

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CURSO DE LICENCIATURA EM MATEMÁTICA

FABIO AUGUSTO FORTUNATO FILHO

CONJUNTO SIMPLÉTICOS E OTIMIZAÇÃO: MÉTODOS
SIMPLEX E NELDER-MEAD

TRABALHO DE CONCLUSÃO DE CURSO

CURITIBA
2020

FABIO AUGUSTO FORTUNATO FILHO

**CONJUNTO SIMPLÉTICOS E OTIMIZAÇÃO: MÉTODOS
SIMPLEX E NELDER-MEAD**

Trabalho de Conclusão de Curso apresentada como requisito parcial à obtenção do título Licenciado em Matemática, do Departamento Acadêmico de Matemática da Universidade Tecnológica Federal do Paraná.

Orientador: Prof. Dr. Rodolfo Gotardi Begiato

CURITIBA
2020



Ministério da Educação
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
Câmpus Curitiba
Diretoria de Graduação e Educação Profissional
Departamento Acadêmico de Matemática
Coordenação do curso de **Licenciatura em Matemática**



TERMO DE APROVAÇÃO

“Conjunto Simpléticos e Otimização: Métodos Simplex e Nelder-Mead”

por

“Fabio Augusto Fortunato Filho”

Este Trabalho de Conclusão de Curso foi apresentado às **14h** do dia **3** de **dezembro** de **2020** na sala <https://meet.google.com/gho-iovb-bvc> como requisito parcial à obtenção do grau de Licenciado em Matemática na Universidade Tecnológica Federal do Paraná - UTFPR - Câmpus Curitiba. O(a) estudante foi arguido pela Banca de Avaliação abaixo assinados. Após deliberação, de acordo com o parágrafo 2º do art. 24 do Regulamento do Trabalho de Conclusão de Curso para os Cursos de Graduação da UTFPR, a Banca de Avaliação considerou o trabalho aprovado.

<hr/> Prof Dr Rodolfo Gotardi Begiato (Presidente - UTFPR/Curitiba)	<hr/> Prof. Dr. Adriano Verdério (Avaliador 2 - UTFPR/Curitiba)
<hr/> Profª Drª Priscila Savulski Ferreira (Avaliador 3 - UTFPR/Curitiba)	<hr/> Profª Drª Priscila Savulski Ferreira (Professor Responsável pelo TCC – UTFPR/Curitiba)
<hr/> Profª Drª Neusa Nogas Tocha (Coordenação do curso de Licenciatura em Matemática – UTFPR/Curitiba)	

“A Folha de Aprovação assinada encontra-se na Coordenação do Curso”

RESUMO

FILHO, Fabio Augusto Fortunato. Conjunto Simpléticos e Otimização: Métodos Simplex e Nelder-Mead. 2020. 79 f. Trabalho de Conclusão de Curso – Curso de Licenciatura em Matemática, Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

É comum a utilização de técnicas numéricas para a solução de problemas de minimizar ou maximizar funções matemáticas. Baseado nisso, o presente trabalho faz uma explanação de dois métodos simpléticos de otimização livres de derivação. Inicialmente foi feito um breve estudo no conceito da Análise Convexa para definirmos um conjunto simplético. Em seguida, foi realizada uma introdução aos problemas de Programação Linear e suas características. Para a resolução de tais problemas, foi abordado um estudo do Método Simplex, que consiste em minimizar uma função utilizando apenas operações matriciais básicas. Com um algoritmo do Método Simplex feito, o próximo passo da pesquisa foi estudar problemas de Programação Não Linear, para assim, introduzir o algoritmo do Método de Nelder-Mead. Com base neste algoritmo, foi apresentado um estudo teórico do método, assim como demonstrado algumas propriedades. Destaca-se, neste estudo, a abordagem das vantagens e desvantagens de se utilizar o Método de Nelder-Mead, com uma pequena variação do método clássico da literatura.

Palavras-chave: Programação Linear. Programação Não Linear Irrestrita. Método Simplex. Método de Nelder-Mead.

ABSTRACT

FILHO, Fabio Augusto Fortunato. Symplectics Sets and Optimization: Simplex and Nelder-Mead Method. 2020. 79 f. Trabalho de Conclusão de Curso – Curso de Licenciatura em Matemática, Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

It's common to use numerical techniques for solving problems to minimize, or maximize mathematical functions. Based on this, the present work makes an explanation of two symplectical methods of derivative free optimization. Initially a brief study was made on the concept in convex analysis to define a symplectic set. Then, an introduction was made to the problems of linear programming and their characteristics. To solve these problems a study of the Simplex Method was approached, which consists of minimization an objective function using only basic matrix operations. With the algorithm of the Simplex Method done, the next step of this work was to study nonlinear programming problems to introduce the algorithm of the Nelder-Mead Method. About this algorithm, we present a theoretical study of the method, proving some properties and bringing their advantages and disadvantages to use the Nelder-Mead Method with a small variation of the classic method in the literature.

Keywords: Linear Programming. Unconstrained Nonlinear Programming. Simplex Method. Nelder-Mead Method.

LISTA DE FIGURAS

Figura 1 – Exemplo do Conjunto Simplético de dimensão 3, o 3 – <i>simplex</i>	8
Figura 2 – Iterações do Método de Nelder-Mead.	9
Figura 3 – Exemplos de conjuntos convexos.	11
Figura 4 – Exemplos de conjuntos não convexos.	12
Figura 5 – Exemplos de fecho de conjunto convexo e não convexo.	13
Figura 6 – Um subespaço paralelo a um conjunto afim.	16
Figura 7 – Vetores ortogonais em \mathbb{R}^2	17
Figura 8 – Vetor ortogonal a um subespaço no \mathbb{R}^2	17
Figura 9 – Intersecção de conjuntos afins.	21
Figura 10 – Fecho afim de um conjunto.	22
Figura 11 – Exemplo de simplex em \mathbb{R}^2 e \mathbb{R}^3	23
Figura 12 – Exemplo geométrico das restrições.	34
Figura 13 – Intersecções de inequações.	41
Figura 14 – Intersecções de inequações com vetores.	41
Figura 15 – Função objetivo não linear	57
Figura 16 – Exemplos de centróides em \mathbb{R}^2 e \mathbb{R}^3	58
Figura 17 – Exemplos do processo de nós no \mathbb{R}^2	59
Figura 18 – Exemplo de encolhimento em \mathbb{R}^2	59
Figura 19 – Função de Himmeblau.	62
Figura 20 – Convergência do Método de Nelder-Mead na função de Himmeblau. . .	63
Figura 21 – Comparação do volume de um simplex usando contração externa no \mathbb{R}^2 . .	67
Figura 22 – Convergência do Método de Nelder-Mead em f dada em (23).	72
Figura 23 – Identificação em imagem utilizando o Método de Nelder-Mead.	74
Figura 24 – Intersecção de Inequações com vetores do PPL.	76

SUMÁRIO

1 – INTRODUÇÃO	7
2 – INTRODUÇÃO À GEOMETRIA SIMPLÉTICA	11
2.1 Conjuntos Convexos	11
2.2 Conjuntos Afins	13
2.3 Simplex	23
3 – PROBLEMAS DE PROGRAMAÇÃO LINEAR	25
3.1 O Problema de Programação Linear	25
3.2 Soluções Básicas	28
3.3 O Teorema Fundamental da Programação Linear	29
3.4 Método Simplex	34
3.4.1 Pivoteamento	34
3.4.2 Interpretação Geométrica	40
3.4.3 Encontrar o Mínimo da Função Objetivo	42
3.4.4 Método Simplex de Duas Fases	47
3.4.5 Introdução ao Algoritmo do Método Simplex Revisado	50
3.4.5.1 Representação Matricial do Método Simplex	50
3.4.5.2 Algoritmo do Método Simplex Revisado	52
3.4.6 Aplicação do Método Simplex	54
4 – PROBLEMAS IRRESTRITOS DE PROGRAMAÇÃO NÃO LINEAR	56
4.1 Método de Nelder-Mead	57
4.1.1 Minimizando Funções com o Algoritmo do Método de Nelder-Mead	60
4.1.2 Nelder-Mead e Simplex	63
4.1.3 Ordem Lexicográfica Decrescente	70
4.1.4 Convergência do Método de Nelder-Mead	72
4.1.5 Aplicações do Método de Nelder-Mead	73
5 – CONSIDERAÇÕES FINAIS	75
Referências	78

1 INTRODUÇÃO

Otimização é uma área da Matemática Aplicada onde se desenvolve o estudo de problemas em que se busca minimizar ou maximizar uma função, com alguns critérios predefinidos, dentro de um conjunto preestabelecido. Diante disso, utilizaremos a Pesquisa Operacional (PO) como um recurso para analisar sistemas complexos do mundo real, a fim de otimizar a performance. A ideia central da PO utiliza modelos matemáticos e estatísticos, que com as técnicas adequadas, ajudam à tomada de decisão. Os dois principais problemas da PO, que serão abordados ao longo desse texto são: o Problema de Programação Linear e o Problema Irrestritos de Programação Não Linear. Um problema de otimização pode ser representado da seguinte maneira:

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & x \in D \end{array} \quad (1)$$

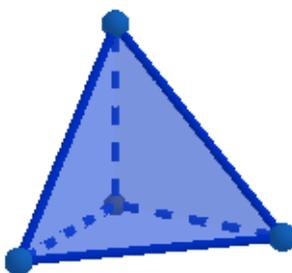
sendo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $D \subseteq \mathbb{R}^n$. Tal formulação é chamada de um problema de otimização ou um problema de programação matemática. Muitos problemas do mundo real e teóricos podem ser modelados nessa estrutura geral. A função f é chamada de função objetivo (ou função de custo), uma vez que temos o objetivo de maximizá-la ou minimizá-la. Por convenção, a forma padrão de um problema de otimização é definida em termos de minimização¹. O conjunto D é chamado de conjunto factível ou conjunto viável. Um ponto que pertence ao conjunto D é chamado solução factível ou ponto factível, uma solução que minimiza a função objetivo será chamada de uma solução ótima factível ou simplesmente solução ótima. Desse modo, os Problemas de Programação Linear, consistem em problemas com f uma função linear e D um conjunto definido por um sistema de equações lineares. Já os Problemas Irrestritos de Programação Linear, a função objetivo f pode, ou não, ser linear. E os tais problemas são considerados irrestritos, pois $D = \mathbb{R}^n$.

Segundo Friedlander (1994 apud BAZARAA; SHETTY; SHERALI, 1979), em problemas irrestritos, se a função for continuamente diferenciável, temos que os pontos de máximo ou mínimo tem o gradiente da função avaliado no ponto igual a zero. Podemos mencionar dois métodos que utilizando fortemente a busca pelo gradiente igual a zero, sendo eles o Método de Newton ou o Método do Gradiente. Existem diversas funções que, ou não possuem derivadas, ou as derivadas são complicadas de serem utilizadas. Para isso são utilizados os métodos livres de derivadas, neste trabalho, abrangearemos dois métodos de resolução de problemas dessa forma. Sendo eles, o Método Simplex e o Método de Nelder-Mead. Ambos os métodos utilizam a mesma geometria para o embasamento teórico, denominada Geometria Simplética.

¹Maximizar a função f é equivalente a minimizar a função $-f$.

A Geometria Simplética trata de simpléticos (ou simplex) que é uma generalização, para n dimensões, do conceito de triângulo em \mathbb{R}^2 . Por exemplo, utilizaremos um conjunto com 4 pontos disjuntos, a fim de construir um conjunto simplético de dimensão 3. Chamaríamos de 3 – *simplex*. Cada ponto do conjunto é um vértice de um tetraedro, conforme a Figura 1, generalizando, utilizaremos um conjunto de $n + 1$ pontos para construir um simplético de dimensão n .

Figura 1 – Exemplo do Conjunto Simplético de dimensão 3, o 3 – *simplex*.



Fonte: Autoria própria.

sendo assim, a Geometria Simplética é muito estudada por suas aplicações na Geometria Complexa, Topologia e Física Matemática. Utilizando a Geometria Simplética, além de grandes técnicas de otimização, pode ser construída uma relação intermediária entre a Geometria Diferencial e a Álgebra Não-Comutativa (IMPA, 2020).

Os métodos livres de derivações são extremamente importantes, uma vez que existem diversos casos de funções em que a derivada não é acessível, quer pelo alto custo computacional quer por sua indisponibilidade, mas mesmo assim temos o interesse de encontrar máximos ou mínimos de tais funções. Segundo Conn,

[...] consideramos a otimização sem derivadas uma das mais importantes, abertas e desafiadoras áreas na ciência computacional e engenharia, com um enorme potencial prático. O motivo pelo qual é desafiador é que, do ponto de vista da otimização, desiste-se de tanta informação por não ter derivadas (CONN; SCHEINBERG; VICENTE, 2009, p. 1, tradução nossa)

Vimos então como o campo de pesquisa em otimização livre de derivadas é uma área com grande demanda.

