano, $v(D_L) = v(P_{\lambda^*})$, seja igual à solução ótima do problema primal, $v(P)$, em otimização inteira, isto é, quando $v(P) = v(P_{\lambda^*})$. Esta idéia é formalizada com o resultado a seguir.

Teorema 4.2. *(Condições de Otimalidade Lagrangiana) Se, para um dado vetor $\lambda^* \geq 0$, uma solução x^* satisfizer as seguintes condições:*

(i) *x^* é ótimo para P_{λ^*};*

(ii) *x^* é factível para P: $Ax^* \leq b$;*

(iii) *Folgas complementares satisfeitas: $(\lambda^*)^T(b - Ax^*) = 0$*

então x^* também é uma solução ótima para o problema primal P. Se x^* satisfaz (i) e (ii) mas não satisfaz (iii), então x^* é chamada de solução ξ-ótima de P com gap de dualidade $\xi = v(P_{\lambda^*}) - v(P)$.

Demonstração: Para provar este resultado, será mostrado que $v(P) = v(P_{\lambda^*})$, pois o maior valor que $v(P)$ poderá assumir será um valor igual a $v(P_{\lambda^*})$. Por hipótese, tem-se $\lambda^* \geq 0$ e x^* satisfazendo (i), (ii) e (iii).

Por (i) e pelo Teorema 4.1, x^* é solução ótima para P_{λ^*} e

$$v(P) \leq v(P_{\lambda^*}) = c^T x^* + (\lambda^*)^T(b - Ax^*) \tag{4.1}$$

Por (iii),

$$c^T x^* + \underbrace{(\lambda^*)^T(b - Ax^*)}_{=0} = c^T x^* \overset{(4.1)}{\Longrightarrow} v(P) \leq c^T x^* \tag{4.2}$$

Por (ii), x^* factível a P implica em

$$c^T x^* \leq v(P) \tag{4.3}$$

Por (4.2) e (4.3), obtém-se:

$$c^T x^* = v(P) \tag{4.4}$$

isto é, x^* é solução ótima para P implicando em $v(P_{\lambda^*}) = v(P)$. ■

Isto quer dizer que se, para determinado λ, a solução ótima da relaxação lagrangiana (que é a solução ótima do dual lagrangiano) for factível a P e satisfizer as folgas complementares, então podemos afirmar se tratar de uma solução ótima também para o problema primal, ou seja, $v(P) = v(D_L)$. Como nem sempre tais condições são satisfeitas, o interesse recai sobre a obtenção de λ que forneça para a relaxação lagrangiana o limitante superior mais próximo possível do valor ótimo do problema primal P. O exemplo abaixo irá ilustrar o último resultado.

Exemplo 4.2.1. *Considere a classe dos problemas de otimização inteira binários apresentados em Parker e Rardin [49]:*

$$\begin{array}{lll} (Q) & \textit{maximizar} & c^T x \\ & \textit{sujeito a:} & Ax \leq b,\ x \in B = \{0,1\} \end{array}$$

onde todas as variáveis são inteiras e restritas aos valores 0 ou 1 (Para mais classes de problemas em otimização inteira, veja Wolsey [50]). Dualizando todas as restrições lineares, a relaxação lagrangiana para $\lambda \geq 0$ é descrita por:

$$\begin{array}{lll} (Q_\lambda) & \textit{maximizar} & c^T x + \lambda^T(b - Ax) \\ & \textit{sujeito a:} & x \in B \end{array}$$

Como a função objetivo de (Q_λ) pode ser reescrita como $\sum_j (c_j - \lambda^T a_j)^T x_j + \lambda^T b$, onde a_j é a j-ésima coluna da matriz A, então (Q_λ) pode ser resolvido por inspeção da seguinte forma:

$$v(Q_\lambda) = \lambda^T b + \sum_j \max\{0, c_j - \lambda^T a_j\}$$

Para todo j, se 0 for o máximo em $\max\{0, c_j - \lambda^T a_j\}$, então implicitamente estará sendo tomado $x_j = 0$. Caso contrário, se $c_j - \lambda^T a_j$ for o máximo, então estará sendo tomado, implicitamente, $x_j = 1$. Considere o caso particular:

$$\begin{array}{lll} (Q) & \textit{maximizar} & f(x) = 3x_1 + 4x_2 \\ & \textit{sujeito a:} & \\ & & 6x_1 + 2x_2 \leq 3 \\ & & 5x_1 + 6x_2 \leq 6 \\ & & x_1, x_2 \in B \end{array}$$

com relaxação lagrangiana associada:

$$\begin{array}{lll}(Q_\lambda) & \textit{maximizar} & 3x_1 + 4x_2 + \lambda_1(3 - 6x_1 - 2x_2) + \lambda_2(6 - 5x_1 - 6x_2) \\ & \textit{sujeito a:} & x_1, x_2 \in B\end{array}$$

podendo ser reescrita por:

$$\begin{array}{lll}(Q_\lambda) & \textit{maximizar} & (3 - 6\lambda_1 - 5\lambda_2)x_1 + (4 - 2\lambda_1 - 6\lambda_2)x_2 + 3\lambda_1 + 6\lambda_2 \\ & \textit{sujeito a:} & x_1, x_2 \in B\end{array}$$

ou ainda por:

$$\begin{array}{lll}(Q_\lambda) & \textit{maximizar} & (\lambda_1, \lambda_2)(3,6)^T + [3 - (\lambda_1, \lambda_2)(6,5)^T]x_1 + [4 - (\lambda_1, \lambda_2)(2,6)^T]x_2 \\ & \textit{sujeito a:} & x_1, x_2 \in B\end{array}$$

Geralmente, atribuir valores aleatórios para λ em Q_λ fornecem apenas limitantes superiores para Q, $v(Q_\lambda)$. Por exemplo, tomando $\lambda = (1,1)$, obtém-se:

$$\begin{aligned} v(Q_{(1,1)}) &= (1,1)(3,6)^T + \max\{0, 3 - (1,1)(6,5)^T\} + \max\{0, 4 - (1,1)(2,6)^T\} \\ &= 9 + \max\{0, -8\} + \max\{0, -4\} \\ &= 9 + 0 + 0 = 9, \end{aligned}$$

onde a solução ótima para a relaxação lagrangiana será: $x_1^* = x_2^* = 0$.

O valor $v(Q_{(1,1)}) = 9$ não é igual a $v(Q)$, pois, apesar dos itens (i) e (ii) do Teorema 4.2 estarem sendo satisfeitos (devido à factibilidade de x^* em Q), as folgas complementares de (iii) não são satisfeitas:

$$\lambda^T(b - Ax^*) = \begin{pmatrix} 1 & 1 \end{pmatrix} (\begin{pmatrix} 3 \\ 6 \end{pmatrix} - \begin{pmatrix} 6 & 2 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}) = 9 \neq 0$$

Daí, o que se pode dizer é que $x^* = (0,0)$ fornece um limitante superior para Q, $v(Q_{(1,1)}) = 9$.

Entretanto, ao tomar $\lambda_1 = 0$ e $\lambda_2 \in [\frac{3}{5}, \frac{2}{3}]$, obtém-se $\lambda = \lambda^*$ ótimo e x^*, que é solução ótima de Q_{λ^*}, também ótima para Q. Por exemplo, se $\lambda^* = (0, \frac{13}{20})$ então:

$$\begin{aligned} v(Q_{(0,\frac{13}{20})}) &= (0, \frac{13}{20})(3,6)^T + \max\{0, 3 - (0, \frac{13}{20})(6,5)^T\} + \max\{0, 4 - (0, \frac{13}{20}(2,6)^T)\} \\ &= \frac{39}{10} + \max\{0, -\frac{1}{4}\} + \max\{0, \frac{1}{10}\} \\ &= \frac{39}{10} + 0 + \frac{1}{10} = 4, \end{aligned}$$

Implicitamente, foi obtido $x^* = (0,1)$ que é factível a Q e satisfaz as folgas complementares:

$$\begin{aligned} (\lambda^*)^T(b - Ax^*) &= \begin{pmatrix} 0 & \frac{13}{20} \end{pmatrix} (\begin{pmatrix} 3 \\ 6 \end{pmatrix} - \begin{pmatrix} 6 & 2 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}) \\ &= \begin{pmatrix} 0 & \frac{13}{20} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0 \end{aligned}$$

Logo, os três itens do Teorema 4.2 são satisfeitos e $x^* = (0,1)$ *é solução ótima para* Q *com* $v(Q) = 4$. *Observe que os únicos pontos factíveis a* Q *são* $(0,0)$ *e* $(0,1)$. *Como a função objetivo é de maximização então* $(0,1)$ *é, de fato, a solução ótima de* Q.

A necessidade de que encontrar bons limitantes (o mais próximo possível da solução ótima) para um problema primal depende de uma boa escolha de λ é constatada pelo Exemplo 4.2.1. Tal fato remete a necessidade de se resolver o problema dual lagrangiano, que fornece o valor dos multiplicadores lagrangianos ótimos, bem como a melhor aproximação para o problema primal, P.

Algumas propriedades foram desenvolvidas para auxiliar na resolução de certos problemas duais lagrangianos; a principal delas, chamada *propriedade de integralidade*, está descrita na Seção 4.2.1 a seguir. Porém, tal propriedade é melhor entendida quando são fornecidos pré-requisitos envolvendo relações existentes entre os limitantes superiores fornecidos pelas diferentes relaxações e problemas duais (mesmo que a propriedade de integralidade não seja necessariamente verificada). Sendo assim, considere o resultado a seguir.

Teorema 4.3. $v(P(\bar{S})) = v(D_L(\bar{S})) \geq v(P_{\tilde{\lambda}}(S)) \geq v(D_L(S)) = v(D_L([S])) = v(P([S]))$, *onde* $\tilde{\lambda}$ *é o vetor associado à solução ótima da relaxação linear* $P(\bar{S})$.

A demonstração deste teorema encontra-se no trabalho de Geoffrion [35]. O fato de $v(P(\bar{S})) = v(D_L(\bar{S}))$ implica diretamente que resolver o dual lagrangiano com o conjunto de restrições, S, relaxado não é mais eficiente do que resolver a relaxação linear de P. $v(D_L(\bar{S})) \geq v(P_{\tilde{\lambda}}(S)) \geq v(D_L(S))$ implica que o multiplicador ótimo para $P(\bar{S})$, embora não necessariamente ótimo para $P_\lambda(S)$, fornece ao menos e, possivelmente, um melhor limitante, ou tão bom quanto, $v(P(\bar{S}))$. Além disso, por $v(D_L(S)) = v(D_L([S]))$, o conjunto S pode ser substituído pelo seu politopo $[S]$ e nenhuma mudança no valor da solução dual lagrangiana ocorrerá. Por fim, $v(D_L(S)) = v(P([S]))$ indica que o valor dual lagrangiano pode ser obtido ao resolver P substituindo S por $[S]$. Esta última igualdade pode ser facilmente verificada, se for observado que:

$$\begin{aligned} v(D_L) &= v[\min_{\lambda \geq 0} v[\max_{x \in S} c^T x + \lambda^T (b - Ax)]] \\ &= v[\min_{\lambda \geq 0} v[\max_{x \in [S]} c^T x + \lambda^T (b - Ax)]] \end{aligned}$$

pois um problema de otimização inteira com função objetivo e restrições lineares possui como solução ótima (se existir), um ponto extremo de seu politopo. Desde que o conjunto de restrições formam um conjunto de pontos inteiros que, ao serem combinados de maneira convexa, formam um politopo, é irrelevante a ordem com que os problemas acima são resolvidos:

$$\begin{aligned} v(D_L) &= v[\max_{x \in [S]} c^T x + v[\min_{\lambda \geq 0} \lambda^T (b - Ax)]] \\ &= v \left[\begin{array}{ll} \max & c^T x + v[\min_{\lambda \geq 0} \lambda^T (b - Ax)] \\ \text{s.a} & Ax \leq b \\ & x \in [S] \end{array} \right] \end{aligned}$$

pois as restrições $Ax \leq b$ são implicitamente satisfeitas pela relaxação e, conseqüentemente, pelo dual lagrangiano na penalização da função objetivo. Desde que $(b - Ax) \geq 0$ implica em $\lambda = 0$

ser sempre solução ótima para P, $\lambda^T(b - Ax) = 0$ e daí, segue a igualdade:

$$v(D_L) = v\left[\begin{array}{ll} \max & c^T x \\ \text{s.a} & Ax \leq b \\ & x \in [S] \end{array}\right] = v(P[S])$$

4.2.1 A Propriedade de Integralidade

Um elemento chave do desenvolvimento de Geoffrion [35] em dualidade lagrangiana para problemas de otimização inteira é a propriedade de integralidade na relaxação lagrangiana. Como citado anteriormente, considere $P(\bar{S})$, a relaxação linear do problema P:

$$\begin{array}{lll} (P(\bar{S})) & \text{maximizar} & c^T x \\ & \text{sujeito a:} & Ax \leq b,\ x \in \bar{S} \end{array}$$

onde $\bar{S} = \{x \geq 0 : Gx \leq h\}$.

Definição 4.2.1. *O problema P_λ possui a propriedade de integralidade se $v(P_\lambda(S)) = v(P_\lambda(\bar{S}))$ para todo $\lambda \geq 0$, isto é, se a relaxação lagrangiana pode ser resolvida como um problema de otimização linear, para todo λ escolhido.*

Isto que dizer que se uma relaxação lagrangiana para o problema P possui a propriedade de integralidade então o limitante fornecido pela relaxação linear de P_λ é igual àquele fornecido pelo próprio P_λ. Logo, não será necessário resolver o problema lagrangiano, que geralmente possui maior complexidade computacional, para obtenção de melhores limitantes para P; basta resolver sua relaxação linear, $P_\lambda(\bar{S})$, como um problema de otimização linear. Neste caso, resolver P_λ ao invés de $P_\lambda(\bar{S})$ será a melhor opção exclusivamente nos casos em que $P_\lambda(\bar{S})$ é muito grande e, geralmente, difícil de ser resolvido em um tempo computacional viável. O significado geométrico da propriedade de integralidade é que os limitantes fornecidos por estes dois problemas, $v(P_\lambda(S))$ e $v(P_\lambda(\bar{S}))$, são iguais, pois os pontos extremos da região factível de P_λ são todos inteiros, isto é, a região factível de P_λ, S, é um politopo de pontos extremos inteiros.

Uma análise sobre os limitantes fornecidos por $P(\bar{S})$ e por D_L, quando é satisfeita a propriedade de integralidade para P_λ, também pode ser estabelecida. A solução ótima de problemas de otimização inteira são sempre dados por pelo menos um dos vértices do maior politopo de pontos inteiros que está contido ou é igual à sua região de factibilidade, S. Se S coincide com este politopo então, como visto acima, a propriedade de integralidade é satisfeita e pode-se desprezar as condições de integralidade de x, obtendo assim, o seguinte problema lagrangiano equivalente:

$$\begin{array}{lll} (P'_\lambda = P_\lambda) & \text{maximizar} & c^T x + \lambda^T(b - Ax) \\ & \text{sujeito a:} & x \in \bar{S} \end{array}$$

Será provado que $v(P(\bar{S})) = v(D_L(S))$, ou seja, que o problema dual lagrangiano e a relaxação linear fornecem o mesmo limitante superior para P, isto é, possuem mesma solução ótima, desde que o problema lagrangiano possua a propriedade de integralidade. Observe que foi denotado anteriormente $v(D_L(S)) = v(P_{\lambda^*}(S))$, onde λ^* é o multiplicador dual ótimo de P_λ.

Teorema 4.4. *Quando P_λ possui a propriedade de integralidade então $v(P(\bar{S})) = v(D_L(S))$.*

Demonstração: Pelo Teorema 4.3, sempre é satisfeita a desigualdade:

$$v(P(\bar{S})) \geq v(D_L) \tag{4.5}$$

Além disso, de acordo a definição de relaxação de Nemhauser e Wolsey [33] disposta na Seção 4.1, o problema lagrangiano P'_λ, que possui a propriedade de integralidade, é uma relaxação do problema $P(\bar{S})$. Então, verifica-se $v(P(\bar{S})) \leq v(P'_\lambda), \forall\ \lambda \geq 0$. Em particular, também é válido para $\lambda = \lambda^*$:

$$v(P(\bar{S})) \leq v(P'_{\lambda^*}) = v(D_L) \tag{4.6}$$

De (4.5)-(4.6), conclui-se a demonstração do teorema. ■

O Teorema 4.4 é importante, pois indica que se P_λ possui a propriedade de integralidade então o dual lagrangiano não fornece um limitante mais potente do que àquele fornecido pela relaxação linear. Portanto, pela Definição 4.3.1, bem como pelo próprio Teorema 4.4, resolver o dual lagrangiano quando P_λ possui a propriedade de integralidade será a melhor opção apenas caso $P(\bar{S})$ ou $P_\lambda(\bar{S})$ sejam problemas complicados de serem resolvidos; encontrar solução dual lagrangiana, em geral, possui maior complexidade computacional que encontrar solução para uma relaxação linear. Alguns métodos para encontrar os multiplicadores lagrangianos ótimos e solução(ões) ótima(s) para D_L estão descritos no Capítulo 5.

4.3 Dualidade *Surrogate*

A *relaxação surrogate* para o problema de otimização inteira P, associado a qualquer vetor $\mu \geq 0$, pode ser definida da seguinte forma:

$$\begin{array}{lll} (P^\mu) & \text{maximizar} & c^T x \\ & \text{sujeito a:} & \mu^T(Ax - b) \leq 0,\ x \in S \end{array}$$

onde $S = \{x \geq 0 : Gx \leq h,\ x \text{ satisfaz algumas restrições discretas}\}$ e μ, denotado pelo vetor contendo os chamados multiplicadores *surrogate* μ_i, $i = 1, \ldots, m$. Assim como no caso lagrangiano, pode-se escrever ainda $v(P^\mu) = \sup_{x \in S}\{c^T x : \mu(Ax - b) \leq 0\}$.

A estratégia do problema *surrogate*, P^μ, é substituir as restrições complicadas $Ax \leq b$, no sentido de que tornam o problema P difícil de ser resolvido, por apenas uma única restrição $\mu^T(Ax - b) \leq 0$, chamada restrição *surrogate*. Para que tal substituição ocorra, a multiplicação entre os vetores $(Ax - b)$ e μ, de dimensões $m \times 1$, é efetuada fazendo surgir a nova restrição *surrogate* como combinação linear não-negativa das restrições complicadas do problema P^μ em questão.

Teorema 4.5. *(Dualidade Surrogate Fraca) Para qualquer escolha do vetor multiplicador surrogate, $\mu \geq 0$, $v(P) \leq v(P^\mu)$.*

Demonstração: A equação $\mu^T(Ax - b) \leq 0$ torna-se redundante quando $Ax - b \leq 0$, pois, por hipótese, $\mu \geq 0$. Então:

$$\begin{aligned} v(P^\mu) &= v[\max_{\mu \geq 0}\{c^T x : \mu^T(Ax - b) \leq 0,\ x \in S\}] \\ &\geq v[\max_{\mu \geq 0}\{c^T x : \mu^T(Ax - b) \leq 0,\ Ax - b \leq 0,\ x \in S\}] \\ &\qquad\qquad \text{(mais restrições implica em menor região factível)} \\ &= v[\max_{\mu \geq 0}\{c^T x : Ax - b \leq 0,\ x \in S\}] \\ &= v(P) \end{aligned}$$

■

Desde que o problema *surrogate* fornece um limitante superior para P, o Teorema 4.5 garante que P^μ é, de fato, uma relaxação para P, com μ não-negativo. A aproximação de $v(P^\mu)$ a $v(P)$ será tanto melhor quanto $\mu^T(Ax - b) \leq 0$ seja uma melhor representação das restrições $Ax \leq b$, pois x estará tornando-se factível a um número cada vez maior de restrições de P. Observa-se, por exemplo, $v(P^\mu) \leq v(P^0)$ para todo $\mu \neq 0$, já que $\mu = 0$ simplesmente elimina a restrição $Ax \leq b$; desta forma, zero pode ser excluído como possível valor para μ. Encontrar o vetor μ que torna a proximidade entre $v(P^\mu)$ e $v(P)$ a menor possível, isto é, escolher o melhor (menor) valor possível para $v(P^\mu)$, é obter qualidade do limitante superior, bem como fornecer restrições *surrogate*, cada vez melhores para P^μ através da resolução do seguinte *problema dual surrogate*:

$$(D_S) \qquad \text{minimizar}_{\ \mu \geq 0}\ \{v(P^\mu)\}$$

que fornece ambos, multiplicador *surrogate* ótimo, μ^*, e melhor solução para a relaxação *surrogate*, $v(P^{\mu^*})$.

Corolário 4.2. *Para qualquer escolha do vetor multiplicador surrogate, $\mu \geq 0$, temos que $v(P) \leq v(D_S)$*

Demonstração: O valor ótimo do problema dual *surrogate* pode ser denotado por $v(D_S) = v(P^{\mu^*}) = \inf_{\mu \geq 0}\{v(P^\mu)\}$. Pela estrutura de D_S e pelo Teorema 4.5, $v(P) \leq v(P^\mu)$ acontece para todo μ factível a D_L, em particular, para $\inf\limits_{\mu \geq 0} v(P^\mu)$. O que implica diretamente em $v(D_S) \stackrel{\triangle}{=} \inf\limits_{\mu \geq 0}\{v(P^\mu)\} \geq v(P)$. ■

Tal como no dual lagrangiano, não se pode garantir obtenção de um vetor μ tal que $v(D_S) = v(P)$, isto é, tal que $v(P^{\mu^*}) = v(P)$. Entretanto, condições necessárias e suficientes para que tal igualdade aconteça podem ser estipuladas como a seguir, no próximo teorema.

Teorema 4.6. *(Condições de Otimalidade Surrogate) Se, para dado vetor μ^*, uma solução x^* satisfizer as seguintes condições:*

(i) *x^* é ótimo para P^{μ^*};*

(ii) *x^* é factível para P: $Ax^* \leq b$;*

então x^* também é uma solução ótima para o problema primal P. Se x^* satisfaz (i) mas não satisfaz (ii) então x^* é solução ótima de $v(P^{\mu^*})$ com gap de dualidade surrogate $v(P^{\mu^*}) - v(P)$. Mais ainda, se x^* satisfizer a condição (iii) $(\mu^*)^T(Ax^* - b) = 0$, então o Teorema possui Condições de Otimalidade Surrogate Fortes.

Demonstração: Será mostrado que $v(P) = v(P^{\mu^*})$ quando x^* satisfaz as hipóteses (i) e (ii) do teorema acima. Considere o problema primal, P, definido anteriormente:

$$\begin{array}{lll}(P) & \text{maximizar} & c^T x \\ & \text{sujeito a:} & Ax - b \leq 0,\ x \in S\end{array}$$

1. Pelo Teorema 4.5, $v(P) \leq v(P^{\mu})$ para todo $\mu \geq 0$. Em particular, para $\mu = \mu^*$, tem-se:

$$v(P) \leq v(P^{\mu^*}).$$

2. Por (i) e (ii), a solução ótima de P^{μ^*} satisfaz $Ax^* - b \leq 0$, isto é, x^* é factível a P. Então:

$$\begin{aligned} v(P) &\stackrel{\triangle}{=} \sup_{x \in S}\{c^T x :\ Ax - b \leq 0\} \\ &\geq c^T x^* = v(P^{\mu^*}) \\ &\Rightarrow v(P^{\mu^*}) \leq v(P). \end{aligned}$$

Por 1 e 2, conclui-se que $v(P) = v(P^{\mu^*})$. Como pela hipótese (i), $v(P^{\mu^*}) = c^T x^*$ então $v(P) = c^T x^*$, isto é, x^* é solução ótima para P. ■

Exemplo 4.3.1. *Para ilustrar as implicações do Teorema 4.6, considere a relaxação surrogate do problema inteiro binário:*

$$\begin{array}{lll}(Q) & \textit{maximizar} & c^T x \\ & \textit{sujeito a:} & Ax \leq b,\ x \in B = \{0,1\}\end{array}$$

dada por:

$$\begin{array}{lll}(Q^{\mu}) & \textit{maximizar} & c^T x \\ & \textit{sujeito a:} & \mu^T(Ax - b) \leq 0,\ x \in B\end{array}$$

Desde que $x \in B$ e a restrição $\mu^T(Ax - b)$ pode ser reescrita por $\sum_j \mu^T a_j x_j \leq \mu^T b$, em que a_j é a j-ésima coluna da matriz A, então Q^{μ} é um problema da mochila 0-1 unidimensional (veja Wolsey [50]):

$$\begin{array}{lll}(Q^{\mu}) & \textit{maximizar} & c^T x \\ & \textit{sujeito a:} & \sum_j a'_j x_j \leq V,\ x \in B\end{array}$$

onde $a'_j = \mu^T a_j$ e $V = \mu^T b$. Considerando as dimensões iguais a do problema P, como o vetor custos possui n coordenadas, então o problema da mochila é definido sobre n itens. Assim, para μ fixo, Q^{μ} pode ser resolvido pelo algoritmo branch and bound (veja, por exemplo, Arenales et al. [51]).

Considerando o caso particular:

$$
\begin{array}{lll}
(Q) & \textit{maximizar} & f(x) = 4x_1+3x_2 \\
& \textit{sujeito a:} & \\
& & 6x_1 + 2x_2 \leq 3 \\
& & 5x_1 + 6x_2 \leq 6 \\
& & x_1, x_2 \in B
\end{array}
$$

então a relaxação surrogate associada é dada pelo problema:

$$
\begin{array}{lll}
(Q^\mu) & \textit{maximizar} & 4x_1 + 3x_2 \\
& \textit{sujeito a:} & \mu_1(6x_1 + 2x_2 - 3) + \mu_2(5x_1 + 6x_2 - 6) \leq 0 \\
& & x_1, x_2 \in B
\end{array}
$$

podendo ser reescrita por:

$$
\begin{array}{lll}
(Q^\mu) & \textit{maximizar} & 4x_1 + 3x_2 \\
& \textit{sujeito a:} & (6\mu_1 + 5\mu_2)x_1 + (2\mu_1 + 6\mu_2)x_2 \leq 3\mu_1 + 6\mu_2 \\
& & x_1, x_2 \in B
\end{array}
$$

Escolhendo aleatoriamente $\mu^T = (1, 6)$ *obtém-se:*

$$
\begin{array}{lll}
(Q^{(1,6)}) & \textit{maximizar} & 4x_1 + 3x_2 \\
& \textit{sujeito a:} & 36x_1 + 38x_2 \leq 39 \\
& & x_1, x_2 \in B
\end{array}
$$

que possui solução ótima $x^* = (1, 0)^T$. *Como* x^* *não satisfaz a condição* (ii) *do Teorema 4.6, então não será solução ótima para* Q. *O que se pode dizer é que* $v(Q^{(1,6)}) = 4 \times 1 + 3 \times 0 = 4$ *é um limitantes superior para* Q.

Entretanto, ao tomar qualquer vetor μ *tal que* $\mu_1 > \frac{\mu_2}{3}$, *obtém-se* $v(Q^\mu) = v(Q)$, *pois as condições* (i) *e* (ii) *deste mesmo teorema serão verificadas. Por exemplo, se* $\mu^* = (2, 1)$, *então:*

$$
\begin{array}{lll}
(Q^{(2,1)}) & \textit{maximizar} & 4x_1 + 3x_2 \\
& \textit{sujeito a:} & 17x_1 + 10x_2 \leq 12 \\
& & x_1, x_2 \in B
\end{array}
$$

que possui solução ótima $x^* = (0, 1)^T$. *Logo, como* x^* *é factível a* Q *então* $v(Q) = v(Q^{(2,1)}) = 4 \times 0 + 3 \times 1 = 3$. *Mais uma vez, pode-se verificar que* $(0, 1)$ *é, de fato, solução ótima para* Q *se for observado que os únicos pontos factíveis a este problema são os pontos* $(0, 0)$ *e* $(0, 1)$.

Ao contrário do caso lagrangiano, não é necessário que sejam verificadas as folgas complementares para que a solução ótima do problema *surrogate*, x^*, seja também solução ótima de P. De acordo com o Teorema 4.6, se a solução ótima do problema dual *surrogate* for factível a P, para algum $\mu \geq 0$, então também será uma solução ótima para P: $v(P) = v(D_S)$. Métodos para determinação de μ^* e x^* ótimos para o caso *surrogate* foram pouco estudados se comparados ao caso lagrangiano. No Capítulo 5 deste presente trabalho, a resolução de D_S foi amplamente estudada e desenvolvida tendo como base as idéias do algoritmo de Sarin *et al.* do trabalho [5].

4.3.1 A Propriedade de Integralidade *Surrogate*

Definição 4.3.1. *O problema P^μ possui a propriedade de integralidade surrogate se $v(P^\mu(S)) = v(P^\mu(\bar{S}))$ para todo $\mu \geq 0$, isto é, se a relaxação surrogate pode ser resolvida como um problema de programação linear para todo μ.*

De acordo com o trabalho de Karwan e Rardin [12], a propriedade de integralidade para o caso *surrogate* é mais difícil de ser verificada ou mesmo, de existir, pois gira em torno do vetor custos c, ao contrário do caso lagrangiano onde a propriedade de integralidade está baseada no conjunto de restrições $Ax \leq b$. Assim, o estudo encontrado na literatura sobre a propriedade de integralidade *surrogate* possui grande foco sobre relações existentes entre os limitantes superiores obtidos pelos duais lagrangiano e *surrogate*, $v(D_L(S))$ e $v(D_S(S))$, quando tal propriedade é verificada. Algumas destas serão desenvolvidas na Seção 4.4 juntamente com outras relações entre as soluções ótimas dos duais mesmo quando nem a propriedade de integralidade lagrangiana e nem a *surrogate* são verificadas.

