unesp

UNIVERSIDADE ESTADUAL PAULISTA

PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Programação Linear e as Condições de Karush-Kuhn-Tucker

Ari Gomes da Mota Filho

INSTITUTO DE GEOCIÊNCIAS E CIÊNCIAS EXATAS

RIO CLARO
2021

Universidade Estadual Paulista "Júlio de Mesquita Filho"
Instituto de Geociências e Ciências Exatas
Câmpus de Rio Claro

Programação Linear e as Condições de Karush-Kuhn-Tucker

Ari Gomes da Mota Filho

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Pós-Graduação em Matemática, mestrado profissional, do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de Rio Claro.

Orientador
Prof. Dr. Wladimir Seixas

Rio Claro
2021

M917p

Mota Filho, Ari Gomes da
Programação Linear e as Condições de Karush-Kuhn-Tucker / Ari Gomes da Mota Filho. -- Rio Claro, 2021
71 p.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Instituto de Geociências e Ciências Exatas, Rio Claro
Orientador: Wladimir Seixas

1. Métodos Matemáticos. 2. Otimização. 3. Programação Linear. 4. Método Simplex. 5. Programação Não Linear. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Geociências e Ciências Exatas, Rio Claro. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

TERMO DE APROVAÇÃO

Ari Gomes da Mota Filho

Programação Linear e as Condições de Karush-Kuhn-Tucker

Dissertação APROVADA como requisito parcial para a obtenção do grau de Mestre no Curso de Pós-Graduação em Matemática, mestrado profissional, do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", pela seguinte banca examinadora:

Prof. Dr. Wladimir Seixas
Orientador

Prof. Dr. Luis Antonio da Silva Vasconcellos
Departamento de Matemática - Faculdade de Ciências
Universidade Estadual Paulista Júlio de Mesquita Filho

Profa. Dra. Marta Cilene Gadotti
Departamento de Matemática - IGCE
Universidade Estadual Paulista Júlio de Mesquita Filho

Rio Claro, 26 de Janeiro de 2021

Dedico este trabalho a Deus, a minha família e a todas as forças universais que conspiram para que o melhor sempre aconteça.

Agradecimentos

A Deus por ter me concedido esta oportunidade em minha vida. A minha esposa pelo amor, companheirismo, incentivo e apoio. A minha filha que está a caminho com muito amor neste momento. A minha mãe e irmãs que me dão suporte nesta trajetória. Aos amigos que me ensinam bastante e aos professores da UNESP-RC que me acompanharam durante o mestrado, em especial aos professores Wladimir Seixas, Marta Cilene Gadotti e Carina Alves Severo, por todo incentivo, paciência e confiança que tiveram comigo. Agradeço a todos que participaram de alguma forma desta conquista.

Os hábitos angulares dizem que o sucesso não depende de acertar cada mínimo detalhe, mas, em vez disso, baseia-se em identificar umas poucas prioridades centrais e transformá-las em poderosas alavancas.

Charles Duhigg

Resumo

Esta dissertação se propõe a organizar, discutir e redigir de maneira precisa e rigorosa as ideias matemáticas envolvidas na Programação Linear, em especial o Método Simplex, e na Programação Não Linear através do estudo de Máximos e Mínimos, Multiplicadores de Lagrange e das Condições Ótimas de Karush-Kuhn-Tucker.

Palavras-chave: Métodos Matemáticos, Otimização, Programação Linear, Método Simplex, Programação Não Linear.

Abstract

The aim of this work is to organize, discuss and write the mathematical ideas involved in Linear Programming, in particular the Simplex Method, and Nonlinear Programming through of study of Maxima and Minima, Lagrange Multipliers and Karush-Kuhn-Tucker optimality conditions.

Keywords: Mathematical Methods, Optimization, Linear Programming, Simplex Method, Nonlinear Programming.

Lista de Figuras

1.1 Fluxograma do processo de modelagem. 24
1.2 Representação geométrica para o exemplo 1.2. 28
1.3 Regiões delimitadas por inequações do exemplo 1.2. 29
1.4 Curvas de nível exemplo 1.2 30
1.5 Gradiente normal à curva de nível 31
1.6 Múltiplas soluções 32
1.7 Exemplo de solução ilimitada 32
1.8 Exemplo de solução infactível 33

2.1 Fluxograma do Método Simplex 45
2.2 Maior círculo inscrito em um polígono 46
2.3 Círculo limitado por retas 46
2.4 Reta normal e ponto de interseção 47
2.5 Região limitada pelas retas 49
2.6 Gráfico solução - Exemplo 2.5 54

3.1 Máximo e mínimo - Exemplo 3.11. 60

Lista de Tabelas

1.1 Valores dos pontos extremos aplicados na função objetivo 29

2.1 Quadro inicial geral para um problema de Programação Linear 38
2.2 Quadro inicial para o exemplo 1.2 39
2.3 Segundo quadro para o exemplo 1.2 42
2.4 Quadro inicial para o exemplo 1.2 43
2.5 Segundo quadro para o exemplo 1.2 43
2.6 Quadro final para o exemplo 1.2 44
2.7 Quadro inicial para o exemplo 2.5 50
2.8 Iteração 1 51
2.9 Iteração 2 52
2.10 Iteração 3 52
2.11 Iteração 4 53
2.12 Quadro Final - Exemplo 2.5 53

Sumário

Introdução **21**

1 Propriedades da Programação Linear **23**
1.1 Modelagem Matemática na Pesquisa Operacional 23
1.2 Um problema padrão de Programação Linear 25
1.2.1 Exemplos e o Método Gráfico 27
1.2.2 Conjuntos Convexos 33

2 Método Simplex **37**
2.1 Adicionando variáveis de folga 37
2.2 Selecionando a variável de entrada 40
2.3 Escolhendo a variável de saída 41
2.4 O Método Simplex 44
2.5 Exemplo: maior círculo inscrito em um polígono 45

3 Programação Não Linear e as Condições de Karush-Kuhn-Tucker **55**
3.1 Noções preliminares 55
3.2 Máximos e Mínimos 57
3.2.1 Condições necessárias para que um ponto interior seja um extremante local 57
3.2.2 Condição suficiente para um ponto crítico ser extremante local . . . 58
3.3 Método dos Multiplicadores de Lagrange 58
3.4 Condições Ótimas de Karush-Kuhn-Tucker 60

Considerações finais **69**

Referências **71**

Introdução

Segundo Arenales et. al. [1], a Pesquisa Operacional surgiu na Inglaterra, em momento pré Segunda Guerra Mundial (1936) e foi um termo empregado à invenção do radar aéreo. O intuito de estudar as tecnologias relacionadas ao radar aéreo visava a interceptação de aviões inimigos. Em 1941, houve a inauguração da Seção de Pesquisa Operacional do Comando da Força Aérea de Combate, que tinha como objetivo solucionar problemas de operações em guerra, como inspeções de aviões, escolha do tipo de avião para uma missão, melhoria na probabilidade de destruições de submarinos, além de controle de artilharia antiaérea e dimensionamento de comboios de frota.

Após a Segunda Guerra Mundial a Pesquisa Operacional evoluiu rapidamente e deu origem ao projeto SCOOP (Scientific Computation of Optimal Programs) no Pentágono. A ideia consistia em apoiar as decisões de operações da força aérea americana. Neste momento surgem os primeiros trabalhos precursores para resoluções de problemas em otimização linear.

Em 1952, foi fundada a Sociedade Científica Americana de Pesquisa Operacional e em 1953 a Sociedade Inglesa de Pesquisa Operacional. Em 1957 ocorreu a primeira conferência internacional em Pesquisa Operacional, em Oxford, na Inglaterra. Neste momento foi observado os diferentes focos dos trabalhos desenvolvidos pelos ingleses e americanos. A ênfase dos trabalhos desenvolvidos pelos ingleses eram estudos de caso ou problema específicos, enquanto os trabalhos desenvolvidos pelos americanos abordavam modelos ou métodos matemáticos em diversos temas como teoria dos estoques, substituições de equipamentos, teoria das filas, programação de tarefas em máquinas, teoria dos jogos, fluxo em redes e otimização linear.

Entre o início de 1950 e o fim de 1960, a Pesquisa Operacional foi aplicada em uma variedade de problemas oriundos dos setores públicos e privados. Em um levantamento feito em 1953 com 160 organizações da Grã-Bretanha, 45 responderam que possuíam um departamento de Pesquisa Operacional ou pelo menos uma pessoa nessa atividade.

A Pesquisa Operacional no Brasil teve início em 1960. O primeiro simpósio brasileiro de Pesquisa Operacional foi realizado em 1968 no ITA, em São José dos Campos/SP. Em seguida (1969) foi fundada a Sociedade Brasileira de Pesquisa Operacional (SOBRAPO).

Após este breve relato sobre o contexto em que surgiram os primeiros trabalhos sobre Pesquisa Operacional, podemos concluir que é um tema de bastante relevância no cenário mundial e é até hoje um tema presente em diversas áreas. A busca pela otimização de recursos é um tema importante, pois vários recursos no planeta são limitados.

O intuito deste trabalho é organizar, discutir e redigir as ideias matemáticas envolvidas na Programação Linear e Não Linear, relacionando com os conteúdos iniciais da matemática do Ensino Superior. Este trabalho está dividido em três capítulos, as considerações finais e as referências.

No primeiro capítulo trataremos das Propriedades da Programação Linear, iniciando com a modelagem de problemas de otimização, em seguida a descrição de um problema padrão da Programação Linear, exemplos e o método gráfico. No segundo capítulo, abordaremos o Método Simplex, que é uma técnica utilizada para determinar uma solução ótima de um modelo de Programação Linear; e no terceiro capítulo a Programação Não Linear, onde iremos discorrer sobre Máximos e Mínimos, Multiplicadores de Lagrange e as Condições Ótimas de Karush-Kuhn-Tucker (KKT).

As condições KKT generalizam o Método de Multiplicadores de Lagrange e permitem as restrições com desigualdades. Elas são condições de primeira ordem para que uma solução de um problema de programação não linear seja ótima, desde que, sejam válidas as condições de qualificação.

1 Propriedades da Programação Linear

A Programação Linear é uma parte da Matemática que tem aplicações em diversas áreas do conhecimento. Ela é considerada uma ferramenta essencial para ciência da Administração e da Pesquisa Operacional. O objetivo da Programação Linear é a otimização através da maximização ou minimização. A ideia consiste, inicialmente, em formular modelos matemáticos e descrever um método computacional de solução através do estudo de equações lineares com desigualdades. Sua origem data da Segunda Guerra Mundial (1939-1945) e teve como objetivo a solução de problemas de deslocamento, logística e manutenção dos grandes exércitos.

Em 1947, o matemático George Dantizig (1914-2005) desenvolveu um dos métodos para resolução de problemas envolvendo otimização linear, o Método Simplex. Este método é considerado até hoje uma técnica fundamental para a Pesquisa Operacional.

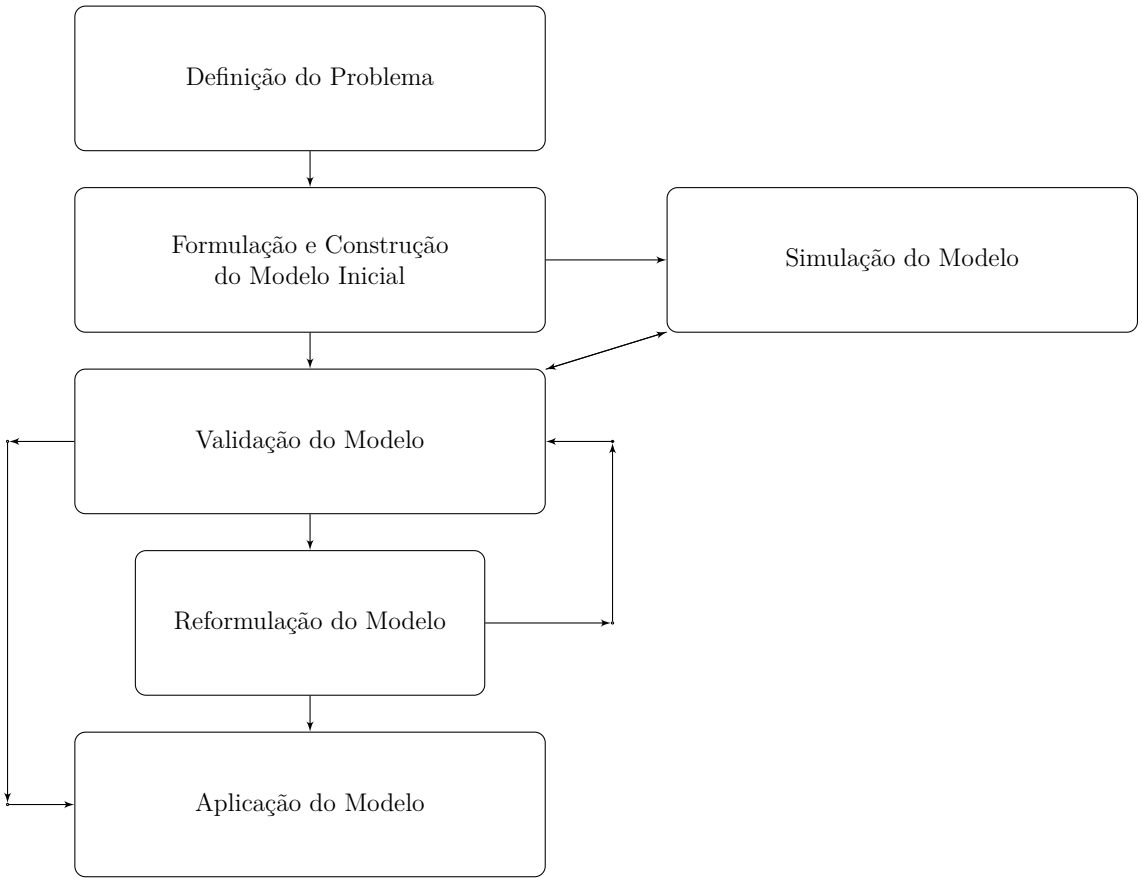
Neste capítulo serão demonstrados alguns teoremas que garantem que a solução ótima está em um extremo definido pelas interseções das retrições. Utilizaremos um exemplo da indústria para ilustrar cada etapa do processo: modelagem, formulação matemática e solução. As principais referências utilizadas neste capítulo são Goldbarg et. al. [2], Kolman [3] e Pedregal [4].

1.1 Modelagem Matemática na Pesquisa Operacional

A modelagem é um processo que alia teoria e prática e que motiva o usuário na busca de um entendimento da realidade que o cerca. Além disso, esta teoria busca explicar processos considerando parâmetros formais através de um sistema artificial que é chamado de modelo ou modelagem. O termo modelagem pode ser entendido de diversas maneiras e a ênfase nesta seção será a modelagem matemática na Pesquisa Operacional.

Segundo Goldbarg et al.(2015, p.13), é possível resumir o Processo de Modelagem ou de construção de modelos na ótica operacional através dos passos sugeridos pelo fluxograma apresentado na figura 1.1 a seguir.

Figura 1.1: Fluxograma do processo de modelagem.



Fonte: Goldbarg et. al. (2015, p.13)

Conforme observamos no fluxograma, uma série de etapas devem ser cumpridas na construção de um modelo na Pesquisa Operacional.

De acordo com Goldbarg et. al. (2015, p.16), podemos destacar que

> "[...] Os modelos [da Pesquisa Operacional –] PO são estruturados de forma lógica e formal, objetivando a otimização do funcionamento dos sistemas representados. Os modelos de Programação Matemática estão entre os principais da Pesquisa Operacional e constituem uma das mais importantes variedades dos modelos quantitativos.".

Cada etapa descrita no fluxograma da figura 1.1 é definida da seguinte forma:

Definição do Problema: A definição do problema compreende a identificação do problema colocado. Ele deve ser traduzido através do objetivo, das variáveis de decisão e níveis de detalhes (restrições).

Formulação e Construção do Modelo Inicial: A formulação e Construção do Modelo Inicial é o momento em que definimos os tipos de variáveis, assim como o nível apropriado de importância de cada variável. É quando formalizamos e surgem as restrições do problema e a função a ser otimizada.

Simulação do Modelo: Nesta etapa verificamos as situações de incerteza ou complexidade do sistema. Na simulação do Modelo podemos compreender melhor as variáveis do sistema e contornar possíveis dificuldades.

Validação do Modelo: É o processo onde se verifica a representatividade do modelo, é uma etapa indispensável em qualquer procedimento científico.

Reformulação do Modelo: No processo iterativo há necessidade, algumas vezes, de um ajuste entre o modelo e a realidade, sendo assim necessário a reformulação do modelo.

Aplicação do modelo: Concluído o modelo, é necessário o teste do problema do real, então, é quando o modelo e o sistema representado são confrontados.

1.2 Um problema padrão de Programação Linear

De acordo com Pedregal (2004), a melhor maneira de tratar sobre otimização é através de situações práticas que podem ser resolvidas através de técnicas conhecidas, isto é, mostrar na prática as ferramentas básicas para resolução de problemas de Programação Linear.

Pedregal (2004) cita como exemplo o problema da dieta. Este problema consiste em mensurar o conteúdo nutritivo de certos alimentos bem como seus preços, visando o mínimo diário necessário de cada nutriente e a busca pelo menor custo total possível de uma determinada dieta.

No decorrer desta seção trataremos de forma genérica um problema padrão de Programação Linear com intuito de familiarizar o tratamento algébrico.