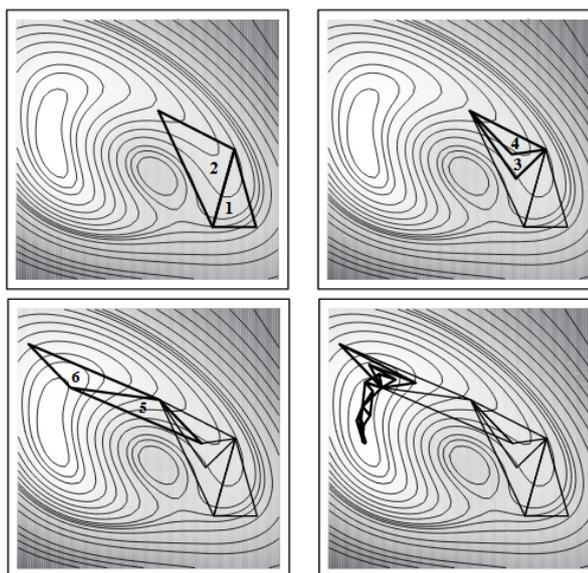
O desenvolvimento metodológico mais importante do período pós-guerra foi o Método Simplex, criado por George B. Dantzig, em 1947. Este método busca resolução de Problemas de Programação Linear. Isto é, problemas de planejamento nos quais são utilizados modelos de otimização lineares. Segundo Dantzig (1987), as ideias do Método Simplex já estavam em sua cabeça durante a Segunda Guerra Mundial, mas só foi

melhor elaborado em 1947 quando Dantzig viajou buscando aprimoramento em outras universidades. Dessa forma, o primeiro artigo publicado utilizando o Método Simplex foi em 1947 (DANTZIG, 1990).

Para o Método Simplex, é necessário que f seja uma função linear e D um conjunto formado por igualdades e desigualdades lineares que, geometricamente, corresponde a intersecções de semi-hiperplanos e retas. Esse método tem uma grande importância, pois resolve problemas de grandes dimensões apenas utilizando operações matriciais básicas. Uma interpretação do método é dada pelo resultado que a intersecção dos semi-hiperplanos geram. Como o conjunto viável D é um polítopo², se o problema tem solução mínima, é possível mostrar que uma solução mínima pode ser encontrada em um dos vértices do polítopo. Dessa forma, veremos que o Método Simplex avalia, a cada iteração, o valor da função em um dos vértices até encontrar o vértice que gera o menor valor da função.

O segundo método que estudaremos é o Método de Nelder-Mead, o qual busca a resolução de Problemas Irrestritos de Programação Não Linear. Este método consiste em avaliar a função em $n + 1$ pontos iniciais dados, tais que esses pontos geram um $n - simplex$. Organizando os $n + 1$ valores obtidos na função f de forma crescente, o ponto responsável pelo maior valor avaliado será substituído criando um novo conjunto de pontos. Logo, criará um novo $n - simplex$. A Figura 2, obtida de Givens e Hoeting (2013) mostra intuitivamente como o processo iterativo ocorre.

Figura 2 – Iterações do Método de Nelder-Mead.



Fonte: Givens e Hoeting (2013)

Perceba que, a cada iteração do processo, um novo triângulo é gerado, em que esse triângulo tem seus vértices mais próximos do mínimo da função. Outra observação

²Um polítopo é uma região contida em \mathbb{R}^n , o qual é resultante da intersecção de um conjunto de semi-espacos.

considerável é que mesmo tendo um $2 - simplex$ próximo de um mínimo local, o Método de Nelder-Mead pode convergir para outro mínimo da função objetivo.

O artigo de Nelder e Mead (1965) traz um algoritmo que é capaz de encontrar mínimos de uma função objetivo. Entretanto, a demonstração da convergência do método para o mínimo da função não é realizada no artigo e, além disso, é um dos problemas ainda em estudo referente ao método. Segundo Gonçalves (2013), existem variações do Método de Nelder-Mead em que a convergência é garantida, onde as hipóteses iniciais excluem algumas famílias de funções. Contudo, não abordaremos essas variações.

Neste contexto, o objetivo principal deste presente trabalho é desenvolver um estudo baseado na geometria simplética e avaliar como os elementos ali encontrados permitiram o desenvolvimeto do Método Simplex e no Método de Nelder-Mead, para resolução de Problemas de Programação Linear e Não Linear, respectivamente. Os dois métodos tem grande utilização nas mais diversas áreas de conhecimentos e são bem consolidados como alternativa para resolver problemas de minimização. Segundo Dantzig (1987), o Método Simplex, além de ser um dos mais utilizados métodos computacionais em diferentes áreas de estudo, é um dos métodos mais utilizados em otimização empresariais, uma vez que o objetivo do método é minimizar custos. A autora Gonçalves (2013) afirma que o Método de Nelder-Mead é o mais utilizado para a resolução de Problemas de Otimização Não Lineares. Afirmando, assim, a importância do estudo dos dois métodos clássicos de busca direta, livres de derivadas. O presente trabalho também tem o diferencial de introduzir um breve estudo da Análise Convexa para a definição formal de um Conjunto Simplex, juntamente com os métodos que utilizam tal geometria.

A monografia está formada por cinco capítulos. No Capítulo 2, introduziremos a Geometria Simplética. Isto é, definiremos os conjuntos convexos, os conjuntos afim e, por último, os conjuntos simpléticos. Abordaremos os Problemas de Programação Linear no Capítulo 3, apresentando algumas soluções básicas. Demonstraremos, assim, o Teorema Fundamental da Programação Linear. Ainda neste capítulo, apresentaremos o Método Simplex, assim como duas variações desse método: o Método Simplex de Duas Fases e o Método Simplex Revisado. Para este último método, apresentaremos um algoritmo. Finalizaremos o capítulo apresentando uma sistematização do processo obtido pelo Método Simplex. No Capítulo 4, discutiremos os Problemas de Programação Não Linear e a resolução de seus problemas através do Método de Nelder-Mead. Apresentaremos, dessa forma, o algoritmo do método com o objetivo de minimizar funções, assim como algumas propriedades, mas não afirmaremos a convergência do método para o mínimo da função objetivo. Concluiremos este quarto capítulo com algumas aplicações do Método de Nelder-Mead. Por fim, as considerações finais deste trabalho são apresentadas no Capítulo 5.

2 INTRODUÇÃO À GEOMETRIA SIMPLÉTICA

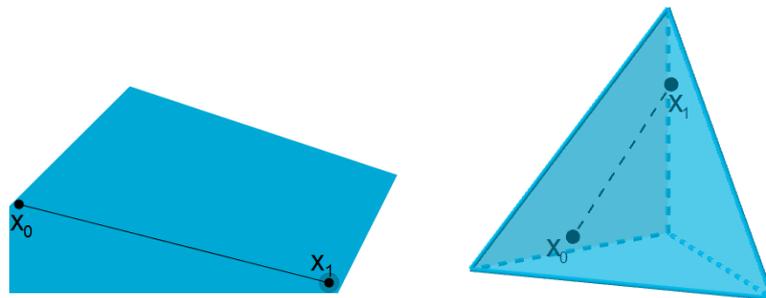
A definição do conceito de simplético exige a introdução de alguns conceitos analíticos, tais como convexidade e conjunto afim. Neste sentido, vamos destinar este capítulo para a exposição de elementos básicos da Análise Convexa, os quais fornecem um material teórico utilizado para desenvolver os métodos computacionais de otimização que utilizam simpléticos. Essa introdução teórica está baseada no livro *Convex Analysis* de Rockafellar (1972).

2.1 Conjuntos Convexos

Definição 1 (Conjunto Convexo). *Um conjunto $C \subset \mathbb{R}^n$ é dito convexo se, para quaisquer dois pontos de C , obtemos um segmento de reta que está contido em C . De outra forma, dados $x_0, x_1 \in C$ tem-se*

$$(1 - \delta)x_0 + \delta x_1 \in C \text{ para todo } \delta \in [0,1].$$

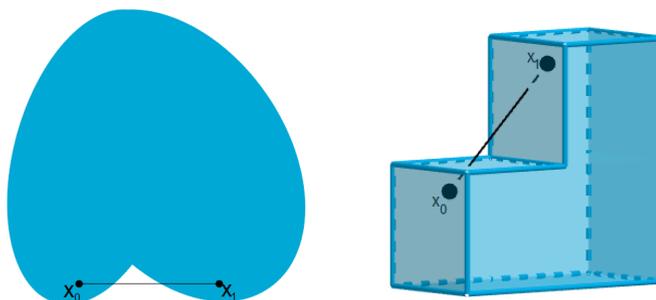
Figura 3 – Exemplos de conjuntos convexos.



Fonte: A autoria própria.

A Figura 16 apresenta dois conjuntos que satisfazem a definição de conjunto convexo. Os exemplos clássicos de conjuntos convexos são: os intervalos da reta real \mathbb{R} (abertos e/ou fechados), retas, planos, semiplanos e qualquer espaço vetorial que podem facilmente ser demonstrados utilizando a definição aqui apresentada. Já os conjuntos não convexos são conjuntos em que é possível obter ao menos dois pontos $x_0, x_1 \in C$ e um escalar δ , tais que $(1 - \delta)x_0 + \delta x_1 \notin C$. A Figura 4 abaixo ilustra exemplos desses conjuntos. A união de duas retas não coincidentes, a união de dois intervalos disjuntos e a união de um plano com uma reta que não a contém, são exemplos de conjuntos não convexos.

Figura 4 – Exemplos de conjuntos não convexos.



Fonte: Autoria própria.

Para prosseguir em nosso estudo, necessitaremos introduzir o conceito de fecho convexo. Para isso precisaremos do seguinte Teorema 2, o qual garante que a interseção de convexos resulta num conjunto convexo.

Teorema 2. *A intersecção de uma coleção arbitrária de conjuntos convexos é um conjunto convexo.*

Demonstração. Defina $M = \bigcap_{i=1}^{\infty} H_i$, com H_i conjuntos convexos para todo $i \in \mathbb{N}$. Precisamos mostrar que M é convexo.

Se M é formado por um único ponto ou M é um conjunto vazio, então é trivial verificar que o teorema é válido. Caso contrário, temos que se $x, y \in M$ então $x, y \in H_i$ para todo $i \in \mathbb{N}$. Da convexidade de cada H_i , temos que $\{(x - y)\lambda + y \mid \lambda \in \mathbb{R}\} \subset H_i$ para todo $i \in \mathbb{N}$.

Logo, podemos concluir que para quaisquer $x, y \in M$, segue que $\{(x - y)\lambda + y \mid \lambda \in \mathbb{R}\} \subset M$. Portanto, M é convexo. \square

Observação 3. *Notemos que, diferentemente do que ocorre com a intersecção, a união de convexos não necessariamente resulta em um conjunto convexo. Um exemplo, que podemos utilizar para concluir isso, é tomar os intervalos $(0,1)$ e $(2,3)$ da reta real. Notadamente, tomando $x = \frac{1}{2}$ e $y = \frac{5}{2}$ temos que $x, y \in (0,1) \cup (2,3)$, mas $\{(x - y)\lambda + y \mid \lambda \in [0,1]\} \not\subset (0,1) \cup (2,3)$.*

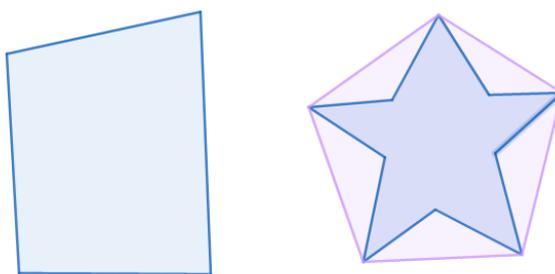
Definiremos a seguir o fecho convexo, que tem forte relação com o Teorema anterior, pois é possível provar que o fecho convexo de um conjunto S qualquer é a interseção de todos os conjuntos convexos que contenha S . Pode-se ainda provar que o fecho convexo é o menor conjunto convexo que contém S , no sentido que qualquer outro convexo que contenha S contém, também o fecho.

Definição 4 (Fecho Convexo). *O fecho convexo de um conjunto $S \subset \mathbb{R}^n$ é o conjunto*

$$\{x \in \mathbb{R}^n \mid x = \lambda_1 x_1 + \dots + \lambda_n x_n, x_i \in S, i = 1, \dots, n, \lambda_i \geq 0, i = 1, \dots, n, \sum_{i=1}^n \lambda_i = 1\}.$$

Observemos que foi necessário o teorema anterior para introduzir essa nova definição, uma vez que não faria sentido em chamar de fecho convexo, se não tivéssemos a certeza que tal conjunto seria convexo. Vejamos na Figura 5 exemplos de fecho de um conjunto convexo e um não convexo.

Figura 5 – Exemplos de fecho de conjunto convexo e não convexo.



Fonte: Autoria própria.

Perceba na Figura 5 que o quadrilátero é um polígono convexo. Logo, é igual ao seu fecho convexo. Entretanto, isso não ocorre no segundo polígono, pois observe que o fecho convexo é um pentágono que contém o polígono não convexo.

Além disso, se considerarmos os intervalos $(0,1)$ e $(2,3)$ da reta real, temos que o fecho convexo do conjunto $(0,1) \cup (2,3)$ é dado por $[0,3]$, isto é, um conjunto convexo.

2.2 Conjuntos Afins

Como queremos lidar com Conjuntos Simpléticos, precisaremos introduzir o conceito de hiperplanos, como sendo uma generalização para espaços de dimensões maiores do conceito de plano que temos no \mathbb{R}^3 . Assim, relacionaremos esses hiperplanos com o conceito clássico de espaço vetorial. A grosso modo, o que impede um hiperplano de ser um espaço vetorial é a ausência do elemento neutro da adição, o que possibilita traçar uma relação de paralelismo destes com um (único) espaço vetorial. Não somente os hiperplanos têm essa característica/possibilidade, outros conjuntos também gozam dessas particularidades e o conjunto de todos eles nos levam a definir o conceito de conjuntos afins.

Neste sentido, nesta seção iremos trabalhar com conjuntos afins, relacionando-os com espaços vetoriais, estabelecendo entre eles uma relação de paralelismo.

