

CARLOS ALBERTO DA SILVA JUNIOR

**UMA CONTRIBUIÇÃO AO ESTUDO DA
PROGRAMAÇÃO LINEAR**



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA MECÂNICA

**Uberlândia - MG
2008**

CARLOS ALBERTO DA SILVA JUNIOR

**UMA CONTRIBUIÇÃO AO ESTUDO DA
PROGRAMAÇÃO LINEAR**

Dissertação apresentada ao Programa de Pós Graduação da Faculdade de Engenharia Mecânica como parte dos requisitos exigidos para obtenção do Título de **Mestre em Engenharia Mecânica**.

Área de Concentração: Mecânica dos Sólidos e Vibrações Mecânicas.

Orientadora: Profa. Dra. Sezimária de Fátima P. Saramago

Uberlândia - MG
2008

Dados Internacionais de Catalogao na Publicao (CIP)

S586c	Silva Junior, Carlos Alberto da, 1975 - Uma contribuição ao estudo da programação linear / Carlos Alberto da. Silva Junior.- 2008. 191 f. : il. Orientadora: Sezimária de Fátima P. Saramago. Dissertação (mestrado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Engenharia Mecânica. Inclui bibliografia. 1. Programação linear - Teses. 2. Água - Distribuição - Teses. I. Saramago, Sezimária de Fátima Pereira. II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Engenharia Mecânica. IV. Título. CDU: 519.852
-------	---

Elaborada pelo Sistema de Bibliotecas da UFU / Setor de Catalogação e Classificação

CARLOS ALBERTO DA SILVA JUNIOR

**UMA CONTRIBUIÇÃO AO ESTUDO DA PROGRAMAÇÃO
LINEAR**

Dissertação **APROVADA** pelo programa de Pós
Graduação em Engenharia Mecânica da Universidade
Federal de Uberlândia.

Área de Concentração: Mecânica dos Sólidos e
Vibrações Mecânicas.

Orientadora: **Profa. Dra. Sezimária de Fátima P. Saramago**

Prof. Dr. Peter Zörnig - UNB

Prof. Dr. César Guilherme de Almeida - FAMAT - UFU

Prof. Dr. Valder Steffen Júnior - FEMEC - UFU

**Uberlândia - MG
10 Fevereiro 2008**

DEDICATÓRIA

À minha esposa Carolina por todo amor, apoio e atenção que dela eu tenho.

Às minhas irmãs Carla Prado e Lígia Prado, por sempre acreditarem que eu seria capaz.

À todos os meus familiares que, de uma forma ou de outra, me ajudaram a chegar até aqui.

Aos professores Márcio Dantas e Lúcia Bonfim, pelo incentivo e por terem mostrado o quanto eu poderia ir mais longe.

À todos os outros que também acreditaram que seria possível.

À minha afilhada Nina Prado.

AGRADECIMENTOS

À CAPES por ter financiado este projeto.

À Universidade Federal de Uberlândia e, principalmente, à Faculdade de Engenharia Mecânica pela oportunidade concedida de me tornar mestre.

Aos meus colegas pela paciência e por todos os incentivos que me deram desde o início deste trabalho.

Faço aqui um agradecimento mais do que especial à minha orientadora Sezimária, por sempre confiar, acreditar e estar disposta, me atendendo em todos os momentos e estar presente ensinando, ajudando e fazendo com que eu chegasse ao fim.

À Gilvane Gonçalves Corrêa pela revisão técnico-científica deste trabalho.

SILVA JUNIOR, C. A. **UMA CONTRIBUIÇÃO AO ESTUDO DA PROGRAMAÇÃO LINEAR**. 2008. 185 f. Dissertação de Mestrado, Universidade Federal de Uberlândia, Uberlândia-MG.

Resumo

Programação Linear é bastante útil para obter soluções ótimas de problemas relacionados com várias áreas da ciência, tais como, engenharia de produção, indústria de alimentos, ciências da saúde, mineração, transporte, entre outros. Este trabalho faz uma revisão sobre o assunto, produzindo um texto didático-científico que apresenta uma fundamentação teórica consistente dos principais conceitos da programação linear. Três códigos computacionais, Lindo, Linprog e Mosek, são comparados considerando o tempo computacional, o número de iterações e os valores ótimos da função objetivo e das variáveis de projeto. O problema do planejamento ótimo do transporte de cargas rodoviárias e a otimização de uma rede de distribuição de água são formulados e solucionados, comparando os valores obtidos com os três códigos estudados e também com resultados da literatura.

Palavras Chaves: Otimização, Programação Linear, Método Simplex, Problema de Transporte, Rede de Distribuição de Água.

SILVA JUNIOR. C. A. **A CONTRIBUTION TO THE STUDY OF LINEAR PROGRAMMING**. 2008. 185 f. M. Sc. Dissertation, Universidade Federal de Uberlândia, Uberlândia-MG.

Abstract

The Linear Programming is quite useful to obtain optimal solutions for problems related to several areas of science, such as, industrial engineering, food industry, health science, mining, transport, among others. This work makes a review about linear programming, producing a didactic-scientific text that presents a consistent theoretical base of the main concepts of linear programming. In this study three computer codes, Lindo, Linprog and Mosek, are compared considering the computational time, the number of iterations, the optimal values of the objective function and the design variables. The problem of the optimal planning of goods transportation and the optimization of water distribution networks are formulated and solved, by comparing the values obtained with the three studied codes and also with results found in the literature.

Keywords: Optimization, Linear Programming, Simplex Method, Transport Problem, Water Distribution Network.

Sumário

Lista de Figuras	xii
-------------------------	------------

Lista de Tabelas	xv
-------------------------	-----------

1 Definições e Notações	21
--------------------------------	-----------

1.1 Conjuntos	21
1.2 Vetores e Matrizes	23
1.3 Matriz Inversa	26
1.4 Bases e Subespaços	27
1.5 Subespaço Afim	29
1.6 Aplicação Linear	30
1.7 Função Convexa, Máximo e Mínimo (Local e Global)	30

2 Problemas de Programação Linear	35
--	-----------

2.1 Problema Geral de Otimização	35
2.2 O Problema de Programação Linear	36
2.3 Forma Padrão de um Problema de Programação Linear	39
2.4 Função Linear Convexa por Partes	42
2.5 Problema de Otimização envolvendo Função Módulo	43
2.6 Exemplos de Problemas de Programação Linear	45
2.7 Representação Gráfica e Soluções	52

3	A Geometria da Programação Linear	57
3.1	Hiperplanos, Semi-Planos e Poliedro	57
3.2	Pontos Extremos, Vértices e Soluções Básicas Viáveis	61
3.3	Poliedro na Forma Padrão	68
3.4	Degeneração	74
3.5	Existência de Pontos Extremos	78
3.6	Condições para que Pontos Extremos sejam Ótimos	81
3.7	Representação de Poliedros Limitados	84
4	O método Simplex	87
4.1	Condição para a Existência de Pontos Ótimos	87
4.2	Desenvolvimento do Método Simplex	94
4.2.1	O Método Simplex para Problemas Degenerados	100
4.2.2	Seleção do Pivô	101
4.3	Implementação do Método Simplex	102
4.3.1	Implementação Simples	102
4.3.2	O Método Simplex Revisado	103
4.3.3	Implementação do Quadro Simplex	106
4.4	O Método Simplex Duas Fases	113
4.5	Ciclagem em Problemas de Programação Linear	120
5	O Método Simplex Dual	125
5.1	Motivação	125
5.2	O Problema Dual do Problemas Linear	129
5.3	O Teorema Fundamental para o Problema Dual	136
5.4	Uma Interpretação Econômica para o Dual	141
5.5	Uma Representação Gráfica do Problema Dual	143
6	Um Problema de Transporte Rodoviário de Cargas	145
6.1	Número de Variáveis de Projeto e de Restrições do PPL	150

6.2	Desenvolvimento Simplificado do Problema I	153
6.2.1	Simulação Numérica para o Problema de Transporte I	160
6.3	Simulação Numérica para o PPL de Transporte II	167
7	Otimização de uma Rede de Distribuição de Água	175
7.1	Descrição do Problema	176
7.2	Solução Numérica do Problema de Distribuição de Água	182
8	Conclusão	185
A	Tabelas de Símbolos	187

Lista de Figuras

1.1	<i>Representação Gráfica de uma função convexa.</i>	31
1.2	<i>Representação Gráfica de uma função côncava.</i>	31
2.1	<i>Solução Gráfica - Exemplo 2.7.1.</i>	52
2.2	<i>Representação da região viável - Exemplo 2.7.1.</i>	52
2.3	Representação das retas perpendiculares ao vetor $c^T = [-1 \ -1]$ - Exemplo 2.7.1	53
2.4	Representação da região viável - Exemplo 2.7.2.	53
2.5	Representação da região viável - Exemplo 2.7.3.	54
2.6	Representação da região viável com uma nova restrição - Exemplo 2.7.3.	55
3.1	Representação de um hiperplano e dos dois semi-espacos determinados por ele.	59
3.2	Representação de um poliedro formado pela intersecção de 5 semiplanos.	59
3.3	Representação de pontos extremos num poliedro.	61
3.4	Representação de vértice.	62
3.5	Representação do poliedro ilustrado na Definição 3.2.3	63
3.6	Representação de um poliedro ilustrando a Definição 3.2.4	65
3.7	Representação do problema na forma padrão com $n = 4$ e $m = 2$.	72
3.8	Representação de uma Solução Básica Degenerada no plano.	75
3.9	Representação de uma Solução Básica Degenerada no espaço.	75
3.10	Representação de uma restrição degenerada que tornou-se não degenerada, através de uma pequena carga.	76
3.11	Representação de solução degenerada em poliedros na forma padrão.	77
3.12	Representação de um poliedro que contém uma reta e de um que não contém.	78

3.13	Representação da interpretação geométrica do Teorema 3.5.1.	81
3.14	Representação da interpretação geométrica do Teorema 3.6.1.	82
3.15	Representação da idéia geométrica para o Teorema 3.7.1.	84
4.1	Representação da ilustração da Definição 4.1.1.	88
4.2	Representação da ilustração da direção básica em soluções degeneradas.	90
4.3	Representação de um PPL com uma solução degenerada.	101
4.4	Representação da solução viável do PPL - Exemplo 4.3.3	110
5.1	Representação de uma aplicação para o Teorema Dual no Exemplo 5.3.1.	141
5.2	Representação da região viável do problema primal e do dual - Exemplo 5.5.1. . .	144
7.1	Descrição da Rede utilizada para estudo	176

Lista de Tabelas

2.1	Característica dos cinco sistemas desenvolvidos pela DEC, em 1988	46
3.1	Procedimento para Construção de Soluções Básicas:	70
4.1	Uma iteração do Método Simplex:	98
4.2	Uma Iteração do Método Simplex Revisado.	106
4.3	A Tabela Simplex	108
4.4	A Tabela Simplex escrita de maneira detalhada.	108
4.5	Uma Iteração da Implementação da Tabela Simplex.	109
4.6	Tabela do Exemplo 4.3.3 - Passo Inicial	111
4.7	Tabela do Exemplo 4.3.3 - 1º Passo	112
4.8	Tabela Simplex do Exemplo 4.3.3 - 2º Passo	113
4.9	Tabela Simplex do Exemplo 4.3.3 - 3º Passo	113
4.10	Uma Iteração do Método Simplex Duas Fases:	116
4.11	Tabela Inicial Fase 01 do Exemplo 4.4.1	118
4.12	Tabela do Exemplo 4.4.1 - 1º Passo Fase 01	118
4.13	Tabela do Exemplo 4.4.1 - 2º Passo Fase 01	119
4.14	Tabela do Exemplo 4.4.1 - 3º Passo Fase 01	119
4.15	Tabela Simplex Inicial Fase 02 - Exemplo 4.4.1	119
4.16	Tabela Simplex Inicial Fase 02 Completa - Exemplo 4.4.1	120
4.17	Primeira Tabela do PPL - Exemplo 4.5.1	121
4.18	Segunda Tabela do PPL - Exemplo 4.5.1	121
4.19	Terceira Tabela do PPL - Exemplo 4.5.1	121

4.20	Quarta Tabela do PPL - Exemplo 4.5.1	122
4.21	Quinta Tabela do PPL - Exemplo 4.5.1	122
4.22	Sexta Tabela do PPL - Exemplo 4.5.1	122
4.23	Sétima Tabela do PPL - Exemplo 4.5.1	123
5.1	Representação do Problema Primal e do seu Dual, de um PPL	130
5.2	Resumo apresentando a relação entre o Problema Primal e o seu Dual	130
6.1	Dados de entrada para o Problema de Transporte I (Simplificado)	160
6.2	Resultados obtidos para o Problema de Transporte I	165
6.3	Vetor X_w obtido na solução do Problema de Transporte I.	166
6.4	Vetor X_t obtido na solução do Problema de Transporte I.	166
6.5	Vetor X_z obtido na solução do Problema de Transporte I.	166
6.6	Vetor X_x obtido na solução do Problema de Transporte I.	167
6.7	Vetor X_y obtido na solução do Problema de Transporte I.	167
6.8	Dados de entrada para o Problema de Transporte II (271 Variáveis)	168
6.9	Resultados obtidos para o Problema de Transporte II	171
6.10	Soluções não nulas obtidas para a variável w (Problema II)	172
6.11	Soluções obtidas para a variável t (Problema II)	173
6.12	Soluções obtidas para a variável y (Problema II)	173
7.1	Dados relativos aos nós da rede	176
7.2	Dados relativos aos nós da rede	177
7.3	Diâmetros comerciais utilizado na programação linear	178
7.4	Coeficientes $J_{i,k}$ para o PPL	179
7.5	Preço e velocidade máxima na tubulação segundo seus diâmetros	179
7.6	Resultados obtidos para a rede de distribuição de água.	182
7.7	Resultados obtidos para a rede de distribuição de água.	183
7.8	Resultados obtidos para a rede de distribuição de água.	184

Introdução

A Programação Linear surgiu na última metade do século vinte, sendo uma técnica de otimização significativamente completa, apresentando códigos capazes de resolver problemas com grande número de restrições e variáveis de projeto. O problema de otimizar uma função linear sujeita a restrições lineares teve a sua origem nos estudos de Fourier sobre Sistemas de Inequações Lineares, em 1826. No entanto, apenas em 1939 Kantorovich mostrou a importância prática destes problemas, apresentando um algoritmo para sua resolução. Além de apresentar um documento com fundamentação teórica, Kantorovich apresentou exemplos para a aplicação da Programação Linear, sendo que a idéia fundamental para cada exemplo era maximizar a produção utilizando de forma ótima os recursos disponíveis. Infelizmente, durante vários anos, o trabalho de Kantorovich foi insuficientemente conhecido por toda a Europa. A divulgação deste trabalho no ocidente apenas ocorreu depois de 1950. O problema de otimizar uma função linear, sujeita a restrições lineares, teve o seu auge com George Dantzig na década de 1940, consultor de matemática do US Air Force Comptroller, e com o prêmio Nobel da Economia George Stigler, que formulou o problema das dietas como um problema de mistura de componentes. Dantzig não só formulou o problema de programação linear, mas também criou o Algoritmo Simplex para a solução do mesmo, em 1947. Ainda em 1947, Koopmans mostra que a programação linear é um modelo apropriado para a análise da teoria econômica clássica. Entretanto, nos EUA, Frank L. Hitchcock apresentou o que é hoje a formulação base do problema de transporte. Independentemente, o professor Koopmans formula o mesmo problema relacionando-o com o seu trabalho na "Combined Shipping Adjustment Board". Por isso, o problema de transporte é referido, na literatura científica, como o problema de transporte de Hitchcock ou como problema de transporte de Hitchcock-Koopmans. Em 1956, Alex Orden propôs a generalização do modelo de transporte em que eram permitidos pontos de transbordo de carga. Esta formulação é conhecida hoje como um problema de transbordo sem limite de capacidade. Ao mesmo tempo, o problema de fluxos máximos e o problema do fluxo do custo mínimo em rede foram formulados e investigado pela famosa equipe de Lester Ford e Delbert Fulkerson. Entre 1950 a 1965 muita atividade foi dirigida para o desenvolvimento de algoritmos

para modelos de programação linear em rede. Os algoritmos desenvolvidos podem ser classificados em duas classes: 1) Especialização do Método Simplex; 2) Método Primal-Dual. A especialização primal simplex começou com o trabalho de Dantzig e atingiu o apogeu com o documento de Ellis Johnson. O método primal-dual, originado com o algoritmo do húngaro Harold Kuhn para o problema de atribuição, é concluído com o algoritmo da condição de Delbert Fulkerson em 1961. Como resultado desta atividade de investigação, desenvolveu-se uma eficiente implementação das técnicas básicas, a particularização dos códigos de programação linear em rede, melhorou o desempenho dos procedimentos para a resolução de problemas de programação linear em rede. O primeiro problema de tamanho considerável, resolvido pelo Algoritmo Simplex, foi o das dietas de Stigler com nove equações e setenta e sete variáveis não negativas, gastando 120 horas nas calculadoras existentes na época e menos de um segundo nos atuais computadores pessoais. Em 1975, a Academia Real de Ciência atribuiu o prêmio Nobel da Ciência em Economia a Kantorovich e Koopmans pelas suas contribuições para a teoria da alocação de recursos, considerando a contribuição de Dantzig mais no âmbito matemático. No entanto, Danzig permanecerá na história da formulação da programação linear como um de seus arquitetos fundamentais. Outras informações pode ser obtida em Sousa (2003).

Atualmente, são estudados na programação linear, problemas ligados a diversos setores. Por exemplo, na mineração Costa e Pinto (2005) determinam o ritmo de lavra de cada frente, considerando a alocação de equipamentos de carga e transporte, de modo a fornecer à usina de beneficiamento uma alimentação adequada. Problemas ligados à indústria de alimentos, como visto em Melo et al. (2004), onde são construídos modelos em programação linear e não-linear visando otimizar o processo de fabricação do biscoito cracker. Podem ser aplicados a problemas médicos, como o apresentado por Barboza e Oliveira (2006), que desenvolveu o estudo na implementação de métodos de pontos interiores específicos para o problema de planejamento do tratamento de câncer por radioterapia. Outro aspecto importante é o estudo do fenômeno de ciclagem em problemas de programação linear, como visto em Zörnig (2006a) e Zörnig (2006b), onde são apresentadas condições necessárias e suficientes para a ciclagem. Esta teoria apresenta a construção sistemática de exemplos de ciclagem. Estas são algumas, dentre tantas outras, aplicações da programação linear.

Para resolver problemas de programação linear o uso de códigos comerciais é de grande utilidade. Existem vários códigos livres quanto comerciais, como por exemplo o MOSEK (Mathematical Optimization Software Specified, Mosek (2006)) e o Linprog (Linear Programming), ambos utilizando o Software MatLab, MathWorks (2006), o código LINDO (Linear, Interactive and Discrete Optimizer, Prado (1999)), o LINGO, o CVX, o ProLin V1.0, entre outros. É importante ressaltar, também, que a programação linear é a fundamentação básica de alguns métodos diretos (que trabalham diretamente com as restrições) da programação não linear como, por exemplo, o método das direções viáveis ou o método da programação linear seqüencial (SLP).

Vários métodos de otimização estão sendo estudados na Universidade Federal de Uberlândia, por pesquisadores do programa de pós-graduação em Engenharia Mecânica nos últimos 20 anos. A seguir alguns destes trabalhos são citados: Faria (1991) apresenta uma contribuição aos procedimentos de otimização aplicados a sistemas mecânicos; Assis (1993) faz a otimização multi-critérios de máquinas rotativas; Braga (1998) apresenta um estudo sobre o uso de algoritmos genéticos para aplicação em problemas de otimização de sistemas mecânicos; a otimização de sistemas mecânicos, como uma ferramenta de engenharia de concepção é apresentada por Choze (1998); a otimização de trajetórias de robôs manipuladores na presença de obstáculos foi estudado por Saramago (1998); Pereira (2003) estuda técnicas de otimização discreto-contínuas aplicadas ao controle de vibrações em estruturas inteligentes; Oliveira (2005) faz uma contribuição ao estudo dos métodos de otimização multi-objetivo; Oliveira (2006) apresenta um estudo, com aplicações, da evolução diferencial; Santana (2007) trabalha com modelagem numérica e otimização de Shunts Piezelétricos aplicados ao controle passivo de vibrações, entre outros. Porém, ainda não tinha sido feito um trabalho relacionado à programação linear. Como este é um assunto importante, este estudo procurará fazer uma contribuição sobre o assunto.

Este trabalho apresenta uma revisão sobre programação linear, fazendo uma fundamentação teórica consistente do assunto. Um dos objetivos deste estudo é a produção de um texto didático-científico, de forma a contribuir para a produção bibliográfica em língua portuguesa sobre o tema. Neste sentido, o estudo da obra de Bertsimas e Tsitsiklis (1997) foi essencial, pois trata-se de um texto que preocupa-se em apresentar os conceitos de forma completa e bem fundamentada. Outra contribuição desta pesquisa é comparar três códigos: Lindo, Linprog e Mosek, escolhidos

por serem de fácil acesso e simples implementação. Com estas finalidades são estudados dois problemas práticos: o planejamento ótimo de transporte de cargas rodoviárias e a otimização de uma rede de distribuição de água. Os resultados obtidos são comparados considerando o tempo computacional, o número de iterações e os valores ótimos encontrados para a função objetivo e suas variáveis de projeto.

De forma a facilitar a compreensão do leitor, o texto é organizado da maneira a seguir:

✓ **Capítulo 01:** é apresentado os conceitos iniciais sobre Teoria de Conjuntos e Álgebra Linear, além de mostrar a definição de aplicação linear, funções côncavas e convexas;

✓ **Capítulo 02:** faz uma breve introdução aos problemas de Programação Linear; apresentando as definições e teoremas, junto com alguns exemplos que são importantes para o desenvolvimento do conhecimento;

✓ **Capítulo 03:** trata das propriedades geométricas básicas do Poliedro, focando-se na definição de pontos extremos e de sua existência, e também enfatizando a interação entre os pontos de vistas algébricos e geométricos;

✓ **Capítulo 04:** completa o estudo do Método Simplex, apresentado um algoritmo que representa uma iteração do método, também é apresentado a Tabela Simplex, problemas relacionados com a escolha do Pivô, método simplex revisado e das duas fases;

✓ **Capítulo 05:** trata do Método Simplex Dual, fazendo a construção do dual através do primal e apresentando os conceitos básicos para a formulação do problema;

✓ **Capítulo 06:** apresenta um problema relacionado a transporte de determinados tipos de demandas, onde são estudados todos os detalhes do problema e onde também são feitas comparações das soluções obtidas pelos códigos em estudo;

✓ **Capítulo 07:** apresenta um problema relacionado a redes de distribuição de água, comparando com os resultados obtidos com os códigos estudados e com a literatura.

Para encerrar são apresentadas as conclusões do estudo, como também apontamentos para novas pesquisas.

Capítulo 1

Definições e Notações

Na primeira parte deste trabalho são apresentadas as notações, definições e, até mesmo, são apresentados alguns teoremas simples envolvendo a teoria de conjuntos, matrizes, determinantes e álgebra linear, que são necessários para desenvolver a base deste trabalho. É claro que não é apresentado um curso completo da teoria de conjuntos, nem mesmo de álgebra linear, porém acredita-se que os conceitos necessários para um bom entendimento do texto são apresentados nesta introdução.

Como um dos objetivos deste trabalho é apresentar um texto que sirva de referência para os mais diversos tipos de leitores, o excesso de detalhe apresentado neste capítulo não pode ser omitido.

1.1 Conjuntos

Um conjunto é uma reunião de todos os elementos que apresentaram, ao mesmo tempo, um certo tipo de característica. Por exemplo, se é tomado $\mathbf{S} = \{a, e, i, o, u\}$ tem-se o conjunto das vogais do alfabeto arábico. Por convenção, todo conjunto é representado por uma letra maiúscula. Neste texto, geralmente, é denotado por $\mathbf{S}$, e qualquer um dos seus elementos é representado por uma letra minúscula, geralmente por x . Quase sempre, como se trabalha com conjuntos numéricos, não é possível listar todos os elementos como foi feito acima, por isto, pode-se representar um conjunto por sua lei de formação, no exemplo, usando esta notação, obtém-se o conjunto:

$\mathbf{S} = \{x | x \text{ seja vogal do alfabeto arábico}\}$. Quando um conjunto tem um número finito de elementos representa-se a sua cardinalidade (o número de seus elementos) por $|\mathbf{S}|$. Entre as operações envolvendo conjuntos, a união, intersecção e a diferença necessitam de referência.

Definição 1.1.1. Diz-se que um conjunto é a união de $\mathbf{S}$ com $\mathbf{T}$, e escreve-se $\mathbf{S} \cup \mathbf{T}$, se ele é o conjunto formado por todos os elementos que ou pertencem a $\mathbf{S}$ ou pertencem a $\mathbf{T}$, isto é, $\mathbf{S} \cup \mathbf{T} = \{x | x \in \mathbf{S} \text{ ou } x \in \mathbf{T}\}$.

Definição 1.1.2. Diz-se que um conjunto é a intersecção de $\mathbf{S}$ com $\mathbf{T}$, e escreve-se $\mathbf{S} \cap \mathbf{T}$, se ele é o conjunto formado por todos os elementos que pertencem a $\mathbf{S}$ e pertencem a $\mathbf{T}$ ao mesmo tempo, isto é, $\mathbf{S} \cap \mathbf{T} = \{x | x \in \mathbf{S} \text{ e } x \in \mathbf{T}\}$.

Definição 1.1.3. Diz-se que um conjunto é a diferença de $\mathbf{S}$ com $\mathbf{T}$, e escreve-se $\mathbf{S} \setminus \mathbf{T}$, se ele é o conjunto formado por todos os elementos que pertencem a $\mathbf{S}$ e não pertencem a $\mathbf{T}$, isto é, $\mathbf{S} \setminus \mathbf{T} = \{x | x \in \mathbf{S} \text{ e } x \notin \mathbf{T}\}$.

Muitas vezes trabalha-se com conjuntos que representam partes de um dado conjunto, por isto, a notação $\mathbf{S} \subseteq \mathbf{T}$ representa o fato que $\mathbf{S}$ é um *subconjunto* de $\mathbf{T}$, ou seja, todo elemento de $\mathbf{S}$ também é elemento de $\mathbf{T}$. Observe que a igualdade acontece quando todos os elementos de $\mathbf{T}$ também são elementos de $\mathbf{S}$. Caso a igualdade não aconteça, diz-se que $\mathbf{S}$ é um *subconjunto próprio* de $\mathbf{T}$, e a notação usada é $\mathbf{S} \subset \mathbf{T}$. Ainda existem conjuntos que não possuem elementos, e a notação usada para representar este conjunto é $\emptyset$. Tem-se também que os símbolos $\exists$ e $\forall$ representam “existe” e “para todo”, respectivamente.

O conjunto dos números reais é representado por $\mathbb{R}$, sendo que quando é necessário representar apenas um intervalo, usa-se $\mathbf{S} = [a, b]$ para representar um conjunto fechado $\mathbf{S} = \{x \in \mathbb{R} | a \leq x \leq b\}$ e usa-se $\mathbf{S} = (a, b)$ para representar um conjunto aberto $\mathbf{S} = \{x \in \mathbb{R} | a < x < b\}$. Se existe a necessidade de representar um intervalo misto, pode-se usar as duas notações em

conjunto. Por fim, o conjunto de todas as n -uplas $(x_1, x_2, \dots, x_n)$ de números reais é o conjunto $\mathbb{R}^n$.

1.2 Vetores e Matrizes

Um *vetor* é definido como sendo qualquer segmento de reta orientado. Logo, é fácil ver que todo vetor possui uma magnitude, também chamada de norma, uma direção e um sentido. Neste texto é feito um estudo algébrico dos vetores, por isto os vetores são representados por coordenadas, ou seja, pensa-se que cada vetor v tem como origem o ponto $\mathbf{0} = (0, 0, \dots, 0)$ e como extremidade o ponto $(x_1, x_2, \dots, x_n)$, daí conclui-se que cada ponto do $\mathbb{R}^n$ representa um vetor.

Como os vetores estão diretamente relacionados com matrizes, a seguinte notação é adotada:

$$v = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad (1.1)$$

e para o vetor transposto:

$$v^T = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}. \quad (1.2)$$

Uma *matriz* é uma tabela ordenada onde cada elemento possui uma posição de linha e uma posição de coluna. A *dimensão*, ou a *ordem*, de uma matriz é dada pelo seu número de linhas juntamente com o seu número de colunas. Portando, uma matriz $\mathcal{A}_{m,n}$ é uma matriz que possui m linhas e n colunas, e ela é representada de uma maneira genérica da seguinte forma:

$$\mathcal{A}_{m,n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \quad (1.3)$$

e a sua ordem é $m \times n$.

Uma matriz que tem apenas uma linha é chamada de *matriz linha* e a matriz que possui apenas uma coluna é chamada de *matriz coluna*, daí conclui-se que um vetor é representado por uma matriz coluna. Quando $m = n$, diz-se que a matriz é uma *matriz quadrada*, e por simplicidade diz-se que a ordem da matriz é n . Uma matriz sempre é representada por letras maiúsculas, enquanto um vetor é representado por uma letra minúscula em itálico. Outra maneira de se representar uma matriz é pela notação $[A]_{ij}$, onde i representa a linha do elemento e o j representa a coluna. O vetor $\mathbf{0}$ é chamado de *vetor nulo*, enquanto que a matriz

$$\mathcal{O} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}, \quad (1.4)$$

é chamada de *matriz nula*. O i -ésimo *vetor unitário* é o vetor onde a i -ésima coordenada vale 1 e as outras coordenadas valem zero, assim,

$$\mathbf{e}_i^T = \begin{bmatrix} 0 & \cdots & 0 & \underbrace{1}_{i\text{-ésima posição}} & 0 & \cdots & 0 \end{bmatrix}. \quad (1.5)$$

Por outro lado, a *matriz identidade* é a matriz cujos elementos da diagonal principal valem 1 (elementos da diagonal principal são os elementos onde $i = j$) e os demais elementos valem zero. Para exemplificar, a matriz identidade de ordem 4, que é representada por $\mathcal{I}_4$, é a matriz:

$$\mathcal{I}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (1.6)$$

Outro conceito muito usado, é o de *matriz transposta*. Define-se a transposta de uma matriz $\mathcal{A}$, e escreve-se $\mathcal{A}^T$, como sendo a matriz onde cada linha da primeira é representada pela coluna da segunda, isto é, $[\mathcal{A}^T]_{ij} = [\mathcal{A}]_{ji}$. Logo, pode-se observar facilmente que na Eq. (1.2) foi usado a notação de transposta de uma matriz para o vetor v .

Agora, se x e y são dois vetores no $\mathbb{R}^n$, então deve-se tomar o *produto interno* entre x e y como sendo o produto matricial da matriz linha x^T com a matriz coluna y , ou seja:

$$\langle x, y \rangle = x^T \cdot y = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^n x_i y_i. \quad (1.7)$$

A notação usada para representar o produto interno é $\langle x, y \rangle$.

Se dois vetores são *ortogonais*, então tem-se que o seu produto interno vale zero. É fácil perceber que $\langle x, x \rangle \geq 0$, e a igualdade ocorre se, e somente se, $x = 0$. Para calcular a *norma* de um vetor x , que é representado por $\|x\|$, é usada a norma euclidiana, ou seja, é usado a norma

$$\|x\| = \sqrt{\langle x, x \rangle}. \quad (1.8)$$

Uma desigualdade também muito utilizada é a *desigualdade de Schwartz* (para ver uma demonstração desta desigualdade consulte Coelho e Lourenço (2005)), sendo dada por:

$$|x^T y| \leq \|x\| \|y\|. \quad (1.9)$$

com igualdade acontecendo se, e somente se, um dos vetores é múltiplo escalar do outro.

Muitas vezes é necessário falar de uma determinada coluna de uma matriz ou de uma determinada linha, por isto estabelece-se a seguinte notação: quando estiver sendo feita alguma referência a j -ésima coluna de uma matriz $\mathcal{A}$, escreve-se $\mathcal{A}_j$, e quando estiver sendo feita referência a i -ésima linha de $\mathcal{A}$, escreve-se a_i .

Para finalizar, diz-se que um vetor $x \geq 0$ (respectivamente, $x > 0$) quando $x_i \geq 0$ (respectivamente $x_i > 0$), $\forall i \in \{1, \dots, n\}$. A mesma notação é usada para matrizes, ou seja, $\mathcal{A} \geq 0$ (respectivamente, $\mathcal{A} > 0$) quando $a_{ij} \geq 0$ (respectivamente $a_{ij} > 0$), $\forall i \in \{1, \dots, m\}$ e $\forall j \in \{1, \dots, n\}$. Outros detalhes sobre o estudo de matrizes são dados em Hoffman e Kunze (1979).

1.3 Matriz Inversa

Uma matriz quadrada de ordem n é dita ser invertível, se existe uma matriz $\mathcal{B}$, também de ordem n , tal que $\mathcal{A} \cdot \mathcal{B} = \mathcal{B} \cdot \mathcal{A} = \mathcal{I}$. Tal matriz $\mathcal{B}$ é chamada de a *inversa* de $\mathcal{A}$ e é representada por $\mathcal{A}^{-1}$. É fácil ver que $(\mathcal{A}^{-1})^{-1} = \mathcal{A}$, e que $(\mathcal{A}\mathcal{B})^{-1} = \mathcal{B}^{-1} \cdot \mathcal{A}^{-1}$. Os métodos de se obter a matriz inversa pode ser encontrado em Hoffman e Kunze (1979).

Diz-se que os vetores $x_1, x_2, \dots, x_t$ são *Linearmente Dependentes* (LD) se existem escalares $a_1, a_2, \dots, a_n$, nem todos nulos, tais que

$$\sum_{j=1}^t a_j x_j = \mathbf{0}, \quad (1.10)$$

caso contrário eles são chamados de *Linearmente Independentes* (LI).

Seja $\mathcal{A}$ uma matriz invertível e quadrada. Usando as fórmulas de Cramer, que pode ser visto em Hoffman e Kunze (1979), é fácil obter a solução do sistema

$$\mathcal{A}x = b \quad (1.11)$$

pois

$$x = \mathcal{A}^{-1}b. \quad (1.12)$$

Existe uma ampla bibliografia disponível a respeito de solução de sistemas lineares, por exemplo Hoffman e Kunze (1979), Coelho e Lourenço (2005), Iezzi e Hazzan (2001).

1.4 Bases e Subespaços

Um conjunto $\mathbf{S}$ do $\mathbb{R}^n$ é chamado de *subespaço de $\mathbb{R}^n$* , se $\forall x, y \in \mathbf{S}$ e cada $a, b \in \mathbb{R}$ têm-se que $ax + by \in \mathbf{S}$. Se $\mathbf{S} \neq \mathbb{R}^n$, então diz-se que $\mathbf{S}$ é um *subespaço próprio* do $\mathbb{R}^n$. Uma observação importante é que qualquer subespaço obrigatoriamente possui o vetor nulo, pois basta tomar a e b como sendo zero.

Um conjunto $\mathbf{S}$ do $\mathbb{R}^n$ é dito ser *um subespaço finitamente gerado* pelos vetores $x_1, \dots, x_t$, e usa-se a notação $\langle \mathbf{S} \rangle$, se todos os vetores de S são da forma:

$$\sum_{j=1}^t a_j x_j, \quad (1.13)$$

onde cada a_j é um número real qualquer, ou seja, cada vetor $y \in \mathbf{S}$ é uma combinação dos vetores $x_1, \dots, x_t$, e por isto o vetor y é chamado de *uma combinação linear* de $x_1, \dots, x_t$. Para completar esta definição, é apresentado um dos conceitos mais importante sobre subespaços, que é o conceito de base. Uma *base* de um subespaço $\mathbf{S}$ do $\mathbb{R}^n$, com $\mathbf{S} \neq \{\mathbf{0}\}$, é uma coleção de vetores LI $\{x_1, \dots, x_t\}$, que gere também um subespaço igual a $\mathbf{S}$, ou seja,

$$\mathbf{S} = \langle \{x_1, \dots, x_t\} \rangle. \quad (1.14)$$

Qualquer base de um subespaço dado tem o mesmo número de vetores LI e como neste trabalho é considerado apenas subespaço finitamente gerado, este número de vetores é chamado de *dimensão do subespaço*. Para exemplificar isto, tem-se que a dimensão do espaço $\mathbb{R}^n$ é n e, por isto, todo subespaço próprio do $\mathbb{R}^n$ tem dimensão menor do que n . Pode-se tomar qualquer reta que passa pela origem como sendo um subespaço próprio do $\mathbb{R}^n, n > 1$. Facilmente pode-se

observar que o conjunto $\mathbf{S} = \{\mathbf{0}\}$ é um subespaço, e ele é chamado de *subespaço nulo* e a sua dimensão, por definição, é zero. Se $\mathbf{S}$ é um subespaço próprio do $\mathbb{R}^n$, então existe um vetor não nulo $\mathbf{a}$ que é ortogonal a $\mathbf{S}$, isto é, $\mathbf{a}^T \mathbf{x} = 0, \forall \mathbf{x} \in \mathbb{R}^n$. Generalizando, se $\mathbf{S}$ tem dimensão $m < n$, então existem $n - m$ vetores LI que são ortogonais a $\mathbf{S}$. Nas referências Coelho e Lourenço (2005) e Hoffman e Kunze (1979) pode-se encontrar um estudo completo envolvendo bases e subespaços.

Teorema 1.4.1. *Suponha que o subespaço $\mathbf{S}$ gerado pelos vetores $\mathbf{x}_1, \dots, \mathbf{x}_t$, tenha dimensão m , então:*

- a) existe uma base de $\mathbf{S}$ constituída de m dos vetores $\mathbf{x}_1, \dots, \mathbf{x}_t$;*
- b) se $t < m$ e $\mathbf{x}_1, \dots, \mathbf{x}_t$ são LI, pode-se formar uma base de $\mathbf{S}$ iniciada por $\mathbf{x}_1, \dots, \mathbf{x}_t$ e acrescentando $m - t$ vetores $\mathbf{x}_{t+1}, \dots, \mathbf{x}_m$.*

Demonstração 1.4.1. *a) Suponha que $t > m$ e seja $\mathbf{x}_{i_1}$ um dos vetores geradores de $\mathbf{S}$. Se não existe outro vetor que seja LI com este tem-se que todos os vetores $\mathbf{v}$ de $\mathbf{S}$ são da forma $\mathbf{v} = a_1 \mathbf{x}_{i_1}$, onde a_1 é uma constante, logo a dimensão de $\mathbf{S}$ é 1. Se $m \neq 1$, então existe um vetor $\mathbf{x}_{i_2}$, que é LI com $\mathbf{x}_{i_1}$, e supondo que não existe outro vetor que seja LI com estes, tem-se que todos os vetores $\mathbf{v}$ de $\mathbf{S}$ são da forma $\mathbf{v} = a_1 \mathbf{x}_{i_1} + a_2 \mathbf{x}_{i_2}$, onde $a_i, i \in \{1, 2\}$ são constantes, e conseqüentemente $m = 2$. Logo, repetindo este argumento m vezes tem-se que existe m vetores LI de $\mathbf{x}_1, \dots, \mathbf{x}_t$ que geram $\mathbf{S}$, logo estes formam uma base para $\mathbf{S}$. Se $t = m$, então tem-se que m vetores geram um subespaço $\mathbf{S}$ de dimensão m , logo estes vetores são LI pois, caso contrário, eles não geram $\mathbf{S}$. Por isto, tem-se que $\mathbf{x}_1, \dots, \mathbf{x}_m$ é uma base para $\mathbf{S}$. Agora se $t < m$, aplica-se (b) e o resultado estará completo.*

b) Como $\mathbf{x}_1, \dots, \mathbf{x}_t$ são LI e $m > t$ tem-se que $\mathbf{S}$ não é gerado por $\mathbf{x}_1, \dots, \mathbf{x}_t$ pois, caso contrário, a dimensão de $\mathbf{S}$ seria t . Então, existe um vetor $\mathbf{y}_1$ em $\mathbf{S}$ tal que o conjunto $\{\mathbf{x}_1, \dots, \mathbf{x}_t, \mathbf{y}_1\}$ seja LI, logo se $t + 1 = m$, então este conjunto é uma base para $\mathbf{S}$. Caso contrário, existe um vetor $\mathbf{y}_2$ em $\mathbf{S}$ tal que o conjunto $\{\mathbf{x}_1, \dots, \mathbf{x}_t, \mathbf{y}_1, \mathbf{y}_2\}$ seja LI. Se $t + 2 = m$, então este conjunto é uma base para $\mathbf{S}$. Caso contrário, pode-se repetir o argumento $m - t$ vezes até encontrar a base desejada, o que termina a demonstração.

□

Pensando em forma de matriz pode-se tomar as seguintes definições: o *espaço coluna* de uma matriz $\mathcal{A}$ de dimensão $m \times n$ é o subespaço do $\mathbb{R}^m$ gerado pelas colunas de $\mathcal{A}$. De maneira análoga, tem-se que o *espaço linha* de uma matriz $\mathcal{A}$ de dimensão $m \times n$ é o subespaço do $\mathbb{R}^n$ gerado pelas linhas de $\mathcal{A}$. A dimensão do espaço linha e o espaço coluna é a mesma, sendo este fato demonstrado em Hoffman e Kunze (1979), e este número é chamado de *posto* da matriz $\mathcal{A}$. É obvio que $\text{posto}(\mathcal{A}) \leq \{m, n\}$.

Para finalizar, é dado a definição de núcleo e imagem: o conjunto $\mathbf{Nuc} = \{x \in \mathbb{R}^m | \mathcal{A}x = 0\}$ é chamado de *Núcleo* de $\mathcal{A}$ e o conjunto $\mathbf{Im} = \{y \in \mathbb{R}^n | \mathcal{A}x = y\}$ é chamado de *Imagem* de $\mathcal{A}$. Tem-se que a dimensão do posto de $\mathcal{A}$ é a dimensão da imagem de $\mathcal{A}$ e, pelo teorema do núcleo e da imagem tem-se que $\text{dimensão}(\mathbf{Nuc}) = m - \text{posto}(\mathcal{A})$, em Coelho e Lourenço (2005) esta demonstração é apresentada.

1.5 Subespaço Afim

Seja $\mathbf{S}^0$ um subespaço do $\mathbb{R}^n$ e seja x_0 algum vetor também do $\mathbb{R}^n$. Se é tomado o conjunto

$$\mathbf{S} = \mathbf{S}^0 + x_0 = \{x + x_0, \forall x \in \mathbf{S}^0\}, \quad (1.15)$$

que geralmente não é um subespaço, visto que o vetor nulo nem sempre pertence a este conjunto, então ele é chamado de *subespaço afim* de $\mathbf{S}^0$. A dimensão do subespaço afim $\mathbf{S}$ é definida como sendo a mesma do subespaço $\mathbf{S}^0$.

Só para ilustrar, seja $x_0, x_1, \dots, x_t$ alguns vetores no $\mathbb{R}^n$. Considere o conjunto $\mathbf{S}$ formado por todos os vetores v da forma

$$v = x_0 + \lambda_1 x_1 + \dots + \lambda_t x_t \quad (1.16)$$

onde $\lambda_1, \dots, \lambda_t$ são escalares arbitrários. Se $\mathbf{S}^0 = \langle x_1, \dots, x_t \rangle$, tem-se que $\mathbf{S}$ é o subespaço afim $\mathbf{S}^0 + x_0$, finitamente gerado. Se a dimensão de $\mathbf{S}^0$ é k então a dimensão de $\mathbf{S}$ também é k .

1.6 Aplicação Linear

Seja f uma aplicação de $\mathbf{S}$ em $\mathbf{T}$. Então, se a aplicação f satisfaz a relação

$$f(ax + by) = af(x) + bf(y), \forall x, y \in \mathbb{R}^n, a, b \in \mathbb{R} \quad (1.17)$$

ela é dita ser uma *aplicação linear*. Pode-se pensar que em uma aplicação linear operações no domínio levam as mesmas operações na imagem, por isto, pode-se escolher fazê-las “antes” ou “depois” de se efetuar a aplicação.

Para exemplificar isto, tome como exemplo a aplicação $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ dada por

$$f(x, y, z) = x + y + z. \quad (1.18)$$

É fácil ver que esta aplicação é linear, pois se $a \in \mathbb{R}$, $x = (x_1, y_1, z_1) \in \mathbb{R}^3$ e $y = (x_2, y_2, z_2) \in \mathbb{R}^3$ então,

$$\begin{aligned} f(ax + y) &= f(ax_1 + x_2, ay_1 + y_2, az_1 + z_2) = (ax_1 + x_2) + (ay_1 + y_2) + (az_1 + z_2) = \\ &= a(x_1 + y_1 + z_1) + (x_2 + y_2 + z_2) = af(x_1, y_1, z_1) + f(x_2, y_2, z_2) = af(x) + f(y). \end{aligned} \quad (1.19)$$

Logo, f é uma aplicação linear.

1.7 Função Convexa, Máximo e Mínimo (Local e Global)

No estudo dos métodos de otimização, muitas vezes é necessário trabalhar com um tipo especial de função, que são as chamadas *funções convexas*. Por isto, é apresentada agora, a definição formal da mesma.

Definição 1.7.1. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função. Assim:*

a) *ela é chamada de **função convexa** se, para cada $x, y \in \mathbb{R}^n$ e para todo $\lambda \in [0, 1]$, tem-se que*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad (1.20)$$

b) ela é chamada de **função côncava** se, para cada $x, y \in \mathbb{R}^n$ e para todo $\lambda \in [0, 1]$, tem-se que

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y). \quad (1.21)$$

Note que se $x, y \in \mathbb{R}^n$ e se λ varia em $[0, 1]$, então os pontos da forma $\lambda x + (1 - \lambda)y$ são os pontos que pertencem ao segmento de reta com extremidades x e y . É fácil ver que se f é uma função linear a desigualdade da parte (a) da definição deve ser tomada como igualdade. A desigualdade, portanto, quer dizer que quando se restringe a atenção aos segmentos de reta ligando qualquer dois pontos da imagem da função, o gráfico da referida função está sempre abaixo do gráfico que representa o segmento de reta ligando os mesmos dois pontos. Para um exemplo, tome a Fig. 1.1. Observe que quaisquer dois pontos que forem tomados no gráfico da função, o segmento de reta ligando os mesmos dois pontos está acima do gráfico da função, mostrando que ela é uma função convexa. Já, na Fig. 1.2, observe que quaisquer dois pontos que forem tomados no gráfico da função, o segmento de reta ligando os mesmos dois pontos está abaixo do gráfico da função, por isto ela é uma função côncava.