4.4 Relações entre os Duais Lagrangiano e *Surrogate*

É natural surgir o questionamento de qual das relaxações, linear, lagrangiana ou *surrogate*, fornecem melhores limitantes duais (para o caso de maximização, superiores) para um problema primal de otimização inteira, P. Esta comparação foi estudada por diversos autores no decorrer dos anos. Karwan e Rardin [12] mostraram que uma condição para que nenhum *gap* exista entre o dual lagrangiano e o dual *surrogate* é que P^μ tenha a propriedade de integralidade *surrogate* para todo vetor custos c. Greenberg e Pierskalla [52] provaram, para um problema primal de minimização P, que $v(\bar{P}) \leq v(D_L) \leq v(D_S) \leq v(P)$. Mais ainda, eles desenvolveram condições para que não exista nenhum *gap* entre D_S e P. Karwan e Rardin [12] também mostraram que, em geral, há existência de *gap* entre o dual lagrangiano e o dual *surrogate*, isto é, $v(D_L) < v(D_S)$, com excessão de alguns casos especiais. Então, para o caso de maximização, também é possível denotar o resultado análogo a seguir:

Teorema 4.7. $v(P) \leq v(D_S) \leq v(D_L) \leq v(\bar{P})$

A demonstração do teorema acima é análoga à proposta por Greenberg e Pierskalla [52] para problemas de minimização em otimização inteira. ■

É claro que $v(D_S)$ e $v(D_L)$ fornecem limitantes superiores para P como previamente mostrado pelos Corolários 4.1 e 4.2, respectivamente. Para demonstrar um resultado que interessa a este presente trabalho, $v(D_S) \leq v(D_L)$, observe que $v(P^\mu)$ é definido sobre o conjunto $\{x \in S : \mu^T(Ax - b) \leq 0\}$ que é mais restritivo que o conjunto $\{x \in S\}$ relativo ao qual $v(P_\lambda)$ foi definido. Assim, a região de factibilidade de P^μ é menor que a de P_λ fornecendo $v(P^\mu) \leq v(P_\lambda)$ e, desde que $v(D_S) = \inf_{\mu \geq 0} v(P^\mu)$ e $v(D_L) = inf_{\lambda \geq 0} v(P_\lambda)$, tem-se $v(D_S) \leq v(D_L)$. Além disso, mesmo que $v(P_\lambda)$ fosse definido sobre o conjunto de restrições $\{x \in S : \mu^T(Ax - b) \leq 0\}$ então os multiplicadores $\mu \geq 0$, agora lagrangianos, garantiriam que:

$$\begin{aligned}
\mu^T(b - Ax) \geq 0 &\Rightarrow \max c^T x + \mu^T(b - Ax) \geq \max c^T x \\
&\Rightarrow v(\max c^T x + \mu^T(b - Ax)) \geq v(\max c^T x) \\
&\Rightarrow v(P_\mu) \geq v(P^\mu) \\
&\Rightarrow \inf_{\mu \geq 0}\{v(P_\mu)\} \geq \inf_{\mu \geq 0}\{v(P^\mu)\} \\
&\Rightarrow v(D_L) \geq v(D_S).
\end{aligned}$$

Portanto, o Teorema 4.7 formaliza o resultado de que o dual *surrogate* fornece um limitante superior melhor ou, no máximo, igual ao do dual lagrangiano. Mais ainda, é possível concluir que se as relaxações lagrangiana e *surrogate* não possuem necessariamente a propriedade de integralidade, então os duais lagrangiano e *surrogate* podem fornecer melhores limitantes superiores que o da relaxação linear. Observe que quanto maior for o limitante encontrado pelas relaxações, pior será tal limitante, pois mais distante ele estará da solução ótima de P. Os Teoremas 4.9, 4.10 e 4.11 fornecerão uma caracterização da existência ou não de *gap* entre os valores $v(D_L)$ e $v(D_S)$. Entretanto, considere anteriormente o seguinte resultado.

Teorema 4.8. *1.* $v(D_L(S)) = v(D_S([S]))$

2. $v(P(\bar{S})) = v(D_L(\bar{S})) = v(D_S(\bar{S}))$

3. Se P^μ *tem a propriedade de integralidade surrogate então:*

$$v(P(\bar{S})) = v(P_{\tilde{\lambda}}(S)) = v(D_L(S)) = v(D_S(S))$$

onde $\tilde{\lambda}$ *é o vetor multiplicador associado à solução ótima da relaxação linear* $P(\bar{S})$.

Demonstração:

1. Inicialmente, observe que

$$\begin{aligned}
v(D_S([S])) &= v[\min_{\mu \geq 0} \max\{c^T x : \mu^T A x \leq \mu^T b,\ x \in [S]\}] \\
&= v[\min_{\mu \geq 0} P^\mu([S])].
\end{aligned}$$

Ao considerar $P^\mu([S])$ um problema primal e observar, pela última igualdade do Teorema 4.3, que $v(P^\mu([S]))$ é igual ao(s) valor(es) ótimo(s) de seu dual lagrangiano, onde $x \in [S]$, tem-se:

$$\begin{aligned}
v(D_S([S])) &= v[\min_{\mu \geq 0} \max\{c^T x : \mu^T A x \leq \mu^T b,\ x \in [S]\}] && \text{(Teorema 4.3)} \\
&= v[\min_{\mu \geq 0} \min_{\alpha \geq 0} \max\{c^T x + \alpha(\mu^T b - \mu^T A x) : x \in [S]\}] \\
&= v[\min_{\mu \geq 0} \min_{\alpha \geq 0} \max\{c^T x + \alpha\mu^T(b - Ax) : x \in [S]\}] \\
&= v[\min_{\lambda \geq 0} \max_{x \in [S]}\{c^T x + \lambda^T(b - Ax)\}] \\
&= v(D_L([S])) \\
&\overset{T.4.3}{=} v(D_L(S))
\end{aligned}$$

onde $\lambda = \alpha\mu$.

2. Analogamente à prova do item anterior:

$$\begin{aligned} v(D_S(\bar{S})) &= v[\min_{\mu\geq 0} \max\{c^T x : \mu^T Ax \leq \mu^T b,\ x \in \bar{S}\}] \\ &= v[\min_{\mu\geq 0} \min_{\alpha\geq 0} \max\{c^T x + \alpha(\mu^T b - \mu^T Ax) : x \in \bar{S}\}] \qquad \text{(Teorema 4.3)} \\ &= v[\min_{\mu\geq 0} \min_{\alpha\geq 0} \max\{c^T x + \alpha\mu^T(b - Ax) : x \in \bar{S}\}] \\ &= v[\min_{\lambda\geq 0} \max_{x\in\bar{S}} \{c^T x + \lambda^T(b - Ax)\}] \\ &= v(D_L(\bar{S})) \overset{T.4.3}{=} v(P(\bar{S})) \end{aligned}$$

onde $\lambda = \alpha\mu$.

3. Se $P^\mu(S)$ possui a propriedade de integralidade *surrogate*, então para qualquer $\mu \geq 0$:

$$v(P^\mu(S)) = v(P^\mu(\bar{S})) \text{ implicando em } v(D_S(S)) = v(D_S(\bar{S})) \tag{4.7}$$

pois, em particular, também será válido para μ ótimo. Do Teorema 4.7 e por (4.7):

$$v(D_L(S)) \geq v(D_S(S)) = v(D_S(\bar{S})) \tag{4.8}$$

Pelo Teorema 4.3, tem-se:

$$v(P(\bar{S})) = v(D_L(\bar{S})) \geq v(P_{\tilde{\lambda}}(S)) \geq v(D_L(S)) \tag{4.9}$$

Pelas desigualdades (4.8) e (4.9), obtém-se:

$$v(P(\bar{S})) = v(D_L(\bar{S})) \geq v(P_{\tilde{\lambda}}(S)) \geq v(D_L(S)) \geq v(D_S(S)) = v(D_S(\bar{S})) \tag{4.10}$$

Pela parte 2 deste Teorema, os dois valores extremos da desigualdade acima são iguais: $v(P(\bar{S})) = v(D_S(\bar{S}))$. Assim, pode-se tomar a igualdade por todo conjunto de inequações (4.10) e escolher, em particular, o que era desejado demonstrar:

$$v(P(\bar{S})) = v(P_{\tilde{\lambda}}(S)) = v(D_L(S)) = v(D_S(S)) \qquad \blacksquare$$

Existem várias implicações importantes do teorema anterior. Primeiramente, se $P^\mu(S)$ tem a propriedade de integralidade *surrogate*, nem $D_L(S)$ nem $D_S(S)$ podem melhorar o limitante obtido por $P(\bar{S})$ (embora eles possam fornecer um método mais eficiente de calcular $v(P(\bar{S}))$). Além disso, relaxar $P^\mu(S)$ e resolvê-lo como uma programação linear quando tal problema não possui a propriedade de integralidade *surrogate*, ainda fornece $v(P(\bar{S}))$ como um valor dual (Parte 2 do Teorema 4.8).

A compreensão mais importante deste último resultado, juntamente com o Teorema 4.4, é concluir que se um problema de programação inteira necessita ser resolvido e a propriedade de integralidade lagrangiana e/ou *surrogate* pode(m) ser verificada(s), então os limitantes fornecidos pelo(s) dual(is) lagrangiano e/ou *surrogate* serão idênticos àqueles fornecidos pela relaxação linear. Quando a verificação destas duas propriedades de integralidade não for possível (ou, como dito anteriormente, a relaxação linear for um problema computacionalmente intratável), vale à pena investir em técnicas que resolvam os duais lagrangiano e *surrogate* como será visto no próximo capítulo.

O Teorema a seguir formaliza a idéia de que se os duais lagrangiano e *surrogate* fornecem o mesmo limitante superior para o problema primal P (não há *gap* entre $v(D_L)$ e $v(D_S)$) e se λ é um vetor multiplicador ótimo obtido da resolução de D_L, então λ é também um vetor multiplicador ótimo para D_S e o conjunto das soluções ótimas do problema *surrogate*, P^λ, é formado por todo x ótimo do problema lagrangiano, P_λ, que satisfaz as folgas complementares $\lambda^T(b - Ax) = 0$.

Teorema 4.9. *Se $v(D_L) = v(D_S)$ e $\lambda \in \Omega(D_L)$, então $\lambda \in \Omega(D_S)$ e*

$$\Omega(P^\lambda) = \{x \in \Omega(P_\lambda) : \lambda^T(b - Ax) = 0\} = \Gamma$$

Demonstração:

1. Inicialmente, será mostrado que $\lambda \in \Omega(D_S)$.

   Observe que para todo $x \in S$, se $c^T x > v(D_L) = c^T x + \lambda^T(b - Ax)$ então é claro que $\lambda^T(b - Ax) < 0$, implicando em $\lambda^T(Ax - b) > 0$, não satisfazendo a restrição *surrogate*. Logo, todo $x \in S$ que satisfaz $c^T x > v(D_L)$ será infactível à relaxação *surrogate*, P^λ.

   Desde que a negativa desta implicação é verdadeira, então todo $\tilde{x}$ factível a P^λ deve satisfazer necessariamente $c^T\tilde{x} \leq v(D_L)$, onde $c^T\tilde{x}$ é o valor da relaxação *surrogate* fornecido por $\tilde{x}$, para algum λ fixo. Em particular, também é válido para a solução ótima de P^λ: $v(P^\lambda) \leq v(D_L)$. Unindo esta desigualdade às hipóteses do teorema ($v(D_L) = v(D_S)$ e λ é multiplicador lagrangiano ótimo) e ao resultado do Teorema 4.7 ($v(D_S) \leq v(D_L)$), obtém-se:

   $$\begin{aligned} v(D_L) = v(D_S) = \inf_{\mu \geq 0}\{v(P^\mu)\} \leq v(P^\lambda) \leq v(D_L) \Rightarrow \\ \Rightarrow v(D_L) = v(P^\lambda) = v(D_S) \\ \Rightarrow \text{ a igualdade é válida para todas as desigualdades acima} \\ \Rightarrow \lambda \text{ é multiplicador } \textit{surrogate} \text{ ótimo para } v(D_S) \\ \Rightarrow \lambda \in \Omega(D_S) \end{aligned}$$

2. Agora, será mostrado que $\Omega(P^\lambda) = \Gamma$.

   - Suponha $x \in \Omega(P^\lambda)$.

     Desde que as restrições *surrogate* são satisfeitas para x ótimo a P^λ, então $\lambda^T(Ax - b) \leq 0$ implicando em $\lambda^T(b - Ax) \geq 0$. Como $v(D_L) = v(D_S)$ e $\lambda \in \Omega(D_L)$, então para $x \in S$,

     $$\begin{aligned} v(D_L) = v(D_S) \leq v(P^\lambda) = c^T x \leq c^T x + \lambda^T(b - Ax) = v(P_\lambda) \leq v(D_L) \Rightarrow \\ \Rightarrow \text{ a igualdade é válida para todas as desigualdades acima} \\ \Rightarrow x \in \Omega(P_\lambda) \text{ e } \lambda^T(Ax - b) = 0 \\ \Rightarrow x \in \{x \in \Omega(P_\lambda) : \lambda^T(Ax - b) = 0\} = \Gamma \\ \Rightarrow \Omega(P^\lambda) \subset \Gamma. \end{aligned}$$

- Suponha agora que $x \in \Gamma$.

  Se $x \in \Omega(P_\lambda)$ e $\lambda^T(b - Ax) = 0$ então obtém-se, x, uma solução ótima para o dual *surrogate* $v(D_L) = c^T x + 0 = c^T x = v(D_S)$ que é factível à relaxação *surrogate*: $\lambda^T(Ax - b) \leq 0$. Logo, $x \in \Omega(P^\lambda)$ e, portanto, $\Gamma \subset \Omega(P^\lambda)$. ■

O que se pode concluir dos últimos dois teoremas é que, se P^μ possui a propriedade de integralidade *surrogate*, então D_S fornecerá mesmo limitante superior e mesmo vetor de multiplicadores ótimos que o dual lagrangiano, D_L. De acordo com Karwan e Rardin [12], também é possível verificar condições necessárias para existência de *gap* entre $v(D_S)$ e $v(D_L)$, bem como aplicar um teste construtivo para sua determinação. Os dois teoremas a seguir comprovam tal afirmação e suas respectivas demonstrações podem ser encontradas em [12].

Teorema 4.10. *Ou $v(D_S) < v(D_L)$ ou então, para cada $\lambda \in \Omega(D_L)$, deve existir $x \in \Omega(P_\lambda)$ tal que $\lambda^T(b - Ax) = 0$.*

Teorema 4.11. *Sejam λ um multiplicador ótimo para D_L e o conjunto (possivelmente vazio) Γ. Se ao resolver o problema*

$$
\begin{array}{lll}
(R) & \textit{minimizar} & \alpha \\
 & \textit{sujeito a:} & d(b - Ax) \leq \alpha, \quad \textit{para todo} \quad x \in \Gamma,\ d \geq 0
\end{array}
$$

for obtido como valor ótimo $\alpha = 0$, então $v(D_L) = v(D_S)$. Mais ainda, se S é um conjunto finito e $\alpha \neq 0$, então $v(D_S) < v(D_L)$.

Capítulo 5

Métodos de Solução para os Duais

5.1 Introdução

Como discutido na Seção 4.4, quando as propriedades de integralidade lagrangiana ou *surrogate* não são verificadas, os limitantes duais fornecidos por $v(D_L)$ e $v(D_S)$ estarão mais próximos de $v(P)$ que àquele fornecido por $v(\bar{P})$. Mesmo quando a propriedade de integralidade *surrogate* for verificada e concluir-se que a resolução da relaxação linear for computacionalmente inviável, então tais limitantes ainda poderão ser obtidos para aproximação de $v(P)$. Em ambos os casos, o interesse na resolução de um problema primal de otimização inteira, P, está relacionado ao interesse em resolver os duais lagrangiano e *surrogate* por serem mais fáceis de serem resolvidos que o próprio P. Técnicas de resolução para o dual lagrangiano já foram amplamente estudadas na literatura e algumas delas foram provadas convergir para a solução ótima; como é o caso do *método de otimização do subgradiente* que será apresentado neste capítulo. Segundo Espejo e Galvão [48], devido à dificuldade em trabalhar com as restrições duais *surrogate*, o mesmo estudo sobre técnicas que forneçam multiplicadores ótimos para o caso *surrogate* não acontece, mesmo tendo sido mostrado na literatura D_S fornecer melhores limitantes duais que D_L (veja Teorema 4.7). A adaptação do método de otimização do subgradiente para o dual *surrogate* vem sendo utilizada com bastante freqüência na literatura para a resolução de D_S. Entretanto, tal técnica não garante convergência para os multiplicadores *surrogate* ótimos.

Neste capítulo será apresentado um algoritmo para a resolução do dual *surrogate* proposto por Sarin *et al.* [5] para posterior adaptação ao PCP do Produtor no Capítulo 6. Inicialmente, idéias fundamentais para a resolução dos problemas duais lagrangiano e *surrogate* serão exibidas e acompanhadas por justificativas sobre a dificuldade encontrada em se resolver o dual *surrogate*. O método de otimização do subgradiente também é apresentado para o caso lagrangiano, bem como sua adaptação para o caso *surrogate*.

5.2 A Busca de Multiplicadores Duais Ótimos

O valor da solução ótima de uma relaxação pode ser visto como uma função implícita de seus vetores multiplicadores, isto é, $v(P_\lambda)$ e $v(P^\mu)$ são funções implícitas de λ e μ, respectivamente. Assim, resolver o problema dual, sendo ele lagrangiano ou *surrogate*, é encontrar o melhor vetor,

λ ou μ, respectivamente, que fornece o limitante dual (superior, para o caso de maximização) mais próximo possível de $v(P)$, dado pelo valor ótimo da relaxação lagrangiana ou *surrogate*. O problema crítico no uso efetivo das relaxações é a derivação de um bom conjunto de multiplicadores para os problemas duais correspondentes. Nos Exemplos 4.2.1 e 4.3.1, foi constatada essa necessidade de se encontrar bons vetores multiplicadores para a obtenção de bons limitantes para um problema de otimização inteira, isto é, a necessidade de se resolver o problema dual.

Como será discutido no decorrer deste capítulo, a dificuldade em se trabalhar com as restrições *surrogate* tornaram o dual *surrogate* menos estudado na literatura que o dual lagrangiano, mesmo D_S fornecendo limitantes duais mais próximos para $v(P)$. Inclusive a propriedade de integralidade *surrogate* é mais difícil de ser encontrada (e provada) que a propriedade de integralidade lagrangiana porque, de acordo com a p. 325 do trabalho de Karwan e Rardin [12], todos os exemplos da literatura que possuem a propriedade de integralidade para P_λ são baseados na estrutura do conjunto de restrições $Ax \leq b$ enquanto que para o caso *surrogate*, baseiam-se no vetor custos c (Teorema 2.3 do respectivo trabalho).

As características das funções $v(P^\mu)$ e $v(P_\lambda)$ determinam a existência de bons procedimentos de busca para encontrar os multiplicadores duais ótimos ou quase-ótimos. O valor ótimo da relaxação lagrangiana, $v(P_\lambda)$, por exemplo, é uma função de λ, convexa e linear por partes e torna o desenvolvimento de procedimentos de busca para multiplicadores lagrangianos ótimos mais fáceis que para o caso *surrogate* (para uma revisão bibliográfica completa, veja [53]). Tal fato acontece principalmente porque, como provado por Greenberg e Pierskalla [52], o valor ótimo da relaxação *surrogate*, $v(P^\mu)$, é uma função quase-convexa de μ implicando na possível existência de partes onde a função se mantém constante e que tendem a complicar o processo de busca de multiplicadores duais *surrogate*.

A literatura evidencia que o *método de otimização do subgradiente*, apresentado na seção a seguir, vem sendo utilizado com freqüência para resolver os duais das relaxações lagrangiana, *surrogate* e lagrangiana-*surrogate* (esta última relaxação não será discutida neste presente trabalho, mas referências podem ser encontradas em [48]). Tal método tem se mostrado efetivo para a otimização lagrangiana graças à convexidade da função $v(P_\lambda)$ para o problema de maximização P_λ citada acima. Logo, como visto no Apêndice B, um ponto de mínimo local será um ponto de mínimo global e o método de otimização do subgradiente converge para o multiplicador ótimo. Entretanto, não existem garantias de convergência para o problema P^μ ao utilizar-se tal método para obter os multiplicadores *surrogate*, pois a função $v(P^\mu)$ não possui necessariamente subgradiente que aponta para onde a função cresce em todos os pontos.

Além do método de otimização do subgradiente, diversos algoritmos foram propostos para resolver os duais lagrangiano e *surrogate*. Serão citados os principais algoritmos para o caso *surrogate* devido ao amplo desenvolvimento de técnicas para o caso lagrangiano presentes na literatura. Para referências envolvendo aplicações lagrangianas e lagrangiano-*surrogate*, veja [48]. Os trabalhos de Banerjee [54] e Karwan e Rardin [12] utilizaram o procedimento de Benders que aborda o problema dual *surrogate* indiretamente, resolvendo reformulações de D_S. Karwan e Rardin [55] propuseram utilizar um algoritmo *branch and bound*. Glover [56] desenvolveu um algoritmo que pode ser utilizado para encontrar multiplicadores ótimos para problemas com duas restrições. Dyer [57] propôs dois algoritmos que são extensões naturais de métodos utilizados na relaxação lagrangiana e provou suas convergências. Gavish e Pirkul [58] propuseram um algo-

ritmo para problemas com duas restrições e depois incorporaram este algoritmo a uma heurística para calcular multiplicadores para problemas com mais de duas restrições. Sarin *et al.* [5] desenvolveram um procedimento para a busca dos multiplicadores do dual *surrogate* utilizando uma sucesão de buscas no dual lagrangiano. Lorena e Lopes [46] e Lorena e Narciso [47] propuseram a estratégia de relaxar as restrições de integralidade das variáveis de decisão no problema *surrogate* e utilizar o método de otimização do subgradiente para resolver o dual *surrogate* para o problema de recobrimento de conjuntos e o problema de alocação generalizado. Também são encontradas algumas aplicações da dualidade *surrogate* para problemas especialmente estruturados: a aplicação de Karwan e Pan [59] para o problema de transporte com carga fixa, a implementação de Dinkel e Kochenberger [60] para problemas de programação não-linear, a aplicação de Gavish e Pirkul [58] para o problema da mochila 0-1 multi-dimensional e a aplicação de Glover *et al.* [61] para o problema de planejamento de pessoal em larga escala. Para referências sobre os principais resultados teóricos envolvendo restrições surrogate, veja [5].

5.2.1 Método de Otimização do Subgradiente Aplicado ao Dual Lagrangiano

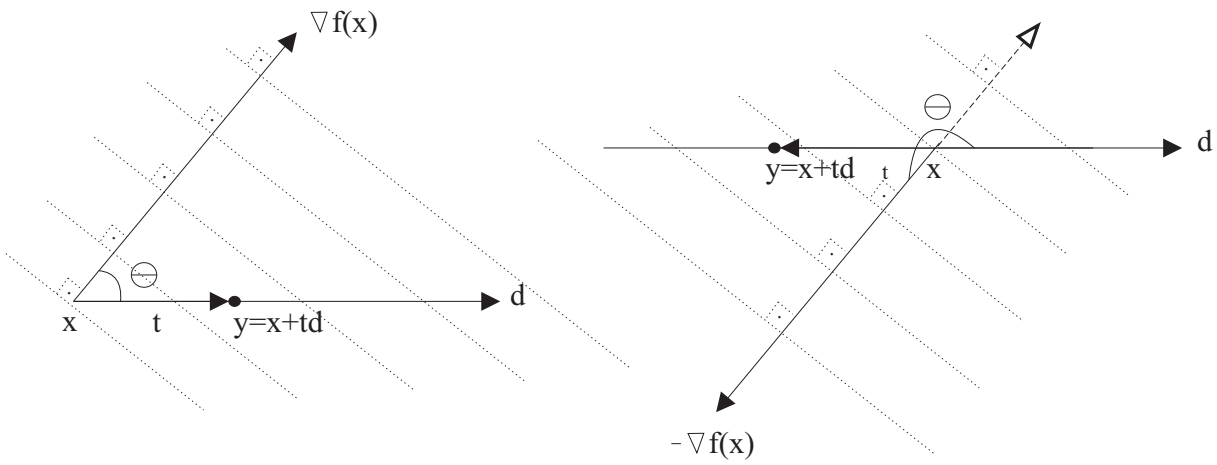
O *método do gradiente* é bem conhecido na literatura e é utilizado para otimizar funções que são diferenciáveis em todos os pontos de seu domínio (para mais detalhes veja [37]). Sabe-se, por definição, que o gradiente de uma função no ponto x é dado por:

$$\nabla f(x) = \left(\begin{array}{cccc} \frac{\partial f(x)}{\partial x_1} & \frac{\partial f(x)}{\partial x_2} & \cdots & \frac{\partial f(x)}{\partial x_n} \end{array} \right)$$

e aponta para onde a função $f(x)$ cresce. Cada parcela $\frac{\partial f(x)}{\partial x_i}$ representa a derivada parcial da função $f(x)$ em relação à x_i no ponto x. Assim, para problemas de minimização da função objetivo $f(x)$ de um dado problema, $-\nabla f(x)$ aponta para onde tal função decresce seu valor.

Analogamente a Valério de Carvalho [62], para um dado ponto x e uma direção d, tal que que $-\nabla f(x) \cdot d > 0$, a idéia da otimização pelo método do gradiente baseia-se no fato de que um pequeno passo, de valor t, dado na direção d produz uma solução com menor valor para a função objetivo, isto é, para o ponto $y = x + td$, tem-se que $f(y) < f(x)$. Observe que o produto interno entre os vetores $\nabla f(x)$ e d ser menor que zero implica diretamente no fato de que o ângulo formado entre eles é obtuso, ou seja, se Θ é tal ângulo, então $\frac{\pi}{2} < \Theta < \frac{3\pi}{2}$. A Figura 5.1 representa informalmente o que acontece com o método do gradiente para que haja o melhoramento da função de $f(x)$ para $f(y) = f(x + td)$ tanto para o caso análogo de maximização, em que $\nabla f(x) \cdot d > 0$ implica em $\frac{3\pi}{2} < \Theta < \frac{\pi}{2}$ (para que $f(y) = f(x+td) > f(x)$), como para o caso de minimização.

O *método de otimização do subgradiente* é uma adaptação do método do gradiente para funções não diferenciáveis em todos os pontos do domínio. É um método iterativo proposto por Held e Karp em 1971 [34] e validado por Held *et al.* em 1974 [63] para problemas de otimização que utilizam uma relaxação R, e o problema dual associado, para sua resolução. Para problemas duais de minimização, a idéia do método é considerar passos tomados, t, na direção negativa de um subgradiente a fim de que os valores da relaxação, que formam uma função convexa, diminuam a cada iteração. Assim, o método de otimização do subgradiente fornece um multiplicador α^k



(a) Problemas de Maximização

(b) Problemas de Minimização

Figura 5.1: Otimização da função f ao caminhar um passo t na direção d de um ponto x para o ponto y

para a relaxação atual em cada iteração k, partindo de um vetor inicial α^0 (por exemplo, o vetor nulo). Como discutido anteriormente, atualização de cada multiplicador atual deve, portanto, ser efetuada na direção oposta a um subgradiente atual, G^K, de acordo com um dado passo, t^k, da seguinte forma:

$$\alpha^{k+1} = \alpha^k - t^k G^k$$

onde

$$t^k = \frac{\varepsilon(v(R\backslash_{\alpha^k}) - v(H))}{\|G^k\|^2}$$

O valor $v(R\backslash_{\alpha^k})$ é o valor da função objetivo da relaxação utilizada, resolvida com o vetor α^k atual; e o valor $v(H)$ é dado pelo cálculo da função objetivo resultante da aplicação de uma heurística, isto é, uma solução factível para P. Portanto, $(v(R\backslash_{\alpha^k}) - v(H))$ é obtido como a diferença entre um limitante superior e um limitante inferior para o problema primal P. O parâmetro ε ajusta o tamanho do passo à medida que pretende-se escolher com maior minúcia os valores de α. Para mais detalhes sobre o desdobramento desde algoritmo, veja [64] para problemas primais de maximização. Observe que, para G^k ser um subgradiente válido, ou seja, apontar para onde a função objetivo da relaxação realmente decresce, $v(R)$ deve ser uma função convexa sobre um problema dual de minimização (a relaxação e o problema primal devem portanto, ser um problema de maximização). Apesar de teoria análoga poder ser realizada para um problema primal de minimização, ela não será tratada no presente trabalho por ser encontrada em grande quantidade na literatura e por não ser utilizada no desenvolvimento dos demais capítulos.

Segundo Oliveira e Morabito [21], variações do método de otimização do subgradiente acima, importantes no presente trabalho, podem ser encontradas na literatura. É o caso de Camerini *et al.* [65] que sugeriram modificações na atualização do tamanho do passo t^k em cada iteração k. Na primeira iteração, o tamanho é dado por um valor fixo ($t^1 = 10^{-5}$) e nas demais iterações, t^k é definido pela seguinte expressão tal que os subgradientes utilizados dependem

diretamente dos subgradientes das iterações anteriores:

$$t^k = \frac{\varepsilon(v(R\backslash_{\alpha^k}) - v(H))}{\|s^k\|^2}$$

onde $s^k = G^k + \beta_k s^{k-1}$, $s^{k-1} = 0$ para $k = 0$ e, para $\gamma = 1.4142$,

$$\beta_k = \begin{cases} -\gamma \frac{s^{k-1}G^k}{\|s^{k-1}\|^2}, & \text{se } s^{k-1}G^k < 0; \\ 0, & \text{caso contrário .} \end{cases} \tag{5.1}$$

Este método de otimização do subgradiente especializado é essencialmente equivalente a utilizar uma soma ponderada de todas as direções subgradiente anteriores e possui uma aplicação mais detalhada no trabalho de Crowder [66].

Características da Relaxação Lagrangiana

Um dos resultados mais importantes sobre a relaxação lagrangiana é apresentado no seguinte teorema.