Definição 5 (Conjunto Afim). *Um subconjunto $M \subset \mathbb{R}^n$ é chamado de conjunto afim se $(1 - \lambda)x + \lambda y \in M$ para todo $x, y \in M$ e $\lambda \in \mathbb{R}$.*

O conjunto vazio e os conjuntos unitários são exemplos de conjuntos afins, os quais podem ser trivialmente demonstrados. Além desses conjuntos, no espaço \mathbb{R}^n podemos verificar que são afins as retas, os planos e os hiperplanos. Veremos agora que todo espaço

vetorial também é um conjunto afim. Posteriormente, iremos avaliar como se relaciona os conjuntos afins com espaços vetoriais.

Teorema 6. *Todo espaço vetorial é um conjunto afim contendo o vetor 0, isto é, o elemento neutro da adição. Ademais, todo subconjunto afim de um espaço vetorial será subespaço se contiver o vetor 0.*

Demonstração. Por definição, todo espaço vetorial contém o vetor 0, no que resta provar que ele é um conjunto afim. De fato, espaços vetoriais são fechados para a adição e para a multiplicação por escalar. Assim sendo, considerando um espaço vetorial M , dados $x, y \in M$, temos que, para quaisquer escalares λ, γ , vale $\lambda x + \gamma y \in M$. Logo, particularmente para $\lambda \in \mathbb{R}^n$ e $\gamma = 1 - \lambda$, temos que $\lambda x + (1 - \lambda)y \in M$, o que prova que M é afim.

Para provarmos a segunda parte, lembremos que todo subespaço vetorial deve conter o vetor nulo e deve ser fechado para soma e multiplicação por escalar.

Seja então M um conjunto afim contendo o vetor 0, temos que para qualquer $x \in M$ e para qualquer escalar λ vale:

$$\lambda x = (1 - \lambda)0 + \lambda x \in M.$$

Portanto, M é fechado para a multiplicação por escalar. Agora, se tomarmos o seguinte ponto de M , temos que

$$\frac{1}{2}x + \left(1 - \frac{1}{2}\right)y \in M \Rightarrow \frac{1}{2}(x + y) \in M,$$

para quaisquer $x, y \in M$. Como demonstramos anteriormente que M é fechado para a multiplicação por escalar, segue que

$$2\frac{1}{2}(x + y) \in M \Rightarrow x + y \in M.$$

Portanto, M é um subespaço de \mathbb{R}^n . □

Observação 7. *Cabe notar aqui que todo espaço vetorial é um conjunto afim, mas nem todo conjunto afim é um espaço vetorial. De fato, falta garantir a propriedade do elemento neutro da soma, ou seja, não é possível concluir que o 0 irá pertencer ao conjunto afim.*

Para o nosso estudo posterior, será muito importante relacionar conjuntos afins com espaços vetoriais, a fim de aproveitar algumas propriedades e resultados deste último. A forma de relacionar espaços afim com espaços vetoriais é através da extensão de paralelismo entre retas e planos para outros espaços afim. Neste sentido, vamos iniciar o nosso trabalho apresentando o conceito de conjunto de transladação.

Definição 8 (Conjunto de Transladação). *Dado $a \in \mathbb{R}^n$, um conjunto de transladação de um conjunto afim $M \subset \mathbb{R}^n$ é dado por*

$$M + a = \{x + a | x \in M\}.$$

Com base nessa definição, temos a seguinte proposição.

Proposição 9. *Um conjunto de transladação de um conjunto afim $M \subset \mathbb{R}^n$ também é um conjunto afim.*

Demonstração. Deja $L = M + a$, com $a \in \mathbb{R}^n$. Se $x, y \in L$ então existem \bar{x} e \bar{y} tais que $x = \bar{x} + a$ e $y = \bar{y} + a$. Dessa maneira, vamos ter para todo $\lambda \in \mathbb{R}$ que:

$$(1 - \lambda)x + \lambda y = (1 - \lambda)(\bar{x} + a) + \lambda(\bar{y} + a) = (1 - \lambda)\bar{x} + \lambda\bar{y} + a \in L,$$

o que prova o que queríamos. □

Definição 10 (Conjunto afim paralelo). *Um conjunto afim M é dito paralelo ao conjunto afim L quando $M = L + a$ para algum a .*

Ou seja, no caso em que M for um conjunto de transladação de L . Sabendo da natureza dos conjuntos, podemos dizer apenas que " M é paralelo a L ". Ademais, isso é uma relação de equivalência entre os subconjuntos afins de \mathbb{R}^n .

Observação 11. *Note que o conceito de paralelismo definido para conjuntos afins é mais restrito do que aquele conhecido na geometria analítica de \mathbb{R}^2 e \mathbb{R}^3 , pois não considera, por exemplo, uma reta paralela a um plano. Por sua vez, uma reta pode ser paralela a uma reta, assim como um plano ser paralelo a um plano.*

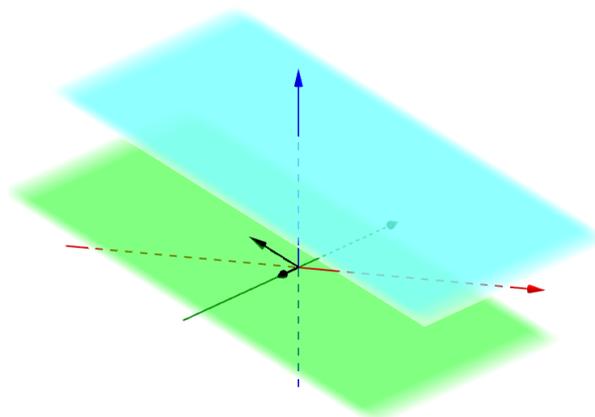
Teorema 12. *Em \mathbb{R}^n , cada conjunto afim não vazio M é paralelo a um único subespaço L . Além disso, L é dado por*

$$L = M - M = \{x - y \mid x, y \in M\}.$$

Demonstração. Inicialmente, vamos mostrar que um conjunto afim M não pode ser paralelo a dois subespaços. Dessa maneira, supomos que M é paralelo aos subespaços L_1 e L_2 . Neste caso, podemos afirmar que $L_2 = L_1 + a$ para algum $a \in \mathbb{R}^n$. Porém, $0 \in L_2$, ou seja, $-a \in L_1$ e também $a \in L_1$, pois L_1 é um subespaço. Sendo assim, $L_2 = L_1 + a \subset L_1$. De maneira análoga, podemos afirmar que $L_1 \subset L_2$, e assim $L_1 = L_2$. Agora, observe que para todo $a \in M$ temos que $L = M - a = M + (-a)$ é uma transladação de M que conterà a origem. Como a transladação é um espaço afim, temos pelo Teorema 6 e da primeira parte demonstrada que L é o subespaço que procuramos. Como a foi escolhido de maneira arbitrária, podemos escrever $L = M - M$. □

Na Figura 6, temos a representação de um conjunto afim que é um plano. Além disso, o subespaço paralelo é um plano contendo a origem.

Figura 6 – Um subespaço paralelo a um conjunto afim.



Fonte: Autoria própria.

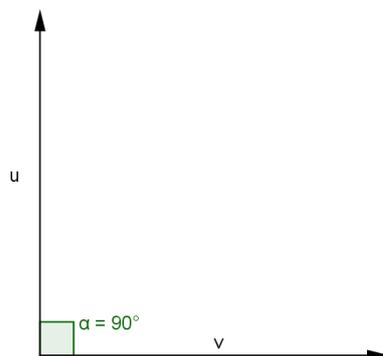
O Teorema 12 nos permitirá utilizar conjuntos afins contendo o 0, um espaço vetorial, ao invés de utilizar conjuntos afins quaisquer. Note que não terá ambiguidade, pois demonstrando que existe um único conjunto afim que é espaço vetorial, relacionado com um conjunto afim que não necessariamente satisfaz as condições de espaço vetorial, é possível estabelecer uma relação de equivalência. Para evitar confusões, a partir de agora, chamaremos os conjuntos afins que são espaços vetoriais de *espaços afins*. Esse teorema permite também que estabeleçamos a definição de dimensão do conjunto afim, como se segue.

Definição 13 (Dimensão de um Conjunto Afim). *A dimensão de um conjunto afim não vazio será dada através da dimensão do subespaço paralelo a ele. Além disso, por convenção, será definido como -1 a dimensão de \emptyset . Os conjuntos afins de dimensão 0, 1 e 2 são chamados de pontos, retas e planos respectivamente e, para um conjunto afim de dimensão $(n-1)$ pertencentes a \mathbb{R}^n , chamaremos de hiperplanos.*

Agora, vamos relembrar conceitos de ortogonalidade oriundos da Álgebra Linear. É importante definirmos tais conceitos, uma vez que nos dará um grande ferramental matemático para definir os hiperplanos de uma maneira mais simples de ser trabalho.

Definição 14 (Vetores Ortogonais). *Dados os vetores $x, y \in \mathbb{R}^n$ dizemos que x é ortogonal a y ($x \perp y$) se $\langle x, y \rangle = 0$.*

Observemos o significado geométrico dessa definição. Por exemplo, sejam dois vetores $u, v \in \mathbb{R}^2$, a ortogonalidade é representada conforme a Figura 7.

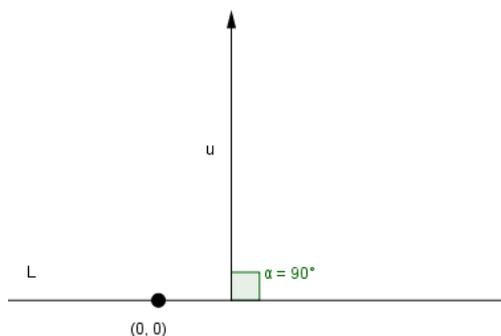
Figura 7 – Vetores ortogonais em \mathbb{R}^2 .

Fonte: Autoria própria.

Vejam agora a definição de ortogonalidade entre um vetor e um subespaço.

Definição 15 (Vetor ortogonal a um conjunto). *Seja um vetor $x \in \mathbb{R}^n$ e um subespaço $L \subset \mathbb{R}^n$, dizemos que x é ortogonal a L se $x \perp y$ para todo $y \in L$.*

Um exemplo simples de se analisar é se considerarmos o subespaço L sendo uma reta e um vetor u tal que qualquer vetor $x \in L$ é ortogonal a u . Na Figura 8 a seguir, podemos observar a representação geométrica dessa definição.

Figura 8 – Vetor ortogonal a um subespaço no \mathbb{R}^2 .

Fonte: Autoria própria.

Também é muito utilizado um vetor ortogonal a um plano no \mathbb{R}^3 , tal vetor é denominado de vetor normal.

Proposição 16. *Todo subespaço de dimensão $(n - 1)$ é um conjunto da forma $\{x | x \perp b\}$, com $b \in \mathbb{R}^n$ e $x \in L$, com L sendo um conjunto afim. Além disso, as translações desse conjunto geram um hiperplano.*

Demonstração. Seja $M \subset \mathbb{R}^n$ um subespaço de dimensão $n - 1$. Então, temos que existem $n - 1$ vetores linearmente independentes de \mathbb{R}^n que geram M . Isto é

$$M = [m_1, m_2, \dots, m_{n-1}],$$

com $m_1, m_2, \dots, m_{n-1} \in \mathbb{R}^n$. Como a dimensão de M é $n - 1$, existe $b \in \mathbb{R}^n$ tal que b é ortogonal a todos os vetores da base de M . Ou seja, $\langle m_i, b \rangle = 0$ para $i = 1, 2, \dots, n - 1$. Além disso, todo vetor $m \in M$ é ortogonal a b pois, se $m = m_1\alpha_1 + \dots + m_{n-1}\alpha_{n-1}$ então segue que

$$\begin{aligned} \langle m, b \rangle &= \langle m_1\alpha_1 + \dots + m_{n-1}\alpha_{n-1}, b \rangle \\ &= \langle m_1\alpha_1, b \rangle + \dots + \langle m_{n-1}\alpha_{n-1}, b \rangle \\ &= \alpha_1\langle m_1, b \rangle + \dots + \alpha_{n-1}\langle m_{n-1}, b \rangle \\ &= 0. \end{aligned}$$

Logo, $M = \{x \in \mathbb{R}^n \mid \langle x, b \rangle = 0\}$, pois se tivermos um y ortogonal a M e também a b . Como $M \cup [b] = \mathbb{R}^n$, temos que $y \notin \mathbb{R}^n$. Além disso, se transladarmos o conjunto M por a , então

$$\{x \mid x \perp b\} + a = \{x + a \mid \langle x, b \rangle = 0\} = \{y \mid \langle y - a, b \rangle = 0\} = \{y \mid \langle y, b \rangle = \beta\},$$

com $y = x + a$ e $\beta = \langle a, b \rangle$. Portanto, os hiperplanos são translações do conjunto M . \square

A Proposição 16 acima permite estabelecer uma nova forma de definir hiperplanos. De fato, dado o vetor de variáveis $x = (x_1, x_2, \dots, x_n)$, um vetor conhecido $a = (a_1, a_2, \dots, a_n)$ e $\beta \in \mathbb{R}$, um hiperplano pode ser escrito como o conjunto $H = \{a_1x_1 + a_2x_2 + \dots + a_nx_n = \beta\}$ ou ainda $H = \{x, a \in \mathbb{R}^n \text{ e } \langle x, a \rangle = \beta\}$.

Teorema 17. *Dados $\beta \in \mathbb{R}$, e um vetor diferente de zero $b \in \mathbb{R}^n$. Então*

$$H = \{x \mid \langle x, b \rangle = \beta\}$$

é um hiperplano em \mathbb{R}^n . Além disso, todo hiperplano pode ser definido dessa forma, com b e β únicos a menos de um múltiplo diferente de zero.