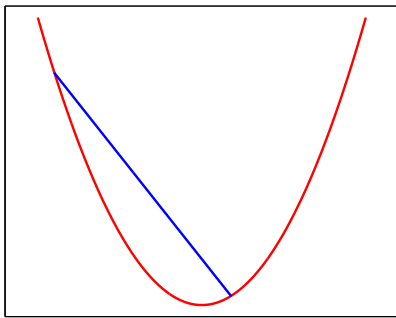


Figura 1.1: *Representação Gráfica de uma função convexa.*

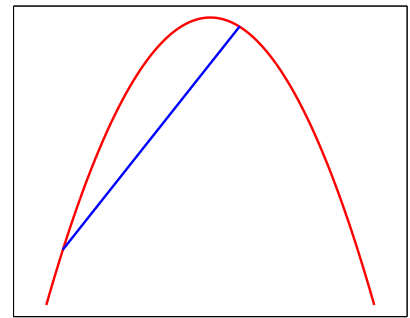


Figura 1.2: *Representação Gráfica de uma função côncava.*

Facilmente pode-se demonstrar que uma função f é convexa se, e somente se, a função

$-f$ é côncava. Agora, seja f a função dada por

$$f(x) = a_0 + \sum_{i=1}^n a_i x_i \quad (1.22)$$

onde $a_0, a_1, \dots, a_n$ são escalares. Esta função é chamada de *função afim*. É fácil ver que a função afim é linear se $a_0 = 0$ e, por consequência, ela é ao mesmo tempo côncava e convexa. Mais ainda, as funções lineares são as únicas que gozam desta propriedade.

Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a função definida por

$$f(x) = \max_i (C_i^T x + d_i). \quad (1.23)$$

onde $C_1, \dots, C_m$ são vetores do $\mathbb{R}^n$, e $d_1, \dots, d_m$ são escalares. Para provar que tal função é convexa é proposto o resultado a seguir.

Teorema 1.7.1. *Sejam $f_1, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}$ funções convexas. Então a função definida como*

$$f(x) = \max_i f_i(x) \quad (1.24)$$

também é convexa.

Demonstração 1.7.1. *Seja x e y no $\mathbb{R}^n$ e seja $\lambda \in [0, 1]$. Daí, tem-se:*

$$f(\lambda x + (1 - \lambda)y) = \max_i f_i(\lambda x + (1 - \lambda)y) \leq \max_i (\lambda f_i(x) + (1 - \lambda)f_i(y)) = \lambda f(x) + (1 - \lambda)f(y).$$

□

Uma função dada pela Eq. (1.23) é chamada de *função linear convexa por partes*, e esta função é usada para aproximar uma função geral. Talvez o exemplo mais simples deste tipo de função seja a função módulo, definida pela Eq. (1.25).

$$f(x) = |x| = \max\{x, -x\}. \quad (1.25)$$

Para finalizar esta seção, juntamente com este capítulo, é apresentado o conceito de mínimos e máximos, tanto global como local.

Definição 1.7.2. Diz-se que um vetor $x \in \mathbb{R}^n$ é um ponto de mínimo local de uma função f se $f(x) \leq f(y)$, $\forall y$, numa vizinhança de x . Ele é chamado de um ponto de mínimo global de uma função f se $f(x) \leq f(y)$, $\forall y \in \mathbb{R}^n$.

Definição 1.7.3. Diz-se que um vetor $x \in \mathbb{R}^n$ é um ponto de máximo local de uma função f se $f(x) \geq f(y)$, $\forall y$, numa vizinhança de x . Ele é chamado de um ponto de máximo global de uma função f se $f(x) \geq f(y)$, $\forall y \in \mathbb{R}^n$.

Capítulo 2

Problemas de Programação Linear

Nesta seção é apresentado o Problema de Programação Linear, alguns exemplos e uma maneira de obter, graficamente, soluções de problemas de programação linear.

2.1 Problema Geral de Otimização

Um *problema geral de otimização* visa minimizar ou maximizar uma função, chamada de *função objetivo*, sujeita ou não a restrições que podem ser restrições de igualdade e desigualdade. De uma maneira geral, tem-se que a função objetivo e as funções de restrições podem ser lineares ou não-lineares, funções de uma única variável ou de várias variáveis, podem ser explícitas ou até mesmo implícitas. Assim, para construir a definição geral de um problema de otimização, considere a definição de uma *função objetivo*.

Definição 2.1.1. *Seja F uma função tal que $F : \Phi \rightarrow \Psi$, onde $\Phi \subset \mathbb{R}^m$ e $\Psi \subset \mathbb{R}^n$. Assim, tem-se que F , chamada de função objetivo, é representada através das suas funções coordenadas, isto é,*

$$F(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) & f_2(\mathbf{x}) & \cdots & f_n(\mathbf{x}) \end{bmatrix}^T, \quad (2.1)$$

sendo que cada uma das suas funções coordenadas é uma função $f_i : \Phi \rightarrow \Omega, i \in \{1, \dots, n\}$,

onde $\Phi \subset n\mathbb{R}^m$ e $\Omega \subset \mathbb{R}$.

Tem-se que cada uma das funções coordenadas f_i , $i \in \{1, \dots, n\}$ também é uma função com qualquer lei de formação e, se $i \neq j$, não é necessário que f_i e f_j sejam de mesma natureza.

Definição 2.1.2. Sejam $F : \Phi_1 \rightarrow \Psi_1$, $G_j : \Phi_2 \rightarrow \Psi_2$, $H_k : \Phi_3 \rightarrow \Psi_3$, onde F, G_j e H_k são funções como na Eq. (2.1), sendo $\Phi_1, \Phi_2, \Phi_3 \subset \mathbb{R}^m$ e $\Psi_1, \Psi_2, \Psi_3 \subset \mathbb{R}^n$. Então um problema geral de otimização é definido como na Eq. (2.2),

$$\begin{aligned} &\text{minimizar ou maximizar:} && F(\mathbf{x}) = [f_1(\mathbf{x}) \ \cdots \ f_n(\mathbf{x})]^T \text{ (Vetor das Funções Objetivo)} \\ &\text{sujeito a:} && \left\{ \begin{array}{ll} G_j(\mathbf{x}) \leq 0 & j = 1, \dots, s \quad (\text{restrições de desigualdades}) \\ H_k(\mathbf{x}) = 0 & k = 1, \dots, t \quad (\text{restrições de igualdades}) \\ x_i^l \leq x_i \leq x_i^u & i = 1, \dots, m \quad (\text{restrições laterais}) \end{array} \right. \end{aligned} \quad (2.2)$$

sendo que o vetor das variáveis de projeto, ou vetor das variáveis de decisão, é dado pela Eq. (2.3).

$$\mathbf{x}^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \end{bmatrix} \quad (2.3)$$

2.2 O Problema de Programação Linear

Um *Problema de Programação Linear* (PPL) visa otimizar uma função objetivo, também chamada de *função custo*, sujeita a restrições de igualdade e de desigualdade, sendo que todas as funções envolvidas, seja a função objetivo ou as funções de restrição, são funções lineares.

O vetor das variáveis de decisão, dado pela Eq. (2.3), que satisfaz a todas as restrições é chamado *uma solução viável* ou *um vetor viável*. O conjunto de todas as soluções viáveis é chamado de *conjunto viável* ou *região viável*. Se $N = \{1, \dots, n\}$ e $j \notin N$ então diz-se que x_j é uma *variável livre* ou que ela é *não-restrita*.

De um modo geral, num PPL é dado um vetor, chamado de *vetor custo*

$$c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}, \quad (2.4)$$

e tenta-se minimizar a função

$$F(x) = c^T x = \sum_{i=1}^m c_i x_i, \quad (2.5)$$

sobre todos os vetores m -dimensionais

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}, \quad (2.6)$$

sujeita ao conjunto de restrições lineares de igualdades e desigualdades. Daí, a função $F(x) = c^T x$ é chamada de *função custo*. Uma solução viável x^* que minimiza a função custo, isto é, uma solução que forneça $c^T x^* \leq c^T x$ para todo x viável, é chamada de *solução viável ótima*, ou simplesmente, uma *solução ótima* e o valor $c^T x^*$ é chamado de *custo ótimo*. Por outro lado, se para cada $k \in \mathbb{R}$ for possível achar uma solução viável x cujo custo seja menor do que k , então diz-se que o custo ótimo é $-\infty$, ou que ele é ilimitado inferiormente. Deve-se observar que não existe necessidade de se estudar problemas de maximização separadamente, visto que maximizar uma função custo $F(x) = c^T x$ é o mesmo que minimizar a função custo $G(x) = -c^T x$.

Para escrever um PPL em linguagem matricial, observe que uma restrição de igualdade $a_i x = b_i$ é equivalente a duas restrições de desigualdades $a_i x \geq b_i$ e $a_i x \leq b_i$. Ainda, tem-se que, toda restrição da forma $a_i x \leq b_i$ pode ser reescrita como $-a_i x \geq -b_i$. Por fim, pode-se observar

que, todas as restrições da forma $x_i \leq 0$ e $x_i \geq 0$ são casos especiais de restrições $a_i x \geq b_i$, onde a_i é o i -ésimo vetor unitário e $b_i = 0$. Com isto, é possível representar um PPL na forma geral exclusivamente em termos de restrições de desigualdades $a_i x \geq b_i$ e, conseqüentemente, na forma de representação matricial. De fato, suponha que exista um total de m restrições do tipo $a_i x \geq b_i$, seja

$$b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \quad (2.7)$$

e seja $\mathcal{A}$ a matriz de dimensão $m \times n$ cujas linhas são os vetores de restrição, ou seja,

$$\mathcal{A} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}. \quad (2.8)$$

Então, as restrições $a_i x \geq b_i$ podem ser expressas da maneira compacta $\mathcal{A}x \geq b$, e o PPL é reescrito como:

$$\begin{array}{ll} \text{minimizar} & F(x) = c^T x \\ \text{sujeito a} & \mathcal{A}x \geq b \end{array}. \quad (2.9)$$

Uma interpretação das desigualdades do tipo $\mathcal{A}x \geq b$ é que para cada $i \in \{1, \dots, m\}$, a i -ésima componente do vetor $\mathcal{A}x$, que é $a_i x$, é sempre maior ou igual a i -ésima componente b_i do vetor b . Para um melhor entendimento do que foi exposto, considere o exemplo a seguir.

Exemplo 2.2.1. *Seja o seguinte PPL:*

$$\begin{array}{ll}
 \text{minimizar :} & F(\mathbf{x}) = 2x_1 - x_2 + 4x_3 \\
 \text{sujeito a} & \left\{ \begin{array}{l} x_1 + x_2 + x_4 \leq 2 \\ 3x_2 - x_3 = 5 \\ x_3 + x_4 \geq 4 \\ x_1 \geq 0 \\ x_3 \leq 0 \end{array} \right.
 \end{array} \quad (2.10)$$

Escrevendo de forma matricial, conforme Eq. (2.9), obtém-se:

$$\mathbf{c}^T = \begin{bmatrix} 2 & -1 & 4 & 0 \end{bmatrix}; \quad (2.11)$$

a matriz $\mathcal{A}$ formado pelas restrições é dada por:

$$\mathcal{A} = \begin{bmatrix} -1 & -1 & 0 & -1 \\ 0 & 3 & -1 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}; \quad (2.12)$$

e o vetor dos recursos disponíveis $\mathbf{b}$:

$$\mathbf{b}^T = \begin{bmatrix} -2 & 5 & -5 & 3 & 0 & 0 \end{bmatrix}. \quad (2.13)$$

2.3 Forma Padrão de um Problema de Programação Linear

Um PPL está na forma padrão quando é representado por:

$$\begin{array}{ll}
\text{minimizar} & F(x) = c^T x \\
\text{sujeito a} & \left\{ \begin{array}{l} \mathcal{A}x = b, \quad b \geq \mathbf{0} \\ x \geq \mathbf{0} \end{array} \right. .
\end{array} \quad (2.14)$$

Veja uma interpretação para problemas na forma padrão: seja x um vetor de dimensão n e seja $\mathcal{A}_1, \dots, \mathcal{A}_n$ as colunas de $\mathcal{A}$. Então, a restrição $\mathcal{A}x = b$ é representada por

$$\sum_{i=1}^n \mathcal{A}_i x_i = b. \quad (2.15)$$

Intuitivamente, existem n vetores $\mathcal{A}_1, \dots, \mathcal{A}_n$ à disposição e um vetor objetivo b . Deseja-se “sintetizar” o vetor b usando o valor não-negativo de x_i de cada vetor $\mathcal{A}_i$, minimizando todo o custo $\sum_{i=1}^n c_i x_i$, onde c_i é o custo unitário do i -ésimo vetor.

Para escrever um PPL qualquer em sua forma padrão, considere que dois problemas são equivalentes quando dado uma solução viável de um problema é possível construir uma solução viável para o outro problema, com o mesmo custo. Em particular, se dois problemas tem o mesmo custo ótimo, dada uma solução ótima para um problema é possível construir uma solução ótima para o outro.

a) Eliminação das Variáveis Livres: Dada uma variável livre x_i num problema na forma geral, deve-se substituí-la por x_i^+ e x_i^- , tal que:

$$x_i = x_i^+ - x_i^-, \text{ onde } x_i^+ \geq 0 \text{ e } x_i^- \geq 0, \quad (2.16)$$

pois qualquer número real pode ser reescrito como sendo a diferença de dois números reais não-negativos.

b) Eliminação das restrições de desigualdade: Dado uma restrição de desigualdade da forma

$$\sum_{j=1}^n a_{ij}x_j \leq b_i, \quad (2.17)$$

adicionando uma nova variável s_i , a restrição é reescrita da forma:

$$\sum_{j=1}^n a_{ij}x_j + s_i = b_i, s_i \geq 0. \quad (2.18)$$

A variável s_i é chamada *variável de folga*. De maneira análoga, restrições da forma $\sum_{j=1}^n a_{ij}x_j \geq b_i$ podem ser colocadas na forma padrão subtraindo uma variável de folga s_i e a restrição torna-se

$$\sum_{j=1}^n a_{ij}x_j - s_i = b_i, s_i \geq 0. \quad (2.19)$$

c) **Ocorrência de $b_i < 0$:** Neste caso, basta multiplicar a restrição i por -1 , pois os coeficientes a_{ij} podem ter qualquer sinal.

Com as afirmações anteriores, concluiu-se que o PPL na forma geral pode ser transformado para a forma padrão e, assim, basta desenvolver métodos para a forma padrão.

Exemplo 2.3.1. *O PPL na forma geral*

$$\begin{array}{ll} \text{minimizar} & F(x) = 2x_1 + 4x_2 \\ \text{sujeito a} & \left\{ \begin{array}{l} x_1 + x_2 \geq 3 \\ 3x_1 + 2x_2 = 14 \\ x_1 \geq 0 \quad x_2 \in \mathbb{R} \end{array} \right. \end{array} \quad (2.20)$$

é equivalente ao PPL na forma padrão

$$\begin{array}{ll}
\text{minimizar} & F(x) = 2x_1 + 4x_2^+ - 4x_2^- \\
\text{sujeito a} & \begin{cases} x_1 + x_2^+ - x_2^- - x_3 = 3 \\ 3x_1 + 2x_2^+ - 2x_2^- = 14 \\ x_1, x_2^+, x_2^-, x_3 \geq 0 \end{cases}
\end{array} \quad (2.21)$$

Veja que dado a solução viável $(x_1, x_2) = (6, -2)$ para o problema original, pode-se obter a solução viável $(x_1, x_2^+, x_2^-, x_3) = (6, 0, 2, 1)$ para o problema na forma padrão, que produz o mesmo custo. Analogamente, dada a solução viável $(x_1, x_2^+, x_2^-, x_3) = (8, 1, 6, 0)$ para o problema na forma padrão, pode-se obter a solução $(x_1, x_2) = (8, -5)$ para o problema original com o mesmo custo.

2.4 Função Linear Convexa por Partes

Considere uma generalização de um problema de otimização, escrito na seguinte forma:

$$\begin{array}{ll}
\text{minimizar} & \max_i (c_i^T x + d_i) \\
\text{sujeito a} & Ax \geq b
\end{array} \quad (2.22)$$

Observe que o valor de $\max_i (c_i^T x + d_i)$ é igual ao menor número z que satisfaz $z \geq c_i^T x + d_i, \forall i$. Por isto, o problema de otimização acima, sob certas considerações, é equivalente ao problema de programação linear:

$$\begin{array}{ll}
\text{minimizar} & z \\
\text{sujeito a} & \begin{cases} z \geq c_i^T x + d_i, & i \in \{1, \dots, n\} \\ Ax \geq b \end{cases}
\end{array} \quad (2.23)$$

onde as variáveis de projeto são z e x .

Resumindo, tem-se que um PPL pode ser usado para resolver problemas de função custo linear convexa por partes e, estas funções, podem ser usadas para aproximar funções custo mais gerais. Por outro lado, uma aproximação linear convexa por partes nem sempre é uma boa idéia,

visto que ela transforma funções “suaves” em funções “não-suaves”, ou seja, ela transforma em funções que tem derivadas descontínuas. Por fim, se for dado uma restrição da forma $f(x) \leq h$, onde f é uma função linear convexa por partes $f(x) = \max_i (f_i^T x + g_i)$, uma outra restrição pode ser escrita

$$f_i^T x + g_i \leq h, i \in \{1, \dots, m\}, \quad (2.24)$$

e novamente a programação linear é aplicada.

2.5 Problema de Otimização envolvendo Função Módulo

Considere um problema de otimização da forma

$$\begin{aligned} \text{minimizar} \quad & F(x) = \sum_{i=1}^n c_i^T |x_i|, \\ \text{sujeito a} \quad & Ax \geq b \end{aligned} \quad (2.25)$$

assumindo que $x^T = [x_1 \cdots x_n]$ e os coeficientes custo c_i são não-negativos. O custo é a soma das funções convexas lineares por parte $c_i^T |x_i|$ que pode ser escrita como uma função linear convexa por partes. De fato, tem-se que

$$c_i^T |x_i| = \begin{cases} c_i^T x_i & \text{se } x_i \geq 0 \\ -c_i^T x_i & \text{se } x_i \leq 0 \end{cases}, \quad (2.26)$$

logo, como cada uma das duas funções obtidas são lineares e cada uma é convexa, então a intersecção delas é linear convexa por partes.

Contudo, expressando este custo na forma $\max_i (c_i^T x + d_i)$, o caminho envolvido é menor e mais direto. Como $|x_i|$ é o maior número z_i que satisfaz $x_i \leq z_i$ e $-x_i \leq z_i$, pode-se obter a seguinte formulação para o PPL:

$$\begin{aligned}
& \text{minimizar} && F(x) = \sum_{i=1}^n c_i z_i \\
& \text{sujeito a} && \begin{cases} \mathcal{A}x \geq b \\ x_i \leq z_i, \quad i \in \{1, \dots, n\} \\ -x_i \leq z_i, \quad i \in \{1, \dots, n\} \end{cases} .
\end{aligned} \tag{2.27}$$

Um método alternativo para se tratar com funções envolvendo valor absoluto seria a de introduzir novas variáveis x_i^+ e x_i^- , obrigando-as a serem positivas e sendo $x_i = x_i^+ - x_i^-$. Assim, substituindo $|x_i|$ por $x_i^+ + x_i^-$ obtém-se a formulação alternativa

$$\begin{aligned}
& \text{minimizar} && \sum_{i=1}^n c_i (x_i^+ + x_i^-) \\
& \text{sujeito a} && \begin{cases} \mathcal{A}x^+ - \mathcal{A}x^- \geq b \\ x^+, x^- \geq 0 \end{cases} ,
\end{aligned} \tag{2.28}$$

onde

$$x^+ = \begin{bmatrix} x_1^+ \\ \vdots \\ x_n^+ \end{bmatrix} \tag{2.29}$$

e

$$x^- = \begin{bmatrix} x_1^- \\ \vdots \\ x_n^- \end{bmatrix} . \tag{2.30}$$

A relação $x_i = x_i^+ - x_i^-$, $x_i^+ \geq 0$, $x_i^- \geq 0$ não é suficiente para garantir que $|x_i| = x_i^+ + x_i^-$ e, por isto, a validade desta reformulação pode não ser tão óbvia. Assuma, por simplicidade, que $c_i > 0, \forall i$. Numa solução ótima, para o problema reformulado, e para todo i , deve-se ter ou

$x_i^+ = 0$ ou $x_i^- = 0$, porque caso contrário deve-se reduzir x_i^+ e x_i^- na soma e deve-se preservar a viabilidade, quando se reduz o custo, o que vai em contradição com o fato dela ser ótima. Tendo garantido que $x_i^+ = 0$ ou $x_i^- = 0$, a relação desejada $|x_i| = x_i^+ + x_i^-$ agora segue. Uma outra informação que deve ser acrescentada é que a suposição da não-negatividade dos coeficientes custo c_i é importantíssima, visto que, de outro modo, o custo critério seria não-convexo.

Para exemplificar o estudo com função módulo, seja o exemplo a seguir.

Exemplo 2.5.1. *Considere o PPL:*

$$\begin{aligned} &\text{minimizar} && F(x) = 2|x_1| + x_2 \\ &\text{sujeito a} && x_1 + x_2 \geq 4 \end{aligned} \quad (2.31)$$

A primeira reformulação produz

$$\begin{aligned} &\text{minimizar} && 2z_1 + x_2 \\ &\text{sujeito a} && x_1 + x_2 \geq 4 \\ &&& x_1 \leq z_1 \\ &&& -x_1 \leq z_1 \end{aligned} \quad (2.32)$$

enquanto que a segunda formulação produz

$$\begin{aligned} &\text{minimizar} && F(x) = 2x_1^+ + 2x_1^- + x_2 \\ &\text{sujeito a} && x_1^+ - x_1^- + x_2 \geq 4 \\ &&& x_1^+ \geq 0 \\ &&& -x_1^- \geq 0 \end{aligned} \quad (2.33)$$

2.6 Exemplos de Problemas de Programação Linear

Os exemplos apresentados nesta seção são encontrados em Bertsimas e Tsitsiklis (1997), sendo apresentados com o objetivo de demonstrar o quanto o estudo do PPL é amplo e variado. Na mesma referência podem ser encontrados vários outros exemplos da aplicação de PPL.

Exemplo 2.6.1. O Problema de Planejamento da Produção de um Computador

Seja o problema que foi estudado pela empresa Digital Equipment Corporation (DEC) no fim de 1988, que ilustra a complexidade e a incerteza das aplicações no mundo real, por exemplo, a dificuldade em se aplicar a modelagem matemática nos modelos reais para se tomar importantes decisões estratégicas. O objetivo deste problema era obter o plano de produção para o primeiro trimestre de 1989.

A DEC introduziu uma nova família de sistemas de computadores GP – 1, GP – 2 e GP – 3 e estações de trabalho WS – 1 e WS – 2, que em geral propunha sistemas de computadores com memória distinta, disco de armazenagem e capacidade de expansão. A Tab. 2.1 apresenta o modelo, o preço, o disco médio e o número de memória usados pelos sistemas. Por exemplo, o sistema GP-1 usa **quatro** memórias internas de 256k e **três** em cada **dez** unidades são fornecidas com um Disco Drive.

Tabela 2.1: Característica dos cinco sistemas desenvolvidos pela DEC, em 1988

Sistema	Preço	Disco Drive	Memória 256k
GP – 1	\$60.000	0,3	4
GP – 2	\$40.000	1,7	2
GP – 3	\$30.000	0,0	2
WS – 1	\$30.000	1,4	2
WS – 2	\$15.000	0,0	1

A distribuição desta nova família de produtos começou no segundo semestre de 1988, sendo que as dificuldades abaixo foram identificadas para o primeiro trimestre de 1989:

- O fornecedor dos gabinetes das CPU's pode fornecer no máximo 7.000 unidades.
- A produção de Disco Drive tem uma estimativa de entrega entre 3.000 e 7.000 unidades.
- O fornecimento da memória de 256k está limitado aos valores de 8.000 à 16.000 unidades.
- O departamento de marketing estabeleceu que a demanda máxima para o primeiro trimestre de 1989 deva ser 1.800 unidades para o sistema GP – 1, 300 unidades para o sistema GP – 3, 3.800 sistemas para toda a família GP e 3.200 sistemas para a família WS. Nesta projeção já foram incluídas as encomendas anteriores, de 500 unidades para o GP – 2, 500 unidades para WS – 1 e 400 unidades para o WS – 2, que devem ser entregue no trimestre

em estudo.

Para modelar o problema da DEC, foi introduzida as variáveis x_1, x_2, x_3, x_4 e x_5 que representam o número de sistemas (por milhares) GP-1, GP-2, GP-3, WS-1 e WS-2, respectivamente, a ser produzido por trimestre. Visto que $1.000x_i$ representa o número de unidades, ele deve ser um número inteiro, mas, pode ser obtido truncando cada x_i depois da terceira casa decimal.

Ainda, no trimestre anterior, a fim de limitar a escassez de Disco Drive, a DEC produziu GP-1, GP-3 e WS-2 sem Disco Drive, mesmo sabendo que 3 de cada 10 clientes para os sistemas GP-1 precisassem de um Disco Drive, e GP-2 e WS-1 com um Disco Drive. Este modo de configuração dos sistemas são chamados de **modo imposto de produção**.

Ainda existia a possibilidade da DEC tratar a falta de Memória de 256k, usando dois pentes de memória alternativa, ao invés dos quatro da memória 256k, no sistema GP-1. A DEC era capaz de produzir 4.000 pentes alternativos para o próximo trimestre.

$$\begin{array}{ll}
 \text{minimizar} & F(x) = 60x_1 + 40x_2 + 30x_3 + 30x_4 + 15x_5 \quad (\text{Rendimento total}) \\
 \text{sujeito a} & \left\{ \begin{array}{ll}
 x_1 + x_2 + x_3 + x_4 + x_5 \leq 7 & (\text{CPU à disposição}) \\
 4x_1 + 2x_2 + 2x_3 + 2x_4 + x_5 \geq 8 & (\text{Memória Mínima}) \\
 4x_1 + 2x_2 + 2x_3 + 2x_4 + x_5 \leq 16 & (\text{Memória Máxima}) \\
 x_2 + x_4 \geq 3 & (\text{Drive Mínimo}) \\
 x_2 + x_4 \leq 7 & (\text{Drive Máximo}) \\
 x_1 \leq 1,8 & (\text{Demanda para GP-1}) \\
 x_3 \leq 0,3 & (\text{Demanda para GP-3}) \\
 x_1 + x_2 + x_3 \leq 3,8 & (\text{Demanda para GP}) \\
 x_4 + x_5 \leq 3,2 & (\text{Demanda para WS}) \\
 x_2 \geq 0,5 & (\text{Demanda para GP-2}) \\
 x_4 \geq 0,5 & (\text{Demanda para WS-1}) \\
 x_5 \geq 0,4 & (\text{Demanda para S-2}) \\
 x_1, x_2, x_3, x_4, x_5 \geq 0 &
 \end{array} \right.
 \end{array} \quad (2.34)$$

A DEC tinha que tomar duas decisões distintas: se usava o modo imposto de produção à respeito ao uso do Disco Drive e se usava a memória alternativa para o sistema GP-1. Como resultado disto, passou a existir quatro combinações diferentes de escolha. O PPL modelado na Eq. (2.34), não utiliza a memória alternativa para o sistema GP-1 e adota o modo imposto de produção.

Note que a função objetivo está em milhares de dólares e o vetor b em milhares de unidades.

Para acomodar as outras três escolhas a DEC tinha que mudar algumas das restrições, como a seguir: se não fosse usado o modo imposto de produção para os Disco Drive, então:

$$\text{troque} \quad \begin{cases} x_2 + x_4 \geq 3 \\ x_2 + x_4 \leq 7 \end{cases} \quad \text{por} \quad \begin{cases} 0.3x_1 + 1.7x_2 + 1.4x_4 \geq 3 \\ 0.3x_1 + 1.7x_2 + 1.4x_4 \leq 7 \end{cases} \quad (2.35)$$

além disso, se for de interesse usar a memória alternativa nos sistemas GP-1, então:

$$\text{troque} \quad \begin{cases} 4x_1 + 2x_2 + 2x_3 + 2x_4 + x_5 \geq 8 \\ 4x_1 + 2x_2 + 2x_3 + 2x_4 + x_5 \leq 16 \end{cases} \quad (2.36)$$

$$\text{por} \quad \begin{cases} 2x_1 \leq 4 \\ 2x_2 + 2x_3 + 2x_4 + x_5 \geq 8 \\ 2x_2 + 2x_3 + 2x_4 + x_5 \leq 16 \end{cases} \quad (2.37)$$

e, assim, cada uma das quatro combinações de escolha levam a quatro diferentes PPL.

Exemplo 2.6.2. Planejando uma escala de Plantões de Trabalho

Neste exemplo a escolha das variáveis não é tão imediata como no caso anterior. Trata-se de uma aplicação numa área bem distinta do Exemplo 2.6.1, pois envolve a distribuição de um certo número de enfermeiras em alguns dias distintos de trabalho.

Num certo hospital é necessário fazer o planejamento do plantão semanal para as suas enfermeiras, que é executado da forma 12pm-8am. O número de enfermeiras para o plantão no dia j é um número inteiro $d_j, j \in \{1, 2, \dots, 7\}$. Cada enfermeira trabalha cinco dias numa semana de plantão. O problema deste Exemplo é achar o número mínimo de enfermeiras que o hospital precisa contratar para os plantões.

Pode-se tentar iniciar usando uma variável y_j , igual ao número de enfermeiras que trabalham no dia j . Mas, com esta definição, pode ser que a restrição de que cada enfermeira trabalhe cinco dias por seção de plantões não seja respeitada. Por esta razão deve-se escolher as variáveis de uma maneira diferente. Por exemplo, defina x_j como sendo o número de enfermeiras que começa sua semana no dia j , assim, uma enfermeira que começa a trabalhar no dia 5, trabalhará nos dias 5,6,7,1,2, ou seja, as enfermeiras que começam a trabalhar no dia 5 trabalharão nos dias 1,2 e depois nos dias 5,6 e 7. Agora pode-se obter a seguinte formulação para o PPL:

$$\begin{array}{ll} \text{minimizar} & F(\mathbf{x}) = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 \\ \text{sujeito a} & \left\{ \begin{array}{l} x_1 + x_4 + x_5 + x_6 + x_7 \geq d_1 \\ x_1 + x_2 + x_5 + x_6 + x_7 \geq d_2 \\ x_1 + x_2 + x_3 + x_6 + x_7 \geq d_3 \\ x_1 + x_2 + x_3 + x_4 + x_7 \geq d_4 \\ x_1 + x_2 + x_3 + x_4 + x_5 \geq d_5 \\ x_2 + x_3 + x_4 + x_5 + x_6 \geq d_6 \\ x_3 + x_4 + x_5 + x_6 + x_7 \geq d_7 \\ x_j \geq 0, \quad x_j \in \mathbb{Z}_+ \end{array} \right. \quad (2.38) \end{array}$$

Então, baseado na restrição de que cada x_i tem que ser um número inteiro, este tipo de problema é conhecido por um **problema de programação linear inteira** (PPLI). Uma maneira de se tratar com este tipo de problema é trabalhar com variáveis reais e assim obter uma solução para o PPL aproximado do problema original. Como o PPL aproximado tem menos restrições do que o problema original, seu custo é menor ou igual ao custo ótimo do

caso original. Se a solução ótima do PPL aproximado for inteira, então ela também é uma solução ótima para o problema original. Se ela não for inteira, então pode-se aproximar cada x_i para um número inteiro, e com isto obter uma solução possível, que não é necessariamente ótima, do problema original.

Exemplo 2.6.3. Escolhendo Caminhos na Rede de Comunicação

Neste exemplo é apresentado um PPL de Fluxo de Rede, que é uma das aplicações mais importante envolvendo Programação Linear.

Considere uma rede de comunicação constituída de n nós. Os nós são conexões de links de comunicação. Um link deixa um caminho de transmissão do nó i ao nó j descrevendo um par ordenado (i, j) . Seja $\mathcal{A}$ o conjunto de todos os links. Assuma que cada link $(i, j) \in \mathcal{A}$ pode transportar até u_{ij} bits por segundo. Existe uma carga positiva c_{ij} para cada bit transmitido por este link. Cada nó k produz dados, na taxa de b^{kl} bits por segundo, que foi transmitido para o nó l , ou por um link direto (k, l) ou seguindo uma seqüência de links. O problema é escolher caminhos onde todos os dados cheguem ao destino desejado, de forma a minimizar o custo total. É admitido que os dados no nó de origem e no nó de destino são os mesmo, porém, eles podem ser transmitidos por caminhos diferentes.

Para resolver este problema, como um PPL, é introduzido variáveis x_{ij}^{kl} indicando a soma dos dados com origem k e destino l que passa pelo link (i, j) . Defina

$$b_i^{kl} = \begin{cases} b^{kl}, & \text{se } i = k \\ -b^{kl}, & \text{se } i = l \\ 0, & \text{caso contrário} \end{cases}. \quad (2.39)$$

Assim, b_i^{kl} é a rede que flui no nó i , de fora da rede, dos dados com origem k e destino l . Tem-se, então, a seguinte formulação:

$$\begin{aligned}
& \text{minimizar} && \sum_{(i,j) \in (A)} \sum_{k=1}^n \sum_{l=i}^n c_{ij} x_{ij}^{kl} \\
& \text{sujeito a} && \left\{ \begin{aligned} & \sum_{\{j|(i,j) \in (A)\}} x_{ij}^{kl} - \sum_{\{j|(i,j) \in (A)\}} x_{ji}^{kl} = b_i^{kl}, i, k, l = \{1, \dots, n\} \\ & \sum_{k=1}^n \sum_{l=1}^n x_{ij}^{kl} \leq u_{ij}, (i, j) \in \mathcal{A} \\ & x_{ij}^{kl} \geq 0, (i, j) \in \mathcal{A}; k, l = \{1, \dots, n\} \end{aligned} \right. . \quad (2.40)
\end{aligned}$$

A primeira restrição está baseada na conservação do fluxo do nó i pelos dados com origem k e destino l , A expressão

$$\sum_{\{j|(i,j) \in (A)\}} x_{ij}^{kl} \quad (2.41)$$

representa a soma de dados com origem k e destino l , respectivamente, que deixa o nó i ao longo de algum link. A expressão

$$\sum_{\{j|(i,j) \in (A)\}} x_{ji}^{kl} \quad (2.42)$$

representa a soma de dados da mesma origem e destino, do caso anterior, que entra no nó i através de algum link. Finalmente, b_i^{kl} é a soma líquida de todos os dados que entram no nó i de fora da rede.

A segunda restrição expressa a necessidade de que o tráfego total de um link (i, j) não pode exceder a capacidade do link. Este problema é conhecido como um problema de **multiconveniência de fluxo**, no qual o tráfego corresponde a cada par origem-destino é visto como uma conveniente diferença.

2.7 Representação Gráfica e Soluções

Nesta seção são apresentados alguns exemplos simples para auxiliar a compreensão geométrica útil de PPL naturais. O primeiro exemplo envolve solução gráfica de um PPL com duas variáveis.

Exemplo 2.7.1. *Considere o seguinte PPL*

$$\begin{array}{ll} \text{minimizar} & F(x) = -x_1 - x_2 \\ \text{sujeito a} & \begin{cases} x_1 + 2x_2 \leq 3 \\ 2x_1 + x_2 \leq 3 \\ x_1, x_2 \geq 0 \end{cases} \end{array} \quad (2.43)$$

Traçando as restrições obtém-se a Fig. 2.1. Por isto, fica fácil observar que o conjunto viável é o conjunto em destaque na Fig. 2.2.

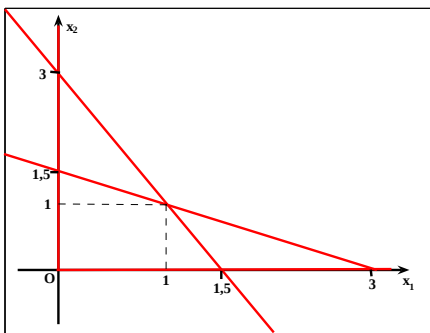


Figura 2.1: Solução Gráfica - Exemplo 2.7.1.

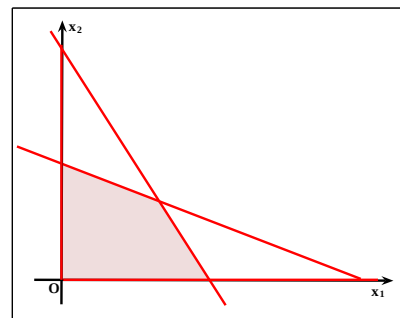


Figura 2.2: Representação da região viável - Exemplo 2.7.1.

Para achar a solução ótima proceda como a seguir: para qualquer escalar z dado, considere o conjunto de todos os pontos cujo custo $c^T x$ é igual a z (esta é a reta descrita pela equação $-x_1 - x_2 = z$). Note que esta reta é perpendicular ao vetor $c^T = [-1 \ -1]$ (veja Fig. 2.3). Valores diferentes de z levam a retas diferentes, todas paralelas entre si, e em particular, o crescimento de z corresponde ao movimento da linha $z = -x_1 - x_2$ sobre a direção do vetor c . Como se deseja minimizar z , deve-se andar com a reta, tanto quanto que possível na direção de $-c$, dentro da região viável. Logo, o ponto ótimo é encontrado quando

$z = -2$, como observado na Fig. 2.3 e, portanto, o ponto $\mathbf{x}^T = [1 \ 1]$ é uma solução ótima. Observe que este ponto é um “canto” do conjunto viável.

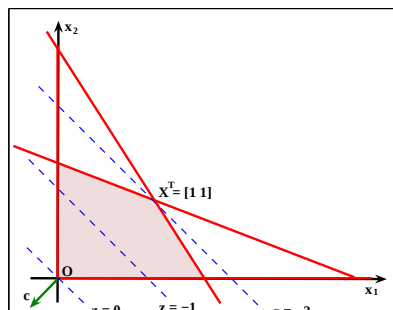


Figura 2.3: Representação das retas perpendiculares ao vetor $\mathbf{c}^T = [-1 \ -1]$ - Exemplo 2.7.1

Para problemas de três dimensões, a mesma argumentação pode ser usada, exceto pelo fato de que o conjunto de pontos com o mesmo valor de $\mathbf{c}^T \mathbf{x}$ é um plano, e não uma reta. Do mesmo modo que no Exemplo 2.7.1 deve-se fazer com que este plano, que é perpendicular ao vetor $\mathbf{c}$, “escorregue” tanto quanto possível, na direção de $-\mathbf{c}$, enquanto não saia do conjunto viável.

Exemplo 2.7.2. Suponha que o conjunto viável seja o cubo unitário, descrito pelas restrições $0 \leq x_i \leq 1, i \in \{1, 2, 3\}$ e que $\mathbf{c}^T = [-1 \ -1 \ -1]$. Então o vetor $\mathbf{x}^T = [1 \ 1 \ 1]$ é uma solução ótima e, como no caso anterior, ela também acontece num “extremo” do conjunto viável, como observado na Fig. 2.4.

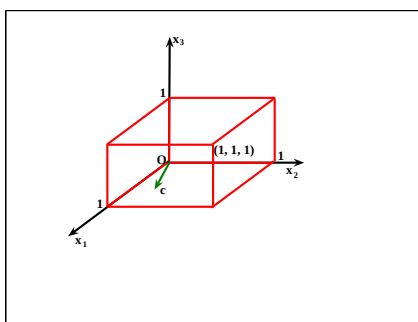


Figura 2.4: Representação da região viável - Exemplo 2.7.2.

No dois exemplos anteriores, o conjunto possível é limitado, isto é, não pode ser estendido até ao infinito, e o problema tem uma única solução ótima. Nem sempre isto acontece e por isto algumas possibilidades adicionais podem ser observadas no exemplo que se segue.

Exemplo 2.7.3. Considere o PPL definido pelas restrições:

$$\begin{aligned} &\text{minimizar} && F(x) = c_1x_1 + c_2x_2, \quad c^T = [c_1 \quad c_2] \\ &\text{sujeito a} && \begin{cases} -x_1 + x_2 \leq 1 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases} \end{aligned} \quad (2.44)$$

cujo conjunto viável pode ser visto na Fig. 2.5.

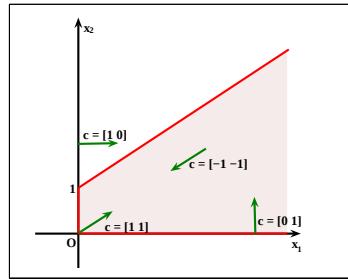


Figura 2.5: Representação da região viável - Exemplo 2.7.3.

a) Para o vetor custo $c^T = [1 \ 1]$, pode ser observado que $x^T = [0 \ 0]$ é a única solução ótima.

b) Para o vetor custo $c^T = [1 \ 0]$ existem múltiplas soluções ótimas, ou seja, cada vetor x da forma $x^T = [0 \ x_2]$, $0 \leq x_2 \leq 1$ é ótima (observe que o conjunto de soluções ótimas é limitado).

c) Para o vetor custo $c^T = [0 \ 1]$ existem múltiplas soluções ótimas, ou seja, cada vetor x da forma $x^T = [x_1 \ 0]$, com $x_1 \geq 0$ é ótima (observe que neste caso o conjunto das soluções ótimas é ilimitado, contendo vetores de magnitudes grandes).

d) Considere o vetor custo $c^T = [-1 \ -1]$. Para qualquer solução viável $x^T = [x_1 \ x_2]$ é sempre possível produzir outra solução viável, com custo menor, por pequenos aumento no

valor de x_1 . Logo a solução viável não é ótima. Além disso, considerando vetores $x^T = [x_1 \ x_2]$ com incremento nos valores de x_1 e x_2 , é possível obter uma seqüência de soluções viáveis cujo custo tende para $-\infty$, por isto, é dito que o custo é $-\infty$.

e) Se é acrescentada uma restrição adicional, por exemplo $x_1 + x_2 \leq -2$, fica evidente (veja Fig. 2.6) que não existe solução viável.

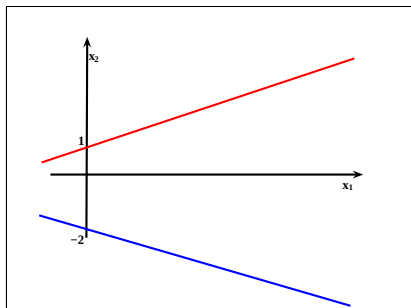


Figura 2.6: Representação da região viável com uma nova restrição - Exemplo 2.7.3.

Observando o exemplo anterior pode-se chegar as seguintes conclusões:

- I) Pode existir uma única solução ótima;
- II) Podem existir múltiplas soluções ótimas, com o conjunto de soluções ótimas podendo ser limitado ou ilimitado;
- III) O custo ótimo é $-\infty$ e a solução viável não é ótima;
- IV) O conjunto viável é vazio.

Ainda é possível acrescentar uma outra possibilidade: uma solução ótima pode não existir mesmo que o problema seja viável e o custo ótimo não seja $-\infty$. Isto pode ser observado quando se deseja minimizar a função $F(x) = x^{-1}$ sujeito a $x > 0$, pois para toda solução possível existe outra solução com custo menor, mas o custo ótimo não é $-\infty$, porém, num PPL este fato não acontece.

Capítulo 3

A Geometria da Programação Linear

Neste capítulo é definido um poliedro como sendo um conjunto descrito por um número finito de equações lineares e de restrições de desigualdades. Em particular, o conjunto viável de um PPL é um poliedro. Por isto, são estudados as propriedades geométricas básicas do Poliedro e alguns detalhes que enfatizam o conceito de “extremo” (vértice). Depois a idéia aplicada ao Poliedro tridimensional é ampliada para poliedros de dimensões maiores. Muitos conceitos, como por exemplo o de vértice, são apresentados tanto geometricamente como algebricamente, apesar de que, a geometria é sempre vista como sendo a mais natural, mas não se pode deixar de ressaltar que a álgebra aproxima-se muito mais do natural no ponto de vista computacional.

3.1 Hiperplanos, Semi-Planos e Poliedro

Aqui são introduzidos alguns conceitos importantes que são úteis para o estudo geométrico de um PPL. Esta seção é iniciada com a definição de Poliedro.

Definição 3.1.1. *Um **poliedro** é um conjunto que pode ser descrito da forma $\{x \in \mathbb{R}^n \mid Ax \geq b\}$, onde A é uma matriz de ordem $m \times n$ e b é um vetor no $\mathbb{R}^n$.*

O conjunto viável de qualquer PPL pode ser descrito pelas restrições de desigualdades da

forma $Ax \geq b$, e é, portanto, um poliedro. Em particular, um conjunto da forma $\{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ também é um poliedro, numa representação na forma padrão.

Um poliedro pode “se estender ao infinito” ou pode estar confinado a uma região limitada. A próxima definição está ligada a esta distinção.

Definição 3.1.2. Um conjunto $S \subset \mathbb{R}^n$ é dito ser **limitado** se existe uma constante K tal que o valor absoluto de cada componente de todo elemento de S seja menor ou igual a K .

A próxima definição relaciona poliedros determinados por uma restrição linear singular.

Definição 3.1.3. Seja a um vetor não nulo do $\mathbb{R}^n$ e seja b um escalar.

a) O conjunto $\{x \in \mathbb{R}^n \mid ax = b\}$ é chamado de um **hiperplano**.

b) O conjunto $\{x \in \mathbb{R}^n \mid ax \geq b\}$ é chamado de um **semi-espaço**.

Observe que um hiperplano é a fronteira de um semi-espaço. Ainda, o vetor a , na definição de hiperplano, é perpendicular ao próprio semiplano. De fato, note que se x e y pertencem ao mesmo hiperplano, então $ax = ay$, logo $a(x - y) = 0$ e, portanto, a é ortogonal a qualquer vetor direcional limitado ao hiperplano. Um hiperplano divide o plano em dois semi-espaços, como pode ser visto na Fig. 3.1. Ainda pode ser observado que um poliedro é igual a intersecção de um número finito de semiplanos, conforme visto na Fig. 3.2, onde o poliedro $\{x \mid a_i x \geq b_i, i \in \{1, 2, 3, 4, 5\}\}$ é a intersecção de 5 semiplanos.

A próxima definição refere-se a importantes propriedades de um número finito de vetores.

Definição 3.1.4. Sejam $x_1, \dots, x_k$ vetores no $\mathbb{R}^n$ e sejam $\lambda_1, \dots, \lambda_k$ escalares cuja soma é um.

a) O vetor $\sum_{i=1}^k \lambda_i x_i$ é dito ser uma combinação convexa dos vetores $x_1, \dots, x_k$.

b) O envoltório convexo dos vetores $x_1, \dots, x_k$ é o conjunto de toda combinação convexa destes vetores.

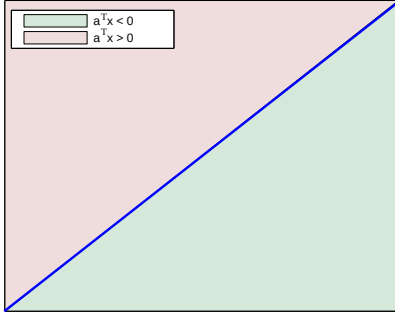


Figura 3.1: Representação de um hiperplano e dos dois semi-espacos determinados por ele.