Teorema 5.1. *Sejam P o problema primal e P_λ a relaxação lagrangiana como definidas anteriormente sobre o conjunto finito de restrições S. Então, $v(P_\lambda)$ é uma função de λ, convexa e linear por partes.*

Demonstração: Pela Definição B.1.5 do Apêndice B, o fato de $[S]$ ser politopo implica necessariamente na existência de uma família de pontos extremos de $[S]$, $\{x_1, x_2, \ldots, x_k\}$, que também são pontos de S e que satisfazem $[S] = [\{x_1, x_2, \ldots, x_k\}]$. Logo:

$$v(P_\lambda) = \max_x \{c^T x + \lambda^T (b - Ax) : x \in S\} = \max_{i=1,\ldots,k} \{c^T x_i + \lambda^T (b - Ax_i)\}$$

Como $v(P_\lambda)$ é definida pela família de funções lineares de λ, $\eta(x_i) = c^T x_i + \lambda^T (b - Ax_i)$, $i = 1, \ldots, k$, então fica provado que $v(P_\lambda)$ é uma função de λ linear por partes. Basta portanto, provar sua convexidade. Observe que todas funções $\eta(x_i)$ são convexas , pois a cada par de pontos extremos tomados, x_i e x_j, $i, j = 1, \ldots, k$ e $i \neq j$, verifica-se para $\rho \in (0, 1)$:

$$\begin{aligned}
\eta(\rho x_i + (1-\rho)x_j) &= c^T(\rho x_i + (1-\rho)x_j) + \lambda^T[b - A(\rho x_i + (1-\rho)x_j)] \\
&= c^T \rho x_i + c^T(1-\rho)x_j + \lambda^T b - \lambda^T A\rho x_i - \lambda^T(1-\rho)Ax_j \\
&= c^T \rho x_i + c^T(1-\rho)x_j + \rho\lambda^T b + (1-\rho)\lambda^T b - \lambda^T A\rho x_i - \lambda^T(1-\rho)Ax_j \\
&= [\rho c^T x_i + \rho\lambda^T b - \rho\lambda^T Ax_i] + [(1-\rho)c^T x_j + (1-\rho)\lambda^T b - (1-\rho)\lambda^T Ax_j] \\
&= \rho(c^T x_i + \lambda^T(b - Ax_i)) + (1-\rho)(c^T x_j + \lambda^T(b - Ax_j)) \\
&= \rho\eta(x_i) + (1-\rho)\eta(x_j),
\end{aligned}$$

satisfazendo a Definição B.1.6.

Desde que o máximo dentre um número finito de funções convexas, $\max_{i=1,\ldots,k} \{c^T x_i + \lambda^T (b - Ax_i)\}$, é uma função convexa, então pode-se concluir que $v(P_\lambda)$ também é convexa. ■

Geometricamente, pode-se interpretar $v(P_\lambda)$ como o envoltório superior de uma família de funções lineares de λ, $\eta = \{c^T x_i + \lambda^T (b - Ax_i)\}$, $i = 1, \ldots, k$ (veja Figura 5.2). A função

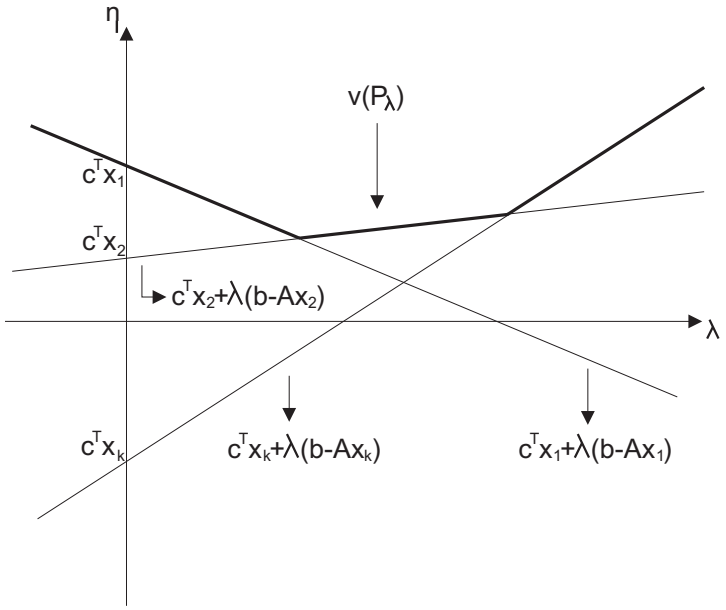


Figura 5.2: Função lagrangiana para o caso de um problema primal de maximização

lagrangiana é convexa e, de fato, é não-diferenciável apenas nos pontos em que a solução ótima não é única, isto é, onde as funções do tipo $c^T x_i + \lambda^T(b - Ax_i)$ se cruzam. Em todos este pontos em que $v(P_\lambda)$ é não-diferenciável, não existe gradiente para a função; entretanto, pelo Teorema B.2, um subgradiente sempre existe desde que a função lagrangiana tenha solução ótima.

Teorema 5.2. *Sejam P o problema primal e P_λ a relaxação lagrangiana como definidas anteriormente sobre o conjunto finito de restrições S. Então a coleção de subgradientes de $v(P_\lambda)$ sobre qualquer λ é chamado de subdiferencial de $v(P_\lambda)$ sobre x, é denotado por $\partial v(P_\lambda)$, e tem a forma:*

$$\{ \sum_{x \in S(\lambda)} \rho_x^T (b - Ax) : \sum_{x \in S(\lambda)} \rho_x = 1,\ \rho_x \geq 0 \quad \textit{para todo} \quad x \in S(\lambda)\} \tag{5.2}$$

onde $S(\lambda) = \{x \in S : x \quad \textit{resolve} \quad P_\lambda\}$.

Demonstração: Esta demonstração foi baseada na idéia encontrada no trabalho de Parker e Rardin [49]. Considere qualquer vetor multiplicador lagrangiano, $\hat{\lambda}$, e seja

$$G = \sum_{x \in S(\hat{\lambda})} \rho_x^T (b - Ax), \ \text{ onde todo } \ \rho_x \geq 0 \ \text{ e } \sum_{x \in S(\hat{\lambda})} \rho_x = 1 \tag{5.3}$$

Note que como $x \in S(\hat{\lambda})$, então o valor $cx + \hat{\lambda}^T(b - Ax)$ deve ser sempre o mesmo. Para provar que (5.2) é a coleção de subgradientes da função $v(P_\lambda)$, primeiramente deve-se mostrar que todo elemento que pertence a esta coleção, é subgradiente de P_λ, ou seja, que G satisfaz a Definição B.1.7, $v(P_\lambda) \geq v(P_{\hat{\lambda}}) + G^T(\lambda - \hat{\lambda})$, para qualquer λ escolhido:

$$
\begin{aligned}
v(P_{\hat{\lambda}}) + G^T(\lambda - \hat{\lambda}) &= \sum_{x \in S(\hat{\lambda})} \rho_x^T [c^T x + \hat{\lambda}^T (b - Ax)] + [\sum_{x \in S(\hat{\lambda})} \rho_x^T (b - Ax)]^T (\lambda - \hat{\lambda}) \\
&= \sum_{x \in S(\hat{\lambda})} \rho_x^T [c^T x + \hat{\lambda}^T (b - Ax) + \lambda^T (b - Ax) - \hat{\lambda}^T (b - Ax)] \\
&= \sum_{x \in S(\hat{\lambda})} \rho_x^T [c^T x + \lambda^T (b - Ax)] \\
&\leq \max_{x \in S(\hat{\lambda})} [c^T x + \lambda^T (b - Ax)] \\
&\leq \max_{x \in S} [c^T x + \lambda^T (b - Ax)] \quad \text{(menos restrições} \Rightarrow \text{melhor solução)} \\
&= v(P_\lambda)
\end{aligned}
$$

Para completar a prova, deve-se mostrar que todo subgradiente de $v(P_\lambda)$ tem a forma (5.2). Tal demonstração é encontrada em [37] na página 192 e requer noções de análise convexa que não serão discutidas no presente trabalho. Geralmente, o desenvolvimento desta prova depende do fato de que subgradientes formam um conjunto convexo com pontos extremos G, que podem ser escritos por $G = (b - Ax)$ para algum $x \in S(\hat{\lambda})$. ■

Uma consequência imediata deste último resultado é que qualquer solução ótima, $\tilde{x}$, fornece um subgradiente, $(b - A\tilde{x})$, para $v(P_\lambda)$ sobre $\tilde{\lambda}$. Quando apenas um $\tilde{x}$ resolve $v(P_{\tilde{\lambda}})$ então existe apenas um subgradiente, $(b - A\tilde{x})$, sobre $\tilde{\lambda}$. Caso contrário, se $\tilde{x}_k$ pontos resolvem $v(P_{\tilde{\lambda}})$ então existem vários subgradientes para a função lagrangiana: $(b - A\tilde{x}_1), (b - A\tilde{x}_2), \ldots, (b - A\tilde{x}_k)$ mais qualquer combinação convexa entre estes vetores. Observe que os vários subgradientes para $v(P_\lambda)$ são encontrados apenas nos pontos em que as retas $\eta(x_i)$ do Teorema 5.1 se cruzam, enquanto que subgradiente único é encontrado em todos os demais pontos de cada uma destas retas. Além disto, como discutido anteriormente, $(b - A\tilde{x})$ é sempre subgradiente válido para $v(P_\lambda)$ desde que a função lagrangiana é convexa (Teorema 5.1) e, por isso, espera-se grande aproveitamento na resolução de D_L através do método de otimização do subgradiente.

Resolução do Dual Lagrangiano pelo Método de Otimização do Subgradiente

Passo 0: *(Inicialização)* Tome qualquer $\lambda^1 \geq 0$ e faça $k \leftarrow 1$, $v^0 \leftarrow \infty$.

Passo 1: *(Relaxação Lagrangiana)* Resolva a relaxação lagrangiana P_{λ^k} obtendo solução ótima $x^k \in S$. Se $(b - Ax^k) \geq 0$ e $\lambda^k(b - Ax^k) = 0$, páre; Teorema 4.2 satifeito com solução ótima x^k resolvendo o problema primal P: $v(P_{\lambda^k}) = v(D_L) = v(P)$.

Passo 2: *(Atualização de v^k)* Se $v^{k-1} \leq v(P_{\lambda^k}) = c^T x^k + \lambda^k(b - Ax^k)$, então faça $v^k \leftarrow v^{k-1}$. Caso contrário, faça $v^k \leftarrow v(P_{\lambda^k})$.

Passo 3: *(Passo Subgradiente)* Calcule o novo multiplicador lagrangiano:

$$\lambda^{k+1} \leftarrow \lambda^k - \rho_k \frac{(b - Ax^k)}{\|b - Ax^k\|} \tag{5.4}$$

onde ρ_k satisfaz

$$\rho_k > 0 \qquad \text{para todo } k,$$

$$\lim_{k\to\infty} \rho_k = 0,$$
$$\sum_{k=1}^{\infty} \rho_k = +\infty$$

PASSO 4: *(Atualização de λ^k)* Projete o novo vetor multiplicador lagrangiano λ^{k+1} sobre $\{\lambda \geq 0\}$, fazendo:

$$\lambda_j^{k+1} \leftarrow \max\{0, \lambda_j^{k+1}\} \qquad \text{para todo } j.$$

Então, faça $k \leftarrow k+1$ e retorne ao Passo 1.

Teorema 5.3. *(Convergência do Método de Otimização do Subgradiente para o Problema Lagrangiano) Seja P_λ a relaxação lagrangiana do problema de otimização discreta P como definida anteriormente sobre o conjunto finito de restrições S. Então, se o método de otimização do subgradiente é aplicado para resolver o dual lagrangiano, D_L, ou ele pára no Passo 1 com $v(P_{\lambda^k}) = v(D_L) = v(P)$, ou gera uma sequência $\{v^k\}$ satisfazendo $\lim_{k\to\infty} v^k = v(D_L)$.*

O método termina no Passo 1, caso as hipóteses do Teorema 4.2 sejam satisfeitas. Então, segue que $v(P_{\lambda^k}) = v(D_L) = v(P)$.

Se o algoritmo não pára, então o Passo 2 garante que $\{v^k\}$ é uma sequência monótona não-crescente. Mais ainda, desde que v^k é gerado por $v(P_{\lambda^k})$ então $v(D_L) \leq v^k$, para todo k. Para demonstrar que $\lim_{k\to\infty} v^k = v(D_L)$, basta mostrar que para qualquer $v(D_L) < \hat{\beta}$ sempre existe um número finito $\hat{k}$ tal que $v^{\hat{k}} \leq \hat{\beta}$, pois, para um valor de $\hat{\beta}$ tão próximo a $v(D_L)$ quanto desejar-se, mostra-se que sempre existe um elemento $k = \hat{k}$ com $v^{\hat{k}}$ ainda menor que $\hat{\beta}$ entre eles: $v(D_L) \leq v^{\hat{k}} \leq \hat{\beta}$. Ou seja, conforme $\hat{\beta}$ aproxima-se de $v(D_L)$, a sequência $v^{\hat{k}}$ fica restrita entre $v(D_L)$ e $\hat{\beta}$, resultando em $v^{\hat{k}} = v(D_L)$ quando k vai para o infinito.

A prova de que $v(D_L) < \hat{\beta}$ está presente na página 220 da referência [49]. ■

5.2.2 Método de Otimização do Subgradiente Aplicado ao Dual *Surrogate*

O método de otimização do subgradiente apresentado na seção anterior também vem sendo utilizado com freqüência na literatura para resolução do problema dual *surrogate*. Entretanto, tal aplicação não é muito eficiente para obtenção de $v(D_S)$, pois, como é mostrado pelo próximo resultado, $v(P^\mu)$ é uma função quase-convexa que permite existência de intervalos tais que $v(P^\mu)$ é constante, complicando o cálculo de subgradientes válidos e não garantindo convergência para o caso *surrogate*.

Teorema 5.4. *Sejam P o problema primal e P^μ a relaxação surrogate como definidas anteriormente sobre o conjunto finito de restrições S. Então, $v(P^\mu)$ é função quase-convexa.*

Demonstração: Esta prova é baseada na demonstração encontrada no trabalho de Greenberg e Pierskalla [52].

Sem perda de generalidade, observe que é possível normalizar arbitrariamente μ ao invés de tomar apenas $\mu \geq 0$, pois:

$$\begin{aligned} v(P^{k\mu}) &= \sup_{x \in S}\{c^T x : (k\mu^T)(Ax - b) \leq 0\} \\ &= \sup_{x \in S}\{c^T x : k(\mu^T(Ax - b)) \leq 0\} \\ &= \sup_{x \in S}\{c^T x : \mu^T(Ax - b) \leq \frac{0}{k}\} \\ &= \sup_{x \in S}\{c^T x : \mu^T(Ax - b) \leq 0\} \\ &= v(P^{\mu}) \end{aligned}$$

Apesar de outras normas poderem ser consideradas, tome $\mu \in \Lambda = \{\mu : \mu \geq 0,\ \Sigma_{i=1}^{m}\mu_i = 1\}$. A vantagem da utilização de Λ ao invés de $\mathbb{R}_+$ encontra-se no fato de Λ ser um conjunto de pontos convexo e limitado. Desde que Λ é convexo então, pela Definição B.1.2, para quaisquer dois pontos $\mu_1, \mu_2 \in \Lambda$ e $0 \leq \rho \leq 1$, sempre existe um segmento de reta que os une contido em Λ, isto é, $\mu = \rho\mu_1 + (1 - \rho)\mu_2 \in \Lambda$.

Sejam x_μ, x_{μ_1} e x_{μ_2} quaisquer soluções ótimas de $v(P^{\mu})$, $v(P^{\mu_1})$ e $v(P^{\mu_2})$, respectivamente. Para qualquer $\rho \in [0, 1]$, pelo menos uma das desigualdades devem ser satisfeitas:

$$(i)\ \mu_1^T(Ax_\mu - b) \leq 0 \qquad\qquad (ii)\ \mu_2^T(Ax_\mu - b) \leq 0$$

pois caso contrário, se ambas desigualdades não fossem verificadas, ou seja, $\mu_1^T(Ax_\mu - b) > 0$ e $\mu_2^T(Ax_\mu - b) > 0$, então seria verdade que:

$$\rho[\mu_1^T(Ax_\mu - b)] + (1 - \rho)[\mu_1^T(Ax_\mu - b)] > 0$$

pois todos os termos da parte esquerda da expressão acima são positivos. Daí, também seria verdade que:

$$\begin{aligned} [\rho\mu_1^T + (1 - \rho)\mu_2^T](Ax_\mu - b) > 0 &\Rightarrow [\rho\mu_1 + (1 - \rho)\mu_2]^T(Ax_\mu - b) > 0 \\ &\Rightarrow \mu^T(Ax_\mu - b) > 0 \end{aligned}$$

o que é um absurdo, pois a solução ótima de P^{μ} seria infactível à suas restrições *surogate*.

Assim, por (i) e (ii), para qualquer $\rho \in [0, 1]$, x_μ é factível a P^{μ_1} e P^{μ_2}, isto é, x_μ fornece um valor pior, ou no máximo igual, para a função objetivo de P^{μ_1} e P^{μ_2} que suas respectivas soluções ótimas x_{μ_1} e x_{μ_2}, tornando satisfeitas as seguintes desigualdades:

$$\begin{aligned} v(P^{\mu}) = c^T x_\mu &\leq \max\{c^T x_{\mu_1}, c^T x_{\mu_2}\} \\ &= \max\{v(P^{\mu_1}), v(P^{\mu_2})\} \end{aligned}$$

Portanto, pela Definição B.1.8 do Apêndice B, o teorema é satisfeito. ■

Assim, o método de otimização do subgradiente pode fornecer soluções ótimas para P_λ mais próximas a P que P^{μ}, mesmo o Teorema 4.7 tendo indicado $v(D_S)$ render limitantes superiores mais eficientes para P que $v(D_L)$. Isto porque o método de otimização do subgradiente aplicado ao dual lagrangiano possui garantias de convergência, enquanto que aplicado ao dual *surrogate* o

mesmo não ocorre; salvos os casos em que P^μ é também função convexa e fornece subgradientes necessariamente válidos nas iterações do algoritmo de otimização do subgradiente.

Por isso, novas formas de atualização dos multiplicadores duais *surrogate* devem ser desenvolvidas a fim de garantir convergência para os multiplicadores ótimos, bem como fornecer melhores limitantes superiores (duais) para P. Uma nova abordagem sobre a resolução de P^μ é apresentada na Seção 5.2.3 com o algoritmo proposto por Sarin *et al.* [5]. Devido à dificuldade encontrada no desenvolvimento de algoritmos com garantias de convergência para os multiplicadores *surrogate* ótimos, trabalhos encontrados na literatura adaptaram o método de otimização do subgradiente para o caso *surrogate* (veja [4], [21], [12]). Por isso, o algoritmo genérico é apresentado abaixo, seguido de algumas condições necessárias para a existência de direções de decrescimento da função *surrogate*, $v(P^\mu)$, que não dependem do subgradiente.

Resolução do Dual *Surrogate* pelo Método de Otimização do Subgradiente

PASSO 0: *(Inicialização)* Tome qualquer $\mu^1 \geq 0$ e faça $k \leftarrow 1$, $v^0 \leftarrow \infty$.

PASSO 1: *(Relaxação Surrogate)* Resolva a relaxação *surrogate* P^{μ^k} obtendo solução ótima $x^k \in S$. Se $(Ax^k - b) \leq 0$, páre; Teorema 4.6 satifeito com solução ótima x^k resolvendo o problema primal P: $v(P^{\mu^k}) = v(D_S) = v(P)$.

PASSO 2: *(Atualização de v^k)* Se $v^{k-1} \leq v(P^{\mu^k}) = c^T x^k$, então faça $v^k \leftarrow v^{k-1}$. Caso contrário, faça $v^k \leftarrow v(P^{\mu^k})$.

PASSO 3: *(Passo Subgradiente)* Calcule o novo multiplicador *surrogate*:

$$\mu^{k+1} \leftarrow \mu^k + \rho_k \frac{(b - Ax^k)}{\|b - Ax^k\|} \tag{5.5}$$

onde ρ_k satisfaz

$$\rho_k > 0 \qquad \text{para todo } k,$$

$$\lim_{k\to\infty} \rho_k = 0,$$

$$\sum_{k=1}^{\infty} \rho_k = +\infty$$

PASSO 4: *(Atualização de μ^k)* Projete o novo vetor multiplicador *surrogate* μ^{k+1} sobre $\{\mu \geq 0\}$, fazendo:

$$\mu_j^{k+1} \leftarrow \max\{0, \mu_j^{k+1}\} \qquad \text{para todo } j.$$

Então, faça $k \leftarrow k + 1$ e retorne ao Passo 1.

Condições Necessárias para Direções de Decrescimento do Problema *Surrogate*

Karwan e Rardin [12] apresentam condições necessárias para direções de crescimento de $v(P^\mu)$ em problemas primais de minimização sobre toda solução ótima $\bar{x}$ para a relaxação *surrogate* com multiplicadores atuais, μ_i. Um contra-exemplo é também apresentado para qualquer tipo de esquema de busca subgradiente a fim de constatar que, de fato, subgradientes para o caso *surrogate* não garantem a geração de direções de crescimento de $v(P^\mu)$.

Condições análogas para o caso de maximização serão aqui desenvolvidas e finalizadas com um contra-exemplo que forneça as mesmas conclusões que o trabalho de Karwan e Rardin. Tal ilustração explica a motivação deste presente trabalho na procura de novas formas de atualização dos multiplicadores *surrogate* que sejam diferentes da fornecida pelo método de otimização do subgradiente aplicado ao dual *surrogate*.

Teorema 5.5. *Seja $\bar{\mu} \geq 0$ um multiplicador surrogate atual. Qualquer direção d para o qual existe $\sigma \geq 0$ satisfazendo $v(P^{\bar{\mu}}) > v(P^{\bar{\mu}+\sigma d})$, isto é, no qual é uma direção de decrescimento para o problema surrogate, deve satisfazer:*

$$d^T(Ax - b) > 0, \quad \textit{para todo} \quad x \in \Omega(P^{\bar{\mu}}). \tag{5.6}$$

Demonstração: Seja x^* qualquer elemento de $\Omega(P^{\bar{\mu}})$. Pela sua factibilidade a $P^{\bar{\mu}}$, $\bar{\mu}^T(Ax^* - b) \leq 0$. Suponha, por absurdo, que $d^T(Ax^* - b) \leq 0$. Então

$$(\bar{\mu} + \sigma d)^T(Ax^* - b) = \underbrace{\bar{\mu}^T(Ax^* - b)}_{\leq 0} + \sigma \underbrace{d^T(Ax^* - b)}_{\leq 0} \leq 0$$

para todo $\sigma \geq 0$. Mas $(\bar{\mu} + \sigma d)^T(Ax^* - b) \leq 0$ implica em x^* factível a $P^{\bar{\mu}+\sigma d}$, isto é, desde que a relaxação *surrogate* é um problema de maximização, $c^T x^*$ é um valor pior (menor), ou no máximo igual, que a solução ótima de $P^{\bar{\mu}+\sigma d}$, $v(P^{\bar{\mu}+\sigma d})$: $v(P^{\bar{\mu}+\sigma d}) \geq c^T x^* = v(P^{\bar{\mu}})$. Portanto, como a hipótese de que $v(P^{\bar{\mu}}) > v(P^{\bar{\mu}+\sigma d})$ foi violada, pode-se concluir que d é direção de decrescimento para o problema *surrogate* sobre μ atual se, e somente se, $d^T(Ax - b) > 0$ para todo $x \in \Omega(P^{\bar{\mu}})$. ■

Corolário 5.1. *Qualquer direção d para o qual existe $\sigma \geq 0$ satisfazendo $v(P^{\bar{u}}) > v(P^{\bar{u}+\sigma d})$ é uma direção de decrescimento para o problema lagrangiano:*

$$\begin{array}{lll} (P_0^{\bar{\mu}}) & \textit{minimizar} & c^T x + 0(b - Ax) \\ & \textit{sujeito a:} & x \in S,\ \bar{\mu}^T(Ax - b) \leq 0 \end{array}$$

Demonstração: $(P_0^{\bar{\mu}})$ é idêntico a $P^{\bar{\mu}}$. Assim, $\Omega(P_0^{\bar{\mu}}) = \Omega(P^{\bar{\mu}})$ e as direções de decrescimento para $(P_0^{\bar{\mu}})$ são idênticas àquelas que satisfazem (5.6). ■

O exemplo a seguir, baseado em [12], mostra que o subgradiente é único e não fornece necessariamente uma direção de decrescimento para o dual *surrogate*, embora esta exista.

Exemplo 5.2.1. *Considere o seguinte problema primal:*

$$\begin{array}{lll} (P) & \textit{maximizar} & -x_1 - x_2 + x_3 \\ & \textit{sujeito a:} & -x_1 + x_2 - 1 \leq 0 \\ & & -x_2 + 1 \leq 0 \\ & & x_1 + x_2 + 11x_3 - 10 \leq 0 \\ & & x \in S \end{array}$$

onde $S = \{x : \ 0 \leq x_i \leq 2, \quad inteiro, \ i = 1, 2, 3\}$. *Tome* $\bar{\mu} = \begin{pmatrix} 5 & 1 & 20 \end{pmatrix}$ *como multiplicador atual surrogate. A restrição surrogate para* $P^{\bar{\mu}}$ *será:*

$$\bar{\mu}^T(Ax - b) \leq 0 \Rightarrow \begin{pmatrix} 5 & 1 & 20 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ 1 & 1 & 11 & -10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix} \leq 0$$

$$\Rightarrow \begin{pmatrix} 15 & 24 & 220 & -204 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix} \leq 0$$

$$\Rightarrow 15x_1 + 24x_2 + 220x_3 - 204 \leq 0$$

É possível resolver a relaxação surrogate atual:

$$\begin{array}{lll} (P^{\bar{\mu}}) & \textit{maximizar} & -x_1 - x_2 + x_3 \\ & \textit{sujeito a:} & 15x_1 + 24x_2 + 220x_3 - 204 \leq 0, \ x \in S \end{array}$$

por inspeção, ao observar que, desde que x_1 *e* x_2 *penalizam a função objetivo, toma-se* $x_1 = x_2 = 0$ *e, para satisfazer a restrição surrogate, toma-se também* $x_3 = 0$ *. Logo, é obtida solução ótima* $x^* = (0, 0, 0)$. *Seja:*

$$(Ax^* - b) = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & 11 \end{pmatrix} \vec{0} - \begin{pmatrix} 1 \\ -1 \\ 10 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 1 & -10 \end{pmatrix}^T$$

Note que $(Ax^* - b)$ *é único, uma vez que* x^* *também é único, e que* $\bar{\mu}^T(Ax^* - b) = -204 \leq 0$, *isto é,* x^* *é factível a* $P^{\bar{\mu}}$. *Tomando como direção de decrescimento para* $v(P^{\bar{\mu}})$ *o vetor subgradiente acima,* $d = (Ax^* - b)$, *será mostrado, por absurdo, que o fato de* $d^T(Ax^* - b) > 0$ *não implica que* $d = (Ax^* - b)$ *seja direção de decrescimento para o dual surrogate, ou seja, que a condição* (5.6) *do Teorema 5.5 não é suficiente; é apenas necessária. Além da volta deste Teorema não ser sempre verificada, também será possível concluir que, de fato, o subgradiente para o problema surrogate nem sempre é direção de decrescimento para* D_S.

1. *Primeiramente, observe que o ponto super-ótimo (infactível) a P,* $\hat{x} = (0, 0, 1)$, *é "cortado" por* $\bar{\mu}$ *uma vez que é infactível a* $P^{\bar{\mu}}$, *isto é:*

$$\bar{\mu}^T(A\hat{x} - b) = \begin{pmatrix} 5 & 1 & 20 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ 1 & 1 & 11 & -10 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 5 & 1 & 20 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

$$= 16 > 0.$$

Em toda iteração de qualquer método de resolução do dual surrogate, deseja-se obter um esquema de atualização de $\bar{\mu}$ para um vetor mutiplicador $\bar{\mu} + \sigma d$ que minimize ainda mais o valor da relaxação surrogate. Uma condição para que isto aconteça é tornar x^*, solução ótima atual, infactível ao novo problema $P^{\bar{\mu}+\sigma d}$ para que x^* seja excluída como uma possível solução ótima para a nova relaxação surrogate. Isto é, deseja-se reduzir a região de factibilidade de $P^{\bar{\mu}}$ para a região de factibilidade de $P^{\bar{\mu}+\sigma d}$ de forma que, pelo menos, x^* seja excluída como possível solução ótima. Mas, "cortar" x^* da região de factibilidade de $P^{\bar{\mu}+\sigma d}$ é torná-lo infactível a este mesmo problema:

$$
\begin{aligned}
(\bar{\mu} + \sigma d)^T (Ax^* - b) > 0 \Rightarrow \bar{\mu}^T (Ax^* - b) + \sigma d^T (Ax^* - b) > 0 \\
\Rightarrow -204 + \sigma \begin{pmatrix} -1 & 1 & -10 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ -10 \end{pmatrix} > 0 \\
\Rightarrow -204 + 102\sigma > 0 \\
\Rightarrow \sigma > \frac{204}{102} \\
\Rightarrow \sigma > 2
\end{aligned}
$$

2. *Além disso, como a nova região de factibilidade de $P^{\bar{\mu}+\sigma d}$ deve estar contida na região de factibilidade de $P^{\bar{\mu}}$ e como $\hat{x}$ era infactível a $P^{\bar{\mu}}$, então $\hat{x}$ também deve ser infactível a $P^{\bar{\mu}+\sigma d}$:*

$$
\begin{aligned}
(\bar{\mu} + \sigma d)^T (A\hat{x} - b) > 0 \Rightarrow \bar{\mu}^T (A\hat{x} - b) + \sigma d^T (Ax^* - b) > 0 \\
\Rightarrow 16 + \sigma \begin{pmatrix} -1 & 1 & -10 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} > 0 \\
\Rightarrow 16 - 8\sigma > 0 \\
\Rightarrow \sigma < 2
\end{aligned}
$$

De 1 e 2, chega-se a um absurdo. Logo, pode-se concluir que mesmo que a equação (5.6) *esteja sendo satisfeita, isto é, $d^T(Ax^* - b) = 102 > 0$, d, tomado como o vetor subgradiente, não é direção de decrescimento para P^{μ} quando $\mu = \bar{\mu}$, pois atualizar $\bar{\mu}$ para $\bar{\mu} + \sigma d$ através de d não minimiza $v(P^{\mu})$, já que σ não pode ser, simultaneamente, estritamente maior que 2 e estritamente menor que 2. Portanto, a volta do Teorema 5.5 não é necessariamente satisfeita.*

5.2.3 Método Sarin *et al.* Aplicado ao Dual *Surrogate*

Baseado no desenvolvimento de novas relações entre a dualidade lagrangiana e *surrogate*, Sarin *et al.* [5] apresentaram um procedimento de busca para os multiplicadores duais *surrogate* que permite utilizar diretamente métodos de busca lagrangiano já conhecidos, como por exemplo, o método de otimização do subgradiente descrito na Seção 5.2.1. Até então, o algoritmo de relaxação linear relatado em Karwan e Rardin [45] era o procedimento de busca dual *surrogate* mais eficiente. Entretanto, o método Sarin *et al.* modificou essa realidade ao ser comparado

empiricamente com o algoritmo de relaxação linear para problemas de "receita de capital" randomicamente gerados (para mais detalhes, veja [5]) e por esta razão, o estudo desta seção estará focado neste novo procedimento de Sarin *et al.*.