Antes de iniciar as definições de um modelo pradrão na Programação Linear é importante observarmos que sua principal característica é que todas as funções envolvidas (a função objetivo e as que expressam as restrições) devem ser lineares. O surgimento de

uma única função não linear, seja na função objetivo ou nas restrições é suficiente para não tratar o problema como um problema de Programação Linear.

Um problema padrão na Programação Linear pode ser definido da seguinte forma:

Definição 1.1.

$$\text{Maximize } z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

sujeito a

$$\begin{cases} a_{11}x_1 + & a_{12}x_2 + & \cdots + & a_{1n}x_n \leqslant & b_1 \\ a_{21}x_1 + & a_{22}x_2 + & \cdots + & a_{2n}x_n \leqslant & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 + & a_{m2}x_2 + & \cdots + & a_{mn}x_n \leqslant & b_m \end{cases}$$

com $x_j \geqslant 0$ para $1 \leqslant j \leqslant n$ e a_{ij}, b_j e c_i para $1 \leqslant i \leqslant n$ e $1 \leqslant j \leqslant m$ são dados do problema.

Quando tratamos de sistemas de inequações/equações podemos fazer uso da notação matricial. Sendo assim, nosso problema padrão de Programação Linear poderá ser escrito da seguinte forma: Determine o vetor x em $\mathbb{R}^n$ que irá maximizar a função

$$z = c^T \cdot x \tag{1.1}$$

sujeito a

$$Ax \leqslant b \tag{1.2}$$

$$x \geqslant 0 \tag{1.3}$$

onde

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \; x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \; b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \; c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Observamos que dados os vetores u, $v \in \mathbb{R}^n$ com $u = (u_1, \ldots, u_n)^T$ e $v = (v_1, \ldots, v_n)^T$ definimos

$$u \leqslant v \Leftrightarrow u_i \leqslant v_i, \text{ para todo } 1 \leqslant i \leqslant n.$$

Um vetor x em $\mathbb{R}^n$ que satisfaz (1.2) e (1.3) é chamado de uma **solução viável** do problema dado. A solução viável que maximiza a função objetivo (1.1) é chamada de **solução ótima**.

Quando tratamos de soluções de problemas em otimização, devemos ter bem claro nossa função objetivo e as suas retrições. Além disso, devemos observar se o problema atende aos critérios ou se há necessidade de adequações.

É importante ressaltar que podemos também definir um problema de minimização como um problema maximização, para isso devemos considerar que o mínimo de $c_1x_1 + c_2x_2 + \cdots + c_nx_n$ é

$$\max\{cx : Ax \leqslant b, x \geqslant 0\} = -\min\{(-c)x : Ax \leqslant b, x \geqslant 0\}.$$

Além disso, é possível inverter algumas desigualdades (restrições), quando necessário, com intuito de transformá-lo em um problema padrão conforme a definição 1.1, restrições do tipo $\leqslant$.

Um problema pode ser escrito de diferentes maneiras e é importante analisar a forma mais eficiente sobre como as informações são declaradas e analisar a relação entre as variáveis e as restrições.

A formulação dos problemas nos ajuda a entender a conexão entre o problema original e a sua tradução em equações, desigualdades, igualdades, etc. Esse processo é a primeira etapa da modelagem na Pesquisa Operacional, sendo ela a definição do problema. É o momento que determinamos a função objetivo e as retrições ao qual o problema está limitado.

Por fim, após verificarmos se o problema de Programação Linear possui o formato de um problema padrão poderemos seguir para próxima etapa.

1.2.1 Exemplos e o Método Gráfico

Nesta subseção abordaremos um exemplo aplicado à indústria. Iremos com isso, ilustrar cada um dos pontos da teoria tratados até o momento.

Exemplo 1.2. Um fabricante de produtos de limpeza prepara dois tipos de polidores de metais por dia, tipo 1 e tipo 2, usando como matéria-prima as soluções A e B. Sabemos que cada litro do polidor tipo 1 contém duas medidas da solução A e uma medida da solução B, enquanto que o polidor tipo 2 contém uma medida da solução A e duas medidas da solução B. Suponhamos também que o lucro em cada litro do polidor tipo 1 é 8 reais, enquanto o lucro em cada litro do polidor tipo 2 é de 10 reais. Sabemos que a indústria tem disponíveis 50 medidas da solução A e 70 da solução B por dia. Supondo que é vendido tudo que se produz pergunta-se: quantos litros o fabricante deve produzir de cada tipo de polidor por dia para maximizar o lucro?

Inicialmente temos a definição das variáveis do problema. Sejam x e y o número de litros dos polidores do tipo 1 e 2 produzido por dia respectivamente.

Em seguida, temos as restrições do problema. Sabemos que cada litro do polidor tipo 1 contém duas medidas da solução A e cada litro do polidor tipo 2 contém uma medida da solução A. Assim, a quantidade total da solução A necessária é, $2x + y$.

Analogamente, como cada litro de polidor tipo 1 contém uma medida da solução B e cada litro do polidor tipo 2 contém duas medidas da solução B. A quantidade total da solução B necessária é $x + 2y$.

Por outro lado, temos apenas 50 medidas da solução A e 70 medidas da solução B. Logo, obtemos as seguintes restrições para o problema, $2x + y \leqslant 50$ e $x + 2y \leqslant 70$.

Sabemos ainda que o lucro em cada litro do polidor tipo 1 e do polidor tipo 2 é de 8 e 10 reais, respectivamente. O lucro total em reais será de $z = 8x + 10y$.

O problema de Programação Linear será dado por: Determinar os valores de x e y que maximizam

$$z = 8x + 10y \tag{1.4}$$

Sujeitos as seguintes retrições:

$$\begin{cases} 2x + y \leqslant 50 \\ x + 2y \leqslant 70 \\ x \geqslant 0 \\ y \geqslant 0 \end{cases} \tag{1.5}$$

Em notação matricial: Determinar um vetor $x = \begin{bmatrix} x \\ y \end{bmatrix}$ em $\mathbb{R}^2$ que irá maximizar $z = \begin{bmatrix} 8 & 10 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$ sujeito a

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \leqslant \begin{bmatrix} 50 \\ 70 \end{bmatrix}$$

com

$$\begin{bmatrix} x \\ y \end{bmatrix} \geqslant \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

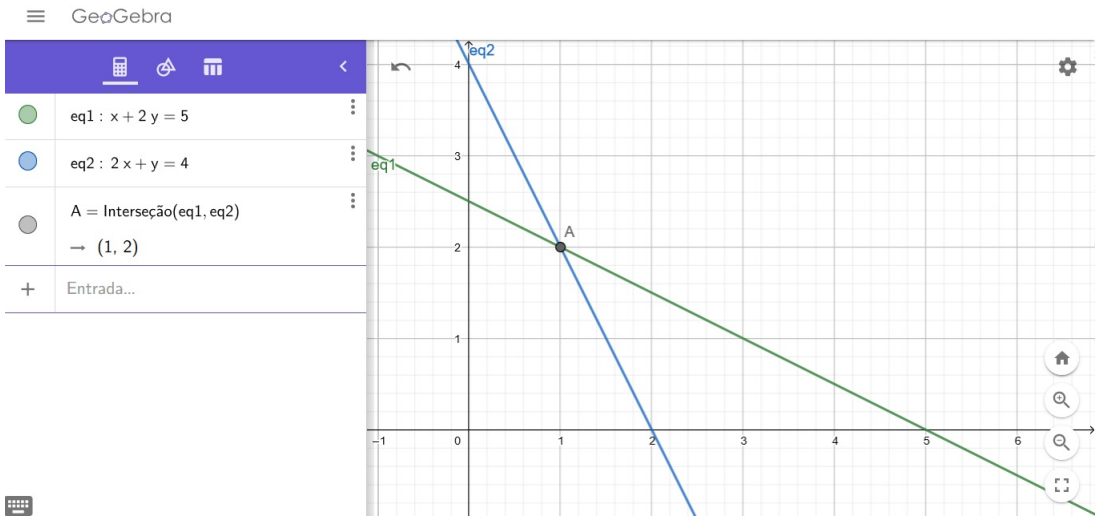
Quando tratamos de sistema com duas variáveis de decisão podemos facilmente identificar um conjunto de candidatos a solução ótima de maneira geométrica. Para isso, representamos o problema de Programação Linear como regiões no plano cartesiano. A resolução geométrica não deve ser vista como uma ferramenta prática, pois é adequada somente para problemas que envolvem duas ou três variáveis. No entanto, apesar da não praticidade, podemos utilizar este tipo de resolução em aplicações no nível do Ensino Médio para auxiliar na compreensão dos estudos sobre sistemas de equações/inequações e suas soluções.

Consideremos o sistema de equações:

$$\begin{cases} x + 2y = 5 \\ 2x + y = 4 \end{cases} \tag{1.6}$$

Ao representar geometricamente as equações (1.6) no plano xy obtemos a figura 1.2 a seguir.

Figura 1.2: Representação geométrica para o exemplo 1.2.



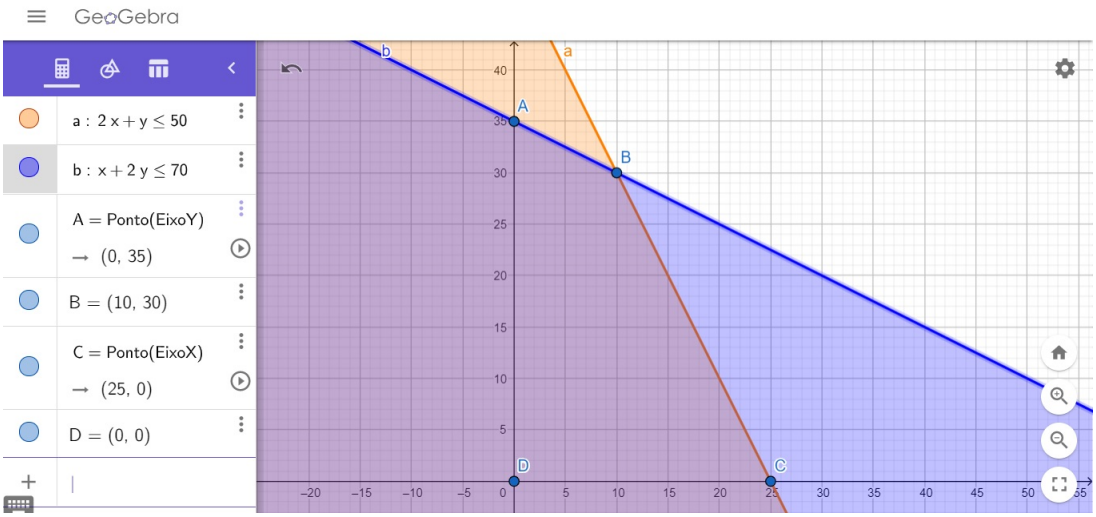
Fonte: Elaborado pelo autor.

Verificamos que a partir da representação geométrica das equações, o conjunto solução para o sistema de equações (1.6) é dado pela interseção de duas retas no plano xy. O ponto de interseção A na figura 1.2 tem coordenadas $x = 1$ e $y = 2$.

Apropriando-se da representação geométrica e retomando ao exemplo 1.2, através das inequações (1.4) e (1.5), podemos dizer que o conjunto de pontos pertencentes a região

limitada pelas retas a e b da figura 1.2 (onde $x \geqslant 0$ e $y \geqslant 0$) representam possíveis candidatos que podem maximizar a função objetivo $z = 8x + 10y$, conforme mostra a figura 1.3 a seguir.

Figura 1.3: Regiões delimitadas por inequações do exemplo 1.2.



Fonte: Elaborado pelo autor.

Ao tabular os valores dos pares (x, y) para os pontos extremos A, B, C e D da figura 1.3 na função objetivo obtemos a tabela a seguir.

Tabela 1.1: Valores dos pontos extremos aplicados na função objetivo

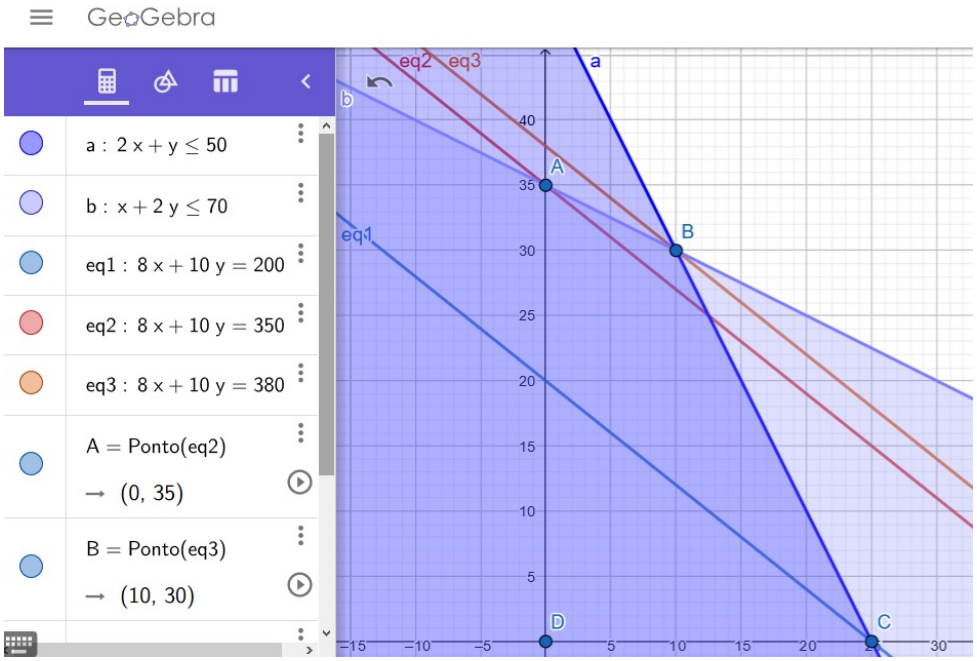
Ponto	Valor na função
$A(0, 35)$	350
$B(10, 30)$	380
$C(25, 0)$	200
$D(0, 0)$	0

Fonte: Elaborado pelo autor.

Com base na tabela 1.1, não podemos garantir que no ponto $B(10, 30)$ a função objetivo irá assumir seu maior valor para a região delimitada pelas restrições do problema. Por outro lado, como não é possível verificarmos para cada um dos pontos da região (temos uma infinidade de pontos), devemos utilizar algumas noções auxiliares que possam garantir que o resultado obtido produz a sua maximização.

A função objetivo no exemplo 1.2 é uma função linear ($z = 8x + 10y$), ou seja, será um plano em $\mathbb{R}^3$ passando pela origem $O = (0, 0, 0)$. As interseções deste plano com planos $z =$ constante serão retas paralelas que quando projetadas ortogonalmente sobre o plano xy determinam retas denominadas **curvas de nível**. O deslocamento paralelo destas retas dizem se a função objetivo cresce ou decresce na região viável. Na figura 1.4 a seguir, verificamos que as retas $eq1$, $eq2$ e $eq3$, são as curvas de nível para os níveis, $z = 200$, $z = 350$ e $z = 380$, respectivamente.

Figura 1.4: Curvas de nível exemplo 1.2



Fonte: Elaborado pelo autor.

Consideremos as seguintes definições:

Definição 1.3 (Derivada Parcial)**.** Seja f uma função definida em um conjunto aberto $U \subset \mathbb{R}^2$ em $\mathbb{R}$ e seja $(x_0, y_0) \in U$. A derivada parcial de f em relação a x no ponto (x_0, y_0) é definida por

$$f_x(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}.$$

Analogamente, define-se derivada parcial de f em relação a y no ponto (x_0, y_0) como

$$f_y(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = \lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}.$$

Geometricamente $\dfrac{\partial f}{\partial x}(x_0, y_0)$ será a inclinação da reta tangente à curva $C = [z = f(x, y)] \cap [y = y_0]$ no ponto $(x_0, y_0, f(x_0, y_0))$.

Definição 1.4 (Diferenciabilidade)**.** Seja $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ uma função definida em um conjunto aberto $U \subset \mathbb{R}^2$. Diremos que f é diferenciável no ponto $(x_0, y_0) \in U$ se existirem valores reais A e B tais que

$$\lim_{(h,k)\to(0,0)} \frac{f(x_0 + h, y_0 + k) - f(x_0, y_0) - A.h - B.k}{\sqrt{h^2 + k^2}} = 0.$$

A função f é dita diferenciável no aberto U se f for diferenciável em todos os pontos de U.

Definição 1.5 (Vetor Gradiente)**.** Seja $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ diferenciável no conjunto aberto U. O vetor gradiente de f no ponto $(x_0, y_0) \in U$ é definido como sendo $\nabla f(x_0, y_0) = \left(\dfrac{\partial f}{\partial x}(x_0, y_0), \dfrac{\partial f}{\partial y}(x_0, y_0)\right)$.