Demonstração. Sabemos que um hiperplano pode ser representado por uma equação da forma $\alpha_1x_1 + \dots + \alpha_{n-1}x_{n-1} = \beta$. Logo

$$b_1x_1 + \dots + b_{n-1}x_{n-1} = \beta \Leftrightarrow (x_1, \dots, x_{n-1}) \cdot (b_1, \dots, b_{n-1}) = \beta \Leftrightarrow \langle x, b \rangle = \beta,$$

para $x = (x_1, \dots, x_{n-1})$ e $b = (b_1, \dots, b_{n-1})$. Logo, o hiperplano pode ser escrito como

$$H = \{x \mid \langle x, b \rangle = \beta\}.$$

Falta mostrar que b e β são únicos a menos de um escalar diferente de zero. Suponha que

$$H = \{x \mid \langle x, b \rangle = \beta\} = \{x \mid \langle x, a \rangle = \alpha\},$$

isto indica que H é o conjunto solução de dois sistemas lineares equivalentes, cada qual com uma única equação. A teoria de sistemas lineares garante que podemos chegar de um sistema a outro através de operações elementares. No entanto, devido ao fato de termos somente uma única equação, a única possibilidade aqui é multiplicar a linha por um escalar não nulo. Dessa forma, provamos o que foi pedido, uma vez que se multiplicarmos ambos os lados da equação $\langle x, b \rangle = \beta$ pelo escalar $\gamma \neq 0$ teremos a equação $\langle x, a \rangle = \alpha$, de onde podemos concluir que a é múltiplo de b e na mesma proporção que α é múltiplo de β . \square

Vamos agora estabelecer uma relação de equivalência entre os conjuntos afins e os conjuntos soluções de sistemas lineares.

Teorema 18. *Dado um vetor $b \in \mathbb{R}^m$ e uma matriz $B \in \mathbb{R}^{m \times n}$. Então o conjunto*

$$M = \{x \in \mathbb{R}^n | Bx = b\}$$

é um conjunto afim em \mathbb{R}^n . Ademais, todo conjunto afim pode ser representado dessa forma.

Demonstração. Sejam $x, y \in M$ e $z = (1 - \lambda)x - \lambda y$, então temos que

$$\begin{aligned} Bz &= B((1 - \lambda)x - \lambda y) \\ &= B(1 - \lambda)x - B\lambda y \\ &= (1 - \lambda)Bx - \lambda By, \text{ como } x \text{ e } y \text{ são soluções do sistema} \\ &= (1 - \lambda)b + \lambda b \\ &= b - \lambda b + \lambda b \\ &= b \Leftrightarrow (1 - \lambda)b + \lambda b \\ &= b. \end{aligned}$$

Logo, $Bz \in M$. Então M é um espaço afim.

Agora precisamos mostrar que todo conjunto afim pode ser representado da forma $\{x \in \mathbb{R}^n | Bx = b\}$. Para isso, tomaremos um conjunto afim arbitrário M , e L um subespaço paralelo a M . Sejam b_1, b_2, \dots, b_m base do complemento ortogonal¹ de L , dado por L^\perp , então

$$\begin{aligned} L &= (L^\perp)^\perp \\ &= \{x | x \perp b_1, x \perp b_2, \dots, x \perp b_m\} \\ &= \{x | \langle b_i, x \rangle = 0, \quad i = 1, 2, \dots, m\} \\ &\Leftrightarrow \{b_1 x^T + b_2 x^T + \dots + b_m x^T = 0\} \\ &= \{Bx = 0\}, \end{aligned}$$

¹O complemento ortogonal de um subespaço L de \mathbb{R}^n é dado por todos os vetores de \mathbb{R}^n que são ortogonais a todos os vetores de L . Tal conjunto é descrito por L^\perp .

com $B \in \mathbb{R}^{m \times n}$. Como M é paralelo a L então existe $a \in \mathbb{R}^n$ tal que

$$M = L + a = \{x + a | Bx = 0\} = \{x | B(x - a) = 0\} = \{x | Bx = Ba\} = \{x | Bx = b\},$$

com $Ba = b$. Portanto, temos que L é o único subespaço paralelo ao conjunto afim M . Com isso, M pode ser representado por um sistema de equação, ou seja,

$$M = \{x | Bx = b\}.$$

□

Observe que podemos representar os conjuntos \emptyset e \mathbb{R}^n dessa maneira. Para o conjunto \emptyset , tomamos B como sendo uma matriz com todas as entradas iguais a zero, e b como sendo um vetor diferente de zero. Para o conjunto \mathbb{R}^n , tomaremos B como sendo a matriz com todas as entradas zero, e b como sendo o vetor nulo.

Se considerarmos b_i a i -ésima linha de B , e β_i como sendo a i -ésima entrada de b temos que

$$\begin{aligned} M &= \{x | \langle x, b_1 \rangle = \beta_1, \langle x, b_2 \rangle = \beta_2, \dots, \langle x, b_m \rangle = \beta_m\} \\ &= \{x | \langle x, b_i \rangle = \beta_i, \quad i = 1, 2, \dots, m\}. \end{aligned}$$

Portanto, se chamarmos $H_i = \{x | \langle x, b_i \rangle = \beta_i\}$, temos que H_i é um hiperplano para $i = 1, 2, \dots, m$. Assim,

$$M = \{x | \langle x, b_i \rangle = \beta_i, \quad i = 1, 2, \dots, m\} = \bigcap_{i=1}^m H_i.$$

Ou seja, M pode ser descrito como uma intersecção finita de hiperplanos.

Corolário 19. *Todo subconjunto afim de \mathbb{R}^n é uma intersecção de uma quantidade finita de hiperplanos.*

Demonstração. Do Teorema 18, temos que todo conjunto afim pode ser representado como uma equação matricial. Portanto, se M é um conjunto afim então $M = \{x | Bx = \beta\}$. Isto é,

$$Bx = \beta \Leftrightarrow \begin{bmatrix} b_{11} & \dots & b_{1n-1} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn-1} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \Leftrightarrow \begin{cases} b_{11}x_1 + \dots + b_{1n-1}x_{n-1} = \beta_1 \\ \vdots \\ b_{m1}x_1 + \dots + b_{mn-1}x_{n-1} = \beta_m \end{cases}$$

Logo,

$$M = \{x = (x_1, \dots, x_{n-1}) | \langle b'_1, x \rangle = \beta_1 \text{ e } \dots \text{ e } \langle b'_m, x \rangle = \beta_m\},$$

sendo b'_1, b'_2, \dots, b'_m as linhas de B . Sabemos que $\langle b_i, x \rangle = \beta_i$ são hiperplanos de \mathbb{R}^n . Além disso, sabemos que a intersecção de conjuntos afins ainda é um conjunto afim. Portanto, M é formado por intersecções de hiperplanos. □

Embora não seja complicado mostrar que a intersecção de conjuntos afins ainda é afim, este resultado é muito importante para a discussão como um todo. Para isso, iremos demonstrar a proposição a seguir.

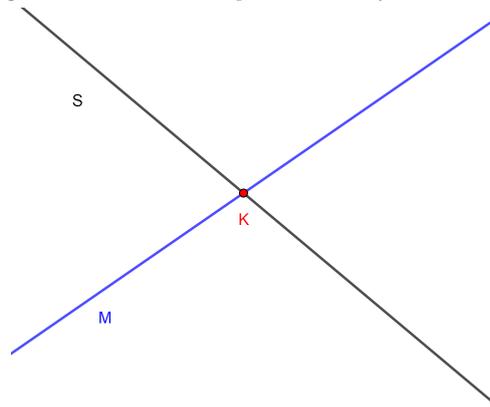
Proposição 20. *A intersecção arbitrária de conjuntos afins também é um conjunto afim.*

Demonstração. Sabemos que se $x, y \in \bigcap_{i=1}^{\infty} A_i$, com A_i um conjunto afim para todo $i \in \mathbb{N}$, então teremos que $x, y \in A_i$ para todo i . Como A_i é afim, para todo $\lambda \in \mathbb{R}$, então segue que

$$\lambda x + y(1 - \lambda) \in A_i.$$

Como isso ocorre para todo A_i e para todo λ , então $x, y \in \bigcap_{i=1}^{\infty} A_i$ é um conjunto afim. \square

Figura 9 – Intersecção de conjuntos afins.



Fonte: Autoria própria.

Se $S \subset \mathbb{R}^n$, então há um menor conjunto afim² que contém S . Tal conjunto é o fecho afim que iremos definir a seguir, pode-se provar ainda que o fecho afim é dado pela intersecção de todos os conjuntos afins que contenha S .

Definição 21 (Fecho Afim). *O fecho afim de um conjunto $S \subset \mathbb{R}^n$ é o conjunto*

$$aff(S) = \{x \in \mathbb{R}^n \mid x = \lambda_1 x_1 + \dots + \lambda_n x_n, x_i \in S, i = 1, \dots, n, \sum_{i=1}^n \lambda_i = 1\}$$

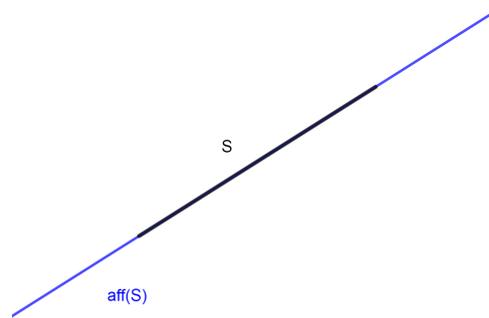
Observação 22. *A diferença básica entre o fecho afim e o fecho convexo é que na composição da combinação $\lambda_1 x_1 + \dots + \lambda_n x_n$ o fecho afim aceita valores negativos e o fecho convexo não aceita, assim sendo todo elemento do fecho convexo é elemento do fecho afim (o inverso não é verdade) ou seja, o fecho convexo de S está contido no fecho afim de S .*

²Neste caso, um conjunto A ser o menor conjunto afim que contenha S significa que, se existir outro conjunto afim M , tal que $S \subset M$ então $A \subset M$.

Notemos que existem quatro conjuntos afins possíveis em \mathbb{R}^2 , sendo eles: o conjunto vazio, os conjuntos formados por pontos, os conjuntos formados por retas e o próprio \mathbb{R}^2 . Podemos também analisar esse fato geometricamente. Se tomarmos S e M como conjuntos afins, a intersecção dos dois será o ponto K , que também é um conjunto afim. Vejamos a Figura 9.

Neste caso, também podemos representar geometricamente. Se considerarmos S como sendo um segmento de reta no plano, teremos então que $\text{aff}(S)$ será dado por uma reta, conforme a Figura 10 abaixo.

Figura 10 – Fecho afim de um conjunto.



Fonte: Autoria própria.

Definição 23 (Conjunto independentemente afim). *O conjunto $\{b_0, b_1, \dots, b_m\}$ é chamado de conjunto independentemente afim se $\text{aff}(b_0, b_1, \dots, b_m)$ tem dimensão m .*

Para determinar a dimensão de $\text{aff}(b_0, b_1, \dots, b_m)$, podemos utilizar a estratégia de determinar a dimensão do espaço afim paralelo $L = \text{aff}(0, b_1 - b_0, \dots, b_m - b_0)$, o qual é um espaço vetorial. Neste sentido, para que o conjunto $\{b_0, b_1, \dots, b_m\}$ seja independentemente afim, podemos simplesmente averiguar se $\{b_1 - b_0, \dots, b_m - b_0\}$ é linearmente independente. O Teorema 24 que se segue prova isso.

Teorema 24. *Os pontos b_0, b_1, \dots, b_m são independentemente afim se, e somente se, os vetores $b_1 - b_0, \dots, b_m - b_0$ são linearmente independentes.*

Demonstração. Iremos supor primeiramente que os pontos são independentemente afim. Seja $L + b_0 = \text{aff}(b_0, b_1, \dots, b_m)$. Sabemos que $L + b_0$ tem dimensão m . Além disso, sabemos também que L é o único espaço vetorial paralelo a $L + b_0$. Logo, L tem dimensão m . Também temos que $L = [b_1 - b_0, \dots, b_m - b_0]$. Portanto, os vetores $b_1 - b_0, \dots, b_m - b_0$ são linearmente independentes.

Para a recíproca, temos que os vetores $b_1 - b_0, \dots, b_m - b_0$ são linearmente independentes. Logo, o conjunto afim paralelo a esse espaço vetorial tem dimensão m . Dessa forma, o conjunto afim $L + b_0$ tem dimensão m . Portanto, b_0, b_1, \dots, b_m são independentemente afim. \square

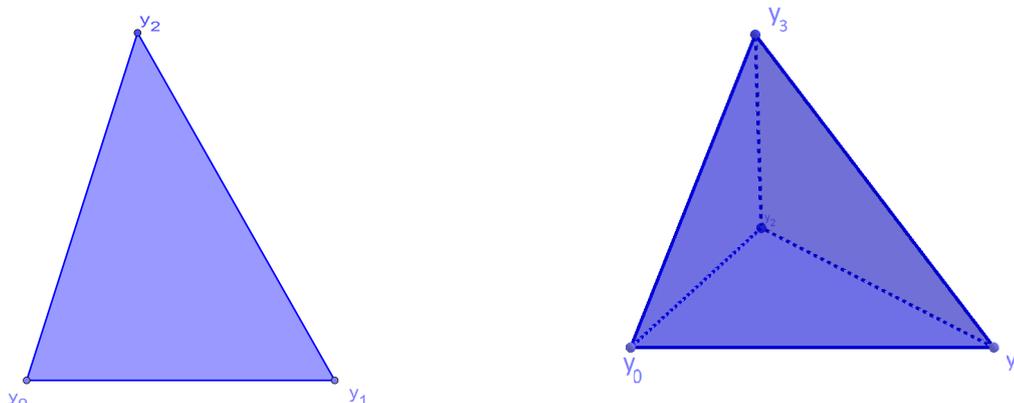
Finalmente, definiremos o conceito simplético. Para isso, iremos utilizar o resultado da construção de conjunto afim e fecho de conjunto que foram desenvolvidas anteriormente. Nessa parte do trabalho, o enfoque será dado conforme o livro *Introduction to derivative-free optimization* (CONN; SCHEINBERG; VICENTE, 2009).