A teoria apresentada nesta seção é amplamente análoga àquela encontrada em Karwan e Rardin [45] e Sarin *et al.* [5]. A grande modificação encontra-se na apresentação da teoria para problemas de maximização, raramente encontrada na literatura, e na exibição de mais detalhes nas demonstrações e compreensão dos resultados apresentados. Para facilitar o entendimento de futuras comparações com a relaxação lagrangiana, sem perda de generalidade, é possível reescrever a relaxação *surrogate* da seguinte forma:

$$\begin{array}{lll} (P^{\mu}) & \text{maximizar} & c^T x \\ & \text{sujeito a:} & \mu^T(b - Ax) \geq 0,\ x \in S \end{array}$$

onde $S = \{x \geq 0 : Gx \leq h, x \text{ satisfaz algumas restrições discretas}\}$.

Em geral, enquanto os métodos de busca lagrangiana contam com a idéia de penalização da função objetivo através do termo do vetor custos de (P_λ), $\lambda^T(b - Ax)$, os métodos de busca *surrogate* tornam um $x \in S$ super-ótimo infactível aos subproblemas P^{μ} (veja a ilustração de uma iteração dual no Exemplo 5.2.1). Mais especificamente, eles requerem vetores multiplicadores μ satisfazendo o seguinte sistema de inequações:

$$(I_{S(\alpha)}) \qquad \mu^T(b - Ax) \leq \delta \qquad \text{para todo } x \in S(\alpha),\ \mu \geq 0,\ \delta < 0$$

onde α é uma constante dada e $S(\alpha) = \{x \in S : \alpha \leq c^T x\}$. Qualquer μ satisfazendo $(I_{S(\alpha)})$ implicará que todo $x \in S(\alpha)$ é infactível em P^{μ}. Daí, tem-se $v(P^{\mu}) < \alpha$, pois desde que, por definição, $\alpha \leq c^T x$, então $v(P^{\mu}) < \alpha$ corta todo $x \in S(\alpha)$, já que $c^T x$ será um limitante superior pior que α. O seguinte resultado formaliza a idéia de que o maior α que não admite solução para $(I_{S(\alpha)})$ deve ser igual a $v(D_S)$ (basta tomar a negativa da implicação do Teorema 5.6).

Teorema 5.6. *Sejam P^{μ} e $(I_{S(\alpha)})$ como definidos anteriormente. Então $(I_{S(\alpha)})$ tem uma solução (μ, δ) se, e somente se, $v(D_S) < \alpha$.*

Demonstração:

- Considere que para $\alpha = \hat{\alpha}$, $(\hat{\mu}, \hat{\delta})$ é uma solução para $I_{S(\hat{\alpha})}$.

  Então todo $x \in I_{S(\hat{\alpha})}$ é infactível em $P^{\hat{\mu}}$, pois contraria $\hat{\mu}^T(b - Ax) \geq 0$. E, como discutido previamente, obtém-se $v(P^{\hat{\mu}}) < \hat{\alpha}$. Desde que $v(D_S) \stackrel{\triangle}{=} \inf_{\mu \geq 0}\{v(P^{\mu})\}$, então $v(D_S) \leq v(P^{\mu})$; em particular, para $\mu = \hat{\mu}$, $v(D_S) \leq v(P^{\hat{\mu}})$, provando assim, a primeira parte deste Teorema.

- Considere que para $\alpha = \hat{\alpha}$, $v(D_S) < \hat{\alpha}$.

  Desde que foi assumido, no início deste trabalho, P limitado e factível, então $v(D_S)$ é finito. Como $\inf_{\mu \geq 0}\{v(P^{\mu})\} \stackrel{\triangle}{=} v(D_S) < \hat{\alpha}$, deve existir um $\hat{\mu} \geq 0$, tal que $v(P^{\hat{\mu}}) < \hat{\alpha}$. Mas isto implica, mais uma vez, que todo $x \in I_{S(\alpha)}$ é infactível a $P^{\hat{\mu}}$; o que implica que $\hat{\mu}^T(b - Ax) < 0$ para todo $x \in I_{S(\alpha)}$. Além disso, como também foi assumido S um politopo, então $S(\hat{\alpha}) = \{x \in S : \hat{\alpha} \leq c^T x\} \subseteq S$ também o será e, assim, deve existir um

$\hat{x} \in S(\hat{\alpha})$ específico tal que:

$$\sup_{x \in S(\hat{\alpha})} \{\hat{\mu}^T(b - Ax)\} = \hat{\mu}^T(b - A\hat{x}) < 0.$$

Escolhendo $\hat{\delta} = \hat{\mu}^T(b - A\bar{x})$, completa-se a construção de uma solução $(\hat{\mu}, \hat{\delta})$ para $I_{S(\hat{\alpha})}$, uma vez que $\hat{\mu}^T(b - Ax) \leq \hat{\mu}^T(b - A\hat{x}) = \hat{\delta}$. ■

A implicação do Teorema 5.6 é que resolver $I_{S(\alpha)}$ para valores cada vez menores de α pode levar a um algoritmo para a busca dos multiplicadores duais *surrogate*, pois α estará se tornando cada vez mais próximo de $v(D_S)$.

A informação análoga de que $\alpha < v(D_S)$ para problemas primais de minimização é utilizada por Karwan e Rardin [45] para propor um algoritmo de busca *surrogate* genérico que interage entre resolver relaxações *surrogate* P^μ e modificar μ para aumentar o valor da próxima relaxação *surrogate*. Três variantes deste algoritmo foram implementadas, diferindo apenas em como μ foi atualizado. O procedimento envolvendo relaxação linear mostrou-se o mais efetivo dentre tais variantes e tornou-se o método mais poderoso para encontrar multiplicadores *surrogate* até ser comparado com o algoritmo de Sarin *et al.* que aqui será apresentado.

Enfatizando o resultado anterior, pode-se ainda afirmar que qualquer α para o qual $I_{S(\alpha)}$ admite solução será um limitante superior para $v(D_S)$; e da mesma forma, pode-se afirmar que para o maior α fixo tal que $I_{S(\alpha)}$ não possui nenhuma solução tem-se $v(D_S) = \alpha$. Assim, desde que esta teoria gira em torno de resolver $I_{S(\alpha)}$ e verificar se existe pelo menos uma solução para tais inequações, é possível então resolver o seguinte problema:

$$\begin{array}{lll} (P_{\alpha,\mu}) & \text{maximizar} & f(x) = \mu^T(b - Ax) \\ & \text{sujeito a:} & x \in S(\alpha) \end{array}$$

Se $f(x) < 0$, então $I_{S(\alpha)}$ possui solução e, conseqüentemente, $v(D_S) < \alpha$. Esta idéia será formalizada com o resultado a seguir.

Corolário 5.2. *Se para qualquer $\mu \geq 0$, for obtido $v(P_{\alpha,\mu}) < 0$, então $v(D_S) < \alpha$.*

Demonstração: Tem-se que $v(P_{\alpha,\mu}) < 0$ se e somente se $\mu^T(b - Ax) < 0$, para todo $x \in S(\alpha)$. Mas $\mu^T(b - Ax) < 0$ implica que $I_{S(\alpha)}$ possui uma solução (basta pensar em $\delta = 0$), o que implica, pelo Teorema 5.6, que $v(D_S) < \alpha$. ■

Além disso, $(P_{\alpha,\mu})$ pode ser pensado como a relaxação lagrangiana do problema a seguir:

$$\begin{array}{lll} (P_\alpha) & \text{maximizar} & z \stackrel{\triangle}{=} 0 \\ & \text{sujeito a:} & Ax \leq b,\ x \in S(\alpha) \end{array}$$

com o conjunto de restrições $Ax \leq b$ relaxado pelos multiplicadores, agora lagrangianos, $\mu \geq 0$. O dual lagrangiano pode ser então escrito por:

$$(D_{\alpha,L}) \qquad \text{minimizar}_{\ \mu \geq 0}\{v(P_{\alpha,\mu})\}$$

onde $v(D_{\alpha,L}) = \inf_{\mu \geq 0}\{v(P_{\alpha,\mu})\}$. Uma característica muito particular da função de μ, $v(D_{\alpha,L})$, é que seu valor é sempre igual a zero ou a menos infinito, para qualquer α fixo.

Corolário 5.3. *Para qualquer α fixo, $v(D_{\alpha,L})$ possui valor zero ou então é ilimitado inferiormente.*

Demonstração: $v(D_{\alpha,L}) > 0$ é impossível, desde que $v(D_{\alpha,L})$ é o ínfimo de uma função e pode-se sempre escolher os multiplicadores $\mu = (0, 0, \ldots, 0)$ tornando $v(D_{\alpha,L}) = 0$. Se $v(D_{\alpha,L}) < 0$, então existe um vetor multiplicador μ tal que:

$$v(P_{\alpha,\mu}) = \mu^T(b - Ax) < 0$$

Como, por hipótese, $\mu \geq 0$, é possível escolher um múltiplo de μ arbitrariamente grande, tornando assim $v(P_{\alpha,\mu})$, e conseqüentemente, $v(D_{\alpha,L})$, arbitrariamente pequeno. ■

Combinando os resultados anteriores, chega-se à seguinte relação entre os duais $v(D_{\alpha,L})$ e D_S.

Corolário 5.4. *Para um dado α, se $v(D_{\alpha,L}) \to -\infty$, então $v(D_S) < \alpha$.*

Demonstração: Se $v(D_{\alpha,L}) \to -\infty$, como $v(D_{\alpha,L}) = \inf_{\mu \geq 0}\{v(P_{\alpha,\mu})\}$, então $v(P_{\alpha,\mu}) \to -\infty$. Logo, $v(P_{\alpha,\mu}) = \mu^T(b - Ax) < 0$ implica, pelo Corolário 5.2, que $v(D_S) < \alpha$ para qualquer $\mu \geq 0$. ■

A seguinte análise da teoria apresentada até aqui pode ser explicitada para melhor compreensão da idéia de Sarin *et al.*:

$$\begin{aligned} v(D_{\alpha,L}) \neq 0 &\Rightarrow v(D_{\alpha,L}) = -\infty \\ &\Rightarrow v(P_{\alpha,\mu}) < 0 \\ &\Rightarrow v(D_S) < \alpha \\ &\Rightarrow I_{S(\alpha)} \text{ possui solução} \\ &\Rightarrow \text{Todo } x \in S(\alpha) \text{ é infactível a } P^\mu. \end{aligned}$$

Ou seja, a principal idéia desta teoria traduz a idéia básica dos algoritmos de busca dual *surrogate* descrita no início desta seção: tornar um $x \in S$ super-ótimo infactível aos subproblemas P^μ. Entretanto, a estratégia do algoritmo de Sarin *et al.* é diferente; ao invés de resolver D_S por repetidas resoluções de P^μ, considera-se α fixo e resolve-se $D_{\alpha,L}$ com subproblemas $P_{\alpha,\mu}$. O custo da utilização desta estratégia recai, portanto, em como escolher e atualizar os valores de α, pois quanto menor for seu valor, mais valores super-ótimos de x estarão sendo eliminandos como possíveis soluções ótimas (pois, por hipótese, $\alpha \leq c^T x$). Desde que dentro do algoritmo estarão sendo resolvidos problemas lagrangianos, qualquer técnica de busca subgradiente poderá ser empregada, evitando-se assim a maior complexidade em trabalhar-se com procedimentos de busca *surrogate*. A análise abaixo, análoga àquela feita acima, também tem importante significado dentro do algoritmo de Sarin *et al.* e será utilizada para a demonstração da parte (*ii*) do Teorema 5.7.

$$\begin{aligned} v(D_{\alpha,L}) = 0 &\Rightarrow v(P_{\alpha,\mu}) = 0 \\ &\Rightarrow \mu^T(b - Ax) = 0 \\ &\Rightarrow I_{S(\alpha)} \text{ não possui solução, para todo } x \in S(\alpha) \\ &\qquad (\text{pois } 0 = \mu^T(b - Ax) \leq \delta \text{ é absurdo já que, por definição, } \delta < 0) \\ &\overset{T5.6}{\Longrightarrow} v(D_S) \geq \alpha \end{aligned}$$

Outra importante observação para compreensão de algumas partes do Algoritmo de Sarin *et al.* é apresentada a seguir. Como mostrado na demonstração do Corolário 5.3, se $v(P_{\alpha,\mu}) < 0$ então $v(D_{\alpha,L}) = -\infty$ implicando diretamente em $v(D_S) < \alpha$; ou seja, α deve ser reduzido para a eliminação de mais pontos $x \in S$ super-ótimos. Além disso, como $v(D_{\alpha,L}) \leq v(P_{\alpha,\mu})$ (pois para α fixo, $v(D_{\alpha,L}) \stackrel{\triangle}{=} \inf_{\mu \geq 0}\{v(P_{\alpha,\mu})\}$), se $v(P_{\alpha,\mu}) \geq 0$ então não se pode afirmar que $v(D_{\alpha,L}) = 0$ ou que $v(D_{\alpha,L}) = -\infty$. Deve-se aplicar um algoritmo de otimização do subgradiente para encontrar o valor do respectivo problema dual até que $v(D_{\alpha,L}) = -\infty$ (Passo 2(b) do algoritmo) ou até que α seja um limitante inferior para P_α (Passo 2(c) do algoritmo) para posterior atualização de α.

Observe ainda que, desde que $v(D_{\alpha,L}) = 0$ ou $D_{\alpha,L}$ é ilimitado, de acordo com Sarin *et al.*, é possível resolver o dual $D_{\alpha,L}$ utilizando um algoritmo dual descendente como a seguir: Inicialize com os multiplicadores $\mu = (0, 0, \ldots, 0)$. $v(P_{\alpha,\mu}) = \mu^T(b - Ax) = 0$ para todo $x \in S(\alpha)$. Dada a solução atual μ, se existe direção descendente então obviamente pode-se encontrar uma nova solução $\hat{\mu}$ tal que $v(P_{\alpha,\mu}) < 0$, implicando em $v(D_{\alpha,L}) = -\infty$. Caso contrário, se não existe direção descendente, então sempre $v(D_{\alpha,L}) = 0$ e um critério de parada no algoritmo de Sarin *et al.*, a seguir, deve ser utilizado para que a solução do dual não entre em *looping* infinito (no Passo 3). Assim a busca dual descendente lagrangiana para resolver $v(D_{\alpha,L})$ é reduzida a um problema de determinar a existência ou não de uma direção descendente para o dual lagrangiano, isto é, de determinar se em algum momento $v(P_{\alpha,\mu}) < 0$ ou se $v(D_{\alpha,L}) = 0$, respectivamente. É claro que, quando $v(P_{\alpha,\mu}) < 0$ então o método de otimização do subgradiente deve convergir para um solução ótima, uma vez que sempre existe direção descendente. Tal resolução do dual $v(D_{\alpha,L})$ é realizada implicitamente nos Passos 2 e 3 do método descrito a seguir.

Método Sarin *et al.*

Passo 0: *(Inicialização)* Escolha um valor máximo de iterações duais $klim$, limitantes inferior e superior iniciais para $v(D_S)$, ll e ul, uma tolerância de aproximação $\epsilon > 0$, um valor negativo $\hat{z}$ e um fator $\theta \in (0, 1)$. Faça $\hat{\alpha} \leftarrow ul - \theta(ul - ll)$ e tome qualquer $\hat{x} \in S(\hat{\alpha})$.

Passo 1: *(Inicialização de $D_{\alpha,L}$)* Inicialize o contador de iterações duais $k \leftarrow 1$, uma solução dual inicial $\hat{\mu} \geq 0$, o vetor subgradiente suavizado $\hat{G} = 0$ e o tamanho inicial $t = 10^{-5}$.

Passo 2: *(Resolução de $P_{\alpha,\mu}$)* Se $\hat{\mu}^T(b - A\hat{x}) \geq 0$, "término antecipado" pois $v(P_{\hat{\alpha},\hat{\mu}}) \geq 0$; vá para o Passo 3 modificar $\hat{\mu}$. Caso contrário, aplique um método enumerativo para resolver $(P_{\hat{\alpha},\hat{\mu}})$, utilizando $\hat{x}$ como solução incumbente.

2(a): Se em algum estágio do método enumerativo, $\hat{\mu}^T(b - A\hat{x}) \geq 0$, "término mais cedo" de $v(P_{\hat{\alpha},\hat{\mu}})$; vá para o Passo 3 modificar $\hat{\mu}$. Caso contrário, prossiga até provar otimalidade de $\hat{x}$ para $P_{\hat{\alpha},\hat{\mu}}$:

2(b): Se $\hat{\mu}^T(b - A\hat{x}) \stackrel{\triangle}{=} v(P_{\hat{\alpha},\hat{\mu}}) < 0$, termine as iterações duais tomando $ul \leftarrow \hat{\alpha}$ e vá para o Passo 4 decrescer $\hat{\alpha}$.

2(c): Se $v(P_{\hat{\alpha},\hat{\mu}}) \geq 0$ e $(b - A\hat{x}) \geq 0$, $\hat{x}$ é uma nova solução incumbente para P e $c^T\hat{x}$, um novo limitante inferior para $v(D_S)$. Assim, termine as iterações duais tomando $ll \leftarrow c^T\hat{x}$ e vá para o Passo 5 aumentar $\hat{\alpha}$.

2(D): Se $v(P_{\hat{\alpha},\hat{\mu}}) \geq 0$ e $(b - A\hat{x}) \ngeq 0$, continue as iterações duais: vá para o Passo 3 modificar $\hat{\mu}$.

PASSO 3: *(Modificando μ)* Se $k \geq klim$, termine as iterações duais indo para o Passo 5 aumentar $\hat{\alpha}$. Caso contrário:

3(A): Atualize o subgradiente suavizado e o multiplicador *surrogate* como realizado em Camerini *et al.* [65]:

$$\beta = \begin{cases} -\gamma \frac{\hat{G}(b-A\hat{x})}{\|\hat{G}\|^2}, & \text{se } \hat{G}(b - A\hat{x}) < 0; \\ 0, & \text{caso contrário .} \end{cases}$$

$$\text{onde, caso } \hat{G}(b - A\hat{x}) < 0,\ \gamma \leftarrow -\frac{\|\hat{G}\|\|b - A\hat{x}\|}{\hat{G}(b - A\hat{x})}$$

$$t \leftarrow \frac{\mu^T(b - A\hat{x}) - \hat{z}}{\|\hat{G}\|^2} \quad \text{quando } k > 1$$

$$\hat{G} \leftarrow (b - A\hat{x}) + \beta\hat{G}$$

$$\hat{\mu} \leftarrow \hat{\mu} - t\hat{G}$$

3(B): Projete o novo $\hat{\mu}$ sobre $\{\mu \geq 0\}$ fazendo $\hat{\mu}_j \leftarrow \max\{0, \hat{\mu}_j\}$ para todo j. Então, faça $k \leftarrow k + 1$ e retorne ao Passo 2 para continuar a resolução do dual $D_{\hat{\alpha},L}$.

PASSO 4: *(Decrescendo α)* Até o momento, temos $v(D_S) < \hat{\alpha}$. Se $|ul - ll| < \epsilon$, pare; $v(D_S) \cong \hat{\alpha}$. Caso contrário, aplique um "processo de decrescimento de α" para escolher um novo $\hat{\alpha} \in (ll, ul)$ e retorne ao Passo 1.

PASSO 5: *(Aumentando α)* Até o momento, temos $\hat{\alpha} \leq v(D_S)$. Se $|ul - ll| < \epsilon$, pare; $v(D_S) \cong \hat{\alpha}$. Caso contrário, escolha $\hat{\alpha} \leftarrow ul - \theta(ul - ll)$, $\hat{x} \in S(\hat{\alpha})$ e retorne ao Passo 1.

O fluxograma 5.3 sintetiza o método de Sarin *et al.*.

Importantes orientações sobre o método devem ser comentadas. O Passo 2 inicializa-se com algum $\hat{x} \in S(\hat{\alpha})$ e algum $\hat{\mu} \geq 0$. Enquanto for obtido $\hat{\mu}^T(b - A\hat{x}) \geq 0$ para esse $\hat{x}$, continua-se atualizando $\hat{\mu}$ no Passo 3 até que $\hat{\mu}^T(b - A\hat{x}) < 0$. Em seguida, aplica-se, caso necessário, qualquer método enumerativo que resolva $P_{\hat{\alpha},\hat{\mu}}$, para $\hat{\alpha}$ e $\hat{\mu}$ fixos, de maneira que $\hat{x}$ vai sendo modificado nos estágios da árvore enumerativa do método. Se em algum momento for obtido $\hat{\mu}^T(b - A\hat{x}) \geq 0$, então o algoritmo recai automaticamente para o Passo 3 para modificação de $\hat{\mu}$ (Passo 2(A)). Posteriormente, é recomeçado o Passo 2, resolvendo novamente $P_{\hat{\alpha},\hat{\mu}}$ atual, para um novo $\hat{\mu}$ fixo ($\hat{\alpha}$ não é modificado durante todo o processo de resolução do dual $D_{\alpha,L}$ nos Passos 2 e 3). Este ciclo repete-se até que em nenhum estágio da árvore enumerativa seja encontrado $\hat{\mu}^T(b - A\hat{x}) \geq 0$ (Passo 2(B)). Se, de fato, não o for encontrado, então o método procede até resolver completamente $P_{\hat{\alpha},\hat{\mu}}$. Daí, é verificada qual das condições do algoritmo de Sarin *et al.* são satisfeitas, 2(B), 2(C) ou 2(D), procedendo em seguida, como descrito a seguir (e, implicitamente, no algoritmo):

- Se 2(B) é satisfeito, é porque foi obtido $v(P_{\hat{\alpha},\hat{\mu}}) < 0$ implicando em $v(D_{\alpha,L}) \to -\infty$ (veja demonstração do Corolário 5.3). Daí, pelo Corolário 5.4, $v(D_S) < \hat{\alpha}$ e é possível reduzir ainda mais o valor de $\hat{\alpha}$. Note que, desde que $\hat{\alpha}$ atual é um limitante superior para $v(D_S)$, ul é atualizado para $\hat{\alpha}$ antes mesmo de atualizar o próprio $\hat{\alpha}$.

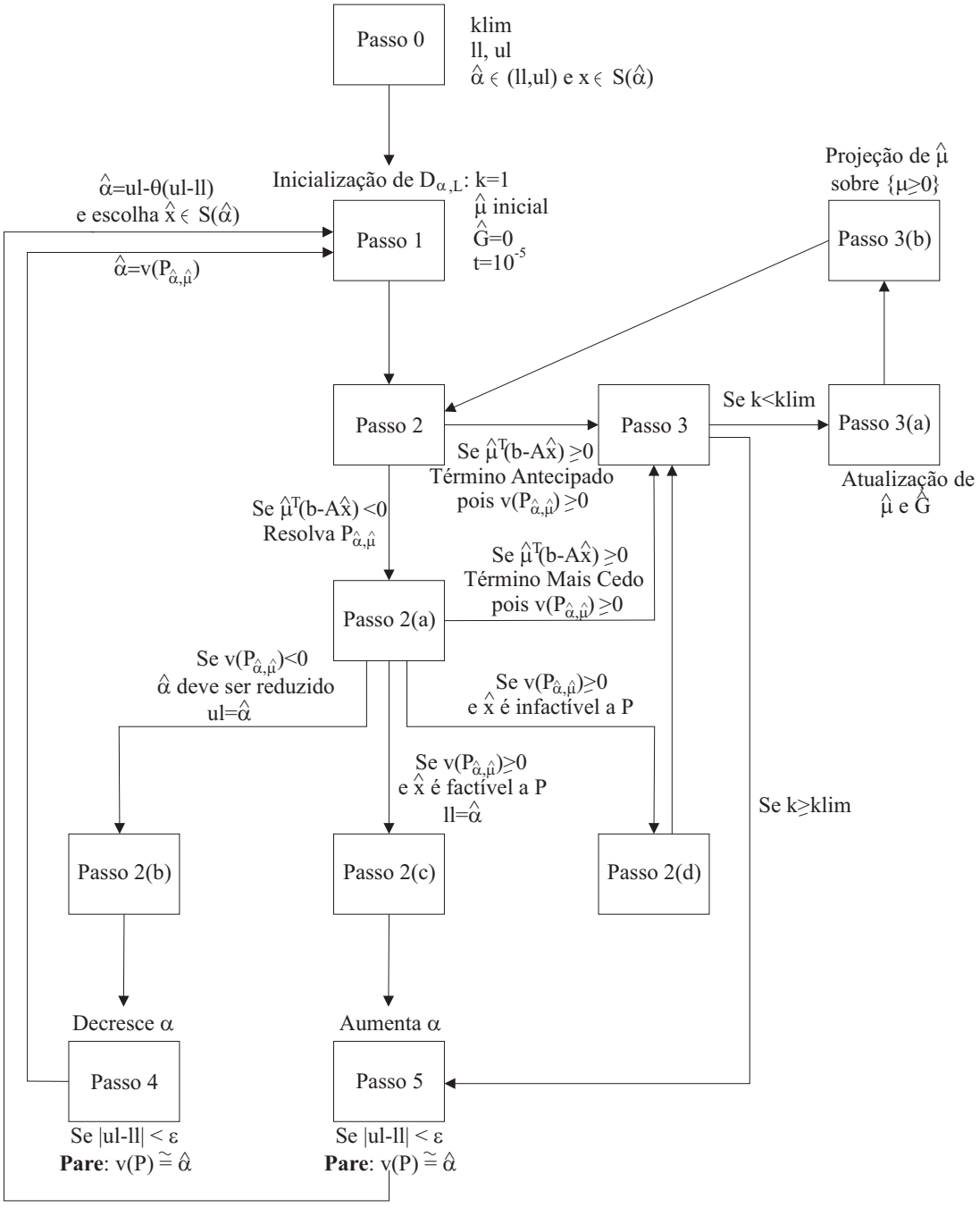


Figura 5.3: Fluxograma do Algoritmo de Sarin *et al.*

- Se 2(C) é satisfeito, é porque foi obtido $v(P_{\hat{\alpha},\hat{\mu}}) \geq 0$ e $(b - A\hat{x}) \geq 0$. Ou seja, também foi obtido $\hat{x}$ factível a P, isto é, $c^T\hat{x} \leq v(P)$. Pela definição de $S(\hat{\alpha})$, $\hat{\alpha} \leq c^T\hat{x}$ e, pelo Teorema 4.2, $v(P) \leq v(D_S)$. Assim, considerando tais conceitos, obtém-se:

$$\hat{\alpha} \leq c^T\hat{x} \leq v(P) \leq v(D_S),$$

ou seja, $\hat{\alpha}$ é limitante inferior para $v(D_S)$. O algoritmo, portanto, atualiza o novo limitante inferior $ll = \hat{\alpha}$ e segue para o Passo 5, para aumentar o valor de $\hat{\alpha}$.

- Se 2(D) é satisfeito, é porque foi obtido $v(P_{\hat{\alpha},\hat{\mu}}) \geq 0$ e $(b - A\hat{x}) \ngeq 0$. Desde que $\hat{x}$ é infactível a P, nada se pode afirmar; e então o algoritmo concentra-se na informação de que $v(P_{\hat{\alpha},\hat{\mu}}) \geq 0$ para melhorar o valor de $\hat{\mu}$ no Passo 3. Note que, conseqüentemente, um novo ciclo de resolução é iniciado pelo Passo 2.

O Processo de Modificação de μ

No Passo 3, μ é atualizado para um valor melhor na tentativa de se resolver $D_{\alpha,L}$. Como este problema é o dual lagrangiano de P_α, qualquer procedimento de busca de multiplicadores duais lagrangianos pode ser empregado. Nesta etapa do algoritmo, Sarin *et al.* [5] implementaram um método de busca subgradiente especializado (apresentado em Camerini *et al.* [65]), cuja idéia é utilizar uma soma ponderada de todas as direções subgradiente anteriores. Observe que, no Passo 3(A), β representa tal soma ponderada, t calcula o tamanho do passo a ser dado na direção subgradiente e que é dado pela diferença entre um limitante superior para P, $\mu^T(b-A\hat{x})$, e um limitante inferior dado por uma heurística, $\hat{z}$ (discutido posteriormente). Observe ainda que, enquanto $v(D_{\alpha,L}) = 0$, é estipulada uma quantidade máxima de vezes, $klim$, em que o algoritmo pode, de fato, atualizar μ; assim ganha-se tempo computacional para novas atualizações de α.

Término Antecipado de $P_{\alpha,\mu}$

Suponha que $\hat{x}$ seja a solução atual de $P_{\alpha,\mu}$. Se para algum $\hat{\mu}$ for obtido $\hat{\mu}^T(b - A\hat{x}) \geq 0$, então nem será necessário calcular $v(P_{\alpha,\mu})$, pois como $P_{\alpha,\mu}$ é um problema de maximização e sua região de factibilidade não se modifica durante o "processo de modificação de μ" (pois α é fixo), então a solução ótima também fornecerá $v(P_{\alpha,\mu}) \geq 0$. Daí, como discutido anteriormente, deve-se aplicar um algoritmo de otimização do subgradiente que gera um novo vetor multiplicador $\hat{\mu}$ para tentar melhorar o valor de $v(P_{\alpha,\mu})$. Entretanto, caso não exista direção de decrescimento que aponta para onde a função lagrangiana decresce, sempre será obtido $v(D_{\alpha,L}) = 0$, isto é, o dual assumirá o pior valor possível, e a solução ótima de $P_{\alpha,\mu}$ (que nem será necessário calcular) não será solução ótima de P_α (para complementar esta discussão, veja "definição dos parâmetros de entrada" ao final seção). O "término antecipado" de $P_{\alpha,\mu}$ é um teste barato que pode render grande ganho de tempo computacional e que fornece a direção $(b - A\hat{x})$ como novo vetor subgradiente.

Término Mais Cedo de $P_{\alpha,\mu}$

Baseando-se na continuação da idéia acima, caso no teste inicial do Passo 2, $\hat{\mu}^T(b-A\hat{x}) < 0$, então não existirão garantias de que $v(P_{\alpha,\mu}) \geq 0$. Logo, será necessário aplicar um método enumerativo para resolver $P_{\alpha,\mu}$ e verificar se, ainda assim, tal desigualdade é satisfeita para alguma solução incumbente $\hat{x}$ no desenvolvimento do método. Se em algum estágio da árvore enumerativa encontra-se $v(P_{\alpha,\mu}) \geq 0$, não será necessário realizar enumeração completa e resolver $P_{\alpha,\mu}$, pois, mais uma vez, desde que $P_{\alpha,\mu}$ é um problema de maximização, então a solução ótima também fornecerá $v(P_{\alpha,\mu}) \geq 0$; basta aplicar o "término mais cedo" de $P_{\alpha,\mu}$ e prosseguir para o Passo 3 com o novo subgradiente $b - A\hat{x}$ como feito no "término antecipado" descrito acima.