Seja a curva $\gamma : I \subset \mathbb{R} \to \mathbb{R}^2$ com o intervalo aberto I considerado apropriadamente de maneira que o traço de γ esteja contido no aberto U. Considere a curva $\alpha(t) = f(\gamma(t))$. Aplicando a Regra da Cadeia, temos que

$$\frac{d\alpha}{dt}(t) = \frac{\partial f}{\partial x}(x(t), y(t)).\frac{dx}{dt}(t) + \frac{\partial f}{\partial y}(x(t), y(t)).\frac{dy}{dt}(t).$$

Para $z = f(x, y) = c =$ constante segue que a curva de nível será dada por $\alpha(t) = c$ para todo $t \in I$. Assim,

$$\frac{\partial f}{\partial x}(x(t), y(t)).\frac{dx}{dt}(t) + \frac{\partial f}{\partial y}(x(t), y(t)).\frac{dy}{dt}(t) = 0$$

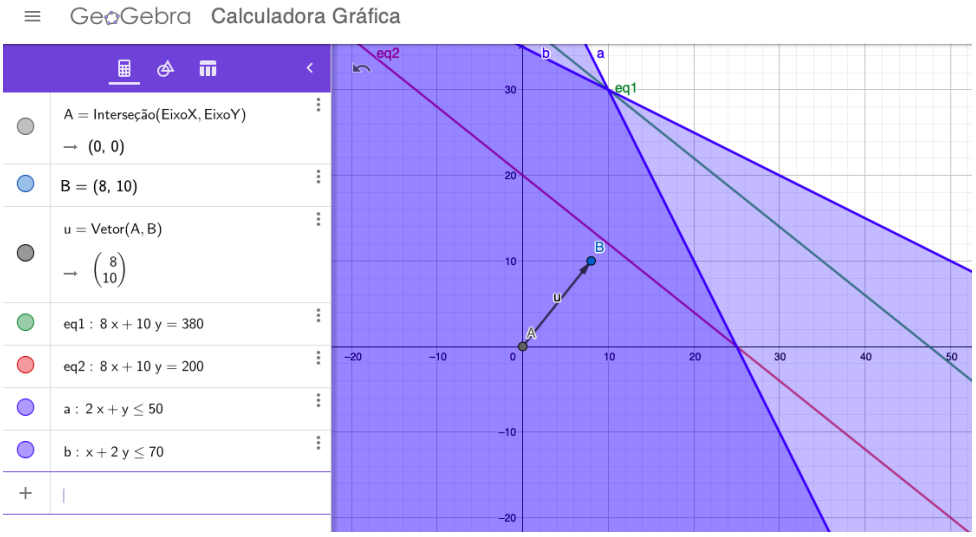
para todo $t \in I$. Ou seja,

$$\langle \nabla f(\gamma(t)), \gamma'(t) \rangle = 0 \tag{1.7}$$

onde $\langle , \rangle$ denota o produto interno usual de $\mathbb{R}^2$. Se restringirmos nosso estudo ao domínio U da função f temos que a expressão (1.7) nos diz que o vetor gradiente da função f será sempre perpendicular ao vetor velocidade da curva de nível no ponto $\gamma(t)$, indicando a direção e sentido de crescimento da função objetivo.

Retornando ao exemplo 1.2 verificamos que o gradiente da função objetivo $z = 8x+10y$ é o vetor $u = (8, 10)$, e este é perpendicular às curvas de nível, indicando a direção e sentido de crescimento da função objetivo.

Figura 1.5: Gradiente normal à curva de nível



Fonte: Elaborado pelo autor.

Neste caso, o ponto de coordenadas $x = 10$ e $y = 30$ será ponto de máximo.

O Método Gráfico possibilita discutirmos outros tipos de soluções para problemas de Programação Linear. Veremos a seguir alguns exemplos que dão origem a soluções múltiplas, ilimitadas e infactíveis.

Exemplo 1.6. Maximizar $x + 2y$

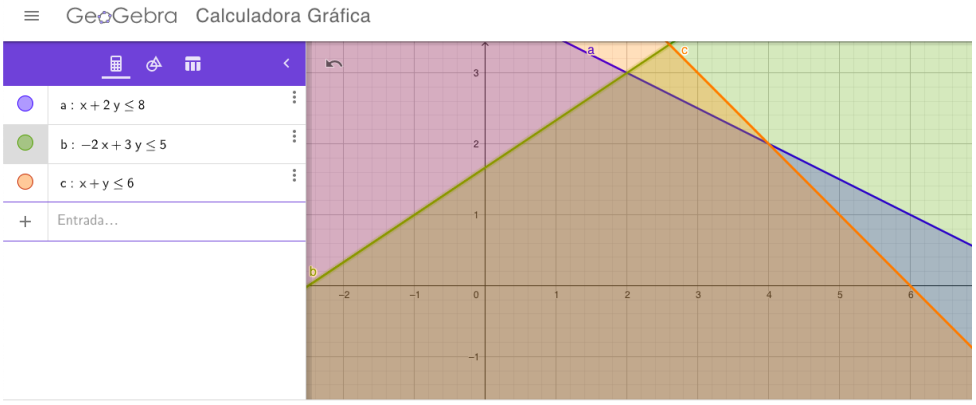
Sujeito a

$$\begin{cases} x + 2y \leqslant 8 \\ -2x + 3y \leqslant 5 \\ x + y \leqslant 6 \end{cases}$$

com $x \geqslant 0$ e $y \geqslant 0$.

Podemos observar que a função objetivo é paralela à inequação apresentada na primeira restrição. Sendo assim, o problema possui um conjunto (de múltiplas soluções). É o segmento da reta $x + 2y = 8$ entre os pontos $(2, 3)$ e $(4, 2)$ conforme mostra a figura 1.6 a seguir.

Figura 1.6: Múltiplas soluções



Fonte: Elaborado pelo autor.

Exemplo 1.7. Maximizar $2x + 3y$

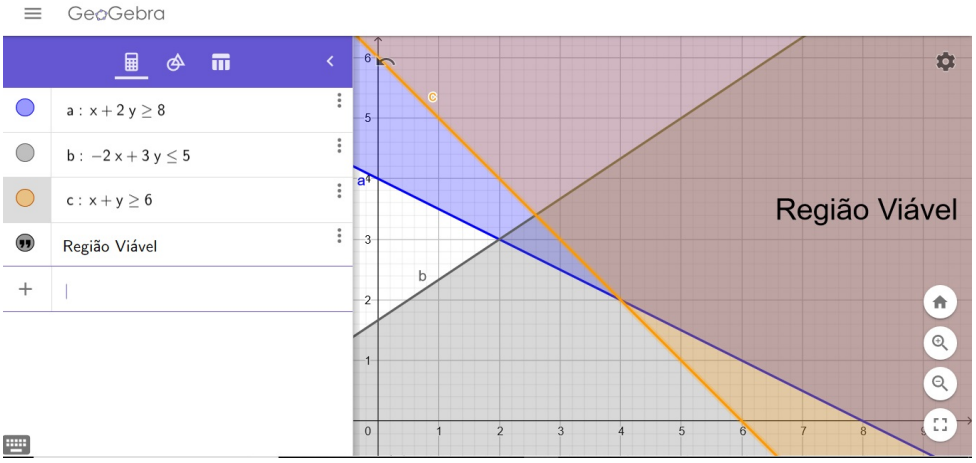
Sujeito a

$$\begin{cases} x + 2y \geqslant 8 \\ -2x + 3y \leqslant 5 \\ x + y \geqslant 6 \end{cases}$$

com $x \geqslant 0$ e $y \geqslant 0$.

A região viável obtida a partir das inequações é mostrada na figura 1.7 a seguir.

Figura 1.7: Exemplo de solução ilimitada



Fonte: Elaborado pelo autor.

Podemos observar na figura 1.7 que o problema tem solução tendendo ao infinito, pois a função objetivo pode se deslocar para a direita sem limitações e assim aumentar o valor de z indefinidamente.

Exemplo 1.8. Maximizar $2x + 3y$

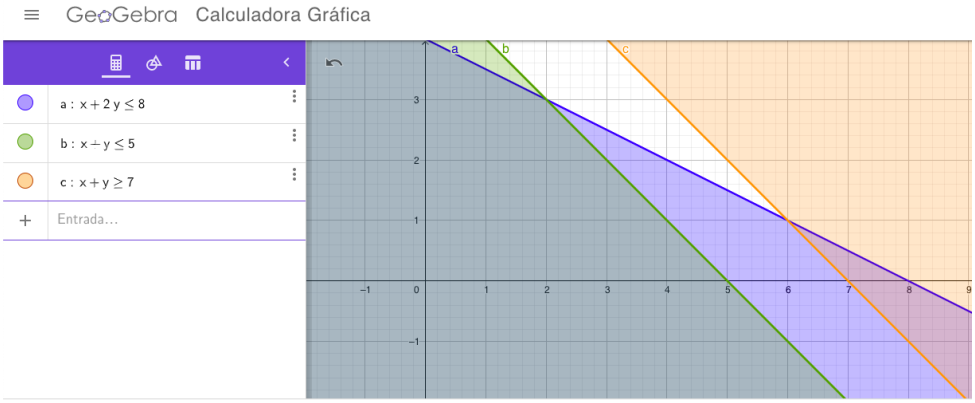
Sujeito a

$$\begin{cases} x + 2y \leqslant 8 \\ x + y \leqslant 5 \\ x + y \geqslant 7 \end{cases}$$

com $x \geqslant 0$ e $y \geqslant 0$.

Como não é possível satisfazer $x + y \leqslant 5$ e $x + y \geqslant 7$ simultaneamente, temos que o conjunto solução é vazio. Isso pode ser observado na figura 1.8 a seguir.

Figura 1.8: Exemplo de solução infactível



Fonte: Elaborado pelo autor.

1.2.2 Conjuntos Convexos

Quando tratamos de um conjunto de soluções de um problema de Programação Linear estamos falando de um conjunto de valores que pertencem a um espaço satisfazendo um conjunto de restrições. Vamos agora estudar as propriedades que este conjunto de soluções viáveis possui e assim nos auxiliar na percepção se um determinado problema tem ou não solução.

Consideremos o sistema $A.x = b$. Se $x_1, \ldots, x_n$ são coordenadas de $\mathbb{R}^n$ temos que cada equação $a_{i1}x_1 + \ldots + a_{in}x_n = b_i$ irá definir um hiperplano em $\mathbb{R}^n$. A solução para o sistema $A.x = b$ será a interseção de m hiperplanos.

Para cada $1 \leqslant i \leqslant m$ definimos os seguintes subconjuntos de $\mathbb{R}^n$.

$$H_i^- = \{x \in \mathbb{R}^n \text{ tal que } a_{i1}x_1 + \ldots + a_{in}x_n \leqslant b_i\}$$

e

$$H_i^+ = \{x \in \mathbb{R}^n \text{ tal que } a_{i1}x_1 + \ldots + a_{in}x_n \geqslant b_i\}$$

Da mesma maneira, uma solução para o problema de Programação Linear $A.x \leqslant b$ (ou $A.x \geqslant b$) irá pertencer ao conjunto interseção dos conjuntos $H_i^-, i = 1, \ldots, m$ (ou $H_i^+, i = 1, \ldots, m$).

Definição 1.9. O segmento de reta ligando os pontos distintos x_A e x_B em $\mathbb{R}^n$ é o conjunto de todos os pontos $P \in \mathbb{R}^n$ da forma $P = \lambda x_B + (1 - \lambda)x_A$ com $0 \leqslant \lambda \leqslant 1$.

Observamos que, se $\lambda = 0$ então $P = x_A$ e se $\lambda = 1$ então $P = x_B$.

Definição 1.10. Um conjunto não-vazio S de pontos de $\mathbb{R}^n$ é dito convexo se o segmento de reta ligando quaisquer dois pontos distintos em S está inteiramente contido em S.

Definição 1.11. Sejam $A_1, \ldots, A_n$ pontos dos conjunto S. O ponto $P \in S$ é dito combinação convexa dos pontos $A_1, \ldots, A_n$ se

$$P = \sum_{i=1}^{n} \lambda_i A_i$$

com $\lambda_i \geqslant 0$ para todo $1 \leqslant i \leqslant n$ e $\sum_{i=1}^{n} \lambda_i = 1$.

Definição 1.12. O ponto P é dito ponto extremo do conjunto convexo S de $\mathbb{R}^n$ se não pertencer a nenhum segmento de reta que une dois pontos quaisquer de S distintos de P.

Equivalentemente, um ponto P é ponto extremo de um conjunto S quando não é combinação convexa de quaisquer pontos de S.

Teorema 1.13. *O conjunto S de todas soluções do problema de Programação Linear $A.x \leqslant b$ é conjunto convexo.*

Demonstração. Sejam $x_A, x_B, \in S$ soluções quaisquer para o problema de Programação Linear $A.x \leqslant b$. Consideremos o segmento de reta que liga as soluções x_A e x_B, isto é, o conjunto de pontos dados por $P = \lambda x_B + (1 - \lambda)x_A$ com $\lambda \in [0, 1]$. Temos que mostrar que $A.P \leqslant b$. De fato,

$$A.P = A.[\lambda x_B + (1 - \lambda)x_A] = \lambda A.x_B + (1 - \lambda)A.x_A.$$

Por hipótese, $A.x_A \leqslant b$ e $A.x_b \leqslant b$. Como $\lambda \in [0, 1]$ segue que,

$$A.P \leqslant \lambda b + (1 - \lambda)b = b.$$

Logo, $A.P \leqslant b$. Portanto, P é solução e S é convexo.

Analogamente, temos o mesmo resultado para $A.x \geqslant b$. □

Consideremos agora S o conjunto de soluções admissíveis para um problema de Programação Linear. O conjunto S é obtido a partir da interseção de um número finito de hiperplanos. Desta forma, teremos as seguintes possibilidades:

1. $S = \emptyset$, ou seja, o problema não tem solução.
2. $S \neq \emptyset$ e S não é limitado. Neste caso, a função objetivo irá crescer indefinidamente.
3. $S \neq \emptyset$ e S é limitado. Para este caso segue o teorema 1.14.

Teorema 1.14. *Se um problema de Programação Linear admite uma solução ótima, então pelo menos um ponto extremo (vértice) do conjunto de pontos viáveis é uma solução ótima do problema.*

Demonstração. De fato, seja S o conjunto de soluções de um problema de Programação Linear e Z a função objetivo associada a ele. Suponhamos que S é não vazio e limitado com $x_1, x_2, \cdots, x_m$ pontos extremos do conjunto S. A função objetivo Z é contínua (função afim) e S é fechado e limitado, segue que a função objetivo admite máximo e mínimo no conjunto S. Assim, temos garantida a existência de um ponto de máximo de Z. Seja x^* uma solução ótima. Então,

$$Z(x^*) \geqslant Z(x) \text{ para todo } x \in S \quad \text{(problema de maximização)}. \tag{1.8}$$

Temos duas possibilidades:

1. x^* é ponto extremo. Então, o teorema está demonstrado.

2. x^* não é ponto extremo.

Como vimos no teorema 1.13, S é convexo. Então, todo $x \in S$ pode ser escrito como combinação convexa dos pontos extremos de S. Assim,

$$x^* = \sum_{i=1}^{m} \lambda_i x_i$$

e

$$Z(x^*) = c^T x^* = \lambda_1 c^T x_1 + \lambda_2 c^T x_2 + \cdots + \lambda_m c^T x_m.$$

Com $\lambda_i \geqslant 0$ para todo $1 \leqslant i \leqslant m$ e $\sum_{i=1}^{m} \lambda_i = 1$.

Precisamos mostrar que em algum ponto extremo x_i de S, $Z(x_i) = Z(x^*)$. Seja $x_{\max}$ o ponto extremo de S que satisfaz $Z(x_{\max}) = \max_{1 \leqslant i \leqslant m} Z(x_i)$. Segue então que,

$$\begin{aligned} Z(x^*) &= \lambda_1 c^T x_1 + \lambda_2 c^T x_2 + \cdots + \lambda_m c^T x_m \\ &\leqslant \lambda_1 Z(x_{\max}) + \lambda_2 Z(x_{\max}) + \cdots + \lambda_m Z(x_{\max}) = Z(x_{\max}) \sum_{i=1}^{m} \lambda_i = Z(x_{\max}). \end{aligned}$$

Logo,

$$Z(x^*) \leqslant Z(x_{\max}). \tag{1.9}$$

De (1.8) e (1.9) segue que $Z(x_{\max}) \leqslant Z(x^*) \leqslant Z(x_{\max})$. Portanto, $Z(x_{\max}) = Z(x^*)$. Concluímos que um dos pontos extremos de S é também ponto ótimo (solução ótima) do problema. □

Teorema 1.15. *Se um problema de Programação Linear admite uma solução ótima em mais de um ponto extremo, então toda combinação convexa destes pontos será também uma solução ótima do problema.*

Demonstração. De fato, sejam $x_1, x_2, \cdots, x_m$ pontos extremos de S. Sem perda de generalidade suponhamos que os pontos extremos $x_1, x_2, \cdots, x_n$, $n \leqslant m$ são pontos ótimos do conjunto S. Seja a função objetivo Z dada por $Z(x) = c^T x$. Se $M = \max_{1 \leqslant i \leqslant m} Z(x_i)$ segue que

$$Z(x_1) = Z(x_2) = \cdots = Z(x_n) = M.$$

Consideremos um ponto $x \in S$ obtido como combinação convexa dos pontos extremos ótimos de S, ou seja,

$$x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n.$$

com $\lambda_i \geqslant 0$ para todo $1 \leqslant i \leqslant n$ e $\sum_{i=1}^{n} \lambda_i = 1$. A função objetivo calculada neste ponto será dada por

$$\begin{aligned} Z(x) &= \lambda_1 Z(x_1) + \lambda_2 Z(x_2) + \cdots + \lambda_n Z(x_n) \\ &= \lambda_1 M + \lambda_2 M + \cdots + \lambda_n M \\ &= M \sum_{i=1}^{n} \lambda_i = M. \end{aligned}$$

Portanto, todo ponto dado como combinação convexa de pontos extremos ótimos será também ponto ótimo. □

Os teoremas 1.14 e 1.15 garantem que os pontos ótimos de um problema de Programação Linear, se existirem, irão ocorrer nos pontos extremos de S. No próximo capítulo iremos abordar um procedimento algorítmico que percorre os pontos extremos em busca do ponto ótimo.