2.3 Simplex

Definição 25 (Simplex). *Dado um conjunto de pontos $Y = \{y_0, \dots, y_m\}$ independentemente afim, o fecho convexo desse conjunto será chamado de simplex de dimensão m . Além disso, os pontos y_0, \dots, y_m são chamados de vértices do simplex.*

Vejamos um exemplo de simplex.

Figura 11 – Exemplo de simplex em \mathbb{R}^2 e \mathbb{R}^3 .



Fonte: Autoria própria.

Como pudemos observar, temos que o simplex de dimensão 2 é um triângulo e o simplex de dimensão 3 é um tetraedro. Notemos ainda que os pontos são simplex de dimensão 0 e os segmentos de reta são simplex de dimensão 1.

Por fim, temos a seguinte observação.

Observação 26. *Para a definição de Simplex, nós tivemos que utilizar os dois conceitos estudados anteriormente: o conceito de conjunto convexo, o qual foi introduzido para definir o conjunto em si (dado pelo fecho convexo); e a noção de conjunto afim, a qual foi utilizada para definir a dimensão do conjunto (dada pela dimensão do fecho afim). Esses dois conceitos muitas vezes causa confusão, uma vez que a diferença entre fecho convexo e fecho afim é que, no caso do fecho convexo, os escalares da combinação linear intrínseca $\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_m x_m$ tem valores de λ_i sempre positivos, o que não é necessário para determinar os elementos do fecho afim.*

Diante do exposto neste capítulo, temos embasamento necessário para iniciar os estudos dos Métodos Simpléticos de Otimização. Nesse trabalho, iremos explorar apenas dois métodos (sendo um método para problemas lineares e outro para problemas não lineares). Porém, existem uma grande quantidade de métodos iterativos de otimização baseados nos conjuntos simpléticos, além de um estudo aprofundado na topologia simplética com diferentes aplicações.

3 PROBLEMAS DE PROGRAMAÇÃO LINEAR

Neste capítulo, estudaremos os Problemas de Programação Linear. Além disso, desenvolveremos o Método Simplex, o qual é um dos métodos mais utilizados para resolver esse tipo de problema. A maior parte do que foi apresentado aqui segue do livro intitulado *Linear and Nonlinear Programming* (LUENBERGER; YE, 2008).

3.1 O Problema de Programação Linear

De forma geral, resolver um Problema de Programação Linear significa encontrar os valores x_1, \dots, x_n tais que satisfazem as seguintes condições:

$$\begin{aligned}
 &\text{minimizar} && c_1x_1 + c_2x_2 + \dots + c_nx_n \\
 &\text{sujeito à} && a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\
 &&& a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\
 &&& \vdots \\
 &&& a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \\
 &&& x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0.
 \end{aligned} \tag{2}$$

Para $j = 1, \dots, n$ e $i = 1, \dots, m$, os valores b_i, c_j, a_{ij} são valores reais fixados. A função a ser minimizada $c(x_1, \dots, x_n) = c_1x_1 + c_2x_2 + \dots + c_nx_n$, em consonância com o que escrevemos na introdução, é chamada de função objetivo e as equações

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$$

, para $i = 1, \dots, m$, são chamadas restrições de igualdade.

Utilizando a notação matricial, podemos reescrever um PPL de uma maneira mais compacta, como se segue

$$\begin{aligned}
 &\text{minimizar} && c^T x \\
 &\text{sujeito a} && Ax = b \\
 &&& \text{e } x \geq 0
 \end{aligned} \tag{3}$$

com $x, c \in \mathbb{R}^n, b \in \mathbb{R}^m, A \in \mathbb{R}^{m \times n}$. Notemos que $x \geq 0$ indica que nenhuma das entradas de x tem valor negativo. Isto é, $x_i \geq 0$, para $i = 1, \dots, n$. Neste caso, chamaremos a matriz A de matriz de restrições.

Embora hajam diferentes maneiras de apresentar um Problema de Programação Linear, utilizando certas ferramentas, sempre conseguimos reduzir ao caso geral do PPL (3). A seguir, iremos apresentar tais ferramentas.

Caso 1 - Variáveis de folga: um Problema de Programação Linear pode ser modelado (ou escrito) de maneira que apresente desigualdades, tendo a seguinte versão:

$$\begin{aligned} \text{minimizar} \quad & c_1x_1 + c_2x_2 + \cdots + c_nx_m \\ \text{sujeito à} \quad & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1 \\ & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leq b_2 \\ & \vdots \\ & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq b_m \\ & x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0 \end{aligned}$$

Criando as variáveis de folga y_1, \dots, y_n , exigindo que elas sejam todas positivas e subtraindo cada variável em uma das inequações, iremos desfazer as desigualdades e criar igualdades nas equações. Voltaremos, assim, ao caso geral (2) inicialmente proposto. Obteremos, portanto

$$\begin{aligned} \text{minimizar} \quad & c_1x_1 + c_2x_2 + \cdots + c_nx_m \\ \text{sujeito à} \quad & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + y_1 = b_1 \\ & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + y_2 = b_2 \\ & \vdots \\ & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + y_m = b_m \\ \text{com} \quad & x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0 \\ \text{e} \quad & y_1 \geq 0, y_2 \geq 0, \dots, y_m \geq 0 \end{aligned}$$

Ao adicionarmos as variáveis y_i , transformamos as inequações em equações, para assim recairmos no primeiro caso apresentado. Temos, dessa forma, um conjunto de restrições com $n + m$ variáveis reais, sendo elas $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m$. As restrições de igualdade agora determinam um sistema linear com $n + m$ variáveis. Logo, podemos escrever o problema como $B \in \mathbb{R}^{(n+m) \times m}$ no qual $B = [A I]$, sendo $A \in \mathbb{R}^{m \times n}$ a matriz das restrições e $I \in \mathbb{R}^{m \times m}$ a matriz identidade.

Caso 2 - Variáveis de excesso: outro caso possível é dado como se segue

$$\begin{aligned} \text{minimizar} \quad & c_1x_1 + c_2x_2 + \cdots + c_nx_m \\ \text{sujeito à} \quad & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \geq b_1 \\ & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \geq b_2 \\ & \vdots \\ & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \geq b_m \\ & x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0 \end{aligned}$$

Logo, precisamos corrigir a desigualdade. Para isso, precisamos subtrair variáveis y_1, y_2, \dots, y_m e transformar a desigualdade em uma igualdade, da seguinte maneira

$$\begin{array}{ll}
 \text{minimizar} & c_1x_1 + c_2x_2 + \dots + c_nx_n \\
 \text{sujeito à} & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n - y_1 = b_1 \\
 & a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n - y_2 = b_2 \\
 & \vdots \\
 & a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n - y_m = b_m \\
 \text{com} & x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0 \\
 \text{e} & y_1 \geq 0, y_2 \geq 0, \dots, y_m \geq 0
 \end{array}$$

Portanto, adquirimos igualdades nos lugares das desigualdades subtraindo as chamadas variáveis de excesso.

Caso 3 - Variáveis livres: se a condição $x_i \geq 0$ não for necessária para algum i , então podemos reescrever a variável x_i como a combinação de duas variáveis positivas dada como se segue

$$x_i = u_i - v_i \text{ com } u_i \geq 0, v_i \geq 0.$$

Notemos que, neste caso, estamos aumentando a dimensão do nosso sistema. Supondo, por exemplo, que somente uma variável tenha esta particularidade, vamos transformar em um sistema com $n + 1$ variáveis, sendo elas $x_1, x_2, \dots, u_i, v_i, \dots, x_n$, e agora, todas as variáveis devem ser positivas.

Outra maneira de lidar com essa particularidade, é escolher uma das equações do sistema tais que o coeficiente relacionado a x_i seja diferente de zero, para expressar x_i em relação as outras variáveis, isto é,

$$x_i = \frac{1}{a_{li}}(b_l - a_{l1}x_1 + a_{l2}x_2 + \dots + a_{l(i-1)}x_{i-1} + a_{l(i+1)}x_{i+1} + \dots + a_{ln}).$$

Com isso, podemos substituir x_i nas outras equações obtendo um problema com $n - 1$ variáveis e $m - 1$ equações, pois a equação utilizada para relacionar x_i às outras variáveis, irá ser desconsiderada do problema já que a mesma será nula.

Caso 4 - Maximizar uma função objetivo: como estudaremos o Método Simplex para a resolução de PPL's, teremos um problema quando buscarmos maximizar a função objetivo, uma vez que o Método Simplex apenas minimiza a função objetivo. Esse problema é facilmente resolvido pois teremos que vale

$$\begin{array}{ll}
 \text{maximizar} & c^T x & \text{minimizar} & -c^T x \\
 \text{sujeito a} & Ax = b & \Leftrightarrow & \text{sujeito a} & Ax = b \\
 \text{e} & x \geq 0 & & \text{e} & x \geq 0
 \end{array}$$

Ou seja, podemos utilizar o Método Simplex para maximizar a função f , modificando o problema em minimizar a função $-f$. Essa estratégia é muito utilizada em problemas de otimização em geral.

3.2 Soluções Básicas

O resultado mais importante da Teoria de Programação Linear é o Teorema Fundamental da Programação Linear. Veremos no próximo capítulo, que o mesmo motiva e valida o Método Simplex que veremos mais adiante. O Teorema Fundamental da Programação Linear se baseia na separação das variáveis do Problema (2) em básicas e não-básicas. É disso que trataremos nesta seção.

Definição 27. *Considerando o Problema de Programação Linear em sua forma matricial (3), se o vetor $x \geq 0$ e é solução para o sistema $Ax = b$ então x é chamado de **solução factível** para o Problema de Programação Linear.*

Perceba que, se x é solução factível, não significa que x irá resolver o problema de minimização, ou seja, que irá minimizar a função objetivo. Uma solução factível que minimiza o Problema de Programação Linear 2 é chamada de **solução ótima factível**, ou simplesmente **solução ótima**. Adiante, veremos como obter uma solução ótima através do Método Simplex.

Definição 28. *Considere o sistema $Ax = b$, onde A é uma matriz $m \times n$ ($m \leq n$) e B uma submatriz invertível de A com m colunas e m linhas. Considere agora que o conjunto $\mathcal{I} = \{\sigma(1), \dots, \sigma(n)\}$ seja o conjunto de índices das colunas de A que compõem B . Se x é uma solução factível para esse sistema linear e toda variável x_i tal que $i \notin \mathcal{I}$ (componentes que não estão relacionados a B) é nula, então a solução para essa equação é chamada de **solução básica**. Além disso, as variáveis de x_i tais que $i \in \mathcal{I}$ (que estão relacionados com às colunas de B) são chamadas de **variáveis básicas**.*

Consideraremos um sistema linear $Ax = b$, onde A é uma matriz $m \times n$, com $m \leq n$, admitindo m colunas linearmente independentes, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ e B é uma submatriz invertível de A , tal que $B \in \mathbb{R}^{m \times m}$. Além disso, as colunas de B são linearmente independentes. Assim, para compreender melhor a definição anterior, vamos supor, sem perda de generalidade, que a matriz B corresponde às m primeiras colunas de A . Assim sendo, teremos que vale

$$Ax = b \Leftrightarrow \begin{bmatrix} B & C \end{bmatrix} \begin{bmatrix} x_B \\ 0 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Notemos que analisando o problema anterior, a matriz C não tem qualquer impacto na determinação da solução básica. Por fim, antes de finalizar esta seção, vamos apresentar um último conceito que será de extrema utilidade para a compreensão dos tópicos que serão abordados posteriormente.

Definição 29. *Se uma ou mais variáveis básicas em uma solução básica são zero, então essa solução é chamada de **solução básica degenerada**.*

3.3 O Teorema Fundamental da Programação Linear

Abordaremos, o Teorema Fundamental da Programação Linear. Este resultado tem uma grande importância, pois é ele que fornece embasamento para garantir o funcionamento do Método Simplex. Antes de enunciá-lo, apresentaremos um lema necessário para a sua demonstração.

Lema 30. *Se A é uma matriz $m \times n$ de posto m , tal que suas p primeiras colunas são linearmente independentes. Então podemos encontrar $m - p$ colunas entre as $n - p$ colunas restantes, de maneira que formamos m colunas linearmente independentes.*

Demonstração. Seja $A = [A_1 \ A_2 \ \dots \ A_n]$ sendo A_1, A_2, \dots, A_n as colunas de A , se ω for o espaço coluna de A , temos que $\dim(\omega) = m$, além disso, podemos escrever $\omega = \omega_1 \oplus \omega_2$, com ω_1 sendo o espaço gerado por A_1, A_2, \dots, A_p (base de ω , pois é linearmente independente).

Diante do fato de que

$$\begin{aligned} \dim(\omega) &= \dim(\omega_1) + \dim(\omega_2) - \dim(\omega_1 \cap \omega_2) \\ &\Leftrightarrow \dim(\omega_2) = m - p \end{aligned}$$

pois a soma direta nos dá que $\omega_1 \cap \omega_2 = 0$.