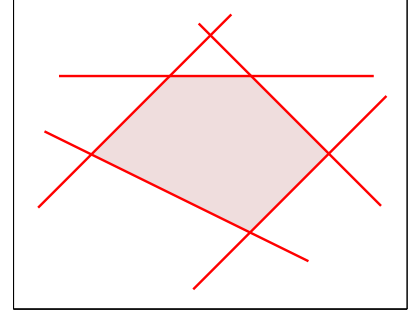


Figura 3.2: Representação de um poliedro formado pela intersecção de 5 semiplanos.

A seguir é apresentado um teorema que estabelece fatos importantes relacionado a convexidade de poliedros.

Teorema 3.1.1. *a) A intersecção de conjuntos convexos é também um conjunto convexo.*

b) Cada poliedro é um conjunto convexo.

c) Uma combinação convexa de um número finito de elementos de um conjunto convexo também pertence ao conjunto.

d) A cobertura convexa de um número finito de vetores é um conjunto convexo.

Demonstração 3.1.1. *a) Sejam $\mathbf{S}_i$ ($i \in I$) conjuntos convexos e seja I um conjunto qualquer de índices. Suponha que $\mathbf{x}$ e $\mathbf{y}$ pertencem à intersecção destes conjuntos, ou seja,*

$$\mathbf{x}, \mathbf{y} \in \bigcap_{i \in I} \mathbf{S}_i. \quad (3.1)$$

Seja $\lambda \in [0, 1]$. Como cada $\mathbf{S}_i$ é um conjunto convexo, então tem-se que $\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}$ também pertence a cada $\mathbf{S}_i$, e, conseqüentemente, pertence à intersecção $\bigcap_{i \in I} \mathbf{S}_i$. Portanto, $\bigcap_{i \in I} \mathbf{S}_i$ é um conjunto convexo.

b) Seja $\mathbf{a}$ um vetor e seja b um escalar. Suponha que $\mathbf{x}$ e $\mathbf{y}$ satisfaçam $\mathbf{a} \mathbf{x} \geq b$ e $\mathbf{a} \mathbf{y} \geq b$, respectivamente, e, portanto, pertencem ao mesmo semiplano. Seja $\lambda \in [0, 1]$. Então

$a(\lambda x + (1 - \lambda)y) \geq \lambda b + (1 - \lambda)b = b$, o que prova que $\lambda x + (1 - \lambda)y$ também pertence ao mesmo semiplano. Logo, um semi plano é convexo. Visto que um poliedro é a intersecção finita de semiplanos, o resultado segue da parte (a).

c) Uma combinação convexa de dois elementos de um conjunto convexo está no conjunto, pela definição de convexidade. Se é assumido, como hipótese de indução, que uma combinação convexa de k elementos de um conjunto convexo pertence ao conjunto, então considere $k + 1$ elementos $x_1, \dots, x_{k+1}$ de um conjunto convexo $\mathbf{S}$ e sejam $\lambda_1, \dots, \lambda_{k+1}$ escalares não-negativos cuja soma seja 1. Assuma, sem perda de generalidade que $\lambda_{k+1} \neq 1$. Assim,

$$\sum_{i=1}^{k+1} \lambda_i x_i = \lambda_{k+1} x_{k+1} + (1 - \lambda_{k+1}) \sum_{i=1}^k \Upsilon_i x_i \quad (3.2)$$

onde

$$\Upsilon_i = \frac{\lambda_i}{1 - \lambda_{k+1}}. \quad (3.3)$$

Os coeficientes Υ_i ($i = 1, \dots, k$) são não-negativos cuja soma é 1, logo, usando a hipótese de indução, tem-se que $\sum_{i=1}^k \Upsilon_i x_i \in \mathbf{S}$. Logo, como $\mathbf{S}$ é convexo tem-se que a Eq. (3.2) implica que $\sum_{i=1}^{k+1} \lambda_i x_i \in \mathbf{S}$, então o processo de indução foi completado.

d) Seja $\mathbf{S}$ a cobertura convexa dos vetores $x_1, \dots, x_k$ e sejam $y = \sum_{i=1}^k \xi_i x_i$, $z = \sum_{i=1}^k \theta_i x_i$ dois elementos de $\mathbf{S}$, onde $\xi_i \geq 0$, $\theta_i \geq 0$, e $\sum_{i=1}^k \xi_i = \sum_{i=1}^k \theta_i = 1$. Seja $\lambda \in [0, 1]$, então:

$$\lambda y + (1 - \lambda)z = \lambda \left(\sum_{i=1}^k \xi_i x_i \right) + (1 - \lambda) \left(\sum_{i=1}^k \theta_i x_i \right) = \sum_{i=1}^k (\lambda \xi_i + (1 - \lambda) \theta_i) x_i. \quad (3.4)$$

Note que os coeficientes $\lambda \xi_i + (1 - \lambda) \theta_i$, ($i = 1, \dots, k$) são não negativos, visto que ele é a soma de números não-negativos, cuja soma é a unidade. De fato:

$$\sum_{i=1}^k (\lambda \xi_i + (1 - \lambda) \theta_i) = \sum_{i=1}^k \lambda \xi_i + \sum_{i=1}^k (1 - \lambda) \theta_i = \lambda \sum_{i=1}^k \xi_i + (1 - \lambda) \sum_{i=1}^k \theta_i = \lambda + (1 - \lambda) = 1. \quad (3.5)$$

Logo, tem-se que $\lambda y + (1 - \lambda)z$ é uma combinação convexa de $x_1, \dots, x_k$ e, portanto, ela pertence a S . Logo, tem-se que S é convexo.

□

3.2 Pontos Extremos, Vértices e Soluções Básicas Viáveis

No capítulo anterior foi dito que num PPL a solução ótima tende a acontecer num “extremo” do poliedro sobre o qual se está otimizando. Nesta seção são dados três modos diferentes de definir o conceito de extremo e é provado que as três definições são equivalentes.

A primeira definição apresenta o ponto extremo de um poliedro como sendo um ponto que não pode ser expresso como sendo uma combinação convexa de dois outros elementos do poliedro, como pode ser visto na Fig. 3.3. Note que esta definição é inteiramente geométrica e ela pode não se referir a representações específicas de um poliedro em termos de restrições lineares. Pode-se observar na Fig. 3.3 que o vetor w não é um ponto extremo de P , visto que ele é uma combinação convexa de v e u e que o vetor x é um ponto extremo.

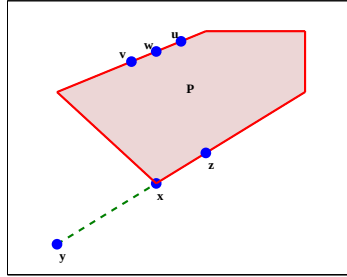


Figura 3.3: Representação de pontos extremos num poliedro.

Definição 3.2.1. Seja P um poliedro. Um vetor $x \in P$ é um **ponto extremo** de P se não é possível encontrar dois vetores diferentes $y, z \in P$ e um escalar $\lambda \in [0, 1]$ tal que $x = \lambda y + (1 - \lambda)z$.

Uma segunda definição geométrica diz que um vértice de um poliedro é uma solução ótima de um PPL cujo conjunto possível é P .

Definição 3.2.2. *Seja P um poliedro. Um vetor $x \in P$ é um **vértice** de P se existe algum c tal que $c^T x < c^T y$, para todo y satisfazendo $x \neq y$ e $y \in P$.*

Em outras palavras, x é um vértice de P se, e somente se, P está de um lado do hiperplano $\{y \mid c^T y = c^T x\}$ que intercepta P somente no ponto x . Observe na Fig. 3.4 que existe uma reta que passa por x e intercepta P num único ponto, logo x é um vértice. Por outro lado, w não é um vértice, pois a intersecção do hiperplano $\{y \mid c^T y = c^T w\}$ com o poliedro P é diferente de w .

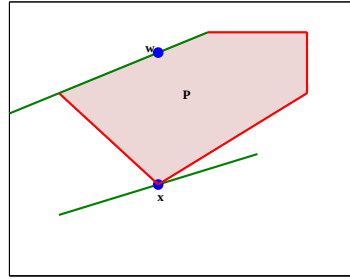


Figura 3.4: Representação de vértice.

As duas definições geométricas anteriores não estão escritas sob a forma de um algoritmo. Porém, deseja-se encontrar uma definição que mostre a representação de um poliedro em função das restrições lineares que o definem, tornando o seu teste apenas algébrico. Para chegar a esta definição algumas notações são apresentadas a seguir.

Considere um poliedro $P \subset \mathbb{R}^n$ definido em relação às restrições de igualdades e desigualdades lineares:

$$\begin{cases} a_i x \geq b_i, & i \in M_1 \\ a_i x \leq b_i, & i \in M_2 \\ a_i x = b_i, & i \in M_3 \end{cases}, \quad (3.6)$$

onde M_1 , M_2 e M_3 são conjuntos finitos de índices, cada a_i é um vetor no $\mathbb{R}^n$ e cada b_i é um escalar.

Definição 3.2.3. Se um vetor x^* satisfaz $a_i x^* = b_i$, para algum i em M_1 , M_2 ou M_3 , então a restrição correspondente é chamada de **ativa** em x^* .

A Fig. 3.5 ilustra a Definição 3.2.3. Seja $P = \{(x, y, z) \mid x + y + z = 1, x, y, z \geq 0\}$. Existem três restrições que são ativas em cada um dos pontos A, B, C e D . Existem somente duas restrições que são ativas no ponto E , que são $x + y + z = 1$ e $y = 0$.

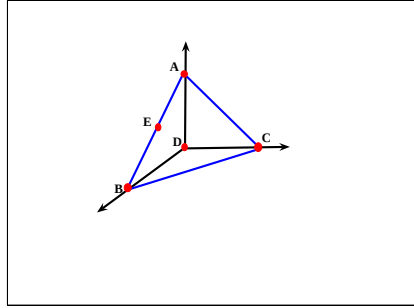


Figura 3.5: Representação do poliedro ilustrado na Definição 3.2.3

Se existir n restrições que são ativas num determinado vetor x^* , então x^* satisfaz um determinado sistema de n equações lineares e n incógnitas. Este sistema tem uma única solução se, e somente se, as n restrições lineares forem LI. O próximo teorema apresenta um significado preciso para esta afirmação, junto com uma ligeira generalização.

Teorema 3.2.1. Seja x^* um elemento do $\mathbb{R}^n$ e seja $I = \{i \mid a_i x^* = b_i\}$ o conjunto dos índices de restrições que são ativas em x^* . Então as seguintes sentenças serão equivalentes:

- Existem n vetores no conjunto $I = \{a_i \mid i \in I\}$ que serão LI.
- A cobertura dos vetores a_i , $i \in I$, é todo o $\mathbb{R}^n$, isto é, todo elemento do $\mathbb{R}^n$ pode ser expresso como uma combinação linear dos vetores a_i , $i \in I$.
- O sistema de equações $a_i x = b_i$, $i \in I$, tem uma única solução.

Demonstração 3.2.1. (a) $\Leftrightarrow$ (b): Suponha que os vetores $a_i, i \in I$ cobre $\mathbb{R}^n$, então a cobertura destes vetores tem dimensão n . Assim, pelo Teorema 1.4.1 tem-se que n destes vetores formam uma base para o $\mathbb{R}^n$ e, conseqüentemente, eles são LI. Reciprocamente, suponha que n dos vetores $a_i, i \in I$ são LI. Então, o subespaço gerado por esses n vetores é n -dimensional e, portanto, ele é igual ao $\mathbb{R}^n$. Assim, cada vetor do $\mathbb{R}^n$ é uma combinação linear dos vetores $a_i, i \in I$, o que estabelece a equivalência.

(b) $\Leftrightarrow$ (c): Se o sistema de equações $a_i x = b_i, i \in I$ tem soluções múltiplas, por exemplo, x^1 e x^2 , então o vetor não-nulo $d = x^1 - x^2$ satisfaz $a_i d = 0, i \in I$, pois $a_i d = a_i(x^1 - x^2) = a_i x^1 - a_i x^2 = b_i - b_i = 0, i \in I$. Visto que o vetor d é ortogonal a todo vetor $a_i, i \in I$, então ele não é uma combinação linear destes vetores e se segue que os vetores $a_i, i \in I$, não pode cobrir $\mathbb{R}^n$. Reciprocamente, se o vetor $a_i, i \in I$ não pode cobrir $\mathbb{R}^n$, escolha um vetor não-nulo d que seja ortogonal ao subespaço coberto por estes vetores. Se x satisfaz $a_i d = 0, \forall i \in I$, então também é verdade que $a_i(x + d) = b_i, \forall i \in I$, logo chegou-se a soluções múltiplas. O que estabelece a equivalência.

□

Para simplificar a linguagem algumas vezes é dito que certas restrições são LI, enquanto que o certo é dizer que os vetores correspondentes a_i são LI. Com esta terminologia, a primeira declaração do Teorema 3.2.1 exige que exista n restrições LI que são ativas em x^* .

Agora já é possível dar uma definição algébrica para um ponto extremo, como sendo uma solução viável onde existam n restrições ativas LI. Assim, como o interesse são as soluções viáveis, toda as restrições de igualdade devem ser ativas. Isto fornece o seguinte modo de se encontrar pontos extremos: primeiro imponha que as restrições de igualdade e algumas restrições adicionais sejam ativas, totalizando n restrições ativas LI. Com estas n restrições ativas LI, um único vetor x^* é determinado, como garante o Teorema 3.2.1. Contudo, este procedimento não dá garantias ao vetor viável x^* , pois algumas das restrições podem ser violadas; neste caso diz-se que ela é uma solução básica, mas não viável.

Definição 3.2.4. Considere um poliedro P definido por restrições de igualdades e de desigualdades, e seja x^* um elemento do $\mathbb{R}^n$.

a) O vetor x^* é uma **solução básica** se:

i) Todas as restrições de igualdade forem ativas;

ii) Dentro destas restrições que são ativas em x^* , existem n delas que são LI.

b) Se x^* é uma solução básica que satisfaça todas as restrições, então diz-se que ela é uma **solução básica viável**.

Olhando na Fig. 3.5, tem-se que os pontos A , B e C são soluções básicas viáveis. O ponto D não é uma solução básica, visto que ela falha na restrição de igualdade $x_1 + x_2 + x_3 = 1$. O ponto E é viável mas não é básico. Se a restrição de igualdade for trocada pelas restrições $x_1 + x_2 + x_3 \leq 1$ e $x_1 + x_2 + x_3 \geq 1$, então o ponto D passa a ser uma solução básica. Portanto, para um ponto ser uma solução básica, ou não, pode depender de como que um poliedro é representado. A Fig. 3.6 também exemplifica a Definição 3.2.4. Seja $P = \{(x, y) \mid -x + y \leq 1; x + y \leq 2; x \geq 0 \text{ e } y \geq 0\}$. Então, os pontos A , B , C , D , E , F são todos soluções básicas pois em cada uma delas existe duas restrições LI que são ativas. Os pontos C , D , E , F são soluções básicas viáveis.

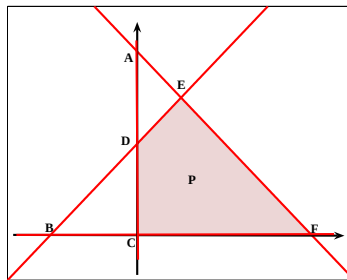


Figura 3.6: Representação de um poliedro ilustrando a Definição 3.2.4

Observe que se o número m de restrições usado para definir um poliedro $P \subset \mathbb{R}^n$ for menor do que n , então o número de restrições ativas em qualquer ponto deve ser também menor do que n e, por isto, não existem soluções básicas ou soluções básicas viáveis.

Até aqui foram apresentadas três diferentes definições que são significantes para um mesmo

conceito, sendo que duas destas foram geométricas (ponto extremo e vértice) e uma foi algébrica (solução básica possível). A seguir é provado que as três definições são equivalentes e, por isto, os três conceitos são usados, de agora em diante, permutavelmente.

Teorema 3.2.2. *Seja P um poliedro não vazio e seja $x^* \in P$. Então, as seguintes sentenças são equivalentes:*

- a) x^* é um vértice;
- b) x^* é um ponto extremo;
- c) x^* é uma solução básica viável.

Demonstração 3.2.2. (a) $\Rightarrow$ (b): *Suponha que $x^* \in P$ seja um vértice. Então pela Definição 3.2.2, existe algum $c \in \mathbb{R}^n$ tal que $c^T x^* < c^T y$, para todo y satisfazendo $y \in P$ e $y \neq x^*$. Se $y \in P$, $z \in P$, $y \neq x^*$, $z \neq x^*$ e $0 \leq \lambda \leq 1$, então $c^T x^* < c^T y$ e $c^T x^* < c^T z$, conseqüentemente tem-se que $c^T x^* = c^T(\lambda x^* + (1 - \lambda)x^*) < c^T(\lambda y + (1 - \lambda)z)$ e, por isto, $x^* \neq \lambda y + (1 - \lambda)z$. Assim, x^* não pode ser expresso como uma combinação convexa de dois outros elementos de P e ele é, portanto, um ponto extremo.*

(b) $\Rightarrow$ (c): *Suponha que $x^* \in P$ não seja uma solução básica viável. Seja $I = \{i \mid a_i x^* = b_i\}$. Como x^* não é uma solução básica viável não existem n vetores LI diferentes na família $a_i, i \in I$. Assim, os vetores $a_i, i \in I$ estão num subespaço próprio do $\mathbb{R}^n$, e por isto existe um vetor $d \in \mathbb{R}^n$ tal que $a_i d = 0, \forall i \in I$. Seja ϵ um número positivo pequeno e considere os vetores $y = x^* + \epsilon d$ e $z = x^* - \epsilon d$. Facilmente tem-se que $a_i y = a_i z = b_i$, para $i \in I$. Além disso, para $i \notin I$, tem-se que $a_i x^* > b_i$ e, como ϵ é um número pequeno, tem-se também que $a_i y > b_i$. Assim, quando ϵ é ínfimo, $y \in P$ e, de maneira análoga, tem-se que $z \in P$. É fácil ver que $x^* = \frac{(y+z)}{2}$, o que implica que x^* não é um ponto extremo.*

(c) $\Rightarrow$ (a): *Seja x^* uma solução básica viável e seja $I = \{i \mid a_i x^* = b_i\}$. Seja $c = \sum_{i \in I} a_i$. Então tem-se:*

$$c^T x^* = \sum_{i \in I} a_i x^* = \sum_{i \in I} b_i. \quad (3.7)$$

Além disso, para qualquer $x \in P$ e para qualquer i , tem-se que $a_i x \geq b_i$ e, por isto,

$$c^T x = \sum_{i \in I} a_i x \geq \sum_{i \in I} b_i. \quad (3.8)$$

Logo, tem-se que x^* é uma solução ótima do problema de minimização de $c^T x$ sobre o conjunto P . Além disto, as igualdades tomadas na Eq. (3.8) acontecem se, e somente se, $a_i x = b_i, \forall i \in I$. Visto que x^* é uma solução básica viável, existem n restrições LI que são ativas em x^* e que x^* é a única solução do sistema de equações $a_i x = b_i, \forall i \in I$ (Teorema 3.2.1). Segue que x^* é o único ponto que minimiza $c^T x$ sobre o conjunto P e, portanto, x^* é um vértice de P .

□

Visto que um vetor é uma solução básica viável se, e somente se, ele é um ponto extremo, e visto que a definição de um ponto extremo não faz referência a qualquer representação particular de um poliedro, então pode-se concluir que esta propriedade é estendida à solução básica viável, isto é, ser uma solução básica viável independe da representação usada.

O próximo resultado é também muito importante.

Corolário 3.2.1. *Dado um número finito de restrições lineares de desigualdade, elas somente podem ter um número finito de soluções básicas ou soluções básicas viáveis.*

Demonstração 3.2.1. *Considere um sistema de m restrições de desigualdades lineares imposta num vetor $x \in \mathbb{R}^n$. Em qualquer solução básica existe n restrições LI ativas. Visto que quaisquer n restrições LI ativas definem um único ponto, segue que soluções básicas diferentes correspondem a diferentes conjuntos de n restrições LI ativas. Portanto, o número de soluções básicas é limitado superiormente pelo número de combinações diferentes que se pode escolher n restrições do total de m , que é um conjunto finito.*

□

Embora o conjunto de soluções básicas, e conseqüentemente o número de soluções básicas viáveis, esteja garantido ser finito, ele pode ser muito grande, por exemplo, o cubo unitário $\{x \in \mathbb{R}^n \mid 0 \leq x_i \leq 1, i = 1, \dots, n\}$ está definido por $2n$ restrições, mas ele tem 2^n soluções básicas viáveis.

Definição 3.2.5. *Duas soluções básicas distintas de um conjunto de restrições lineares são ditas serem **adjacentes** se é possível achar $n - 1$ restrições LI que são ativas em ambas.*

Para exemplificar esta definição, na Fig. 3.6 tem-se que D e E são adjacentes a B , também tem-se que A e C são adjacente a D . Se duas soluções básicas adjacentes são também viáveis, então o segmento de reta que liga as duas é chamado de **borda** do conjunto viável.

3.3 Poliedro na Forma Padrão

A Definição 3.2.4, de soluções básicas viáveis, refere-se a poliedros numa forma geral. As definições e os resultados que agora são apresentados são fundamentais para o desenvolvimento do Método Simplex.

Seja $P = \{x \in \mathbb{R}^n \mid \mathcal{A}x = b, x \geq 0\}$ um poliedro na forma padrão, e seja $\mathcal{A}$ de dimensão $m \times n$, onde m é o número de restrições de igualdade. Sem perda de generalidade é tomado que as m linhas de $\mathcal{A}$ são LI, visto que como as linhas são n -dimensionais, tem-se que $m \leq n$, e como será mostrado posteriormente, linhas LD de $\mathcal{A}$ correspondem a restrições redundantes que podem ser descartadas, o que torna esta suposição aceitável.

Qualquer solução viável tem que ter n restrições LI que sejam ativas. Além disso, cada solução básica deve satisfazer à restrição de igualdade $\mathcal{A}x = b$, o que fornece m restrições ativas, que são LI pela suposição anterior sobre as linhas de $\mathcal{A}$. Assim, para obter um total de n restrições ativas deve-se escolher $n - m$ das variáveis x_i e as tornarem nulas, o que faz a correspondente restrição não-negativa $x_i \geq 0$ ativa. Contudo, para que o conjunto resultante de n restrições ativas seja LI a escolha destas $n - m$ variáveis não pode ser totalmente arbitrária, como é visto a seguir.

Teorema 3.3.1. *Considere a restrição $\mathcal{A}x = b$ e $x \geq 0$, e assuma que a matriz $\mathcal{A}$ de dimensão $m \times n$ tem linhas LI. Um vetor $x \in \mathbb{R}^n$ é uma solução básica se, e somente se, tem-se que $\mathcal{A}x = b$ e existem índices $B(1), \dots, B(m)$ tal que:*

- a) *As colunas $\mathcal{A}_{B(1)}, \dots, \mathcal{A}_{B(m)}$ são LI;*
- b) *Se $i \neq B(1), \dots, B(m)$, então $x_i = 0$.*

Demonstração 3.3.1. *Considere algum $x \in \mathbb{R}^n$ e suponha que existam índices $B(1), \dots, B(m)$ que satisfaz (a) e (b) na afirmação do Teorema. A restrição ativa $x_i = 0, i \neq B(1), \dots, B(m)$, e $\mathcal{A}x = b$ implica que:*

$$\sum_{i=1}^m \mathcal{A}_{B(i)} x_{B(i)} = \sum_{i=1}^n \mathcal{A}_i x_i = \mathcal{A}x = b. \quad (3.9)$$

Visto que as colunas de $\mathcal{A}_{B(i)}, i = 1, 2, \dots, m$ são LI, $x_{B(1)}, \dots, x_{B(m)}$ são determinados de maneira única. Assim, o sistema formado pelas restrições ativas tem uma única solução. Pelo Teorema 3.2.1 existem n restrições ativas LI, e isto implica que x é uma solução básica.

Para demonstrar a recíproca, assuma que x é uma solução básica. Seja $x_{B(1)}, \dots, x_{B(m)}$ as componentes de x que não são nulas. Como x é uma solução básica viável, o sistema de equações formado pelas restrições ativas $\sum_{i=1}^n \mathcal{A}_i x_i = b$ e $x_i = 0, i \neq B(1), \dots, B(m)$ tem uma única solução (conforme Teorema 3.2.1); equivalentemente, a equação $\sum_{i=1}^k \mathcal{A}_{B(i)} x_{B(i)} = b$ tem uma única solução. Daí, as colunas $\mathcal{A}_{B(1)}, \dots, \mathcal{A}_{B(m)}$ são LI, pois caso contrário é possível achar escalares $\lambda_1, \dots, \lambda_n$, nem todos nulos, tal que $\sum_{i=1}^k \mathcal{A}_{B(i)} \lambda_i = 0$, então $\sum_{i=1}^k \mathcal{A}_{B(i)} (x_{B(i)} + \lambda_i) = b$, contrariando a unicidade da solução.

Com isto tem-se que $\mathcal{A}_{B(1)}, \dots, \mathcal{A}_{B(k)}$ são LI, e por isto $k \leq n$. Visto que $\mathcal{A}$ tem m linhas LI, segue que ela também tem m colunas LI, é possível achar $m - k$ colunas adicionais $\mathcal{A}_{B(k+1)}, \dots, \mathcal{A}_{B(m)}$ tal que as colunas $\mathcal{A}_{B(i)}, i \in \{1, \dots, n\}$ sejam LI. Assim, se $i \neq B(1), \dots, B(m)$, então $i \neq B(1), \dots, B(k)$, pois $k \leq m$ e $x_i = 0$. Logo, (a) e (b) do teorema são satisfeitas.

□

Por causa do Teorema 3.3.1 toda solução básica para um poliedro na forma padrão pode ser construído como a seguir:

Tabela 3.1: Procedimento para Construção de Soluções Básicas:

1 - escolha m colunas LI $\mathcal{A}_{B(1)}, \dots, \mathcal{A}_{B(m)}$.
2 - Seja $x_i = 0$ para todo $i \neq B(1), \dots, B(m)$.
3 - Resolva o sistema de m equações $\mathcal{A}x = b$ para os $x_{B(1)}, \dots, x_{B(m)}$ desconhecidos.

Se uma solução básica construída conforme este procedimento não é negativa, então ela é viável, e por isto ela é uma solução básica viável. Reciprocamente, visto que toda solução básica viável é uma solução básica, ela pode ser obtida por este procedimento. Se x for uma solução básica, as variáveis $x_{B(1)}, \dots, x_{B(m)}$ são chamadas **variáveis básicas** e as variáveis restantes são chamadas **variáveis não básicas**. As colunas $\mathcal{A}_{B(1)}, \dots, \mathcal{A}_{B(m)}$ são as **colunas básicas** e por serem LI elas formam uma base para o $\mathbb{R}^m$. Quando for dito que duas bases são distintas, ou diferentes, está sendo feita referencia à duas bases que envolvem conjuntos de índices básicos $\{B(1), \dots, B(m)\}$ diferentes, por isto duas bases que são constituídas pelos vetores em ordens diferentes são vistas como as mesmas.

A matriz $\mathbf{B}$ de dimensão $m \times m$ tomada como sendo a matriz formada pelas colunas básicas, uma ao lado da outra, é chamada de uma **matriz básica**. Como as colunas de $\mathbf{B}$ são LI ela é uma matriz invertível. De uma maneira análoga, tem-se que um vetor $\mathbf{x}_B$ formado pelas variáveis básicas é o **vetor básico**. Assim,

$$\mathbf{B} = \begin{bmatrix} B(1) & \dots & B(m) \end{bmatrix}, \quad \mathbf{x}_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix}. \quad (3.10)$$

As variáveis básicas são determinadas pela resolução da equação $\mathbf{B}\mathbf{x}_B = b$ cuja única solução é dada por

$$\mathbf{x}_B = \mathbf{B}^{-1}b. \quad (3.11)$$

Exemplo 3.3.1. *Seja a restrição $\mathcal{A}x = b$ da forma*

$$\begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 6 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 8 \\ 12 \\ 4 \\ 6 \end{bmatrix} \quad (3.12)$$

Escolha $\mathcal{A}_4, \mathcal{A}_5, \mathcal{A}_6$ e $\mathcal{A}_7$ como as colunas básicas. É fácil concluir que estas colunas são LI e que a matriz básica correspondente é a identidade. Usando o Teorema 3.3.1 tem-se que a solução básica $x = (0, 0, 0, 8, 12, 4, 6)$, que é não negativa, é uma solução básica viável. Por outro lado, escolhendo as colunas $\mathcal{A}_3, \mathcal{A}_5, \mathcal{A}_6$ e $\mathcal{A}_7$ como as colunas básicas, a solução básica correspondente é $x = (0, 0, 4, 0, -12, 4, 6)$ que não é uma solução viável, visto que $x_5 < 0$.

Suponha que tenha uma 8ª coluna que seja idêntica a 7ª coluna. Então os conjuntos de colunas $\{\mathcal{A}_3, \mathcal{A}_5, \mathcal{A}_6, \mathcal{A}_7\}$ e $\{\mathcal{A}_3, \mathcal{A}_5, \mathcal{A}_6, \mathcal{A}_8\}$ coincidem. Por outro lado, os correspondentes conjuntos de índices básicos $\{3, 5, 6, 7\}$ e $\{3, 5, 6, 8\}$ são diferentes, o que fornece duas bases diferentes, conforme a convenção dada.

Para uma visão intuitiva de soluções básicas relembre a interpretação das restrições $\sum_{i=1}^n \mathcal{A}_i x_i = b$ como sendo um requisito para sintetizar o vetor $b \in \mathbb{R}^n$ recorrendo aos vetores $\mathcal{A}_i$. Numa solução básica é preciso recorrer somente ao uso de m vetores, onde estes estão associados com as variáveis básicas. Além disso, numa solução básica viável, isto é ilustrado usando um múltiplo escalar, não negativo, de cada vetor viável, como visto na Fig. 3.7. Considere um problema na forma padrão com $n = 4$, $m = 2$ e os vetores $b, \mathcal{A}_1, \dots, \mathcal{A}_4$. Os vetores $\mathcal{A}_1, \mathcal{A}_2$ formam uma base, mas a solução básica correspondente é inviável, visto que é necessário um múltiplo escalar negativo de x_2 para sintetizar b através de $\mathcal{A}_1, \mathcal{A}_2$. Os vetores $\mathcal{A}_1, \mathcal{A}_3$ formam outra base, e a solução básica correspondente é viável. Os vetores $\mathcal{A}_1, \mathcal{A}_4$ não podem formar uma base pois eles são LD.

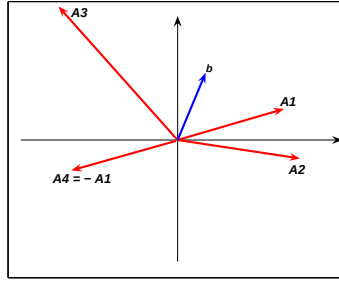


Figura 3.7: Representação do problema na forma padrão com $n = 4$ e $m = 2$.

Diferentes soluções básicas devem corresponder a diferentes bases, pois uma base determina de maneira única uma solução básica. Contudo, duas bases diferentes podem levar as mesmas soluções básicas, por exemplo, sendo $b = \mathbf{0}$, então cada matriz básica leva às mesmas soluções básicas, que é o vetor nulo. Por outro lado, duas bases são adjacentes se elas partem de todas, menos uma, coluna básica, e por isto não é difícil checar que soluções básicas adjacentes sempre podem ser obtidas de duas bases adjacentes. Reciprocamente, se duas bases adjacentes levam a soluções básicas distintas então estas soluções são adjacentes.

Exemplo 3.3.2. Referindo-se ao Exemplo 3.3.1 tem-se que as bases $\{\mathcal{A}_4, \mathcal{A}_5, \mathcal{A}_6, \mathcal{A}_7\}$ e $\{\mathcal{A}_3, \mathcal{A}_5, \mathcal{A}_6, \mathcal{A}_7\}$ são adjacentes, pois todas, menos uma, das colunas são as mesmas. As soluções básicas correspondentes $x_1 = (0, 0, 0, 8, 12, 4, 6)$ e $x_2 = (0, 0, 4, 0, -12, 4, 6)$ são adjacentes (tem-se 7 restrições mas 6 destas restrições ativas são LI, são elas $x_1 \geq 0$, $x_2 \geq 0$ e as quatro restrições de igualdade).

Veja, a seguir, que a hipótese de completar o posto linha de $\mathcal{A}$ não provoca perda de generalidade.

Teorema 3.3.2. Seja $P = \{x \mid \mathcal{A}x = b, x \geq 0\}$ um poliedro não vazio, onde $\mathcal{A}$ é uma matriz de dimensão $m \times n$, com linhas $a_1, \dots, a_m$. Suponha que $\text{posto}(\mathcal{A}) = k < m$ e que as linhas $a_{i_1}, \dots, a_{i_k}$ são LI. Considere o poliedro $Q = \{x \mid a_{i_1}x = b_{i_1}, \dots, a_{i_k}x = b_{i_k}, x \geq 0\}$. Então $P = Q$.

Demonstração 3.3.2. Considere o caso em que $i_1 = 1, \dots, i_k = k$, ou seja, as primeiras k linhas de $\mathcal{A}$ são LI. O caso geral pode ser obtido reorganizando as linhas de $\mathcal{A}$.

Qualquer elemento de P satisfaz todas as condições de Q , visto que os elementos de Q são linhas de $\mathcal{A}$, por isto $P \subset Q$.

Visto que $\text{posto}(\mathcal{A}) = k$ o espaço gerado pelas linhas de $\mathcal{A}$ tem dimensão k e as linhas $a_1, a_2, \dots, a_k$ formam uma base para este espaço. Assim, cada linha de $\mathcal{A}$ pode ser escrita como sendo uma combinação linear das linhas $a_1, a_2, \dots, a_k$, isto é, cada linha a_i é da forma

$$a_i = \sum_{j=1}^k \lambda_{ij} a_j$$

para alguns escalares λ_{ij} . Seja x um elemento de P . Note que

$$b_i = a_i x = \sum_{j=1}^k \lambda_{ij} a_j x = \sum_{j=1}^k \lambda_{ij} b_j, \quad i = 1, \dots, n. \quad (3.13)$$

Considere agora um elemento y de Q . Para qualquer i tem-se que

$$a_i y = \sum_{j=1}^k \lambda_{ij} a_j y = \sum_{j=1}^k \lambda_{ij} b_j = b_j,$$

o que estabelece que $y \in P$ e por isto $Q \subset P$.

□

O Poliedro Q no Teorema 3.3.2 está na forma padrão, isto é, $Q = \{x \mid \mathcal{D}x = f, x \geq 0\}$ onde $\mathcal{D}$ é uma submatriz (de dimensão $k \times m$) da matriz $\mathcal{A}$ de posto k e f é um vetor k -dimensional derivado de b . Com isto, conclui-se que quando o conjunto viável é não vazio, um PPL na forma padrão pode ser reduzido em um outro PPL, também na forma padrão e com o mesmo conjunto viável, onde as restrições de igualdade são LI.

Exemplo 3.3.3. Considere o Poliedro, não vazio, definido pelas restrições:

$$\begin{cases} 2x_1 + x_2 + x_3 = 2 \\ x_1 + x_2 = 1 \\ x_1 + x_3 = 1 \\ x_1, x_2, x_3 \geq 0 \end{cases} \quad (3.14)$$

A matriz correspondente $\mathcal{A}$ tem posto 2, pois as duas últimas linhas de $\mathcal{A}$, $a_2 = [1 \ 1 \ 0]$ e $a_3 = [1 \ 0 \ 1]$, são LI, mas a primeira linha é a soma das outras duas. Com isto a primeira restrição é redundante e pode ser eliminada, mantendo o mesmo poliedro.

3.4 Degeneração

Conforme visto na seção anterior uma solução básica deve ter n restrições ativas LI, mas existe a possibilidade do número de restrições ativas ser maior do que n , mesmo tendo que num espaço n -dimensional apenas n destas restrições são LI. Assim, quando uma solução básica tiver mais do que n restrições ativas ela é uma solução básica degenerada.

Definição 3.4.1. Uma solução básica $x \in \mathbb{R}^n$ é dita ser degenerada se mais do que n das restrições são ativas em x .

Num espaço bidimensional, uma solução básica degenerada está na intersecção de três ou mais retas. No espaço tridimensional uma solução básica degenerada é a intersecção de quatro ou mais planos. Na Fig. 3.8 o ponto C é uma solução básica viável degenerada no plano, já que três restrições são ativas em C , já o ponto E é uma solução básica viável não degenerada no plano e o ponto D é uma solução básica degenerada no plano. Na Fig. 3.9 ponto A é solução básica viável degenerada no $\mathbb{R}^3$ e o ponto B é uma solução básica não degenerada.

Exemplo 3.4.1. Considere o Poliedro P definido pelas restrições:

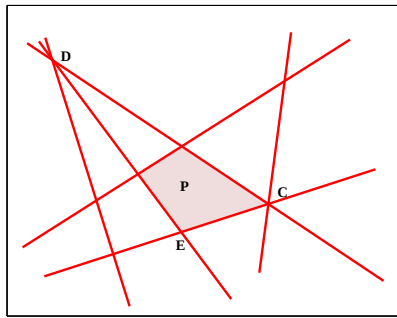


Figura 3.8: Representação de uma Solução Básica Degenerada no plano.

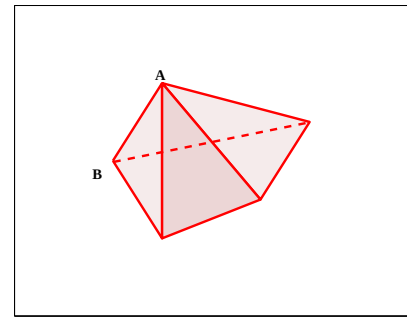


Figura 3.9: Representação de uma Solução Básica Degenerada no espaço.

$$\begin{cases} x_1 + x_2 + 2x_3 \leq 8 \\ x_2 + 6x_3 \leq 12 \\ x_1 \leq 4 \\ x_2 \leq 6 \\ x_1, x_2, x_3 \geq 0 \end{cases}.$$

O vetor $x_1 = (2, 6, 0)$ é uma solução básica viável não degenerada, visto que são exatamente três as restrições ativas em x_1 , que são LI, são elas $x_1 + x_2 + 2x_3 \leq 8$, $x_2 \leq 6$ e $x_3 \geq 0$. O vetor $x_2 = (4, 0, 2)$ é uma solução viável básica degenerada, porque existem quatro restrições ativas em x_2 , sendo três LI, são elas $x_1 + x_2 + 2x_3 \leq 8$, $x_2 + 6x_3 \leq 12$, $x_1 \leq 4$ e $x_2 \geq 0$.

Numa solução básica de um poliedro na forma padrão, as m restrições de igualdade são sempre ativas. Portanto, ter mais do que n restrições ativas é o mesmo que ter mais do que $n - m$ variáveis valendo zero. Com isto temos que:

Definição 3.4.2. Considere o poliedro $P = \{x \in \mathbb{R}^n \mid \mathcal{A}x = b, x \geq 0, b \geq 0\}$ na forma padrão, seja x uma solução básica e seja m o número de linhas de $\mathcal{A}$. O vetor x é uma solução básica degenerada se mais do que $n - m$ das componentes de x valem zero.

Exemplo 3.4.2. Considere novamente o poliedro do Exemplo 3.4.1. Passando o poliedro P para a forma padrão, utilizando as variáveis de folga x_4, x_5, x_6, x_7 , tem-se que $P = \{x = (x_1, x_2, x_3, x_4, x_5, x_6, x_7) \mid \mathcal{A}x = b, x \geq 0, b \geq 0\}$, onde

$$\mathcal{A} = \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 6 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 8 \\ 12 \\ 4 \\ 6 \end{bmatrix}.$$

Considere a base constituída pelas colunas LI $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_7$. Para calcular a solução básica correspondente, primeiro transforme as variáveis não básicas x_4, x_5 e x_6 em zero, e então resolva o sistema $\mathcal{A}x = b$ para as variáveis restantes, obtendo $x_1 = (4, 0, 2, 0, 0, 0, 6)$. Com isto, x_1 é uma solução básica degenerada, visto que ela possui 4 variáveis valendo zero e com isto $n - m = 7 - 4 = 3$. Considere agora a base constituída pelas colunas LI $\mathcal{A}_1, \mathcal{A}_3, \mathcal{A}_4, \mathcal{A}_7$. A solução básica viável correspondente é de novo $x_1 = (4, 0, 2, 0, 0, 0, 6)$.

O Exemplo 3.4.2 sugere que é possível considerar soluções degeneradas do seguinte modo: escolha um solução básica tomando n restrições LI para serem restrições de igualdade. Se as entradas de $\mathcal{A}$ e b são escolhidas ao acaso isto quase nunca acontece. Além disso, a Fig. 3.10 mostra que se houver uma pequena perturbação nas restrições ativas, solução degenerada pode desaparecer. Em problemas práticos, contudo, as entradas de $\mathcal{A}$ e b tem muitas vezes uma estrutura espacial e por isto soluções degeneradas são mais comuns do que o argumento anterior.

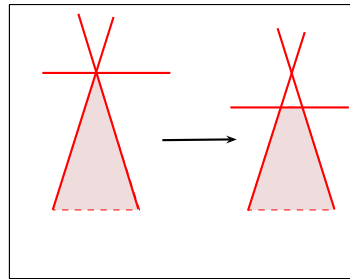


Figura 3.10: Representação de uma restrição degenerada que tornou-se não degenerada, através de uma pequena carga.

Para visualizar soluções degeneradas em poliedros na forma padrão, é assumido que $n - m = 2$ e tem-se que o conjunto viável é um subconjunto do $\mathbb{R}^2$ definido pelas restrições de igualdade $\mathcal{A}x = b$. Na Fig. 3.11 é feita uma ilustração de soluções degeneradas $(n - m)$ -dimensional, onde $n = 6$ e $m = 4$. A solução viável A é não degenerada e as variáveis básicas são x_1, x_2, x_3, x_6 . A solução básica viável B é degenerada. Pode-se escolher x_1, x_6 como sendo as variáveis básicas, ou mesmo x_1, x_5 ou x_5, x_6 . Logo existem três bases possíveis para a mesma solução básica viável B .

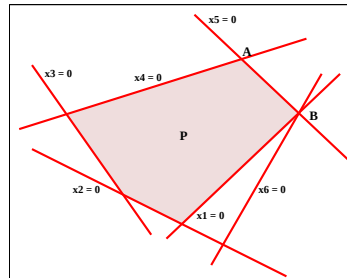


Figura 3.11: Representação de solução degenerada em poliedros na forma padrão.

Numa solução básica não degenerada exatamente $n - m$ das restrições $x_i \geq 0$ são ativas e as variáveis correspondentes são não básicas. No caso de uma solução básica degenerada mais do que $n - m$ destas restrições $x_i \geq 0$ são ativas, e normalmente existem vários modos de escolha das $n - m$ variáveis não-básicas e por isto, geralmente, existem várias bases correspondentes para a mesma solução básica. Contudo, existem exemplos de soluções básicas degeneradas para a qual existe apenas uma base correspondente.

A degeneração de uma solução básica viável não é, em geral, uma propriedade geométrica, mas também ela pode depender de uma representação particular do poliedro. Para observar isto, considere o poliedro $P_1 = \{(x_1, x_2, x_3) \mid x_1 - x_2 = 0; x_1 + x_2 + 2x_3 = 2, x_1, x_2, x_3 \geq 0\}$ na forma padrão. tem-se que $n = 3$, $m = 2$ e $n - m = 1$. O vetor $(1, 1, 0)$ é uma solução não degenerada, pois somente uma variável é zero. A solução $(0, 0, 1)$ é degenerada pois duas variáveis são nulas. Contudo, o mesmo poliedro pode ser reescrito na forma, não padrão, $P_2 = \{(x_1, x_2, x_3) \mid x_1 - x_2 = 0; x_1 + x_2 + 2x_3 = 2, x_1, x_3 \geq 0\}$. Agora a solução $(0, 0, 1)$ é básica viável não degenerada, pois só existem três restrições ativas.

Um outro exemplo pode ser obtido considerando a solução básica viável não degenerada

x^* de um poliedro na forma padrão $P = \{x \mid \mathcal{A}x = b, x \geq 0, b \geq 0\}$, onde $\mathcal{A}$ é de dimensão $m \times n$. Em particular, exatamente $n - m$ das variáveis x_i^* são iguais a zero. Agora, represente P na forma $P = \{x \mid \mathcal{A}x \geq b, -\mathcal{A}x \geq -b, x \geq 0\}$. Então, na solução básica viável x^* $n - m$ variáveis são tomadas como sendo zero e $2m$ restrições de desigualdades sendo transformada em restrições de igualdade. Logo, tem-se $n + m$ restrições ativas em x^* , portanto, ela é degenerada. Assim, na segunda representação, todas as soluções básicas viáveis são degeneradas. Com isto, pode-se concluir que uma solução básica viável é degenerada sob uma certa representação de um poliedro, mas pode ser não degenerada em outra representação.

3.5 Existência de Pontos Extremos

O objetivo desta seção é dar condições necessárias e suficientes para que um poliedro tenha no mínimo um ponto extremo. Primeiro observe que nem todo poliedro apresenta esta propriedade, por exemplo, se $n > 1$, um semiespaço do $\mathbb{R}^n$ é um poliedro que não possui ponto extremo. Além disso, se uma matriz $\mathcal{A}$ tem menos do que n linhas, então o poliedro $\{x \in \mathbb{R}^n \mid \mathcal{A}x \geq b\}$ não pode ter uma solução básica viável.

Na verdade a existência de um ponto extremo em um poliedro depende se ele contém ou não uma reta. Na Fig. 3.12, tem-se que o poliedro P contém uma reta, e por isto ele não tem pontos extremos, enquanto o poliedro Q não contém uma reta e ele possui pontos extremos.

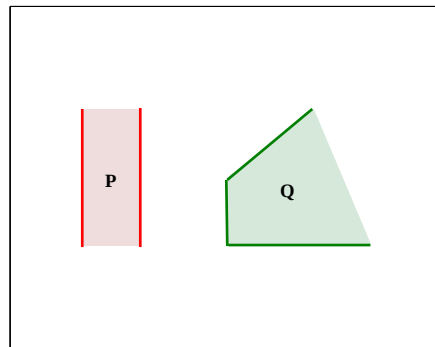


Figura 3.12: Representação de um poliedro que contém uma reta e de um que não contém.

Definição 3.5.1. Um poliedro $P \subset \mathbb{R}^n$ contém uma reta se existe um vetor $x \in P$ e um vetor não nulo $d \in \mathbb{R}^n$ tal que $x + \lambda d \in P$, para todo escalar λ .

Com esta definição chega-se ao seguinte resultado.