2 Método Simplex

O Método Simplex consiste em determinar uma "solução ótima" para uma dada função objetivo, a partir de um conjunto de inequações, conhecidas como restrições. Inicialmente, o conjunto de inequações, cujas incógnitas são as variáveis do problema modelado, é transformado em um conjunto de equações com o acréscimo das chamadas variáveis de folga. Temos então, uma região delimitada pelas equações, onde as interseções das equações são candidatas a "solução ótima". O Método Simplex nos permite sair de um ponto extremo dado (solução viável básica) para um outro ponto extremo adjacente de tal modo que o valor da função objetivo cresce quando nos movemos de um ponto extremo para outro até obtermos uma solução ótima ou descobrirmos que o problema dado não tem uma solução ótima. O objetivo do algoritmo é percorrer todas as interseções e verificar qual valor produz o melhor resultado na nossa função objetivo.

A garantia de que o melhor resultado está em uma interseção das restrições foi visto no capítulo anterior.

O Método Simplex consiste de duas etapas:

1. Uma maneira de verificarmos se determinada solução viável básica é ótima, e
2. Uma maneira de obtermos outra solução viável básica onde a função objetivo tem um valor maior.

Consideremoos um problema padrão na Programação Linear descrito da seguinte forma:

$$\text{Maximize } z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

sujeito a

$$\begin{cases} a_{11}x_1 + & a_{12}x_2 + & \cdots + & a_{1n}x_n \leqslant & b_1 \\ a_{21}x_1 + & a_{22}x_2 + & \cdots + & a_{2n}x_n \leqslant & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 + & a_{m2}x_2 + & \cdots + & a_{mn}x_n \leqslant & b_m \end{cases}$$

com $x_j \geqslant 0$ para $1 \leqslant j \leqslant n$. Os valores a_{ij}, b_j e c_i são dados do problema.

2.1 Adicionando variáveis de folga

Como sabemos, tratar sistemas de desigualdades lineares algebricamente não é algo tão simples. Contudo, trabalhar com sistemas de equações lineares é algo bem conhecido na literatura matemática. Visando transformar nossas restrições em um conjunto de equações, adicionaremos variáveis que irão equilibrar as nossas desigualdades. Assim, consideremos uma restrição genérica dada por:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leqslant b_i \tag{2.1}$$

Observamos que o lado esquerdo de (2.1) não é maior do que o direito. Podemos transformar (2.1) em uma equação adicionando uma variável desconhecida não-negativa u do seu lado esquerdo obtendo

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + u = b_i \tag{2.2}$$

A variável u será denominada **variável de folga**.

Repetimos o procedimento em todas as inequações. Se para cada i-ésima restrição adicionarmos uma variável desconhecida não-negativa x_{i+n} podemos reescrever o problema padrão da seguinte forma: Determinar os valores de $x_1, x_2, \cdots, x_n, x_{n+1}, \cdots, x_{n+m}$ que irão maximizar

$$z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

sujeito a

$$\begin{cases} a_{11}x_1 + & a_{12}x_2 + & \cdots + & a_{1n}x_n + & x_{n+1} & & & = b_1 \\ a_{21}x_1 + & a_{22}x_2 + & \cdots + & a_{2n}x_n & & +x_{n+2} & & = b_2 \\ \vdots & \vdots & & \vdots & & & \ddots & \vdots \\ a_{m1}x_1 + & a_{m2}x_2 + & \cdots + & a_{mn}x_n & & & +x_{n+m} & = b_m \end{cases}$$

com $x_i \geqslant 0$ para $i = 1, \ldots, n+m$.

Em notação matricial temos: Determinar o vetor x que irá maximizar $z = c^Tx$ sujeito a $Ax = b$ e $x \geqslant 0$ onde

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & 0 & 0 & \cdots & 1 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{n+m} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Com o objetivo de facilitar o procedimento iremos organizar o problema de Programação Linear já com as varíaveis de folga adicionadas em um quadro. O quadro terá na primeira coluna as váriáveis ditas variáveis básicas. Nas colunas seguintes, os coeficientas da matriz A e na última coluna o vetor b. Uma linha é anexada abaixo do quadro colocando os coeficientes da função objetivo e o valor que esta assume nas variáveis básicas irá na última coluna. A figura 2.1 a seguir mostra o quadro inicial associado ao problema.

Tabela 2.1: Quadro inicial geral para um problema de Programação Linear

Variável básica	$\mathbf{x_1}$	$\mathbf{x_2}$	$\cdots$	$\mathbf{x_n}$	$\mathbf{x_{n+1}}$	$\mathbf{x_{n+2}}$	$\cdots$	$\mathbf{x_{n+m}}$	$\mathbf{b}$
x_{n+1}	a_{11}	a_{12}	$\cdots$	a_{1n}	1	0	$\cdots$	0	b_1
x_{n+2}	a_{21}	a_{22}	$\cdots$	a_{2n}	0	1	$\cdots$	0	b_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_{n+m}	a_{n1}	a_{n2}	$\cdots$	a_{nn}	0	0	$\cdots$	1	b_m
$Z - c^Tx$	$-c_1$	$-c_2$	$\cdots$	$-c_n$	0	0	$\cdots$	0	0

Fonte: Elaborado pelo autor.

Exemplo 2.1. Consideremos o exemplo 1.2. Inicialmente adicionamos as variáveis de folga, u e v, no problema padrão de Programação Linear.

$$\textbf{Maximizar} \quad z = 8x + 10y \tag{2.3}$$

$$\textbf{sujeito a} \quad \begin{cases} 2x + y + u = 50 \\ x + 2y + v = 70 \end{cases} \tag{2.4}$$

com $x, y, u, v \geqslant 0$.

Podemos escolher todas as variáveis originais (as que não são de folga) como nossas variáveis não-básicas, igualando-as a zero, ou seja, $x = y = 0$. Temos assim, $u = 50$ e $v = 70$. Nossa solução viável básica inicial é o vetor $\begin{bmatrix} 0 \\ 0 \\ 50 \\ 70 \end{bmatrix}$.

O quadro inicial para este problema será dado pelas variáveis x, y, u, v e z escritas na primeira linha como títulos das colunas correspondentes. Os vínculos (2.4) são colocadas nas linhas seguintes. A última linha do quadro, chamada de **linha correspondente a função objetivo**, referente à (2.3) será a equação $z - 8x - 10y = 0$. Ao longo do lado esquerdo do quadro indicamos a variável básica na equação correspondente. Então, u é uma variável básica na primeira equação e v é uma variável básica na segunda equação.

Tabela 2.2: Quadro inicial para o exemplo 1.2

	x	y	u	v	
u	2	1	1	0	50
v	1	2	0	1	70
	-8	-10	0	0	0

Fonte: Elaborado pelo autor.

O quadro inicial da tabela 2.2 acima mostra os valores das variáveis básicas u e v, as variáveis não básicas assumem os valores $x = 0$ e $y = 0$. O valor da função objetivo para essa solução básica é $8(0) + 10(0) + 0(u) + 0(v) = 0$, que é o elemento na última coluna da linha correspondente a função objetivo. Podemos observar que essa solução não é ótima, pois a partir da última linha do quadro inicial escrevemos

$$z = 0 + 8x + 10y - 0u - 0v. \tag{2.5}$$

O valor de z irá crescer aumentando-se o valor de x ou y, uma vez que essas duas variáveis aparecem em (2.5) com coeficientes positivos. Como a linha correspondente a função objetivo do nosso quadro inicial tem elementos negativos nas colunas associadas a variáveis, podemos aumentar z incrementando qualquer variável que tenha um elemento negativo em sua coluna correspondente na linha correspondente a função objetivo.

Podemos assim obter o critério de otimização descrito a seguir para determinar se uma solução viável indicada em um quadro é uma solução ótima, gerando um valor máximo para a função objetivo z.

Critério de otimização : *Se a linha correspondente a função objetivo de um quadro não tem elementos negativos nas colunas associadas a variáveis, então a solução indicada é ótima e podemos parar nossos cálculos.*

2.2 Selecionando a variável de entrada

Se a linha correspondente a função objetivo de um quadro tem elementos negativos nas colunas associadas a variáveis, então a solução indicada não é ótima e temos que continuar a procurar uma solução ótima.

O Método Simplex move-se de um ponto extremo para outro ponto extremo adjacente de modo a aumentar o valor da função objetivo. Isso é feito aumentando-se uma variável de cada vez. Podemos obter o aumento por vez para z escolhendo a variável que corresponde ao número mais negativo na linha correspondente a função objetivo. Se ocorrer empate entre os candidatos a variável de entrada, escolhemos um deles.

Exemplo 2.2. No quadro da tabela 2.2 para o exemplo 1.2, o elemento mais negativo na linha correspondente a função objetivo é -10, na coluna y. Assim, esta será a variável a ser aumentada. A variável a ser aumentada é chamada de **variável de entrada**, pois ela se tornará uma variável básica na próxima iteração, entrando no conjunto de variáveis básicas. Um aumento em y tem que ser acompanhado por uma diminuição em algumas das outras variáveis. Isso pode ser visto se resolvemos a equação em (2.4) para u e v:

$$\begin{aligned} u &= 50 - 2x - y \\ v &= 70 - x - 2y \end{aligned}$$

Como vamos aumentar apenas y, mantemos $x = 0$ e obtemos

$$\begin{cases} u = 50 - y \\ v = 70 - 2y \end{cases} \tag{2.6}$$

de modo que, quando y aumenta, tanto u quanto v diminuem. As equações (2.6) também mostram o quanto podemos aumentar y. Em outras palavras, como u e v não podem ser negativos, temos

$$y \leqslant \frac{50}{1} = 50$$

e

$$y \leqslant \frac{70}{2} = 35.$$

Podemos ver que o aumento permitido para y não pode ser maior do que o menor dos quocientes $\frac{50}{1}$ e $\frac{70}{2}$. Fazendo $y = 35$, obtemos a seguinte solução viável básica:

$$\begin{cases} x = 0 \\ y = 35 \\ u = 15 \\ v = 0 \end{cases}$$

Assim, as variáveis básicas são y e u. As variáveis x e v são não básicas. A função objetivo para essa solução tem agora o valor $z = 8(0) + 10(35) + 0(15) + 0(0) = 350$.

2.3 Escolhendo a variável de saída

Vimos no exemplo 2.2 que variável $v = 0$ não é básica. Esta variável será chamada de **variável de saída**, pois saiu do conjunto de variáveis básicas. A coluna da variável de entrada é chamada de **coluna do pivô** e a linha da variável de saída de **linha do pivô**.

A escolha da variável de saída está relacionada à determinação de quanto queremos aumentar a variável de entrada (no caso do exemplo 2.2, a variável y). Para isso, consideramos os quocientes dos elementos acima da linha correspondente a função objetivo na última coluna à direita do quadro pelos elementos correspondentes na coluna do pivô. O menor desses quocientes nos diz de quanto podemos aumentar a variável de entrada. A variável básica associada à linha para o qual temos o menor quociente (a linha do pivô) é, então, a variável de saída. No exemplo 2.2, $\frac{50}{1}$ e $\frac{70}{2}$ na primeira e segunda linha, respectivamente. O menor desses quocientes, 35, indica que a segunda linha é a linha do pivô, e a variável básica, v, associada é a variável de saída que não será mais básica. Se não escolhermos o menor quociente, então uma das variáveis na nova solução será negativa. Logo, a nova solução não será mais uma solução viável.

O que acontece se existem elementos nulos ou negativos na coluna pivô? Se algum elemento na coluna pivô é negativo, então o quociente correspondente também é negativo, sendo assim a equação associada ao quociente negativo não impõe restrição alguma em quanto podemos aumentar a variável da entrada.

Se algum elemento na coluna do pivô for nulo, o quociente correspondente não pode ser formado (não podemos dividir por zero) e, novamente, a equação associada não limita de quanto podemos aumentar y. Portanto, ao formar os quocientes, usamos apenas os elementos positivos acima da linha do pivô na coluna do pivô.

Se todos os elementos acima da linha correspondente a função objetivo na coluna do pivo são nulos ou negativos, então a variável de entrada pode ser escolhida tão grande quanto quisermos. Isso significa que o problema não tem solução finita ótima.

Exemplo 2.3. Dando continuidade ao exemplo 2.2, obtemos um novo quadro indicando as novas variáveis básicas e as novas soluções viáveis básicas. Resolvendo a segunda equação em (2.4) para y, temos

$$y = 35 - \frac{1}{2}x - \frac{1}{2}v. \tag{2.7}$$

Substituindo a expressão (2.7) na primeira equação em (2.4) temos

$$\frac{3}{2}x + u - \frac{1}{2}v = 15. \tag{2.8}$$

Dividindo por 2 (o coeficiente de y) a segunda equação em (2.4), segue que

$$\frac{1}{2}x + y + \frac{1}{2}v = 35. \tag{2.9}$$

Substituindo y por (2.7) em (2.3) obtemos

$$-3x + 5v + z = 350. \tag{2.10}$$

Como $x = 0$ e $v = 0$, obtemos o valor de z para a solução viável básica atual. Ou seja, $z = 350$. As equações (2.8), (2.9) e (2.10) geram o novo quadro da tabela 2.3 mostrado a seguir.

Tabela 2.3: Segundo quadro para o exemplo 1.2

	x	y	u	v	
u	$\frac{3}{2}$	0	1	$-\frac{1}{2}$	15
y	$\frac{1}{2}$	1	0	$\frac{1}{2}$	35
	-3	0	0	5	350

Fonte: Elaborado pelo autor.

Observemos que colocamos as variáveis básicas no quadro indicando sua linha correspondente.

Comparando o quadro da tabela 2.2 com quadro da tabela 2.3, observamos que podemos transformar o primeiro no último usando as operações elementares nas linhas descritas a seguir.

Passo 1: Localizamos e marcamos com um círculo em volta o elemento na linha e coluna dos pivôs. Esse elemento é o **pivô**. Marcamos a coluna do pivô colocando uma seta $\downarrow$ acima da variável de entrada e marcamos a linha do pivô colocando uma seta $\leftarrow$ à esquerda da variável de saída.

Passo 2: Se o pivô $k \neq 0$ então multiplicamos a linha do pivô por $\frac{1}{k}$, transformando o elemento pivô em 1.

Passo 3: Devemos somar múltiplos apropriados da linha do pivô a todas as outras linhas (inclusive a linha correspondente a função objetivo) de modo a transformar em 0 todos os elementos da coluna pivô exceto o elemento pivô, que continua 1.

Passo 4: No novo quadro, trocamos a variável na linha do pivô colocando a variável de entrada.

Esses quatro passos formam um processo conhecido como **eliminação por pivô**. É um dos procedimentos para transformar uma matriz em uma forma escada reduzida por linhas.

Exemplo 2.4. O quadro da tabela 2.2, agora com setas próximas das variáveis de entrada e de saída e o elemento pivô marcado, é mostrado na tabela 2.4 a seguir.

Tabela 2.4: Quadro inicial para o exemplo 1.2

			↓			
		x	y	u	v	
	u	2	1	1	0	50
←	v	1	(2)	0	1	70
		-8	-10	0	0	0

Fonte: Elaborado pelo autor.

Usando a eliminação por pivô no quadro da tabela 2.2 obtemos o quadro da tabela 2.3. Vamos agora repetir o procedimento usando o quadro da tabela 2.3 como quadro inicial. O elemento mais negativo na linha correspondente a função objetivo é -3 que aparece na primeira coluna. Desta forma, x é a variável de entrada e a primeira coluna é a coluna do pivô.

Para determinar a variável de saída, formamos os quocientes dos elementos na última coluna à direita (exceto para a linha correspondente a função objetivo) pelos elementos correspondentes na coluna do pivô que são positivos e selecionamos o menor desses quocientes. Como ambos os elementos na coluna do pivô são positivos, os quocientes são $\frac{15}{\frac{3}{2}}$ e $\frac{35}{\frac{1}{2}}$. O menor desses, 10, ocorre na primeira linha. Deste modo, a variável de saída é u e a linha do pivô é a primeira.

Tabela 2.5: Segundo quadro para o exemplo 1.2

		↓				
		x	y	u	v	
←	u	$(\frac{3}{2})$	0	1	$-\frac{1}{2}$	15
	y	$\frac{1}{2}$	1	0	$\frac{1}{2}$	35
		-3	0	0	5	350

Fonte: Elaborado pelo autor.

Usando a eliminação por pivô no quadro da tabela 2.5 obtemos o quadro da tabela 2.6 a seguir.

Tabela 2.6: Quadro final para o exemplo 1.2

		x	y	u	v	
	x	1	0	$\frac{2}{3}$	$-\frac{1}{3}$	10
	y	0	1	$-\frac{1}{3}$	$\frac{2}{3}$	30
		0	0	2	4	380

Fonte: Elaborado pelo autor.