A base de ω pode ser extraída do conjunto formado pelas colunas de A , devido à soma direta e à forma como foi construído ω_1 , temos que é possível extrair de $\{A_{p+1}, A_{p+2}, \dots, A_n\}$, o conjunto $X = \{A_{\sigma(1)}, A_{\sigma(2)}, \dots, A_{\sigma(m-p)}\}$ LI e, portanto, uma base de ω_2 e, por fim, novamente devido à soma direta, temos que

$$\{A_1, A_2, \dots, A_p, A_{\sigma(1)}, A_{\sigma(2)}, \dots, A_{\sigma(m-p)}\}$$

é LI, o que completa a demonstração. □

Com esse resultado, podemos demonstrar o principal teorema da Programação Linear, o que irá ser indispensável para encontrar a melhor solução. Nesse trabalho, demonstraremos o teorema em duas partes. Sendo a primeira parte em que demonstramos a existência de uma solução básica factível. Posteriormente, na segunda parte, mostraremos a existência de uma solução ótima básica factível.

Teorema 31 (Teorema Fundamental da Programação Linear - PARTE I). *Seja uma matriz $A \in \mathbb{R}^{m \times n}$,*

$$\begin{aligned} & \text{Maximizar} && c^T x \\ & \text{sujeito à} && Ax = b \\ & && x \geq 0. \end{aligned}$$

Temos que, se há uma solução factível, então há uma solução básica factível.

Demonstração.

Se supnhamos que x seja uma solução factível, efetuando a multiplicação das matrizes, temos

$$\begin{aligned} Ax = b & \Leftrightarrow \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \Leftrightarrow \\ & \Leftrightarrow \begin{bmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \begin{bmatrix} a_{12}x_2 \\ a_{22}x_2 \\ \vdots \\ a_{m2}x_2 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}. \end{aligned}$$

Portanto se chamarmos a i -ésima coluna de A como A_i podemos reescrever o problema de outra forma

$$x_1A_1 + x_2A_2 + \cdots + x_nA_n = b.$$

Supnhamos que a equação tem p variáveis x_i maiores que zero, para assim estudarmos o valor de p , vamos supor, sem perda de generalidade que sejam as p primeiras variáveis, então

$$x_1A_1 + x_2A_2 + \cdots + x_pA_p = b.$$

Agora temos que as colunas de A que aparecem na equação acima podem ser linearmente independente ou linearmente dependente. Por isso, separaremos em dois casos.

Caso 1: As colunas A_1, A_2, \dots, A_p formam um conjunto linearmente independente. Por hipótese. Evidentemente $p \leq m$. Se $p = m$ a solução é básica e portanto está provado. Se $p < m$ temos que, pelo Lema 30 podemos obter $m - p$ colunas entre as $n - p$ colunas restantes de A para formar uma base com m colunas linearmente independente. Podemos atribuir valor 0 para as $m - p$ variáveis construídas, obtendo uma solução degenerada.

Caso 2: Supnhamos que as p primeiras colunas de A são linearmente dependentes, então podemos fazer

$$y_1A_1 + y_2A_2 + \cdots + y_pA_p = 0,$$

uma solução não trivial para tal combinação linear onde pelo menos um y_i é estritamente positivo. Então, para todo ε temos que

$$\varepsilon y_1 A_1 + \varepsilon y_2 A_2 + \cdots + \varepsilon y_p A_p = 0.$$

Podemos subtrair esse valor da solução factível, então temos que

$$(x_1 - \varepsilon y_1)A_1 + (x_2 - \varepsilon y_2)A_2 + \cdots + (x_p - \varepsilon y_p)A_p = b.$$

Dessa maneira, se $y = (y_1, \dots, y_p, 0 \dots 0)$, $x - \varepsilon y$ é uma solução para o sistema linear $Ax = b$, mas não sabemos se para todo $i = 1, \dots, p$ vamos ter $x_i - \varepsilon y_i \geq 0$ e, assim sendo, não podemos afirmar que é uma solução factível para o Problema de Programação Linear 2.

Agora, como pelo menos um y_i é estritamente positivo, podemos definir

$$\varepsilon = \min \left\{ \frac{x_i}{y_i} \mid y_i > 0 \right\}.$$

E, fazendo isso, temos que a equação anterior será reduzida para, no máximo, $p - 1$ variáveis e daí, temos uma nova solução básica factível, com uma variável básica a menos.

O novo conjunto de colunas de A , pode ou não pode ser um conjunto de colunas linearmente independente. Caso o conjunto continue sendo linearmente dependente, podemos repetir o processo até conseguir um conjunto linearmente independente, e assim, podemos aplicar o caso 1, o que prova o Teorema.

□

Teorema 32 (Teorema Fundamental da Programação Linear - PARTE II). *Seja uma matriz $A \in \mathbb{R}^{m \times n}$, dado o PPL,*

$$\begin{array}{ll} \text{Maximizar} & c^T x \\ \text{sujeito à} & Ax = b \\ & x \geq 0. \end{array}$$

Se há uma solução ótima factível, então existe uma solução ótima básica factível.

Demonstração. Tomando $x = (x_1, x_2, \dots, x_n)$ uma solução ótima factível, igualmente ao que foi feito na demonstração anterior, vamos supor que há p variáveis positivas x_1, x_2, \dots, x_p , escrever a seguinte combinação linear

$$A_1 x_1 + A_2 x_2 + \cdots + A_p x_p = b$$

e dividir nos mesmos dois casos: um em que com as p primeiras colunas de A formam um conjunto linearmente independente e outro em que formam um conjunto linearmente dependente.

Caso 1: $\{A_1, A_2, \dots, A_p\}$ é LI. A demonstração aqui segue pelo mesmo caminho que foi demonstrado o Caso 1 do Teorema 31. Isto é, apoiados no Lema 30 podemos obter $m - p$ colunas entre as $n - p$ colunas restantes de A para formar uma base com m colunas linearmente independente e, atribuindo valor 0 para as $m - p$ variáveis restantes, obtemos uma solução ótima degenerada.

Caso 2: $\{A_1, A_2, \dots, A_p\}$ é LD. Aqui também, vamos seguir o que foi demonstrado no Caso 2 do Teorema 31. No entanto, ainda precisamos mostrar que a solução obtida no fim do processo é ótima. Seguindo o procedimento ali descrito, temos sucessivas adições do tipo $-\varepsilon y$ à solução ótima inicial, neste sentido, devemos mostrar que essas adições não irão modificar o valor da função objetivo, isto é, devemos mostrar que $c^T x - \varepsilon c^T y = c^T x$, ou seja, temos de averiguar que $c^T y = 0$. Suponha, por absurdo, que $c^T y \neq 0$ para um determinado valor de ε suficientemente pequeno em módulo, $x - \varepsilon y$ é uma solução factível para valores positivos e negativos. Utilizando o sinal adequado, temos que o valor de função $c^T x - \varepsilon c^T y$ é menor do que o valor $c^T x$ para um ponto factível, o que contradiz a hipótese de que x é uma solução ótima. \square

Para entender o comportamento do Método Simplex para resolução do PPL (2), é de grande utilidade conhecer também as propriedades geométricas das soluções básicas. Neste sentido, vamos generalizar o conceito de vértice como o ponto de um conjunto convexo, em que não é possível estar contido em um segmento determinado por outros dois pontos.

Definição 33. Chamaremos o ponto x de vértice de um conjunto convexo C quando, para todo $x_1, x_2 \in C$ e $x_1 \neq x_2$, não existe $\alpha \in (0,1)$, tal que $x = x_1 \alpha + (1 - \alpha)x_2$.

Em um triângulo, os vértices são os pontos extremos formados pela intersecção das arestas. Já em um círculo, todos os pontos de sua circunferência são vértices. Baseado nessa definição, poderemos demonstrar um teorema que relaciona as soluções básicas aos vértices do conjunto factível.

Teorema 34. Seja A uma matriz $m \times n$ de posto m e $b \in \mathbb{R}^m$ e seja K o conjunto convexo tendo todos os vetores $x \in \mathbb{R}^n$ que satisfazem

$$Ax = b, \quad x \geq 0$$

x é um vértice de K se, e somente se, x for uma solução básica factível.

Demonstração. Suponhamos que $x = (x_1, \dots, x_m, 0, \dots, 0)$ é uma solução básica factível, então

$$x_1 A_1 + x_2 A_2 + \dots + x_m A_m = b$$

sendo que A_1, A_2, \dots, A_m é L.I. Suponhamos agora que x pode ser escrito como $x = \alpha y + (1 - \alpha)z$, sendo $y, z \in K$ e $\alpha \in (0,1)$. Como x, y , e z são não negativos, e y e z também são solução do sistema, logo temos que

$$y_1 A_1 + y_2 A_2 + \dots + y_m A_m = b$$

e

$$z_1 A_1 + z_2 A_2 + \dots + z_m A_m = b$$

portanto, teremos que $x = y = z$. Ou seja, uma contradição. dessa maneira, x é um vértice.

Na recíproca teremos que, por hipótese, x é um vértice de K . Como x é solução, e suponhamos que as k primeiras entradas são diferentes de zero, então

$$x_1 A_1 + x_2 A_2 + \dots + x_k A_k = b.$$

Para mostrar que x é uma solução básica factível, precisamos concluir que $\{A_1, \dots, A_k\}$ é L.I. Portando, por absurdo, suponhamos que exista uma combinação linear não-trivial

$$y_1 A_1 + y_2 A_2 + \dots + y_k A_k = 0.$$

Definindo $y = (y_1, \dots, y_k, 0, \dots, 0) \in \mathbb{R}^n$, como $x_i > 0$, $i = 1, 2, \dots, k$. Tomando ε suficientemente pequeno, para que

$$x_i + \varepsilon y_i \geq 0 \text{ e } x_i - \varepsilon y_i \geq 0, \quad i = 1, 2, \dots, k,$$

teremos então que $x = \frac{1}{2}(x + \varepsilon y) + \frac{1}{2}(x - \varepsilon y)$, logo x está sendo escrito como uma combinação convexa de vetores distintos de K . Contradizendo o fato de x ser um vértice de K . Logo $\{A_1, \dots, A_k\}$ deve ser linearmente independente e, portanto, x é uma solução básica factível. □

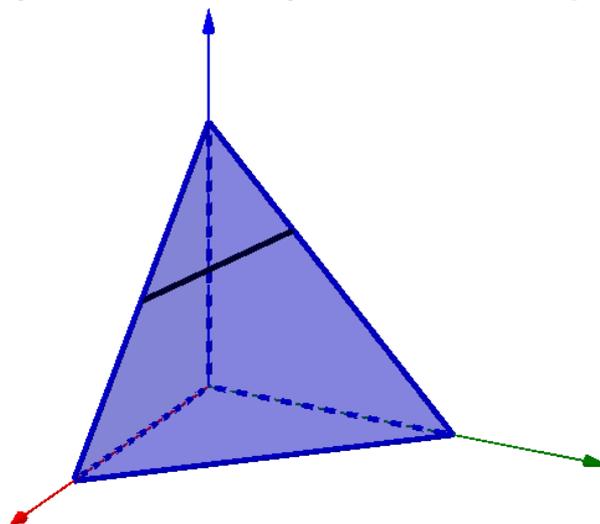
Com base no Teorema 34 e no Teorema 32, obtemos que se existem soluções ótimas factíveis, então elas estão em vértices do conjunto factível. Sendo assim, precisaremos analisar o valor da função objetivo nos vértices do conjunto factível. Como há um número limitado de semiplanos (inequações desigualdades) e retas (equações) que formam o conjunto solução K , é fato de que a quantidade de vértices é finita.

Por exemplo, vamos analisar as restrições do problema de programação linear a seguir

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

As restrições escritas no sistema (4) representam a intersecção de dois planos e três semi-espacos, gerando um segmento no espaço como ilustramos na Figura 12.

Figura 12 – Exemplo geométrico das restrições.



Fonte: Autoria própria.

Agora, iniciaremos um estudo sobre o Método Simplex. Tal método, busca soluções mínimas em funções lineares, baseado na geometria simplética que introduzimos no Capítulo 2. Veremos que para encontrar uma solução mínima, em uma função linear com restrições lineares, utilizaremos apenas processos já conhecidos da Álgebra Linear.

3.4 Método Simplex

Desenvolvido por Gerge B. Dantzig em 1947, o Método Simplex foi um dos mais importantes métodos feito no período pós-guerra. O método foi publicado pela primeira vez na resolução do problema da dieta. O problema consiste em minimizar o custo na alimentação de uma pessoa satisfazendo as necessidades nutricionais básicas. Tal assunto era muito importante no período proposto, uma vez que o custo de um soldado em campo era relativamente alto. Tal método, consiste em uma busca iterativa de soluções básicas de um Problema de Programação Linear. Ou seja, encontrar um ponto extremo no conjunto de restrições. Para isso, o método busca diminuir, a cada iteração, os valores da função objetivo.

3.4.1 Pivoteamento

Para a compreensão do processo, temos que analisar o pivoteamento em um conjunto de equações lineares. Iremos descrever, portanto, o pivoteamento baseado nas colunas da matriz referente as restrições.

Já vimos que um PPL pode ser representado de duas maneiras, sendo a forma matricial dada a seguir

$$Ax = b, \text{ com } x \geq 0$$

no qual $A \in \mathbb{R}^{n \times m}$, $x \in \mathbb{R}^n$ e b um vetor de $b \in \mathbb{R}^m$. Outra maneira de escrever esse problema, seria utilizando um sistema de equações da seguinte forma:

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

ou ainda, podemos utilizar as colunas da matriz A . Se denotarmos por A_i sendo a i -ésima coluna de A , teremos que

$$x_1A_1 + x_2A_2 + \cdots + x_nA_n = b. \tag{5}$$

Isto é, b pode ser escrito como uma combinação linear das colunas A_1, \dots, A_n . Notemos que, se A tem posto m , as colunas geram o subespaço $\mathbb{E} \subset \mathbb{R}^m$ e, se $m < n$, as colunas são necessariamente *L.D.* Logo, temos que há mais de uma solução para a Equação (5).