Convergência do Método

A convergência do Algoritmo de Sarin *at al.* está diretamente relacionada às propriedades do procedimento de otimização do subgradiente do Passo 3 e às propriedades do esquema de decrescimento de α do Passo 4 (estas últimas serão discutidas apenas no próximo tópico).

O processo de modificação de μ do Passo 3 requer um parâmetro de entrada $\hat{z}$ ($v(D_{\alpha,L}) < \hat{z}$) no cálculo do tamanho do passo t. Um Teorema provado em Camerini *et al.* [65] garante que a seqüência dos μ's produzidos pelo Passo 3 ou irá finalizar em um ponto $\hat{\mu} \in P_{\hat{z}}$ ou irá convergir para um ponto da fronteira de $P_{\hat{z}}$, onde $P_{\hat{z}}$ denota o politopo de soluções factíveis para $\hat{z} \geq v(P_{\alpha,\mu})$, para todo μ. Quando $v(D_{\alpha,L}) = -\infty$, qualquer escolha de $\hat{z} < 0$ resulta em convergência finita do Passo de modificação de μ baseada no subgradiente. Quando $v(D_{\alpha,L}) = 0$, uma regra de parada, que permite apenas um número finito de iterações ($klim$), impede que o método entre em *looping* infinito [5].

Teorema 5.7. *Para o método de Sarin et al.:*

(i) *Se o algoritmo pára quando a solução y para um problema surrogate, P^{μ}, torna-se factível a P (Passo 2(C)), então y resolve P.*

(ii) *Se klim é suficientemente grande, o conjunto S é finito e se o algoritmo termina quando $|ul - ll| < \epsilon$, então $v(D_S) \in [ul - \epsilon, ul]$. Além disso, se todo c_j é inteiro, então $v(D_S) = ul$.*

Demonstração:

(i) Teorema 4.6 (Condições de Otimalidade *Surrogate*).

(ii) Esta demonstração é baseada na prova realizada em Sarin *et al.*. Pelo Teorema 5.3, o método de otimização do subgradiente para a busca dual lagrangiana é finitamente convergente. Desde que $klim$ seja suficientemente grande, se $v(D_S)$ é menor que α, então nos Passos 2 e 3 do algoritmo acima $v(D_{\alpha,L})$ será diferente de zero. Isto irá decrescer o valor de ul. Caso contrário, se $v(D_{\alpha,L}) = 0$ então $v(D_S)$ deve ser maior ou igual a α. Então um crescimento no valor de ll é realizado. Em ambos os casos, o intervalo de incerteza $[ll, ul]$ contendo $v(D_S)$ vai sendo reduzido a intervalos menores com extremos finitos. Desde que $[ll, ul]$ é reduzido até uma distância arbitrariamente pequena (dada por ϵ), então $v(D_S)$ será igual a ul. ■

O Processo de Decrescimento de α

No método Sarin *et al.*, um elemento indefinido é estipulado no Passo 4: a forma como α é reduzido ao concluir-se que $v(D_{\alpha,L}) = -\infty$, isto é, $v(D_S) < \alpha$. Em seu trabalho, Sarin *et al.* utilizaram quatro variações do algoritmo descrito acima, modificando apenas a forma como a atualização de α foi implementada.

Se o intervalo $[ll, ul]$ é pequeno, o algoritmo termina com valor ótimo para a função objetivo, ul. Caso contrário, deseja-se decrescer α de seu valor atual ul. O método que mostrou render

melhor solução dentre os quatro implementados em [5], denotados por S1, S2, S3 e S4, foi aquele que possui o esquema de decrescimento para α descrito a seguir.

Esquema S2: *(Relaxação Surrogate)* Se $v(D_{\alpha,L}) = -\infty$ e $v(P_{\alpha,\mu}) < 0$ é porque $\{x \in S : \alpha \leq c^T x,\ \mu^T(b - Ax) \geq 0\} = \varnothing$, pois $I_{S(\alpha)}$ possui solução. Mas isto implica que $v(P^\mu) = \sup\{c^T x : x \in S,\ \mu^T(b - Ax) \geq 0\} < \alpha$, pois todo x super-ótimo estará sendo excluído do conjunto de soluções factíveis para P^μ (por fornecer um limitante superior pior: $v(P^\mu) = \sup\{c^T x : x \in S,\ \mu^T(b - Ax) \geq 0\} \leq \alpha \leq c^T x$). De fato, já foi mostrado anteriormente que $v(D_S) < \alpha$, que também implica nesta mesma desigualdade. Assim, o esquema S2 implementa esta idéia, resolvendo uma relaxação *surrogate* P^μ em cada Passo 4 e tomando $\alpha \leftarrow v(P^\mu)$.

Ao comparar esta alternativa do método Sarin *et al.* (com o esquema S2 para decrescer α) com o melhor algoritmo até então implementado na literatura para encontrar $v(D_S)$ para problemas de "receita de capital" (*capital budgeting problems*), desenvolvido em [45] e denotado por "Algoritmo LR", Sarin *et al.* encontraram resultados empíricos superiores em seu trabalho e atestaram, assim, a eficácia do método proposto. Daí, a motivação e justificativa do presente trabalho em aplicar este algoritmo na resolução do PCP do Produtor.

Definição dos Parâmetros de Entrada

O parâmetro $\hat{z}$ é utilizado para computar os tamanhos dos passos, t, durante a busca lagrangiana baseada no subgradiente do Passo 3. Como discutido em "Convergência do Método", $\hat{z}$ é um estimador (superior) de $D_{\alpha,L}$ e qualquer escolha de $\hat{z} < 0$ irá resultar na convergência finita do procedimento de otimização do subgradiente. Os valores testados para $\hat{z}$ no trabalho de Sarin *et al.* [5] (mantendo os outros parâmetros fixados) não influenciaram a eficácia do método. Por isso, espera-se que para o método adaptado por este presente trabalho, para problemas de maximização, o valor escolhido para $\hat{z}$ também não influencie o desenvolvimento do algoritmo Sarin *et al.* em relação ao tempo computacional em uma possível implementação, desde que $\hat{z}$ seja um valor negativo; todas as escolhas provavelmente obterão o mesmo desempenho, com mesmo número de iterações.

No algoritmo também é estipulado um número máximo de iterações duais do método de otimização do subgradiente, $klim$, durante a resolução de $D_{\alpha,L}$, para α fixo. Escolher valores muito grandes ou muito pequenos para $klim$ pode fornecer um algoritmo pobre em desempenho. Se $v(D_{\alpha,L}) = 0$, como é escolhido $\hat{z} < 0$ e desde que é sabido que $v(D_{\alpha,L}) < \hat{z}$, escolher $klim$ muito grande levaria o algoritmo a um desperdício de tentativas, em vão, de mostrar que $v(D_{\alpha,L}) < \hat{z} < 0$ (ou seja $v(D_{\alpha,L}) \neq 0$). De outro modo, caso $v(D_{\alpha,L}) < 0$ e $klim$ seja um valor muito pequeno, existirá o risco de terminar o método de otimização do subgradiente antes do dual lagrangiano ótimo ser encontrado. Após resultados empíricos, em seu trabalho, Sarin *et al.* decidiram por $klim = 15$; entretanto, para outros tipos de problema primais de maximização, devem ser testados outros valores (mantendo os outros fixados) a fim de encontrar o valor ideal para $klim$.

Em cada tipo de problema primal de maximização, diferentes valores para θ devem ser testados para uma escolha mais apropriada (mantendo os outros parâmetros fixados). Entretanto, desde que $\theta \in (0, 1)$, a convergência do método Sarin *et al.* não será influenciada; apenas a velocidade com que ele irá convergir.

Além disso, como definido anteriormente, ϵ é a tolerância máxima permitida para determinar se o intervalo de incerteza $[ll, ul]$, que contém $v(D_S)$, é pequeno o suficiente, isto é, se $ul - ll < \epsilon$ então $v(D_S)$ pode ser aproximado para ul. Se, além disso, os problemas a serem resolvidos pelo método Sarin *et al.* possuírem coeficientes inteiros, então escolher $\epsilon = 1$ será o suficiente para a conclusão de que $v(D_S) = ul$, pois ul e ll também serão inteiros (implicando em $ul - ll$ também ser um valor inteiro).

Capítulo 6

Heurística *Surrogate* para o PCP do Produtor

6.1 Introdução

Dentre os trabalhos que desenvolveram métodos exatos para o PCP do Produtor, Oliveira e Morabito [21] apresentaram um método *branch and bound* em que, a cada nó, utiliza limitantes derivados de relaxações lagrangiana e/ou *surrogate* de uma formulação de otimização linear 0-1. Algoritmos de otimização do subgradiente são utilizados para otimizar estes limitantes, inclusive para o caso *surrogate* que não possui garantias de convergência. Assim, os resultados mais eficientes obtidos em [21] basearam-se no método exato com resolução da relaxação lagrangiana em cada nó da árvore enumerativa. Como o problema dual *surrogate* fornece limitantes duais mais próximos da solução ótima de um problema de otimização inteira, provavelmente o algoritmo de otimização do subgradiente implícito no método exato deve ter divergido em pelo menos alguns dos exemplos teste utilizados por Oliveira e Morabito.

Devido à dificuldade em se trabalhar com as restrições *surrogate*, poucos métodos de resolução para D_S que possuam garantias de convergência para os multiplicadores *surrogate* ótimos foram desenvolvidos. Entretanto, o desenvolvimento destas estratégias pode render resultados muito superiores que métodos de solução envolvendo o dual lagrangiano ou a relaxação linear; principalmente quando P^{μ} não possui a propriedade de integralidade (isto é, os limitantes fornecidos por D_S e D_L não são necessariamente iguais - veja Teoremas 4.7 e 4.8). Mesmo provando a existência desta propriedade para o caso *surrogate* (fato quase sempre não consumado), há interesse no desenvolvimento de técnicas de resolução *surrogate* para possíveis comparações entre métodos a partir dos resultados obtidos em termos do tempo computacional de execução dos programas que implementem estas idéias. Assim como para o PCP do Produtor, espera-se que, mesmo que os limitantes duais obtidos por D_L e D_S sejam iguais, um algoritmo que possua convergência para multiplicadores *surrogate* ótimos tenda a fornecer melhores limitantes duais a cada iteração, acelerando a convergência de tal método.

Na próxima seção serão brevemente discutidas as propriedades de integralidade lagrangiana e *surrogate* para o PCP do Produtor modelado com a formulação de Beasley. O modelo (2.3)-(2.6) foi escolhido devido às experiências anteriores da literatura com o problema de corte genérico

no trabalho de Beasley [8], com a resolução do PCP do Produtor nos trabalhos de Farago e Morabito [15], de Oliveira [4] e de Oliveira e Morabito [21] e à experiência anterior absorvida e descrita no Capítulo 3, quando foram resolvidos exemplos teste por meio da implementação do modelo (2.3)-(2.6) no pacote AMPL/CPLEX.

E seguida, a formulação de Beasley, considerada como o problema primal de otimização inteira, é relacionada à teoria desenvolvida nos Capítulos 4 e 5, sendo apresentados os duais lagrangiano e *surrogate* aplicados ao PCP do Produtor, bem como uma variante do Método de Sarin *et al.* para o mesmo problema. Um exemplo numérico finaliza o capítulo, ilustrando o desenvolvimento da teoria de Sarin *et al.* aplicado ao PCP do Produtor.

6.2 Resolução do PCP do Produtor

6.2.1 As Propriedades de Integralidade Lagrangiana e *Surrogate*

De acordo com a Seção 2.2, a formulação de Beasley para o PCP do Produtor é dada da seguinte forma:

$$\text{maximizar} \quad \sum_{i=1}^{2} \sum_{\{p\in X | p\leq L-l_i\}} \sum_{\{q\in Y | q\leq W-w_i\}} x_{ipq} \tag{6.1}$$

sujeito a:

$$\sum_{i=1}^{2} \sum_{\{p\in X | p\leq L-l_i\}} \sum_{\{q\in Y | q\leq W-w_i\}} a_{ipqrs} x_{ipq} \leq 1$$

$$\text{para todo } r \in X \text{ e } s \in Y \tag{6.2}$$

$$P \leq \sum_{i=1}^{2} \sum_{\{p\in X | p\leq L-l_i\}} \sum_{\{q\in Y | q\leq W-w_i\}} x_{ipq} \leq Q \tag{6.3}$$

$$x_{ipq} \in \{0,1\},\ i=1,2,\ p\in X,\ q\in Y \text{ tais que } p\leq L-l_i, q\leq W-w_i \tag{6.4}$$

Baseando-se na Seção 5.2.3, a formulação a cima será considerada o problema primal P com relaxação *surrogate* (P^{μ}), dual *surrogate* (D_L), relaxação lagrangiana (P_{λ}) e dual lagrangiano (D_L) associados:

$$(P^{\mu}) \qquad \text{maximizar} \quad \sum_{i=1}^{2} \sum_{\{p\in X | p\leq L-l_i\}} \sum_{\{q\in Y | q\leq W-w_i\}} x_{ipq} \tag{6.5}$$

sujeito a:

$$\sum_{r\in X} \sum_{s\in Y} \mu_{rs} (1 - \sum_{i=1}^{2} \sum_{\{p\in X | p\leq L-l_i\}} \sum_{\{q\in Y | q\leq W-w_i\}} a_{ipqrs} x_{ipq}) \geq 0 \tag{6.6}$$

$$x \in S \tag{6.7}$$

onde $S = \{x :\ P \leq \sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} x_{ipq} \leq Q,\ x_{ipq} \in \{0,1\},\ i = 1,2,\ p \in X$ tal que $p \leq L - l_i,\ q \in Y$ tal que $q \leq W - w_i\}$

$$(D_S) \qquad \text{minimizar}_{\mu \geq 0}\{v(P^{\mu})\} \tag{6.8}$$

$$(P_\lambda) \qquad \text{maximizar} \quad \sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} x_{ipq} + \\ + \sum_{r \in X} \sum_{s \in Y} \lambda_{rs} (1 - \sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} a_{ipqrs} x_{ipq}) \tag{6.9}$$

$$\text{sujeito a:} \quad x \in S \tag{6.10}$$

$$(D_L) \qquad \text{minimizar}_{\lambda \geq 0}\{v(P_\lambda)\} \tag{6.11}$$

- A propriedade de integralidade lagrangiana pode ser facilmente verificada. Basta notar que a região de factibilidade de P_λ, $x \in S$, possui apenas duas desigualdades:

$$\sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} x_{ipq} \geq P \quad \text{e} \quad \sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} x_{ipq} \leq Q$$

  Assim, como P e Q são valores inteiros e os coeficientes das variáveis x_{ipq} de ambas desigualdades são iguais a 1, então a região $\sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} x_{ipq}$ cruza os eixos cartesianos em pontos inteiros formando, portanto, um politopo de pontos extremos inteiros (note que as duas regiões não possuem interseções). Logo, pode-se concluir que os limitantes superiores fornecidos por P_λ e $P(\bar{S})$ serão idênticos.

- Já a propriedade de integralidade *surrogate* é mais difícil de ser provada, pois seria necessário mostrar que a interseção entre o politopo formado pelas duas desigualdades acima com a região descrita pela restrição *surrogate* (6.6) também forma um politopo de pontos extremos inteiros. O exemplo presente no final deste capítulo mostrará a dificuldade existente na prova desta propriedade para o PCP do Produtor. O que se pode dizer é que ao utilizar $P = 0$ e Q como um limitante de Barnes (não discutido pelo presente trabalho), Oliveira e Morabito [21] não conseguiram encontrar um contra-exemplo mostrando que P^{μ} anterior não possui a propriedade de integralidade. Também não conseguiram encontrar, ao longo de seus experimentos computacionais, um contra-exemplo cujo limitante gerado pela relaxação *surrogate* fosse melhor do que o limitante gerado pela relaxação lagrangiana (veja a parte 3 do Teorema 4.8 - a negativa da volta da implicação). Mas desde que

o processo de atualização dos multiplicadores *surrogate* foram gerados pelo algoritmo de otimização do subgradiente, não é correto afirmar heuristicamente que P^{μ} provavelmente possui a propriedade de integralidade *surrogate* para o PCP do Produtor; pode-se afirmar apenas que os resultados obtidos por Oliveira e Morabito sugerem sua existência.

6.2.2 O Método de Sarin *et al.* para o PCP do Produtor

Conforme mencionado anteriormente, o motivo de se utilizar o método de Sarin *et al.* na resolução do problema dual *surrogate* é que, ao contrário do algoritmo de otimização do subgradiente aplicado à D_S, tal método possui garantias de convergência para os multiplicadores duais ótimos. Assim, baseando-se na teoria desenvolvida na Seção 5.2.3, o conjunto de inequações $I_{S(\alpha)}$ pode ser definido da seguinte forma para o PCP do Produtor:

$$(I_{S(\alpha)}) \qquad \sum_{r\in X}\sum_{s\in Y}\mu_{rs}\Big(1-\sum_{i=1}^{2}\sum_{\{p\in X|p\le L-l_i\}}\sum_{\{q\in Y|q\le W-w_i\}}a_{ipqrs}x_{ipq}\Big)\le\delta$$
$$\text{para todo } x\in S(\alpha),\ \mu\ge 0,\ \delta<0$$

onde $S(\alpha)=\{x\in S:\alpha\le\sum_{i=1}^{2}\sum_{\{p\in X|p\le L-l_i\}}\sum_{\{q\in Y|q\le W-w_i\}}x_{ipq}\}$; bem como os respectivos problemas relacionados à resolução do dual *surrogate* no algoritmo de Sarin *et al.*:

$$(P_{\alpha}) \qquad \text{maximizar} \quad z\stackrel{\triangle}{=}0 \tag{6.12}$$

sujeito a:

$$\sum_{i=1}^{2}\sum_{\{p\in X|p\le L-l_i\}}\sum_{\{q\in Y|q\le W-w_i\}}a_{ipqrs}x_{ipq}\le 1$$
$$\text{para todo } r\in X \text{ e } s\in Y \tag{6.13}$$

$$x\in S(\alpha) \tag{6.14}$$

$$(P_{\alpha,\mu}) \qquad \text{maximizar} \quad \sum_{r\in X}\sum_{s\in Y}\mu_{rs}\Big(1-\sum_{i=1}^{2}\sum_{\{p\in X|p\le L-l_i\}}\sum_{\{q\in Y|q\le W-w_i\}}a_{ipqrs}x_{ipq}\Big) \tag{6.15}$$

sujeito a:

$$x\in S(\alpha) \tag{6.16}$$

onde, agora, os multiplicadores duais *surrogate* de P^{μ}, μ_{rs}, serão os multiplicadores duais lagrangianos de $P_{\alpha,\mu}$ para o problema primal P_{α}. O dual lagrangiano associado é então definido por:

$$(D_{\alpha,L}) \qquad \text{minimizar}_{\mu\ge 0}\{v(P_{\alpha,\mu})\} \tag{6.17}$$

Método Sarin *et al.* adaptado ao PCP do Produtor

PASSO 0: *(Inicialização)* Gere uma solução factível para o PCP do Produtor; por exemplo, a solução homogênea com todas as caixas arranjadas sob a mesma orientação. Faça Z_{LI} igual ao valor desta solução homogênea e no qual é um limitante inferior para o problema $v(D_S)$. Fixe uma solução dual inicial $\hat{\mu}_{rs}$, para todo $r \in X$ e $s \in Y$, como valores iniciais para os multiplicadores duais *surrogate*. Resolva o problema *surrogate* $P^{\hat{\mu}}$ adaptado ao PCP do produtor, (6.5)-(6.7), obtendo $\tilde{x}$ e tomando o valor ótimo da função objetivo de $P^{\hat{\mu}}$ como um limitante superior para $v(D_S)$ e denote-o por Z_{LS}. Se

$$\sum_{r\in X}\sum_{s\in Y}\hat{\mu}_{rs}(1-\sum_{i=1}^{2}\sum_{\{p\in X|p\leq L-l_i\}}\sum_{\{q\in Y|q\leq W-w_i\}}a_{ipqrs}\tilde{x}_{ipq})\geq 0,$$

e

$$(1-\sum_{i=1}^{2}\sum_{\{p\in X|p\leq L-l_i\}}\sum_{\{q\in Y|q\leq W-w_i\}}a_{ipqrs}\tilde{x}_{ipq})\geq 0, \quad \forall\ r\in X,\ s\in Y,$$

então páre; $\tilde{x}$ é a solução ótima para (6.1)-(6.4) e $v(P) = Z_{LS}$, pois o Teorema 4.6 é satisfeito.

Caso contrário, tome uma tolerância de aproximação ϵ e escolha um número máximo de iterações duais K_{MAX}, um valor $\hat{z}$ negativo e um fator $\theta \in (0,1)$. Faça $\hat{\alpha} \leftarrow Z_{LS} - \theta(Z_{LS} - Z_{LI})$ e tome qualquer $\hat{x}$ que satisfaça:

$$\lceil \max\{P, \hat{\alpha}\} \rceil \leq \sum_{i=1}^{2}\sum_{\{p\in X|p\leq L-l_i\}}\sum_{\{q\in Y|q\leq W-w_i\}}\hat{x}_{ipq} \leq Q,$$

onde $\hat{x}_{ipq} \in \{0,1\}, i = 1,2,\ p \in X$ tal que $p \leq L - l_i,\ q \in Y$ tal que $q \leq W - w_i$ e onde o vetor solução $\hat{x}$ possui coordenadas $\hat{x}_{ipq}$.

PASSO 1: *(Inicialização de $D_{\alpha,L}$)* Inicialize o contador de iterações duais $k \leftarrow 1$ e considere $\hat{\mu}_{rs}$ atual, como valores iniciais para os multiplicadores duais lagrangianos, para todo $r \in X$ e $s \in Y$; onde $\hat{\mu}_{rs}$ denotam todas as coordenadas do multiplicador lagrangiano $\hat{\mu}$. Tome $\hat{G} = 0$ como vetor subgradiente suavizado inicial, $t = 10^{-5}$ e considere $\hat{x}$ atual.

PASSO 2: *(Resolução de $P_{\alpha,\mu}$)* Se

$$\sum_{r\in X}\sum_{s\in Y}\hat{\mu}_{rs}(1-\sum_{i=1}^{2}\sum_{\{p\in X|p\leq L-l_i\}}\sum_{\{q\in Y|q\leq W-w_i\}}a_{ipqrs}\hat{x}_{ipq})\geq 0,$$

realize o "término antecipado" deste Passo, pois, implicitamente, está sendo satisfeito $v(P_{\hat{\alpha},\hat{\mu}}) \geq 0$; vá para o Passo 3 modificar $\hat{\mu}$. Caso contrário, resolva por inspeção o problema $(P_{\hat{\alpha},\hat{\mu}})$, tal que $\hat{x}$ é a solução ótima de $(P_{\hat{\alpha},\hat{\mu}})$.

2(B): Se

$$\sum_{r\in X}\sum_{s\in Y}\hat{\mu}_{rs}(1-\sum_{i=1}^{2}\sum_{\{p\in X|p\leq L-l_i\}}\sum_{\{q\in Y|q\leq W-w_i\}}a_{ipqrs}\hat{x}_{ipq}) \stackrel{\triangle}{=} v(P_{\hat{\alpha},\hat{\mu}}) < 0,$$

termine as iterações duais, atualizando $Z_{LS} \leftarrow \hat{\alpha}$, e vá para o Passo 4 decrescer $\hat{\alpha}$.

2(C): Se

$$v(P_{\hat{\alpha},\hat{\mu}}) \geq 0$$

e

$$(1 - \sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} a_{ipqrs} \hat{x}_{ipq}) \geq 0, \quad \forall \ r \in X, \ s \in Y,$$

então $\hat{x}$ é uma nova solução incumbente para (6.1)-(6.4) e

$$\sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} \hat{x}_{ipq},$$

um novo limitante inferior para $v(D_S)$. Assim, termine as iterações duais tomando $Z_{LI} \leftarrow \sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} \hat{x}_{ipq}$ e vá para o Passo 5 aumentar $\hat{\alpha}$.

2(D): Se

$$v(P_{\hat{\alpha},\hat{\mu}}) \geq 0$$

e

$$(1 - \sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} a_{ipqrs} \hat{x}_{ipq}) \ngeq 0, \quad \forall \ r \in X, \ s \in Y,$$

continue as iterações duais: vá para o Passo 3 modificar $\hat{\mu}$.

PASSO 3: *(Modificando μ)* Se $k \geq K_{MAX}$, termine as iterações duais indo para o Passo 5 aumentar $\hat{\alpha}$. Caso contrário:

3(A): Atualize o subgradiente suavizado e o multiplicador *surrogate*:

$$G_{rs} = 1 - \sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} a_{ipqrs} \hat{x}_{ipq}$$

$$\gamma \leftarrow -\frac{\|\hat{G}\| \|G\|}{\hat{G}^T G} \text{ quando } \hat{G}^T G < 0$$

$$\beta = \begin{cases} -\gamma \frac{\hat{G}^T G}{\|\hat{G}\|^2}, & \text{se } \hat{G}^T G < 0; \\ 0, & \text{caso contrário .} \end{cases}$$

$$t \leftarrow \frac{\sum_{r \in X} \sum_{s \in Y} \hat{\mu}_{rs} G_{rs} - \hat{z}}{\|\hat{G}\|^2} \quad \text{quando } k > 1$$

$$\hat{G} \leftarrow G + \beta \hat{G}$$
$$\hat{\mu} \leftarrow \hat{\mu} - t\hat{G}$$

para todo $r \in X$ e $s \in Y$ e onde G é um vetor de $|X||Y|$ coordenadas com respectivos valores G_{rs}.

3(B): Projete o novo $\hat{\mu}$ sobre $\{\mu \geq 0\}$ fazendo $\hat{\mu}_{rs} \leftarrow \max\{0, \hat{\mu}_{rs}\}$ para todo $r \in X$ e todo $s \in Y$. Então, faça $k \leftarrow k+1$ e retorne ao Passo 2 para continuar a resolução do dual $D_{\hat{\alpha},L}$.

PASSO 4: *(Decrescendo α)* Até o momento, tem-se $v(D_S) < \hat{\alpha}$. Se $|Z_{LS} - Z_{LI}| < \epsilon$, pare; $v(D_S) \cong Z_{LS}$. Caso contrário, para escolher um novo $\hat{\alpha} \in (Z_{LS}, Z_{LI})$, resolva $P^{\hat{\mu}}$ com os multiplicadores duais atuais, agora *surrogate*, faça $\hat{\alpha} = v(P^{\hat{\mu}})$ e retorne ao Passo 1.

PASSO 5: *(Aumentando α)* Até o momento, tem-se $\hat{\alpha} \leq v(D_S)$. Se $|Z_{LS} - Z_{LI}| < \epsilon$, pare; $v(D_S) \cong Z_{LS}$. Caso contrário, escolha $\hat{\alpha} \leftarrow Z_{LS} - \theta(Z_{LS} - Z_{LI})$, $\hat{x} \in S(\hat{\alpha})$ e retorne ao Passo 1.

Resolução por Inspeção de P^{μ}

No início da execução do algoritmo anterior, é calculado o valor da solução ótima de P^{μ} como limitante superior inicial para $v(D_S)$. Será apresentada a resolução da relaxação *surrogate* aplicada ao PCP do Produtor, como desenvolvida em Oliveira e Morabito [21].

Dados μ_{rs}, o modelo (6.5)-(6.7) resulta em um problema 0-1 com coeficientes reais fácil de se resolver por inspeção, pois os coeficientes da função objetivo são todos iguais a 1. Note que tal modelo pode ser reescrito por:

$$\text{maximizar} \quad \sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} x_{ipq} \tag{6.18}$$

sujeito a:

$$\sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} V_{ipq} x_{ipq} \leq \sum_{r \in X} \sum_{s \in Y} \mu_{rs} \tag{6.19}$$

$$P \leq \sum_{i=1}^{2} \sum_{\{p \in X | p \leq L - l_i\}} \sum_{\{q \in Y | q \leq W - w_i\}} x_{ipq} \leq Q \tag{6.20}$$

$$x_{ipq} \in \{0,1\},\ i = 1,2,\ p \in X, q \in Y \text{ tais que } p \leq L - l_i,\ q \leq W - w_i \tag{6.21}$$

onde $V_{ipq} = \sum_{r \in X} \sum_{s \in Y} \mu_{rs} a_{ipqrs}$. Assim, sendo n o número de variáveis x_{ipq}:

- Se $n \geq Q$, então escolha as Q variáveis com menores valores V_{ipq}, não ultrapassando o valor de $\sum_{r \in X} \sum_{s \in Y} \mu_{rs}$, e fixe-as em 1. Se ultrapassar este limite, escolha o maior número possível de variáveis com menores valores V_{ipq} não ultrapassando o limite, e fixe-as em 1.

- Se $P \leq n \leq Q$, então escolha as n variáveis com menores valores V_{ipq} não ultrapassando o valor de $\sum_{r \in X} \sum_{s \in Y} \mu_{rs}$, e fixe-as em 1. Se ultrapassar este limite, escolha o maior número possível de variáveis com menores valores V_{ipq} não ultrapassando o limite, e fixe-as em 1.

- Se $n < P$, então escolha as P variáveis com menores valores V_{ipq}, não ultrapassando o valor de $\sum_{r \in X} \sum_{s \in Y} \mu_{rs}$, e fixe-as em 1. Se ultrapassar este limite o problema é infactível (oque não ocorrerá caso P seja definido como um limitante inferior para o problema).

As demais variáveis são fixadas em 0.