A linha correspondente a função objetivo no quadro da tabela 2.6 não tem elementos negativos nas colunas associadas a variáveis, concluimos pelo critério de otimização, que terminamos o processo e a solução obtida é ótima. A solução $\{x = 10,\ y = 30,\ u = 0,\ v = 0\}$ é ótima e corresponde ao ponto extremo $(10, 30)$. Podemos observar que o Método Simplex começou com o ponto extremo $(0, 0)$, moveu para o ponto extremo adjacente $(0, 35)$ e por fim, para o ponto extremo $(10, 30)$, adjacente a $(0, 35)$. O valor da função objetivo cresceu de 0 para 350 e então para 380, respectivamente.

2.4 O Método Simplex

O procedimento discutido na seção anterior pode ser resumido nos seguintes passos:

Passo 1 Faça o quadro inicial.

Passo 2 Aplique o teste de otimização: *Se a linha correspondente a função objetivo não tem elementos negativos nas colunas associadas a variáveis, então a solução indicada é ótima e paramos nossos cálculos.*

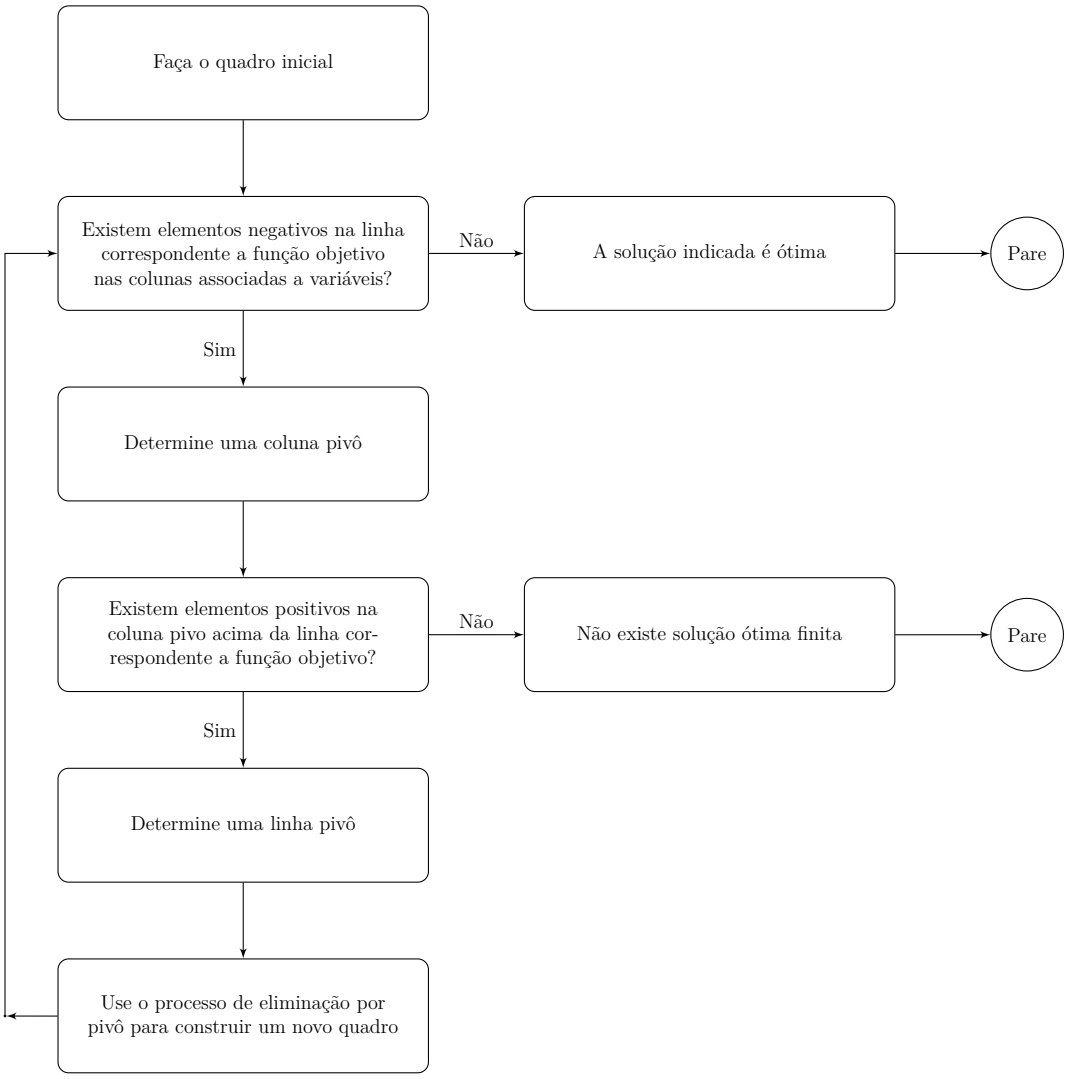
Passo 3 Escolha uma coluna do pivô determinado a coluna que tem o elemento mais negativo na linha correspondente a função objetivo. Se houver mais de um candidato para coluna pivô, escolha um.

Passo 4 Determine os quocientes dos elementos acima da linha correspondente a função objetivo na última coluna à direita pelos elementos positivos correspondentes na coluna do pivô. A linha do pivô é a que corresponde ao menor desses quocientes. Se existe um empate, de modo que o menor quociente acontece em mais de uma linha, escolha qualquer uma das linhas. Se nenhum dos elementos da coluna do pivô acima da linha correspondente a função objetivo é positivo, o problema não tem solução ótima finita. O processo se encerra.

Passo 5 Use o processo de eliminação por pivô para construir um novo quadro e volte para passo 2.

A figura 2.1 a seguir exibe o fluxograma para o Método Simplex.

Figura 2.1: Fluxograma do Método Simplex



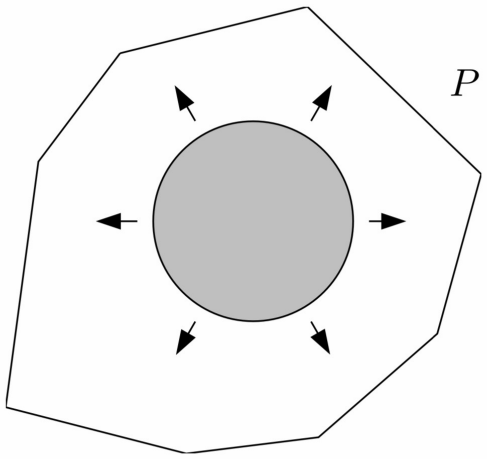
Fonte: Elaborado pelo autor.

2.5 Exemplo: maior círculo inscrito em um polígono

Encerramos este capítulo fazendo uma aplicação do Método Simplex a um problema de empacotamento geométrico. A ideia consiste em determinar o maior círculo em uma região dada por um polígono convexo qualquer. A princípio este problema não parece ser um problema de Programação Linear, contudo, podemos adaptá-lo transformando-o em um problema padrão de Programação Linear.

Consideremos um polígono convexo P de n lados. Como dissemos, queremos determinar o maior círculo contido em P. Entendemos por maior círculo o círculo de maior área contido na região interior ao polígono convexo. Ver a figura 2.2 a seguir.

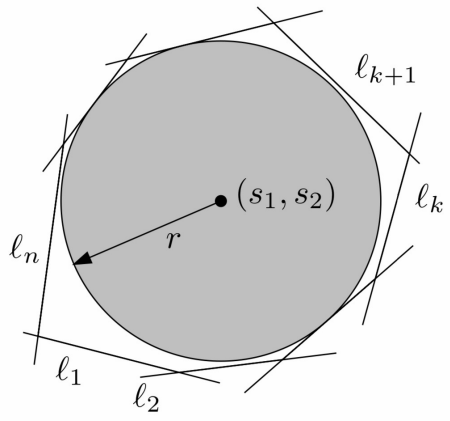
Figura 2.2: Maior círculo inscrito em um polígono



Fonte: Matousek et. al. (2007, p.24)

Inicialmente, sem perda de generalidade, supomos que nenhum dos lados de P seja vertical e que cada lado l_i de P é dado pela equação $y = a_i x + b_i$, com $i = 1, 2, ..., n$. Podemos escolher a numeração dos lados de forma que o primeiro, segundo, até o k-ésimo lado do limite P, estejam abaixo do centro da circunferência e os lados $l_{(k+1)}$ até o n-ésimo lado estejam acima do centro da circunferência, conforme mostra a figura 2.3 a seguir.

Figura 2.3: Círculo limitado por retas



Fonte: Matousek et. al. (2007, p.24)

Consideremos r_i a reta que determina a i-ésima aresta ($y = a_i x + b_i$) do polígono. Sua equação paramétrica dada por

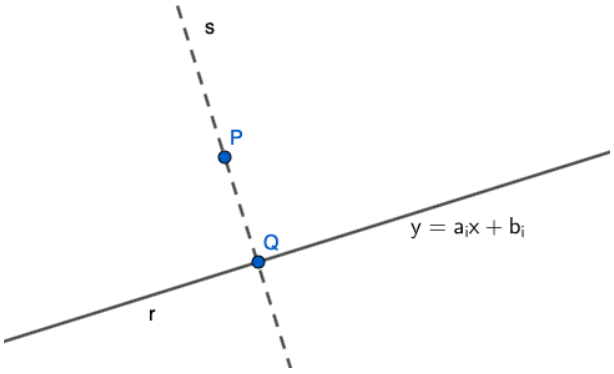
$$\begin{cases} x = s \\ y = b_i + a_i s, s \in \mathbb{R} \end{cases}$$

Assim, o vetor direção da reta será $(1, a_i)$. Sabemos que o vetor normal à reta r_i é dado por $(a_i, -1)$. Então, a equação paramétrica da reta perpendicular à reta r_i passando por um ponto $P = (x_0, y_0)$ será

$$\begin{cases} x = x_0 + ta_i \\ y = y_0 + (-1)t, t \in \mathbb{R}. \end{cases} \tag{2.11}$$

Seja Q o ponto de interseção da reta r_i e a reta normal a ela pelo ponto P. Ver figura 2.4 a seguir.

Figura 2.4: Reta normal e ponto de interseção



Fonte: Elaborado pelo autor.

Para determinamos as coordenadas do ponto Q basta impormos

$$a_i(x_0 + t_Q a_i) + b_i = y_0 - t_Q \tag{2.12}$$

E resolver a equação (2.12) em t_Q. Deste modo, segue que

$$\begin{aligned} t_Q + t_Q {a_i}^2 &= y_0 - a_i x_0 - b_i \\ t_Q({a_i}^2 + 1) &= y_0 - a_i x_0 - b_i \\ t_Q &= \frac{y_0 - a_i x_0 - b_i}{{a_i}^2 + 1}. \end{aligned}$$

Substituindo t_Q na equação (2.11) obtemos as coordenadas do ponto Q.

A distância do ponto $P = (x_0, y_0)$ a reta r_i é dada pela distância do ponto P ao ponto $Q = (x_Q, y_Q)$. Ou seja,

$$\begin{aligned} d(P,Q) &= \sqrt{(x_0 - x_Q)^2 - (y_0 - y_Q)^2} \\ &= \sqrt{(x_0 - (x_0 + t_Q a_i))^2 - (y_0 - (y_0 - t_Q))^2} \\ &= \sqrt{t_Q^2 {a_i}^2 + t_Q^2} = |t_Q|\sqrt{{a_i}^2 + 1} = \left|\frac{y_0 - a_i x_0 - b_i}{{a_i}^2 + 1}\right| \cdot \sqrt{{a_i}^2 + 1} \\ &= \frac{|y_0 - a_i x_0 - b_i|}{\sqrt{{a_i}^2 + 1}}. \end{aligned}$$

Desta maneira, a distância do centro $C = (x, y)$ de um círculo de raio r ao i-ésimo lado do polígono convexo será dada por

$$d_i = \frac{|y - a_i x - b_i|}{\sqrt{a_i^2 + 1}}. \tag{2.13}$$

A expressão $y - a_i x - b_i$ será positiva quando a reta estiver "abaixo" do centro C do círculo de raio r, e negativa quando a linha estiver "acima".

Sob estas condições nossas restrições serão:

$$\frac{y - a_i x - b_i}{\sqrt{a_i^2 + 1}} \geqslant r, \text{ para } i \leqslant k,$$
$$\frac{y - a_i x - b_i}{\sqrt{a_i^2 + 1}} \leqslant -r, \text{ para } i > k.$$

Reorganizando as desigualdades temos:

$$\left(\frac{1}{\sqrt{a_i^2 + 1}}\right) y - \left(\frac{a_i}{\sqrt{a_i^2 + 1}}\right) x - r \geqslant \left(\frac{b_i}{\sqrt{a_i^2 + 1}}\right), \text{para } i \leqslant k, \tag{2.14}$$
$$\left(\frac{1}{\sqrt{a_i^2 + 1}}\right) y - \left(\frac{a_i}{\sqrt{a_i^2 + 1}}\right) x + r \leqslant \left(\frac{b_i}{\sqrt{a_i^2 + 1}}\right), \text{para } i > k. \tag{2.15}$$

O círculo procurado será aquele de maior raio r. Ressaltamos que os valores entre os parenteses são constantes uma vez que fazem parte da descrição do polígono. Enunciamos o problema de Programação Linear da seguinte forma:

Maximizar r

$$\textbf{Sujeito a} \begin{cases} \left(\frac{1}{\sqrt{a_i^2 + 1}}\right) y - \left(\frac{a_i}{\sqrt{a_i^2 + 1}}\right) x - r \geqslant \left(\frac{b_i}{\sqrt{a_i^2 + 1}}\right), \text{para } i \leqslant k \\ \left(\frac{1}{\sqrt{a_i^2 + 1}}\right) y - \left(\frac{a_i}{\sqrt{a_i^2 + 1}}\right) x + r \leqslant \left(\frac{b_i}{\sqrt{a_i^2 + 1}}\right), \text{para } i > k \end{cases}$$

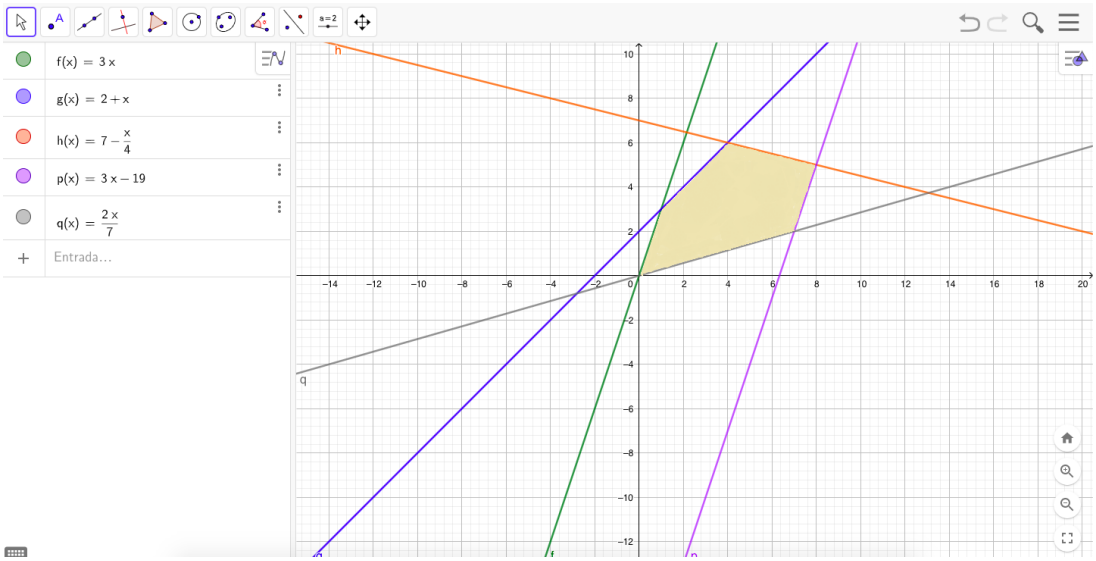
com x, y e $r \geqslant 0$.

Exemplo 2.5. Consideremos a região poligonal determinada pelos pontos $A = (0, 0)$, $B = (1, 3)$, $C = (4, 6)$, $D = (8, 5)$, $E = (7, 2)$. As cinco retas no plano cartesiano que passam por estes pontos são dadas por:

$$y = 3x,$$
$$y = x + 2,$$
$$y = -\frac{x}{4} + 7,$$
$$y = 3x - 19,$$
$$y = \frac{2}{7}x.$$

A região limitada pode ser vista na figura 2.5 a seguir.

Figura 2.5: Região limitada pelas retas



Fonte: Elaborado pelo autor.

Como vimos anteriormente, podemos descrever as retas acima em termos dos pares (a_i, b_i), ou seja, (3, 0), (1,2), $\left(-\frac{1}{4}, 7\right)$, (3,-19), $\left(\frac{2}{7}, 0\right)$.

Aplicando os pontos que descrevem as retas nas equações (2.14) e (2.15), obtemos as restrições para o problema.