Suponhamos agora que começaremos o sistema com as m primeiras colunas da forma canônica. Portanto teríamos que:

$$\begin{array}{cccccccc} x_1 & + & 0 & + & \cdots & + & 0 & + & a_{1\ m+1}x_{m+1} & + & \cdots & + & a_{1\ n}x_n & = & b_1 \\ 0 & + & x_2 & + & \cdots & + & 0 & + & a_{2\ m+1}x_{m+1} & + & \cdots & + & a_{2\ n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots & & \ddots & & \vdots & & \\ 0 & + & 0 & + & \cdots & + & x_m & + & a_{m\ m+1}x_{m+1} & + & \cdots & + & a_{m\ n}x_n & = & b_m \end{array} \Leftrightarrow Ax = b$$

com $a_{ij}, b_i \in \mathbb{R}$ fixados. Podemos representar toda coluna A_i , com $m + 1 \leq i \leq n$, como uma combinação linear das m primeiras colunas de A , conforme se segue:

$$A_j = a_{1,j}A_1 + a_{2,j}A_2 + \cdots + a_{m,j}A_m \tag{6}$$

sendo os vetores A_j , com $1 \leq j \leq m$, as m primeiras colunas de A e $a_{ji} \in \mathbb{R}$ para $1 \leq j \leq m$. Para facilitar a compreensão, iremos construir um *Tableau* com as informações do sistema.

$$\begin{array}{ccccccc} 1 & 0 & \cdots & 0 & a_{1,m+1} & \cdots & a_{1,n} & b_1 \\ 0 & 1 & \cdots & 0 & a_{2,m+1} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{m,m+1} & \cdots & a_{m,n} & b_m \end{array}$$

Podemos substituir a coluna A_p , com $1 \leq p \leq m$, pela coluna A_q , com $m + 1 \leq q \leq n$, mantendo a independência com as outras $m - 1$ primeiras colunas de A . Observe no *Tableau* que podemos representar A_q da seguinte maneira:

$$A_q = \sum_{i=1 \text{ e } i \neq p}^m a_{i,q} A_i + a_{p,q} A_p.$$

Com isso podemos escrever A_p como se segue

$$A_p = \frac{1}{a_{p,q}} A_q - \sum_{i=1 \text{ e } i \neq p} \frac{a_{i,q}}{a_{p,q}} A_i. \quad (7)$$

Observe que $a_{p,q} \neq 0$, então podemos substituir a Equação (7) na Equação (6). Dessa forma, teremos que

$$\begin{aligned} A_j &= \sum_{i=1 \text{ e } i \neq p}^m a_{i,j} A_i + \left(\frac{a_{p,j}}{a_{p,q}} A_q - \sum_{i=1 \text{ e } i \neq p} \frac{a_{p,j} a_{i,q}}{a_{p,q}} A_i \right) \Leftrightarrow \\ \Leftrightarrow A_j &= \sum_{i=1 \text{ e } i \neq p}^m \left(a_{i,j} - \frac{a_{i,q} a_{p,j}}{a_{p,q}} \right) A_i + \frac{a_{p,j}}{a_{p,q}} A_q \end{aligned}$$

Com isso, temos uma maneira de calcular cada entrada da nova matriz. Assim, denotaremos cada entrada da nova matriz por $a'_{i,j}$, conforme é dado a seguir.

$$\begin{cases} a'_{i,j} = a_{i,j} - \frac{a_{i,q} a_{p,j}}{a_{p,q}} & \text{se } i \neq p \\ a'_{i,j} = \frac{a_{p,j}}{a_{p,q}} & \text{se } i = p. \end{cases}$$

Veremos a seguir um exemplo, conforme em Luenberger e Ye (2008), para analisarmos como é feita a troca de pivôs. Para isso, seja o seguinte sistema de equações lineares:

$$\begin{cases} x_1 + x_2 - x_3 \leq 5 \\ 2x_1 - 3x_2 + x_3 \leq 3 \\ -x_1 + 2x_2 - x_3 \leq -1 \end{cases}$$

Adicionando as variáveis de folga, que neste momento chamaremos de e_i , teremos que:

$$\begin{cases} e_1 & +x_1 + x_2 - x_3 + & = 5 \\ e_2 & +2x_1 - 3x_2 + x_3 + & = 3 \\ e_3 & -x_1 + 2x_2 - x_3 + & = -1 \end{cases}$$

e, para facilitar o processo, colocaremos os valores em um *Tableau*

e_1	e_2	e_3	x_1	x_2	x_3	b
1	0	0	1	1	-1	5
0	1	0	2	-3	1	3
0	0	1	-1	2	-1	-1

sendo assim podemos aplicar a mudança de coeficientes que vimos anteriormente. Se pivotarmos a primeira entrada da coluna x_1 , que está circulado, obtemos o seguinte *Tableau*:

$$\begin{array}{ccccccc}
 e_1 & e_2 & e_3 & x_1 & x_2 & x_3 & b \\
 1 & 0 & 0 & 1 & 1 & -1 & 5 \\
 -2 & 1 & 0 & 0 & \textcircled{-5} & 3 & -7 \\
 1 & 0 & 1 & 0 & 3 & -2 & 4
 \end{array}$$

repetindo o processo, mas agora pivotando a segunda entrada de x_2 , que também está circulado, obteremos o novo *Tableau*

$$\begin{array}{ccccccc}
 e_1 & e_2 & e_3 & x_1 & x_2 & x_3 & b \\
 3/5 & 1/5 & 0 & 1 & 0 & -2/5 & 18/5 \\
 2/5 & -1/5 & 0 & 0 & 1 & -3/5 & 7/5 \\
 -1/5 & 3/5 & 1 & 0 & 0 & \textcircled{-1/5} & -1/5
 \end{array}$$

por fim, o próximo pivô será a última entrada da coluna x_3 .

$$\begin{array}{ccccccc}
 e_1 & e_2 & e_3 & x_1 & x_2 & x_3 & b \\
 1 & -1 & -2 & 1 & 0 & 0 & 4 \\
 1 & -2 & -3 & 0 & 1 & 0 & 2 \\
 1 & -3 & -5 & 0 & 0 & 1 & 1.
 \end{array}$$

Logo, para esta nova forma canônica, teremos uma nova solução básica, com $x_1 = 4$, $x_2 = 2$ e $x_3 = 1$, conforme dada a seguir.

$$x = (0,0,0,4,2,1).$$

Perceba que apenas mudamos de uma solução básica para outra e isso será de suma importância no procedimento do Método Simplex.

Anteriormente, vimos que se há uma solução ótima factível então existe solução ótima básica factível. Agora, provamos que o pivoteamento gera novas soluções básicas, baseado na solução básica anterior, apenas trocando uma variável básica por uma variável não básica. Porém, a não negatividade da solução não necessariamente é mantida. Para tal, temos que utilizar condições especiais, em que se feitas na ordem correta, a factibilidade é mantida.

Embora não possamos dizer certamente qual o par de variáveis deve sair para manter a não negatividade, é possível, arbitrariamente, escolher qual variável não básica deve se tornar básica e, então, qual variável básica deve se tornar não básica.

Vamos considerar a próxima definição por pura conveniência, pois todos os argumentos podem ser utilizados para incluir as soluções degeneradas. O Método Simplex em si, pode ser facilmente modificado para explicá-la.

Definição 35. *Toda solução básica factível é uma solução básica factível não degenerada.*

Isto é, nenhuma variável básica irá ter o valor zero. Com isso, definiremos qual vetor irá deixar a base em um PPL.

Suponha que $x = (x_1, x_2, \dots, x_m, 0, \dots, 0)$ seja uma solução básica factível. Pela definição anterior, temos que $x_i > 0$ para todo $0 \leq i \leq m$. Representaremos o vetor A_q , com $q > m$, em função dos vetores A_i com $0 \leq i \leq m$, conforme dado a seguir.

$$A_q = a_{1q}A_1 + a_{2q}A_2 + \dots + a_{mq}A_m \Leftrightarrow \varepsilon A_q - \varepsilon a_{1q}A_1 - \varepsilon a_{2q}A_2 - \dots - \varepsilon a_{mq}A_m = 0.$$

Sabemos que $x_1A_1 + x_2A_2 + \dots + x_mA_m = b$. Logo, para todo ε , temos que

$$(x_1 - \varepsilon a_{1q})A_1 + (x_2 - \varepsilon a_{2q})A_2 + \dots + (x_m - \varepsilon a_{mq})A_m + \varepsilon A_q = b. \quad (8)$$

Para todo $\varepsilon \geq 0$ temos uma combinação linear de b com no máximo $m + 1$ vetores. Se $\varepsilon = 0$ teremos a solução básica de x definida anteriormente. Agora, se $\varepsilon > 0$, o vetor A_q será considerado e, mesmo sendo factível, a solução não será básica, uma vez que $\{A_1, \dots, A_m, A_q\}$ é *L.D.* Logo, conforme o valor de ε é modificado, o coeficiente relacionado pode diminuir ou aumentar. Considerando ε correto, podemos tornar a Equação (8) uma solução básica factível. Para isso, definiremos:

$$\varepsilon = \min_i \left\{ \frac{x_i}{a_{iq}} \mid a_{iq} > 0 \right\}. \quad (9)$$

Assim, teremos uma nova solução básica factível com o vetor A_q , substituindo o vetor A_p , com p o índice minimizador de (9). Caso exista mais de um índice que possa ser escolhido, então qualquer um desses vetores poderão deixar a base.

Se a Equação (8) tem todos os a_{iq} que não são positivos, mesmo modificando o valor de ε , não teremos novas soluções básicas factíveis. Este é um caso especial no qual existem infinitas soluções factíveis possíveis para o sistema.

Veremos como pode ser utilizado o pivoteamento em um problema genérico, isto é,

$$\begin{aligned} Ax &= b \\ x &\geq 0. \end{aligned}$$

Teremos o seguinte *Tableau* relacionada ao problema:

$$\begin{array}{cccccccc} 1 & 0 & \cdots & 0 & a_{1,m+1} & \cdots & a_{1,n} & b_1 \\ 0 & 1 & \cdots & 0 & a_{2,m+1} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{m,m+1} & \cdots & a_{m,n} & b_m \end{array}$$

Esse *Tableau* será utilizado para calcular o pivô de saída. Temos que $x_1 = b_1, \dots, x_m = b_m$ é uma solução factível. Para substituir um vetor A_q com $q > m$, calcularemos b_i/a_{iq} para $i = 1, \dots, m$. Logo, teremos que o pivô será o elemento a_{iq} , com q sendo o vetor que deixará a base e i o índice referente ao menor valor não negativo de b_i/a_{iq} .

Dessa forma, será feita a derivação do método que foi apresentado anteriormente para encontrar soluções básicas factíveis. Para isso, depois de escolher o pivô a_{iq} , dividiremos a i –ésima linha por a_{iq} para tornar o pivô unitário. Em seguida, subtraímos cada linha do *Tableau* pela nova i –ésima linha, buscando zerar todos os elementos da q –ésima coluna. Ou seja, se escolhessemos o pivô $a_{1,m+1}$ do *Tableau* anterior, teremos o seguinte novo *Tableau*:

$$\begin{array}{cccccccc} 1/a_{1,m+1} & 0 & \cdots & 0 & 1 & \cdots & a_{1,n}/a_{1,m+1} & b_1/a_{1,m+1} \\ a'_{2,1} & 1 & \cdots & 0 & 0 & \cdots & a'_{2,n} & b'_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a'_{2,m} & 0 & \cdots & 1 & 0 & \cdots & a'_{m,n} & b'_m \end{array}$$

gerando uma nova solução básica factível.