Teorema 3.5.1. Suponha que o poliedro $P = \{x \in \mathbb{R}^n \mid a_i x \geq b_i, i = 1, \dots, m\}$ seja não vazio. Então, as seguintes sentenças são equivalentes:

- a) O poliedro P tem, no mínimo, um ponto extremo.
- b) O poliedro P não contém uma reta.
- c) Existem n vetores sobre a família $a_1, \dots, a_m$ que são LI.

Demonstração 3.5.1. (b) $\Rightarrow$ (a): Seja x um elemento de P e seja $I = \{i \mid a_i x = b_i\}$. Se n destes vetores a_i ($i \in I$) correspondentes às restrições ativas são LI, então x é, por definição, uma solução básica viável e, portanto, existe uma solução básica viável. Se este não é o caso, então todos os vetores a_i ($i \in I$) estão num subespaço próprio do $\mathbb{R}^n$, logo existe um vetor não nulo $d \in \mathbb{R}^n$ tal que $a_i d = 0$, para todo $i \in I$. Considerando a reta formada por todos os pontos da forma $y = \lambda d$, onde λ é um escalar arbitrário, para todo $i \in I$ tem-se que $a_i y = a_i x + \lambda a_i d = a_i x = b_i$. Assim, estas restrições, que são ativas em x , permanecem ativas em todos os pontos da reta. Contudo, como é assumido que o poliedro não contém retas, segue que algumas das restrições estão sendo violadas, visto que vale para todo λ . No ponto onde alguma restrição foi violada uma nova restrição deve tornar-se ativa e assim conclui-se que existe algum λ^* e alguns $j \notin I$ tais que $a_j(x + \lambda^* d) = b_j$.

Observe que a_j não é uma combinação linear dos vetores a_i ($i \in I$), visto que, como $a_j x \neq b_j$, pois $j \notin I$, e $a_j(x + \lambda^* d) = b_j$, pela definição de λ^* , e então $a_j d \neq 0$. Logo, como $a_i d = 0$, $\forall i \in I$, pela definição de d , tem-se que d é ortogonal a qualquer combinação linear dos vetores a_i , $i \in I$. Portanto, como d não é ortogonal a a_j , tem-se que a_j não é uma combinação linear dos vetores a_i ($i \in I$). Assim, movendo x para $x + \lambda^* d$, o número de restrições ativa LI aumentou em pelo menos uma. Repetindo o mesmo argumento, tanto quanto necessário, chega-se no ponto onde existem n restrições ativas LI. Logo, tal ponto é,

por definição, uma solução básica viável, visto que ela foi construída no conjunto viável.

(a) $\Rightarrow$ (c): Se P tem um ponto extremo x , então x é também uma solução básica viável, pelo Teorema 3.2.1, e por isto existem n restrições ativas que são LI em x , que correspondem aos vetores a_i , sendo LI.

(c) $\Rightarrow$ (b): Suponha que n dos vetores a_i são LI e, sem perda de generalidade, assumamos que $a_1, \dots, a_n$ são os vetores LI. Suponha que P contenha uma reta $x + \lambda d$, onde d é um vetor não nulo. Assim, tem-se $a_i(x + \lambda d) \geq b_i$, para todo i e para todo λ . Deste modo tem-se que $a_i d = 0$, para todo i , pois tomando $a_i d < 0$ pode-se violar a restrição tomando λ suficientemente grande, e de maneira análoga tem-se que $a_i d$ não pode ser maior do que zero. Visto que os vetores a_i são LI, conclui-se que $d = 0$, o que gera um absurdo. Portanto P não contém uma reta.

□

Um poliedro limitado não contém uma reta, visto que qualquer reta contida neste poliedro viola uma das restrições. De maneira análoga, tem-se que o octante positivo $\{x \mid x \geq 0\}$ não contém uma reta. Visto que um poliedro na forma padrão está contido no octante positivo, então ele não contém uma reta.

Corolário 3.5.1. *Cada poliedro não vazio e limitado e cada poliedro não vazio na forma padrão tem no mínimo uma solução básica viável.*

Demonstração 3.5.1. *Um poliedro limitado não contém uma reta, e pelo Teorema 3.5.1 ele possui, no mínimo um ponto extremo, conseqüentemente possui uma solução básica viável (Teorema 3.2.1). Se um poliedro está na forma padrão, então, ele está contido no octante positivo $\{x \mid x \geq 0\}$, logo, possui um vértice e, novamente, pode-se concluir que possui uma solução básica viável.*

□

Uma interpretação geométrica para o Teorema 3.5.1 pode ser dada pela Fig. 3.13. Tome um ponto arbitrário de um poliedro, escolha uma direção onde toda restrição ativa permaneça ativa. Mova sobre esta direção até que uma nova restrição quase seja violada. Neste novo ponto

o número de restrições ativas LI aumentou em pelo menos uma unidade. Repita este processo até encontrar num ponto n restrições ativas LI, em tais pontos tem-se uma solução básica viável.

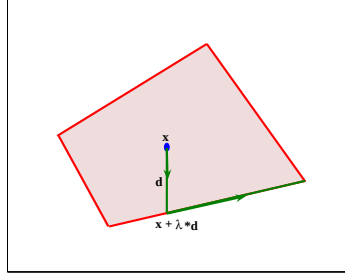


Figura 3.13: Representação da interpretação geométrica do Teorema 3.5.1.

3.6 Condições para que Pontos Extremos sejam Ótimos

Na seção anterior foi estabelecido condições para a existência de pontos extremos. Agora é estabelecido que num PPL sempre é possível achar uma solução ótima dentro do conjunto de pontos extremos do conjunto viável.

Teorema 3.6.1. *Considere o PPL de minimizar $c^T x$ sobre um poliedro P . Suponha que P tenha no mínimo um ponto extremo e que exista uma solução ótima. Então, existe uma solução ótima que é um ponto extremo de P .*

Demonstração 3.6.1. *Seja Q o conjunto de todas as soluções ótimas, que é assumido ser não vazio. Seja P da forma $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ e seja ϑ o valor ótimo do custo $c^T x$. Então $Q = \{x \in \mathbb{R}^n \mid Ax \geq b, c^T x = \vartheta\}$ também é um poliedro. Como $Q \subset P$ e como P não contém uma reta, conforme Teorema 3.5.1, então Q não contém uma reta. Logo, Q possui um ponto extremo.*

Seja x^* um ponto extremo de Q . É necessário mostrar que x^* também é um ponto extremo de P . Suponha, por absurdo, que x^* não é um ponto extremo de P . Então, existe $y \in P$ e $z \in P$ tal que $y \neq x^*$, $z \neq x^*$ e existe algum escalar λ tal que $x^* = \lambda y + (1 - \lambda)z$. Segue

que $\vartheta = c^T x = \lambda c^T y + (1 - \lambda)z$. Além disso, como ϑ é o custo ótimo, $c^T y \geq \vartheta$ e $c^T z \geq \vartheta$, por isto, tem-se que $c^T y = c^T z = \vartheta$. Logo, x^* não é um ponto extremo de Q , o que produz um absurdo (pela definição de x^*). Portanto, x^* é um ponto extremo de P e, sendo um elemento de Q , ele é um ponto ótimo.

□

Observe a Fig. 3.14 como uma ilustração do Teorema 3.6.1. Se Q é o conjunto das soluções ótimas, um ponto extremo x^* de Q também é um ponto extremo de P .

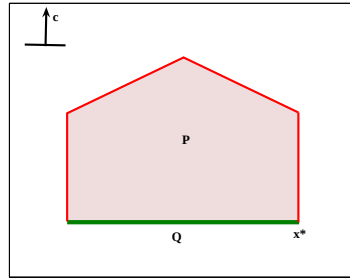


Figura 3.14: Representação da interpretação geométrica do Teorema 3.6.1.

O Teorema 3.6.1 é aplicado à poliedros na forma padrão ou a poliedros limitados, visto que ele não contém uma reta. O próximo resultado é mais forte, pois ele mostra que a existência de uma solução ótima pode ser admitida quando o custo ótimo for finito.

Teorema 3.6.2. *Considere o PPL de minimizar $c^T x$ sobre um poliedro P . Suponha que P tenha pelo menos um ponto extremo. Então, o custo ótimo é igual a $-\infty$ ou existe um ponto extremo que é ótimo.*

Demonstração 3.6.2. *Para esta demonstração é usada a seguinte terminologia: um determinado elemento x de P tem posto k se for possível achar k , mas não mais do que k , restrições LI que são ativas em x .*

Assuma que o custo ótimo é finito. Seja $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ e considere que algum $x \in P$ seja de posto $k < n$. Agora é mostrado que existe algum $y \in P$ que tem posto

maior e que satisfaz $c^T y \leq c^T x$. Seja $I = \{i \mid a_i x = b_i\}$, onde a_i é a i -ésima linha de A . Visto que $k < n$, o vetor a_i , $i \in I$ está num subespaço próprio do $\mathbb{R}^n$ e, por isto, é possível escolher algum $d \in \mathbb{R}^n$, não nulo, ortogonal a cada a_i , $i \in I$. Além disso, pode-se assumir que $c^T d \leq 0$, visto que se $c^T d \geq 0$, então pode-se usar $d_1 = -d$.

Suponha que $c^T d < 0$. Considere a semi-reta $y = x + \lambda d$, onde λ é um escalar positivo. Todos os pontos desta semi-reta satisfaz a relação $a_i y = b_i$, $i \in I$. Se a semi-reta completa está contida em P , o custo ótimo é $-\infty$, que não é o caso, já que por hipótese o custo ótimo é considerado finito. Logo, a semi-reta “sai” de P . Assim, onde isto acontece, existe algum $\lambda^* > 0$ e $j \notin I$ tal que $a_j(x + \lambda^* d) = b_j$. Tome $y = x + \lambda^* d$ e note que $c^T y < c^T x$. Tem-se que a_j é LI com a_i , $i \in I$, e o posto de y é no mínimo $k + 1$.

Suponha agora que $c^T d = 0$. Considere a reta $y = x + \lambda d$, onde λ é um escalar arbitrário. Visto que P não contém uma reta, a reta y vai sair de P e quando isto acontecer tem-se um vetor y cujo posto é maior do que o posto de x . Além disso, como $c^T d = 0$, tem-se que $c^T y = c^T x$.

Assim, criou-se um vetor y tal que $c^T y \leq c^T x$, cujo posto é maior do que o posto de x . Repetindo este argumento, o quanto for necessário, obtém-se um vetor w de posto n e, por isto, w é uma solução básica viável, tal que $c^T w \leq c^T x$.

Sejam $w^1, \dots, w^r$ as soluções básicas viáveis em P e seja w^* uma solução básica viável tal que $c^T w^* \leq c^T w^i, \forall i$. É fácil perceber que para cada x existe algum $i \in I$ tal que $c^T w^i \leq c^T x$. Logo, tem-se que $c^T w^* \leq c^T x$ para todo $x \in P$ e a solução básica viável w^* é ótima.

□

Para um PPL geral o conjunto viável não tem ponto extremo e, por isto, o Teorema 3.6.2 não pode ser aplicado diretamente. No sentido contrário, qualquer PPL pode ser transformado num PPL na forma padrão equivalente, para o qual o Teorema 3.6.2 pode ser aplicado. Isto estabelece o Corolário a seguir.

Corolário 3.6.1. *Considere o PPL de minimizar $c^T x$ sobre um poliedro não vazio. Então,*

ou o custo ótimo é $-\infty$ ou existe uma solução ótima.

O Corolário 3.6.1 deve ser entendido para se perceber as diferenças com uma função custo não linear. Por exemplo, no problema de minimizar $\frac{1}{x}$, sujeito a $x \geq 1$, o custo ótimo não é $-\infty$ e uma solução ótima pode não existir.

3.7 Representação de Poliedros Limitados

Até agora, os poliedros foram apresentados em termos das suas definições de desigualdade. Agora, nesta seção, é dada uma representação alternativa, mostrando que um poliedro limitado pode ser representado como sendo o envoltório convexo de seus pontos extremos. A prova desta afirmação é dada de maneira elementar e construtiva, sendo que a idéia pode ser retirada da Fig. 3.15. Dado um vetor z , expresse-o como sendo a combinação convexa de y e u . O vetor u é um elemento do poliedro Q , cuja dimensão é menor do que a dimensão de P . Usando indução na dimensão, expressa-se u como sendo uma combinação convexa dos pontos extremos de Q , que também são pontos extremos de P .

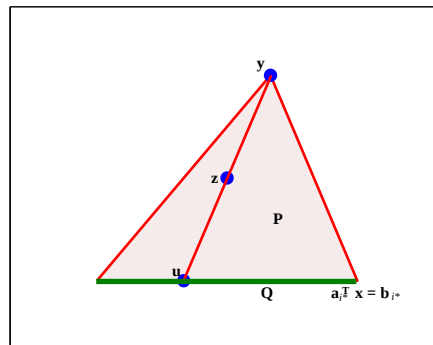


Figura 3.15: Representação da idéia geométrica para o Teorema 3.7.1.

Teorema 3.7.1. *Um poliedro limitado e não vazio é a casca convexa de seus pontos extremos.*

Demonstração 3.7.1. *Cada combinação convexa dos pontos extremos de um poliedro é um elemento do poliedro, visto que os poliedros são conjuntos convexos. Assim, somente é*

necessário mostrar a recíproca e que cada elemento de um poliedro limitado pode ser representado como sendo a combinação convexa de seus pontos extremos.

Defina a dimensão de um poliedro $P \subset \mathbb{R}^n$ como sendo o menor inteiro k tal que P está contido em algum subespaço afim k -dimensional do $\mathbb{R}^n$. Para demonstrar o teorema é usado indução na dimensão do poliedro P . Se P é poliedro de dimensão zero, ele é um ponto singular, logo os seus pontos são um ponto extremo de P e por isto o resultado é verdadeiro.

Agora assumamos que o resultado é verdadeiro para todo poliedro de dimensão menor do que k . Seja $P = \{x \in \mathbb{R}^n \mid a_i \geq b_i, i = 1, \dots, m\}$ um poliedro k -dimensional limitado e não vazio. Então, P está contido num subespaço afim k -dimensional do $\mathbb{R}^n$, que é assumido ser da forma

$$S = \{x^0 + \lambda_1 x^1 + \lambda_2 x^2 + \dots + \lambda_k x^k \mid \lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}\}, \quad (3.15)$$

onde $x^1, \dots, x^k$ são alguns vetores do $\mathbb{R}^n$. Seja $f_1, \dots, f_{n-k}$, $n-k$ vetores LI que são ortogonais a $x^1, \dots, x^k$. Seja $g_i = f_i^T x^0$, para $i = 1, \dots, n-k$. Então, cada elemento $x \in S$ satisfaz

$$f_i^T x = g_i, i = 1, \dots, n-k. \quad (3.16)$$

Visto que $P \subset S$, o mesmo acontecer com cada elemento de P .

Seja z um elemento de P . Se z é um ponto extremo de P , então z é uma combinação convexa trivial dos pontos extremos de P e não existe mais nada para ser provado. Se z não for um ponto extremo de P , escolha um ponto extremo y arbitrário de P e construa a semireta composta por todos os pontos da forma $z + \lambda(z - y)$, onde λ é um escalar não negativo. Visto que P é limitado, esta semireta sai de P e por isto ela viola uma das restrições, digamos $a_i^* x \geq b_i^*$. Considerando que esta violação acontece quando a semireta toca esta restrição é possível encontrar algum λ^* e um $u \in P$, tal que $u = z + \lambda^*(z - y)$ e $a_i^* u = b_i^*$. Visto que a restrição $a_i^* x \geq b_i^*$ é violada se λ ficar maior dos que λ^* , segue daí que $a_i^*(z - y) < 0$.

Seja Q o poliedro definido por $Q = \{x \in P \mid a_i^* x = b_i^*\} = \{x \in \mathbb{R}^n \mid a_i \geq b_i, i = 1, \dots, m; a_i^* x = b_i^*\}$. Visto que $z, y \in P$, tem-se que $f_i^T z = g_i = f_i^T y$, mostrando que $z - y$ é ortogonal a cada vetor f_i , para $i = 1, \dots, n-k$. Inversamente, mostra-se que $a_i^*(z - y) < 0$, implicando que o vetor a_i^* não é uma combinação linear dos vetores f_i e, por isto, ele é LI com os mesmos. Note que $Q \subset \{x \in \mathbb{R}^n \mid a_i x = b_i, f_i^T x = g_i, i = 1, \dots, n-k\}$, visto que a

Eq. (3.16) é tomada para cada elemento de P . Um conjunto adequado é definido por $n - k + 1$ restrições LI de igualdade. Contudo, ele é um subespaço afim de dimensão $k - 1$. Portanto, Q tem dimensão, no máximo, $k - 1$.

Aplicando a hipótese de indução em Q e u , conclui-se que u pode ser expresso como sendo uma combinação convexa $u = \sum_i \lambda_i v^i$ dos pontos extremos v^i de Q , onde λ_i são escalares não negativos cuja soma é um. Note que num ponto extremo y de Q , tem-se $a_i y = b_i$ para n vetores LI a_i , portanto, y também é um ponto extremo de P . Usando a definição de λ^* tem-se, também, $z = \frac{u + \lambda^* y}{1 + \lambda^*}$. Portanto, $z = \frac{\lambda^* y}{1 + \lambda^*} + \sum_i \frac{\lambda_i}{1 + \lambda^*} v^i$, mostrando que z é uma combinação convexa dos pontos extremos de P .

□

O Exemplo 3.7.1 mostra uma aplicação da Demonstração do Teorema 3.7.1. Os coeficientes $(\frac{1}{3}, \frac{1}{3}, \frac{1}{4}, \frac{1}{12})$, como exigidos na demonstração, tem soma 1.

Exemplo 3.7.1. Considere o poliedro $P = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z \leq 1, x, y, z \geq 0\}$.

Ele tem quatro pontos extremos, são eles $x^1 = (1, 0, 0)$, $x^2 = (0, 1, 0)$, $x^3 = (0, 0, 1)$ e $x^4 = (0, 0, 0)$. O ponto $x = (\frac{1}{3}, \frac{1}{3}, \frac{1}{4})$ pertence a P . Ele pode ser representado como sendo

$$x = \frac{1}{3}x^1 + \frac{1}{3}x^2 + \frac{1}{4}x^3 + \frac{1}{12}x^4. \quad (3.17)$$

Capítulo 4

O método Simplex

No capítulo anterior viu-se que se um PPL na forma padrão tem uma solução viável, então existe uma solução básica viável que é ótima. O método Simplex é baseado neste fato e busca por uma solução ótima movendo-se de um vértice à outro, da região viável, sempre numa direção que reduza o custo. Neste capítulo é desenvolvido detalhadamente o Método Simplex, discutindo implementações, como o Quadro Simplex e o Método do Simplex Revisado, falando também de algumas das dificuldades que podem surgir na presença de soluções degeneradas.

Considere o problema na forma padrão

$$\begin{array}{ll} \text{minimizar} & c^T x \\ \text{sujeito a} & \left\{ \begin{array}{l} \mathcal{A}x = b \\ x \geq 0 \end{array} \right. \end{array} \quad (4.1)$$

e P o conjunto viável. Também é assumido que a matriz $\mathcal{A}$ tem dimensão $m \times n$ e que suas linhas são LI. Ainda é mantida a notação $\mathcal{A}_i$ como sendo a i -ésima coluna da matriz $\mathcal{A}$ e a_i como sendo a sua i -ésima linha.

4.1 Condição para a Existência de Pontos Ótimos

Qualquer algoritmo de otimização tem a seguinte estrutura: dada uma solução viável deve-se buscar na sua vizinhança uma solução viável aproximada que reduza o custo da função. Se a

solução viável aproximada não leva a um custo melhor chega-se ao fim do algoritmo e então tem-se uma solução ótima localmente. Para problemas gerais de otimização, uma solução ótima local não precisa ser ótima, globalmente. Para PPL, se um ponto é ótimo localmente ele também é ótimo global, visto que se minimiza uma função convexa sobre um conjunto convexo.

Suponha que se esteja num ponto $x \in P$ e que seja considerado movimentos na direção de um vetor $d \neq 0$. Claramente é preciso escolher somente d que não leve imediatamente para fora do conjunto viável, fato este ilustrado pela Fig. 4.1, pois tem-se várias “direções viáveis” em vários pontos de P .

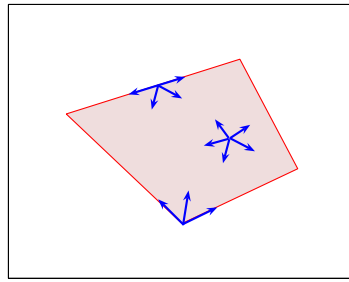


Figura 4.1: Representação da ilustração da Definição 4.1.1.

Definição 4.1.1. *Seja x um elemento de um poliedro P . Um vetor $d \in \mathbb{R}^n$ é dito ser uma **direção viável** em x se existe um escalar $\theta > 0$ para o qual $x + \theta d \in P$.*

Considere a possibilidade de movimento fora de x , por um novo vetor $x + \theta d$, selecionando uma variável não básica x_j , que inicialmente é levada a zero, acrescentando a ela um valor positivo θ , que conserva a variável não básica sendo zero. Algebricamente, $d_j = 1$ e $d_i = 0$ para todo índice não básico i tal que $i \neq j$. Ao mesmo tempo, o vetor x_B , das variáveis básicas, transforma-se no vetor $x_B + \theta d_B$, onde $d_B^T = \begin{bmatrix} d_{B(1)} & d_{B(2)} & \cdots & d_{B(m)} \end{bmatrix}$ é o vetor das componentes de d que corresponde as variáveis básicas.

Dado que o interesse está apenas nas soluções viáveis, deve-se exigir que $\mathcal{A}(x + d\theta) = b$. Como x é viável, também tem-se que $\mathcal{A}x = b$. Deste modo as restrições de igualdade são satisfeitas, para $\theta > 0$, se $\mathcal{A}d = 0$. Como $d_j = 1$ e que $d_i = 0$, para todos os outros índices não básicos i , tem-se

$$0 = \mathcal{A}d = \sum_{i=1}^n \mathcal{A}_i d_i = \sum_{i=1}^m \mathcal{A}_{B(i)} d_{B(i)} + \mathcal{A}_j = \mathcal{B}d_B + \mathcal{A}_j. \quad (4.2)$$

Sabendo que a matriz básica $\mathcal{B}$ é invertível, obtém-se

$$d_B = -\mathcal{B}^{-1} \mathcal{A}_j. \quad (4.3)$$

O vetor direção d que foi construído na Eq. (4.3) é chamado de **a j -ésima direção básica**. Assim, garante-se que as restrições de igualdade são respeitadas quando se move fora de x na direção básica d . Observe que as variáveis x_j sofre um pequeno aumento e que todas as outras variáveis não básicas são tomadas como sendo zero. Então, sobre a não-negatividade das restrições, é preciso se preocupar apenas com as variáveis básicas. Por isto, aparecem dois casos:

a) Suponha que x seja uma solução básica viável não degenerada. Então, $x_B > 0$, do qual segue que $x_B + \theta d_B \geq 0$, mantendo o fato de ser viável quando θ é suficientemente pequeno. Em particular, d é uma direção viável.

b) Agora suponha que x é degenerada. Então, d nem sempre é uma direção viável. Na verdade, ela é viável se uma variável básica $x_{B(i)}$ for nula, enquanto a componente correspondente $d_{B(i)}$ de $d_B = \mathcal{B}^{-1} \mathcal{A}_j$ for negativa. Neste caso, se seguir a j -ésima direção básica, a não negatividade da restrição para $x_{B(i)}$ é violada imediatamente, o que nos leva a uma solução inviável.

Por exemplo, na Fig. 4.2 tem-se $n = 5$ restrições e, como a representação do PPL é no plano, tem-se que $n - m = 2$ e, por isto, é possível visualizar o conjunto viável numa representação plana. Na solução viável não degenerada E , as variáveis x_1 e x_3 são nulas, não básicas, e x_2 , x_4 e x_5 são variáveis básicas positivas. A primeira direção básica é obtida aumentando x_1 , enquanto conserva-se a variável não básica x_3 sendo nula. Esta direção corresponde à borda EF . Agora considere a solução básica viável degenerada F e seja x_3 e x_5 as variáveis não básicas. Note que x_1 é uma solução básica viável nula. Uma direção básica é obtida aumentando x_3 , enquanto

conserva-se a direção básica x_5 nula. Esta é a reta FG , que pode conduzir para fora da região viável. Assim, esta direção básica não é uma direção viável.

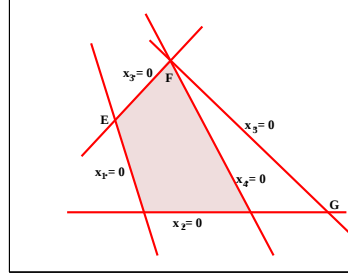


Figura 4.2: Representação da ilustração da direção básica em soluções degeneradas.

Agora veja o que acontece com a função custo quando há um movimento sobre a direção básica. Seja d a i -ésima direção básica. Então, a taxa $c^T d$ do custo que se altera ao longo da direção d é dada por $c_B^T d_B + c_j$, onde $c_B^T = \begin{bmatrix} c_1 & c_2 & \dots & c_m \end{bmatrix}$. Pela Eq. (4.3) tem-se que isto é o mesmo que $c_j - c_B^T \mathcal{B}^{-1} \mathcal{A}_j$. Uma interpretação intuitiva diz que c_j é o custo aumentado por unidade da variável x_j e o termo $-c_B^T \mathcal{B}^{-1} \mathcal{A}_j$ é o custo de se alterar a compensação na variável básica necessária para a restrição $\mathcal{A}x = b$.

Definição 4.1.2. *Seja x uma solução básica, seja $\mathcal{B}$ a matriz básica associada e seja c_B o vetor custo da variável básica. Para cada j , é definido como sendo o custo reduzido $\bar{c}_j$ da variável x_j o valor dado por*

$$\bar{c}_j = c_j - c_B^T \mathcal{B}^{-1} \mathcal{A}_j \quad (4.4)$$

O exemplo a seguir apresenta uma aplicação para a Definição 4.1.2.

Exemplo 4.1.1. *Considere o PPL*

$$\begin{aligned}
& \text{minimizar:} && F(x) = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 \\
& \text{sujeito a} && \begin{cases} x_1 + x_2 + x_3 + x_4 = 2 \\ 2x_1 + 3x_3 + 4x_4 = 2 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}
\end{aligned} \tag{4.5}$$

As duas primeiras colunas da matriz $\mathcal{A}$ são $\mathcal{A}_1 = (1, 2)$ e $\mathcal{A}_2 = (1, 0)$. Como elas são LI, escolha x_1 e x_2 como sendo as variáveis básicas. Logo, a matriz básica correspondente é dada por

$$\mathcal{B} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}. \tag{4.6}$$

Tomando $x_3 = x_4 = 0$ e resolvendo o sistema para x_1 e x_2 obtém-se que $x_1 = 1$ e $x_2 = 1$, que é uma solução básica viável não degenerada. Uma direção básica correspondente para um aumento na variável não básica x_3 pode ser construída da seguinte forma: tome $d_3 = 1$ e $d_4 = 0$. Utilizando a Eq. (4.3), uma mudança de direção viável é dada por

$$\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} d_{B(1)} \\ d_{B(2)} \end{bmatrix} = d_B = -\mathcal{B}^{-1}\mathcal{A}_3 = \mathcal{B}^{-1} \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ \frac{1}{2} \end{bmatrix}. \tag{4.7}$$

O custo sobre esta direção básica é dado por $c^T d = -\frac{3}{2}c_1 + \frac{1}{2}c_2 + c_3$.

Considere agora a Definição 4.1.2 para o caso de uma variável básica. Visto que $\mathcal{B}$ é a matriz $[\mathcal{A}_{B(1)} \cdots \mathcal{A}_{B(m)}]$, tem-se que $\mathcal{B}^{-1}[\mathcal{A}_{B(1)} \cdots \mathcal{A}_{B(m)}] = \mathcal{I}$, onde $\mathcal{I}$ é a matriz identidade de ordem m . Em particular, $\mathcal{B}^{-1}\mathcal{A}_{B(i)}$ é a i -ésima coluna da matriz identidade, que é o i -ésimo vetor unitário e_i . Logo, para cada variável básica $x_{B(i)}$, tem-se:

$$\bar{c}_{B_i} = c_{B_i} - \mathbf{c}_{B_i}^T \mathcal{B}_{-1} \mathcal{A}_{B(i)} = c_{B_i} - \mathbf{c}_{B_i}^T \mathbf{e}_i = c_{B_i} - c_{B_i} = 0, \quad (4.8)$$

o que mostra que o custo reduzido de cada variável básica é zero.

O próximo resultado fornece condições para que um ponto seja ótimo. Dada a interpretação da redução do custo como sendo a taxa da mudança de custo sobre certa direção, este resultado fica intuitivo.

Teorema 4.1.1. *Considere uma solução básica viável $\mathbf{x}$ associada com a matriz básica $\mathcal{B}$ e seja $\bar{\mathbf{c}}$ o vetor custo reduzido correspondente.*

a) *Se $\bar{\mathbf{c}} \geq 0$ então $\mathbf{x}$ é ótimo.*

b) *Se $\mathbf{x}$ é uma solução ótima e $\mathbf{x}$ é não degenerada então $\bar{\mathbf{c}} \geq 0$.*

Demonstração 4.1.1. a) *Assuma que $\bar{\mathbf{c}} \geq 0$ e seja $\mathbf{y}$ uma solução viável arbitrária. Defina $\mathbf{d} = \mathbf{y} - \mathbf{x}$. Por serem soluções possíveis tem-se que $\mathcal{A}\mathbf{y} = \mathcal{A}\mathbf{x} = \mathbf{b}$ e, por isto, $\mathcal{A}\mathbf{d} = 0$. Assim, reescrevendo a última igualdade obtém-se:*

$$\mathcal{B}\mathbf{d}_B + \sum_{i \in N} \mathcal{A}_i \mathbf{d}_i = 0, \quad (4.9)$$

onde N é o conjunto das índices correspondente às variáveis dadas não básicas. Como $\mathcal{B}$ é invertível obtém-se:

$$\mathbf{d}_B = - \sum_{i \in N} \mathcal{B}^{-1} \mathcal{A}_i \mathbf{d}_i, \quad (4.10)$$

e

$$\mathbf{c}^T \mathbf{d} = \mathbf{c}_B^T \mathbf{d}_B + \sum_{i \in N} \mathbf{c}_i \mathbf{d}_i = \sum_{i \in N} (\mathbf{c}_i - \mathbf{c}_B^T \mathcal{B}^{-1} \mathcal{A}_i) \mathbf{d}_i = \sum_{i \in N} \bar{\mathbf{c}}_i \mathbf{d}_i. \quad (4.11)$$

Para qualquer índice $i \in N$, não básico, tem-se que $\mathbf{x}_i = 0$ e, visto que $\mathbf{y}$ é viável, $\mathbf{y}_i \geq 0$.

Assim, $d_i \geq 0$ e $\bar{c}_i d_i \geq 0$, para todo $i \in N$. Por isto, conclui-se que $c^T(y - x) = c^T d \geq 0$. Visto que y é uma solução viável arbitrária, tem-se que x é ótima.

b) Suponha que x seja uma solução básica viável não degenerada e que $\bar{c}_j < 0$ para algum j . Visto que o custo reduzido de uma variável básica é sempre nulo, x_j deve ser uma variável não básica e, $\bar{c}_j$, é a taxa de mudança do custo sobre a j -ésima direção básica. Como x é não degenerada, a j -ésima direção básica é uma direção viável na redução do custo. Movendo-se nesta direção, obtém-se soluções viáveis cujo custo é menor do que o de x e, por isto, x não é uma solução ótima, o que gera uma contradição.

□

Note que o Teorema 4.1.1 admite a possibilidade de que x seja uma solução básica viável (degenerada) ótima, mesmo com $\bar{c}_j < 0$ para algum índice não básico j . Ainda, segundo o Teorema 4.1.1, quando uma solução básica viável não degenerada é ótima todos os custos reduzidos são não negativos, e por isto, é necessário analisar apenas as $n - m$ direções básicas. Se x é uma solução básica viável não degenerada, um simples teste computacional de comparação é necessário para determinar se x é ótimo ou não. O Método Simplex, como é desenvolvido na seção seguinte, é capaz de obter a solução ótima desviando-se desta dificuldade.

Note que para usar o Teorema 4.1.1 e para afirmar que uma certa solução é ótima, é necessário satisfazer duas condições: ser viável e não degenerada. Isto conduz à definição a seguir.

Definição 4.1.3. Uma matriz básica $\mathcal{B}$ é dita ser ótima se:

- (a) $\mathcal{B}^{-1}b \geq 0$; e
- (b) $\bar{c}^T = c^T - c_B^T \mathcal{B}^{-1} \mathcal{A} \geq 0^T$.

É claro que, se for achado uma base ótima, então a solução básica correspondente é viável e, ainda, satisfaz as condições de ser ótima. Por outro lado, no caso degenerado, tendo uma solução básica viável não significa necessariamente que o custo reduzido é não negativo.

4.2 Desenvolvimento do Método Simplex

Agora é completado o desenvolvimento do método simplex. A principal tarefa é apresentar os detalhes de como se mover para uma solução básica viável melhor, sempre que uma direção básica vantajosa é encontrada.

Assuma que cada solução básica viável é não degenerada. Esta suposição é útil no início, porém esta hipótese é enfraquecida posteriormente. Suponha que se esteja numa solução básica viável x e que se tenha calculado os custos reduzidos $\bar{c}_j$ das variáveis não básicas. Se todos estes custos são não negativos, então, pelo Teorema 4.1.1 tem-se que esta solução é ótima e por isto finalizou-se o procura. Caso contrário, pelo menos um dos custos reduzidos $\bar{c}_j$, das variáveis não básicas, é negativo. Se j é este índice, então esta direção é uma direção viável para reduzir o custo. A direção é obtida tomando $d_j = 1$, $d_i = 0$, para $i \neq B(1), B(2), \dots, B(m), j$ e $d_B = -B^{-1}A_j$. Quando se move nesta direção d a variável não básica torna-se positiva e todas as outras variáveis não básicas permanecem nulas. Esta situação foi descrita para dizer que x_j (ou A_j) “entra” na base.

Uma vez que se começa o movimento contrário de x sobre a direção d , estão sendo traçados pontos da forma $x + \theta d$, onde $\theta \geq 0$. Como o custo decresce sobre a direção d , deseja-se mover-se sobre ela o quanto for possível. Isto é tomado no ponto $x + \theta^* d$, onde

$$\theta^* = \max\{\theta \geq 0 \mid x + \theta d \in P\}. \quad (4.12)$$

O custo reduzido alterado é $\theta^* c^T d$, que é o mesmo valor de $\theta^* \bar{c}_j$.

Agora é apresentada uma maneira de se encontrar θ^* . Dado que $Ad = 0$, tem-se que $A(x + \theta d) = Ax = b$, para todo θ , e as restrições de igualdade nunca devem ser violadas. Assim, $x + \theta d$ torna-se impossível somente quando uma das suas componentes torna-se negativa. Por isto, aparecem dois casos:

a) Se $d \geq 0$, então $x + \theta d \geq 0$, para todo $\theta \geq 0$. Assim, o vetor $x + \theta d$ nunca torna-se impossível e, por isto, $\theta^* = \infty$.

b) Se $d_i < 0$, para algum i , a restrição $x_i + \theta d_i \geq 0$ torna-se $\theta \leq -\frac{x_i}{d_i}$. Esta restrição em θ deve ser satisfeita para cada i tal que $d_i < 0$. Assim, o valor máximo possível para θ é

$$\theta^* = \min_{\{i|d_i < 0\}} \left\{ -\frac{x_i}{d_i} \right\} \quad (4.13)$$

Note que $\theta^* > 0$, visto que $x_{B_i} > 0$ para todo i , pois x é não degenerada. A seguir é apresentado um exemplo de como encontrar uma direção viável, saindo de uma solução viável x .

Exemplo 4.2.1. Considere o PPL apresentado no Exemplo 4.1.1:

$$\begin{array}{ll} \text{minimizar:} & F(x) = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 \\ \text{sujeito a} & \left\{ \begin{array}{l} x_1 + x_2 + x_3 + x_4 = 2 \\ 2x_1 + 3x_3 + 4x_4 = 2 \\ x_1, x_2, x_3, x_4 \geq 0 \end{array} \right. \end{array} \quad (4.14)$$

No Exemplo 4.1.1 tomou-se as duas primeiras colunas para formar a matriz básica $\mathcal{B}$. A solução básica viável associada com $\mathcal{B}$ é $x = (1, 1, 0, 0)$ e, também visto no Exemplo 4.1.1, o custo reduzido $\bar{c}_3$, da variável não básica x_3 , é

$$-\frac{3c_1}{2} + \frac{c_2}{2} + c_3. \quad (4.15)$$

Agora, suponha que o vetor custo é dado por $c = (2, 0, 0, 0)$. Neste caso, usando a Eq. (4.15), tem-se que $\bar{c}_3 = -3$. Visto que $\bar{c}_3$ é negativo, tem-se uma direção básica, correspondendo a esta variável. Assim, como $d_B = (-\frac{3}{2}, \frac{1}{2})$, $d_3 = 1$ e $d_4 = 0$, tem-se que $d = (-\frac{3}{2}, \frac{1}{2}, 1, 0)$. Agora, considere vetores da forma $x + \theta d$, com $\theta \geq 0$. Aumentando o valor de θ só a componente x_1 de x que decresce, pois $d_1 < 0$. O valor possível máximo de θ é dado por $\theta^* = -\frac{x_1}{d_1} = \frac{2}{3}$. Com este θ , obtém-se o ponto $y = x + \frac{2}{3}d = (0, \frac{4}{3}, \frac{2}{3}, 0)$. Do sistema dado pela Eq. (4.14), tomando as colunas $\mathcal{A}_2 = (1, 0)$ e $\mathcal{A}_3 = (1, 3)$, correspondendo às variáveis

não nulas do novo vetor y , tem-se que elas também são LI. Portanto, elas formam uma base e por isto o vetor y é uma solução básica viável. Em particular, a variável x_3 entrou na base e a variável x_1 saiu da base.

Visto que θ^* é escolhido e assumido ser finito, move-se para a nova solução viável $y = x + \theta^* d$. Visto que $x_j = 0$ e $d_j = 1$, tem-se que $y_j = \theta^* > 0$. Seja l o índice tal que

$$-\frac{x_{B(l)}}{d_{B(l)}} = \min_{\{i=1, \dots, m \mid d_{B(i)} < 0\}} \left(\frac{x_{B(i)}}{d_{B(i)}} \right) = \theta^* \quad (4.16)$$

em particular,

$$d_{B(l)} < 0, \quad (4.17)$$

e

$$x_{B(l)} + \theta^* d_{B(l)} = 0. \quad (4.18)$$

Observa-se que a variável básica $x_{B(l)}$ passou a valer zero, enquanto que a variável não básica x_j passou a ser um número positivo. Este fato sugere que x_j deve substituir $x_{B(l)}$ na base. Conseqüentemente, toma-se a nova matriz básica $\bar{\mathcal{B}}$ substituindo $\mathcal{A}_{B(l)}$ por $\mathcal{A}_j$. Assim, obtém-se a matriz

$$\bar{\mathcal{B}} = \begin{bmatrix} \mathcal{A}_{B(1)} & \cdots & \mathcal{A}_{B(l-1)} & \mathcal{A}_{B(j)} & \mathcal{A}_{B(l+1)} & \cdots & \mathcal{A}_{B(m)} \end{bmatrix}. \quad (4.19)$$

De forma equivalente, o conjunto de índices básicos $\{B(1), \dots, B(m)\}$ está sendo substituído por um novo conjunto de índices $\{\bar{B}(1), \dots, \bar{B}(m)\}$ dado por

$$\bar{B}(l) = \begin{cases} B(i), & i \neq l, \\ j, & i = l. \end{cases} \quad (4.20)$$

O resultado a seguir mostra que esta matriz $\bar{\mathcal{B}}$ é uma base e que a solução básica $y = x + \theta^* d$ é a solução básica viável associada com esta matriz.

Teorema 4.2.1. *a) As colunas $\mathcal{A}_{B(j)}$, para $j \neq l$, e $\mathcal{A}_j$ são LI e, portanto, $\bar{\mathcal{B}}$ é uma matriz básica.*

b) O vetor $y = x + \theta^* d$ é uma solução básica viável associada com a matriz básica $\bar{\mathcal{B}}$.

Demonstração 4.2.1. *a) Se os vetores $\mathcal{A}_{\bar{B}(j)}$, $i = 1, \dots, m$ são LD, então existem coeficientes $\lambda_1, \dots, \lambda_m$, nem todos nulos, tais que*

$$\sum_{i=1}^m \lambda_i \mathcal{A}_{\bar{B}(j)} = 0, \quad (4.21)$$

o que implica que

$$\sum_{i=1}^m \lambda_i \mathcal{B}^{-1} \mathcal{A}_{\bar{B}(j)} = 0, \quad (4.22)$$

e por isto os vetores $\mathcal{B}^{-1} \mathcal{A}_{\bar{B}(j)}$ também são LD. Por isto, mostrando que os vetores $\mathcal{A}_{B(j)}$, para $i \neq l$, e $\mathcal{A}_j$ são LI, chega-se a um absurdo, provando o resultado. Tem-se que $\mathcal{B}^{-1} \mathcal{B} = \mathcal{I}$. Visto que $\mathcal{A}_{B(i)}$ é a i -ésima coluna de $\mathcal{B}$, segue que os vetores $\mathcal{B}^{-1} \mathcal{A}_{B(i)}$, para $i \neq l$, são todos vetores unitários, exceto, provavelmente, o vetor na posição l e, por isto, todas as coordenadas na posição l dos vetores $\mathcal{B}^{-1} \mathcal{A}_{B(i)}$, para $i \neq l$, valem zero. Agora, tem-se que $\mathcal{B}^{-1} \mathcal{A}_j = -d_B$. A sua entrada de posição $-d_{B(l)}$ é diferente de zero, por definição de l . Por isto, $\mathcal{B}^{-1} \mathcal{A}_j$ é LI com os vetores $\mathcal{B}^{-1} \mathcal{A}_{B(i)}$, $i \neq l$. Então $\bar{\mathcal{B}}$ é uma matriz básica.

b) Tem-se que $y \geq 0$, $\mathcal{A}y = b$ e $y_i = 0$, para $i \neq \bar{B}(1), \dots, \bar{B}(m)$. Além disso, as colunas

$\mathcal{A}_{\bar{B}(1)}, \dots, \mathcal{A}_{\bar{B}(m)}$ são LI, como mostrado em (a). Então, tem-se que y é uma solução básica viável associada com a matriz básica $\bar{\mathcal{B}}$.

□

Visto que θ^* é um número positivo, a nova solução básica viável $x + \theta^*d$ é diferente de x . Como d tem a direção de decrescimento do custo, o custo desta nova solução básica viável é estritamente menor do que o custo da variável x . Portanto, o objetivo de diminuir o valor do custo foi alcançado.

Agora é possível resumir uma iteração do Método Simplex. Para isto, defina um vetor $u = (u_1, \dots, u_m)$ como sendo

$$u = -d_B = \mathcal{B}^{-1}\mathcal{A}_j, \quad (4.23)$$

onde $\mathcal{A}_j$ é a coluna correspondendo a variável que entra na base. Em particular, $u_i = -d_{B(i)}$, para $i = 1, \dots, m$.

Tabela 4.1: Uma iteração do Método Simplex:

-
- 1 - Inicie o processo com uma base constituída das colunas básicas $\mathcal{A}_{B(1)}, \dots, \mathcal{A}_{B(m)}$ e uma solução básica viável associada com esta base.
 - 2 - Calcule o custo reduzido $\bar{c}_j = c_j - c_B^T \mathcal{B}^{-1} \mathcal{A}_j$, para todo índice não básico j . Se todos forem não negativos, a presente solução básica é ótima, por isto, o algoritmo se encerra. Caso contrário, escolha o j para o qual $\bar{c}_j < 0$ e que $\bar{c}_j$ seja o menor possível.
 - 3 - Calcule $u = \mathcal{B}^{-1} \mathcal{A}_j$. Se nenhuma das componente de u é positiva, tem-se que $\theta^* = -\infty$ e que o custo ótimo também é $-\infty$ e, por isto, o algoritmo se finaliza.
 - 4 - Se alguma componente de u é positiva, seja
$$\theta^* = \min_{\{i=1, \dots, m | u_i > 0\}} \left\{ \frac{x_{B(i)}}{u_i} \right\}.$$
 - 5 - Seja l tal que $\theta^* = \frac{x_{B(l)}}{u_l}$. Forme uma nova base trocando $\mathcal{A}_{B(l)}$ com $\mathcal{A}_j$. Se y é a nova solução básica viável, o valor da nova variável básica é $y_j = \theta^*$ e $y_{B(i)} = x_{B(i)} - \theta^* u_i$, para $i \neq l$.
-

O Método Simplex é inicializado com uma solução básica viável e arbitrária, que para problemas na forma padrão, é garantido sua existência. O próximo resultado garante que, para

problemas não degenerado, o Método Simplex funciona corretamente e se finaliza num número finito de iterações.

Observação 4.2.1. *O algoritmo apresentado na Tab. 4.1 é o algoritmo original de Dantzig, mas outras modificações podem ser utilizadas para diminuir o custo computacional. Por exemplo, na etapa 2 você pode escolher o índice j de maneira que ele seja o de menor índice.*

Teorema 4.2.2. *Assuma que o conjunto viável é não vazio e que cada solução básica viável é não degenerada. Então, o Método Simplex termina depois de um número finito de iterações. Ao terminar, uma das possibilidades a seguir acontece:*

- a) *Tem-se uma base ótima $\mathcal{B}$ e uma solução básica viável associada com este ponto ótimo.*
- b) *Acha-se um vetor d satisfazendo $\mathcal{A}d = 0$, $d \geq 0$, $c^T d < 0$ e que o custo ótimo é $-\infty$.*

Demonstração 4.2.2. *Se o algoritmo, apresentado na Tab. 4.1 termina devido ao critério de parada do passo 2, então a condição de um ponto ser ótimo do Teorema 4.1.1 fica satisfeita, por isto $\mathcal{B}$ é uma base ótima e esta solução básica viável é ótima.*

Se o algoritmo termina porque o passo 3 foi satisfeito, então a solução x é básica viável e, por isto, descobriu-se uma variável não básica x_j tal que $\bar{c}_j < 0$ e que a direção básica correspondente d satisfaz $\mathcal{A}d = 0$ e $d \geq 0$. Em particular, $x + \theta d \in P$ para todo $\theta > 0$. Visto que $c^T d = \bar{c}_j < 0$, fazendo θ arbitrariamente grande, o custo pode ser tomado como sendo um número negativo arbitrariamente pequeno e, por isto, o custo ótimo é $-\infty$.