Para o par (3, 0), temos:

$$\left(\frac{1}{\sqrt{(3)^2+1}}\right)y - \left(\frac{3}{\sqrt{(3)^2+1}}\right)x + r \leqslant \frac{0}{\sqrt{(3)^2+1}} \Leftrightarrow \frac{1}{\sqrt{10}}y - \frac{3}{\sqrt{10}}x + r \leqslant 0$$

Para o par (1,2), temos:

$$\left(\frac{1}{\sqrt{1^2+1}}\right)y - \left(\frac{1}{\sqrt{1^2+1}}\right)x + r \leqslant \left(\frac{2}{\sqrt{1^2+1}}\right) \Leftrightarrow \frac{1}{\sqrt{2}}y - \frac{1}{\sqrt{2}}x + r \leqslant \frac{2}{\sqrt{2}}$$

Para o par $\left(-\frac{1}{4}, 7\right)$, temos:

$$\frac{1 \cdot y}{\sqrt{\left(-\frac{1}{4}\right)^2+1}} - \frac{\left(-\frac{1}{4}\right) \cdot x}{\sqrt{\left(-\frac{1}{4}\right)^2+1}} + r \leqslant \frac{7}{\sqrt{\left(-\frac{1}{4}\right)^2+1}} \Leftrightarrow \frac{4}{\sqrt{17}}y + \frac{1}{\sqrt{17}}x + r \leqslant \frac{28}{\sqrt{17}}$$

Para o par (3,-19), temos:

$$\left(\frac{1}{\sqrt{3^2+1}}\right)y - \left(\frac{3}{\sqrt{3^2+1}}\right)x - r \geqslant \frac{-19}{\sqrt{3^2+1}} \Leftrightarrow -\frac{1}{\sqrt{10}}y + \frac{3}{\sqrt{10}}x + r \leqslant \frac{19}{\sqrt{10}}$$

Para o par $\left(\frac{2}{7}, 0\right)$, temos:

$$\frac{1}{\sqrt{\left(\frac{2}{7}\right)^2+1}}\cdot y - \frac{\left(\frac{2}{7}\right)}{\sqrt{\left(\frac{2}{7}\right)^2+1}}\cdot x - r \geqslant \frac{0}{\sqrt{\left(\frac{2}{7}\right)^2+1}} \Leftrightarrow -\frac{7}{\sqrt{53}}y + \frac{2}{\sqrt{53}}x + r \leqslant 0$$

Sabendo as restrições e a função objetivo formulamos o problema, agora em valores aproximados em duas casas decimais, da seguinte forma:

Maximizar r

$$\textbf{Sujeito a} \begin{cases} 0,32y - 0,95x + r \leqslant 0 \\ 0,71y - 0,71x + r \leqslant 1,41 \\ 0,97y + 0,24x + r \leqslant 6,79 \\ -0,32y + 0,95x + r \leqslant 6,01 \\ -0,96y + 0,27x + r \leqslant 0 \end{cases}$$

com x, y e $r \geqslant 0$.

Seguindo os passos do Método Simplex descrito neste capítulo, devemos montar o quadro inicial e determinar a coluna pivô. A coluna pivô é aquela que possui a variável mais negativa na linha correspondente a função objetivo. Como a função objetivo do nosso problema possui apenas a variável r a maximizar, facilmente determinamos a coluna pivô. A seguir, determinamos o elemento pivô. Ao dividir cada termo indepente (b) pelo correspondente elemento da coluna pivô, o menor valor resultante das divisões, não negativo, será chamado de elemento pivô. No nosso problema há empate, em ambas divisões o resultado é zero, podemos então escolher qualquer um.

Tabela 2.7: Quadro inicial para o exemplo 2.5

	y	x	r	f_1	f_2	f_3	f_4	f_5	
f_1	0,32	-0,95	1	1	0	0	0	0	0
f_2	0,71	-0,71	1	0	1	0	0	0	1,41
f_3	0,97	0,24	1	0	0	1	0	0	6,79
f_4	-0,32	0,95	1	0	0	0	1	0	6,01
f_5	-0,96	0,27	1	0	0	0	0	1	0
	0	0	-1	0	0	0	0	0	0

Fonte: Elaborado pelo autor.

Analisando a Tabela 2.7, temos que o menor quociente entre o termo independente e o elemento correspondente na coluna pivô será zero. Iremos então somar múltiplos convenientes de forma a zerar o elementos da coluna pivô, exceto o elemento pivô que deverá ser 1. Obtemos a Tabela 2.8 a seguir.

Tabela 2.8: Iteração 1

				↓						
		y	x	r	f_1	f_2	f_3	f_4	f_5	
←	f_1	0,32	-0,95	1	1	0	0	0	0	0
	f_2	0,71	-0,71	1	0	1	0	0	0	1,41
	f_3	0,97	0,24	1	0	0	1	0	0	6,79
	f_4	-0,32	0,95	1	0	0	0	1	0	6,01
	f_5	-0,96	0,27	1	0	0	0	0	1	0
		0	0	-1	0	0	0	0	0	0

Fonte: Elaborado pelo autor.

Podemos observar que uma vez transformada uma linha pivô, esta não deve ser novamente considerada para fins de análise da escolha do novo elemento pivô na próxima iteração.

O processo será repetido até que não haja elementos negativos na linha correspondente a função objetivo.

Tabela 2.9: Iteração 2

			↓							
		y	x	r	f_1	f_2	f_3	f_4	f_5	
	r	0,32	-0,95	1	1	0	0	0	0	0
	f_2	0,39	0,24	0	-1	1	0	0	0	1,41
	f_3	0,65	1,19	0	-1	0	1	0	0	6,79
	f_4	-0,63	1,90	0	-1	0	0	1	0	6,01
←	f_5	-1,28	1,22	0	-1	0	0	0	1	0
		0,32	-0,95	0	1	0	0	0	0	0

Fonte: Elaborado pelo autor.

Tabela 2.10: Iteração 3

		↓								
		y	x	r	f_1	f_2	f_3	f_4	f_5	
	r	-0,67	0	1	0,22	0	0	0	0,78	0
←	f_2	0,64	0	0	-0,80	1	0	0	-0,20	1,41
	f_3	1,90	0	0	-0,03	0	1	0	-0,97	6,79
	f_4	1,35	0	0	0,55	0	0	1	-1,55	6,01
	x	-1,04	1	0	-0,82	0	0	0	0,82	0
		-0,67	0	0	0,22	0	0	0	0,78	0

Fonte: Elaborado pelo autor.

Tabela 2.11: Iteração 4

					↓					
		y	x	r	f_1	f_2	f_3	f_4	f_5	
	r	0	0	1	-0,62	1,05	0	0	0,57	1,48
	y	1	0	0	-1,25	1,55	0	0	-0,31	2,20
←	f_3	0	0	0	2,34	-2,95	1	0	-0,39	2,62
	f_4	0	0	0	2,23	-2,10	0	1	-1,14	3,04
	x	0	1	0	-2,12	1,62	0	0	0,50	2,30
		0	0	0	-0,62	1,05	0	0	0,57	1,48

Fonte: Elaborado pelo autor.

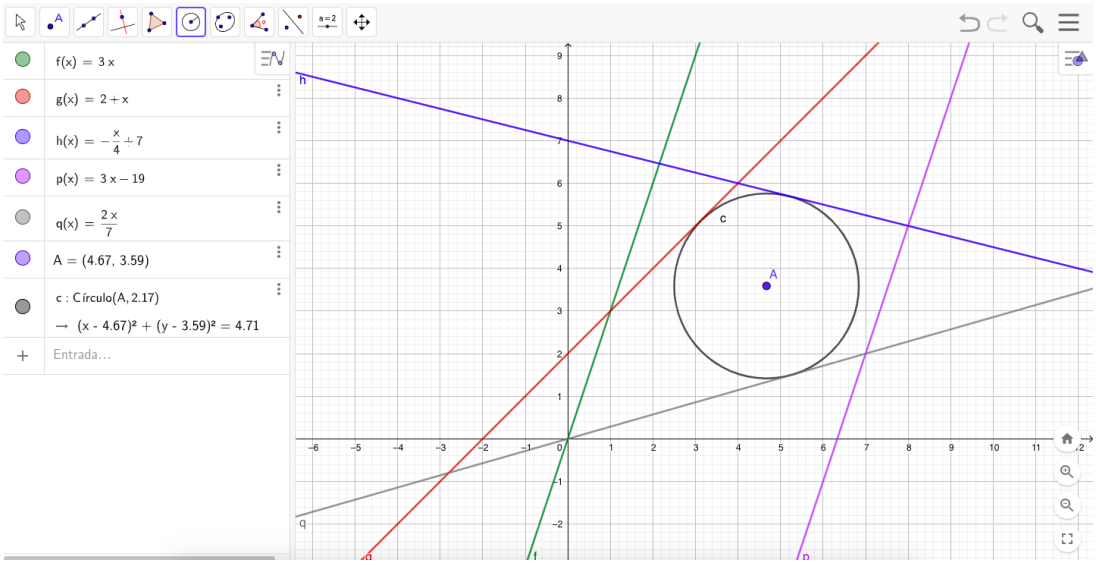
Tabela 2.12: Quadro Final - Exemplo 2.5

	y	x	r	f_1	f_2	f_3	f_4	f_5	
r	0	0	1	0	0,27	0,26	0	0,47	2,17
y	1	0	0	0	-0,02	0,53	0	-0,52	3,59
f_1	0	0	0	1	-1,26	0,43	0	-0,17	1,12
f_4	0	0	0	0	0,72	-0,95	1	-0,76	0,54
x	0	1	0	0	-1,05	0,91	0	0,14	4,67
	0	0	0	0	0,27	0,26	0	0,47	2,17

Fonte: Elaborado pelo autor.

Através da Tabela 2.12 verificamos que o algoritmo concluiu seu objetivo. Temos então os valores aproximados $r = 2,17$, $y = 3,59$ e $x = 4,67$.

Figura 2.6: Gráfico solução - Exemplo 2.5



Fonte: Elaborado pelo autor.

3 Programação Não Linear e as Condições de Karush-Kuhn-Tucker

A Programação Linear e o Método Simplex permitem-nos estudar conteúdos importantes da matemática. Nos capítulos anteriores, abordamos vários conceitos matemáticos como: resolução de sistemas; soluções geométricas; gradiente; curvas de nível; máximo e mínino e funções lineares com restrições.

O objetivo deste capítulo é mais uma vez conectar conteúdos e expandir os conceitos iniciais para funções com mais de duas variáveis e funções não lineares. A modelagem de problemas reais são muitas vezes melhor representados por equações que geralmente não são lineares. Podemos encontrar diversos exemplos na Física, Engenharia e Economia.

Quando falamos no Método dos Multiplicadores de Lagrange estamos tratando de problemas para a determinação de candidatos extremantes condicionados, onde as restrições são do tipo igualdade. Ao final deste capítulo abordaremos as condições de Karush-Kuhn-Tucker ou KKT para a determinação de candidatos extremantes condicionados, sendo que as restrições passam a ser desigualdades.

O modelo empregado na Programação Não Linear segue a mesma motivação apresentada na Programação Linear, isto é, buscamos determinar o vetor x que irá

Maximizar (ou minimizar) $f(x)$

$$\textbf{Sujeito a} \begin{cases} g_1(x) \leqslant b_1 \\ g_2(x) \leqslant b_2 \\ \vdots \\ g_m(x) \leqslant b_m \\ x \geqslant 0 \end{cases}$$

$x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ e f e $g_i \, (1 \leqslant i \leqslant m)$ são funções não lineares.

Na Programação Não Linear o princípio de busca pela solução ótima, na maioria dos métodos existentes, segue na direção da maior taxa de variação da função (dada pelo vetor gradiente) escolhendo arbitrariamente um conjunto de soluções viáveis e caminhando para uma solução teoricamente melhor.

As principais referências utilizadas neste capítulo são Guidorizzi (2002, vol. 2) e Pedregal (2004).

3.1 Noções preliminares

Seja f uma função definida em um conjunto aberto $U \subset \mathbb{R}^n$ em $\mathbb{R}$. As funções derivadas parciais de f em relação às variáveis $x_1, \ldots x_n$ no ponto $a = (a_1, \ldots, a_n)$ são

definidas por

$$\frac{\partial f}{\partial x_i}(a) = \lim_{\Delta x_i \to 0} \frac{f(a_1, \ldots, a_i + \Delta x_i, \ldots, a_n) - f(a_1, \ldots, a_i, \ldots, a_n)}{\Delta x_i},$$

para $i = 1, \ldots, n$.

Se denotarmos por $f_i(a) = \dfrac{\partial f}{\partial x_i}(a)$ segue que as funções derivadas parciais de f_i em relação a variável x_j no ponto $a = (a_1, \ldots, a_n)$, ditas derivadas parciais de segunda ordem, serão definidas por

$$f_{ij}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a) = \frac{\partial f_i}{\partial x_j}(a) = \lim_{\Delta x_j \to 0} \frac{f_i(a_1, \ldots, a_j + \Delta x_j, \ldots, a_n) - f_i(a_1, \ldots, a_j, \ldots, a_n)}{\Delta x_j},$$

para $i, j = 1, \ldots, n$.

De maneira análoga definimos as derivadas parciais de ordem k, com $k \geqslant 3$.

Definição 3.1 (Diferenciabilidade)**.** Seja $f : U \subset \mathbb{R}^n \to \mathbb{R}$ uma função definida em um conjunto aberto $U \subset \mathbb{R}^n$. Diremos que f é diferenciável no ponto $a = (a_1, \ldots, a_n) \in U$ se

(a) Existem as derivadas parciais $\dfrac{\partial f}{\partial x_i}(a)$ para todo $i = 1, \ldots n$;

(b) Para todo $v = (v_1, \ldots, v_n)$ satisfazendo $a + tv \in U$, $t \in \mathbb{R}$ temos

$$f(a + tv) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a).tv_i + r(tv), \text{ com } \lim_{t \to 0} \frac{r(tv)}{\| tv \|} = 0.$$

A função f é dita diferenciável no aberto U se f for diferenciável em todos os pontos de U.

Definição 3.2 (Derivada Direcional)**.** Sejam $f : A \subset \mathbb{R}^n \to \mathbb{R}$, A aberto e $\vec{v} \in \mathbb{R}^n$ um vetor qualquer. A derivada direcional de f no ponto $a \in A$ na direção do vetor $\vec{v}$ é dada por:

$$\frac{\partial f}{\partial \vec{v}}(a) = \lim_{t \to 0} \frac{f(a + tv) - f(a)}{t} \text{ (t } \in \mathbb{R}) \text{ quando esse limite existir.}$$

Definição 3.3 (Vetor Gradiente - Generalização)**.** Seja $f : U \to \mathbb{R}$, definida e diferenciável no conjunto aberto $U \subset \mathbb{R}^n$. O vetor gradiente de f no ponto $a \in U$ é definido como sendo $\text{grad} f(a) = \nabla f(a) = \left(\dfrac{\partial f}{\partial x_1}(a), \cdots, \dfrac{\partial f}{\partial x_n}(a) \right)$.

Propriedades do vetor gradiente Sejam f uma função definida em um conjunto aberto $U \subset \mathbb{R}^n$ em $\mathbb{R}$ e $a \in U$. Segue que:

Propriedade I: O gradiente sempre irá apontar para a direção segundo a qual a função f cresce.

Propriedade II: A direção dada pelo ∇f será aquela de mais rápido crescimento da função f dentre todas as direções nas quais a função f cresce,

Propriedade III: O vetor gradiente da função f no ponto $a \in U$ será sempre ortogonal à curva de nível de f que passa pelo ponto a.

Para demonstração dessas propriedades ver [5].

Teorema 3.4. *Se $f : U \subset \mathbb{R}^n \to \mathbb{R}$ é diferenciável em $a \in U$, com U aberto, então f possui derivada direcional em qualquer direção $\vec{v} = (v_1, \ldots, v_n)$ e vale $\dfrac{\partial f}{\partial \vec{v}}(a) = \langle \nabla f(a), v \rangle$.*

Demonstração. Como f diferenciável no ponto a, para t suficientemente pequeno de forma que $a + tv \in U$, segue que

$$f(a+tv) = f(a) + \frac{\partial f}{\partial x_1}(a)tv_1 + \ldots + \frac{\partial f}{\partial x_n}(a)tv_n + r(tv), \text{ com } \lim_{t\to 0} \frac{r(tv)}{||tv||} = 0. \tag{3.1}$$

Dividindo (3.1) por t, temos:

$$\frac{f(a+tv) - f(a)}{t} = \frac{\partial f}{\partial x_1}(a)v_1 + \ldots + \frac{\partial f}{\partial x_n}(a)v_n + \frac{r(tv)}{t}. \tag{3.2}$$

Calculando limite de (3.2) quando t tende a zero temos:

$$\lim_{t\to 0} \frac{f(a+tv) - f(a)}{t} = \lim_{t\to 0} \langle \nabla f(a), v \rangle + \lim_{t\to 0} \frac{r(tv)}{t}.$$

Observamos que

$$\lim_{t\to 0} \frac{r(tv)}{t} = \lim_{t\to 0} \frac{r(tv)\ \|v\|}{t\ \|v\|} = \|v\| \lim_{t\to 0} \frac{r(tv)}{\|tv\|} = 0.$$

Portanto, $\dfrac{\partial f}{\partial \vec{v}}(a) = \langle \nabla f(a), v \rangle$ □

3.2 Máximos e Mínimos

Considere a função $f : U \subset \mathbb{R}^n$ em $\mathbb{R}$.

Definição 3.5 (Máximo e mínimo local)**.** Dizemos que f tem máximo (mínimo) local no ponto $a \in U$ se existir $\delta > 0$ tal que para todo $v \in U$ com $\|v\| < \delta$ segue que $f(a+v) \leqslant f(a)$ (mínimo $f(a) \leqslant f(a+v)$).