Resolução por Inspeção de $P_{\alpha,\mu}$

A resolução por inspeção de $P_{\alpha,\mu}$ é desenvolvida por este trabalho para utilização no método de Sarin *et al.* aplicado ao PCP do Produtor. Assim como no caso *surrogate* acima, o modelo (6.15)-(6.16) pode ser reescrito por:

$$\text{maximizar} \quad \sum_{i=1}^{2} \sum_{\{p\in X | p\leq L-l_i\}} \sum_{\{q\in Y | q\leq W-w_i\}} V'_{ipq} x_{ipq} + \sum_{r\in X} \sum_{s\in Y} \mu_{rs} \tag{6.22}$$

sujeito a:

$$\lceil \max\{P,\alpha\} \rceil \leq \sum_{i=1}^{2} \sum_{\{p\in X | p\leq L-l_i\}} \sum_{\{q\in Y | q\leq W-w_i\}} x_{ipq} \leq Q \tag{6.23}$$

$$x_{ipq} \in \{0,1\},\ i=1,2,\ p\in X, q\in Y \text{ tais que } p\leq L-l_i,\ q\leq W-w_i \tag{6.24}$$

onde $V'_{ipq} = -\sum_{r\in X}\sum_{s\in Y} \mu_{rs} a_{ipqrs}$. Desde que V'_{ipq} é sempre não positivo, quanto mais variáveis x_{ipq} forem iguais a zero, menos será penalizada a função objetivo (6.22). Dessa forma, dados os multiplicadores μ_{rs}, agora lagrangianos, ordene as variáveis x_{ipq} conforme a ordenação crescente dos elementos V'_{ipq}, fixando em 1 as primeiras k variáveis x_{ipq}, com $k = \lceil \max\{P,\alpha\} \rceil$, e em 0 as demais variáveis.

Caso existam l valores $V'_{ipq} = 0$, se $l \leq k$, as l respectivas variáveis x_{ipq} serão as de menor valor na ordenação descrita anteriormente, e assim, seus valores serão obrigatoriamente iguais a 1. Se $l > k$, então fixe em 1 quaisquer k variáveis que possuem valores $V'_{ipq} = 0$ e fixe em 0 as demais variáveis.

Considerações sobre o Método de Sarin *et al.* Aplicado ao PCP do Produtor

Na heurística *surrogate* para o PCP do Produtor, descrita como o método adaptado de Sarin *et al.* para o PCP do Produtor, modificações tiveram de ser efetuadas quando tal heurística foi aplicada, pois alguns erros foram encontrados na adaptação direta do algoritmo de [5]. Além disso, outras modificações foram realizadas na tentativa de tornar mais eficaz a heurística *surrogate* para a resolução do PCP do Produtor. Estas adaptações são descritas a seguir e podem ser melhor visualizadas no exemplo numérico ao final deste capítulo.

Primeiramente, observe que quando o método executa o Passo 4 (ou o Passo 5), como o critério de parada é $|Z_{LS} - Z_{LI}| < \epsilon$, então, de acordo com o algoritmo genérico da página 63 e 64, o algoritmo deveria atualizar $v(D_S) \cong \hat{\alpha}$. Entretanto, um critério de atualização de $\hat{\alpha}$ deveria ser anteriormente estipulado, pois, caso contrário, $\hat{\alpha}$ possui um valor pior que o atual limitante superior (ou inferior) proveniente do Passo 2(B) (ou 2(C)). Como assume-se neste mesmo algoritmo genérico, $v(D_S) \cong \hat{\alpha}$, a heurística *surrogate* para o PCP do Produtor renderia um valor mais pobre do que poderia render, uma vez que é sabido $v(D_S) \in [Z_{LS} - \epsilon, Z_{LS}]$ (Veja Teorema 5.7). No caso do PCP do Produtor, desde que os coeficientes das variáveis da função objetivo são inteiros, nos Passos 4 e 5, o critério de parada na heurística *surrogate* implica em

$v(D_S) \cong Z_{LS}$. Entretanto, em outros problemas, a nova atualização de α deveria ser efetuada antes da aproximação $v(D_S) \cong \alpha$.

Outra adaptação deve ser realizada ao se resolver $D_{\alpha,L}$. Se o método obtiver no Passo 3, $k = K_{MAX}$, e não tiver atualizado os limitantes superior e/ou inferior nos Passos 2(B) e/ou 2(C), então $\hat{\alpha}$ não será atualizado nos Passos 4 e/ou 5. Daí, haverá a possibilidade do algoritmo tomar o mesmo elemento $\hat{x}$ no novo Passo 1 e repetir os mesmo Passos novamente. Assim, deve-se adaptar o algoritmo considerando uma das seguintes possibilidades: (i) Garantir que o novo $\hat{x}$ seja diferente do $\hat{x}$ previamente considerado no Passo 1 anterior, ou (ii) Forçar uma atualização de $\hat{\alpha}$.

No Passo 0, um teste prévio é realizado sobre a solução fornecida pela resolução de $P^{\hat{\mu}}$, para verificação do Teorema 4.6. Tal teste exclui a não necessidade do algoritmo continuar a procurar pela solução ótima, pois já encontra-se nela, para algum μ inicial: $\tilde{x}$. Observe que, caso tal verificação não fosse efetuada, haveria a possibilidade de considerar $\hat{x} = \tilde{x}$ ao final do Passo 0. Daí, o Passo 2 do algoritmo realizaria o "término antecipado", indo diretamente para o Passo 3 resolver, sem necessidade, $D_{\alpha,L}$.

O Passo 2(A) do algoritmo original, descrito no Capítulo 5, não foi computado no algoritmo adaptado, pois a resolução por inspeção de $(P_{\hat{\alpha},\hat{\mu}})$ é realizada diretamente, sem necessitar de uma árvore enumerativa para fornecer solução ótima; apenas uma solução, a solução ótima $\hat{x}$ (e não a solução incumbente a algum estágio da árvore enumerativa), é encontrada neste tipo de resolução.

Também é aplicado diretamente o esquema S2 no Passo 4 por ter se mostrado o esquema mais efetivo dentre àqueles propostos por Sarin *et al.* [5]. O valor de θ deve ser testado para vários valores entre 0 e 1 a fim de encontrar-se o que melhor se adequa ao PCP do Produtor.

Desde que o vetor de custos do problema (6.1)-(6.4) são todos inteiros (todos iguais a 1), então os limitantes superiores e inferiores obtidos no algoritmo para $v(D_S)$, Z_{LS} e Z_{LI}, também serão inteiros e, conseqüentemente, $|Z_{LS} - Z_{LI}|$ também será um valor inteiro. Portanto, considerar o intervalo de tolerância entre eles ser menor que 1 ($\epsilon = 1$) é o suficiente para garantir que $v(D_S) = Z_{LS}$, uma vez que o critério de parada do algoritmo será $|Z_{LS} - Z_{LI}| < 1$, implicando em $|Z_{LS} - Z_{LI}| = 0$. Entretanto, desde que o tempo computacional de uma possível implementação seja inviável, outros valores para ϵ devem ser testados.

6.3 Exemplo Numérico

Considere o exemplo trivial desenvolvido na Seção 3.3.2. Inicialmente, será mostrada a propriedade de integralidade lagrangiana para este caso em particular e a dificuldade em verificar a propriedade de integralidade *surrogate* para o mesmo caso. Em seguida, será desenvolvida a aplicação da teoria de Sarin *et al.* para o PCP do Produtor descrito anteriormente neste capítulo.

6.3.1 As Propriedades de Integralidade Lagrangiana e *Surrogate*

De acordo com a Seção 3.3.2, quando $(L, W, l, w) = (5, 4, 3, 2)$, o modelo de Beasley para o PCP do Produtor, considerado como o problema primal P, é dado por:

$$\text{maximizar} \quad x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \tag{6.25}$$

sujeito a:

$$x_{100} + x_{200} \leq 1 \tag{6.26}$$
$$x_{102} + x_{200} \leq 1 \tag{6.27}$$
$$x_{100} + x_{120} + x_{220} \leq 1 \tag{6.28}$$
$$x_{102} + x_{122} + x_{220} \leq 1 \tag{6.29}$$
$$x_{120} + x_{220} + x_{230} \leq 1 \tag{6.30}$$
$$x_{122} + x_{220} + x_{230} \leq 1 \tag{6.31}$$

$$x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \leq 3 \tag{6.32}$$
$$x_{ipq} \in \{0, 1\} \tag{6.33}$$

- Então, pela Seção 6.2.1, a relaxação lagrangiana associada será:

$$\begin{aligned} \text{maximizar} \quad & x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} + \\ & \lambda_{00}(1 - x_{100} - x_{200}) + \lambda_{02}(1 - x_{102} - x_{200}) + \\ & \lambda_{20}(1 - x_{100} - x_{120} - x_{220}) + \lambda_{22}(1 - x_{102} - x_{122} - x_{220}) + \\ & \lambda_{30}(1 - x_{120} - x_{220} - x_{230}) + \lambda_{32}(1 - x_{122} - x_{220} - x_{230}) \end{aligned} \tag{6.34}$$

  sujeito a:

$$x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \leq 3 \tag{6.35}$$
$$x_{ipq} \in \{0, 1\} \tag{6.36}$$

  onde $\lambda^T = (\lambda_{00}, \lambda_{02}, \lambda_{20}, \lambda_{22}, \lambda_{30}, \lambda_{32})$ é o vetor multiplicador lagrangiano associado ao conjunto de restrições. A propriedade de integralidade lagrangiana é satisfeita, pois a restrição $x_{100}+x_{102}+x_{120}+x_{122}+x_{200}+x_{220}+x_{230} \leq 3$, que na verdade toma $P = 0$ e $Q = 3$, forma uma região que, com respeito a Q, corta os seis eixos correspondentes às variáveis acima nos pontos inteiros $x_{ipq} = 3$, $i = 1, 2$, $p = 0, 2, 3$ e $q = 0, 2$ e que, com respeito a P corta os eixos apenas na origem. Desde que valores negativos não podem ser considerados, a região forma um politopo de pontos extremos inteiros $(0, 0, 0, 0, 0, 0)$, $(3, 0, 0, 0, 0, 0)$, $(0, 3, 0, 0, 0, 0)$, $(0, 0, 3, 0, 0, 0)$, $(0, 0, 0, 3, 0, 0)$, $(0, 0, 0, 0, 3, 0)$ e $(0, 0, 0, 0, 0, 3)$.

- Também pela Seção 6.2.1, a relaxação *surrogate* pode ser escrita da seguinte forma:

$$\text{maximizar} \quad x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \tag{6.37}$$

  sujeito a:

$$\begin{aligned} & \mu_{00}(x_{100} + x_{200}) + \mu_{02}(x_{102} + x_{200}) + \ \mu_{20}(x_{100} + x_{120} + x_{220}) + \\ & \mu_{22}(x_{102} + x_{122} + x_{220}) + \mu_{30}(x_{120} + x_{220} + x_{230}) + \\ & \mu_{32}(x_{122} + x_{220} + x_{230}) \leq \mu_{00} + \mu_{02} + \mu_{20} + \mu_{22} + \mu_{30} + \mu_{32} \end{aligned} \tag{6.38}$$
$$x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \leq 3 \tag{6.39}$$
$$x_{ipq} \in \{0, 1\} \tag{6.40}$$

Reescrevendo o problema acima, ao se colocar as variáveis x_{ipq} em evidência, obtém-se o modelo:

$$(P^{\mu}) \qquad \text{maximizar} \quad x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \tag{6.41}$$

sujeito a:

$$\begin{aligned}
&(\mu_{00} + \mu_{20})x_{100} + (\mu_{02} + \mu_{22})x_{102} + (\mu_{20} + \mu_{30})x_{120} + (\mu_{22} + \mu_{32})x_{122} + \\
&(\mu_{00} + \mu_{02})x_{200} + (\mu_{20} + \mu_{22} + \mu_{30} + \mu_{32})x_{220} + \\
&(\mu_{30} + \mu_{32})x_{230} \leq \mu_{00} + \mu_{02} + \mu_{20} + \mu_{22} + \mu_{30} + \mu_{32} && (6.42)\\
&x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \leq 3 && (6.43)\\
&x_{ipq} \in \{0, 1\} && (6.44)
\end{aligned}$$

onde $\mu^T = (\mu_{00}, \mu_{02}, \mu_{20}, \mu_{22}, \mu_{30}, \mu_{32})$ é o vetor de multiplicadores *surrogate*.

Para provar a propriedade de integralidade *surrogate* para o problema acima deve-se mostrar que a interseção entre as regiões formadas pelas restrições (6.42), (6.43) e $x_{ipq} \in [0, 1]$, forma necessariamente um politopo de pontos inteiros. Lembrando que os valores μ_{rs}, para $r = 0, 2, 3$ e $s = 0, 2$, são dados por valores reais não-negativos, para o conjunto de restrições (6.44) relaxadas, também deseja-se encontrar os pontos extremos

$$x^T = (x_{100}, x_{102}, x_{120}, x_{122}, x_{200}, x_{220}, x_{230})$$

da interseção entre (6.42) e (6.43) que, além de não serem combinação linear de dois pontos que satisfaçam estas duas restrições, satisfaçam necessariamente esta interseção:

$$\begin{aligned}
&(\mu_{00} + \mu_{20})x_{100} + (\mu_{02} + \mu_{22})x_{102} + (\mu_{20} + \mu_{30})x_{120} + (\mu_{22} + \mu_{32})x_{122} + (\mu_{00} + \mu_{02})x_{200} + \\
&(\mu_{20} + \mu_{22} + \mu_{30} + \mu_{32})x_{220} + (\mu_{30} + \mu_{32})x_{230} - (\mu_{00} + \mu_{02} + \mu_{20} + \mu_{22} + \mu_{30} + \mu_{32}) = \\
&= x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} - 3
\end{aligned}$$

Ou seja:

$$\begin{aligned}
&(\mu_{00} + \mu_{20} - 1)x_{100} + (\mu_{02} + \mu_{22} - 1)x_{102} + (\mu_{20} + \mu_{30} - 1)x_{120} + (\mu_{22} + \mu_{32} - 1)x_{122} + \\
&(\mu_{00} + \mu_{02} - 1)x_{200} + (\mu_{20} + \mu_{22} + \mu_{30} + \mu_{32} - 1)x_{220} + (\mu_{30} + \mu_{32} - 1)x_{230} = \\
&= \mu_{00} + \mu_{02} + \mu_{20} + \mu_{22} + \mu_{30} + \mu_{32} - 3
\end{aligned}$$

Portanto, uma das provas a serem feitas para mostrar a propriedade de integralidade *surrogate* para P^{μ} é demonstrar que, para $x_{ipq} \in [0, 1]$, os pontos descritos por $x_{ipq} = 0$ ou 1, para todo $i = 1, 2,\ p = 0, 2, 3$ e $q = 0, 2$, que satisfazem a soma dada na última equação não são combinação linear de quaisquer outros pontos que satisfazem esta mesma a equação, para qualquer vetor multiplicador μ. Note que tal demonstração implicaria que os pontos extremos da região de interseção formada entre (6.42) e (6.43) são pontos inteiros, para qualquer que seja a escolha de μ. Por exemplo, se μ^T atual é dado pelo vetor $(0.43, 1, 1, 1, 1, 1)$, então deve-se mostrar que todo vetor x de componentes binárias satisfazendo:

$$0.43x_{100} + x_{102} + x_{120} + x_{122} + 0.43x_{200} + 3x_{220} + x_{230} = 2.43$$

não pode ser escrito como combinação linear de quaisquer outros pontos que satisfazem esta mesma igualdade e que possuem componentes no intervalo $[0, 1]$. Provar esta afirmação para qualquer que seja μ^T, não é uma tarefa trivial e, até o presente momento não foi encontrada em nenhum trabalho da literatura.

6.3.2 Aplicação do Método de Sarin *et al.* para o PCP do Produtor

Seguindo com o desenvolvimento de Sarin *et al.* aplicado ao PCP do Produtor das Seções 6.2.1 e 6.2.2 e observando que D_L e D_S são os problemas (6.11) e (6.8), respectivamente, o problema *surrogate* aplicado ao PCP do Produtor quando $(L, W, l, w) = (5, 4, 3, 2)$ pode ser escrito por:

$$(P^{\mu}) \qquad \text{maximizar} \quad x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \tag{6.45}$$

sujeito a:

$$\begin{aligned} &\mu_{00}(1 - x_{100} - x_{200}) + \mu_{02}(1 - x_{102} - x_{200}) + \\ &\mu_{20}(1 - x_{100} - x_{120} - x_{220}) + \mu_{22}(1 - x_{102} - x_{122} - x_{220}) + \\ &\mu_{30}(1 - x_{120} - x_{220} - x_{230}) + \mu_{32}(1 - x_{122} - x_{220} - x_{230}) \geq 0 \end{aligned} \tag{6.46}$$

$$x \in S \tag{6.47}$$

com $S = \{x = (x_{100}, x_{102}, x_{120}, x_{122}, x_{200}, x_{220}, x_{230})^T : x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \leq 3,\ x_{ipq} \in \{0, 1\},\ i = 1, 2,\ p = 0, 2, 3,\ q = 0, 2\}$, então:

$$\begin{aligned} (I_{S(\alpha)}) \qquad &\mu_{00}(1 - x_{100} - x_{200}) + \mu_{02}(1 - x_{102} - x_{200}) + \\ &\mu_{20}(1 - x_{100} - x_{120} - x_{220}) + \mu_{22}(1 - x_{102} - x_{122} - x_{220}) + \\ &\mu_{30}(1 - x_{120} - x_{220} - x_{230}) + \mu_{32}(1 - x_{122} - x_{220} - x_{230}) \leq \delta \\ &\qquad\qquad \text{para todo } x \in S(\alpha),\ \mu \geq 0,\ \delta < 0 \end{aligned} \tag{6.48}$$

onde $S(\alpha) = \{x \in S : \alpha \leq x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230}\}$,

$$(P_{\alpha}) \qquad \text{maximizar} \quad z \stackrel{\triangle}{=} 0 \tag{6.49}$$

sujeito a:

$$x_{100} + x_{200} \leq 1 \tag{6.50}$$

$$x_{102} + x_{200} \leq 1 \tag{6.51}$$

$$x_{100} + x_{120} + x_{220} \leq 1 \tag{6.52}$$

$$x_{102} + x_{122} + x_{220} \leq 1 \tag{6.53}$$

$$x_{120} + x_{220} + x_{230} \leq 1 \tag{6.54}$$

$$x_{122} + x_{220} + x_{230} \leq 1 \tag{6.55}$$

$$x \in S(\alpha) \tag{6.56}$$

$$(P_{\alpha,\mu}) \qquad \text{maximizar} \quad \mu_{00}(1 - x_{100} - x_{200}) + \mu_{02}(1 - x_{102} - x_{200}) + \mu_{20}(1 - x_{100} - x_{120} - x_{220}) + \mu_{22}(1 - x_{102} - x_{122} - x_{220}) + \mu_{30}(1 - x_{120} - x_{220} - x_{230}) + \mu_{32}(1 - x_{122} - x_{220} - x_{230}) \tag{6.57}$$

sujeito a:

$$\lceil \max\{\alpha, P\} \rceil \leq x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230}\} \leq Q \tag{6.58}$$

$$x_{ipq} \in \{0, 1\} \tag{6.59}$$

$$(D_{\alpha,L}) \qquad \text{minimizar}_{\mu \geq 0}\{v(P_{\alpha,\mu})\} \tag{6.60}$$

Exemplo do método Sarin *et al.* adaptado ao PCP do Produtor

Passo 0: *(Inicialização)* Considere

$$\tilde{x}^T = (\tilde{x}_{100}, \tilde{x}_{102}, \tilde{x}_{120}, \tilde{x}_{122}, \tilde{x}_{200}, \tilde{x}_{220}, \tilde{x}_{230}) = (1, 0, 0, 0, 0, 0, 0)$$

solução factível a P tal que $Z_{LI} = 1$. Tomando $\hat{\mu}^T = (\hat{\mu}_{00}, \hat{\mu}_{02}, \hat{\mu}_{20}, \hat{\mu}_{22}, \hat{\mu}_{30}, \hat{\mu}_{32}) = (1, 1, 1, 1, 1, 1)$, obtém-se:

$$(P^{\hat{\mu}}) \qquad \text{maximizar} \quad x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230}$$

sujeito a:

$$2x_{100} + 2x_{102} + 2x_{120} + 2x_{122} + 2x_{200} + 4x_{220} + 2x_{230} \leq 6$$

$$x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} \leq 3$$

$$\hat{x}_{ipq} \in \{0, 1\}$$

que possui fácil resolução por inspeção. Considere as variáveis que maximizem o valor da função objetivo do problema anterior respeitando o conjunto de restrições. Ou seja, ordene as variáveis x_{ipq} de acordo com a ordenação crescente de V_{ipq}, por exemplo:

$$x_{100} \leq x_{102} \leq x_{120} \leq x_{122} \leq x_{200} \leq x_{230} \leq x_{220}$$

Em seguida, fixe em 1 as primeiras $Q = 3$ variáveis que não ultrapassem $\sum_{r \in X} \sum_{s \in Y} \hat{\mu}_{rs} = 6$, $\tilde{x}_{100} = \tilde{x}_{102} = \tilde{x}_{120} = 1$, e fixe as demais variáveis em 0. Logo, é obtida solução ótima $\tilde{x}^T = (1, 1, 1, 0, 0, 0, 0)$ com $Z_{LS} = 3$. Como $\tilde{x}$ é infactível a P, a condição (ii) do Teorema 4.6 não é satisfeita.

Seja $\epsilon = 1.1$ e considere, por exemplo, $K_{MAX} = 10$, $\hat{z} = -10$ e $\theta = 0.6$. Então $\hat{\alpha} = 3 - 0.6(3 - 1) = 1,8$. Tome, por exemplo, $\hat{x} = (1, 0, 1, 0, 0, 0, 0)^T$ pertencente a $S(\hat{\alpha})$, isto é, tal que satisfaz:

$$2 = \lceil 1.8 \rceil \leq \hat{x}_{100} + \hat{x}_{102} + \hat{x}_{120} + \hat{x}_{122} + \hat{x}_{200} + \hat{x}_{220} + \hat{x}_{230} \leq 3$$

Observe que, desde que foi escolhido um valor $\epsilon \neq 1$, quando o método parar no Passo 4 ou no Passo 5, Z_{LS} final será apenas uma aproximação para $v(D_S)$; não existirão garantias de

que $v(D_S) = Z_{LS}$. Tal valor ϵ foi escolhido para que o algoritmo convergisse mais rapidamente para um valor $v(D_S)$.

Passo 1: *(Inicialização de $D_{\alpha,L}$)* Inicialize o contador de iterações duais $k \leftarrow 1$ e considere $\hat{\mu}_{rs}$ atual, como valores iniciais para os multiplicadores duais lagrangianos, para todo $r \in X$ e $s \in Y$ ($\hat{\mu}_{rs}$ denotam todas as coordenadas do multiplicador lagrangiano $\hat{\mu}$). Tome $\hat{G} = 0$ como vetor subgradiente suavizado inicial, $t = 10^{-5}$ e considere $\hat{x}$ atual.

Passo 2: *(Resolução de $P_{\alpha,\mu}$)* Como

$$\hat{\mu}_{00}(1 - \hat{x}_{100} - \hat{x}_{200}) + \hat{\mu}_{02}(1 - \hat{x}_{102} - \hat{x}_{200}) + \hat{\mu}_{20}(1 - \hat{x}_{100} - \hat{x}_{120} - \hat{x}_{220}) +$$
$$\hat{\mu}_{22}(1 - \hat{x}_{102} - \hat{x}_{122} - \hat{x}_{220}) + \hat{\mu}_{30}(1 - \hat{x}_{120} - \hat{x}_{220} - \hat{x}_{230}) + \hat{\mu}_{32}(1 - \hat{x}_{122} - \hat{x}_{220} - \hat{x}_{230}) =$$
$$= 0 + 1 - 1 + 1 + 0 + 0 = 1 \geq 0,$$

realize o "término antecipado" e vá para o Passo 3 modificar $\hat{\mu}$.

Passo 3: *(Modificando μ)* Desde que $k < 10$:

3(a): Atualize o subgradiente suavizado e o multiplicador *surrogate*:

$$\begin{aligned} G_{00} &= 1 - \hat{x}_{100} - \hat{x}_{200} = 0 \\ G_{02} &= 1 - \hat{x}_{102} - \hat{x}_{200} = 1 \\ G_{20} &= 1 - \hat{x}_{100} - \hat{x}_{120} - \hat{x}_{220} = -1 \\ G_{22} &= 1 - \hat{x}_{102} - \hat{x}_{122} - \hat{x}_{220} = 1 \\ G_{30} &= 1 - \hat{x}_{120} - \hat{x}_{220} - \hat{x}_{230} = 0 \\ G_{32} &= 1 - \hat{x}_{122} - \hat{x}_{220} - \hat{x}_{230} = 1 \end{aligned}$$

onde $G^T = (0, 1, -1, 1, 0, 1)$. Como $\hat{G} = 0$, então $\hat{G}^T G = 0$ e não há necessidade em se calcular γ, pois, pelo algoritmo, $\beta = 0$. Assim, dado que para $k = 1$, tem-se $t = 10^{-5}$, calcule:

$$\begin{aligned} \hat{G} \leftarrow G + \beta\hat{G} &= G + \overrightarrow{0} = (0, 1, -1, 1, 0, 1)^T \\ \hat{\mu} \leftarrow \hat{\mu} - t\hat{G} &= (1, 1, 1, 1, 1, 1)^T - 10^{-5}(0, 1, -1, 1, 0, 1)^T \\ &= (1, 1 - 10^{-5}, 1 + 10^{-5}, 1 - 10^{-5}, 1, 1 - 10^{-5})^T \end{aligned}$$

3(b): Projete o novo $\hat{\mu}$ sobre $\{\mu \geq 0\}$ fazendo $\hat{\mu}_{rs} \leftarrow \max\{0, \hat{\mu}_{rs}\}$ para todo $r \in X$ e todo $s \in Y$: $\hat{\mu} \leftarrow (1, 1 - 10^{-5}, 1 + 10^{-5}, 1 - 10^{-5}, 1, 1 - 10^{-5})^T$. Então, faça $k \leftarrow 2$ e retorne ao Passo 2 para continuar a resolução do dual $D_{\hat{\alpha},L}$.

Passo 2': *(Resolução de $P_{\alpha,\mu}$)* Como

$$\hat{\mu}_{00}(1 - \hat{x}_{100} - \hat{x}_{200}) + \hat{\mu}_{02}(1 - \hat{x}_{102} - \hat{x}_{200}) + \hat{\mu}_{20}(1 - \hat{x}_{100} - \hat{x}_{120} - \hat{x}_{220}) +$$
$$\hat{\mu}_{22}(1 - \hat{x}_{102} - \hat{x}_{122} - \hat{x}_{220}) + \hat{\mu}_{30}(1 - \hat{x}_{120} - \hat{x}_{220} - \hat{x}_{230}) + \hat{\mu}_{32}(1 - \hat{x}_{122} - \hat{x}_{220} - \hat{x}_{230}) =$$
$$= 0 + (1 - 10^{-5}) - (1 + 10^{-5}) + (1 - 10^{-5}) + 0 + (1 - 10^{-5}) = 2 - 4 \times 10^{-5} \geq 0,$$

realize o "término antecipado" e vá para o Passo 3 modificar μ.

Passo 3': *(Modificando μ)* Desde que $k < 10$:

3(A)': Atualize o subgradiente suavizado e o multiplicador *surrogate*:

$$\begin{aligned}
G_{00} &= 1 - \hat{x}_{100} - \hat{x}_{200} = 0 \\
G_{02} &= 1 - \hat{x}_{102} - \hat{x}_{200} = 1 \\
G_{20} &= 1 - \hat{x}_{100} - \hat{x}_{120} - \hat{x}_{220} = -1 \\
G_{22} &= 1 - \hat{x}_{102} - \hat{x}_{122} - \hat{x}_{220} = 1 \\
G_{30} &= 1 - \hat{x}_{120} - \hat{x}_{220} - \hat{x}_{230} = 0 \\
G_{32} &= 1 - \hat{x}_{122} - \hat{x}_{220} - \hat{x}_{230} = 1
\end{aligned}$$

onde mantém-se $G^T = (0, 1, -1, 1, 0, 1)$. Como $\hat{G}^T G = 3 > 0$, então não há necessidade em se calcular γ, pois, pelo algoritmo, $\beta = 0$. Assim, calcule:

$$\begin{aligned}
t &\leftarrow \frac{(1, 1-10^{-5}, 1+10^{-5}, 1-10^{-5}, 1, 1-10^{-5})(0, 1, -1, 1, 0, 1)^T - (-10)}{\|\hat{G}\|^2} \\
&= \frac{12 - 4 \times 10^{-5}}{3} = 4 - \frac{4}{3}10^{-5}
\end{aligned}$$

$$\begin{aligned}
\hat{G} &\leftarrow G + \beta\hat{G} = (0, 1, -1, 1, 0, 1)^T + \overrightarrow{0} = (0, 1, -1, 1, 0, 1)^T \\
\hat{\mu} &\leftarrow \hat{\mu} - t\hat{G} = (1, 1-10^{-5}, 1+10^{-5}, 1-10^{-5}, 1, 1-10^{-5})(0, 1, -1, 1, 0, 1)^T \\
&\qquad - (4 - \frac{4}{3}10^{-5})(0, 1, -1, 1, 0, 1)^T \\
&\qquad = (1, -3 + \frac{10^{-5}}{3}, 5 - \frac{10^{-5}}{3}, -3 + \frac{10^{-5}}{3}, 1, -3 + \frac{10^{-5}}{3})^T
\end{aligned}$$

3(B)': Projete o novo $\hat{\mu}$ sobre $\{\mu \geq 0\}$ fazendo $\hat{\mu}_{rs} \leftarrow \max\{0, \hat{\mu}_{rs}\}$ para todo $r \in X$ e todo $s \in Y$: $\hat{\mu} \leftarrow (1, 0, 5 - \frac{10^{-5}}{3}, 0, 1, 0)^T$. Então, faça $k \leftarrow 3$ e retorne ao Passo 2 para continuar a resolução do dual $D_{\hat{\alpha},L}$.