Em cada iteração o algoritmo faz mover-se numa distância de $\theta^* > 0$ sobre uma direção d que satisfaça $c^T d < 0$. Portanto, o custo de cada solução básica viável apresentada pelo algoritmo é estritamente menor do que o custo da anterior e, além disso, soluções básicas viáveis não são apresentadas pelo algoritmo duas vezes. Visto que existe um número finito de soluções básicas viáveis, o algoritmo vai finalizar-se.

□

4.2.1 O Método Simplex para Problemas Degenerados

Foi apresentado até aqui o desenvolvimento do método simplex sob a suposição de que toda solução básica viável é não degenerada. Agora, suponha que o mesmo algoritmo seja usado para PPL com solução degenerada. Então, as seguintes possibilidades podem aparecer no decorrer do algoritmo:

- a) Se a solução básica viável presente x é degenerada, θ^* pode ser nulo e, neste caso, a nova solução básica viável y é a mesma que x . Isto acontece se alguma variável básica $x_{B(l)}$ é igual a zero e a componente correspondente $d_{B(l)}$ do vetor direcional d é negativa. Contudo, definindo uma nova base $\bar{B}$, pela troca de $A_{B(l)}$ com A_j , o Teorema 4.2.1 continua sendo válido.
- b) Mesmo que θ^* seja positivo, é possível que aconteça de mais do que uma das variáveis básicas originais se torne nulas no novo ponto $x + \theta^*d$. Visto que somente uma delas deixa a base, as outras continuam na base com custo valendo zero e, por isto, a nova solução básica viável é degenerada.

Mudando de base e encontrando a mesma solução básica viável não é um problema. Na Fig. 4.3, pode-se observar que uma seqüência de tal mudança de bases pode levar a eventual descoberta de uma direção viável para reduzir o custo. Por outro lado, uma seqüência de mudanças de bases pode levar, de novo, à base inicial e, neste caso, o algoritmo pode ficar indefinidamente num “laço”. Degeneração pode causar vários problemas computacionais, como ciclagem ou problemas na análise pós-ótima. Por exemplo, suponha que o PPL da Fig. 4.3 está na forma padrão, com $n - m = 2$, posicionado no plano bidimensional definido pelas restrições de igualdade $Ax = b$. A solução básica viável x é degenerada. Se x_4 e x_5 são as variáveis não básicas, então as duas direções básicas correspondentes são os vetores $\mathbf{f}$ e $\mathbf{h}$. Para ambas as direções, tem-se $\theta^* = 0$. Contudo, se é executada uma mudança de base, com x_4 entrando na base e x_6 saindo, as duas direções básicas serão $\mathbf{h}$ e $-\mathbf{g}$, onde a direção $-\mathbf{g}$ é tomada quando x_5 está variando e x_6 é conservado em zero. Em particular, pode-se seguir na direção $\mathbf{h}$ para alcançar uma nova solução básica viável y com custo menor.

Caso o algoritmo entre num laço, ele apresenta um fenômeno que é chamado de *ciclagem*. Muitas vezes afirma-se, de maneira equivocada, que a ciclagem é um fenômeno raro e excepcional.

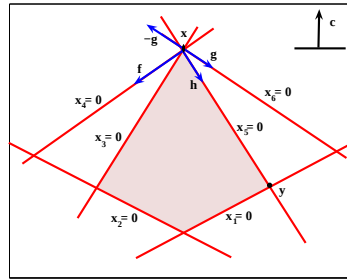


Figura 4.3: Representação de um PPL com uma solução degenerada.

Contudo, para PPL altamente estruturado, as soluções básicas viáveis são degeneradas e, por isto, ciclagem se torna bem provável de ocorrer. Uma maneira de se evitar ciclagem é fazer uma escolha sensata das variáveis que vão entrar e sair da base.

4.2.2 Seleção do Pivô

O algoritmo simplex, como está sendo descrito, tem certos graus de liberdade: na Etapa 2, da Tab. 4.1, pode-se escolher qualquer j cujo custo reduzido $\bar{c}_j$ seja negativo; já na Etapa 5 pode-se obter vários índices l que atinja o mínimo na definição de θ^* , e se está livre para escolher qualquer um deles. Regras para fazer tal escolha são chamadas de Regra de Pivoteamento.

Com respeito à escolha das colunas de entrada na base, as seguintes regras apresentam candidatos naturais:

- a) Escolha uma coluna $\mathcal{A}_j$, com $\bar{c}_j < 0$, de modo que o custo reduzido seja o menor. Visto que o custo reduzido é a taxa de variação da função custo, esta regra escolhe uma direção pela qual o custo decresce rapidamente com a taxa. Contudo, o custo atual decresce dependendo de até onde se move ao longo da direção escolhida.
- b) Escolha uma coluna com $\bar{c}_j < 0$ para o qual o custo correspondente decrescente $\theta^* |\bar{c}_j|$ seja mínimo. Esta regra apresenta uma possibilidade de alcançar o ponto ótimo num número mínimo de iterações. Por outro lado, o peso computacional em cada iteração é grande, visto que é necessário calcular θ^* para cada coluna com $\bar{c}_j < 0$.

Para PPL com grande números de restrições, até mesmo a regra para escolher $\bar{c}_j$ como sendo o de menor custo, pode apresentar um peso computacional grande, pois exige o cálculo do custo reduzido para cada variável. Por isto, podem ser adotadas regras simples para diminuir o custo computacional, tal como a regra do *menor sub-índice*, que escolhe o menor j para o qual $\bar{c}_j$ é negativo. Outro critério que tem sido usado para melhorar o tempo gasto global é a Regra da Máxima Descida, que pode ser vista em Vanderplaats (1999).

Para fazer a escolha da coluna de saída, uma alternativa simples também é dada pela regra do menor sub-índice, ou seja, dentre todas as variáveis elegíveis para sair da base, escolha a de menor sub-índice.

4.3 Implementação do Método Simplex

Nesta seção são discutidas alguns modos de aplicação do Método Simplex. Inicialmente deve estar claro, para a exposição do algoritmo, que os vetores $B^{-1}\mathcal{A}_j$ são pontos chave no método. Se estes vetores são avaliados então, o custo reduzido, a direção de movimento e o tamanho do passo θ^* são calculados facilmente. Assim, a principal diferença entre as implementações alternativas está no modo que os vetores $B^{-1}\mathcal{A}_j$ são calculados e na soma de informações que são carregadas para uma próxima iteração.

4.3.1 Implementação Simples

Esta implementação é a mais simples, visto que nenhuma informação auxiliar é carregada para a próxima iteração. No início de uma iteração típica, tem-se os índices $B(1), \dots, B(m)$ das variáveis básicas correspondentes. É formada uma matriz básica B e é calculado o vetor $p^T = c_B^T B^{-1}$, resolvendo o sistema $p^T B = c_B^T$ para o vetor desconhecido p . Este vetor p é chamado de *vetor do múltiplo simplex* associado com a base B . O custo reduzido $\bar{c}_j = c_j - c_B^T B^{-1} \mathcal{A}_j$ de qualquer variável x_j é obtido segundo a relação

$$\bar{c}_j = c_j - p^T \mathcal{A}_j. \quad (4.24)$$

Dependendo da regra de pivoteamento utilizada, é possível calcular todos os custos reduzidos ou apenas calcular o custo reduzido até que apareça uma variável com custo reduzido negativo. Visto que uma coluna $\mathcal{A}_j$ é selecionada para entrar na base, resolve-se o sistema linear $\mathcal{B}u = \mathcal{A}_j$, para obter o vetor u , dado por $u = \mathcal{B}^{-1}\mathcal{A}_j$. Assim, formou-se a direção para o qual se está movendo para fora da solução básica viável atual. Finalmente determine θ^* e a variável que vai sair da base, para que possa ser construída a nova solução básica viável.

4.3.2 O Método Simplex Revisado

O maior peso computacional que aparece na implementação do método simplex se deve à necessidade de resolver dois sistemas lineares. Por isto, numa implementação alternativa, os vetores $c_B^T \mathcal{B}^{-1}$ e $\mathcal{B}^{-1}\mathcal{A}_j$ são calculados pela multiplicação entre vetores e matrizes. Assim, é necessário um modo prático de se calcular cada entrada da matriz $\mathcal{B}^{-1}$, cada vez que a base seja alterada.

Seja

$$\mathcal{B} = \begin{bmatrix} \mathcal{A}_{B(1)} & \cdots & \mathcal{A}_{B(m)} \end{bmatrix} \quad (4.25)$$

a matriz básica no início de uma iteração e seja

$$\bar{\mathcal{B}} = \begin{bmatrix} \mathcal{A}_{B(1)} & \cdots & \mathcal{A}_{B(l-1)} & \mathcal{A}_j & \mathcal{A}_{B(l+1)} & \cdots & \mathcal{A}_{B(m)} \end{bmatrix} \quad (4.26)$$

a matriz básica no início da próxima iteração. Estas duas matrizes básicas tem as mesmas colunas, exceto que a l -ésima coluna $\mathcal{A}_{B(l)}$, que é a única que saiu da base, foi substituída por $\mathcal{A}_j$. Assim, é normal pensar que a matriz $\mathcal{B}^{-1}$ contenha informações que possam ser exploradas para se obter a nova matriz $\bar{\mathcal{B}}^{-1}$. A seguir é apresentada algumas termologias e resultados necessários para demonstrar que este pensamento está correto.

Definição 4.3.1. *Dada uma matriz, não necessariamente quadrada, a substituição de uma linha pela adição entre duas ou mais linhas, ou ainda, a multiplicação de uma linha por um escalar é chamada de **uma operação elementar sobre linha**.*

O exemplo a seguir sugere que a execução de uma operação elementar sobre linha de uma matriz $\mathcal{C}$ é equivalente a formar uma matriz $\mathcal{QC}$, onde $\mathcal{Q}$ é uma matriz quadrada construída de maneira conveniente.

Exemplo 4.3.1. *Seja $\mathcal{Q} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ e $\mathcal{C} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$. Note que $\mathcal{QC} = \begin{bmatrix} 11 & 14 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$.*

Em particular, a multiplicação a esquerda da matriz $\mathcal{Q}$ foi o mesmo que substituir a primeira linha por uma linha que é dada pela soma da primeira linha com o dobro da terceira linha da matriz $\mathcal{C}$.

Para generalizar o Exemplo 4.3.1, dada uma matriz $\mathcal{C}$, de ordem $m \times n$, deve-se observar que ao se multiplicar a j -ésima linha de $\mathcal{C}$ por β e somá-la com a i -ésima linha da mesma, sendo $i \neq j$, é o mesmo que multiplicar à esquerda de $\mathcal{C}$ uma matriz $\mathcal{Q} = \mathcal{I} + \mathcal{D}_{ij}$, onde $\mathcal{D}_{ij}$ é a matriz cujo os elementos valem zero, exceto, o (i, j) -ésimo elemento que vale β . Como o determinante de $\mathcal{Q}$ vale um, então esta matriz é invertível.

Suponha, agora, que uma seqüência de k operações elementares seja aplicada a uma matriz $\mathcal{C}$ e que cada operação corresponda a multiplicar à esquerda de $\mathcal{C}$ por uma certa matriz invertível $\mathcal{Q}_i$, $1 \leq i \leq k$. Então, a seqüência dessas operações elementares é o mesmo que multiplicar à esquerda por uma matriz invertível $\mathcal{Q} = \mathcal{Q}_k \mathcal{Q}_{k-1} (\cdots) \mathcal{Q}_2 \mathcal{Q}_1$. Assim, conclui-se que executando uma seqüência de operações elementares sobre a linha de uma matriz é o mesmo que multiplicar esta matriz, à esquerda, por uma certa matriz invertível.

Visto que $\mathcal{B}^{-1}\mathcal{B} = \mathcal{I}$, tem-se que $\mathcal{B}^{-1}\mathcal{A}_B(i)$ é o i -ésimo vetor unitário e_i . Com esta

observação, tem-se que

$$\mathcal{B}^{-1}\bar{\mathcal{B}} = \begin{bmatrix} e_1 & \cdots & e_{l-1} & u & e_{l+1} & \cdots & e_m \end{bmatrix} = \begin{bmatrix} 1 & & & u_1 & & & \\ & \ddots & & \vdots & & & \\ & & & u_l & & & \\ & & & \vdots & & \ddots & \\ & & & u_m & & & 1 \end{bmatrix}, \quad (4.27)$$

onde $u = \mathcal{B}^{-1}\mathcal{A}_j$. Agora, uma seqüência de operações elementares sobre linha é aplicada para mudar a matriz da Eq. (4.27) na matriz identidade. Para isto, considere a seguinte seqüência de operações elementares sobre linha:

- Para cada $i \neq l$, some $-\frac{u_i}{u_l}$ vezes a linha l com a linha i , como $u_l > 0$, tem-se que cada u_i se torna nulo.
- Divida a linha l por u_l , o que substitui u_l por 1.

Em outras palavras, soma-se um múltiplo escalar da linha l em cada linha i , com $i \neq l$, para substituir a l -ésima coluna u pelo l -ésimo vetor unitário e_l . Esta seqüência de operações elementares é equivalente a multiplicar, à esquerda, $\mathcal{B}^{-1}\bar{\mathcal{B}}$ por uma certa matriz invertível $\mathcal{Q}$. Visto que $\mathcal{Q}\mathcal{B}^{-1}\bar{\mathcal{B}} = \mathcal{I}$ então, tem-se que $\mathcal{Q}\mathcal{B}^{-1} = \bar{\mathcal{B}}^{-1}$. A última igualdade mostra que se é aplicada a mesma seqüência de operações elementares de linha para a matriz $\mathcal{B}^{-1}$, isto é, multiplicar à esquerda de $\mathcal{B}^{-1}$ a matriz $\mathcal{Q}$, obtém-se a matriz $\bar{\mathcal{B}}^{-1}$. Generalizando, para se obter qualquer $\bar{\mathcal{B}}^{-1}$, basta iniciar com uma matriz $\mathcal{B}^{-1}$ e aplicar a seqüência de operações elementares descrita acima para chegar à $\bar{\mathcal{B}}^{-1}$. O próximo exemplo, mostra como é feita esta construção.

Exemplo 4.3.2. Seja $\mathcal{B}^{-1} = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 3 & 1 \\ 4 & -3 & -2 \end{bmatrix}$ e $u = \begin{bmatrix} -4 \\ 2 \\ 2 \end{bmatrix}$.

Suponha que $l = 3$. Assim, o objetivo passa a ser transformar o vetor u no vetor unitário $e_3 = (0, 0, 1)$. Multiplique a linha 3 por 2 e some com a linha 1. Subtraia a linha 3 da linha

2. Divida a linha 3 por 2. Assim obtém-se

$$\bar{\mathcal{B}}^{-1} = \begin{bmatrix} 9 & -4 & -1 \\ -6 & 6 & 3 \\ 2 & -\frac{3}{2} & -1 \end{bmatrix} \quad (4.28)$$

Quando a matriz $\bar{\mathcal{B}}^{-1}$ é processada do modo descrito acima, obtém-se uma implementação do método simplex conhecido como sendo **O Método Simplex Revisado**, conforme algoritmo apresentado na Tab. 4.2.

Tabela 4.2: Uma Iteração do Método Simplex Revisado.

1 - Inicia-se com uma base constituída das colunas básicas $\mathcal{A}_{B(1)}, \dots, \mathcal{A}_{B(m)}$, uma solução básica viável associada com esta base e a inversa $\mathcal{B}^{-1}$ desta matriz básica.
2 - Calcule o vetor linha $p_j^T = c_B^T \mathcal{B}^{-1}$ e os valores dos custos reduzidos $\bar{c}_j = c_j - p_j^T \mathcal{A}_j$. Se todos forem não negativos, a presente solução básica é ótima, por isto, o algoritmo se encerra. Caso contrário, escolha algum j para o qual $\bar{c}_j < 0$.
3 - Calcule $u = \mathcal{B}^{-1} \mathcal{A}_j$. Se nenhuma das componente de u é positiva, tem-se que $\theta^* = -\infty$ e que o custo ótimo também é $-\infty$ e, por isto, o algoritmo se finaliza.
4 - Se alguma componente de u é positiva, seja
$\theta^* = \min_{\{i=1, \dots, m u_i > 0\}} \left\{ \frac{x_{B(i)}}{u_i} \right\}.$
5 - Seja l tal que $\theta^* = \frac{x_{B(l)}}{u_l}$. Forme uma nova base trocando $\mathcal{A}_{B(l)}$ com $\mathcal{A}_j$. Se y é a nova solução básica viável, o valor da nova variável básica é $y_j = \theta^*$ e $y_{B(i)} = x_{B(i)} - \theta^* u_i$, para $i \neq l$.
6 - Forme a $m \times (m+1)$ matriz $[\mathcal{B}^{-1} \mid u]$. Some a cada uma de suas linhas um múltiplo da l -ésima linha para fazer da última coluna o vetor unitário e_l . As primeiras m colunas do resultado obtido é a matriz $\bar{\mathcal{B}}^{-1}$.

4.3.3 Implementação do Quadro Simplex

Finalmente é possível descrever a implementação do Quadro Simplex, também chamado de Tabela Simplex ou de Quadro Completo. No quadro simplex ao invés de se trabalhar com a matriz $\mathcal{B}^{-1}$, trabalha-se com a matriz de ordem $m \times (n+1)$

$$\mathcal{B}^{-1} [b \mid \mathcal{A}] \quad (4.29)$$

com as colunas $\mathcal{B}^{-1}b, \mathcal{B}^{-1}\mathcal{A}_1, \dots, \mathcal{B}^{-1}\mathcal{A}_n$. A coluna $\mathcal{B}^{-1}b$ é chamada de **a coluna zero** e ela contém os valores das variáveis básicas. A coluna $\mathcal{B}^{-1}\mathcal{A}_i$ é chamada de **a i -ésima coluna da Tabela Simplex**. A coluna $u = \mathcal{B}^{-1}\mathcal{A}_j$, correspondente à variável que entra na base, é chamada de **a coluna pivô**. Se a l -ésima variável básica sai da base, a l -ésima linha da Tabela Simplex é chamada de **a linha pivô**. Finalmente, o elemento que pertence a intersecção da linha pivô com a coluna pivô é o elemento u_l é chamado de **o elemento pivô**. Tem-se que o elemento pivô é sempre positivo pois, caso contrário, o algoritmo termina, pelo passo três do algoritmo.

As informações contidas nas linhas da Tabela Simplex admitem a seguinte interpretação: as restrições de igualdade são dadas inicialmente na forma $b = \mathcal{A}x$. Dada a atual matriz básica $\mathcal{B}$, as restrições de igualdade podem ser expressas na forma

$$\mathcal{B}^{-1}b = \mathcal{B}^{-1}\mathcal{A}x \quad (4.30)$$

que é a informação contida na Tabela Simplex. Em outras palavras, as linhas da Tabela Simplex fornece os coeficientes das restrições de igualdade $\mathcal{B}^{-1}b = \mathcal{B}^{-1}\mathcal{A}x$.

No final de cada iteração, precisa-se atualizar os dados da tabela $\mathcal{B}^{-1} [b \mid \mathcal{A}]$, sendo calculado $\bar{\mathcal{B}}^{-1} [b \mid \mathcal{A}]$. Isto pode ser realizado multiplicando a esquerda da Tabela Simplex uma matriz $\mathcal{Q}$, que satisfaça $\mathcal{Q}\mathcal{B}^{-1} = \bar{\mathcal{B}}^{-1}$.

A respeito da determinação da coluna de saída $\mathcal{A}_{B(l)}$ e do tamanho do passo θ^* , os passos 4 e 5, que aparecem no resumo do Método Simplex tem o seguinte significado: $\frac{x_{B(l)}}{u_i}$ é a proporção da i -ésima entrada na coluna zero da Tabela Simplex pela i -ésima entrada da coluna pivô da Tabela. Somente é considerado os índices i 's onde $u_i > 0$. O menor valor encontrado nestas proporções é θ^* e o i correspondendo a este valor passa a ser l .

É usual estender a Tabela Simplex adicionando uma linha acima, chamada de **a linha zero**. A entrada no canto esquerdo do topo contém o valor $-c_B^T x_B$ que é o valor negativo do custo presente. O restante da linha zero é o vetor linha do custo reduzido das variáveis básicas, isto é, o vetor $\bar{c}^T = c^T - c_B^T \mathcal{B}^{-1}\mathcal{A}$. Assim, a estrutura da Tabela Simplex é:

Tabela 4.3: A Tabela Simplex

$-c_B^T \mathcal{B}^{-1} b$	$c^T - c_B^T \mathcal{B}^{-1} \mathcal{A}$
$\mathcal{B}^{-1} b$	$\mathcal{B}^{-1} \mathcal{A}$

e, escrevendo de forma detalhada,

Tabela 4.4: A Tabela Simplex escrita de maneira detalhada.

$-c_B^T x_B$	$\bar{c}_1$	$\dots$	$\bar{c}_n$
$x_{B(1)}$	$\left \begin{array}{c} \mathcal{B}^{-1} \mathcal{A}_1 \end{array} \right $	$\dots$	$\left \begin{array}{c} \mathcal{B}^{-1} \mathcal{A}_n \end{array} \right $
$\vdots$			
$x_{B(n)}$	$\left \begin{array}{c} \mathcal{B}^{-1} \mathcal{A}_1 \end{array} \right $	$\dots$	$\left \begin{array}{c} \mathcal{B}^{-1} \mathcal{A}_n \end{array} \right $

As regras utilizadas na linha zero são as mesmas usadas nas outras linhas da Tabela Simplex. Para provar que esta regra produz o resultado correto, para a linha zero, observe que no início de uma iteração típica, da Tabela Simplex, a linha zero é dada por

$$[0 \mid c^T] - g^T [b \mid \mathcal{A}], \quad (4.31)$$

onde $g^T = c_B^T$. A linha zero é igual a $[0 \mid c^T]$, mais uma combinação linear das linhas de $[b \mid \mathcal{A}]$. Seja j o índice da coluna pivô e l o índice da linha pivô. Assim, a linha pivô é da forma $h^T [b \mid \mathcal{A}]$, onde h^T é a l -ésima linha de $\mathcal{B}^{-1}$. Depois que um múltiplo da linha pivô é adicionado na linha zero, tal linha fica igual à $[0 \mid c^T]$ mais uma combinação, provavelmente diferente, das linhas de $[b \mid \mathcal{A}]$, digamos

$$[0 \mid c^T] - p^T [b \mid \mathcal{A}], \quad (4.32)$$

para algum vetor p . Lembre-se que a regra usada transforma o elemento correspondente em zero na linha zero:

$$c_{\bar{B}(l)} - p^T \mathcal{A}_{\bar{B}(l)} = c_j - p^T \mathcal{A}_j = 0. \quad (4.33)$$

Considere, agora, a $\bar{B}(i)$ -ésima coluna para $i \neq l$, onde esta coluna corresponde a uma variável básica que ficou na base. O valor correspondente a esta variável, na linha zero, é zero. Como $\mathcal{B}^{-1}\mathcal{A}_{B(i)}$ é o i -ésimo vetor unitário e $i \neq l$, então o valor correspondente a esta entrada, na linha pivô, também é zero. Portanto, somando um múltiplo da linha pivô com a linha zero da tabela tem-se que a entrada correspondente continua valendo zero. Logo, pode-se concluir que o vetor p satisfaz $c_{\bar{B}(i)} - p^T \mathcal{A}_{\bar{B}(i)} = 0$, para cada coluna $\mathcal{A}_{\bar{B}(i)}$ na nova base. Isto implica que $c_B^T - p^T \mathcal{B} = 0$, isto é, $p^T = c_B^T \mathcal{B}^{-1}$. Assim, os dados da linha zero da Tabela Simplex ficam igual à

$$[\mathbf{0} \mid c^T] - c_B^T \mathcal{B}^{-1} [b \mid \mathcal{A}], \quad (4.34)$$

como desejado.

A Tab. 4.5 apresenta uma implementação de uma iteração da Tabela Simplex.

Tabela 4.5: Uma Iteração da Implementação da Tabela Simplex.

1 - Inicia-se com uma tabela associada com uma matriz básica $\mathcal{B}$ e a solução básica viável x correspondente.
2 - Examine o custo reduzido na linha zero da Tabela Simplex. Se todos eles forem não negativos, a presente solução básica é ótima, por isto, o algoritmo se encerra. Caso contrário, escolha algum j para o qual $\bar{c}_j < 0$.
3 - Considere o vetor $u = \mathcal{B}^{-1}\mathcal{A}_j$, que é a coluna pivô da Tabela Simplex. Se nenhuma das componente de u é positiva, tem-se que $\theta^* = -\infty$ e que o custo ótimo também é $-\infty$ e, por isto, o algoritmo se finaliza.
4 - Para cada i tal que u_i é positivo, calcule o passo $\frac{x_{B(i)}}{u_i}$. Seja l o índice da linha que corresponde ao menor passo. A coluna $\mathcal{A}_{B(l)}$ sai da base e a coluna $\mathcal{A}_j$ entra na base.
5 - Some cada linha da Tabela Simplex com um múltiplo da linha pivô, de modo que cada entrada da linha pivô, exceto o elemento pivô, se transforme em zero e o elemento pivô em um.

O Exemplo a seguir, apresenta a solução de um PPL, pela Tabela Simplex, apresentando todos os passos de uma implementação.

Exemplo 4.3.3. Considere o PPL

$$\begin{array}{ll}
 \text{minimizar} & -10x_1 - 12x_2 - 12x_3 \\
 \text{sujeito a} & \left\{ \begin{array}{l} x_1 + 2x_2 + 2x_3 \leq 20 \\ 2x_1 + x_2 + 2x_3 \leq 20 \\ 2x_1 + 2x_2 + x_3 \leq 20 \\ x_1, x_2, x_3 \geq 0 \end{array} \right. .
 \end{array} \quad (4.35)$$

A Fig. 4.4 apresenta o conjunto viável deste problema.

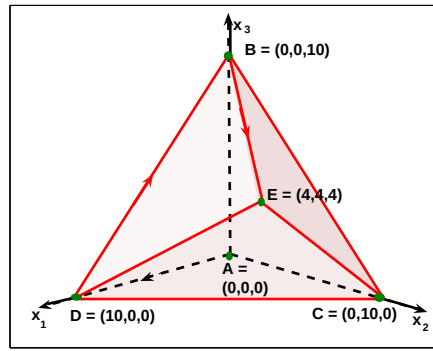


Figura 4.4: Representação da solução viável do PPL - Exemplo 4.3.3

Introduzindo as variáveis de folga, obtém-se o seguinte PPL na forma padrão

$$\begin{array}{ll}
 \text{minimizar} & -10x_1 - 12x_2 - 12x_3 \\
 \text{sujeito a} & \left\{ \begin{array}{l} x_1 + 2x_2 + 2x_3 + x_4 = 20 \\ 2x_1 + x_2 + 2x_3 + x_5 = 20 \\ 2x_1 + 2x_2 + x_3 + x_6 = 20 \\ x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{array} \right.
 \end{array} \quad (4.36)$$

Tomando x_4, x_5, x_6 , como sendo os vetores da base, a solução $\mathbf{x} = (0, 0, 0, 20, 20, 20)$ é a solução básica viável, e com ela é possível iniciar o algoritmo. Assim, tome $B(1) = 4$,

$B(2) = 5$ e $B(3) = 6$, com isto a matriz básica correspondente é a matriz identidade $\mathcal{I}$. Para obter a linha zero da Tabela Simplex inicial, observe que $c_B = 0$ e, portanto, $c_B^T x_B = 0$ e $\bar{c} = c$. Portanto, tem-se que a Tabela inicial é a Tab. 4.6.

Tabela 4.6: Tabela do Exemplo 4.3.3 - Passo Inicial

		x_1	x_2	x_3	x_4	x_5	x_6
	0	-10	-12	-12	0	0	0
$x_4 =$	20	1	2	2	1	0	0
$x_5 =$	20	2*	1	2	0	1	0
$x_6 =$	20	2	2	1	0	0	1

Note algumas convenções no formato da Tab. 4.6: o título x_i na primeira linha indica a posição que cada uma das variáveis ocupa na Tabela Simplex. O título $x_i =$, na primeira coluna, diz quais são as três variáveis básicas e em qual ordem elas ocupam a base. Por exemplo, a primeira variável básica $x_{B(1)}$ é x_4 , e o seu valor é 20. Analogamente, tem-se que $x_{B(2)} = x_5 = 20$ e $x_{B(3)} = x_6 = 20$. É claro que estes “rótulos”, para as variáveis, não são totalmente necessários, já que a coluna na Tabela associada com a primeira variável básica é o primeiro vetor unitário. Assim, como x_4 é o primeiro vetor unitário, tem-se que x_4 é a primeira variável básica.

O custo reduzido de x_1 é negativo, e usando o fato que ele tem o menor sub-índice, ele é colocado na base. Assim, a coluna pivô é $u = (1, 2, 2)$. Forme o passo $\frac{x_{B(i)}}{u_i}$, com i variando de 1 a 3. Os valores obtidos são $(20, 10, 10)$, e por isto os menores valores estão associados com $i = 2$ e $i = 3$. Escolha $l = 2$. Assim, o elemento pivô está na linha 2 e colina 1, e este é o elemento marcado com asterisco, na Tab. 4.6. A segunda variável básica, $x_5 =$, sai da base. A nova base é dada por $\bar{x}_{B(1)} = x_4$, $\bar{x}_{B(2)} = 1$ e $\bar{x}_{B(3)} = 6$. Para transformar a linha pivô no segundo vetor unitário, faça:

- 1º) multiplique a linha pivô por 5 e some com a linha zero;
- 2º) multiplique a linha pivô por $-\frac{1}{2}$ e some com a linha um;
- 3º) multiplique a linha pivô por -1 e some com a linha três;
- 4º) Divida a linha dois por 2.

Assim, a nova Tabela Simplex é dada pela Tab. 4.7:

Tabela 4.7: Tabela do Exemplo 4.3.3 - 1º Passo

		x_1	x_2	x_3	x_4	x_5	x_6
	100	0	-7	-2	0	5	0
$x_4 =$	10	0	3/2	1	1	-1/2	0
$x_1 =$	10	1	1/2	1	0	1/2	0
$x_6 =$	0	0	1*	-1	0	-1	1

Assim, a solução básica viável correspondente a nova base é $\mathbf{x} = (10, 0, 0, 10, 0, 0)$. Em relação as variáveis originais, x_1 , x_2 e x_3 e a Fig. 4.4, tem-se que saiu do ponto $A = (0, 0, 0)$ e deslocou-se para o ponto $D = (10, 0, 0)$. Está é uma solução básica degenerada, pois $x_6 = 0$. Novamente, na Fig. 4.4 este fato pode ser observado, pois tem-se quatro restrições que são ativa no ponto D .

As linhas da Tabela Simplex, exceto a linha zero, apresenta uma representação das restrições de igualdade $\mathcal{B}^{-1}\mathcal{A}\mathbf{x} = \mathcal{B}^{-1}\mathbf{b}$, que é o mesmo que $\mathcal{A}\mathbf{x} = \mathbf{b}$. Logo, usando a Tab. 4.7, as restrições de igualdade podem ser reescritas na forma equivalente:

$$\begin{cases} 10 = & \frac{3}{2}x_2 + x_3 + x_4 - \frac{1}{2}x_5 \\ 10 = x_1 + & \frac{1}{2}x_2 + x_3 + \frac{1}{2}x_5 \\ 0 = & x_2 - x_3 - x_5 + x_6 \end{cases} \quad (4.37)$$

Para continuar a resolver o PPL, observe que as variáveis x_2 e x_3 tem custo negativo. Novamente, escolhendo a de menor sub-índice x_2 para entrar na base, a coluna pivô fica sendo $\mathbf{u} = (\frac{3}{2}, \frac{1}{2}, 2)$. Como $\frac{x_{B(1)}}{u_1} = \frac{20}{3}$, $\frac{x_{B(2)}}{u_2} = 20$ e $\frac{x_{B(3)}}{u_3} = 0$, segue que o menor valor corresponde à variável x_6 , logo ele sai da base. O elemento pivô, que está na linha três e coluna dois, está marcado com asterisco na Tab. 4.7. Executando as operações elementares sobre linhas necessárias, obtém-se a Tab. 4.8, que é a Tabela Simplex no segundo passo.

Observando a Fig. 4.4, continua-se no ponto $D = (10, 0, 0)$, apenas a direção básica foi alterada. Neste ponto, x_3 é a única variável com custo negativo, logo esta é a única variável que pode entrar na base. Fazendo a análise $\frac{x_{B(i)}}{u_i}$, para $i = 1$ ou $i = 4$, tem-se que a variável

Tabela 4.8: Tabela Simplex do Exemplo 4.3.3 - 2º Passo

		x_1	x_2	x_3	x_4	x_5	x_6
	100	0	0	-9	0	-2	7
$x_4 =$	10	0	0	$5/2^*$	1	1	$-1/2$
$x_1 =$	10	1	0	$3/2$	0	1	$-1/2$
$x_2 =$	0	0	1	-1	0	-1	1

x_4 sai da base, a Tabela Simplex resultante é dada pela Tab. 4.9.

Tabela 4.9: Tabela Simplex do Exemplo 4.3.3 - 3º Passo

		x_1	x_2	x_3	x_4	x_5	x_6
	136	0	0	0	$18/5$	$8/5$	$26/5$
$x_3 =$	4	0	0	1	$2/5$	$2/5$	$-1/5$
$x_1 =$	4	1	0	0	$2/5$	$2/5$	$-1/5$
$x_2 =$	4	0	1	0	$-3/5$	$-3/5$	$4/5$

Com este movimento, alcança-se o ponto $E = (4, 4, 4)$ na Fig. 4.4. Este é realmente um ponto ótimo, visto que todos os custos das variáveis são não negativos.

Neste exemplo, foi necessário a mudança de base três vezes para que fosse alcançado o ponto ótimo. O caminho utilizado na Fig. 4.4 foi $A - D - D - E$. Usando outras regras de obtenção de pivô, outro caminho pode ser traçado. Uma pergunta que aparece é: o Método Simplex poderia resolver o sistema traçando o caminho $A - D - E$? A resposta é não, já que a base final e a base inicial tem três colunas diferentes, o que exige que seja feita, no mínimo, três mudanças de bases.

4.4 O Método Simplex Duas Fases

Para iniciar o Método Simplex, é preciso achar uma solução básica viável inicial. Algumas vezes este trabalho é fácil. Veja o seguinte exemplo: suponha um problema envolvendo restrições da forma $Ax \leq b$, onde $b \geq \mathbf{0}$. Pode-se, então, introduzir variáveis de folga, não negativas, e reescrever as restrições da seguinte maneira $Ax + y = b$. O Vetor (x, y) definido por $x = \mathbf{0}$ e $y = b$, é uma solução básica viável e a matriz básica correspondente é a matriz identidade. Em geral, esta facilidade de achar uma solução básica viável não ocorre e, por isto, é necessário resolver um PPL

auxiliar, como é mostrado a seguir.

Considere o PPL

$$\begin{array}{ll} \text{minimizar} & c^T x \\ \text{sujeito a} & \left\{ \begin{array}{l} Ax = b \\ x \geq 0 \end{array} \right. \end{array} \quad (4.38)$$

Assuma, sem perda de generalidade, que $b \geq 0$ pois, caso contrário, multiplique as restrições onde $b_i < 0$ por -1 . Agora, introduza um vetor de variáveis artificiais $x^a \in \mathbb{R}^n$ e use o método simplex para resolver o PPL auxiliar

$$\begin{array}{ll} \text{minimizar} & x_1^a + \cdots + x_m^a \\ \text{sujeito a} & \left\{ \begin{array}{l} Ax + x^a = b \\ x \geq 0 \\ x^a \geq 0 \end{array} \right. \end{array} \quad (4.39)$$

A inicialização do PPL auxiliar fica fácil, pois basta tomar $x = 0$ e $x^a = b$ que uma solução básica viável é formada, cuja matriz básica é a matriz identidade.

Se x é uma solução básica viável para o problema original, esta escolha de x , junto com $x^a = 0$ produz uma solução com custo zero para o problema auxiliar. Portanto, se o custo ótimo no problema auxiliar é não nulo, conclui-se que o problema original é inviável. Por outro lado, se o custo obtido no PPL auxiliar é nulo, se é satisfeito a relação $x^a = 0$, x é uma solução viável para o problema original.

Até aqui o problema foi resolvido apenas parcialmente, pois este método só é capaz de detectar se o problema original é inviável ou se é possível achar uma solução viável para o problema. Contudo, para que o Método Simplex seja iniciado, no problema original, necessita-se de uma solução básica viável, associada com a matriz básica B , ou até mesmo a Tabela Simplex associada. Tudo isto é realizado se o Método Simplex é aplicado no problema auxiliar, terminando

com uma matriz básica $\mathcal{B}$ consistindo exclusivamente das colunas de $\mathcal{A}$. Pode-se simplesmente “abandonar” as colunas que correspondem às variáveis artificiais e continuar o Método Simplex no problema original, usando $\mathcal{B}$ como sendo a matriz básica inicial.

O processo de tirar as variáveis artificiais da base é um pouco mais complexo se o problema original é viável, já que o Método Simplex aplicado ao problema auxiliar termina com uma solução viável x^* para o problema original, mas algumas das variáveis artificiais podem estar na base final. Visto que o valor final das variáveis artificiais é zero, isto implica que se tem uma solução básica viável degenerada para o problema auxiliar. Seja k ($k < m$) o número de colunas de $\mathcal{A}$ que pertencem à base final e, sem perda de generalidade, assuma que essas são as colunas $\mathcal{A}_{B(1)}, \dots, \mathcal{A}_{B(k)}$. Note que as colunas $\mathcal{A}_{B(1)}, \dots, \mathcal{A}_{B(k)}$ devem ser LI, visto que elas são parte de uma matriz básica. Sobre a suposição inicial de que a matriz $\mathcal{A}$ tem posto completo, as colunas de $\mathcal{A}$ podem ser estendidas ao $\mathbb{R}^m$, e por isto é possível escolher $m - k$ colunas adicionais $\mathcal{A}_{B(k+1)}, \dots, \mathcal{A}_{B(m)}$ de $\mathcal{A}$, para obter um conjunto de m colunas LI, isto é, uma base consistindo exclusivamente das colunas de $\mathcal{A}$. Com esta base, todas as variáveis não básicas são levadas à zero, e segue que x^* é uma solução básica viável associada com esta nova base também. Neste ponto, as variáveis artificiais e as suas correspondentes colunas podem ser retiradas da Tabela.

Para aplicar a retirada da variável artificial da base, processa como a seguir: suponha que a l -ésima variável básica é uma variável artificial, que está na base. Examine a l -ésima linha da Tabela e ache algum j tal que a l -ésima entrada de $\mathcal{B}^{-1}\mathcal{A}_j$ seja não nulo. Tem-se que $\mathcal{A}_j$ é LI com as colunas $\mathcal{A}_{B(1)}, \dots, \mathcal{A}_{B(k)}$. De fato, note que $\mathcal{B}^{-1}\mathcal{A}_{B(i)} = e_i$, para $i = 1, \dots, k$ e visto que $k < l$, a l -ésima entrada desses vetores é zero. Segue que a l -ésima entrada de qualquer combinação linear dos vetores $\mathcal{B}^{-1}\mathcal{A}_{B(1)}, \dots, \mathcal{B}^{-1}\mathcal{A}_{B(k)}$ também é zero. Como a l -ésima entrada de $\mathcal{B}^{-1}\mathcal{A}_j$ é diferente de zero, ele não pode ser uma combinação linear dos vetores $\mathcal{B}^{-1}\mathcal{A}_{B(1)}, \dots, \mathcal{B}^{-1}\mathcal{A}_{B(k)}$. Com isto, $\mathcal{A}_j$ é levada para a base enquanto a l -ésima variável sai da base. Isto é realizado de maneira usual: aplique as operações elementares sobre linha para substituir $\mathcal{B}^{-1}\mathcal{A}_j$ pelo l -ésimo vetor unitário. A única diferença que aparece do mecanismo usual do Método Simplex é que o elemento pivô pode ser negativo. Sabendo que a l -ésima variável básica vale zero, adicionando um múltiplo da l -ésima linha a outras linha não altera o valor das outras variáveis básicas. Este processo mantém a mesma solução básica viável do problema auxiliar, mas reduz o número variáveis

básicas artificiais em um. Repetindo este processo, quantas vezes seja necessário, todas as variáveis artificiais sairão da base, como desejado.

Assuma agora que a l -ésima linha de $\mathcal{B}^{-1}\mathcal{A}$ é zero, neste caso o processo descrito acima falha. Note que a l -ésima linha de $\mathcal{B}^{-1}\mathcal{A}$ é igual a $\mathbf{g}^T\mathcal{A}$, onde $\mathbf{g}^T$ é a l -ésima linha de $\mathcal{B}^{-1}$. Portanto, $\mathbf{g}^T\mathcal{A} = \mathbf{0}$, para algum vetor não nulo $\mathbf{g}$ e, por isto, a matriz $\mathcal{A}$ possui linha LD. Visto que apenas problemas possíveis são tratados, deve-se ter também que $\mathbf{g}^T\mathbf{b} = 0$. Assim, a restrição $\mathbf{g}^T\mathcal{A}\mathbf{x} = \mathbf{g}^T\mathbf{b}$ é redundante e pode ser eliminada. Visto que esta restrição é a informação dada pela l -ésima linha da Tabela, esta linha pode ser eliminada e o processo continua.

O processo descrito acima é chamado de *Método Simplex Duas Fases*, um algoritmo é apresentado na Tab. 4.10.

Tabela 4.10: Uma Iteração do Método Simplex Duas Fases:

FASE 01:	
1 - Multiplique algumas das restrições por -1 , mudando o problema de modo que $\mathbf{b} \geq \mathbf{0}$.	
2 - Introduza variáveis artificiais $x_1^a, \dots, x_m^a$, se necessário, e aplique o Método Simplex para o problema auxiliar com custo $\sum_{i=1}^m x_i^a$.	
3 - Se o custo ótimo no problema auxiliar é positivo, então o problema original é inviável e, por isto, o algoritmo termina.	
4 - Se o custo ótimo no problema auxiliar é zero, uma solução viável para o problema original pode ser achada. Se nenhuma variável artificial pertence a base final, então as variáveis de folga e as suas colunas correspondentes são eliminadas. Com isto, uma base viável para o problema original está disponível.	
5 - Se a l -ésima variável básica é uma variável artificial, observe a l -ésima entrada da coluna $\mathcal{B}^{-1}\mathcal{A}_j$, $j = 1, \dots, n$. Se todas estas entradas são nulas, então a l -ésima linha representa uma restrição redundante e ela é eliminada. Caso contrário, se a l -ésima entrada, da j -ésima coluna é não nulo, faça uma mudança de base, tomando o elemento (l, j) como sendo o elemento pivô. Assim, a l -ésima variável sai da base e a x_j variável entra na base. Repita este argumento até que todas as variáveis artificiais saiam da base.	
FASE 02:	
1 - Considere a base final e a Tabela Simplex final obtida na Fase 01 como sendo a base inicial e a Tabela inicial para a Fase 02.	
2 - Calcule o custo reduzido de toda variável para esta base inicial, usando os coeficientes do custo do problema original.	
3 - Aplique o Método Simplex para o problema original.	

O algoritmo do Método Simplex Duas Fases é um método completo, no sentido que ele pode tratar de todos os resultados possíveis. Se a ciclagem for evitada, ou por ausência de degeneração, ou por alguma regra anti-ciclagem, ou mesmo sorte, uma das seguintes possibilidades acontece:

- a) Se o problema é inviável, isto é detectado no fim da Fase 01;
- b) Se o problema é viável, mas as linhas de $\mathcal{A}$ são LD, isto também é detectado e corrigido no final da Fase 01, eliminando as restrições de igualdade redundantes;
- c) Se o custo ótimo é $-\infty$, no início da Fase 02 isto é detectado.
- d) Caso contrário, no final da Fase 02 é encontrado uma solução ótima.

Exemplo 4.4.1. *Considere o PPL*

$$\begin{array}{ll} \text{minimizar} & x_1 + x_2 + x_3 \\ \text{sujeito a} & \left\{ \begin{array}{l} x_1 + 2x_2 + 3x_3 = 3 \\ -x_1 + 2x_2 + 6x_3 = 2 \\ 4x_2 + 9x_3 = 5 \\ 3x_3 + x_4 = 1 \\ x_1, x_2, x_3, x_4 \geq 0 \end{array} \right. \end{array} \quad (4.40)$$

Com o objetivo de achar uma solução viável, forme um problema auxiliar

$$\begin{array}{ll} \text{minimizar} & x_5^a + x_6^a + x_7^a \\ \text{sujeito a} & \left\{ \begin{array}{l} x_1 + 2x_2 + 3x_3 + x_5^a = 3 \\ -x_1 + 2x_2 + 6x_3 + x_6^a = 2 \\ 4x_2 + 9x_3 + x_7^a = 5 \\ 3x_3 + x_4 = 1 \\ x_1, \dots, x_4, x_5^a, \dots, x_7^a \geq 0 \end{array} \right. \end{array} \quad (4.41)$$

Uma solução básica viável para o problema auxiliar é obtida tomando $(x_5^a, x_6^a, x_7^a, x_4) = b = (3, 2, 5, 1)$. A matrix básica correspondente é a matriz identidade. Portanto, $c_B = (1, 1, 1, 1)$. Tem-se que calcular o custo reduzido de cada uma das variáveis original x_i , que é $-c_B A_i$, para formar a Tab. 4.11.

Tabela 4.11: Tabela Inicial Fase 01 do Exemplo 4.4.1

		x_1	x_2	x_3	x_4	x_5^a	x_6^a	x_7^a
	-11	0	-8	-21	0	0	0	0
$x_5^a =$	3	1	2	3	0	1	0	0
$x_6^a =$	2	-1	2*	6	0	0	1	0
$x_7^a =$	5	0	4	9	0	0	0	1
$x_4 =$	1	0	0	3	1	0	0	0

Tem-se que x_2 é a variável de menor índice com custo negativo, então $u = (2, 2, 4, 0)$. Assim, como $\frac{x_{B(1)}}{u_1} = 1, 5$, $\frac{x_{B(2)}}{u_2} = 1$ e $\frac{x_{B(3)}}{u_3} = 1, 2$, então o menor valor corresponde a variável $x_{B(2)} = x_6^a$, por isto ela sai da base. Assim, o elemento pivô é o elemento $(B(2), 2)$. Com isto, a Tab. 4.12 é a tabela no primeiro passo para o problema auxiliar.

Tabela 4.12: Tabela do Exemplo 4.4.1 - 1º Passo Fase 01

		x_1	x_2	x_3	x_4	x_5^a	x_6^a	x_7^a
	-3	-4	0	3	0	0	4	0
$x_5^a =$	1	2*	0	-3	0	1	-1	0
$x_2 =$	1	-1/2	1	3	0	0	1/2	0
$x_7^a =$	1	2	0	-3	0	0	-2	1
$x_4 =$	1	0	0	3	1	0	0	0

O único custo negativo corresponde a variável x_1 . Novamente, fazendo a análise $\frac{x_i}{u_i}$ tem-se que a variável que sai da base é a x_5^a . Logo, a Tab. 4.13 é a Tabela Simplex, no 2º passo, para o problema auxiliar.