Definição 3.6 (Máximo e mínimo global)**.** Dizemos que um ponto $a \in U$ é máximo (mínimo) global de f se $f(v) \leqslant f(a)$ (mínimo $f(a) \leqslant f(v)$) para todo $v \in U$.

Se $a \in U$ é ponto de máximo (mínimo), o valor $f(a)$ será o valor máximo (mínimo) de f em U. Assim, os pontos de máximo (mínimo) de f denominam-se extremantes de f.

3.2.1 Condições necessárias para que um ponto interior seja um extremante local

O teorema a seguir fornece um críterio para selecionarmos, dentre os pontos interiores do domínio da f, os candidatos a extremantes.

Teorema 3.7. *Sejam a um ponto interior de $U \subset \mathbb{R}^n$ e $f : U \to \mathbb{R}$ uma função. Suponhamos que todas as derivadas parciais de f existam, isto é, $\dfrac{\partial f}{\partial x_i}(a)$ existe para $i = 1, \ldots, n$. Assim, uma condição necessária para que $a \in U$ seja um extremante local de f é que $\dfrac{\partial f}{\partial x_i}(a) = 0$ para todo $i = 1, \ldots, n$.*

Demonstração. Sejam $f : U \subset \mathbb{R}^n \to \mathbb{R}$ e $a = (a_1, \ldots, a_n) \in U$ um ponto máximo (ou mínimo) local de f. Para todo $i = 1, \ldots, n$ existe $\delta > 0$ tal que a função $g_i : (-\delta, \delta) \to \mathbb{R}$ definida por $g_i(x) = f(a_1, \ldots, a_{i-1}, a_i + x, a_{i+1}, \ldots, a_n) = f(a + xe_i)$. É fácil verificar que

(a) g_i é bem definida.

(b) g_i é derivável em $x = x_i$, pois $g_i'(x_i) = \dfrac{\partial f}{\partial x_i}(a_1, \ldots, a_{i-1}, a_i + x_i, \ldots, a_n) = \dfrac{\partial f}{\partial x_i}(a + xe_i)$.

(c) a é ponto de máximo (ou mínimo) local no ponto $x = 0$ de g_i ($g_i(0) = f(a)$). Segue que $g_i'(0) = 0$ e consequentemente $\dfrac{\partial f}{\partial x_i}(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) = \dfrac{\partial f}{\partial x_i}(a) = 0$.

Analogamente para as demais derivadas parciais. Portanto, se a é ponto de máximo (ou mínimo) local segue que $\dfrac{\partial f}{\partial x_i}(a) = 0$ para todo $i = 1, \ldots, n$. □

Definição 3.8 (Ponto Crítico)**.** Um ponto $a \in U$ é chamado de ponto crítico de $f : U \to \mathbb{R}$ quando $\dfrac{\partial f}{\partial x_1}(a) = \dfrac{\partial f}{\partial x_2}(a) = \cdots = \dfrac{\partial f}{\partial x_n}(a) = 0$.

3.2.2 Condição suficiente para um ponto crítico ser extremante local

Seja f de U em $\mathbb{R}$ uma função de classe C^2 definida no conjunto aberto $U \subset \mathbb{R}^n$. A forma quadrática hessiana da função f no ponto $x \in U$ é dada por

$$H(x) \cdot v^2 = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x).v_i v_j,$$

para todo $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$.

A condição suficiente para um ponto crítico de f ser um extremante local de f será dada pelo resultado a seguir.

Teorema 3.9. *Sejam $f : U \subset \mathbb{R}^n \to \mathbb{R}$ de classe C^2 e $a \in U$ um ponto interior do domínio de f ponto crítico de f Então,*

(a) Se a forma quadrática hessiana $H(a) > 0$ então a será ponto de mínimo local de f.

(b) Se a forma quadrática hessiana $H(a) < 0$ então a será ponto de máximo local de f.

(c) Se a forma quadrática hessiana $H(a)$ for indefinida então nada podemos dizer.

Para a demonstração ver [5].

3.3 Método dos Multiplicadores de Lagrange

O Método dos Multiplicadores de Lagrange tem por objetivo determinar os pontos críticos de uma função $f : U \subset \mathbb{R}^{n+1} \to \mathbb{R}$ de classe C^1 definida no aberto U restrita a pontos de um conjunto $M = \varphi^{-1}(c)$ sendo que c é a imagem inversa do valor regular da função $\varphi : U \to \mathbb{R}$ também de classe C^1. Como Lima (2013, p.91) ressalta "não se trata de determinar os pontos críticos de $f : U \to \mathbb{R}$ que estão localizados sobre a hiperfície M, mas sim os pontos críticos da função $f|M : M \to \mathbb{R}$".

Teorema 3.10. *Seja $f : U \subset \mathbb{R}^{n+1} \to \mathbb{R}$ de classe C^1 definida no aberto U restrita a pontos de um conjunto $M = \varphi^{-1}(c)$ com c a imagem inversa do valor regular da função $\varphi : U \to \mathbb{R}$ também de classe C^1. O ponto $x \in U$ será um ponto crítico de $f|M : M \to \mathbb{R}$ com $M = \varphi^{-1}(c)$ se, e somente se,*

(a) $\varphi(x) = c$ *e*

(b) $\nabla f(x) = \lambda \cdot \nabla\varphi(x)$ *para algum* $\lambda \in \mathbb{R}$.

λ é denominado de multiplicador de Lagrange.

Para a demonstração ver [5].

O Método dos Multiplicadores de Lagrange para problemas de otimização não linear não é muito utilizado em problemas reais que envolvem funções não lineares, por existirem algoritmos mais eficientes. Contudo, é uma ferramenta fundamental para o entendimento da Programação Não Linear.

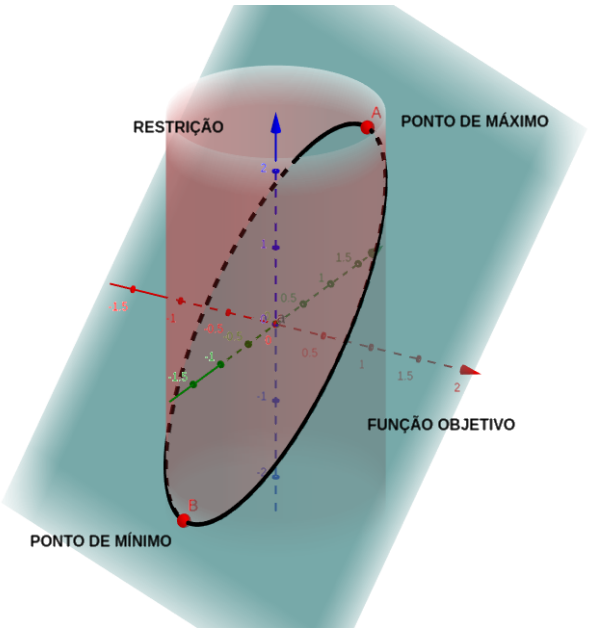
Exemplo 3.11. Determine os candidatos a extremantes de $f(x, y) = x + 2y$ com a restrição $x^2 + y^2 = 1$.

Seja $g(x, y) = x^2 + y^2 - 1$. Queremos determinar os extremantes de f em $B = \{(x, y) \in \mathbb{R}^2 / g(x, y) = 0\}$. Como g é de classe C^1 e $\nabla g(x, y) = (2x, 2y) \neq (0, 0)$ em B, resulta que os candidatos a extremantes locais são os pontos (x, y) que tornam compatível o sistema

$$\begin{cases} (1, 2) = \lambda(2x, 2y) \\ g(x, y) = 0 \end{cases} \quad \text{que é equivalente a} \quad \begin{cases} 1 = 2\lambda x \\ 2 = 2\lambda y \\ x^2 + y^2 = 1 \end{cases}$$

Como $\lambda \neq 0$ temos $x = \dfrac{1}{2\lambda}$ e $y = \dfrac{1}{\lambda}$. Substituindo em $x^2 + y^2 = 1$ segue que $\dfrac{1}{4\lambda^2} + \dfrac{1}{\lambda^2} = 1$. Logo, $\lambda = \pm\dfrac{\sqrt{5}}{2}$. Assim, $\left(\dfrac{\sqrt{5}}{5}, \dfrac{2\sqrt{5}}{5}\right)$ e $\left(-\dfrac{\sqrt{5}}{5}, -\dfrac{2\sqrt{5}}{5}\right)$ são os candidatos a extremantes locais. Como B é compacto e $f\left(\dfrac{\sqrt{5}}{5}, \dfrac{2\sqrt{5}}{5}\right) > f\left(-\dfrac{\sqrt{5}}{5}, -\dfrac{2\sqrt{5}}{5}\right)$, resulta que $\left(\dfrac{\sqrt{5}}{5}, \dfrac{2\sqrt{5}}{5}\right)$ é ponto de máximo e $\left(-\dfrac{\sqrt{5}}{5}, -\dfrac{2\sqrt{5}}{5}\right)$ é ponto de mínimo de f em B.

Figura 3.1: Máximo e mínimo - Exemplo 3.11.



Fonte: Elaborado pelo autor.

3.4 Condições Ótimas de Karush-Kuhn-Tucker

Nesta seção trataremos problemas de otimização do tipo:

$$\text{Minimize } f(x) \text{ sujeito a } g(x) \leqslant 0$$

onde $f : U \subset \mathbb{R}^n \to \mathbb{R}$, U conjunto aberto e $g(x) = [g_1(x), \ldots, g_m(x)]^T$ com $g_i : U \subset \mathbb{R}^n \to \mathbb{R}$ para $i = 1, \ldots, m$. Suponhamos também que f e g são de classe C^1.

As condições necessárias de otimização, conhecidas como Condições de Karush-Kuhn-Tucker (KKT) são enunciadas no teorema a seguir. Vale observar que a demonstração deste resultado não é imediata e não será tratada neste trabalho.

Teorema 3.12. *Sejam f e $g_i, i = 1, \ldots, m$ funções reais definidas e diferenciáveis no aberto $U \subset \mathbb{R}^n$ no problema de otimização*

$$\textit{Minimizar } f(x) \textit{ sujeito a } g(x) \leqslant 0$$

onde $g(x) = [g_1(x), \ldots, g_m(x)]^T$. Suponhamos que a é uma solução ótima não singular do problema. Mais ainda, seja M o conjunto de índices $M = \{1, 2, ..., m\}$ e definimos

$$J(a) = \{j \in M : g_j(a) = 0\}.$$

tal que, o conjunto $\{\nabla g_j(a) = 0; j \in J(a)\}$ é linearmente independente. Então, existe um vetor de multiplicadores λ_j $j \in J(a)$ tal que

$$\nabla f(a) + \sum_{j \in J(a)} \lambda_j \nabla g(a) = 0.$$

Sempre que o problema de otimização for de maximização podemos fazer a seguinte alteração

$$\text{Maximizar } f(x) \text{ sujeito a } g(x) \leqslant 0 \Leftrightarrow \text{Minimizar } [-f(x)] \text{ sujeito a } g(x) \leqslant 0$$

Além disso, qualquer restrição não negativa $g_i(x) \geqslant 0$ pode ser convertida em uma restrição não positiva $-g_i(x) \leqslant 0$.

Vejamos alguns exemplos.

Exemplo 3.13. Minimizar $f(x,y) = x^2 + y^2 - 4x$

$$\text{sujeito a } \begin{cases} g_1(x,y) = 2 - xy \leqslant 0 \\ g_2(x,y) = x^2 + y - 5 \leqslant 0 \\ g_3(x,y) = -x \leqslant 0 \\ g_4(x,y) = -y \leqslant 0 \end{cases}$$

Inicialmente verificamos que $x = y = 0$ satisfaz a segunda, terceira e quarta restrição, mas não a primeira. Descartando a terceira e quarta restrição e considerando agora apenas a primeira e segunda restrição, segue que

$$\begin{cases} g_1(x,y) = 2 - xy = 0 \\ g_2(x,y) = x^2 + y - 5 = 0 \end{cases} \Leftrightarrow \begin{cases} 2 - xy = 0 \\ x^3 + xy - 5x = 0 \end{cases}$$

se, e somente se, $x^3 - 5x + 2 = 0$ que tem por soluções: $x = 2$ ou $x = \sqrt{2} - 1$ ou $x = -\sqrt{2} - 1$. A última solução não satisfaz a terceira restrição logo não será considerada.

Consideremos $x = 2$. Segue que $y = 1$. Assim, o ponto $a = (2,1)$ satisfaz $g_1(2,1) = g_2(2,1) = 0$. Vamos agora verificar que este ponto satisfaz as condições KKT. De fato, partindo das expressões:

$$\begin{aligned} \nabla f(x,y) &= (2x - 4, 2y) \\ \nabla g_1(x,y) &= (-y, -x) \\ \nabla g_2(x,y) &= (2x, 1) \end{aligned}$$

e substituindo na expressão

$$\nabla f(x,y) + \lambda_1 \nabla g_1(x,y) + \lambda_2 \nabla g_2(x,y) = 0.$$

Obtemos o sistema de equações

$$\begin{cases} 2x - 4 - \lambda_1 y + 2\lambda_2 x = 0 \\ 2y - \lambda_1 x + \lambda_2 = 0 \end{cases}$$

Substituindo os valores de $x = 2$ e $y = 1$ obtemos que $\lambda_1 = \dfrac{8}{7}$ e $\lambda_2 = \dfrac{2}{7}$. Desta maneira, o ponto $a = (2,1)$ satisfaz também as condições KKT. Analogamente, para $x = \sqrt{2} - 1$ temos que $y = \dfrac{2}{\sqrt{2} - 1}$. Substituindo esses valores no sistema acima obtemos

$$\lambda_1 = \frac{17\sqrt{2} - 25}{13\sqrt{2} - 18} \text{ e } \lambda_2 = \frac{-22\sqrt{2} + 27}{13\sqrt{2} - 18}.$$

Este ponto também satisfaz as condições de KKT.

Substituindo na função objetivo os pontos determinados verificamos que:

$$\begin{cases} f(2,1) = -3 \\ f(\sqrt{2} - 1, \dfrac{2}{\sqrt{2} - 1}) = \dfrac{32\sqrt{2} - 49}{2\sqrt{2} - 3} \end{cases}$$

Portanto, $a = (2,1)$ é ponto de mínimo da função f.

Exemplo 3.14. Consideremos agora uma corda de comprimento a que será utilizada para amarrar uma caixa em formato de paralelepípedo de cima para baixo ao longo das duas direções perpendiculares. Queremos determinar as dimensões de uma caixa que possua o maior volume possível.

Sejam x_1, x_2 e x_3 as dimensões relativas à largura, profundidade e altura, respectivamente. O volume da caixa é dado por

$$V(x_1, x_2, x_3) = x_1 x_2 x_3$$

e a corda deverá satisfazer as condições de restrições:

$$\begin{cases} x_1 \geqslant 0 \\ x_2 \geqslant 0 \\ x_3 \geqslant 0 \\ 2x_1 + 2x_2 + 4x_3 \leqslant a. \end{cases}$$

sendo a última condição referente a amarração da caixa. Vamos agora transformar o problema em nosso formato padrão. Sendo assim, temos

Minimizar $f(x_1, x_2, x_3) = -x_1 x_2 x_3$

$$\text{sujeito a } \begin{cases} g_1(x_1, x_2, x_3) = -x_1 \leqslant 0 \\ g_2(x_1, x_2, x_3) = -x_2 \leqslant 0 \\ g_3(x_1, x_2, x_3) = -x_3 \leqslant 0 \\ g_4(x_1, x_2, x_3) = 2x_1 + 2x_2 + 4x_3 - a \leqslant 0 \end{cases}$$

A primeira, segunda e terceira restrição acima não podem se anular simultaneamente, pois em qualquer caso nulo não definiria um volume. Sendo assim, devemos aplicar as condições KKT apenas para a quarta restrição. Partindo das expressões:

$$\nabla f(x_1, x_2, x_3) = (-x_2 x_3, -x_1 x_3, -x_1 x_2),$$
$$\nabla g_4(x_1, x_2, x_3) = (2, 2, 4),$$

vamos considerar as condições KKT, ou seja,

$$\nabla f(x_1, x_2, x_3) + \lambda_4 \nabla g_4(x_1, x_2, x_3) = 0.$$

Obtendo o seguinte sistema de equações:

$$\begin{aligned} -x_2 x_3 + 2\lambda_4 &= 0 \\ -x_1 x_3 + 2\lambda_4 &= 0 \\ -x_1 x_2 + 4\lambda_4 &= 0 \end{aligned}$$

Multiplicando a primeira equação por x_1 e a segunda equação por x_2 e subtraindo uma expressão da outra obtemos

$$2\lambda_4(x_2 - x_1) = 0 \Leftrightarrow \lambda_4 = 0 \text{ ou } x_2 = x_1.$$

Analogamente, multiplicando a primeira equação por x_1 e a terceira equação por x_3 e subtraindo uma expressão da outra obtemos

$$2\lambda_4(x_1 - 2x_3) = 0 \Leftrightarrow \lambda_4 = 0 \text{ ou } x_3 = \frac{x_1}{2}.$$

Como $\lambda_4 \neq 0$ substituimos as expressões de x_2 e x_3 em termos de x_1 na condição de restrição g_4, isto é,

$$2x_1 + 2x_2 + 4x_3 = a \Leftrightarrow 6x_1 = a \Leftrightarrow x_1 = x_2 = \frac{a}{6} \text{ e } x_3 = \frac{a}{12}.$$

Sendo este o ponto ótimo, pois satisfaz as condições KKT.