PASSO 2": *(Resolução de $P_{\alpha,\mu}$)* Como

$$\begin{aligned}
&\hat{\mu}_{00}(1 - \hat{x}_{100} - \hat{x}_{200}) + \hat{\mu}_{02}(1 - \hat{x}_{102} - \hat{x}_{200}) + \hat{\mu}_{20}(1 - \hat{x}_{100} - \hat{x}_{120} - \hat{x}_{220}) + \\
&\quad \hat{\mu}_{22}(1 - \hat{x}_{102} - \hat{x}_{122} - \hat{x}_{220}) + \hat{\mu}_{30}(1 - \hat{x}_{120} - \hat{x}_{220} - \hat{x}_{230}) + \hat{\mu}_{32}(1 - \hat{x}_{122} - \hat{x}_{220} - \hat{x}_{230}) = \\
&= 0 + 0 + (5 - \frac{10^{-5}}{3})(-1) + 0 + 0 + 0 = -5 + \frac{10^{-5}}{3} < 0,
\end{aligned}$$

resolva por inspeção o problema:

$$\begin{aligned}
(P_{\hat{\alpha},\hat{\mu}}) \qquad \text{maximizar} \quad & \hat{\mu}_{00}(1 - x_{100} - x_{200}) + \hat{\mu}_{02}(1 - x_{102} - x_{200}) + \\
& \hat{\mu}_{20}(1 - x_{100} - x_{120} - x_{220}) + \hat{\mu}_{22}(1 - x_{102} - x_{122} - x_{220}) + \\
& \hat{\mu}_{30}(1 - x_{120} - x_{220} - x_{230}) + \hat{\mu}_{32}(1 - x_{122} - x_{220} - x_{230}) \qquad (6.61)
\end{aligned}$$

sujeito a:

$$\lceil \max\{\hat{\alpha}, P\} \rceil \leq x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230}\} \leq Q \qquad (6.62)$$

$$x_{ipq} \in \{0, 1\} \qquad (6.63)$$

que é dado por:

$$\text{maximizar} \quad (7-\frac{10^{-5}}{3})-(6-\frac{10^{-5}}{3})x_{100}-(6-\frac{10^{-5}}{3})x_{120}-x_{200} \\ -(6-\frac{10^{-5}}{3})x_{220}-x_{230} \tag{6.64}$$

sujeito a:

$$2 \leq x_{100}+x_{102}+x_{120}+x_{122}+x_{200}+x_{220}+x_{230} \leq 3 \tag{6.65}$$

$$x_{ipq} \in \{0,1\} \tag{6.66}$$

Ordenando as variáveis x_{ipq} conforme a ordenação crescente dos coeficientes V'_{ipq}, obtém-se:

$$x_{102} \leq x_{122} \leq x_{100} \leq x_{120} \leq x_{220} \leq x_{200} \leq x_{230}$$

Escolhendo as 2 primeiras variáveis iguais a 1, $\hat{x}_{102}=\hat{x}_{122}=1$, e as demais iguais a zero, a solução ótima será: $\hat{x}^T=(0,1,0,1,0,0,0)$.

Passo 2(d)": Como $v(P_{\hat{\alpha},\hat{\mu}})=1+5-\frac{10^{-5}}{3}+1=7-\frac{10^{-5}}{3}\geq 0$ e $\hat{x}$ não é factível a P, vá para o Passo 3 modificar μ.

Passo 3": *(Modificando μ)* Desde que $k<10$:

3(a)": Atualize o subgradiente suavizado e o multiplicador *surrogate*:

$$\begin{aligned} G_{00} &= 1-\hat{x}_{100}-\hat{x}_{200}=1 \\ G_{02} &= 1-\hat{x}_{102}-\hat{x}_{200}=0 \\ G_{20} &= 1-\hat{x}_{100}-\hat{x}_{120}-\hat{x}_{220}=1 \\ G_{22} &= 1-\hat{x}_{102}-\hat{x}_{122}-\hat{x}_{220}=-1 \\ G_{30} &= 1-\hat{x}_{120}-\hat{x}_{220}-\hat{x}_{230}=1 \\ G_{32} &= 1-\hat{x}_{122}-\hat{x}_{220}-\hat{x}_{230}=0 \end{aligned}$$

onde $G^T=(1,0,1,-1,0,-1)$. Como $\hat{G}^TG=-2<0$, então:

$$\gamma \leftarrow -\frac{\|(0,1,-1,1,0,1)\|\|(1,0,1,-1,1,0)\|}{-2}=-\frac{2\times 2}{-2}=2$$

$$\beta=-\gamma\frac{\hat{G}^TG}{\|\hat{G}\|^2}=-2\frac{(-2)}{4}=1$$

$$t \leftarrow \frac{(1,0,5-\frac{10^{-5}}{3},0,1,0)(1,0,1,-1,1,0)^T-(-10)}{\|\hat{G}\|^2}=\frac{17-\frac{10^{-5}}{3}}{4}$$

$$\hat{G} \leftarrow G+\beta\hat{G}=(1,0,1,-1,1,0)^T+(0,1,-1,1,0,1)^T=(1,1,0,0,1,1)^T$$

$$\begin{aligned} \hat{\mu} \leftarrow \hat{\mu}-t\hat{G} &= (1,0,5-\frac{10^{-5}}{3},0,1,0)^T-(\frac{17-\frac{10^{-5}}{3}}{4})(1,1,0,0,1,1)^T \\ &= (1-(\frac{17-\frac{10^{-5}}{3}}{4}),-(\frac{17-\frac{10^{-5}}{3}}{4}),5-\frac{10^{-5}}{3},0,1-(\frac{17-\frac{10^{-5}}{3}}{4}),-(\frac{17-\frac{10^{-5}}{3}}{4}))^T \end{aligned}$$

3(b)": Projete o novo $\hat{\mu}$ sobre $\{\mu\geq 0\}$ fazendo $\hat{\mu}_{rs}\leftarrow\max\{0,\hat{\mu}_{rs}\}$ para todo $r\in X$ e todo $s\in Y$: $\hat{\mu}\leftarrow(0,0,5-\frac{10^{-5}}{3},0,0,0)^T$. Então, faça $k\leftarrow 4$ e retorne ao Passo 2 para continuar a resolução do dual $D_{\hat{\alpha},L}$.

Passo 2"': *(Resolução de $P_{\alpha,\mu}$)* Como

$$\hat{\mu}_{00}(1-\hat{x}_{100}-\hat{x}_{200})+\hat{\mu}_{02}(1-\hat{x}_{102}-\hat{x}_{200})+\hat{\mu}_{20}(1-\hat{x}_{100}-\hat{x}_{120}-\hat{x}_{220})+$$
$$\hat{\mu}_{22}(1-\hat{x}_{102}-\hat{x}_{122}-\hat{x}_{220})+\hat{\mu}_{30}(1-\hat{x}_{120}-\hat{x}_{220}-\hat{x}_{230})+\hat{\mu}_{32}(1-\hat{x}_{122}-\hat{x}_{220}-\hat{x}_{230})=$$
$$=0+0+(5-\frac{10^{-5}}{3})+0+0+0=5-\frac{10^{-5}}{3}\geq 0,$$

realize o "término antecipado" e vá para o Passo 3 modificar $\hat{\mu}$.

Passo 3"': *(Modificando μ)* Desde que $k<10$:

3(a)"': Atualize o subgradiente suavizado e o multiplicador *surrogate*:

Como $\hat{x}$ foi mantido, assim como calculado no Passo 3(A)', mantém-se $G^T=(1,0,1,-1,1,0)$. E, como

$$\hat{G}^TG=(1,1,0,0,1,1)(1,0,1,-1,1,0)^T=2>0,$$

então $\beta=0$ e não será necessário calcular γ. Assim, calcule

$$t\leftarrow\frac{(0,0,5-\frac{10^{-5}}{3},0,0,0)(1,0,1,-1,1,0)^T-(-10)}{\|\hat{G}\|^2}=\frac{15-\frac{10^{-5}}{3}}{4}$$
$$\hat{G}\leftarrow G+\beta\hat{G}=(1,0,1,-1,1,0)^T+\overrightarrow{0}=(1,0,1,-1,1,0)^T$$
$$\hat{\mu}\leftarrow\hat{\mu}-t\hat{G}=(0,0,5-\frac{10^{-5}}{3},0,0,0)^T-(\frac{15-\frac{10^{-5}}{3}}{4})(1,0,1,-1,1,0)^T$$

3(b)"': Projete o novo $\hat{\mu}$ sobre $\{\mu\geq 0\}$ fazendo $\hat{\mu}_{rs}\leftarrow\max\{0,\hat{\mu}_{rs}\}$ para todo $r\in X$ e todo $s\in Y$:

$$\hat{\mu}\leftarrow(0,0,0,\frac{15-\frac{10^{-5}}{3}}{4},0,0)^T.$$

Então, faça $k\leftarrow 5$ e retorne ao Passo 2 para continuar a resolução do dual $D_{\hat{\alpha},L}$.

Passo 2"": *(Resolução de $P_{\alpha,\mu}$)* Como

$$\hat{\mu}_{00}(1-\hat{x}_{100}-\hat{x}_{200})+\hat{\mu}_{02}(1-\hat{x}_{102}-\hat{x}_{200})+\hat{\mu}_{20}(1-\hat{x}_{100}-\hat{x}_{120}-\hat{x}_{220})+$$
$$\hat{\mu}_{22}(1-\hat{x}_{102}-\hat{x}_{122}-\hat{x}_{220})+\hat{\mu}_{30}(1-\hat{x}_{120}-\hat{x}_{220}-\hat{x}_{230})+\hat{\mu}_{32}(1-\hat{x}_{122}-\hat{x}_{220}-\hat{x}_{230})=$$
$$=0+0+0-(\frac{15-\frac{10^{-5}}{3}}{4})+0+0=-(\frac{15-\frac{10^{-5}}{3}}{4})<0,$$

resolva por inspeção o problema $P_{\hat{\alpha},\hat{\mu}}$:

$$\text{maximizar}\quad(\frac{15-\frac{10^{-5}}{3}}{4})(1-x_{102}-x_{122}-x_{220}) \tag{6.67}$$

sujeito a:

$$2\leq x_{100}+x_{102}+x_{120}+x_{122}+x_{200}+x_{220}+x_{230}\leq 3 \tag{6.68}$$
$$x_{ipq}\in\{0,1\} \tag{6.69}$$

Desde que os coeficientes V'_{ipq} de $x_{100},\ x_{120},\ x_{200}$ e x_{230} são iguais a zero, de acordo com a resolução por inspeção de $P_{\alpha,\mu}$ dada pelo presente trabalho, fixe em 1 quaisquer 2 destas variáveis, por exemplo, $\hat{x}_{100}=\hat{x}_{230}=1$, e as demais em zero. Ou seja, a solução ótima será: $\hat{x}^T=(1,0,0,0,0,0,1)$.

Passo 2(c)"": Como $v(P_{\hat{\alpha},\hat{\mu}}) = \frac{15-\frac{10^{-5}}{3}}{4} \geq 0$ e $\hat{x}$ é factível a P, atualize o limitante inferior para D_S, $Z_{LI} = x_{100} + x_{102} + x_{120} + x_{122} + x_{200} + x_{220} + x_{230} = 2$, e vá para o Passo 5 aumentar $\hat{\alpha}$

Passo 5"": *(Aumentando α)* Até o momento, tem-se: $\alpha = 1.8 \leq D_S$. Como $|Z_{LS} - Z_{LI}| = 3 - 2 = 1 < 1.1 = \epsilon$. Páre; aproxime $v(D_S) = Z_{LS} = 3$

Capítulo 7

Conclusões e Propostas Futuras

7.1 Conclusões

Este trabalho abordou o PCP do Produtor que é definido pelo problema bidimensional de arranjar ortogonalmente o máximo número de retângulos (l, w) (faces das caixas) dentro de um retângulo maior (L, W) (superfície do palete). Devido à escala e extensão de certos sistemas logísticos, um pequeno aumento no número de produtos carregados sobre o palete pode resultar em economias substanciais. A complexidade computacional destes problemas ainda é desconhecida mas a dificuldade na resolução aumenta proporcionalmente à razão entre as dimensões do palete e as dimensões da caixa. Os algoritmos exatos conhecidos apresentam complexidade exponencial sobre o tempo computacional, pois utilizam métodos de busca em árvore que podem chegar a 2^n resoluções de subproblemas, onde n é o número de caixas como solução ótima.

Algumas formulações matemáticas foram apresentadas para esse problema e, dentre elas, foi utilizada a Formulação de Beasley adaptada ao PCP do Produtor por Oliveira e Morabito [21]. Testes computacionais foram realizados com a implementação deste modelo no pacote AMPL/CPLEX a fim de que fossem efetuadas comparações com os resultados obtidos em [21].

O principal objetivo deste trabalho foi desenvolver uma heurística *surrogate* para o PCP do Produtor que tivesse garantia de convergência para os multiplicadores *surrogate* ótimos. Para tanto, o Capítulo 4 forneceu todo o tratamento teórico necessário para a compreensão de métodos de resolução aplicados ao PCP do Produtor no Capítulo 5. As relaxações lagrangiana e *surrogate*, bem como os principais resultados encontrados na literatura para problemas minimização, foram adaptados para problemas gerais de maximização e, em sua maioria, foram demonstrados com ou sem adaptações de outros trabalhos dispostos na literatura. Exemplos ilustrativos também foram desenvolvidos para o caso de maximização, pois são pouco encontrados tanto em dualidade lagrangiana como em dualidade *surrogate*. Um contra-exemplo baseado no trabalho de Karwan e Rardin [12] também foi adaptado para problemas de maximização para mostrar que subgradientes não fornecem necessariamente direções de decrescimento da função *surrogate*, mesmo quando elas existem.

O método de otimização do subgradiente, proposto por Held e Karp em 1971 [34] para resolução do dual lagrangiano, foi apresentado juntamente com as propostas realizadas por Camerini

et al. [65]. Tal método possui garantia de convergência, pois, como a relaxação lagrangiana forma uma função de λ que é convexa, então o negativo dos subgradientes $b - Ax$ apontam necessariamente para onde tal função decresce seu valor. Para o caso *surrogate*, a mesma afirmação não é sempre verificada, pois, como mostrado no presente trabalho, a relaxação *surrogate* é uma função quase-convexa, permitindo a existência de intervalos onde tal função é constante. Assim, subgradientes não são necessariamente válidos para o caso *surrogate* e o método de otimização do subgradiente para resolução do dual *surrogate* pode não convergir. Esta dificuldade na atualização dos multiplicadores em D_S tornou o problema pouco estudado na literatura, se comparado ao estudo teórico encontrado sobre a resolução de D_L. E, por isso, a maioria dos métodos de resolução do dual *surrogate* são adaptações de métodos de resolução do dual lagrangiano.

O trabalho de Oliveira e Morabito [21], por exemplo, implementou um algoritmo *branch and bound* exato para a resolução do PCP do Produtor. Em cada nó da árvore, heurísticas eram resolvidas para a atualização dos limitantes. Foram implementadas as heurísticas lagrangiana, *surrogate* e lagrangiana/*surrogate* com o método de otimização do subgradiente para a atualização dos multiplicadores, que não possui garantias de convergências para a heurística *surrogate*. Os resultados apresentados foram superiores para o caso lagrangiano, isto é, a média das soluções obtidas para problemas teste do PCP do Produtor aproximaram-se mais da solução ótima no caso lagrangiano. Desde que o dual *surrogate* fornece solução melhor ou, no pior dos casos, igual ao do dual lagrangiano, o método de otimização do subgradiente pode ter, de fato, divergido para o dual *surrogate*, em pelo menos alguns exemplos.

Dado que o método proposto em [5] garante a convergência para o caso *surrogate*, toda a teoria do trabalho de Sarin *et al.* foi adaptada por este trabalho para problemas de maximização. As idéias implícitas em [5] também foram cuidadosamente detalhadas, assim como a adaptação do algoritmo e as demonstrações. Este método, ao invés de resolver repetidas relaxações *surrogate* para atualização dos multiplicadores *surrogate*, estabelece relações entre P^{μ} e a relaxação lagrangiana $P_{\alpha,\mu}$ (isto é, entre os duais D_S e $D_{\alpha,L}$) tais que repetidas resoluções do dual lagrangiano $D_{\alpha,L}$ por um algoritmo de otimização do subgradiente (que, neste caso, possui garantias de convergência) fornece limitantes para D_S. Mais ainda, este método mantém a principal idéia dos algoritmos de busca *surrogate*, ou seja, tornar infactível um x super-ótimo ao problema *surrogate*. Além de demonstrar a convergência deste algoritmo, os resultados obtidos em [5] mostraram-se superiores àqueles obtidos pelo melhor procedimento de busca dual *surrogate* até então conhecido: o algoritmo de relaxação linear relatado em Karwan e Rardin [45]. Por isso o interesse do presente trabalho na aplicação deste método ao PCP do Produtor, até então não realizado na literatura. Então, depois da teoria de Sarin *et al.* ter sido adaptada para problemas gerais de maximização, o algoritmo foi desenvolvido para o PCP do Produtor juntamente com um exemplo trivial, para ilustração do funcionamento do novo método de resolução para o PCP do Produtor.

7.2 Propostas Futuras

A principal proposta de pesquisa futura é implementar a heurística *surrogate* proposta neste trabalho para o PCP do Produtor. Como poucos trabalhos da literatura tentam utilizar procedimentos de resolução que não sejam lagrangianos, a idéia também é adaptar o método de Sarin *et al.* para outros problemas da literatura, principalmente para àqueles que não não possuam a

propriedade de integralidade lagrangiana e/ou *surrogate*, isto é, há possibilidade de obtenção de melhores resultados com a resolução de D_S se comparado à resolução de D_L e $P(\bar{S})$.

Outra contribuição seria implementar o mesmo procedimento *branch and bound* exato proposto por Oliveira e Morabito [21] para o caso lagrangiano, substituindo apenas a heurística lagrangiana pela heurística *surrogate* deste trabalho. Em mesmo ambiente de trabalho, finalmente seria mostrado se a relaxação *surrogate* produz ou não melhores limitantes superiores que a relaxação lagrangiana; ou até mostrar que, mesmo que os limitantes sejam iguais (D_S possui a propriedade de integralidade), qual procedimento resolve o PCP do Produtor em menor quantidade de tempo, isto é, qual heurística converge mais rapidamente para os multiplicadores duais ótimos.

Implementadas estas idéias, alterações poderiam ser realizadas nos procedimentos *branch and bound* tanto para o caso lagrangiano como para o caso *surrogate* para futuras comparações entre os resultados obtidos. Uma idéia para diminuir o tempo computacional é sempre normalizar o vetor multiplicador dual atual antes de prosseguir para as próximas iterações. Isto porque, como visto na demonstração do Teorema 5.4, a solução ótima de D_S não será alterada. A vantagem em se utilizar os vetores normalizados é que o espaço formado por eles, $\{\lambda : \lambda \geq 0, \sum_{i=1}^{m} \lambda_i = 1\}$, é um conjunto de pontos convexo e limitado. Outra alteração que auxiliaria na diminuição do tempo computacional, é implementar dentro dos procedimentos *branch and bound* as idéias de Alvarez-Valdez *et al.* [20] para reduzir o tamanho do modelo de Beasley para o PCP do Produtor, ao considerar a redução de varáveis e restrições nos modelos com os conjuntos *raster points* e considerações de dominância de posições destes conjuntos.

Referências Bibliográficas

[1] UELZE, R. **Logística empresarial**: uma introdução à administração de transportes. São Paulo: Pioneira, 1974. 292 p.

[2] HODGSON, T. *A combined approach to the pallet loading problem.* **IIE Transactions**, v. 14, n. 3, p. 176-182, 1982.

[3] MORALES, S. R.; MORABITO, R.; WIDMER, J. A. Otimização do Carregamento de Produtos Paletizados em Caminhões. **Gestão & Produção**, v. 4, n. 2, p. 234-250, 1997.

[4] OLIVEIRA, L. K. **Métodos exatos baseados em relaxação lagrangiana e surrogate para o problema de carregamento de paletes do produtor**. 2004. 184 f. Tese (Doutorado em Engenharia de Produção) - Centro de Ciências Exatas e de Tecnologia, Universidade Estadual de São Carlos, São Carlos, 2004.

[5] SARIN, S.; KARWAN, M. H.; RARDIN, R. L. *A new surrogate dual multiplier search procedure.* **Naval Research Logistics**, v. 34, p. 431-450, 1987.

[6] DYCKHOFF, H. *A typology of cutting and packing problems.* **European Journal of Operational Research**, v. 44, p. 145-159, 1990.

[7] WÄSCHER, G.; HAUBNER, H.; SCHUMANN, H. *An improved typology of cutting and packing problems.* **European Journal of Operational Research**, v. 183, p. 1109-1130, 2007.

[8] BEASLEY, J. E. *An exact two-dimensional non guillotine cutting tree search procedure.* **Operations Research**, v. 33, p. 49-64, 1985b.

[9] TSAI, R. D.; MALSTROM, E. M.; KUO, W. *Three dimensional palletization of mixed box sizes.* **IEE Transactions**, v. 25, n. 4, p. 64-75, 1993.

[10] HADJICONSTANTINOU, E.; CHRISTOFIDES, N. *An exact algorithm for general, orthogonal, two-dimensional knapsach problems.* **European Journal of Operational Research**, v. 83, p. 39-56, 1995.

[11] CHEN, C. S.; LEE, S. M.; SHEN, Q. S. *An analytical model for the container loading problem.* **European Journal of Operational Research**, v. 80, p. 68-76, 1995.

[12] KARWAN, M. H.; RARDIN, R. L. *Some relationships between lagrangian and surrogate duality in integer programming.* **Mathematical Programming**, v. 17, p. 320-334, 1979.

[13] FACCIO, A. P. **Propostas de solução para o problema de corte de estoque bidimensional de uma fábrica de móveis de pequeno porte.** 2008. 121 f. Dissertação (Mestrado em Matemática Aplicada) - Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2008.

[14] LETCHFORD, A. N.; AMARAL, A. *Analysis of upper bounds for the pallet loading problem.* **European Journal Of Operational Research**, v. 3, n. 132, p. 582-593, 2001.

[15] FARAGO, R.; MORABITO, R. Um método heurístico baseado em relaxação lagrangiana para o problema de carregamento de paletes do produtor. **Pesquisa Operacional**, v. 2, n. 20, p. 197-212, 2000.

[16] MORABITO; R.; FARAGO, R. *A tight lagrangean relaxation bound for the manufacturer's pallet loading problem.* **Studia Informatica Universalis**, v. 2, p. 57-76, 2002.

[17] HERZ, J. C. *A recursive computing procedure for two-dimensional stock cutting.* **I. B. M.**, v. 16, p. 462-469, 1972.

[18] CHRISTOFIDES, N.; WHITLOCK, C. *An algotithm for two-dimensional cutting problems.* **Operations Research**, v. 25, p. 30-45, 1977.

[19] SCHEITHAUER, G.; TERNO, J. *The G4-Heuristic for the pallet loading problem.* **Journal of the Operational Research Society**, v. 47, p. 511-522, 1996.

[20] ALVAREZ-VALDEZ, R.; PARREÑO, F.; TAMARIT, J. M. *A branch-and-cut algorithm for the pallet loading problem.* **Computers & Operations Research**, v. 32, p. 3007-3029, 2005.

[21] OLIVEIRA, L. K.; MORABITO, R. *Métodos exatos baseados em relaxação lagrangiana e surrogate para o problema de carregamento de paletes do produtor.* **Pesquisa Operacional**, v. 26, p. 403-432, 2006.

[22] RIBEIRO, G. M.; LORENA, L. A. N. *Lagrangean relaxation with clusters and column generation for the manufacturer's pallet loading problem.* **Computers & Operations Research**, v. 34, p. 2695-2708, 2005.

[23] DOWSLAND, K. *An exact algorithm for the pallet loading problem.* **European Journal of Operational Research**, v. 31, p. 78-84, 1987.

[24] BHATTACHARYA, R.; ROY, R.; BHATTACHARYA, S. *An Exact Depth-First Algorithm for the pallet loading problem.* **European Journal or Operational Research**, v. 110, p. 610-625, 1998.

[25] YOUNG-GUN, G.; MAING-KYU, K. *A fast algorithm for two-dimensional pallet loading problems of large size.* **European Journal of the Operational Research**, v. 134, p. 193-202, 2001.

[26] MORABITO, R.; MORALES, S. R. *A simple and effective heuristic to solve the manufacturer's pallet loading problem.* **Journal of the Operational Research Society**, v. 49, p. 819-828, 1998.

[27] LINS L.; LINS S.; MORABITO, R. *An L-approach for packing (l,w)-rectangles into rectangular and L-shaped pieces.* **Journal of the Operational Research Society**, v. 54, p. 777-789, 2003.

[28] BIRGIN, E. G.; MORABITO, R.; NISHIHARA, F. H. *A note on an L-approach for solving the manufacturer's pallet loading problem.* **Journal of the Operational Research Society**, v. (to appear), 2005.

[29] PUREZA, V.; MORABITO, R. *Some experiments with a simple tabu search algorithm for the manufacturer's pallet loading problem.* **Computers & Operations Research**, v. 33, n. 3, p. 804-819, 2006.

[30] ALVAREZ-VALDEZ, R.; PARREÑO, F.; TAMARIT T. M. *A tabu search algorithm for pallet loading problem.* **OR Spectrum**, v. 27, n. 1, p. 43-61, 2005.

[31] HERBERT, A.; DOWSLAND, K. *A family of genetic algorithm for the pallet loading problem.* In: **Metaheuristic:** theory and applications. Dordrecht: Kluwer Academic Publishers, 1996. p. 378-406.

[32] GONÇALVES, J. F. *A hybrid genetic algorithm-heuristic for a two-dimensional orthogonal packing problem.* **European Journal of Operational Research**, v. 183, p. 1212-1229, 2007.

[33] NEMHAUSER, G. L.; WOLSEY, L. A. **Integer and Combinatorial Optimization**. New York: John Wiley & Sons, 1999.

[34] HELD, M.; KARP, R. M. *The travelling salesman problem and minimum spanning trees: part II.* **Mathematical Programming**, v. 1, p. 6-25, 1971.

[35] GEOFFRION, A. M. *Lagrangian relaxation and its uses in integer programming.* **Mathematical Programming Study**, v. 2, p. 82-114, 1974.

[36] GLOVER, F. *A multiphase-dual algorithm for the zero-one integer programming problems.* **Operations Research**, v. 23, v. 434-451, 1965.

[37] BAZARAA, M. S.; SHERALI, H. D.; SHETTY, C. M. **Non linear programing:** theory and algorithms. 2 ed. New York: John Wiley & Sons, 1993. 638 p.

[38] BEASLEY, J. E. *Modern heuristic techniques for combinatorial problems.* In: **Lagrangean Relaxation**. New York: John Wiley & Sons, 1993. p. 243-303.

[39] GALVÃO, R. D.; ESPEJO, L. G. A.; BOFFEY, B. *A comparison of lagrangean and surrogate relaxations for the maximal covering location problem.* **European Journal of Operational Research**, v. 124, p. 377-389, 2000.

[40] GALVÃO, R. D.; ESPEJO, L. G. A.; BOFFEY, B. *A hierarchical model for the location of perinatal facilities in the municipality of Rio de Janeiro.* **European Journal of Operational Research**, v. 139, p. 495-517, 2002.

[41] FUMERO, F. *A modified subgradient algorithm for lagrangean relaxation.* **Computers & Operations Research**, v. 28, p. 33-52, 2001.

[42] SRIDHARAN, R. *The capacitated plant location problem.* **European Journal of Operational Research**, v. 87, p. 203-213, 1995.

[43] RARDIN; R. L.; UNGER, V. E. *Surrogate constraints and the strength of bounds derived from 0-1 Benders' Partitioning Procedures.* **Operations Research**, v. 24, p. 1169-1175, 1976.

[44] HOLMBERG, K. *On using approximations of the Benders master problem.* **European Journal of Operational Research**, v. 77, p. 111-125, 1994.

[45] KARWAN, M. H.; RARDIN, R. L. *Surrogate dual multiplier search procedures in integer programming.* **Operations Research**, v. 32, p. 52-69, 1984.

[46] LORENA, L. A. N.; LOPES, F. B. *A surrogate heuristic for set covering problems.* **European Journal of Operational Research**, v. 79, p. 138-150, 1994.

[47] LORENA, L. A. N.; NARCISO, M. G. *Relaxation heuristics for a generalized assigment problem.* **European Journal of Operational Research**, v. 91, p. 600-610, 1996.

[48] ESPEJO, L. G. A.; GALVÃO, R. D. *O uso das relaxações lagrangeana e surrogate em problemas de programação inteira.* **Pesquisa Operacional**, v. 22, p. 387-402, 2002.

[49] PARKER, R. G.; RARDIN, R. L. **Discrete Optimization**. Boston: Academic, 1988. 472 p.

[50] WOLSEY, L. A. **Integer Programming**. New York: Jonh Wiley & Sons, 1998. 264 p.

[51] ARENALES, M.; ARMENTANO, V.; MORABITO, R.; YANASSE, H. H. **Pesquisa Operacional.** Rio de Janeiro: Elsevier, 2007. 523 p.

[52] GREENBERG, H. J.; PIERSKALLA, W. P. *Surrogate mathematical programs.* **Operational Research**, v. 18, p. 924-939, 1970.

[53] FISHER, M. L. *The lagrangian relaxation method of solving integer programming problems.* **Management Science**, v. 27, p. 1-18, 1985.

[54] BANERJEE, K. **Generalized lagrange multipliers in dynamic programming**. Research Report No. ORC 71-12. Operations Research Center, University of California, Berkeley, California, 1971.

[55] KARWAN, M. H.; RARDIN, R. L. *Surrogate duality in a branch and bound procedure.* **Naval Research Logistics Quaterly**, v. 28, p. 93-101, 1981.

[56] GLOVER, F. *Surrogate constraint duality in mathematical programming*, **Operations Research**, v. 23, p. 434-451, 1975.

[57] DYER, M. E. *Calculating surrogate constraints.* **Mathematical Programming**, v. 19, p. 255-278, 1980.

[58] GAVISH, B.; PIRKUL, H. *Efficient algorithms for solving multiconstraint zero-one knapsack problems to otimality.* **Mathematical Programming**, v. 31, p. 78-105, 1985.

[59] KARWAN, M. H.; PAN, S. T. **Surrogate duality and the fixed charge transportation problem**. Research Report No. 79-10. Department of Industrial Engineering, State University of New York at Buffalo, Buffalo, NY, 1979.

[60] DINKEL, J. J.; KOSHENBERGER, G. A. *An implementation of surrogate constraint duality.* **Operations Research**, v. 26, p. 358-364, 1978.

[61] GLOVER, F.; KARNEY, D.; KLINGMAN, D. *A Study of Alternative Relaxation Approaches for a Manpower Planning Problem.* In: **Quantitative Planning and Control**. Academic Orlando, 1979, p. 141-164.

[62] CARVALHO, J. M. V. **Programação Inteira**. 2001. 263 f. Dissertação (Mestrado) - Universidade do Minho, 2001.

[63] HELD, M.; WOLFE, P.; CROWDER, H. *Validation of subgradient optimization.* **Mathematical Programming**, v. 6, p. 62-88, 1974.

[64] GUINGARD, M. *Lagrangean relaxation.* **Sociedad de Estadística e Investigación Operativa**, v. 11, n. 2, p. 151-228, 2003.

[65] CAMERINI, P. M.; FRATTA, L.; MAFFIOLI, F. *On improving relaxation methods by gradient techniques.* **Mathematical Programming Study**, Amsterdam, v. 3, p. 26-34, 1975.