Tem-se que neste passo o único custo negativo corresponde a variável x_3 . Usando a análise dos coeficientes $\frac{x_i}{u_i}$, para todos os índices possíveis, tem-se que a variável que sai da base é a x_4 . Logo, a Tab. 4.14 da a Tabela Simplex no terceiro passo para o problema auxiliar.

Tabela 4.13: Tabela do Exemplo 4.4.1 - 2º Passo Fase 01

		x_1	x_2	x_3	x_4	x_5^a	x_6^a	x_7^a
	-1	0	0	-3	0	2	2	0
$x_1 =$	1/2	1	0	-3/2	0	1/2	-1/2	0
$x_2 =$	5/4	0	1	9/4	0	1/4	1/4	0
$x_7^a =$	0	0	0	0	0	-1	-1	1
$x_4 =$	1	0	0	3*	1	0	0	0

Tabela 4.14: Tabela do Exemplo 4.4.1 - 3º Passo Fase 01

		x_1	x_2	x_3	x_4	x_5^a	x_6^a	x_7^a
	0	0	0	0	1	2	2	0
$x_1 =$	1	1	0	0	1/2	1/2	-1/2	0
$x_2 =$	1/2	0	1	0	-3/4	1/4	1/4	0
$x_7^a =$	0	0	0	0	0	-1	-1	1
$x_3 =$	1/3	0	0	1	1/3	0	0	0

Tem-se que o custo do problema original foi levado à zero, indicando que foi encontrado uma solução viável para o problema original. Contudo, uma variável artificial, a x_7^a , continua na base, com custo zero. Como deseja-se obter uma solução básica viável para o problema original, é preciso tirar x_7^a para fora da base. Tem-se que x_7^a é a terceira variável básica e que a terceira entrada da coluna $\mathcal{B}^{-1}\mathcal{A}_j$, $\forall j = 1, 2, 3, 4$, associada com as variáveis originais é zero. Isto indica que a matriz $\mathcal{A}$ tem linhas LD. Neste ponto, remova a terceira linha da Tab. 4.14, pois ela corresponde à restrição redundante, e também remova todas as variáveis artificiais. Assim, a Tabela Simplex inicial para a Fase 02 do PPL é dado pela Tab. 4.15.

Tabela 4.15: Tabela Simplex Inicial Fase 02 - Exemplo 4.4.1

		x_1	x_2	x_3	x_4
	*	*	*	*	*
$x_1 =$	1	1	0	0	1/2
$x_2 =$	1/2	0	1	0	-3/4
$x_3 =$	1/3	0	0	1	1/3

Para resolver o PPL, tome $x_{B(1)} = x_1$, $x_{B(2)} = x_2$ e $x_{B(3)} = x_3$ como sendo as variáveis básicas, logo a matriz básica correspondente é a matriz identidade. O vetor $c_B = (1, 1, 1)$ e, assim a Tabela Simplex Inicial fica dada pela Tab. 4.16.

Tabela 4.16: Tabela Simplex Inicial Fase 02 Completa - Exemplo 4.4.1

		x_1	x_2	x_3	x_4
	$-11/6$	0	0	0	$1/12$
$x_1 =$	1	1	0	0	$1/2$
$x_2 =$	$1/2$	0	1	0	$-3/4$
$x_3 =$	$1/3$	0	0	1	$1/3^*$

Assim, como o custo da variável não básica é positivo, tem-se que a solução $x = (1, \frac{1}{2}, \frac{1}{3}, 0)$ é uma solução ótima para o PPL.

4.5 Ciclagem em Problemas de Programação Linear

Ciclagem pode ocorrer quando existe uma solução básica degenerada num PPL. Se todas as soluções básicas são não degeneradas, tem-se que o algoritmo Simplex termina com um número finito de iterações. Mas, se a solução for degenerada, tem-se que o algoritmo Simplex pode construir uma seqüência de operações que ocasiona um ciclo de repetições indefinidas, chamada de **Ciclagem**.

Desde o desenvolvimento do método Simplex existe uma pergunta simples: qual a possibilidade de ocorrer ciclagem em PPL? Segundo um estudo apresentado por Zörnig (2006b), nos últimos 50 anos apenas 10 exemplos foram encontrados. A seguir é apresentado um exemplo envolvendo ciclagem que foi estudado por este autor.

Exemplo 4.5.1. Seja o PPL dado pela Eq. (4.42).

$$\begin{array}{ll}
 \text{minimizar} & -14x_1 + 25x_2 - \frac{7}{20}x_3 + 20x_4 \\
 \text{sujeito a} & \left\{ \begin{array}{l} x_1 - 2x_2 - \frac{1}{10}x_3 + 5x_4 \leq 0 \\ \frac{7}{10}x_1 - \frac{3}{10}x_2 - \frac{1}{100}x_3 + \frac{19}{50}x_4 \leq 0 \\ x_1, x_2, x_3, x_4 \geq 0 \end{array} \right.
 \end{array} \quad (4.42)$$

Construindo o primeira Tabela Simplex, acrescentando x_5 e x_6 como sendo as variáveis de folga, obtém-se a Tab. 4.17.

Tabela 4.17: Primeira Tabela do PPL - Exemplo 4.5.1

VB	x_1	x_2	x_3	x_4	x_5	x_6
	-14	25	$-7/20$	20	0	0
x_5	1*	-2	$-1/10$	5	1	0
x_6	$7/10$	$-3/10$	$-1/100$	$19/50$	0	1

Para fazer a escolha do pivô, observe que o menor custo, entre os coeficientes que são negativos, é o correspondente a variável x_1 . Assim, escolhendo o pivô como sendo a coordenada positiva que tem o menor índice, segue que 1 é o elemento procurado. Então, x_1 entra na base e x_5 sai da mesma. Logo, a segunda Tabela Simplex é dada na Tab. 4.18.

Tabela 4.18: Segunda Tabela do PPL - Exemplo 4.5.1

VB	x_1	x_2	x_3	x_4	x_5	x_6
	0	-3	$-7/4$	90	14	0
x_1	1	-2	$-1/10$	5	1	0
x_6	0	$11/10^*$	$3/50$	$-78/25$	$-7/10$	1

Para a segunda Tabela, o menor custo negativo é o relacionado com a variável x_2 . Logo, como a única coordenada positiva é $\frac{11}{10}$, então este é o pivô procurado. Desta forma x_2 entra na base e x_6 sai da mesma. A terceira Tabela Simplex é mostrada na Tab. 4.19.

Tabela 4.19: Terceira Tabela do PPL - Exemplo 4.5.1

VB	x_1	x_2	x_3	x_4	x_5	x_6
	0	0	$-349/220$	$4482/55$	$133/11$	$30/11$
x_1	1	0	$1/110^*$	$-37/55$	$-3/11$	$20/11$
x_2	0	1	$3/55$	$-156/55$	$-7/11$	$10/11$

Agora tem-se apenas um custo negativo, relacionado com a variável x_3 . Porém, as duas coordenadas são positivas, então toma-se a de menor índice como sendo o pivô. Assim, x_3 entra na base e x_1 sai. A Tab. 4.20 apresenta o quarto ciclo de operações.

Na Tab. 4.20 tem-se dois custo negativos, relacionados com a variável x_4 e x_5 , sendo que o menor deles depende da variável x_4 . Como a única coordenada positiva é $\frac{6}{5}$, segue que este é o pivô. Assim, x_4 entra na base e x_2 sai. sendo que a Quinta Tabela Simplex fica dada na Tab. 4.21.

Tabela 4.20: Quarta Tabela do PPL - Exemplo 4.5.1

VB	x_1	x_2	x_3	x_4	x_5	x_6
	$349/2$	0	0	$-359/10$	$-71/2$	320
x_3	100	0	1	-74	-30	200
x_2	-6	1	0	$6/5^*$	1	-10

Tabela 4.21: Quinta Tabela do PPL - Exemplo 4.5.1

VB	x_1	x_2	x_3	x_4	x_5	x_6
	-5	$359/12$	0	0	$-67/12$	$125/6$
x_3	-260	$185/3$	1	0	$95/3^*$	$-1250/3$
x_4	-5	$5/6$	0	1	$5/6$	$-25/3$

Na Tab. 4.21 tem-se dois custos negativo, relacionados às variáveis x_1 e x_5 , sendo este último o de menor valor. Assim, como as duas coordenadas são positivas, tem-se que sai da base a de menor índice, ou seja, x_5 entra na base e x_3 sai. A Tab. 4.22 representa o sexto ciclo de operações.

Tabela 4.22: Sexta Tabela do PPL - Exemplo 4.5.1

VB	x_1	x_2	x_3	x_4	x_5	x_6
	$-966/19$	$775/19$	$67/380$	0	0	$-1000/19$
x_5	$-156/19$	$37/19$	$3/95$	0	1	$-250/19$
x_4	$35/19$	$-15/19$	$-1/38$	1	0	$50/19^*$

Na Tab. 4.22 tem-se dois custos negativo, relacionados com as variáveis x_1 e x_6 , sendo este último é o de menor valor. Como a única coordenada positiva é a $\frac{50}{19}$, tem-se que x_6 entra na base e x_4 sai da mesma. Assim, a Sétima Tabela Simplex é dada na Tab. 4.23. Observe que a Tab. 4.23 é exatamente igual à Tab. 4.17, ou seja, após o sétimo ciclo de operações voltou-se ao problema original, caracterizando a presença de ciclagem.

Algumas regras anti-ciclagem existentes são: regra Bland (*Bland's rule*), que escolhe as variáveis elegíveis com sendo a que tenha o menor índice; a regra Aleatória (*Random rule*), que escolhe as variáveis elegíveis uniformemente de maneira aleatória e a regra Lexicográfica (*Lexicographic Rule*) que faz uma perturbação nas restrições tornando o PPL não degenerado. É importante ressaltar que a Regra Aleatória evita ciclagem com grande probabilidade mas não exclui a possibilidade de ocorrer ciclagem.

Tabela 4.23: Sétima Tabela do PPL - Exemplo 4.5.1

VB	x_1	x_2	x_3	x_4	x_5	x_6
	-14	25	$-7/20$	20	0	0
x_5	1	-2	$-1/10$	5	1	0
x_6	$7/10$	$-3/10$	$-1/100$	$19/50$	0	1

A idéia da regra Lexicográfica é construir soluções não degeneradas visto que assim não aparece ciclagem. Então, para isto, considere o PPL perturbado dado pela Eq. (4.43).

$$\begin{aligned} & \text{maximizar} && c^T x \\ & \text{sujeito a} && Ax = b + \epsilon \end{aligned} \quad (4.43)$$

onde $\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$ e $\epsilon_1 \gg \epsilon_2 \gg \dots \gg \epsilon_n$. No problema perturbado tem-se que $x_B = A_B^{-1}(b + \epsilon)$ é sempre não nulo. A j -ésima componente de x_B é uma combinação linear, não nula, das componentes de $b + \epsilon$ então, ela contém pelo menos um dos termos ϵ_i .

Um estudo completo envolvendo a construção de exemplos envolvendo ciclagem é encontrado em Zörnig (2006a) e Zörnig (2006b).

Capítulo 5

O Método Simplex Dual

Este capítulo inicia-se com um PPL, chamado primal, à partir do qual é construído outro PPL chamado dual. O Método Simplex Dual trata a relação entre estes dois problemas e vai desenvolver estruturas fundamentais para a programação linear. Este método é um instrumento teórico poderoso que traz várias aplicações, apresentando uma nova compreensão geométrica e um novo algoritmo para a programação linear.

5.1 Motivação

A definição de **Método Simplex Dual** é motivado por uma extensão do Método dos Multiplicadores de Lagrange (MML), que pode ser visto em Guidorizzi (2006), que muitas vezes é usado para minimizar uma função não linear sujeita a restrições de igualdade. O Exemplo 5.1.1 apresenta uma ilustração desta idéia.

Exemplo 5.1.1. *Resolver o problema, usando o MML.*

$$\begin{array}{ll} \text{minimizar} & x^2 + y^2 \\ \text{sujeito a} & x + y = 1 \end{array}, \quad (5.1)$$

considere a função Lagrangiana $L(x, y, \lambda)$, definida por

$$L(x, y, \lambda) = x^2 + y^2 + \lambda(1 - x - y) \quad (5.2)$$

Para minimizar o Lagrangiano, obtém-se o seu gradiente encontrando as derivadas parciais em relação a x e y . Então, tem-se que $\frac{\partial L}{\partial x} = 2x - \lambda$ e $\frac{\partial L}{\partial y} = 2y - \lambda$. Assim, fazendo $\nabla L = 0$, tem-se:

$$x = y = \frac{\lambda}{2}, \quad (5.3)$$

que depende de λ . Utilizando a derivada parcial em relação a λ , tem-se que:

$$\frac{\partial L}{\partial \lambda} = 1 - x - y, \quad (5.4)$$

logo o valor de λ é 1, o que fornece $(x, y) = (\frac{1}{2}, \frac{1}{2})$ como solução ótima para o PPL original.

A idéia principal apresentada no Exemplo 5.1.1 é a seguinte: ao invés de impor a difícil restrição $x + y = 1$, ignore inicialmente este fato e associe o MML, ou um peso λ , com a soma $1 - x - y$ para permitir esta violação. Isto leva a um PPL irrestrito, dado por $L(x, y, \lambda) = x^2 + y^2 + \lambda(1 - x - y)$. Quando o peso é convenientemente escolhido, no Exemplo 5.1.1 tomou-se $\lambda = 1$, a solução ótima do problema irrestrito é também solução ótima do problema original. Na verdade, sob tal valor específico de λ , a presença ou ausência da restrição não altera o custo ótimo.

Em PPL a situação é similar: associa-se um peso a cada uma das restrições e inicia-se a busca dos pesos sob os quais a presença ou ausência das restrições não fazem efeito no custo ótimo. Como o custo ótimo pode ser encontrado resolvendo um novo PPL, este novo PPL é chamado de Dual do problema original. Agora, é construído uma motivação para o Dual.

Considere o PPL na forma padrão

$$\begin{aligned}
&\text{minimizar} && c^T x \\
&\text{sujeito a} && Ax = b, \\
&&& x \geq \mathbf{0}
\end{aligned} \tag{5.5}$$

este PPL é chamado de *problema primal*. Assuma que a solução ótima x^* exista. Introduza um problema “relaxado”, no qual a restrição $Ax = b$ é retirada e a penalidade $p^T(b - Ax)$, onde p é um vetor peso de mesma dimensão do vetor b , é utilizada para a construção do Lagrangiano. Assim, o novo problema é escrito como:

$$\begin{aligned}
&\text{minimizar} && c^T x + p^T(b - Ax) \\
&\text{sujeito a} && x \geq \mathbf{0}
\end{aligned} . \tag{5.6}$$

Seja $g(p)$ o custo ótimo para o problema relaxado, como sendo uma função do vetor peso p . Como o problema relaxado permite um número maior de soluções do que o problema primal, já que ele tem um número menor de restrições, tem-se que $g(p)$ não pode ser maior do que o custo primal ótimo $c^T x^*$. De fato:

$$g(P) = \min_{x \geq \mathbf{0}} [c^T x + p^T(b - Ax)] \leq c^T x^* + p^T(b - Ax^*) = c^T x^*, \tag{5.7}$$

sendo que a última desigualdade segue do fato que x^* é uma solução viável para o problema primal, satisfazendo $Ax^* = b$. Logo, cada p transforma $g(p)$ numa cota inferior para o custo primal ótimo $c^T x^*$. O problema

$$\text{minimizar } g(p) \text{ irrestrito}, \tag{5.8}$$

pode ser interpretado como sendo uma busca pela maior cota inferior do PPL dado pela Eq. (5.8). Este processo é conhecido como sendo o *problema dual*. O principal resultado envolvendo o Método Simplex Dual garante que o custo ótimo no problema dual é igual ao custo ótimo $c^T x^*$ do

problema primal. Em outras palavras, quando os pesos são escolhidos de maneira que uma solução ótima para o PPL dual é encontrada, a violação da restrição $\mathcal{A}x = b$ não traz prejuízo para o PPL primal.

Usando a definição do $g(p)$, tem-se

$$g(P) = \min_{x \geq \mathbf{0}} [c^T x + p^T (b - \mathcal{A}x)] = p^T b + \min_{x \geq \mathbf{0}} (c^T - p^T \mathcal{A}) x. \quad (5.9)$$

Note que,

$$\min_{x \geq \mathbf{0}} (c^T - p^T \mathcal{A}) x = \begin{cases} 0 & \text{se } c^T - p^T \mathcal{A} \geq \mathbf{0} \\ -\infty & \text{caso contrário.} \end{cases} \quad (5.10)$$

Para maximizar $g(p)$ é preciso considerar somente os valores de p para o qual $g(p)$ é finito. Assim, o dual equivale ao PPL

$$\begin{aligned} & \text{maximizar} && p^T b \\ & \text{sujeito a} && p^T \mathcal{A} \leq c^T \end{aligned} \quad (5.11)$$

O Exemplo 5.1.1 iniciou-se com a restrição de igualdade $\mathcal{A}x = b$ e o Dual foi criado sem restrição de sinal para o vetor peso p . Se o problema primal tem restrição de desigualdade da forma $\mathcal{A}x \geq b$, ela pode ser substituída por $\mathcal{A}x - s = b$, com $s \geq \mathbf{0}$. A restrição de desigualdade pode ser reescrita na forma

$$\left[\mathcal{A} \mid -\mathcal{I} \right] \cdot \begin{bmatrix} x \\ s \end{bmatrix} = b, \quad (5.12)$$

que leva à restrição dual

$$p^T \left[\mathcal{A} \mid -\mathcal{I} \right] \leq \left[c^T \mid \mathbf{0}^T \right], \quad (5.13)$$

ou seja,

$$p^T \mathcal{A} \leq c^T, \quad p \geq \mathbf{0}. \quad (5.14)$$

Além disso, se o vetor x é livre das restrições de sinal, usa-se o fato que

$$\min_x (c^T - p^T \mathcal{A}) x = \begin{cases} 0 & \text{se } c^T - p^T \mathcal{A} \geq \mathbf{0} \\ -\infty & \text{caso contrário.} \end{cases} \quad (5.15)$$

e a restrição $p^T \mathcal{A} = c^T$ aparece no dual. Estas considerações são as que motivarão a forma geral do problema dual que é apresentada na próxima seção.

Em resumo, o problema Dual de um problema de minimização primal consiste em: partindo de um vetor peso, que são as variáveis dual p , obtém-se uma cota inferior do custo primal ótimo. O problema dual é um problema de maximização que olha para a maior cota inferior. Para algum vetor p a cota inferior correspondente é $-\infty$, como esta solução não traz informações relevantes, é abandonada. Assim, é necessário apenas maximizar o dual sobre os vetores p que levam a cotas inferiores não triviais, e por isto fazem aparecer restrições no dual.

5.2 O Problema Dual do Problemas Linear

Seja $\mathcal{A}$ uma matriz com linhas a_i e colunas $\mathcal{A}_j$. Dado um problema primal com a estrutura mostrada na 1ª coluna da Tab. 5.1, então o seu dual é definido como sendo o problema de maximização mostrado na 2ª coluna da Tab. 5.1.

Note que, para cada restrição do Primal, é introduzida uma restrição lateral no problema Dual. Para cada restrição lateral no problema Primal é introduzida uma restrição no primal. Se

Tabela 5.1: Representação do Problema Primal e do seu Dual, de um PPL

Primal		Dual	
minimizar	$c^T x$	maximizar	$p^T b$
sujeito à	$a_i x \geq b_i, \quad i \in M_1$	sujeito à	$p_i \geq 0, \quad i \in M_1$
	$a_i x \leq b_i, \quad i \in M_2$		$p_i \leq 0, \quad i \in M_2$
	$a_i x = b_i, \quad i \in M_3$		p_i livre, $i \in M_3$
	$x_j \geq 0, \quad j \in N_1$		$p^T \mathcal{A}_j \leq c_j, \quad j \in N_1$
	$x_j \leq 0, \quad j \in N_2$		$p^T \mathcal{A}_j \geq c_j, \quad j \in N_2$
	x_j livre, $j \in N_3$		$p^T \mathcal{A}_j = c_j, \quad j \in N_3$

as restrições no Primal são de igualdade ou de desigualdade, as variáveis correspondentes no Dual são livres ou possuem restrições de sinal. Do mesmo modo, se as variáveis do problema Primal são livres ou possuem restrições de sinal, cria-se restrições de igualdade ou desigualdade no problema Dual. Assim, na Tab. 5.2 é apresentado um resumo destas observações.

Tabela 5.2: Resumo apresentando a relação entre o Problema Primal e o seu Dual

Primal	minimizar	maximizar	Dual
restrições	$\geq b_i$	≥ 0	variáveis
	$\leq b_i$	≤ 0	
	$= b_i$	livre	
variáveis	≥ 0	$\leq c_j$	restrições
	≤ 0	$\geq c_j$	
	livre	$= c_j$	

Um problema e seu Dual podem ser representados de maneira mais compacta, usando a notação de matriz, quando o Dual for representado de uma forma particular. Assim, tem-se nas Eq. (5.16) e Eq. (5.17) a representação de dois problemas do tipo Primal, à esquerda, com o seus respectivos Duais, à direita.

$$\begin{array}{ll}
 \text{Primal} & \text{Dual} \\
 \text{minimizar} & c^T x \\
 \text{sujeito a} & Ax = b \\
 & x \geq 0
 \end{array}
 \quad
 \begin{array}{ll}
 \text{maximizar} & p^T b \\
 \text{sujeito a} & p^T A \leq c^T
 \end{array}
 , \quad (5.16)$$

$$\begin{array}{ll}
\text{minimizar} & c^T x \\
\text{sujeito a} & Ax \geq b
\end{array}
\qquad
\begin{array}{ll}
\text{maximizar} & p^T b \\
\text{sujeito a} & p^T A = c^T \\
& p \geq 0
\end{array}
\quad (5.17)$$

O próximo exemplo tenta criar uma idéia de que o Dual de um Dual é o próprio primal, fato este que é confirmado pelo Teorema 5.2.1.

Exemplo 5.2.1. *Considere o Primal, mostrado à esquerda, e seu respectivo Dual, mostrado à direita, na Eq. (5.18).*

$$\begin{array}{ll}
\text{minimizar} & x_1 + 2x_2 + 3x_3 \\
\text{sujeito a} & -x_1 + 3x_2 = 5 \\
& 2x_1 - x_2 + 3x_3 \geq 6 \\
& x_3 \leq 4 \\
& x_1 \geq 0 \\
& x_2 \leq 0 \\
& x_3 \text{ livre}
\end{array}
\qquad
\begin{array}{ll}
\text{maximizar} & 5p_1 + 6p_2 + 4p_3 \\
\text{sujeito a} & p_1 \text{ livre} \\
& p_2 \geq 0 \\
& p_3 \leq 0 \\
& -p_1 + 2p_2 \leq 1 \\
& 3p_1 - p_2 \geq 2 \\
& 3p_2 + p_3 = 3
\end{array}
\quad (5.18)$$

Escreva o Dual da Eq. (5.18) como um problema de minimização equivalente, transformando as variáveis p_1 , p_2 e p_3 em x_1 , x_2 e x_3 , respectivamente, e multiplicando as três últimas restrições por -1 . O problema resultante é mostrado à esquerda, enquanto que à direita é apresentado o seu respectivo Dual.

$$\begin{array}{ll}
\text{minimizar} & 5x_1 - 6x_2 - 4x_3 \\
\text{sujeito a} & x_1 \text{ livre} \\
& x_2 \geq 0 \\
& x_3 \leq 0 \\
& x_1 - 2x_2 \geq -1 \\
& -3x_1 + x_2 \leq -2 \\
& -3x_2 - x_3 = -3
\end{array}
\qquad
\begin{array}{ll}
\text{maximizar} & -p_1 - 2p_2 - 3p_3 \\
\text{sujeito a} & p_1 - 3p_2 = -5 \\
& -2p_1 + p_2 - 3p_3 \leq -6 \\
& -p_3 \geq -4 \\
& p_1 \geq 0 \\
& p_2 \leq 0 \\
& p_3 \text{ livre}
\end{array}
\tag{5.19}$$

Observa-se que o problema Dual da Eq. (5.19) é equivalente ao problema Primal da Eq. (5.18), pois as três primeiras restrições no último Dual são as mesmas três restrições do Primal original, apenas multiplicadas por -1 . Para a transformação do problema de maximização em problema de minimização, basta multiplicar a função objetivo por -1 , transformando o Dual no problema original.

O problema Primal considerado no Exemplo 5.2.1 apresentava todos os elementos de um PPL geral. Por isto, a suposição de que o Dual do Dual é o próprio Primal, parece ser verdadeira. Assim, a apresentação do Teorema 5.2.1 torna-se aceitável.

Teorema 5.2.1. *Se o problema Dual for transformado em um problema de minimização equivalente, então o seu Dual é um problema equivalente ao problema original.*

Demonstração 5.2.1. *A demonstração deste resultado é muito simples, basta utilizar as relações da Tab. 5.2 e, quando necessário, multiplique as equações por -1 , seguindo a mesma idéia do Exemplo 5.2.1.*

□

Qualquer PPL pode ser transformado em qualquer uma das formas equivalentes de um primal. Basta, por exemplo, acrescentar variáveis de folga, ou usar a diferença entre duas variáveis

não negativas para substituir uma variável livre. Cada uma das formas equivalentes leva a alguma forma de problema Dual, mesmo assim, o Exemplo 5.2.2 sugere que o Dual de problemas equivalentes são equivalentes.

Exemplo 5.2.2. Considere o Primal, mostrado à esquerda, e o seu respectivo Dual, mostrado à direita, na Eq. (5.20).

$$\begin{array}{ll}
 \text{minimizar} & \mathbf{c}^T \mathbf{x} \\
 \text{sujeito a} & \mathbf{Ax} \geq \mathbf{b} \\
 & \mathbf{x} \text{ livre}
 \end{array}
 \qquad
 \begin{array}{ll}
 \text{maximizar} & \mathbf{p}^T \mathbf{b} \\
 \text{sujeito a} & \mathbf{p} \geq \mathbf{0} \\
 & \mathbf{p}^T \mathbf{A} = \mathbf{c}^T
 \end{array}
 \quad (5.20)$$

Utilizando variáveis de folga, o problema Primal e o seu Dual passam a ser dados por:

$$\begin{array}{ll}
 \text{minimizar} & \mathbf{c}^T \mathbf{x} + \mathbf{0}^T \mathbf{s} \\
 \text{sujeito a} & \mathbf{Ax} - \mathbf{s} = \mathbf{b} \\
 & \mathbf{x} \text{ livre} \\
 & \mathbf{s} \geq \mathbf{0}
 \end{array}
 \qquad
 \begin{array}{ll}
 \text{maximizar} & \mathbf{p}^T \mathbf{b} \\
 \text{sujeito a} & \mathbf{p} \text{ livre} \\
 & \mathbf{p}^T \mathbf{A} = \mathbf{c}^T \\
 & -\mathbf{p} \leq \mathbf{0}
 \end{array}
 \quad (5.21)$$

Se, ao invés de variáveis de folga, forem trabalhadas as restrições de sinal, os problemas podem ser escritos como:

$$\begin{array}{ll}
 \text{minimizar} & \mathbf{c}^T \mathbf{x}^+ - \mathbf{c}^T \mathbf{x}^- \\
 \text{sujeito a} & \mathbf{Ax}^+ - \mathbf{Ax}^- \geq \mathbf{b} \\
 & \mathbf{x}^+ \geq \mathbf{0} \\
 & \mathbf{x}^- \geq \mathbf{0}
 \end{array}
 \qquad
 \begin{array}{ll}
 \text{maximizar} & \mathbf{p}^T \mathbf{b} \\
 \text{sujeito a} & \mathbf{p} \geq \mathbf{0} \\
 & \mathbf{p}^T \mathbf{A} \leq \mathbf{c}^T \\
 & -\mathbf{p}^T \mathbf{A} \leq -\mathbf{c}^T
 \end{array}
 \quad (5.22)$$

Aqui foram apresentadas três formas equivalentes para o problema primal. Ainda, é possível observar que a restrição $\mathbf{p} \geq \mathbf{0}$ é equivalente à restrição $-\mathbf{p} \leq \mathbf{0}$, a restrição

$p^T \mathcal{A} = c^T$ é equivalente a outras duas restrições $p^T \mathcal{A} \geq c^T$ e $p^T \mathcal{A} \leq c^T$. Por isto o Dual das três variações do problema Primal são equivalentes.

O próximo exemplo tem o mesmo espírito do Exemplo 5.2.2, estudando os efeitos resultantes da mudança de restrições de igualdade no problema primal na forma padrão.

Exemplo 5.2.3. Considere um PPL na forma padrão e o seu Dual.

$$\begin{array}{ll} \text{minimizar} & c^T x \\ \text{sujeito a} & Ax = b \\ & x \geq 0 \end{array} \qquad \begin{array}{ll} \text{maximizar} & p^T b \\ \text{sujeito a} & p^T \mathcal{A} \leq c^T. \end{array} \quad (5.23)$$

Sejam $a_1, a_2, \dots, a_m$ as linhas de $\mathcal{A}$ e suponha que $a_m = \sum_{i=1}^{m-1} \lambda_i a_i$, para alguns escalares $\lambda_1, \lambda_2, \dots$ e λ_{m-1} . Ou seja, a restrição a_m é redundante e pode ser eliminada. Sobre uma solução arbitrária viável x , tem-se:

$$b_m = a_m^T x = \sum_{i=1}^{m-1} \lambda_i a_i x = \sum_{i=1}^{m-1} \lambda_i b_i. \quad (5.24)$$

Note que, as restrições do Dual são da forma $\sum_{i=1}^m p_i a_i \leq c^T$, e podem ser reescritas como sendo

$$\sum_{i=1}^{m-1} (p_i + \lambda_i p_m) a_i x \leq c^T. \quad (5.25)$$

Além disso, usando a Eq. (5.24) o custo do Dual $\sum_{i=1}^m p_i b_i$ é igual a

$$\sum_{i=1}^{m-1} (p_i + \lambda_i p_m) b_i. \quad (5.26)$$

Tomando $q_i = p_i + \lambda_i p_m$, o problema Dual passa a ser equivalente à

$$\begin{aligned} &\text{maximizar} && \sum_{i=1}^{m-1} q_i b_i \\ &\text{sujeito a} && \sum_{i=1}^{m-1} q_i a_i \leq c^T. \end{aligned} \quad (5.27)$$

Observe que este Dual é exatamente igual ao que seria obtido se fosse eliminada a restrição redundante do primal e formado o seu Dual.

As idéias obtidas nos exemplos anteriores são sintetizadas e generalizadas pelo próximo resultado.

Teorema 5.2.2. *Suponha que tenha sido transformado um PPL Π_1 num outro PPL Π_2 , por uma seqüência de transformações como a seguir:*

- a) troca de uma variável livre pela diferença de duas variáveis não negativas;*
- b) troca de uma restrição de desigualdade por uma restrição de igualdade, envolvendo uma variável de folga não negativa;*
- c) se alguma linha da matriz $\mathcal{A}$ num problema na forma padrão é uma combinação linear das outras linhas, elimine a restrição correspondente.*

Então, o Dual de Π_1 e Π_2 são equivalentes, isto é, ou ambos são inviáveis ou ambos tem o mesmo custo ótimo.

Demonstração 5.2.2. *A demonstração deste teorema também é simples, basta seguir os passos utilizados nos Exemplos 5.2.2 e 5.2.3.*

□

5.3 O Teorema Fundamental para o Problema Dual

No início deste capítulo foi visto que, para PPL na forma padrão, o custo $g(p)$ de qualquer solução Dual é uma cota inferior para o custo ótimo. O resultado a seguir garante que esta propriedade é geral.

Teorema 5.3.1. (Teorema Fraco da Dualidade): *Se x é uma solução viável para o problema Primal e p é uma solução viável para o problema Dual, então:*

$$p^T b \leq c^T x. \quad (5.28)$$

Demonstração 5.3.1. *Para qualquer vetor x e p , defina*

$$u_i = p_i(a_i x - b_i), \quad (5.29)$$

$$v_j = (c_j - p^T A_j)x_j. \quad (5.30)$$

Suponha que x e p são soluções do problema Primal e Dual, respectivamente. Pela definição do Dual, tem-se que o sinal de p_i é o mesmo sinal de $a_i x - b_i$, e o sinal de $c_j - p^T A_j$ é o mesmo sinal de x_j . Assim, para que o problema Primal e o problema Dual sejam viáveis é necessário que

$$u_i \geq 0, \forall i \quad e \quad v_j \geq 0, \forall j. \quad (5.31)$$

Note que

$$\sum_i u_i = p^T A x - p^T b, \quad (5.32)$$

e

$$\sum_j v_j = c^T x - p^T Ax. \quad (5.33)$$

Somando as Eq. (5.32) e Eq. (5.33), e usando o fato de que u_i e v_j são não negativos, tem-se que

$$0 \leq \sum_i u_i + \sum_j v_j = c^T x - p^T b. \quad (5.34)$$

□

O Teorema Fraco da Dualidade não é um resultado muito profundo, mas ele fornece algumas informações úteis, envolvendo a relação entre o Primal e o Dual. Ainda, baseado no Teorema 5.3.1, tem-se o seguinte Corolário.

Corolário 5.3.1. *a) Se o custo ótimo do problema Primal é $-\infty$, então o problema Dual é inviável.*

b) Se o custo ótimo do Dual é $+\infty$, então o problema Primal é inviável.

Demonstração 5.3.1. *a) Suponha que o custo ótimo do problema Primal seja $-\infty$ e que o problema Dual tenha uma solução viável p . Pelo Teorema Fraco da Dualidade tem-se que p satisfaz a relação $p^T b \leq c^T x$, para cada solução viável do Primal x , o que gera um absurdo, pois o custo do Primal não é limitado inferiormente. Logo, tem-se (a).*

b) Agora, suponha que o custo ótimo do Problema Dual é $+\infty$ e que o Problema Primal tenha uma solução viável. Então, pelo Teorema Fraco da Dualidade, $p^T b \leq c^T x$, para cada solução viável do Dual, novamente chega-se num absurdo, visto que o problema Dual não é limitado superiormente. Logo, tem-se (b).

□

Outro Corolário importante do Teorema Fraco da Dualidade é apresentado abaixo.

Corolário 5.3.2. *Sejam x e p soluções básicas viáveis para o problema Primal e Dual, respectivamente, e suponha que $p^T b = c^T x$. Então, x e p são soluções ótimas para o problema Primal e do Dual, respectivamente.*

Demonstração 5.3.2. *Sejam x e p como nas hipóteses do Corolário 5.3.2. Para cada solução viável primal y , e pelo Teorema Fraco da Dualidade, tem-se que $c^T x = p^T b \leq c^T y$, o que mostra que x é uma solução ótima para o problema primal. Agora tome s uma solução viável para o Dual. Pelo Teorema Fraco da Dualidade tem-se que $p^T s \leq c^T x = p^T b$, o que prova que p é uma solução ótima para o problema Dual.*

□

O resultado central na Teoria da Dualidade é dado abaixo, resultado este que garante a igualdade da solução do problema Primal e do problema Dual.

Teorema 5.3.2. (Teorema Forte da Dualidade): *Se um PPL tem uma solução ótima, então o custo ótimo do PPL e do seu respectivo dual são iguais.*

Demonstração 5.3.2. *Considere o PPL na forma padrão*

$$\begin{aligned} &\text{minimizar} && c^T x \\ &\text{sujeito a} && Ax = b \quad . \\ &&& x \geq 0 \end{aligned} \tag{5.35}$$

Assuma temporariamente que as linhas de A são LI e que existe uma solução ótima. Aplique o Método Simplex para este problema. Supondo que ciclagem está sendo evitada, usando uma regra de pivoteamento, o método Simplex encontra uma solução ótima viável x e uma matriz básica ótima B . Seja $x_B = B^{-1}b$ o vetor correspondente das variáveis básicas. Quando o Método Simplex termina, o custo reduzido deve ser não negativo e obtendo-se

$$c^T - c_B^T \mathcal{B}^{-1} \mathcal{A} \geq 0^T, \quad (5.36)$$

onde c_B^T é o vetor com os custos das variáveis básicas. Defina um vetor p satisfazendo $p^T = c_B^T \mathcal{B}^{-1}$. Então, tem-se $p^T \mathcal{A} \leq c^T$, mostrando que p é uma solução viável para o problema Dual

$$\begin{array}{ll} \text{maximizar} & p^T b \\ \text{sujeito a} & p^T \mathcal{A} \leq c^T \end{array} \quad (5.37)$$

Mais ainda,

$$p^T b = c_B^T \mathcal{B}^{-1} b = c_B^T x_B = c^T x. \quad (5.38)$$

Segue que p é uma solução ótima para o problema Dual, conforme o Corolário 5.3.2, e o custo ótimo do Dual é igual ao custo ótimo do Primal.

Suponha agora que o PPL Π_1 tenha uma representação geral e que ele tenha uma solução ótima. Faça uma primeira transformação do PPL num PPL na forma padrão Π_2 , que tem o mesmo custo ótimo e no qual as linhas da matriz $\mathcal{A}$ são LI. Seja D_1 e D_2 os duais dos problemas Π_1 e Π_2 , respectivamente. Pelo Teorema 5.2.2 os problemas D_1 e D_2 tem o mesmo custo ótimo. Como já foi mostrado que um PPL na forma padrão e o seu Dual tem o mesmo custo, segue que Π_1 e D_1 também tem o mesmo custo.

□

Segundo a demonstração anterior uma solução ótima do problema Dual é obtido como sendo um sub-produto do Método Simplex aplicado à um problema Primal na forma padrão. Assumindo o fato de que o Método Simplex tem um número máximo de iterações, quando existe uma regra de pivoteamento que evite ciclagem.

O próximo exemplo apresenta uma aplicação da obtenção de uma solução básica viável do problema Primal, a partir do Dual.

Exemplo 5.3.1. *Considere uma bola sólida restrita a um poliedro definido pelas restrições de desigualdades da forma $a_i x \geq b_i$. Se ela está sujeita apenas à força da gravidade, esta bola acha equilíbrio no canto mais baixo x^* do poliedro, conforme a Fig. 5.1. Este canto é uma solução ótima do problema*

$$\begin{aligned} &\text{minimizar} && c^T x \\ &\text{sujeito a} && a_i x \geq b_i, \quad \forall i. \end{aligned} \quad (5.39)$$

onde c é um vetor vertical partindo do “bico”. Em equilíbrio, a gravidade é contrabalanceada pelas forças exercidas na bola pela “parede” do poliedro. As forças são perpendiculares às paredes, isto é, elas são alinhadas com os vetores a_i . Assim, conclui-se que $c = \sum_i p_i a_i$, para alguns coeficientes não negativos p_i . Em particular, o vetor p é uma solução viável para o problema Dual

$$\begin{aligned} &\text{maximizar} && p^T b \\ &\text{sujeito a} && p^T A = c^T \\ &&& p \geq 0. \end{aligned} \quad (5.40)$$

Dado que as forças só podem ser exercidas pelas paredes que tocam na bola, deve-se ter $p_i = 0$, sempre que $a_i x^* > b_i$. Conseqüentemente, $p(b_i - a_i x^*) = 0$, para todo i . Tem-se, portanto, $p^T b = \sum_i p_i b_i = \sum_i p_i a_i x^* = c^T x^*$. Segue do Corolário 5.3.2 que p é uma solução ótima para o problema Dual e o custo ótimo Dual é igual ao custo ótimo Primal.

Relembre que num PPL uma das três possibilidades devem ocorrer:

a) Existe uma solução ótima.

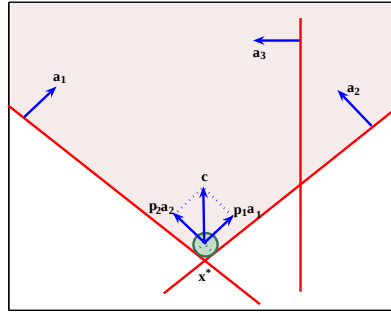


Figura 5.1: Representação de uma aplicação para o Teorema Dual no Exemplo 5.3.1.

- b) O PPL é “ilimitado”, isto é, o custo ótimo é $-\infty$, para o problema de minimização, ou $+\infty$, para problemas de maximização.
- c) O problema é inviável.

5.4 Uma Interpretação Econômica para o Dual

A interpretação econômica dada neste trabalho é apresentada por Bregalda (1988), sendo que o autor ressalta que esta é a interpretação de um problema particular, fazendo com que a análise perca em generalidade.

Seja o PPL dado pela Eq. (5.41)

$$\begin{aligned} & \text{minimizar} && F(x) = c^T x \\ & \text{sujeito a} && Ax \leq b, \\ & && x_j \geq 0, \forall j \end{aligned} \quad (5.41)$$

e seja o seu dual dado pela Eq. (5.42).

$$\begin{aligned} & \text{maximizar} && Q(p) = p^T b \\ & \text{sujeito a} && pA \leq c^T, \\ & && p_i \geq 0, \forall i \end{aligned} \quad (5.42)$$

Agora, suponha que x^* e p^* sejam soluções básicas ótimas do PPL e o seu dual, dados pela Eq. (5.41) e Eq. (5.42), respectivamente, e seja B^* a matriz básica associada com as soluções. Assim, tem-se que

$$F(x^*) = Q(p^*) = p^* b. \quad (5.43)$$

Agora, supondo que b é variável e derivando a Eq. (5.43) em relação a b chega-se na Eq. (5.44).

$$\frac{\partial F(x^*)}{\partial b} = p^*. \quad (5.44)$$

Então, pode-se interpretar p^* como sendo a taxa de variação do valor ótimo da função objetivo $F(x^*)$ por unidade de variação de b . Como $p_i^* \geq 0$, $\forall i$, segue que $F(x^*)$ cresce à medida que b_i cresce.

Para tornar esta interpretação mais precisa, imagine que o problema dado pela Eq. (5.41) seja a alocação de recursos, onde se tenha m recursos disponíveis, nas quantidades $b_1, b_2, \dots, b_m$, com os quais se deseja fabricar n produtos, nas quantidades $x_1, x_2, \dots, x_n$ a serem determinadas. Cada unidade do produto j consome a_{ij} unidades do recurso i , trazendo um retorno de c_j unidades monetárias. Assim, deseja-se determinar a quantidade a ser fabricada de cada produto, de modo a se obter um lucro máximo.

Agora, suponha que a quantidade disponível do recurso k seja aumentada em uma unidade, isto é, suponha que existam $b_k + 1$ unidades. Evidentemente tem-se, agora, que obter uma nova solução ótima $\bar{x}^*$. Suponha que a base associada a $\bar{x}^*$ seja a mesma base B^* associada com a solução básica x^* . Isto significa que a nova solução básica $\bar{p}^*$ continua sendo p^* , visto que $\bar{x}^* = x^* = (c^{B^*})^T (B^*)^{-1}$. Então,

$$F(\bar{x}^*) = \sum_{\substack{i=1 \\ i \neq k}}^m b_i u_i^* + (b_k + 1)u_k^* = \sum_{i=1}^m b_i u_i^* + u_k^* = F(x^*) + u_k^*. \quad (5.45)$$

Ou seja, $F(\bar{x}^*) - F(x^*) = u_k^*$, em outras palavras, u_k^* é o incremento no lucro, trazido pelo aumento de uma unidade da matéria disponível k . Pode-se também tomar a seguinte interpretação: u_k^* é o preço que se deseja pagar para aumentar de uma unidade a quantidade disponível da matéria prima k . Assim, u_k^* é o valor do incremento de uma unidade de matéria prima k . Por isto, u_k^* é chamado de **valor incremental**, ou **valor implícito**.

Com isto é possível interpretar um PPL e o seu dual, dados pela Eq. (5.41) e Eq. (5.42), respectivamente, onde o PPL dado pela Eq. (5.41) representa a melhor utilização dos recursos e o seu dual, dado pela Eq. (5.42), representa a determinação dos valores dos recursos. Obviamente, os valores desses recursos correspondem à sua melhor utilização, o que explica correspondência entre as soluções ótimas do primal e do dual.

5.5 Uma Representação Gráfica do Problema Dual

Graficamente é possível ver uma relação entre o ponto ótimo do problema primal com o ponto ótimo do problema dual. Para isto, observe o exemplo abaixo.

Exemplo 5.5.1. *Resolva o PPL dado pela Eq. 5.46:*

$$\begin{aligned} \text{minimizar} \quad & F(x_1, x_2) = 2x_1 - 3x_2 \\ \text{sujeito a} \quad & 2x_1 + 3x_2 \geq 4 \\ & 2x_1 + 3x_2 \leq 6 \\ & x_1, x_2 \geq 0 \end{aligned} \quad (5.46)$$

Tem-se que o seu dual é dado pela Eq. 5.47:

$$\begin{aligned}
&\text{maximizar} && G(y_1, y_2) = 4y_1 - 6y_2 \\
&\text{sujeito a} && 2y_1 - 2y_2 \leq 2 \\
&&& -y_1 - 3y_2 \leq -3 \\
&&& x, y \geq 0
\end{aligned} \tag{5.47}$$

Tem-se que as regiões viáveis do primal e do dual são dadas na Fig. 5.2.

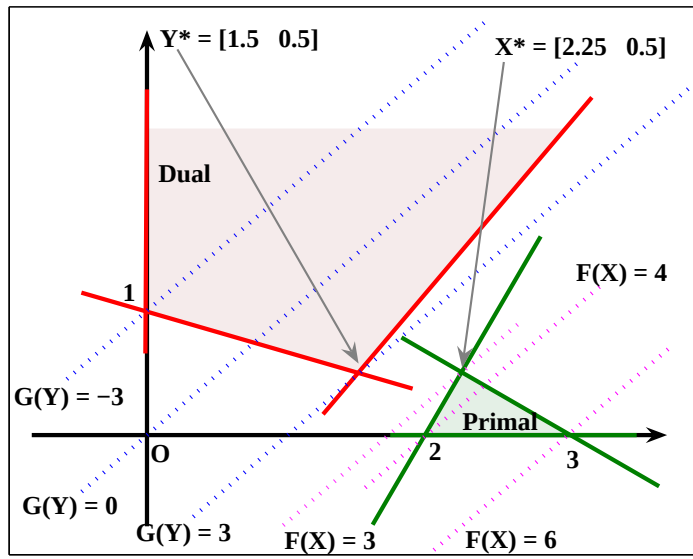


Figura 5.2: Representação da região viável do problema primal e do dual - Exemplo 5.5.1.

Observe que traçando retas da forma $2x_1 - 3x_2 = z$, fica fácil observar que o menor valor que o problema primal assume é quando $z = 3$, que é o ponto $X^* = [2.25 \ 0.5]$. Por outro lado, tem-se que o ponto de máximo do problema dual é $Y^* = [1.5 \ 0.5]$ que produz um valor máximo de $G(Y^*) = 3$, que é o mesmo valor obtido na minimização da função F . Portanto, o problema primal e o problema dual tem o mesmo valor funcional.