Exemplo 3.15. (Pedregal, 2004, p. 84) Determinar o mínimo de

$$f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3$$

sobre a região determinada pelas restrições

$$\begin{cases} x_1^2 + x_2^2 + x_3^2 \leqslant 4 \\ x_1 + x_2 + x_3 \leqslant 1 \end{cases}$$

Inicialmente transformamos o problema em nosso formato padrão.

Minimizar $f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3$

$$\text{sujeito a } \begin{cases} g_1(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 4 \leqslant 0, \\ g_2(x_1, x_2, x_3) = x_1 + x_2 + x_3 - 1 \leqslant 0. \end{cases}$$

E determinamos

$$\begin{aligned} &\nabla f(x_1, x_2, x_3) = (3x_1^2, 3x_2^2, 3x_3^2), \\ &\nabla g_1(x_1, x_2, x_3) = (2x_1, 2x_2, 2x_3), \\ &\nabla g_2(x_1, x_2, x_3) = (1, 1, 1). \end{aligned}$$

Em seguida escrevemos as condições KKT, ou seja,

$$\nabla f(x_1, x_2, x_3) + \lambda_1 \nabla g_1(x_1, x_2, x_3) + \lambda_2 \nabla g_2(x_1, x_2, x_3) = 0.$$

obtendo o seguinte sistema de equações:

$$\begin{aligned} 3x_1^2 + \lambda_1 2x_1 + \lambda_2 &= 0, \\ 3x_2^2 + \lambda_1 2x_2 + \lambda_2 &= 0, \\ 3x_3^2 + \lambda_1 2x_3 + \lambda_2 &= 0, \\ x_1^2 + x_2^2 + x_3^2 &\leqslant 4, \\ x_1 + x_2 + x_3 &\leqslant 1. \end{aligned}$$

Vamos discutir as possíveis soluções para o sistema acima considerando os seguintes casos:

(1) $\lambda_1 = \lambda_2 = 0$: Neste caso, temos $3x_1^2 = 3x_2^2 = 3x_3^2 = 0$ implicando na solução $x_1 = x_2 = x_3 = 0$, que é admissível tanto para o máximo como para o mínimo;

(2) $\lambda_1 = 0$: Segue das condições KKT que

$$\begin{aligned} 3x_1^2 + \lambda_2 &= 0, \\ 3x_2^2 + \lambda_2 &= 0, \\ 3x_3^2 + \lambda_2 &= 0, \end{aligned}$$

ou seja, $x_1^2 = x_2^2 = x_3^2 = -\dfrac{\lambda_2}{3}$.

Analisemos agora os subcasos:

(a) Se $x_1^2 + x_2^2 + x_3^2 = 4$ segue que $3\left(-\frac{\lambda_2}{3}\right) = 4 \Leftrightarrow \lambda_2 = -4$. Temos então todas as possíveis combinações de valores e sinais entre $\pm\frac{2}{\sqrt{3}}$ para as coordenadas x_1, x_2 e x_3. Observamos, considerando a restrição $x_1 + x_2 + x_3 \leqslant 1$, que:

- $\frac{2}{\sqrt{3}} + \frac{2}{\sqrt{3}} + \frac{2}{\sqrt{3}} = 2\sqrt{3} > 1$ não satifaz;
- $\frac{2}{\sqrt{3}} + \frac{2}{\sqrt{3}} - \frac{2}{\sqrt{3}} = \frac{2}{\sqrt{3}} > 1$ não satisfaz;
- $\frac{2}{\sqrt{3}} - \frac{2}{\sqrt{3}} - \frac{2}{\sqrt{3}} = -\frac{2}{\sqrt{3}} < 1$ satisfaz;
- $-\frac{2}{\sqrt{3}} - \frac{2}{\sqrt{3}} - \frac{2}{\sqrt{3}} = -2\sqrt{3}$ satisfaz.

Logo, os pontos serão:

$$\left(\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}\right),$$
$$\left(-\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}, \frac{2}{\sqrt{3}}\right),$$
$$\left(-\frac{2}{\sqrt{3}}, \frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}\right),$$
$$\left(-\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}\right).$$

(b) Supomos que $x_1 + x_2 + x_3 = 1$: Vimos que $x_i = \pm\sqrt{-\frac{\lambda_2}{3}}$ para $i = 1, 2, 3$. Novamente, consideramos todas as possíveis combinações de valores e sinais para as coordenadas x_1, x_2 e x_3 teremos:

- $\sqrt{-\frac{\lambda_2}{3}} + \sqrt{-\frac{\lambda_2}{3}} + \sqrt{-\frac{\lambda_2}{3}} = 3\sqrt{-\frac{\lambda_2}{3}} = 1 \Leftrightarrow \lambda_2 = -\frac{1}{3}$. Neste caso, $x_1 = x_2 = x_3 = \frac{1}{3}$ e a restrição $x_1^2 + x_2^2 + x_3^2 = \frac{1}{3} \leqslant 4$ é satisfeita. O ponto é dado por:

$$\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right).$$

- $\sqrt{-\frac{\lambda_2}{3}} + \sqrt{-\frac{\lambda_2}{3}} - \sqrt{-\frac{\lambda_2}{3}} = 1$. Logo, $\lambda_2 = -3$ e $x_i^2 = 1$ para $i = 1, 2, 3$. Logo, a restrição $x_1^2 + x_2^2 + x_3^2 = 3 \leqslant 4$ está satisfeita. Logo, os pontos possíveis serão:

$$(1, 1, -1),$$
$$(1, -1, 1),$$
$$(-1, 1, 1).$$

- $\sqrt{-\frac{\lambda_2}{3}} - \sqrt{-\frac{\lambda_2}{3}} - \sqrt{-\frac{\lambda_2}{3}} = 1$. Este caso é impossível.
- $-\sqrt{-\frac{\lambda_2}{3}} - \sqrt{-\frac{\lambda_2}{3}} - \sqrt{-\frac{\lambda_2}{3}} = 1$. Este caso também é impossível.

(3) $\lambda_2 = 0$. Somando as três primeiras restrições de KKT segue que

$$3(x_1^2 + x_2^2 + x_3^2) + 2\lambda_1(x_1 + x_2 + x_3) = 0.$$

(a) Se $x_1^2 + x_2^2 + x_3^2 = 4$ segue que

$$12 + 2\lambda_1(x_1 + x_2 + x_3) = 0.$$

Observamos que esta identidade descarta a possibilidade de que $x_1 + x_2 + x_3 = 0$. Logo,

$$\lambda_1 = -\frac{6}{x_1 + x_2 + x_3}.$$

Retornando com o valor de λ_1 em cada uma das três primeiras restrições obtemos

$$3x_i^2 - \frac{12}{x_1 + x_2 + x_3}x_i = 0 \Leftrightarrow 3x_i\left(x_i - \frac{4}{x_1 + x_2 + x_3}\right) = 0$$

cuja solução será $x_i = 0$ ou $x_i = \dfrac{4}{x_1 + x_2 + x_3}$ para $i = 1, 2, 3$.

Analisando as possibilidades segundo a restrição $x_1 + x_2 + x_3 \leqslant 1$ segue:

- $x_1 = x_2 = x_3 = 0$ não satisfaz.
- Duas componentes nulas e a outra $x_i = \dfrac{4}{x_1 + x_2 + x_3} = \dfrac{4}{x_i} \Leftrightarrow x_i^2 = 4$. Logo, $x_i = \pm 2$. Somente $x_i = -2$ satisfaz a restrição. Os pontos válidos serão:

$$(-2, 0, 0),$$

$$(0, -2, 0),$$

$$(0, 0, -2).$$

- Uma componente nula e as outras duas $x_i = x_j = \dfrac{4}{x_1 + x_2 + x_3} = \dfrac{4}{2x_i} \Leftrightarrow x_i^2 = 2$. Logo, $x_i = x_j = \pm\sqrt{2}$. Somente $x_i = -\sqrt{2}$ satisfaz a restrição. Os pontos serão:

$$(-\sqrt{2}, -\sqrt{2}, 0),$$
$$(-\sqrt{2}, 0, -\sqrt{2}),$$
$$(0, -\sqrt{2}, -\sqrt{2}).$$

- $x_1 = x_2 = x_3 = \dfrac{4}{x_1 + x_2 + x_3}$. Assim, $x_i = \dfrac{4}{3x_i} \Leftrightarrow x_i^2 = \dfrac{4}{3}$ teremos que $x_1 = x_2 = x_3 = \pm\dfrac{2}{\sqrt{3}}$. Somente $x_1 = x_2 = x_3 = \dfrac{2}{\sqrt{3}}$ satisfaz a restrição.O ponto será:

$$\left(\frac{-2}{\sqrt{3}}, \frac{-2}{\sqrt{3}}, \frac{-2}{\sqrt{3}}\right).$$

(b) Se $x_1 + x_2 + x_3 = 1$ segue que

$$3(x_1^2 + x_2^2 + x_3^2) + 2\lambda_1 = 0.$$

E assim, $\lambda_1 = -\frac{3}{2}(x_1^2 + x_2^2 + x_3^2)$. Retornando o valor de λ_1 nas três primeiras restrições KKT temos que:

$$x_1 = x_2 = x_3 = (x_1^2 + x_2^2 + x_3^2).$$

E considerando a restrição $x_1 + x_2 + x_3 = 1$ segue que

$$x_1^2 + x_2^2 + x_3^2 = \frac{1}{3}.$$

Logo, temos uma única possibilidade, $x_1 = x_2 = x_3 = \frac{1}{3}$. O ponto será dado por

$$\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

que satisfaz a restrição $x_1^2 + x_2^2 + x_3^2 \leqslant 4$.

(4) A última possibilidade a ser considerada será $x_1 + x_2 + x_3 = 1$ e $x_1^2 + x_2^2 + x_3^2 = 4$. Este caso envolve a igualdade de duas restrições e aplicamos o Método dos Multiplicadores de Lagrange, pois trata-se de um problema do tipo:

Minimizar $f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3$

$$\text{sujeito a } \begin{cases} g_1(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 = 4 \\ g_2(x_1, x_2, x_3) = x_1 + x_2 + x_3 = 1. \end{cases}$$

Assim, partindo de

$$\nabla f(x_1, x_2, x_3) + \lambda_1 \nabla g_1(x_1, x_2, x_3) + \lambda_2 \nabla g_2(x_1, x_2, x_3) = 0.$$

obtemos o seguinte sistema de equações:

$$\begin{aligned} 3x_1^2 + \lambda_1 2x_1 + \lambda_2 &= 0, \\ 3x_2^2 + \lambda_1 2x_2 + \lambda_2 &= 0, \\ 3x_3^2 + \lambda_1 2x_3 + \lambda_2 &= 0, \\ x_1^2 + x_2^2 + x_3^2 &= 4, \\ x_1 + x_2 + x_3 &= 1. \end{aligned}$$

Temos um sistema contendo cinco equações com cinco incógnitas: x_1, x_2, x_3, λ_1 e λ_2. Reescrevendo na forma matricial temos

$$\begin{pmatrix} 3x_1^2 & 2x_1 & 1 \\ 3x_2^2 & 2x_2 & 1 \\ 3x_3^2 & 2x_3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\begin{aligned} x_1^2 + x_2^2 + x_3^2 &= 4, \\ x_1 + x_2 + x_3 &= 1. \end{aligned}$$

Para determinarmos λ_1 e λ_2 não simultaneamente nulos devemos impor que

$$\det \begin{pmatrix} 3x_1^2 & 2x_1 & 1 \\ 3x_2^2 & 2x_2 & 1 \\ 3x_3^2 & 2x_3 & 1 \end{pmatrix} = 0$$

Pela propriedade de determinante podemos fatorar 3 na primeira coluna e o valor 2 na segunda coluna obtendo o determinante de Vandermonde que sabemos é dado por

$$6 \cdot \det \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{pmatrix} = 6 \cdot (x_1 - x_2)(x_1 - x_3)(x_2 - x_3) = 0.$$

Logo, teremos as seguinte possibilidades:

(a) $x_2 = x_1$. Impondo a última restrição temos que $x_3 = 1 - 2x_1$. Substituindo esses valores na penúltima restrição determinamos o valor de x_1. De fato,

$$x_1^2 + x_2^2 + x_3^2 = 4 \Leftrightarrow 2x_1^2 + (1 - 2x_1)^2 = 4 \Leftrightarrow 6x_1^2 - 4x_1 - 3 = 0$$

que tem solução dada por

$$x_1 = \frac{1}{3} \pm \frac{1}{6}\sqrt{22}.$$

Os pontos serão dados por

$$\left(\frac{1}{3} + \frac{1}{6}\sqrt{22}, \frac{1}{3} + \frac{1}{6}\sqrt{22}, \frac{1}{3} - \frac{1}{3}\sqrt{22}\right),$$

$$\left(\frac{1}{3} - \frac{1}{6}\sqrt{22}, \frac{1}{3} - \frac{1}{6}\sqrt{22}, \frac{1}{3} + \frac{1}{3}\sqrt{22}\right).$$

(b) É imediato que os demais casos, $x_3 = x_1$ e $x_3 = x_2$ serão permutações do caso anterior. Assim, para $x_3 = x_1$ os pontos serão dados por

$$\left(\frac{1}{3} + \frac{1}{6}\sqrt{22}, \frac{1}{3} - \frac{1}{3}\sqrt{22}, \frac{1}{3} + \frac{1}{6}\sqrt{22}\right),$$

$$\left(\frac{1}{3} - \frac{1}{6}\sqrt{22}, \frac{1}{3} + \frac{1}{3}\sqrt{22}, \frac{1}{3} - \frac{1}{6}\sqrt{22}\right)$$

e para $x_3 = x_2$ os pontos serão dados por

$$\left(\frac{1}{3} - \frac{1}{3}\sqrt{22}, \frac{1}{3} + \frac{1}{6}\sqrt{22}, \frac{1}{3} + \frac{1}{6}\sqrt{22}\right),$$

$$\left(\frac{1}{3} + \frac{1}{3}\sqrt{22}, \frac{1}{3} - \frac{1}{6}\sqrt{22}, \frac{1}{3} - \frac{1}{6}\sqrt{22}\right).$$

Por fim, substituindo todos os pontos determinados na função f e comparando-os concluimos que o mínimo da função f será obtido nos pontos $(-2, 0, 0)$, $(0, -2, 0)$ e $(0, 0, -2)$.

Considerações finais

Neste texto buscamos organizar e discutir alguns conceitos matemáticos envolvidos na Programação Linear e Não Linear, em especial o Método Simplex e as Condições de Karush-Kuhn-Tucker.

Observamos que a resolução de problemas de otimização possibilita explorar importantes conceitos como: sistemas de equações e de inequações, representações gráficas, curvas de nível, gradiente, diferenciabilidade, máximos e mínimos, entre outros.

Percebemos que o Método Simplex simplifica bastante a solução de problemas de Programação Linear e nos permite uma abordagem diferente sobre alguns temas da Matemática.

Vimos que grande parte dos problemas de otimização não são lineares e aproveitamos para nos aprofundar em conceitos importantes das funções de várias variáveis, como curvas de nível, as condições necessárias para que um ponto interior seja extremante local, os Multiplicadores de Lagrange e as condições de KKT, pouco abordadas nos cursos de graduação.

O trabalho tratou de alguns temas matemáticos com intuito de conectar conteúdos da Matemática do Ensino Superior, podendo ser utilizado como referência em estudos futuros sobre o tema e aprofundado através de outras abordagens. Como sugestão, verificar em [4] as condições suficientes de KKT.

Para verificação de algumas das soluções dos problemas enunciados foram utilizadas algumas ferramentas computacionais disponíveis na literatura.

As referências bibliográficas estão ordenadas de acordo com as citações.

Referências

[1] ARENALES, M. et al. *Pesquisa Operacional*. 1. ed. Rio de Janeiro: Elsevier Editora Ltda., 2007.

[2] GOLDBARG, E.; GOLDBARG, M. *Programação Linear e Fluxo em Redes*. 1. ed. Rio de Janeiro: Elsevier Editora, 2015.

[3] KOLMAN, B. *Introdução à Algebra Linear com Aplicações*. 6. ed. Rio de Janeiro: LTC Editora S.A., 1998.

[4] PEDREGAL, P. *Introduction to Optimization*. New York: Springer-Verlag, 2004.

[5] LIMA, E. L. *Análise Real - Volume 2*. 6. ed. Rio de Janeiro: IMPA, 2013.

[6] BASSANEZI, R. C. *Ensino - aprendizagem com Modelagem matemática*. 3. ed. São Paulo: Editora Contexto, 2002.

[7] GUIDORIZZI, H. L. *Um curso de Cálculo - Volume 1*. 5. ed. Rio de Janeiro: Livros Técnicos e Científicos Editora S.A., 2002.

[8] GUIDORIZZI, H. L. *Um curso de Cálculo - Volume 2*. 5. ed. Rio de Janeiro: Livros Técnicos e Científicos Editora S.A., 2002.

[9] MARINS., F. A. S. *Introdução à Pesquisa Operacional*. São Paulo: Cultura Acadêmica (Universidade Estadual Paulista), 2011.

[10] MATOUSEK, J. *Understanding and Using Linear Programming*. 6. ed. Rio de Janeiro: Springer-Verlag Berlin Heidelberg, 2007.

[11] MCLONE, R. R. *Mathematical Modelling - The art of applying mathematics, in Math. Modelling*. London: Butterwords, 1976.