[66] Crowder, H. *Computational Improvements for Subgradient Optimization.* **Symposio Matematica**, Rome, v. 17, p. 357-372, 1976.

[67] BIRGIN, M.; LOBATO, R. D.; MORABITO, R. *An effective recursive partitioning approach for the packing of identical rectangles in a rectangle.* **Journal of the Operational Research Society**, 2009. DOI: 10.1057/jors.2008.141.

Apêndice A

Implementação no AMPL/CPLEX

A.1 Exemplo Numérico da Seção 3.3.2

param geracao integer>= 0;

param p >= 0;

param q >= 0;

param T integer;

param $Vetor\{0..T\}$;

param $Vetory\{0..K\}$;

option randseed $'180'$;

option showstats 1;

cplexoptions 'timing=1 mipdisplay=3 ';

'iisfind=1 varselect=-1 startalgorithm=1 ';

set orientacao;

set X;

set Y;

param L integer >= 0;

W integer >= 0;

param nmin *integer* >= 0;

param nmax *integer* >= 0;

param $l\{$orientacao$\}$ >= 0;

param $w\{$orientacao$\}$ >= 0;

param $a\{$orientacao,X,Y,X,$Y\}$ >= 0;

var caixa$\{$orientacao,X,$Y\}$ binary;

maximize numcaixas:

sum$\{i$ in orientacao, j in X, k in $Y : j <= (L-l[i])$ and $k <= (W-w[i])\}$ caixa$[i,j,k]$;

subject to sobreposicao $\{rinX, sinY\}$:

sum$\{i$ in orientacao, j in X, k in $Y : j <= (L - l[i])$ and $k <= (W - w[i])\}$ $a[i,j,k,r,s]*$caixa$[i,j,k] <= 1$;

subject to mincaixas:

```
        sum{i in orientacao, j in X, k in Y : j <= (L - l[i]) and k <= (W - w[i])}
caixa[i,j,k] >= nmin;
    subject to maxcaixas:
      sum{i in orientacao, j in X, k in Y : j <= L - l[i] and k <= W - w[i]} caixa[i,j,k]
<= nmax;

set orientacao := 12;
param L := 5;
param W := 4;
param l :=          13          22;
param w :=          12          23;
param nmin := 0;

let nmax := (L * W) div (l[1] * w[1]);
if geracao= 1 then
{
  let T := L - w[1];
  for {t in 0..T}
  {
  letVetor[t] := -1;
  }
  printf "setX :=">
    ("dados.dat");
  for {i in 0..(L - w[1])}
  {
    let p := 0;
    for {j in 0..(L - w[1])}
    {
      if p <= L - w[1] then
      {
        let p := l[1] * i + l[2] * j;
        if p <= L - w[1] then
        {
          let Vetor[p] := p;
        }
      }
    }
  }

  for {t in 0..T}
  {
    if Vetor[t] >= 0 then
    {
```

```
        printf "td",Vetor[t] >("dados.dat");
      }
    }
    printf ";»("dados.dat");

    let K := W − w[1];
    for {t in 0..K}
    {
      let Vetory[t] := −1;
    }
    printf "nset Y :=»
     ("dados.dat");

    for {i in 0..(W − w[1])}
    {
      let q := 0;
      for {j in 0..(W − w[1])}
      {
        if q <= W − w[1] then
        {
          let q := w[1] ∗ i + w[2] ∗ j;
          if q <= W − w[1] then
          {
            let Vetory[q] := q;
          }
        }
      }
    }

    for {t in 0..K}
    {
      if Vetory[t] >= 0 then
      {
        printf "td",Vetory[t] >("dados.dat");
      }
    }
    pr in  tf ";»("dados.dat");

    data ("dados.dat");
  }
  else
  {
    data ("dados.dat");
```

```
for {i in orientacao}
{
  for {j in X}
  {
    for {k in Y}
    {
      for {r in X}
      {
        for {s in Y}
        {
          if j <= r and r <= j + l[i] - 1 and j + l[i] - 1 <= L - 1 then
          {
            if k <= s and s <= k + w[i] - 1 and k + w[i] - 1 <= W - 1 then
            {
              let a[i, j, k, r, s] := 1;
            }
            else
            {
              let a[i, j, k, r, s] := 0;
            }
          }
          else
          {
            let a[i, j, k, r, s] := 0;
          }
        }
      }
    }
  }
}

solve;
display caixa;
numcaixas;
}
```

Apêndice B

Conceitos Primários

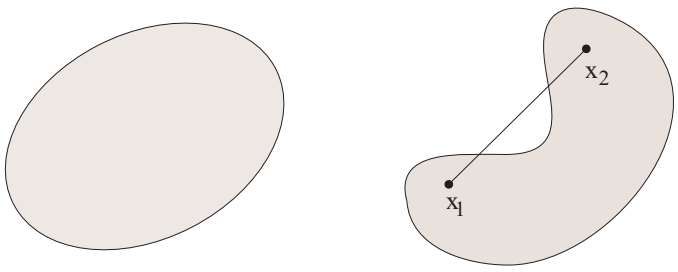
Algumas definições são fundamentais em dualidade lagrangiana e *surrogate* por serem amplamente utilizadas no decorrer de sua teoria e, por isso, algumas noções importantes serão também apresentadas. Os conceitos aqui estabelecidos são tratados apenas sobre o corpo dos $\mathbb{R}^n$, uma vez que para o presente trabalho, será suficiente. Para generalização deste conteúdo, veja [37].

B.1 Análise de Convexidade

Definição B.1.1. *Qualquer ponto da forma* $\sum_{j=1}^{k} \rho_j x_j$*, tais que* $x_1, \ldots, x_k \in S$ *e* $\sum_{j=1}^{k} \rho_j = 1$*, é uma combinação convexa dos pontos* $x_1, \ldots, x_k$*. Se* $\rho_j \neq 1$*, para todo* $j = 1, \ldots, k$*, então a combinação convexa é chamada estrita.*

A interpretação geométrica da definição acima, utilizada no decorrer desta seção, recai sobre o fato de que para cada $\rho \in [0, 1]$, $\rho x_1 + (1 - \rho)x_2$ representa um ponto do segmento de reta que liga x_1 a x_2.

Definição B.1.2. *Um conjunto* S *de* $\mathbb{R}^n$ *é convexo se, para quaisquer dois pontos de* S*, existe um segmento de reta contido em* S *que os une.*



(a) Conjunto convexo (b) Conjunto não convexo

Figura B.1: Ilustração em $\mathbb{R}^2$ de conjuntos convexos e conjuntos não-convexos

Em outras palavras, um conjunto S é convexo se, para x_1, $x_2 \in S$, a combinação convexa destes pontos, $\rho x_1 + (1 - \rho)x_2$, também está contida em S, para todo $\rho \in [0, 1]$ com $\rho_1 = \rho$ e $\rho_2 = 1 - \rho$.

Definição B.1.3. *Um ponto x em um conjunto convexo S é chamado ponto extremo de S, se x não pode ser representado por uma combinação convexa estrita de dois pontos distintos de S.*

Ou seja, x é ponto extremo de S se $x = \rho x_1 + (1-\rho)x_2$, com $\rho \in (0,1)$, implicar diretamente em $x = x_1 = x_2$. A Figura B.2 mostra x_1, ponto extremo de um conjunto convexo, e x_2 e x_3, pontos não-extremos deste mesmo conjunto.

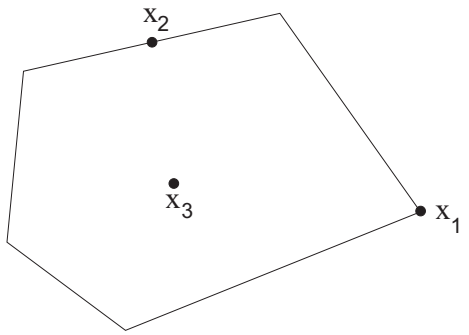


Figura B.2: Pontos extremos e pontos não-extremos de um conjunto convexo

Definição B.1.4. *Seja S um conjunto arbitrário de $\mathbb{R}^n$. O envoltório convexo de S, denotado por $Conv(S)$, é o conjunto de todos os pontos que são combinações convexas dos pontos de S.*

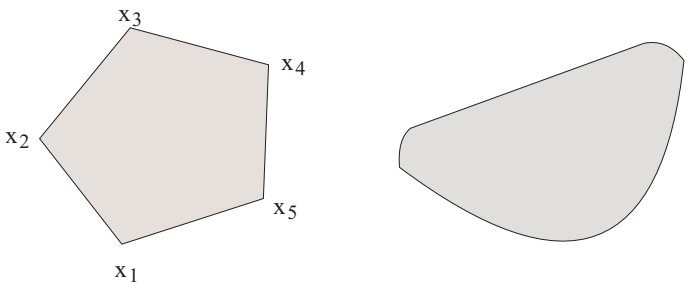
Então, $x \in Conv(S)$ se e somente se x pode ser representado como:

$$x = \sum_{j=1}^{k} \rho_j x_j$$

$$\sum_{j=1}^{k} \rho_j = 1$$

$$\rho_j \geq 0, j = 1, \ldots, k$$

Definição B.1.5. *O envoltório convexo de S contendo um número finito de pontos $x_1, \ldots, x_k$ é chamado politopo e é denotado por $[S]$ (Veja Figura B.3).*



(a) Envoltório convexo que é politopo

(b) Envoltório convexo que não é politopo

Figura B.3: Ilustração em $\mathbb{R}^2$ de politopo e de envoltório convexo que não é politopo

Definição B.1.6. *Seja $f : S \to \mathbb{R}$, onde S é um conjunto convexo não-vazio em $\mathbb{R}^n$. A função f é dita convexa em S se*

$$f[\rho x_1 + (1-\rho)x_2] \leq \rho f(x_1) + (1-\rho)f(x_2)$$

para cada $x_1, x_2 \in S$ e para cada $\rho \in (0,1)$.

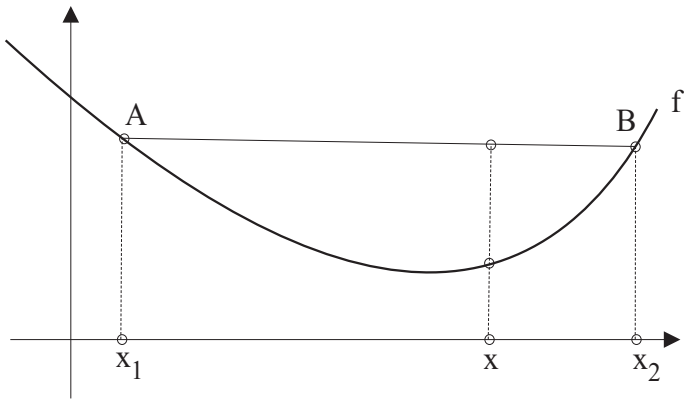


Figura B.4: Função Convexa

Pela Figura B.4, todo ponto do domínio S que está entre x_1 e x_2 é descrito pela combinação convexa de x_1 e x_2, $x = \rho x_1 + (1-\rho)x_2$, bem como $\rho f(x_1) + (1-\rho)f(x_2)$ descreve a corda que liga os pontos $A = (x_1, f(x_1))$ e $B = (x_2, f(x_2))$; ambos quando $\rho \in (0,1)$. Assim, a Definição B.1.6 explicita f como uma função convexa, desde que toda corda que une quaisquer pontos, A e B, esteja acima da imagem de qualquer ponto compreendido entre x_1 e x_2.

Definição B.1.7. *Sejam S um conjunto convexo não-vazio em $\mathbb{R}^n$ e $f : S \to \mathbb{R}$ uma função convexa em S. Então um vetor $G \in R^n$ é chamado um subgradiente de f no ponto $y \in S$ se*

$$f(x) \geq f(y) + G^T(x-y) \quad \textit{para todo} \quad x \in S$$

Pela Figura B.5, observe que a altura $f(x)$ é maior que a soma das alturas $f(y)+G^T(x-y)$.

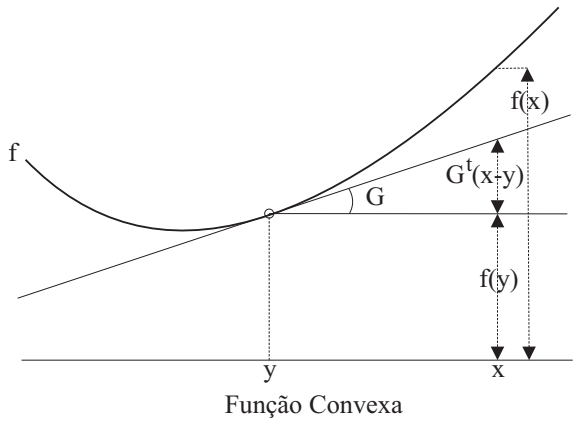


Figura B.5: Interpretração geométrica de subgradientes

Dada $f : \mathbb{R}^n \to \mathbb{R}$, se $y \in S$ é tal que $f(x) \geq f(y)$ para todo $x \in S$, então y é chamado *ponto de mínimo global* para f. Se $y \in S$ e se existe uma vizinhança de y, $V_\gamma(y)$, tal que $f(x) \geq f(y)$ para todo $x \in S \cap V_\gamma(y)$, então y é chamado *ponto de mínimo local* para f.

Teorema B.1. *Sejam S um conjunto convexo não-vazio em $\mathbb{R}^n$ e $f : S \to \mathbb{R}$ uma função convexa em S. Considere o problema de minimizar $f(x)$ sujeito a $x \in S$ e suponha que $y \in S$ é um ponto de mínimo local para o problema. Então y também é um ponto de mínimo global.*

Demonstração: Esta demonstração foi realizada tendo como base a prova do teorema da página 101 de [37].

Por hipótese, y ser mínimo local implica que existe uma vizinhança de y, $V_\gamma(y)$, tal que

$$f(x) \geq f(y) \quad \text{para todo} \quad x \in S \cap V_\gamma(y) \tag{B.1}$$

Suponha, por absurdo, que y não é mínimo global. Então, deve existir algum $x \in S$ tal que $f(x) < f(y)$. Pela convexidade de f, para todo $\rho \in (0, 1)$, tem-se:

$$f[\rho x + (1-\rho)y] \leq \rho f(x) + (1-\rho)f(y) < \rho f(y) + (1-\rho)f(y) = f(y)$$

Mas, para um $\rho > 0$ suficientemente pequeno, tem-se $\rho x + (1-\rho)y \in S \cap V_\gamma(y)$ o que implica pela última desigualdade que existirá um elemento, $\rho x + (1-\rho)y$, que está na vizinhança de y e tal que $f[\rho x + (1-\rho)y] < f(y)$. Por (B.1), isto é um absurdo e portanto, y é mínimo global. ■

Teorema B.2. *Sejam S um conjunto convexo não-vazio em $\mathbb{R}^n$ e $f : S \to \mathbb{R}$ uma função convexa. Considere o problema de minimizar $f(x)$ sujeito a $x \in S$. Então o ponto $y \in S$ é uma solução ótima para este problema se e somente se f possui um subgradiente G no ponto y tal que $G^T(x-y) \geq 0$ para todo $x \in S$.*

Demonstração: Esta demonstração foi realizada tendo como base a prova do teorema da página 102 de [37].

- Suponha G um subgradiente de f em y tal que $G^T(x-y) \geq 0$ para todo $x \in S$.

  Pela convexidade de f

  $$f(x) \geq f(y) + \underbrace{G^T(x-y)}_{\geq 0} \geq f(y) \quad \text{para todo} \quad x \in S$$

  e, portanto, y é uma solução ótima para o problema.

- Suponha y uma solução ótima para o problema.

  Por construção, sejam os conjuntos convexos:

  $$\wedge_1 = \{(x-y, \bar{y}) : x \in \mathbb{R}^n, \bar{y} > f(x) - f(y)\}$$
  $$\wedge_2 = \{(x-y, \bar{y}) : x \in S, \bar{y} \leq 0\}$$

  Note que $\wedge_1 \cap \wedge_2 = \varnothing$, pois, caso contrário, existiria um ponto $(x, \bar{y})$ tal que:

  $$x \in S \quad \text{e} \quad f(x) - f(y) < \bar{y} \leq 0$$

contrariando a hipótese de y ser solução ótima (pois $f(x)$ seria menor que $f(y)$ em um problema de minimização). Então, como pode ser visto por exemplo em [37], existe um hiperplano que separa $\wedge_1$ e $\wedge_2$, isto é, existe um vetor (G_0, η), diferente do vetor nulo, e um escalar α tal que

$$G_0^T(x-y) + \eta\bar{y} \leq \alpha \qquad x \in \mathbb{R}^n,\ \bar{y} > f(x) - f(y) \tag{B.2}$$

$$G_0^T(x-y) + \eta\bar{y} \geq \alpha \qquad x \in S,\ \bar{y} \leq 0 \tag{B.3}$$

Primeiramente, se $x = y$ e $\bar{y} = 0$, por (B.3), $\alpha \leq 0$. Se $x = y$ e $\bar{y} = \varepsilon > 0$, por (B.2), $\eta\varepsilon \leq \alpha$. Se isto é válido para todo $\varepsilon > 0$, então $\eta \leq 0$ e $\alpha \geq 0$. Resumindo, é obtido $\eta \leq 0$ e $\alpha = 0$. Caso $\eta = 0$, por (B.2), $G_0^T(x-y) \leq 0$ para todo $x \in \mathbb{R}^n$.

Agora, se $x = y + G_0$, então $G_0 = x - y$ e pode-se escrever:

$$\|G_0\|^2 = G_0^T(x-y) \leq 0$$

Como $\|G_0\|^2 \geq 0$ sempre, então $\|G_0\|^2 = 0 \Rightarrow G_0 = 0$, já que é tomado o caso em que $x \neq y$. Por hipótese, $(G_0, \eta) \neq (0,0)$; então, para exclusão do vetor nulo, $\eta < 0$. Divindo (B.2) e (B.3) por $-\eta$ e denotando $\frac{-G_0}{\eta}$ por G, obtém-se as seguintes inequações:

$$\bar{y} \geq G^T(x-y) \qquad x \in \mathbb{R}^n,\ \bar{y} > f(x) - f(y) \tag{B.4}$$

$$G^T(x-y) - \bar{y} \geq 0 \qquad x \in S,\ \bar{y} \leq 0 \tag{B.5}$$

Fazendo $\bar{y} = 0$ em (B.5), obtém-se $G^T(x-y) \geq 0$ para todo $x \in S$. Substituindo em (B.4) obtém-se ainda que

$$f(x) \geq f(y) + G^T(x-y) \qquad \text{para todo } x \in \mathbb{R}^n \tag{B.6}$$

Assim, G é um subgradiente de f no ponto y com a propriedade de que $G^T(x-y) \geq 0$ para todo $x \in S$. ■

Definição B.1.8. *Seja $f : S \to \mathbb{R}$, onde S é um conjunto convexo não-vazio em $\mathbb{R}^n$. A função f é chamada quase-convexa se, para todo x_1 e x_2 em S, tem-se:*

$$f[\rho x_1 + (1-\rho)x_2] \leq \max\{f(x_1), f(x_2)\}.$$

Mais ainda, se $f[\rho x_1 + (1-\rho)x_2] < \max\{f(x_1), f(x_2)\}$, com $x_1 \neq x_2$, então f é dita fortemente quase-convexa.

Note que toda função convexa é também quase-convexa, mas que a reciprocidade não é verdadeira. A diferença entre as duas definições é que na função convexa, toda imagem da combinação convexa dos pontos x_1 e x_2 é necessariamente menor ou igual à corda que liga os pontos $(x_1, f(x_1))$ e $(x_2, f(x_2))$, enquanto que na função quase-convexa basta que toda imagem da combinação convexa dos mesmos pontos sejam menores ou iguais que o maior valor entre $f(x_1)$ e $f(x_2)$. Ou seja, a função quase-convexa admite trechos constantes, ao contrário da função convexa.

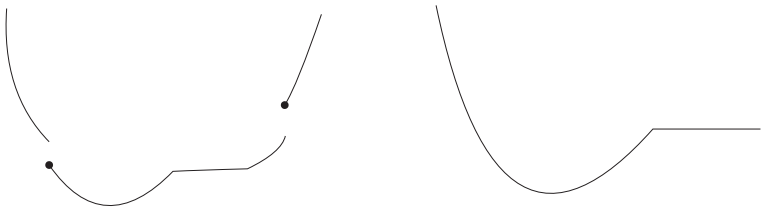


Figura B.6: Exemplos de Função Quase-Convexa

Teorema B.3. *Sejam S um conjunto convexo não-vazio em $\mathbb{R}^n$ e $f : \mathbb{R}^n \to \mathbb{R}$ uma função fortemente quase-convexa. Considere o problema de minimizar $f(x)$ sujeito a $x \in S$ e suponha que y é um ponto de mínimo local para o problema. Então y também é um ponto de mínimo global.*

Demonstração: Por hipótese, y ser mínimo local implica que existe uma vizinhança de y, $V_\gamma(y)$, tal que

$$f(x) \geq f(y) \quad \text{para todo} \quad x \in S \cap V_\gamma(y) \tag{B.7}$$

Suponha, por absurdo, que y não é mínimo global. Então, deve existir algum $x \in S$ tal que $x \neq y$ e $f(x) < f(y)$. Pela quase-convexidade forte de f, para todo $\rho \in (0, 1)$, tem-se:

$$f[\rho x + (1 - \rho)y] < \max\{f(x), f(y)\} = f(y) \tag{B.8}$$

Mas, para um $\rho > 0$ suficientemente pequeno, tem-se $\rho x + (1 - \rho)y \in S \cap V_\gamma(y)$ o que implica pela última desigualdade que existirá um elemento, $\rho x + (1 - \rho)y$, que está na vizinhança de y e tal que $f[\rho x + (1 - \rho)y] < f(y)$. Por (B.7), isto é um absurdo e portanto, y é mínimo global. ■

B.2 Dualidade em Otimização Linear

Além das considerações sobre análise convexa, é importante para este trabalho o conhecimento e entendimento sobre dualidade em otimização linear. Toda teoria apresentada nesta seção foi baseada em Arenales *et al.* [51] e transformada para problemas primais de maximização.

Um *problema de otimização linear* pode ser descrito como a seguir:

$$\begin{array}{lll} (PL) & \text{maximizar} & f(x) = cx \\ & \text{sujeito a:} & Ax = b \\ & & x \geq 0 \end{array}$$

onde seu *problema dual* associado é o problema (veja [51]):

$$\begin{array}{lll} (D) & \text{minimizar} & g(\pi) = b\pi \\ & \text{sujeito a:} & A^T\pi \geq c \end{array}$$

Nesta seção, será mostrado que a solução ótima de D é sempre igual à solução ótima do problema primal, P, em programação linear. São consideradas restrições de igualdade em

P, mas desde que variáveis de folga podem ser consideradas, o resultado torna-se válido para a desigualdade. Teoria análoga pode ser também desenvolvida para restrições de desigualdade para problemas primais de minimização e para casos em que as variáveis são irrestritas, livres de sinal. Para verificação dos problemas duais de todos as combinações de problemas primais possíveis, veja a tabela da página 118 de Arenales *et al.* [51].

Considere A uma matriz $m \times n$ e demais vetores envolvidos com dimensões apropriadas. Adotar $Ax = b$ em PL implica em π irrestrito no problema dual D, uma vez que seu sinal não implica diretamente na penalização da função objetivo.

Definição B.2.1. *(função lagrangiana e função dual) [51] A função lagrangiana é definida por:*

$$\begin{aligned} L(x, \pi) &= c^T x + \pi^T (b - Ax) \\ &= (c^T - \pi^T A)x + \pi^T b \end{aligned}$$

Seja $A = [a_1, a_2, \ldots, a_n]$, onde a_j é a j-ésima coluna da matriz A, e $c = (c_1, c_2, \ldots, c_n)$, então

$$(c^T - \pi^T A) = (c_1 - \pi^T a_1, c_2 - \pi^T a_2, \ldots, c_n - \pi^T a_n)$$

de modo que a função lagrangiana pode ser reescrita por:

$$L(x_1, x_2, \ldots, x_n, \pi) = (c_1 - \pi^T a_1)x_1 + (c_2 - \pi^T a_2)x_2 + \ldots + (c_n - \pi^T a_n)x_n + \pi^T b$$

A função dual é definida por:

$$\begin{aligned} &\max_{x \geq 0} g(\pi) = \{L(x_1, x_2, \ldots, x_n, \pi)\} \\ &= \max_{x \geq 0} \{(c_1 - \pi^T a_1)x_1 + (c_2 - \pi^T a_2)x_2 + \ldots + (c_n - \pi^T a_n)x_n + \pi^T b\} \\ &= \max_{x_1 \geq 0} \{(c_1 - \pi^T a_1)x_1\} + \ldots + \max_{x_n \geq 0} \{(c_n - \pi^T a_n)x_n\} + \pi^T b \end{aligned}$$

Pela Definição B.2.1, é fácil mostrar que o problema dual D fornece um limitante dual (superior, para um problema primal de maximização) para P.

Corolário B.1. *Para todo $\pi \in \mathbb{R}^m$ e para todo x tal que $Ax = b,\ x \geq 0$, então $f(x) \leq g(\pi)$.*

Demonstração: Para todo x é factível a P e $\pi \in \mathbb{R}^m$, então:

$$\begin{aligned} g(\pi) &= \max_{x \geq 0} c^T x + \pi^T (b - Ax) \\ &\geq \max_{x \geq 0 : Ax = b} c^T x + \pi^T (b - Ax) \\ &= \max\{c^T x, \text{ sujeito a: } Ax = b,\ x \geq 0\} \\ &\geq f(x) \end{aligned}$$

■

Observação 1. *Para que $g(\pi)$ forneça bons limitantes superiores para $f(x)$ deve-se ter $c_i - \pi^T a_i \leq 0$ para todo $i = 1, \ldots, n$, pois, caso exista algum i tal que $c_i - \pi^T a_i > 0$, então a parcela $(c_i - \pi^T a_i)x_i$ da função dual da Definição B.2.1 seria tão grande quanto possível, pois $x_i \geq 0$, por hipótese. E isso implicaria num limitante superior inócuo, $g(\pi) = \infty$, isto é, um valor que é limitante superior para qualquer número real.*

Observação 2. *Se $c_i - \pi^T a_i < 0$ então $(c_i - \pi^T a_i)x_i \leq 0$ e para maximizar esta parcela na função dual, deve-se ter $(c_i - \pi^T a_i)x_i = 0$, isto é, $x_i = 0$. Caso $c_i - \pi^T a_i = 0$, analogamente à explicação anterior, x_i poderá assumir qualquer valor não-negativo e, mais uma vez, $(c_i - \pi^T a_i)x_i = 0$. Logo, ao escolher π tal que $c_i - \pi^T a_i \leq 0$, para todo $i = 1, \ldots, n$, tem-se, novamente pela função dual, que $g(\pi) = \pi^T b$ e*

$$\begin{aligned} c_1 \leq \pi^T a_1, \ldots, c_n \leq \pi^T a_n &\Leftrightarrow (\pi^T a_1, \ldots, \pi^T a_n) \geq (c_1, \ldots, c_n) \\ &\Leftrightarrow \pi^T A \geq c^T \\ &\Leftrightarrow (\pi^T A)^T \geq (c^T)^T \\ &\Leftrightarrow A^T \pi \geq c \end{aligned}$$

que são as próprias considerações a serem realizadas para formar o problema dual lagrangiano, D, para PL. Assim, o problema dual D para o problema primal P foi construído a partir da Definição B.2.1.

Corolário B.2. *Sejam x^*, uma solução factível a PL, e π^*, uma solução factível a D. Se $f(x^*) = g(\pi^*)$ então x^* e π^* são as soluções ótimas de PL e D, respectivamente.*

Demonstração: Suponha, por abrsurdo, que x^* não é solução ótima para PL e π^* não é solução ótima para D. Então, desde que PL é um problema de maximização, deve existir pelo menos um $\tilde{x}$ factível em PL tal que $f(x^*) < f(\tilde{x})$. Como, por hipótese, $f(x^*) = g(\pi^*)$ então, $g(\pi^*) < f(\tilde{x})$. Mas isto é um absuro, pois, pelo Corolário (B.1), $f(x) \leq g(\pi)$ para todo x factível e $\pi \in \mathbb{R}^m$; em particular, também é válido para $\tilde{x}$ e $\pi^* : f(\tilde{x}) \leq g(\pi^*)$. ■

Teorema B.4. *(Dualidade Forte - Programação Linear) As soluções x^*, factível a PL, e π^*, factível a D, são ótimas, primal e dual respectivamente, se e somente se satisfazem $f(x^*) = g(\pi^*)$.*

Demonstração:

- A volta do Teorema B.4 é garantida pelo Corolário B.2: se $f(x^*) = g(\pi^*)$ então x^*, factível a PL, e π^*, factível a D, são soluções ótimas, primal e dual, respectivamente.

- Sejam as respectivas soluções ótimas, primal e dual, x^* e π^*.

  Do Corolário B.1, o valor da função objetivo de D será sempre maior ou igual ao valor da função objetivo de P. Pela Obsevação 1, para que D não forneça tais limitantes superiores sendo inócuos, deve-se ter $c_i - {\pi^*}^T a_i \leq 0,\ i = 1, \ldots, n$. Mas isto implica, pela Observação 2, que $(c_i - {\pi^*}^T a_i)x_i^* = 0$. Denotando $\rho_i = c_i - {\pi^*}^T a_i$, segue que:

  $$\rho_1 x_1^* = 0, \rho_2 x_2^* = 0 \ldots \rho_n x_n^* = 0$$

  Portanto, pode-se concluir que, desde que por hipótese, x^* e π^* são soluções ótimas, primal e dual respectivamente, então:

  $$\begin{aligned} &Ax^* = b,\ x \geq 0 && (x^* \text{ é factível primal}) \\ &A^T\pi^* + \rho^T = c,\ \rho \geq 0 && (\pi^* \text{ é factível dual}) \\ &\rho_i x_i^* = 0 && (\text{folgas complementares satisfeitas}) \end{aligned}$$

onde

$$\begin{aligned}\rho_i x_i^* = 0 &\Rightarrow (c^T - {\pi^*}^T A)x^* = 0 \\ &\Rightarrow c^T x = {\pi^*}^T A x^* \\ &\Rightarrow f(x^*) = {\pi^*}^T b \\ &\Rightarrow f(x^*) = g(\pi^*)\end{aligned}$$

■

Autorizo a reprodução xerográfica para fins de pesquisa.

São José do Rio Preto, _____/_____/____

_____
Assinatura