Capítulo 6

Um Problema de Transporte Rodoviário de Cargas

Neste capítulo é apresentado um problema proposto em Couto (2004). O objetivo é maximizar os lucros obtidos na transferência de cargas, entre várias cidades, por uma determinada transportadora, sendo que a mesma possui uma grande frota e uma grande área de cobertura. É importante ressaltar que este problema é similar ao clássico problema do caixeiro viajante.

Neste estudo adotou-se que o custo operacional está diretamente relacionado a quilometragem rodada por cada tipo de carreta e de cavalos mecânicos (caminhão). Por isto, não são levados em conta o peso ou o tipo de produtos transportados já que as Demandas são montadas levando em conta a quantidade e o tipo de carreta. Também não foi considerado a localidade do transporte e os custos particulares de cada carreta e cada cavalo mecânico e as possíveis alterações que ocorreriam devido a restrições climáticas, manutenção de veículos, fechamento de estradas, etc. Da mesma forma que na frota própria, os custos da frota terceirizada estão estimados por quilômetros rodados pelo conjunto que pode atender uma demanda. Por fim, o custo dos motoristas está relacionado, também, com a quilometragem.

O objetivo do problema é definir a frota, isto é, definir a quantidade de cada tipo de veículo que deve ser enviada para cada cidade, em cada uma das datas, definindo quais as cargas a serem terceirizadas para que todas as demandas sejam atendidas, de forma a maximizar os lucros na transferência de cargas entre as cidades. As restrições que tem que ser respeitadas são:

- capacidade da frota agregada;
- atender a todas as demandas;
- limite do comprimento do conjunto cavalo e carreta;
- capacidade do pátio, em números de carreta, de cada filial;
- horário de carregamento;
- viagem acima de 12 horas precisam de dois motoristas;
- prazo de entrega do pedido;
- datas de compras de veículos novos;
- saídas dos cavalos e carretas próprios para manutenção agendada.

Na formulação do problema foi usado o modelo matemático de multifluxos multiperiódicos, com restrições adicionais para controlar o atendimento das demandas e a capacidade das filiais. Para resolver o problema, os dados abaixo são necessários:

- V : o conjunto de todos os tipos de Cavalos;
- R : o conjunto de todos os tipos de Carretas;
- D : o conjunto de todas as demandas;
- N : o conjunto de todos os nós das malhas;
- M : o conjunto de todas as viagens (i, j) que possuem uma duração acima de 12 horas;
- G : o conjunto de todas as viagens (i, j) que possuem uma duração abaixo de 12 horas e que não representam veículos parados nas filiais;
- F_d : frete, ou ganho, para uma carreta atender à demanda d ;
- C_{td} : custo de terceirizar um conjunto, para atender à demanda d ;
- Dem_d^r : quantidade de carretas, do tipo r , requeridas pela demanda d ;
- $C_{i,j}^v$: custo da viagem de i para j , para um cavalo do tipo v ;
- $C_{i,j}^r$: custo da viagem de i para j , para um carreta do tipo r ;
- Q_i^v : quantidade de cavalos do tipo v que estão disponíveis na cidade i no dia de estudo;
- Q_i^r : quantidade de carretas do tipo r que estão disponíveis na cidade i no dia de estudo;

- CM : custo por hora de um motorista;
- $H_{i,j}$: quantidade de horas da viagem de i para j ;
- Cap_i : capacidade, em número de carretas, da filial i .

Com isto, o PPL fica definido pelas seguintes variáveis de projeto:

- $x_{i,j}^v$: número de cavalos do tipo v , indo da cidade i para a cidade j ;
- y_i^r : número de carretas do tipo r , parada na cidade i ;
- $z_{i,j}^{r,v}$: número de conjuntos de cavalos do tipo v e carreta do tipo r , vazios, indo da cidade i para a cidade j ;
- $w_{i,j}^{r,v,d}$: número de conjuntos formados por cavalos do tipo v e carretas do tipo r , atendendo a demanda d , indo da cidade i para a cidade j ;
- t^d : número de conjuntos terceirizados, atendendo a demanda d .

A Função Objetivo que representa o PPL modelado é escrita da seguinte forma:

$$Max : FO = FO_1 + FO_2 - FO_3 - FO_4 - FO_5 - FO_6 \quad (6.1)$$

sendo,

$$FO_1 = \sum_{i \in N} \sum_{j \in N} \sum_{r \in R} \sum_{v \in V} \sum_{d \in D} (F_d - C_{i,j}^r - C_{i,j}^v) \cdot w_{i,j}^{r,v,d} \quad (6.2)$$

a equação que representa o lucro das entregas das demandas obtidos pela frota própria;

$$FO_2 = \sum_{d \in D} (F_d - C_{t_d}) \cdot t^d \quad (6.3)$$

a equação que representa o lucro das entregas das demandas obtidos pela frota terceirizada;

$$FO_3 = \sum_{i \in N} \sum_{j \in N} \sum_{r \in R} \sum_{v \in V} (C_{i,j}^v + C_{i,j}^r) \cdot z_{i,j}^{r,v} \quad (6.4)$$

a equação que representa o custo das viagens dos conjuntos cavalo carreta viajando vazios;

$$FO_4 = \sum_{i \in N} \sum_{j \in N} \sum_{v \in V} C_{i,j}^v x_{i,j}^v \quad (6.5)$$

a equação que representa o custo das viagens dos cavalos, sem carretas;

$$FO_5 = 2CM \sum_{(i,j) \in M} H_{i,j} \cdot \left[\sum_{v \in V} x_{i,j}^v + \sum_{r \in R} \sum_{v \in V} z_{i,j}^{r,v} + \sum_{r \in R} \sum_{v \in V} \sum_{d \in D} w_{i,j}^{r,v,d} \right] \quad (6.6)$$

a equação que representa o custo dos motoristas com viagens acima de 12 horas;

$$FO_6 = CM \sum_{(i,j) \in M} H_{i,j} \cdot \left[\sum_{v \in V} x_{i,j}^v + \sum_{r \in R} \sum_{v \in V} z_{i,j}^{r,v} + \sum_{r \in R} \sum_{v \in V} \sum_{d \in D} w_{i,j}^{r,v,d} \right] \quad (6.7)$$

a equação que representa o custo dos motoristas com viagens de até 12 horas.

O PPL definido ppela Eq. (6.1) está sujeito as seguintes restrições:

$$g_1 : \sum_{j \in N} (x_{i,j}^v + \sum_{r \in R} z_{i,j}^{r,v} + \sum_{r \in R} \sum_{d \in D} w_{i,j}^{r,v,d}) = Q_i^v, \quad i \in S, \quad \forall v \in V \quad (6.8)$$

conjunto de restrições que representa a entrada de fluxo para cada cidade, em um dia, para cavalos do tipo v ;

$$g_2 : \sum_{j \in N} (y_i^r + \sum_{v \in V} z_{i,j}^{r,v} + \sum_{v \in V} \sum_{d \in D} w_{i,j}^{r,v,d}) = Q_i^r, \quad i \in S, \quad \forall r \in R \quad (6.9)$$

conjunto de restrições que representa o fluxo em cada cidade, num dia, representando as carretas do tipo r ;

$$g_3 : \sum_{j \in N} \left(x_{k,j}^v + \sum_{r \in R} z_{k,j}^{r,v} + \sum_{r \in R} \sum_{d \in D} w_{k,j}^{r,v,d} \right) - \sum_{i \in N} \left(x_{i,k}^v + \sum_{r \in R} z_{i,k}^{r,v} + \sum_{r \in R} \sum_{d \in D} w_{i,k}^{r,v,d} \right) = Q_k^r \quad (6.10)$$

$k \notin S, \quad k \notin T, \quad \forall v \in V$, é o conjunto de restrições representando a conservação do fluxo para uma determinada cidade k que não representa nem o primeiro nem o último dia para cavalos do tipo v ;

$$g_4 : \sum_{j \in N} (y_k^v + \sum_{v \in V} z_{k,j}^{r,v} + \sum_{v \in V} \sum_{d \in D} w_{k,j}^{r,v,d}) - \sum_{i \in N} (y_i^v + \sum_{v \in V} z_{i,k}^{r,v} + \sum_{v \in V} \sum_{d \in D} w_{i,k}^{r,v,d}) = Q_k^r \quad (6.11)$$

$k \notin S, k \notin T, \forall r \in R$, é o conjunto de restrições representando a conservação do fluxo para uma determinada cidade k que não representa nem o primeiro nem o último dia para carretas do tipo r ;

$$g_5 : t_d + \sum_{v \in V} w_{i,j}^{r,v,d} = Dem_d^r, \quad (6.12)$$

$\forall d \in D, \forall r \in R$ e o conjunto de restrições representando o atendimento da demanda indo da cidade i para a cidade j ;

$$g_6 : \left(\sum_{k_{in} \in N} \sum_{r \in R} y_{k_{in}}^r + \sum_{k_{in} \in N} \sum_{r \in R} \sum_{v \in V} z_{k_{in},i}^{r,v} + \sum_{k_{in} \in N} \sum_{r \in R} \sum_{v \in V} \sum_{d \in D} w_{k_{in},i}^{r,v,d} \right) - \left(\sum_{k_{out} \in N} \sum_{r \in R} y_{k_{out}}^r + \sum_{k_{out} \in N} \sum_{r \in R} \sum_{v \in V} z_{k_{out},i}^{r,v} + \sum_{k_{out} \in N} \sum_{r \in R} \sum_{v \in V} \sum_{d \in D} w_{i,k_{out}}^{r,v,d} \right) \leq Cap_i; \quad (6.13)$$

$\forall i \in N$ e o conjunto de restrições representando a capacidade de um pátio i , para todas as carretas;

$$g_7 : x_{i,j}^v \geq 0 \quad (6.14)$$

$\forall i \in N, \forall j \in N, \forall v \in V$

$$g_8 : y_i^r \geq 0 \quad (6.15)$$

$\forall i \in N, \forall j \in N, \forall r \in R$

$$g_9 : z_{i,j}^{r,v} \geq 0 \quad (6.16)$$

$\forall i \in N, \forall j \in N, \forall r \in R, \forall v \in V$

$$g_{10} : w_{i,j}^{r,v,d} \geq 0 \quad (6.17)$$

$\forall i \in N, \forall j \in N, \forall r \in R, \forall v \in V, \forall d \in D$

$$g_{11} : t^d \geq 0 \quad (6.18)$$

$\forall d \in D$.

As Eq. (6.14) até Eq. (6.18) representam as restrições laterais, das variáveis de projeto, que devem ser todas não negativas.

6.1 Número de Variáveis de Projeto e de Restrições do PPL

Seja NV_w o número de variáveis w que aparecem no PPL geral. É fácil observar que para cada demanda d e para cada par de cidade (i, j) são necessários k variáveis distintas, onde k representa o número de combinações cavalo-carreta diferentes que existem. Tomando como n o número de pares de cidades envolvidas no problema, segue que o número de variáveis relacionadas com a variável w é dado por:

$$NV_w = D * k * n, \quad (6.19)$$

onde D representa o número de demandas diferentes que precisam ser atendidas.

Para cada tipo de demanda d são necessárias uma nova variável t que representa a quantidade de frota terceirizada que é exigida para completar a frota própria e, assim, ser possível atender a demanda d . Logo, como existem D tipos de demandas distintas, tem-se que o número de variáveis t , representada por NV_t , relacionadas com o PPL é dado por:

$$NV_t = D. \quad (6.20)$$

Seja NV_z o número de variáveis z que aparecem no PPL geral. Para cada par de cidade (i, j) , tem-se que o número de variáveis é k , onde k representa o número de combinações cavalo-carreta diferentes que existem. Novamente, tomando como n o número de pares de cidades envolvidas no problema, segue que o número de variáveis relacionadas com a variável z é dado por:

$$NV_z = k * n. \quad (6.21)$$

Agora considere NV_x como sendo o número de variáveis relacionadas com x que aparece no PPL. Assim, para cada par de cidade, (i, j) são necessárias v variáveis, onde v é o número de

tipos de cavalos diferentes que aparecem na frota. Logo, o número de variáveis relacionadas com a variável x é dado por:

$$NV_x = v * n. \quad (6.22)$$

Por fim, considere NV_y como sendo o número de variáveis y que aparecem no PPL. Para cada cidade i são necessárias r variáveis relacionadas à variável y , onde r é o número de tipos de carretas disponíveis. Como N é o número de cidades envolvidas no problema, segue que o número de variáveis y é dado por:

$$NV_y = r * N. \quad (6.23)$$

Portanto, o número total de variáveis de projeto é dado pela soma da Eq. (6.19) até a Eq. (6.23), ou seja,

$$NV = NV_w + NV_t + NV_z + NV_x + NV_y = D * k * n + D + k * n + v * n + r * N, \quad (6.24)$$

onde, como já explicado, D é o número de tipos diferentes de demandas que tem devem ser atendidas pela frota, k é o número de combinações cavalo-carreta que podem ser formado com a frota própria, N é o número de cidades que são atendidas pela frota e n é o número de pares diferentes de cidades que podem ser formados para atuação da frota.

O primeiro conjunto de restrições é dado pela Eq. (6.8). Ele representa o número de cavalos do tipo v , atrelados ou não a uma carreta, que chegam a uma cidade. Assim, para cada tipo de cavalo v e cada cidade i aparece uma nova restrição. Por isto, considerando NR_{g_1} como sendo o número de restrições que aparecem neste grupo, tem-se que:

$$NR_{g_1} = v * N; \quad (6.25)$$

onde v é o número de tipos de cavalos diferentes existentes na frota.

O segundo conjunto de restrições é dado pela Eq. (6.9). Ele representa o número de carretas do tipo r , atrelados a um cavalo ou paradas no pátio de uma filial, em cada dia numa

determinada cidade. Assim, para cada tipo de carreta do tipo r e para cada cidade i aparece uma nova restrição. Logo, considerando NR_{g_2} como sendo o número de restrições que aparecem neste grupo, então ele é dado por:

$$NR_{g_2} = r * N; \quad (6.26)$$

onde r é o número de tipos de carretas diferentes existentes na frota.

Os próximos grupos de restrições são dados pelas Eq. (6.10) e Eq. (6.11). Elas são análogas ao grupo de restrições dado pelas Eq. (6.8) e Eq. (6.9), sendo que elas são usadas quando são considerados mais do que um dia de planejamento e que não estejam sendo tratados nem o primeiro dia nem o último dia. Estas restrições representam a conservação do fluxo em cada cidade, considerando os cavalos ou carretas que saem da cidade k menos os cavalos ou carretas que entram na cidade k . Assim, o grupo dado pela Eq. (6.10) tem o mesmo número de restrições que o conjunto formado pela Eq. (6.8) e Eq. (6.11) tem o mesmo número de restrições que o conjunto formado pela Eq. (6.9), logo,

$$NR_{g_3} = v * N; \quad e \quad NR_{g_4} = r * N. \quad (6.27)$$

O quinto e o sexto conjunto de restrições estão relacionados com o custo gerado com os motoristas, sendo que no primeiro caso são necessários mais do que um motorista, visto que as viagens tem duração maior do que 12 horas. No segundo caso, as viagens duram no máximo 12 horas e, por isto, precisa-se de apenas um motorista. Para qualquer um dos casos, o número de restrições está relacionado com o número de pares de cidades que podem ser formados, ou seja, para cada conjunto $\{(i, j) \text{ e } (j, i) / i, j \in N\}$ de cidades em estudo é necessário uma nova restrição. Portanto, considerando NR_{g_5} como sendo o número de restrições deste grupo, segue que o número de restrições relacionado com este grupo é:

$$NR_{g_5} = r * (N - 1), \quad (6.28)$$

onde N é o número de cidades envolvidas no projeto. Portanto, como as restrições de igualdade são dadas pelas Eq. (6.8) até Eq. (6.12), segue que o número total de restrições de igualdade é

$$NR_{eq.} = NR_{g_1} + NR_{g_2} + NR_{g_3} + NR_{g_4} + NR_{g_5} = 2(v + r) * N + r * (N - 1). \quad (6.29)$$

As restrições de desigualdade, apresentadas pela Eq. (6.13), estão relacionadas com a capacidade de cada pátio, em cada uma das cidades sedes. Assim, para cada cidade i é acrescentada uma nova restrição. Logo,

$$NR_{de.} = N. \quad (6.30)$$

As restrições laterais, apresentadas pelas Eq. (6.14) até Eq. (6.18), estão relacionadas com o número de variáveis, visto que para cada variável existe uma restrição lateral. Assim o número total de restrições laterais, NR_L , é igual ao número de variáveis de projeto, ou seja,

$$NR_L = NV = D * k * n + D + k * n + v * n + r * N. \quad (6.31)$$

6.2 Desenvolvimento Simplificado do Problema I

Para uma melhor compreensão do PPL definido pela Eq. (6.1) até a Eq. (6.18), considere o seguinte problema simplificado: um tipo de demanda, representado por s , duas cidades que são tanto cidades de saída como cidades de chegada, representadas por I e II . Além disso, suponha que existam 3 tipos de cavalos, representados por $\{a, b, c\}$, e 3 tipos de carretas, representados por $\{1, 2, 3\}$, sendo que apenas uma destas combinações, considere como sendo $(3, c)$, ultrapassa a medida permitida por lei para transitar em rodovias.

A Eq. (6.2) representa o lucro das entregas das demandas obtidos pela frota própria, sendo calculado pela diferença do valor recebido pelo frete da demanda s e a soma dos custos da viagem de cada cavalo com o custo de cada carreta, saindo da cidade i para a cidade j . Assim, para este conjunto de dados, segue que a Eq. (6.2) pode ser reescrita da seguinte forma:

$$\begin{aligned}
FO_1 = & (F_s - C_{I,II}^1 - C_{I,II}^a)w_{I,II}^{1,a,s} + (F_s - C_{I,II}^1 - C_{I,II}^b)w_{I,II}^{1,b,s} + (F_s - C_{I,II}^1 - C_{I,II}^c)w_{I,II}^{1,c,s} \\
& + (F_s - C_{I,II}^2 - C_{I,II}^a)w_{I,II}^{2,a,s} + (F_s - C_{I,II}^2 - C_{I,II}^b)w_{I,II}^{2,b,s} + (F_s - C_{I,II}^2 - C_{I,II}^c)w_{I,II}^{2,c,s} \\
& + (F_s - C_{I,II}^3 - C_{I,II}^a)w_{I,II}^{3,a,s} + (F_s - C_{I,II}^3 - C_{I,II}^b)w_{I,II}^{3,b,s} \\
& + (F_s - C_{II,I}^1 - C_{II,I}^a)w_{II,I}^{1,a,s} + (F_s - C_{II,I}^1 - C_{II,I}^b)w_{II,I}^{1,b,s} + (F_s - C_{II,I}^1 - C_{II,I}^c)w_{II,I}^{1,c,s} \\
& + (F_s - C_{II,I}^2 - C_{II,I}^a)w_{II,I}^{2,a,s} + (F_s - C_{II,I}^2 - C_{II,I}^b)w_{II,I}^{2,b,s} + (F_s - C_{II,I}^2 - C_{II,I}^c)w_{II,I}^{2,c,s} \\
& + (F_s - C_{II,I}^3 - C_{II,I}^a)w_{II,I}^{3,a,s} + (F_s - C_{II,I}^3 - C_{II,I}^b)w_{II,I}^{3,b,s}
\end{aligned} \tag{6.32}$$

Pode-se observar, então, que saindo da Cidade I e chegando na Cidade II , utilizando as combinações cavalo e carreta disponíveis, são necessários 8 variáveis distintas, variáveis que representam o número de cada combinação cavalo-carreta necessário para atender a demanda s . Como existe a necessidade de analisar o caminho inverso, ou seja, o número de conjuntos saindo da cidade II em direção a cidade I , este trajeto também acrescenta 8 novas variáveis. Portanto, para atender a demanda s são necessárias 16 variáveis distintas para a Eq. (6.2), que representa o lucro obtido de entrega da demanda, utilizando a frota própria.

O lucro obtido pela frota terceirizada, representada pela Eq. (6.3), não depende da cidade, sendo calculado como a diferença entre os fretes recebidos e o custo de se terceirizar um conjunto para atender a demanda s . Assim, utilizando os dados considerados neste estudo, a Eq. (6.3) pode ser reescrita como sendo

$$FO_2 = (F_s - C_{t_s}).t^s. \tag{6.33}$$

Logo, para cada demanda, a frota terceirizada representa o acréscimo de uma nova variável t .

O custo das viagens dos conjuntos cavalos do tipo v e carretas do tipo r que viajam vazios da cidade i para a cidade j é representado pela Eq. (6.4). Logo, reescrevendo a Eq. (6.4) tem-se:

$$\begin{aligned}
FO_3 = & (C_{I,II}^1 + C_{I,II}^a)z_{I,II}^{1,a} + (C_{I,II}^1 + C_{I,II}^b)z_{I,II}^{1,b} + (C_{I,II}^1 + C_{I,II}^c)z_{I,II}^{1,c} \\
& + (C_{I,II}^2 + C_{I,II}^a)z_{I,II}^{2,a} + (C_{I,II}^2 + C_{I,II}^b)z_{I,II}^{2,b} + (C_{I,II}^2 + C_{I,II}^c)z_{I,II}^{2,c} \\
& + (C_{I,II}^3 + C_{I,II}^a)z_{I,II}^{3,a} + (C_{I,II}^3 + C_{I,II}^b)z_{I,II}^{3,b} \\
& + (C_{II,I}^1 + C_{II,I}^a)z_{II,I}^{1,a} + (C_{II,I}^1 + C_{II,I}^b)z_{II,I}^{1,b} + (C_{II,I}^1 + C_{II,I}^c)z_{II,I}^{1,c} \\
& + (C_{II,I}^2 + C_{II,I}^a)z_{II,I}^{2,a} + (C_{II,I}^2 + C_{II,I}^b)z_{II,I}^{2,b} + (C_{II,I}^2 + C_{II,I}^c)z_{II,I}^{2,c} \\
& + (C_{II,I}^3 + C_{II,I}^a)z_{II,I}^{3,a} + (C_{II,I}^3 + C_{II,I}^b)z_{II,I}^{3,b}
\end{aligned} \tag{6.34}$$

Novamente, pode-se observar que para estabelecer o números de conjuntos viajando vazios, há uma necessidade de introduzir 8 novas variáveis, indo de I para II e o mesmo número indo de II para I , ou seja, são necessários 16 novas variáveis. Também é importante ressaltar que o custo FO_3 não depende da demanda s , ele depende apenas do tipo de carreta e do tipo de cavalo.

O custo das viagens com cavalos viajando sozinhos, conforme a Eq. (6.5), é obtido multiplicando o custo da viagem de um cavalo do tipo v pelo número de cavalos que estão viajando sozinhos. Assim, reescrevendo a Eq. (6.5) tem-se que:

$$FO_4 = C_{I,II}^a x_{I,II}^a + C_{I,II}^b x_{I,II}^b + C_{I,II}^c x_{I,II}^c + C_{II,I}^a x_{II,I}^a + C_{II,I}^b x_{II,I}^b + C_{II,I}^c x_{II,I}^c. \tag{6.35}$$

Logo, para cada tipo de cavalo aparece um variável indo de uma cidade para outra e, como estão sendo considerados apenas 3 tipos de cavalos, indo da cidade I para a cidade II são necessários 3 novas variáveis, totalizando 6 novas variáveis, considerando o caminho inverso. Portanto, o número de variáveis que aparecem no problema, considerando as Eq.(6.19) até Eq.(6.22), foi

$$NV_w + NV_t + NV_z + NV_x = 16 + 1 + 16 + 6 = 39. \tag{6.36}$$

O custo gerado para pagamento dos motoristas são divididos em duas partes. A primeira parte está relacionado com as viagens que tem duração maior do que 12 horas pois, neste caso, são necessários 2 motoristas distintos para efetuar esta viagem. A segunda está relacionada com as viagens com duração de até 12 horas pois, neste caso, existe a necessidade de ocupar apenas um motorista. Se a distância de I para II for maior do que 12 horas, a Eq. (6.6) representa o custo dos motoristas e, assim, ela pode ser reescrita da seguinte forma:

$$\begin{aligned}
 FO_5 = & 2.CM.H_{I,II} (\quad x_{I,II}^a \quad + x_{I,II}^b \quad + x_{I,II}^c \\
 & + z_{I,II}^{1,a} \quad + z_{I,II}^{1,b} \quad + z_{I,II}^{1,c} \quad + z_{I,II}^{2,a} \quad + z_{I,II}^{2,b} \quad + z_{I,II}^{2,c} \quad + z_{I,II}^{3,a} \quad + z_{I,II}^{3,b} \\
 & + w_{I,II}^{1,a,s} + w_{I,II}^{1,b,s} + w_{I,II}^{1,c,s} + w_{I,II}^{2,a,s} + w_{I,II}^{2,b,s} + w_{I,II}^{2,c,s} + w_{I,II}^{3,a,s} + w_{I,II}^{3,b,s}) \\
 & + 2.CM.H_{II,I} (\quad x_{II,I}^a \quad + x_{II,I}^b \quad + x_{II,I}^c \\
 & + z_{II,I}^{1,a} \quad + z_{II,I}^{1,b} \quad + z_{II,I}^{1,c} \quad + z_{II,I}^{2,a} \quad + z_{II,I}^{2,b} \quad + z_{II,I}^{2,c} \quad + z_{II,I}^{3,a} \quad + z_{II,I}^{3,b} \\
 & + w_{II,I}^{1,a,s} + w_{II,I}^{1,b,s} + w_{II,I}^{1,c,s} + w_{II,I}^{2,a,s} + w_{II,I}^{2,b,s} + w_{II,I}^{2,c,s} + w_{II,I}^{3,a,s} + w_{II,I}^{3,b,s})
 \end{aligned}
 \tag{6.37}$$

Caso a distância de I para II seja no máximo 12 horas, então passa a valer a Eq. (6.7) e ela é reescrita da seguinte forma:

$$\begin{aligned}
 FO_6 = & CM.H_{I,II} (\quad x_{I,II}^a \quad + x_{I,II}^b \quad + x_{I,II}^c \\
 & + z_{I,II}^{1,a} \quad + z_{I,II}^{1,b} \quad + z_{I,II}^{1,c} \quad + z_{I,II}^{2,a} \quad + z_{I,II}^{2,b} \quad + z_{I,II}^{2,c} \quad + z_{I,II}^{3,a} \quad + z_{I,II}^{3,b} \\
 & + w_{I,II}^{1,a,s} + w_{I,II}^{1,b,s} + w_{I,II}^{1,c,s} + w_{I,II}^{2,a,s} + w_{I,II}^{2,b,s} + w_{I,II}^{2,c,s} + w_{I,II}^{3,a,s} + w_{I,II}^{3,b,s}) \\
 & + CM.H_{II,I} (\quad x_{II,I}^a \quad + x_{II,I}^b \quad + x_{II,I}^c \\
 & + z_{II,I}^{1,a} \quad + z_{II,I}^{1,b} \quad + z_{II,I}^{1,c} \quad + z_{II,I}^{2,a} \quad + z_{II,I}^{2,b} \quad + z_{II,I}^{2,c} \quad + z_{II,I}^{3,a} \quad + z_{II,I}^{3,b} \\
 & + w_{II,I}^{1,a,s} + w_{II,I}^{1,b,s} + w_{II,I}^{1,c,s} + w_{II,I}^{2,a,s} + w_{II,I}^{2,b,s} + w_{II,I}^{2,c,s} + w_{II,I}^{3,a,s} + w_{II,I}^{3,b,s})
 \end{aligned}
 \tag{6.38}$$

Observe que, na Eq. (6.37) ou na Eq. (6.38) não são acrescentadas novas variáveis. As últimas variáveis a serem acrescentadas no problema estão relacionadas com as carretas que não

estão em movimento, ou seja, as carretas que estão no pátio sem estar atreladas a um cavalo. Como elas não geram custos, pois estão paradas, elas não aparecem na Eq. (6.1), contudo elas são consideradas nas restrições.

Usando a Eq. (6.8), tem-se que o grupo de restrições dadas por esta equação podem ser reescritas da seguinte forma:

$$\begin{aligned}
 g_1 : \quad & x_{I,II}^v + z_{I,II}^{1,v} + z_{I,II}^{2,v} + z_{I,II}^{3,v} \\
 & + w_{I,II}^{1,v,s} + w_{I,II}^{2,v,s} + w_{I,II}^{3,v,s} = Q_I^v, \quad v = a, b, c \\
 & e \\
 & x_{II,I}^v + z_{II,I}^{1,v} + z_{II,I}^{2,v} + z_{II,I}^{3,v} \\
 & + w_{II,I}^{1,v,s} + w_{II,I}^{2,v,s} + w_{II,I}^{3,v,s} = Q_{II}^v, \quad v = a, b, c
 \end{aligned} \tag{6.39}$$

Como estão sendo considerados três tipos diferentes de cavalos e 2 cidades, segue da Eq. (6.25) que g_1 fornece um conjunto com 6 restrições diferentes.

O segundo conjunto de restrições, dadas por Eq. (6.9), podem ser reescritas da seguinte forma:

$$\begin{aligned}
 g_2 : \quad & y_I^r + z_{I,II}^{r,a} + z_{I,II}^{r,b} + z_{I,II}^{r,c} \\
 & + w_{I,II}^{r,a,s} + w_{I,II}^{r,b,s} + w_{I,II}^{r,c,s} = Q_I^r, \quad r = 1, 2, 3 \\
 & e \\
 & y_{II}^r + z_{II,I}^{r,a} + z_{II,I}^{r,b} + z_{II,I}^{r,c} \\
 & + w_{II,I}^{r,a,s} + w_{II,I}^{r,b,s} + w_{II,I}^{r,c,s} = Q_{II}^r, \quad r = 1, 2, 3
 \end{aligned} \tag{6.40}$$

Como estão sendo considerados três tipos diferentes de carretas e 2 cidades distintas, pela Eq. (6.26) tem-se que g_2 se desdobra em 6 restrições.

O último grupo de variáveis que são necessárias para modelar o PPL, estão relacionadas com número de carretas que não estão em movimento, ou seja, são as variáveis y . Observe que na cidade I podem ter 3 tipos diferentes de carretas e o mesmo ocorre com a cidade II , logo, o

número de variáveis relacionadas com estas carretas é igual ao número 6, totalizando 45 variáveis para o problema. Este número é igual ao número dado pela Eq. (6.24), onde $D = 1$, $k = 8$, $n = 2$, $v = 3$, $r = 3$ e $N = 2$. Assim,

$$NV = 1 * 8 * 2 + 1 + 8 * 2 + 3 * 2 + 3 * 2 = 45. \quad (6.41)$$

O novo grupo de restrições é dado pelas Eq. (6.10) e Eq. (6.11). Elas são usadas quando são considerados mais do que 1 dia de planejamento e que não estejam sendo tratados nem o primeiro dia nem o último dia. Estas restrições representam a conservação do fluxo em cada cidade, considerando os cavalos ou carretas que saem da cidade k menos os cavalos ou carretas que entram na cidade k . Assim, para cada tipo de cavalo v de V e cada cidade k , a Eq. (6.10) pode ser reescrita da seguinte forma:

$$\begin{aligned} g_3 : \quad & (x_{k,II}^v + z_{k,II}^{1,v} + z_{k,II}^{2,v} + z_{k,II}^{3,v} + w_{k,II}^{1,v,II} \\ & + w_{k,II}^{2,v,s} + w_{k,II}^{3,v,s}) - (x_{I,k}^v + z_{I,k}^{1,v} + z_{I,k}^{2,v} \\ & + z_{I,k}^{3,v} + w_{I,k}^{1,v,s} + w_{I,k}^{2,v,s} + w_{I,k}^{3,v,s}) = Q_k^v, \quad v = a, b, c \end{aligned} \quad (6.42)$$

Para cada cidade k , que não estão no conjunto das cidades consideradas como cidades iniciais e as consideradas cidades finais, e cada tipo de cavalo v vai produzir uma nova restrição dada pela Eq. (6.42). Analogamente, para cada tipo de carreta do tipo r cada cidade k , a Eq. (6.11) pode ser reescrita da seguinte forma:

$$\begin{aligned} g_4 : \quad & (y_k^r + z_{k,II}^{r,a} + z_{k,II}^{r,b} + z_{k,II}^{r,c} + w_{k,II}^{r,a,II} \\ & + w_{k,i}^{r,b,s} + w_{k,i}^{r,c,s}) - (y_I^r + z_{I,k}^{r,a} + z_{I,k}^{r,b} \\ & + z_{I,k}^{r,c} + w_{I,k}^{r,a,s} + w_{I,k}^{r,b,s} + w_{I,k}^{r,c,s}) = Q_k^r, \quad r = 1, 2, 3 \end{aligned} \quad (6.43)$$

Como esperado, este conjunto de restrições fornece 6 restrições, para cada cidade k ,

totalizando $6k$ restrições, onde $k \neq S$ e $k \neq T$, sendo que S é o conjunto de índices das cidade de saída e T é o conjunto de índices das cidades finais.

O conjunto de restrições dada pela Eq. (6.12) está relacionado com o atendimento das demandas, que sai da cidade I para a cidade II , e vice e versa, para cada tipo de carreta r . Assim, para cada tipo de demanda s resulta:

$$g_5 = t_s + w_{I,II}^{r,a,s} + w_{I,II}^{r,b,s} + w_{I,II}^{r,c,s} = Dem_d^r, \quad r = 1, 2, 3 \quad (6.44)$$

onde cada demanda s e cada carreta r produz uma nova restrição. Assim, neste grupo são produzidas 3 restrições.

O conjunto de restrições relacionado com g_6 , dados pela Eq. (6.13), está relacionado com a capacidade de estacionamento das carretas no pátio de cada filial situada numa cidade. Para cada cidade i o número de carretas que passam de um dia para o outro é dado pela diferença entre todas as carretas que chegam na cidade no dia menos o o número de carretas que saem da cidade naquele dia. Assim, para cada cidade i a restrição é dada por:

$$\begin{aligned} g_6 : \quad & (y_{k_{in}}^1 + y_{k_{in}}^2 + y_{k_{in}}^3 + z_{k_{in},I}^{1,a} + z_{k_{in},I}^{1,b} + z_{k_{in},I}^{1,c} \\ & + z_{k_{in},I}^{2,a} + z_{k_{in},I}^{2,b} + z_{k_{in},I}^{2,c} + z_{k_{in},I}^{3,a} + z_{k_{in},I}^{3,b} \\ & + w_{k_{in},I}^{1,a,s} + w_{k_{in},I}^{1,b,s} + w_{k_{in},I}^{1,c,s} + w_{k_{in},I}^{2,a,s} + w_{k_{in},I}^{2,b,s} + w_{k_{in},I}^{2,c,s} \\ & + w_{k_{in},I}^{3,a,s} + w_{k_{in},I}^{3,b,s} - (y_I^1 + y_I^2 + y_I^3 \\ & + z_{I,k_{out}}^{1,a} + z_{I,k_{out}}^{1,b} + z_{I,k_{out}}^{1,c} + z_{I,k_{out}}^{2,a} + z_{I,k_{out}}^{2,b} + z_{I,k_{out}}^{2,c} \\ & + z_{I,k_{out}}^{3,a} + z_{I,k_{out}}^{3,b} + w_{I,k_{out}}^{1,a,s} + w_{I,k_{out}}^{1,b,s} + w_{I,k_{out}}^{1,c,s} \\ & + w_{I,k_{out}}^{2,a,s} + w_{I,k_{out}}^{2,b,s} + w_{I,k_{out}}^{2,c,s} + w_{I,k_{out}}^{3,a,s} + w_{I,k_{out}}^{3,b,s} = Cap_I, \end{aligned} \quad (6.45)$$

onde k_{out} são os índices das cidades de estão chegando na cidade i e k_{in} são os índices que estão saindo de i .

As Eq. (6.14) até a Eq. (6.18) são as restrições laterais que indicam que o número de carretas, cavalos e conjuntos carretas-cavalos não podem ser negativos. Logo, para cada variável de projeto aparece uma nova restrição.

O número total de restrições de igualdade é dado pela Eq.(6.29), onde $v = 3$, $r = 3$ e $N = 2$. Assim,

$$NR_{eq.} = 2 * (3 + 3) * 2 + 3 * (2 - 1) = 27, \quad (6.46)$$

e número total de restrições de desigualdade é dado pela Eq.(6.30), onde $N = 2$. Assim,

$$NR_{de.} = 2. \quad (6.47)$$

6.2.1 Simulação Numérica para o Problema de Transporte I

Para efetuar a simulação numérica, é considerado o conjunto de dados apresentados na Tab. 6.1.

Tabela 6.1: Dados de entrada para o Problema de Transporte I (Simplificado)

N	$=$	2;	V	$=$	3;
R	$=$	3;	D	$=$	1;
CM	$=$	R\$10,00;	F_d	$=$	R\$100,00;
C_{td}	$=$	R\$50,00;	$C_{I,II}^v = C_{II,I}^v$	$=$	[10 12 14];
$C_{I,II}^r = C_{II,I}^r$	$=$	[07 09 11];	Dem_d^r	$=$	[40 40 40];
$Q_{I,II}^v$	$=$	[10 10 10];	$Q_{I,II}^r$	$=$	[15 15 15];
Cap_i	$=$	[50 50];	$H_{I,II}$	$=$	[03 03];

É importantes ressaltar que os custos C_v são dados na ordem $\{a, b, c\}$ e o custo C_r são dados na ordem $\{1, 2, 3\}$. As quantidades de demanda a serem atendidas também são apresentadas de forma crescente. É preciso lembrar que apenas a combinação $(3, c)$ carreta-cavalo não pode ser formada.

Para simplificar a apresentação, considere que o vetor custo pode ser escrito da seguinte forma

$$C = \begin{bmatrix} C_w & C_t & -C_z & -C_x & C_y \end{bmatrix}^T; \quad (6.48)$$

onde C_w é formado pelos coeficientes relacionados com a variável w ; C_t é formado pelos coeficientes relacionados com a variável t ; C_z é formado pelos coeficientes relacionados com a variável z ; C_x é formado pelos coeficientes relacionados com a variável x e C_y é formado pelos coeficientes relacionados com a variável y .

Cada coeficiente do vetor C_w tem que ser calculado usando as informações apresentadas nas Eq. (6.32) e Eq. (6.38), visto que está sendo considerado uma distância de 3 horas entre as cidades. Então, com as informações contidas na Tab. 6.1 chega-se a cada um destes coeficientes. Por exemplo, o primeiro coeficiente é obtido da seguinte forma:

$$C_{w_1} = (F_s - C_{I,II}^1 - C_{I,II}^a) - CM.H_{I,II} = (100 - 10 - 7) - 10 * 3 = 100 - 47 = 53. \quad (6.49)$$

De maneira análoga todos os coeficientes são calculados, obtendo o vetor:

$$C_w = \left[53 \ 51 \ 49 \ 51 \ 49 \ 47 \ 49 \ 47 \ 53 \ 51 \ 49 \ 51 \ 49 \ 47 \ 49 \ 47 \right]^T \quad (6.50)$$

Para obter o valor dos coeficientes relacionado com a variável t , usa-se a Eq. (6.33), pois é a única parte da função objetivo que está relacionada com esta variável. Assim,

$$C_{t_1} = F_s - C_{t_s} = 100 - 50 = 50. \quad (6.51)$$

ou seja, tem-se o vetor relacionado com esta variável:

$$C_t = \left[50 \right]^T \quad (6.52)$$

Os coeficiente do vetor C_z devem ser calculados usando as Eq. (6.34) e Eq. (6.38). e com as informações contidas na Tab. 6.1. Por exemplo, o primeiro coeficiente é obtido da seguinte forma:

$$C_{z_1} = (C_{I,II}^1 + C_{I,II}^a) + CM.H_{I,II} = (07 + 10) + 10 * 3 = 17 + 30 = 47. \quad (6.53)$$

De maneira análoga, obtém-se todos os outros:

$$C_z = \left[47 \ 49 \ 51 \ 49 \ 51 \ 53 \ 51 \ 53 \ 47 \ 49 \ 51 \ 49 \ 51 \ 53 \ 51 \ 53 \right]^T \quad (6.54)$$

Por outro lado, os coeficientes do vetor C_x são calculados usando as Eq. (6.35) e Eq. (6.38), com os dados da Tab. 6.1. Novamente, o primeiro coeficiente é obtido como

$$C_{z_1} = C_{I,II}^1 + CM.H_{I,II} = 10 + 10 * 3 = 40. \quad (6.55)$$

Os outros coeficientes são calculados da mesma forma, sendo o vetor dado por:

$$C_x = \left[40 \ 42 \ 44 \ 40 \ 42 \ 44 \right]^T \quad (6.56)$$

Os coeficientes das variáveis y são todos nulos, visto que as carretas paradas não geram custos. Assim, o vetor relacionado a estas variáveis vale:

$$C_y = \left[0 \ 0 \ 0 \ 0 \ 0 \ 0 \right]^T \quad (6.57)$$

O vetor X das variáveis de projeto está disposto da seguinte maneira:

$$X = \begin{bmatrix} X_w & X_t & X_z & X_x & X_y \end{bmatrix}^T \quad (6.58)$$

onde

$$\begin{array}{ccccc} X_w & & X_z & & \\ \parallel & & \parallel & & \\ \left[\begin{array}{c} w_{I,II}^{1,a,s} \\ w_{I,II}^{1,b,s} \\ w_{I,II}^{1,c,s} \\ w_{I,II}^{2,a,s} \\ w_{I,II}^{2,b,s} \\ w_{I,II}^{2,c,s} \\ w_{I,II}^{3,a,s} \\ w_{I,II}^{3,b,s} \\ w_{II,I}^{1,a,s} \\ w_{II,I}^{1,b,s} \\ w_{II,I}^{1,c,s} \\ w_{II,I}^{2,a,s} \\ w_{II,I}^{2,b,s} \\ w_{II,I}^{2,c,s} \\ w_{II,I}^{3,a,s} \\ w_{II,I}^{3,b,s} \end{array} \right] & X_t & \left[\begin{array}{c} z_{I,II}^{1,a} \\ z_{I,II}^{1,b} \\ z_{I,II}^{1,c} \\ z_{I,II}^{2,a} \\ z_{I,II}^{2,b} \\ z_{I,II}^{2,c} \\ z_{I,II}^{3,a} \\ z_{I,II}^{3,b} \\ z_{II,I}^{1,a} \\ z_{II,I}^{1,b} \\ z_{II,I}^{1,c} \\ z_{II,I}^{2,a} \\ z_{II,I}^{2,b} \\ z_{II,I}^{2,c} \\ z_{II,I}^{3,a} \\ z_{II,I}^{3,b} \end{array} \right] & & \\ & \parallel & & & \\ & \left[\begin{array}{c} t^s \end{array} \right] & & & \\ & & & X_x & X_y \\ & & & \parallel & \parallel \\ & & & \left[\begin{array}{c} x_{I,II}^a \\ x_{I,II}^b \\ x_{I,II}^c \\ x_{II,I}^a \\ x_{II,I}^b \\ x_{II,I}^c \end{array} \right] & \left[\begin{array}{c} y_{I,II}^1 \\ y_{I,II}^2 \\ y_{I,II}^3 \\ y_{II,I}^1 \\ y_{II,I}^2 \\ y_{II,I}^3 \end{array} \right] \end{array} \quad (6.59)$$

Usando as Eq. (6.39), Eq. (6.40), Eq. (6.42), Eq. (6.43) e Eq. (6.44), escreve-se a matriz $\mathcal{A}_{eq}$, dado na Eq. (6.60), que é obtida das restrições de igualdade. Usando a Eq. (6.45) tem-se a matriz $\mathcal{A}_{de}$, dada na Eq. (6.61), que é a matriz obtida das restrições de desigualdade.

$$A_{de} = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad (6.61)$$

Com todas as informações contidas na Tab. 6.1, é possível construir o vetor dos recursos disponíveis b_{eq} , que está relacionado com as restrições de igualdade. Este vetor é dado por:

$$b_{eq} = \begin{bmatrix} 10 & 10 & 10 & 10 & 10 & 10 & 10 & 15 & 15 & 15 & 15 & 15 & 15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 40 & 40 & 40 \end{bmatrix}^T. \quad (6.62)$$

O vetor dos recursos disponíveis relacionado com as restrições de desigualdade b_{de} é dado por:

$$b_{de} = \begin{bmatrix} 50 & 50 \end{bmatrix}^T \quad (6.63)$$

Com o objetivo de comparação, serão utilizados três códigos: o MOSEK (Mathematical Optimization Software Specified) apresentado em Mosek (2006), o Linprog (Linear Programming), ambos utilizando o Software MatLab, MathWorks (2006). O código LINDO (Linear, Interactive and Discrete Optimizer), projetado por Prado (1999), também é testado. Estes códigos permitem a escolha dos seguintes métodos de otimização: Simplex ou Ponto Interior. Nesta pesquisa, todos os códigos foram aplicados utilizando o método Simplex. O computador utilizado para efetuar os cálculos tem a seguinte configuração: processador Intel Celeron D315 2.26 Ghz e memória 256 MB DDR PC3200 / 400 Mhz. O valor da função objetivo, o tempo gasto e o número de iterações de cada um dos códigos são apresentados na Tab. 6.2.

Tabela 6.2: Resultados obtidos para o Problema de Transporte I

	Função Objetivo (R\$)	Tempo Computacional (s)	Número de Iterações
Linprog	3940	2.10	24
Mosek	3940	1.00	08
LINDO	3940	3.00	17

Os vetores X_w , X_t , X_z , X_x e X_y são dados, respectivamente, pelas Tab. 6.3, Tab. 6.4, Tab. 6.5, Tab. 6.6 e Tab. 6.7.

Tabela 6.3: Vetor X_w obtido na solução do Problema de Transporte I.

Linprog	$\begin{bmatrix} 5 & 0 & 10 & 0 & 10 & 0 & 5 & 0 & 0 & 0 & 5 & 0 & 5 & 5 & 10 & 5 \end{bmatrix}^T$
Mosek	$\begin{bmatrix} 0 & 0 & 10 & 10 & 0 & 0 & 0 & 10 & 10 & 0 & 0 & 0 & 0 & 10 & 0 & 10 \end{bmatrix}^T$
LINDO	$\begin{bmatrix} 0 & 10 & 0 & 0 & 0 & 10 & 10 & 0 & 0 & 0 & 10 & 10 & 0 & 0 & 0 & 10 \end{bmatrix}^T$

Pode-se observar que os vetores obtidos em cada um dos códigos são diferentes, mas eles apresentaram o mesmo valor da função objetivo. Mesmo sendo distintos, existe uma certa simetria entre eles, isto é, as primeiras coordenadas da Tab. 6.3 saem da cidade *I* em direção à cidade *II* e, nos três códigos, foram necessários 30 conjuntos cavalo-carretas, da frota própria, para atender a demanda. O mesmo ocorre com o trajeto entre a cidade *II* e a cidade *I*. Analisando a Tab. 6.4 tanto no trajeto da cidade *I* para a cidade *II*, quanto no caminho inverso, foram necessário terceirizar 25% da demanda, sendo preciso, então, 20 conjuntos terceirizados para atender a demanda.

Tabela 6.4: Vetor X_t obtido na solução do Problema de Transporte I.

Linprog	$\begin{bmatrix} 20 \end{bmatrix}^T$
Mosek	$\begin{bmatrix} 20 \end{bmatrix}^T$
LINDO	$\begin{bmatrix} 20 \end{bmatrix}^T$

Pode-se observar, também, que todos os códigos desconsideraram o caso onde tanto os conjunto viajavam vazios, formado pelas variáveis z , como o caso onde os cavalos sem carretas, viajavam sozinhos, formado pelas variáveis x , isto é, a solução ótima apresentada em cada um considera que é melhor não ter nenhum conjunto (ou cavalo) viajando vazio.

Tabela 6.5: Vetor X_z obtido na solução do Problema de Transporte I.

Linprog	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$
Mosek	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$
LINDO	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$

Sobre as carretas paradas nos pátios, todos os códigos deixaram o mesmo número de carretas em cada uma das cidades, como visto na Tab. 6.7. O que diferencia uma solução da outra é apenas o tipo de carreta que fica parada, dependendo do tipo de carreta que foi usada para atender a demanda s .

Tabela 6.6: Vetor X_x obtido na solução do Problema de Transporte I.

Linprog	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$
Mosek	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$
LINDO	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$

Tabela 6.7: Vetor X_y obtido na solução do Problema de Transporte I.

Linprog	$\begin{bmatrix} 0 & 5 & 10 & 10 & 5 & 0 \end{bmatrix}^T$
Mosek	$\begin{bmatrix} 5 & 5 & 5 & 5 & 5 & 5 \end{bmatrix}^T$
LINDO	$\begin{bmatrix} 5 & 5 & 5 & 5 & 5 & 5 \end{bmatrix}^T$

É fácil observar que os três códigos obtiveram o mesmo valor da função objetivo, apesar das soluções obtidas serem distintas, o que leva a suspeita de que existem infinitas soluções. A maior diferença observada na aplicação dos códigos foi em relação ao tempo gasto e ao número de iterações, sendo que o código MOSEK mostrou-se mais eficiente (Tab. 6.2). Outra informação importante está relacionada ao fato que, para problemas com alto número de variáveis e restrições, o código LINDO não se torna recomendado, visto que ele não trabalha diretamente com as matrizes e, para sua aplicação, existe a necessidade de escrever todas as equações envolvida nos PPL, fazendo com que este processo não seja eficiente e recomendável para grandes problemas.

É importante ressaltar que todas as restrições foram respeitadas, por todos os códigos, o que garante que as soluções obtidas são ótimas.

6.3 Simulação Numérica para o PPL de Transporte II

Para esta simulação, é considerado o conjunto de dados apresentado na Tab. 6.8.

Neste caso os três tipos de V são dados por $\{a, b, c\}$ e os três tipos de R são dados na ordem $\{1, 2, 3\}$. Os tipos de demanda que serão atendidas são tomadas como sendo $\{1, 2, 3, 4\}$. Neste caso também é considerado que apenas a combinação $(3, c)$ carreta-cavalo não pode ser formada. As cidades consideradas são numeradas como sendo I , II e III .

O vetor custo é escrito como a Eq. (6.64), onde $C_{i,j}^{rw}$, $\forall i$ e $\forall j$, são formado pelos coeficientes relacionados com a variável w ; C_t pelos coeficientes relacionados com a variável t ; $-C_i^z$, $\forall i$, pelos coeficientes relacionados com a variável z ; C_x pelos coeficientes relacionados com a variável x e

Tabela 6.8: Dados de entrada para o Problema de Transporte II (271 Variáveis)

N	$=$	3;	V	$=$	3;
R	$=$	3;	D	$=$	4;
CM	$=$	(R\$) 10,00;	C_{td}	$=$	(R\$) [50 120 230 350];
$C_{i,j}^a$	$=$	$\begin{bmatrix} 0 & 12 & 13 \\ 12 & 0 & 14 \\ 13 & 14 & 0 \end{bmatrix}$;	$C_{i,j}^b$	$=$	$\begin{bmatrix} 0 & 14 & 16 \\ 14 & 0 & 18 \\ 16 & 18 & 0 \end{bmatrix}$;
$C_{i,j}^c$	$=$	$\begin{bmatrix} 0 & 15 & 18 \\ 15 & 0 & 21 \\ 18 & 21 & 0 \end{bmatrix}$;	$C_{i,j}^1$	$=$	$\begin{bmatrix} 0 & 5 & 6 \\ 5 & 0 & 7 \\ 6 & 7 & 0 \end{bmatrix}$;
$C_{i,j}^2$	$=$	$\begin{bmatrix} 0 & 6 & 7 \\ 6 & 0 & 8 \\ 7 & 8 & 0 \end{bmatrix}$;	$C_{i,j}^3$	$=$	$\begin{bmatrix} 0 & 7 & 9 \\ 7 & 0 & 11 \\ 9 & 11 & 0 \end{bmatrix}$;
$H_{I,II}$	$=$	$\begin{bmatrix} 0 & 2 & 13 \\ 2 & 0 & 6 \\ 13 & 6 & 0 \end{bmatrix}$;	F_d	$=$	(R\$)[100 200 400 600];
$Q_{i,j}^v$	$=$	10 $\forall i, \forall j, \forall v$;	$Q_{i,i}^r$	$=$	15 $\forall i, \forall j, \forall r$;
Cap_i	$=$	[50 50 50];	Dem_d^r	$=$	40, $\forall d, \forall r$;

C_y pelos coeficientes relacionados com a variável y .

$$C = \left[C_{I,II}^w \quad C_{I,III}^w \quad C_{II,I}^w \quad C_{II,III}^w \quad C_{III,I}^w \quad C_{III,II}^w \quad C_t \quad -C_I^z \quad -C_{II}^z \quad -C_{III}^z \quad -C_I^x \quad -C_{II}^x \quad -C_{III}^x \quad C_y \right]^T; \quad (6.64)$$

O número total de variáveis é dada pela Eq. (6.24), sendo que neste caso $D = 4$, $r = 3$, $v = 3$, $k = 8$, $n = 6$ e $N = 3$. Assim, para este caso, são necessárias

$$NV_{PPL2} = 4 * 8 * 6 + 4 + 8 * 6 + 3 * 6 + 3 * 3 = 271 \text{ variáveis.} \quad (6.65)$$

Dentre estas 271 variáveis, o número de variáveis relacionadas com w é dado pela Eq.(6.19) e, usando os dados da Tab. 6.8 tem-se que neste problema são necessárias 192 variáveis de projeto w , sendo 32 variáveis para cada par de cidade (i, j) . O número de variáveis relacionadas com t é dado pela Eq.(6.20) e, novamente usando os dados da Tab. 6.8, tem-se que neste problema são necessárias 4 variáveis de projeto t . De modo análogo, usando a Eq.(6.21) tem-se que são necessárias 48 variáveis relacionadas com z , usando a Eq. (6.22) tem-se que são necessárias 18 variáveis relacionadas com x e pela Eq.(6.23) são necessárias 9 variáveis relacionadas a y ,

totalizando as 271 variáveis de projeto.

Cada coeficiente do vetor $C_{i,j}^w$ pode ser calculado usando as informações apresentadas nas Eq. (6.2) e na Eq.(6.6), se a distância entre as cidades for maior do que 12 horas, usando a Eq.(6.7) se a distância entre as cidades for menor ou igual a 12 horas. Então, com as informações contidas na Tab. 6.8 chega-se a cada um dos coeficientes, conforme mostrado na Eq. (6.66).

$$\begin{array}{c} C_{I,II}^w \\ \parallel \\ \left[\begin{array}{c} 63 \\ 163 \\ 363 \\ 563 \\ 61 \\ 161 \\ 361 \\ 561 \\ 60 \\ 160 \\ 360 \\ 560 \\ 62 \\ 162 \\ 362 \\ 562 \\ 60 \\ 160 \\ 360 \\ 560 \\ 59 \\ 159 \\ 359 \\ 559 \\ 61 \\ 161 \\ 361 \\ 561 \\ 59 \\ 159 \\ 359 \\ 559 \end{array} \right] \end{array} \quad \begin{array}{c} C_{I,III}^w \\ \parallel \\ \left[\begin{array}{c} -179 \\ -79 \\ 121 \\ 321 \\ -182 \\ -82 \\ 118 \\ 318 \\ -184 \\ -84 \\ 116 \\ 316 \\ -180 \\ -80 \\ 120 \\ 320 \\ -183 \\ -83 \\ 117 \\ 317 \\ -185 \\ -85 \\ 115 \\ 315 \\ -182 \\ -82 \\ 118 \\ 318 \\ -185 \\ -85 \\ 115 \\ 315 \end{array} \right] \end{array} \quad \begin{array}{c} C_{II,I}^w \\ \parallel \\ \left[\begin{array}{c} 63 \\ 163 \\ 363 \\ 563 \\ 61 \\ 161 \\ 361 \\ 561 \\ 60 \\ 160 \\ 360 \\ 560 \\ 62 \\ 162 \\ 362 \\ 562 \\ 60 \\ 160 \\ 360 \\ 560 \\ 59 \\ 159 \\ 359 \\ 559 \\ 61 \\ 161 \\ 361 \\ 561 \\ 59 \\ 159 \\ 359 \\ 559 \end{array} \right] \end{array} \quad \begin{array}{c} C_{II,III}^w \\ \parallel \\ \left[\begin{array}{c} 19 \\ 119 \\ 319 \\ 519 \\ 15 \\ 115 \\ 315 \\ 515 \\ 12 \\ 112 \\ 312 \\ 512 \\ 18 \\ 118 \\ 318 \\ 518 \\ 14 \\ 114 \\ 314 \\ 514 \\ 11 \\ 111 \\ 311 \\ 511 \\ 15 \\ 115 \\ 315 \\ 515 \\ 11 \\ 111 \\ 311 \\ 511 \end{array} \right] \end{array} \quad \begin{array}{c} C_{III,I}^w \\ \parallel \\ \left[\begin{array}{c} -179 \\ -79 \\ 121 \\ 321 \\ -182 \\ -82 \\ 118 \\ 318 \\ -184 \\ -84 \\ 116 \\ 316 \\ -180 \\ -80 \\ 120 \\ 320 \\ -183 \\ -83 \\ 117 \\ 317 \\ -185 \\ -85 \\ 115 \\ 315 \\ -182 \\ -82 \\ 118 \\ 318 \\ -185 \\ -85 \\ 115 \\ 315 \end{array} \right] \end{array} \quad \begin{array}{c} C_{III,II}^w \\ \parallel \\ \left[\begin{array}{c} 19 \\ 119 \\ 319 \\ 519 \\ 15 \\ 115 \\ 315 \\ 515 \\ 12 \\ 112 \\ 312 \\ 512 \\ 18 \\ 118 \\ 318 \\ 518 \\ 14 \\ 114 \\ 314 \\ 514 \\ 11 \\ 111 \\ 311 \\ 511 \\ 15 \\ 115 \\ 315 \\ 515 \\ 11 \\ 111 \\ 311 \\ 511 \end{array} \right] \end{array} \quad (6.66)$$

Os coeficientes do vetor C_t é obtido usando a Eq. (6.3), e tomando as informações contidas na Tab. 6.8, tem-se:

$$C_t = \begin{bmatrix} 50 \\ 80 \\ 170 \\ 250 \end{bmatrix}. \quad (6.67)$$

Para cada cidade i tem-se que os coeficientes do vetor C_i^z são obtidos usando as Eq. (6.2) e Eq.(6.6), (ou a Eq.(6.7)). Então, tomando as informações contidas na Tab. 6.8, tem-se:

$$-C_I^z = \begin{bmatrix} 37 \\ 39 \\ 40 \\ 38 \\ 40 \\ 41 \\ 39 \\ 41 \\ 279 \\ 282 \\ 284 \\ 280 \\ 283 \\ 285 \\ 282 \\ 285 \end{bmatrix} \quad -C_{II}^z = \begin{bmatrix} 37 \\ 39 \\ 40 \\ 38 \\ 40 \\ 41 \\ 39 \\ 41 \\ 81 \\ 85 \\ 88 \\ 82 \\ 86 \\ 89 \\ 85 \\ 89 \end{bmatrix} \quad -C_{III}^z = \begin{bmatrix} 279 \\ 282 \\ 284 \\ 280 \\ 283 \\ 285 \\ 282 \\ 285 \\ 81 \\ 85 \\ 88 \\ 82 \\ 86 \\ 89 \\ 85 \\ 89 \end{bmatrix} \quad (6.68)$$

Para cada cidade i tem-se que os coeficientes do vetor C_i^x também são obtidos usando as Eq. (6.2) e Eq.(6.6), (ou a Eq.(6.7)). Novamente, tomando as informações contidas na Tab. 6.8, resulta:

$$-C_I^x = \begin{bmatrix} 32 \\ 34 \\ 35 \\ 273 \\ 276 \\ 278 \end{bmatrix} \quad -C_{II}^x = \begin{bmatrix} 32 \\ 34 \\ 35 \\ 74 \\ 78 \\ 81 \end{bmatrix} \quad -C_{III}^x = \begin{bmatrix} 273 \\ 276 \\ 278 \\ 74 \\ 78 \\ 81 \end{bmatrix} \quad (6.69)$$

Como as carretas paradas não geram custos, segue que o vetor C_y é o vetor nulo.

Minimizando a função objetivo, aplicando os códigos Lindo, Linprog e Mosek, obtém-se os seguintes valores da função objetivo, tempo gasto e número de iterações, para o problema proposto:

Tabela 6.9: Resultados obtidos para o Problema de Transporte II

	Função Objetivo (R\$)	Tempo Gasto (s)	Número de Interações
LINDO	71261	1.00	94
Linprog	63550	7.26	142
Mosek	63550	1.00	17

É importante ressaltar que as restrições laterais não foram usadas no código Lindo, visto que a versão disponível apresentava uma limitação sobre a quantidade de linhas (das restrições) que podem ser utilizadas. Além disso, mesmo que o programa esteja utilizando a opção de otimização inteira a solução obtida não foi inteira, sendo tomado o valor inteiro mais próximo do mesmo. Assim, escrevendo o vetor solução da seguinte forma:

$$X = \begin{bmatrix} X_{I,II}^w & X_{I,III}^w & X_{II,I}^w & X_{II,III}^w & X_{III,I}^w & X_{III,II}^w & X_t & -X_I^z & -X_{II}^z & -X_{III}^z & -X_I^x & -X_{II}^x & -X_{III}^x & X_y \end{bmatrix}^T; \quad (6.70)$$

tem-se os seguintes resultados: para cada um dos vetores $X_{i,j}^w$, $\forall i, j \in \{I, II, III\}$ e $i \neq j$,

todas as entradas que não são nulas, ao mesmo tempo nos três códigos, apresentados na Tab. 6.10.

Tabela 6.10: Soluções não nulas obtidas para a variável w (Problema II)

	LINDO	Linprog	Mosek
$w_{I,II}^{1,a,4}$	2	0	0
$w_{I,II}^{1,c,4}$	2	5	10
$w_{I,II}^{2,a,4}$	0	0	5
$w_{I,II}^{2,b,4}$	0	5	0
$w_{I,II}^{2,c,4}$	8	5	0
$w_{I,II}^{3,a,4}$	8	10	5
$w_{I,II}^{3,b,4}$	10	5	10
$w_{II,I}^{1,a,4}$	15	0	0
$w_{II,I}^{1,b,4}$	0	0	5
$w_{II,I}^{1,c,4}$	0	10	0
$w_{II,I}^{2,a,4}$	10	0	0
$w_{II,I}^{2,b,4}$	5	5	0
$w_{II,I}^{2,c,4}$	0	0	10
$w_{II,I}^{3,a,4}$	10	10	10
$w_{II,I}^{3,b,4}$	5	5	5
$w_{III,II}^{1,a,4}$	0	10	5
$w_{III,II}^{1,b,4}$	5	0	0
$w_{III,II}^{1,c,4}$	10	5	10
$w_{III,II}^{2,a,4}$	10	0	5
$w_{III,II}^{2,b,4}$	5	10	10
$w_{III,II}^{2,c,4}$	0	5	0

Pode-se observar uma certa simetria nos resultados obtidos nos três programas, visto que a quantidade de demandas a serem atendidas em cada uma das cidades são valores próximos, variando principalmente os tipos de carretas e cavalos usados para atender as demandas.

Com relação às variáveis t , os três códigos obtiveram o mesmo resultado, para as três primeiras variáveis. Para a quarta variável o Lindo apresentou um valor diferente dos apresentados pelos outros códigos. Estes resultados estão apresentado na Tab. 6.11.

Os três códigos consideraram inviável colocar cavalos, atrelados ou não a uma carreta vazia, viajando de uma cidade para outra, isto é, todos os valores dos vetores X_i^z e X_i^x , para $i \in \{1, 2, 3\}$, são nulos.

Tabela 6.11: Soluções obtidas para a variável t (Problema II)

	LINDO	Linprog	Mosek
t_1	40	40	40
t_2	40	40	40
t_3	40	40	40
t_4	7	10	10

Por fim, pode-se observar na Tab. 6.12, que os códigos LinProg e Mosek obtiveram resultados equivalentes em relação à distribuição das carretas que ficam paradas em cada pátio, mostrando uma boa concordância entre seus valores ótimos.

Tabela 6.12: Soluções obtidas para a variável y (Problema II)

	LINDO	Linprog	Mosek
y_I	12	10	5
	7	5	10
	7	0	0
y_{II}	0	5	10
	0	10	5
	0	0	0
y_{III}	0	0	0
	15	0	0
	0	15	15

A solução proposta pelo Lindo apresentou um valor da função objetivo bem maior do que os apresentados pelos outros códigos, porém ela não satisfaz todas as restrições, o que torna a solução não viável. Por exemplo, observe na Tab. 6.10 os valores obtidos pelo código para as variáveis $w_{II,I}^{1,a,4}$, $w_{II,I}^{2,a,4}$ e $w_{II,I}^{3,a,4}$. Tem-se que a soma destas coordenadas é 35 e o número de carretas disponíveis é 15, o que faz a restrição do número de carretas disponíveis ser violada.

Por outro lado, o valor da função objetivo obtidos pelos outros dois códigos foi o mesmo, mesmo tendo os vetores das variáveis de projeto distintos, sugerindo a existência de infinitas soluções. As soluções encontradas pelos códigos Linprog e Mosek não violaram nenhuma das restrições. É importante ressaltar que o problema foi formulado de tal forma que possa trabalhar com qualquer quantidade de variáveis.

Capítulo 7

Otimização de uma Rede de Distribuição de Água

A distribuição de água, através de redes complexas, pode gerar um custo muito elevado ou pode ser bastante ineficiente, gerando prejuízos econômicos se a rede não for bem planejada. Geralmente as redes são construídas de maneira que forme uma malha. Objetivando a otimização destas malhas, muitas técnicas para seu dimensionamento já foram apresentadas e utilizadas nos últimos anos. Para exemplificar a otimização em redes é considerado o problema apresentado em Curi e Firmino (2004) e que foi formulado na dissertação de mestrado de Formiga (1999).

Segundo Curi e Firmino (2004), alguns métodos tentam apenas dimensionar as redes baseadas nos requisitos mínimos de vazão e pressões para cada anel. A metodologia mais tradicional é a da tentativa e erro para obter a solução do problema, deixando a cargo da experiência do projetista a busca de um dimensionamento mais econômico. Para se levar em consideração os preços das tubulações em função do seu diâmetro e da classe de pressão, buscando estabelecer um custo mínimo para a rede, faz-se necessário o uso de técnicas de otimização. A aplicação de técnicas de programação linear ao problema de redes de distribuição pressurizadas foi introduzida por Karmelli et al. (1968) e o seu emprego vem sendo aperfeiçoado e utilizado por muitos especialistas. A grande limitação está na necessidade de linearizar o problema de otimização. A função objetivo é determinada pela variação do preço das tubulações em função do diâmetro e da classe de pressão.

7.1 Descrição do Problema

A rede utilizada para aplicar a programação linear foi retirada de Curi e Firmino (2004), conforme apresentada na Fig. 7.1.

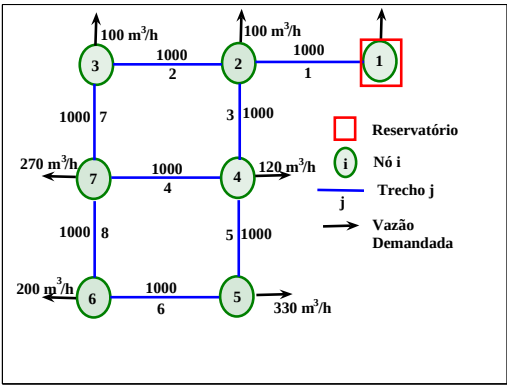


Figura 7.1: Descrição da Rede utilizada para estudo

Os dados relacionados com cada nó é dado pela Tab. 7.1 e os dados relacionados com os trechos da rede são dados pela Tab. 7.2.

Tabela 7.1: Dados relativos aos nós da rede			
Nó	Demanda (m^3/h)	Elevação (m)	Pressão Mínima (mca)
1 (Reservatório)	0	210	30
2	100	150	30
3	100	160	30
4	120	155	30
5	330	165	30
6	200	160	30
7	270	150	30

Fonte Curi e Firmino (2004) (adaptado).

Assim, tem-se que a rede é formada por oito trechos, distribuídos em dois anéis. Ela é abastecida por gravidade a partir de um reservatório que possui carga constante, a uma altura de 210m. Cada nó deve ter uma pressão mínima requerida igual a 30mca, sendo que todos os trechos possuem um comprimento igual a 1000m. Os dados referentes aos custos dos tubos são dados em unidades monetárias por metro (umt/m).

Tabela 7.2: Dados relativos aos nós da rede

Trecho	Nó Ini.	Nó Fim	Comprimento (<i>m</i>)	Diâm. Min. (<i>Pol</i>)	Diâm. Max. (<i>Pol</i>)
1	1	2	1000	1	24
2	2	3	1000	1	24
3	2	4	1000	1	24
4	4	7	1000	1	24
5	4	5	1000	1	24
6	5	6	1000	1	24
7	3	7	1000	1	24
8	6	7	1000	1	24

Fonte Curi e Firmino (2004) (adaptado).

A equação utilizada para o cálculo das perdas de cargas nos trechos foi a de Hazen-Williams. O coeficiente de perda de Hazen-Williams adotado como $C = 130$. Vários coeficientes “ ω ” diferentes são recomendados, pelos mais diversos autores, para a relação que calcula as perdas de cargas:

$$J_i = \omega \frac{1}{0,0254 D_j^{4,87} \left(\frac{Q_j}{C} \right)^{1,852}} L_j \quad (7.1)$$

onde J_j é a perda no trecho j (em *mca*), D_j é o diâmetro do tubo no trecho j (em *pol*), Q_j é a vazão no trecho j (em $\frac{m^3}{s}$), C é o coeficiente de Hazen-Williams e L_j é o comprimento do trecho j . Para este trabalho adotou-se $\omega = 10,5088$, conforme recomendado por Curi e Firmino (2004).

Em Curi e Firmino (2004) foi aplicado o método da programação não linear para descrever o comportamento físico do problema. Então, no primeiro momento as variáveis de decisão foram as vazões e os diâmetros, não nominais, da tubulação para cada trecho da rede. As restrições físicas foram baseadas nas leis de conservação da massa (continuidade) e da energia. Os aspectos operacionais, foram também levados em consideração nas restrições. A função objetivo foi dada pela minimização dos custos envolvidos na seleção dos diâmetros das tubulações. Os resultados obtidos nesta etapa por Curi e Firmino (2004) são mostrados na Tab. 7.3.

Neste trabalho é feito apenas a etapa de otimização utilizando a programação linear. O objetivo é obter os comprimentos ótimos da tubulação considerando dois diâmetros comerciais em

Tabela 7.3: Diâmetros comerciais utilizado na programação linear

Trecho	Vazão (m^3/s)	Obtido (Pol)	Comercial (Pol)	Trecho	Vazão (m^3/s)	Obtido (Pol)	Comercial (Pol)
1	0,3111	19,46	18 20	5	0,1472	14,39	14 16
2	0,1026	11,65	10 12	6	0,0555	9,91	8 10
3	0,1807	15,22	14 16	7	0,0748	10,04	10 12
4	0,0002	1,00	1 2	8	0,0001	1,00	1 2

Fonte Curi e Firmino (2004) (adaptado).

cada trecho, sendo escolhidos o diâmetro comercial imediatamente superior e imediatamente inferior àquele obtido na etapa não-linear, conforme apresentado na Tab. 7.3. Para isto, considera-se as vazões como sendo constantes e iguais àquelas obtidas na etapa não-linear do dimensionamento. O objetivo passa a ser minimizar o custo da rede mantendo-se as vazões e perdas de carga similares às obtidas na etapa não-linear do dimensionamento.

Para descrever o problema, tem-se que a função objetivo é definida pela Eq. (7.2). Neste caso as variáveis de decisão são os comprimentos de dois diâmetros de tubos para cada trecho, considerando as vazões nos trechos da rede são consideradas constantes, assim tem-se:

$$\begin{aligned}
 C(l_{jk}, Q_j) = & [l_{1,18}P(18) + l_{1,20}P(20)] + [l_{2,10}P(10) + l_{2,12}P(12)] + [l_{3,14}P(14) + \\
 & + l_{3,16}P(16)] + [l_{4,1}P(1) + l_{4,2}P(2)] + [l_{5,14}P(14) + l_{5,16}P(16)] + \\
 & + [l_{6,8}P(8) + l_{6,10}P(10)] + [l_{7,10}P(10) + l_{7,12}P(12)] + [l_{8,1}P(1) + l_{8,2}P(2)]
 \end{aligned} \quad (7.2)$$

onde $l_{j,k}$ é o comprimento do trecho j com diâmetro comercial imediatamente inferior, ou superior, ao diâmetro obtido na primeira etapa e $P(k)$ é o preço por unidade de comprimento desta tubulação. Os preços dos tubos por unidade de comprimento, para diâmetros comerciais, são apresentados na Tab. 7.5.

Para calcular cada um dos coeficientes $J_{i,k}$, usando a Eq. (7.1), tem-se que todos os coeficientes $J_{i,k}$ ficam dados pela Tab. 7.4.

Tabela 7.4: Coeficientes $J_{i,k}$ para o PPL

$J_{1,18}$	0.006648	$J_{5,14}$	0.005668
$J_{1,20}$	0.003979	$J_{5,16}$	0.002958
$J_{2,10}$	0.014674	$J_{6,8}$	0.014264
$J_{2,12}$	0.006039	$J_{6,10}$	0.004812
$J_{3,14}$	0.008341	$J_{7,10}$	0.000394
$J_{3,16}$	0.004353	$J_{7,12}$	0.003351
$J_{4,1}$	0.171493	$J_{8,1}$	0.010580
$J_{4,2}$	0.005864	$J_{8,2}$	0.000362

Tabela 7.5: Preço e velocidade máxima na tubulação segundo seus diâmetros

Diâmetro (<i>Pol</i>)	Custo (<i>umt/m</i>)	Velocidade (<i>m/s</i>)	Diâmetro (<i>Pol</i>)	Custo (<i>umt/m</i>)	Velocidade (<i>m/s</i>)
1	2	2	12	50	2, 1
2	5	2	14	60	2, 2
3	8	2	16	90	2, 3
4	11	2	18	130	2, 4
6	16	2	20	170	2, 5
8	23	2	22	300	2, 6
10	32	7	24	550	2, 7

As vazões para cada trecho, que foram obtidas na etapa não linear é apresentada na Tab.

7.3.

A Tab. 7.5 mostra os custos e velocidades máximas para as tubulações, considerando seus diâmetros comerciais.

As restrições do PPL são:

a) pressão mínima no nó tem que ser menor que a diferença entre a cota de cabeceira da rede e a pressão mínima requerida naquele nó. As restrições de conservação da energia garantem que, independente do caminho escolhido entre os trechos, a queda de pressão entre dois nós é sempre a mesma. Ao todo, são seis equações de restrições relacionadas com os requisitos mínimos de pressão, que são dadas nas Eq. (7.3) até Eq. (7.8). É importante ressaltar que o caminho entre os nós deve, de preferência, ser escolhido como sendo o mais curto, de modo a diminuir o tamanho da equação de restrição. Nestas expressões H_i representa a cota do terreno no nó i e pm_i é a pressão mínima requerida no nó i .

$$\text{Nó2} : J_1 < (H_1 - H_2) - pm_2, \quad (7.3)$$

$$\text{Nó3} : J_1 + J_2 < (H_1 - H_3) - pm_3, \quad (7.4)$$

$$\text{Nó4} : J_1 + J_3 < (H_1 - H_4) - pm_4, \quad (7.5)$$

$$\text{Nó5} : J_1 + J_3 + J_5 < (H_1 - H_5) - pm_5, \quad (7.6)$$

$$\text{Nó6} : J_1 + J_3 + J_5 + J_6 < (H_1 - H_6) - pm_6, \quad (7.7)$$

$$\text{Nó7} : J_1 + J_3 + J_4 < (H_1 - H_7) - pm_7, \quad (7.8)$$

onde H_i é a cota do terreno no nó i e pm_i é a pressão mínima requerida no nó i .

b) a conservação da energia nos Anéis, pois visto que não existem estações de bombeamento dentro da rede. Essa restrição deve garantir que a soma algébrica das perdas de carga dos trechos de um anel seja nula. A rede deste exemplo é composta por três anéis, que proporcionarão apenas duas equações de restrição linearmente independentes. Considera-se como positivas as perdas cujo sentido da vazão seja o mesmo que o sentido arbitrado para percurso no anel e, no caso contrário às perdas serão negativas. Os sentidos das vazões são àqueles apontados na Tab. 7.1, via definições de nós de origem e destino, e o percurso no anel é realizado no sentido horário. Dessa forma, tem-se que as restrições são escritas como:

$$\text{Anel1} : J_3 + J_4 - J_2 - J_7 = 0, \quad (7.9)$$

$$\text{Anel2} : J_5 + J_6 + J_8 - J_4 = 0. \quad (7.10)$$

c) o comprimento dos trechos, pois cada trecho possui tubulações com dois diâmetros diferentes, cujo comprimento é variável de decisão do problema. Portanto, a soma destes comprimentos, por trecho, deve ser igual ao comprimento do trecho, e este grupo de restrição é dada pela Eq. (7.11).

$$l_{j\text{dia-min}} + l_{j\text{dia-max}} = L_j, \quad \forall j = 1, \dots, 8 \quad (7.11)$$

d) a não negatividade das variáveis de decisão, pois os comprimentos de cada trecho são positivos.

Com as Eq. (7.9), Eq. (7.10) e Eq. (7.11) tem-se que a matriz de igualdade é dada por:

$$A_{eq} = \begin{bmatrix} 0 & 0 & -J_{2,10} & -J_{2,12} & J_{3,14} & J_{3,16} & J_{4,1} & J_{4,2} & 0 & 0 & 0 & 0 & -J_{7,10} & -J_{7,12} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -J_{4,1} & -J_{4,2} & J_{5,14} & J_{5,16} & J_{6,8} & J_{6,10} & 0 & 0 & J_{8,1} & J_{8,2} \\ J_{1,18} & J_{1,20} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & J_{2,10} & J_{2,12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & J_{4,14} & J_{4,16} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & J_{4,01} & J_{4,02} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & J_{5,14} & J_{5,16} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & J_{6,08} & J_{6,10} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & J_{4,10} & J_{4,12} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & J_{4,01} & J_{4,02} \end{bmatrix} \quad (7.12)$$

e

$$b_{eq} = \begin{bmatrix} 0 & 0 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 \end{bmatrix}^T. \quad (7.13)$$

A matriz de desigualdade, relacionada com Eq. (7.3) até Eq. (7.8), junto com as restrições dada pela Eq. (7.11), pode ser escrita como:

$$A_{de} = \begin{bmatrix} J_{1,18} & J_{1,20} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ J_{1,18} & J_{1,20} & J_{2,10} & J_{2,12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ J_{1,18} & J_{1,20} & 0 & 0 & J_{3,14} & J_{3,16} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ J_{1,18} & J_{1,20} & 0 & 0 & J_{3,14} & J_{3,16} & J_{4,1} & J_{4,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ J_{1,18} & J_{1,20} & 0 & 0 & J_{3,14} & J_{3,16} & 0 & 0 & J_{5,14} & J_{5,16} & 0 & 0 & 0 & 0 & 0 & 0 \\ J_{1,18} & J_{1,20} & 0 & 0 & J_{3,14} & J_{3,16} & 0 & 0 & J_{5,14} & J_{5,16} & J_{6,8} & J_{6,10} & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix} \quad (7.14)$$

Como a pressão mínima é constante, tem-se que $pm_i = 30, \forall i$. Assim, utilizando os dados da Tab. 7.1, chega-se ao vetor dos recursos disponíveis, que está relacionado com as restrições de desigualdade e é dado pela Eq. (7.15).

$$b_{de} = \begin{bmatrix} 30 & 20 & 25 & 15 & 20 & 30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T. \quad (7.15)$$

7.2 Solução Numérica do Problema de Distribuição de Água

Minimizando a função objetivo, aplicando os códigos Lindo, Linprog e Mosek, obtém-se os valores da função objetivo, tempo gasto e número de iterações, para o problema proposto. Estas estão apresentadas na Tab. 7.6. Também são apresentados os resultados obtidos por outros autores para o mesmo problema, que é citado em Curi e Firmino (2004).

Tabela 7.6: Resultados obtidos para a rede de distribuição de água.

	Função Objetivo (R\$)	Tempo Gasto (s)	Número de Interações
Lindo	406980	3.00	8
Linprog	406980	4.02	5
Mosek	406980	4.00	9
Eiger et al. (1994)	402352	-	-
Formiga (1999)	417500	-	-
Curi e Firmino (2004)	408300	-	-

Fonte Curi e Firmino (2004) (adaptado).

O valor da função objetivo obtido pelos três códigos é melhor do que todos os outros valores, exceto o obtido por Eiger et al. (1994). As variáveis de projeto obtidas por cada um dos códigos e os resultados obtidos pela literatura são dados pela Tab. 7.7.

Os valores das variáveis de projeto obtidos pelos três códigos são os mesmos, porém eles são distintos dos valores obtidos por outros autores. É importante ressaltar que neste problema o Linprog precisou de um número menor de iterações do que o Lindo e o Mosek para chegar à solução e que o tempo gasto foi praticamente o mesmo obtido pelos códigos. Não foi apresentado o tempo gasto e nem o número de iterações gastos em cada um dos outros autores, por isto não é possível fazer uma comparação.

Tabela 7.7: Resultados obtidos para a rede de distribuição de água.

<i>Lindo</i>	<i>LinProg</i>	<i>Mosek</i>	<i>Eiter</i>	<i>Formiga</i>	<i>Curi e Firmino</i>
1000	1000	1000	1000	457.2	797.4
0	0	0	0	508	202.6
846.9	846.9	846.9	238.02	254	0
153.1	153.1	153.1	761.98	304.8	1000
0	0	0	0	355.6	0
1000	1000	1000	1000	406.4	1000
21.3	21.3	21.3	1000	25.4	1000
978.7	978.7	978.7	0	50.8	0
384.1	384.1	384.1	628.86	355.6	600.2
615.9	615.9	615.9	371.14	406.4	399.8
19.9	19.9	19.9	989.05	203.2	22.6
980.1	980.1	980.1	10.95	254	977.4
1000	1000	1000	921.86	203.2	921.9
0	0	0	78.14	254	78.1
3.1	3.1	3.1	1000	25.4	1000
996.9	996.9	996.9	0	50.8	0

Fonte Curi e Firmino (2004) (adaptado).

A Tab. 7.8 apresenta nas dez primeiras linhas os valores relacionados com as restrições de igualdade do PPL de distribuição de água e as seis últimas estão relacionadas com as restrições de desigualdade.

Sobre as restrições de igualdade, com os dados da Tab. 7.8, pode-se observar que os resultados obtidos por outros autores, apresentados por Curi e Firmino (2004), não satisfizeram as restrições de igualdade relacionadas com as Eq. (7.9) e Eq. (7.10). Os resultados obtidos pelos três códigos são da ordem de 10^{-3} e podem ser aceitos como válidos. Com relação às restrições de desigualdade, novamente os resultados obtidos por Eiger et al. (1994) e Curi e Firmino (2004) não satisfizeram as restrições estabelecidas pelas Eq. (7.6) até Eq. (7.8). É importante ressaltar que os autores Curi e Firmino (2004) não citaram qual o coeficiente ω utilizado por Eiger et al. (1994), mas os utilizados pelos outros autores foi o mesmo $\omega = 10.5088$.

Desta forma, os resultados obtidos pelos códigos Lindo, LinProg e Mosek minimizaram o custo de instalação da rede de distribuição de água, obedecendo a todas as restrições impostas pelo projetista.

Tabela 7.8: Resultados obtidos para a rede de distribuição de água.

	<i>Lindo</i>			
	<i>LinProg</i>	<i>Eiter</i>	<i>Formiga</i>	<i>Curi e Firmino</i>
	<i>Mosek</i>			
<i>Eq.7.9</i>	-0.0011	167.1266	2.8898	169.1821
<i>Eq.7.10</i>	0.0008	-142.0903	2.9717	-151.3028
<i>Eq.7.11</i>	0	0	-34.8000	0
<i>Eq.7.11</i>	0	0	-441.2000	0
<i>Eq.7.11</i>	0	0	-238.0000	0
<i>Eq.7.11</i>	0	0	-923.8000	0
<i>Eq.7.11</i>	0	0	-238.0000	0
<i>Eq.7.11</i>	0	0	-542.8000	0
<i>Eq.7.11</i>	0	0	-542.8000	0
<i>Eq.7.11</i>	0	0	-923.8000	0
<i>Eq.7.3</i>	-23.3520	-23.3520	-24.9392	-23.8927
<i>Eq.7.4</i>	0	-5.2577	-9.3713	-7.8537
<i>Eq.7.5</i>	-13.9990	-13.9990	-15.2041	-14.5397
<i>Eq.7.6</i>	-9.6071	152.4940	-15.5503	151.9533
<i>Eq.7.7</i>	-0.0001	0.6632	-1.9864	0.0448
<i>Eq.7.8</i>	0	9.8237	-2.8657	0.0704

Capítulo 8

Conclusão

Neste estudo foi verificado que o interesse por métodos de otimização é crescente, visto que existe uma necessidade de cada vez mais otimizar processos em várias áreas da ciência. Dentre os métodos relacionados com a programação linear, observou-se que o método simplex, ou qualquer uma das suas variações, merece destaque, considerando que sua eficiência e simplicidade de implementação é notória, o que torna o procedimento de obtenção de soluções ótimas prático e de baixo custo.

Nesta pesquisa foi realizada uma ampla revisão dos conceitos fundamentais do método simplex. Foram apresentados os principais conceitos e teoremas necessários para a compreensão do assunto. Além disso, algumas das variações do método foram igualmente apresentadas de modo a tornar o trabalho completo. Uma contribuição deste trabalho foi a produção de um texto didático amplo, para que possa ser útil para outros pesquisadores.

Procurou-se produzir um texto que tratasse o tema com uma sólida fundamentação teórica, mas escrito usando uma linguagem acessível, procurando, desta forma, contribuir com pesquisadores da área.

Outra contribuição deste estudo foi comparar a performance dos códigos computacionais Lindo, Linprog e Mosek. Com esta finalidade, duas aplicações em engenharia foram consideradas: um problema de otimização do transporte de cargas rodoviárias e o problema de otimizar uma rede de distribuição de água. Os códigos escolhidos se mostraram bem eficientes chegando ao mesmo valor da função objetivo, tendo uma pequena variação no Problema II de transporte, visto

que existia um limite de restrições a serem utilizadas pelo código Lindo. A maior diferença que foi possível observar em cada um dos códigos foi o tempo computacional gasto e o número de iterações que cada um precisou para obter a solução ótima, sendo observado que o Mosek se mostrou o mais eficiente em cada um dos aspectos observado. Outra vantagem apresentada pelos códigos Mosek e Linprog é a possibilidade de se trabalhar com as matrizes dos coeficientes, sendo que o mesmo não é possível com o Lindo, tornando o trabalho com o código menos apropriado para problemas com um número elevado de variáveis de projeto e de restrições.

Dentre os problemas apresentados, pode-se observar no PPL de Transporte o quanto é complexo um problema do mundo real, onde o número de restrições e o número de variáveis de projeto podem ser extremamente grande, podendo chegar a centenas de milhares. Por isto, a análise dos códigos através de problemas menores, fornece condições para que sejam feitos ajustes e aprimoramentos nos códigos, tornando possível obter conclusões seguras a partir dos resultados obtidos para problemas maiores.

Os estudos atualmente desenvolvidos permitem antever alguns desdobramentos futuros deste trabalho, que são abaixo relacionados:

- utilização de outros códigos computacionais, de forma a comparar a performance em problemas complexos, com elevado número de variáveis de projeto;
- estudo de novos métodos de programação linear, entre eles, o método da penalidade interior;
- modelar novos problemas, inclusive considerar o estudo de complexos problemas na área médica;
- estudo mais completo do problema de ciclagem em programação linear;
- utilização da metodologia híbrida de otimização, aplicando técnicas de otimização não-linear em conjunto com a programação linear.

Apêndice A

Tabelas de Símbolos

Símbolo	Referência
S	Subconjunto do $\mathbb{R}^n$
a	constante real
x	algum elemento do $\mathbb{R}^n$
$\subseteq$	Subconjunto de um conjunto qualquer
$\subset$	Subconjunto próprio de um conjunto qualquer
$\mathcal{A}$	Representação de uma matriz qualquer
0	vetor nulo
a_i	i -ésima linha de uma matriz
$\mathcal{A}_j$	j -ésima coluna de uma matriz
LI	Linearmente Independente
LD	Linearmente Dependente
PPL	Problema de Programação Linear
PPLI	Problema de Programação Linear Inteira
MML	Método dos Multiplicadores de Lagrange

Referências Bibliográficas

- Assis, E. G. d. (1993). Otimização multi-critérios de máquinas rotativas. Dissertação de mestrado, Universidade Federal de Uberlândia.
- Barboza, C. B. e Oliveira, A. R. L. (2006). Planejamento do tratamento por radioterapia através de métodos de pontos interiores. *Pesquisa Operacional*, 26(1).
- Bertsimas, D. e Tsitsiklis, J. N. (1997). *Introduction to Linear Optimization*. Athena Scientific.
- Braga, C. G. (1998). O uso de algoritmos genéticos para aplicação em problemas de otimização de sistemas mecânicos. Dissertação de mestrado, Universidade Federal de Uberlândia.
- Bregalda, P. F. e de Oliveira, A. A. F. e. B. C. T. (1988). *Introdução à Programação Linear*. Editora Campus, 3a edition.
- Choze, S. B. (1998). Otimização de sistemas mecânicos: Ferramenta de engenharia de concepção. Dissertação de mestrado, Universidade Federal de Uberlândia.
- Coelho, F. U. e Lourenço, M. L. (2005). *Um Curso de Álgebra Linear*. EduSP, 2a edition.
- Costa, F. P. e Souza, M. J. F. e Pinto, L. R. (2005). Um modelo de programação matemática para alocação estática de caminhões visando ao atendimento de metas de produção e qualidade. *Revista Escola de Minas*, 58(1).
- Couto, P. T. B. (2004). Resolução de problemas de transporte rodoviário de cargas utilizando programação inteira. Dissertação de mestrado, Pontifícia Universidade Católica do Rio de Janeiro.

- Curi, W. F. e Firmino, M. B. M. (2004). Um método de dimensionamento ótimo de redes pressurizadas malhadas via programações não lineares e linear. *UFPB*.
- Eiger, G., Shamir, U., e Bem-Tal, A. (1994). Optimal design of water distribution systems. *Water Resources Research*, 30(9):2937–2946.
- Faria, M. L. M. (1991). Uma contribuição aos procedimentos de otimização aplicados a sistemas mecânicos. Dissertação de mestrado, Universidade Federal de Uberlândia.
- Formiga, K. T. M. (1999). Metodologia de otimização de redes malhadas através da programação não linear. Dissertação de mestrado, UFPB, CCT, Departamento de Engenharia Civil.
- Guidorizzi, H. L. (2006). *Um curso de Cálculo (4 Volumes)*, volume 2. LTC, 5a edition.
- Hoffman, K. e Kunze, R. (1979). *Álgebra Linear*. LTC, 2a edition.
- Iezzi, G. e Hazzan, S. (2001). *Fundamentos da Matemática Elementar*, volume 04. 6a edition.
- Karmelli, D., Gadish, Y., e Meyers, S. (1968). Design of optimal water distribution networks. *Journal of Pipeline Division*, 94(10):1–10.
- MathWorks (2006). *Optimization Toolbox 3, User's Guide*. The MathWorks, Inc.
- Melo, M. P., Lima, D. P., e Pinheiro, P. R. (2004). Modelos em programação matemática para o processamento do biscoito tipo cracker. *Ciência e Tecnologia de Alimentos*, 24(03).
- Mosek (2006). *The MOSEK optimization toolbox for MATLAB manual. Version 4.0 (Revision 60)*. MOSEK ApS, Denmark.
- Oliveira, G. T. (2006). Estudo e aplicações da evolução diferencial. Dissertação de mestrado, Universidade Federal de Uberlândia.
- Oliveira, L. S. d. (2005). Uma contibuição ao estudo dos métodos de otimização multi-objetivo. Dissertação de mestrado, Universidade Federal de Uberlândia.
- Pereira, C. B. (2003). Técnicas de otimização discreto-contínuas aplicadas ao controle de vibrações em estruturas inteligentes. Dissertação de mestrado, Universidade Federal de Uberlândia.

- Prado, D. (1999). *Programação Linear*, volume 1. Editora de Desenvolvimento Gerencial, 2a edition.
- Santana, D. C. (2007). Modelagem numérica e otimização de shunts piezelétricos aplicados ao controle passivo de vibrações. Tese de doutorado, Universidade Federal de Uberlândia.
- Saramago, S. d. F. P. (1998). Otimização de trajetórias de robôs manipuladores na presença de obstáculos. Tese de doutorado, Universidade Federal de Uberlândia.
- Sousa, J. A. M. d. (2003). *Uma Breve História da Programação Linear - Notas de Aula*. Instituto Superior de Engenharia de Lisboa - Notas de Aula.
- Vanderplaats, G. N. (1999). *Numerical Optimization Techniques For Engineering Design*. Vanderplaats Research & Development, Inc., 3a edition.
- Zörnig, P. (2006a). A note on cycling examples with permutation structure. *UNB, Departamento de Estatística*.
- Zörnig, P. (2006b). Systematic construction of examples for cycling in the simplex method. *Computer & Operations Research*, (33):2247–2262.