Considere as seguintes restrições:

$$\begin{array}{cccccc} e_1 & e_2 & e_3 & x_1 & x_2 & x_3 & b \\ 1 & 0 & 0 & 1 & 1 & -1 & 5 \\ 0 & 1 & 0 & 2 & -3 & 1 & 3 \\ 0 & 0 & 1 & -1 & 2 & -1 & -1 \end{array}$$

se quiséssemos que o vetor x_1 entrasse na base, teríamos que calcular qual seria o pivô referente a tal substituição. Obtemos, assim, as opções

$$5/1 = 1, \quad 3/2 = 1.5, \quad -1/-1 = 1,$$

escolhendo a menor razão não negativa, então o pivô seria o terceiro elemento do vetor x_1 , que está circulado no *Tableau*

$$\begin{array}{cccccc} e_1 & e_2 & e_3 & x_1 & x_2 & x_3 & b \\ 1 & 0 & 0 & 1 & 1 & -1 & 5 \\ 0 & 1 & 0 & 2 & -3 & 1 & 3 \\ 0 & 0 & 1 & \textcircled{-1} & 2 & -1 & -1 \end{array}$$

dividindo a última linha pelo pivô. Neste caso, por -1

$$\begin{array}{cccccc} e_1 & e_2 & e_3 & x_1 & x_2 & x_3 & b \\ 1 & 0 & 0 & 1 & 1 & -1 & 5 \\ 0 & 1 & 0 & 2 & -3 & 1 & 3 \\ 0 & 0 & -1 & 1 & -2 & 1 & 1 \end{array}$$

para zerar os valores acima do pivô, iremos somar a primeira linha por -1 vezes a última linha

$$\begin{array}{cccccc}
 e_1 & e_2 & e_3 & x_1 & x_2 & x_3 & b \\
 1 & 0 & 1 & 0 & 1 & 4 & 4 \\
 0 & 1 & 0 & 2 & -3 & 1 & 3 \\
 0 & 0 & -1 & 1 & -2 & 1 & 1
 \end{array}$$

para finalizar, somaremos a segunda linha por -2 vezes a última linha

$$\begin{array}{cccccc}
 e_1 & e_2 & e_3 & x_1 & x_2 & x_3 & b \\
 1 & 0 & 1 & 0 & 1 & 4 & 4 \\
 0 & 1 & 2 & 0 & 1 & -1 & 1 \\
 0 & 0 & -1 & 1 & -2 & 1 & 1
 \end{array}$$

e assim, finalizando o processo. Os valores de cada coluna do *Tableau* correspondem aos coeficientes que acompanham as variáveis escritas na parte superior. As colunas e_1 , e_2 e x_1 tem apenas uma entrada com valor 1 e as outras 0, por exemplo, na primeira equação a variável e_1 tem coeficiente 1, já as variáveis e_2 e x_1 tem coeficientes *zero*. Portanto, podemos zerar as variáveis e_3 , x_2 e x_3 , impondo que $e_1 = 4$. Além disso, para a primeira linha da equação, os valores de e_2 e x_1 não interferem na igualdade. Podemos repetir o mesmo pensamento para a segunda e terceira equação, obtendo assim uma solução factível $x = (4, 1, 0, 1, 0, 0)$.

3.4.2 Interpretação Geométrica

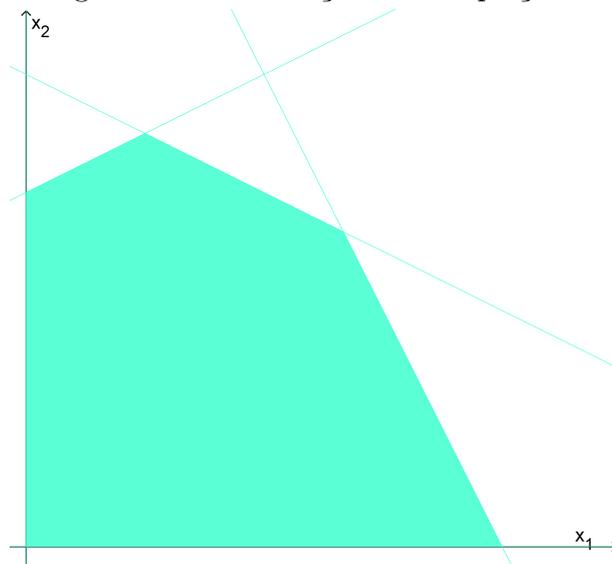
Temos duas maneiras de encontrar soluções básicas factíveis. Entretanto, as duas formas de buscar as tais soluções tem a mesma interpretação geométrica. Vimos anteriormente a definição de um conjunto simplex e sua característica geométrica. Baseado nisso, iremos analisar o Método Simplex junto com as soluções básicas factíveis.

Suponhamos um PPL em \mathbb{R}^2 , com as restrições da seguinte forma:

$$\begin{aligned}
 -x_1 + 2x_2 &\leq 3 \\
 10x_1 + 5x_2 &\leq 20 \\
 x_1 + 2x_2 &\leq 4 \\
 x_1, x_2 &\geq 0
 \end{aligned} \tag{10}$$

Sabemos que cada inequação gera um semiplano e, portanto, a intersecção de todos os semiplanos resultará em um polígono conforme mostra a Figura 13 a seguir.

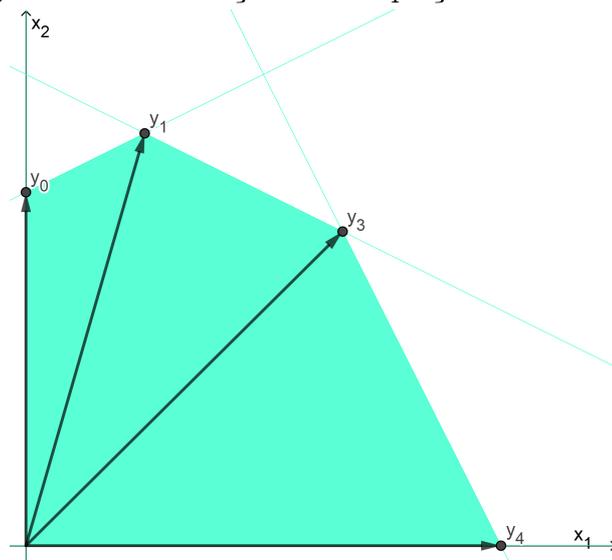
Figura 13 – Intersecções de inequações.



Fonte: Autoria própria.

Este polígono é a representação geométrica do conjunto factível. Ou seja, é o conjunto K das soluções que satisfazem as restrições. No Teorema 31, temos que se há solução ótima factível então há solução ótima básica factível. Além disso, no Teorema 34 mostramos que toda solução básica factível é um vértice do conjunto solução. Portanto, concluímos que a solução ótima encontra-se em um dos vértices do conjunto factível. Dessa forma, queremos encontrar os seguintes vetores:

Figura 14 – Intersecções de inequações com vetores.



Fonte: Autoria própria.

Perceba que os vetores dois a dois geram um 2 – *simplex*. Além disso, a cada iteração, é encontrado um dos vetores y_i .

O Método Simplex busca encontrar um mínimo da função objetivo, analisando apenas os vértices do conjunto factível. Assim, iremos estabelecer uma maneira prática de analisar o valor da função no vértice. E, diante disso, avaliar se o valor encontrado é o mínimo.

3.4.3 Encontrar o Mínimo da Função Objetivo

Vimos anteriormente que podemos definir um PPL como

$$\begin{aligned} &\text{minimizar} && c^T x \\ &\text{sujeito a} && Ax = b \\ &&& e \quad x \geq 0 \end{aligned} \tag{11}$$

sendo $x, c^T \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$. Sabemos que a função objetivo é dada por $z = c_1x_1 + c_2x_2 + \dots + c_nx_n$. Além disso, o PPL (11) corresponde ao seguinte *Tableau*:

$$\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & a_{1 \ m+1} & \cdots & a_{1 \ n} & b_1 \\ 0 & 1 & \cdots & 0 & a_{2 \ m+1} & \cdots & a_{2 \ n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_{m \ m+1} & \cdots & a_{m \ n} & b_m \end{array}$$

uma solução básica será $(x_b, 0) = (b_1, b_2, \dots, b_m, 0, \dots, 0)$. Sendo assim, teremos $z_0 = c_1b_1 + c_2b_2 + \dots + c_mb_m$, sendo o valor da função objetivo avaliada em $(x_b, 0)$. Se tomarmos um vetor $x_c = (x_{m+1}, \dots, x_n)$ com valores de cada entrada arbitrários, teremos que:

$$\begin{aligned} x_1 &= b_1 - \sum_{m+1}^n a_{1,j}x_j \\ x_2 &= b_2 - \sum_{m+1}^n a_{2,j}x_j \\ &\vdots \\ x_m &= b_m - \sum_{m+1}^n a_{m,j}x_j \end{aligned} \tag{12}$$

se substituirmos as variáveis de (12) na função objetivo, então

$$\begin{aligned} z &= c_1x_1 + \dots + c_nx_n \Leftrightarrow \\ \Leftrightarrow z &= c_1(b_1 - \sum_{m+1}^n a_{1,j}x_j) + \dots + c_m(b_m - \sum_{m+1}^n a_{m,j}x_j) + c_{m+1}x_{m+1} + \dots + c_nx_n. \end{aligned}$$

Além disso, se considerarmos $z_j = a_{1j}c_1 + a_{2j}c_2 + \dots + a_{mj}c_m$ para $m + 1 \leq j \leq n$, então concluímos que

$$z = z_0 + (c_{m+1} - z_{m+1})x_{m+1} + \dots + (c_n - z_n)x_n. \tag{13}$$

Essa relação, com a função objetivo, é fundamental para o pivoteamento. Pois se para algum $m + 1 \leq j \leq n$ temos que $c_j - z_j$ for negativo, podemos então alterar o valor de x_j para um valor positivo, diminuindo o resultado da função objetivo avaliada neste novo ponto.

Portanto, agora iremos demonstrar um teorema que garante a melhoria baseada na construção feita anteriormente.

Teorema 36. *Dado uma solução básica factível não degenerada correspondendo a z_0 , suponhamos que para algum j temos que $c_j - z_j < 0$. Então existe uma solução básica factível tal que $z < z_0$. Se a coluna A_j pode ser substituída por algum vetor na base original para produzir uma nova solução básica factível, esta nova solução resultará em $z < z_0$. Se A_j não pode ser substituída para uma nova solução básica factível, então o conjunto solução K é ilimitado e a função objetivo pode ser tão pequena quanto queremos.*

Demonstração. Seja $(x_1, x_2, \dots, x_m, 0, \dots, 0)$ uma solução básica factível, com o valor da função objetivo neste ponto sendo z_0 . Suponhamos $c_{m+1} - z_{m+1} < 0$, podemos escrever uma nova solução factível $(x'_1, x'_2, \dots, x'_m, x'_{m+1}, 0, \dots, 0)$ com $x'_{m+1} > 0$. Pela equação (13), substituindo a nova solução, obtemos que

$$z - z_0 = (c_{m+1} - z_{m+1})x'_{m+1} < 0 \Rightarrow z < z_0.$$

Se aumentarmos o valor x'_{m+1} as outras entradas podem diminuir, aumentar ou permanecer constante. Se existir um valor de x'_{m+1} tal que, $x'_i = 0$ para algum $i \leq m$. Então, esse valor gera uma nova solução básica factível. Caso nenhum dos x'_i diminua, temos que o conjunto solução é ilimitado. Portanto, a função objetivo pode ser tão pequena quanto quisermos. \square

Já sabemos como podemos diminuir o valor da função objetivo utilizando o Teorema 36. Porém, ainda não sabemos quando a solução ótima é encontrada. Para isso, iremos utilizar o teorema a seguir.

Teorema 37 (Condição de Otimalidade). *Se para alguma solução básica factível temos $c_j - z_j \geq 0$ para todo j então será uma solução ótima básica factível.*

Demonstração. A demonstração é imediata pela Equação (13). Pois, se houver qualquer outra solução factível, deveríamos ter $x_i \geq 0$. Se avaliarmos essa nova solução, teremos que $z - z_0 \geq 0$, e portanto $z \geq z_0$. \square

Dessa forma, criamos uma maneira de avaliar a função objetivo nas soluções básicas encontradas. Observamos, então, a importância de $c_j - z_j$ para o desenvolvimento e a aplicação do Método Simplex. Por conveniência, chamaremos de r_j de coeficiente de custo ou coeficiente de redução de custo. Esse coeficiente mensura o custo da entrada relativo à uma dada base.

O Método Simplex como observamos é um método de redução de custo dado em um PPL

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

perceba que a redução de custo não está necessariamente relacionada a ganhos financeiros, como é muito comum em Problemas de Programação Linear, mas em diversas interpretações diferentes da aplicação do Método Simplex.

Um exemplo muito utilizado é o problema da dieta, o qual podemos representar como um PPL, em que cada coluna de A é o valor nutricional de um alimento. Se $A \in \mathbb{R}^{m \times n}$ então estaremos analisando m valores nutricionais em n alimentos. Ainda, x será a quantidade de cada alimento que deverá ser consumido. Por fim, b será o mínimo de nutrientes que uma pessoa deve ingerir.

Se o feijão for uma comida que não está entrando na dieta, então podemos utilizar as informações do *Tableau* do simplex para criar um feijão sintético com valores nutricionais equivalentes ao do feijão propriamente dito. Além disso, esse feijão é uma combinação de outros alimentos em quantidades corretas.

Para analisar se a dieta é a melhor ou não, consideraremos um alimento que não está na base da dieta, como por exemplo o feijão, e observaremos se é vantajoso trazê-lo para a dieta. Se o feijão é o alimento da coluna j do *Tableau*, temos que o custo nutricional do alimento é dado por c_j . Mas o feijão sintético criado anteriormente é descrito por:

$$z_j = \sum_{i=1}^m c_i a_{ij}$$

Notemos que se $c_j - z_j < 0$ então é vantajoso incluir feijão na dieta. Ou seja, a coluna j entra na base.

Os Problemas de Programação Lineares estudados até agora são os problemas da forma $Ax = b$. Além disso, esses problemas tem a forma canônica e, portanto, é fácil de encontrarmos uma solução básica factível inicial. Agora, além do *Tableau* representar $Ax = b$ na forma canônica, adicionaremos mais uma linha correspondente ao custo relativo e o negativo do custo atual, obtendo, assim

A_1	A_2	\dots	A_m	$A_m + 1$	\dots	A_n	b
1	0	\dots	0	a_{1m+1}	\dots	a_{1n}	b_1
0	1	\dots	0	a_{2m+1}	\dots	a_{2n}	b_2
\vdots	\vdots	\ddots	\vdots	\vdots	\ddots	\vdots	\vdots
0	0	\dots	1	a_{mm+1}	\dots	a_{mn}	b_m
0	0	\dots	0	$r_m + 1$	\dots	r_n	$-z_0$

Uma solução básica factível é $(b_1, b_2, \dots, b_m, 0, \dots, 0)$. Podemos escrever esta solução de outra maneira, isto é,

$$x_i = \begin{cases} b_i & \text{se } 0 \leq i \leq m \\ 0 & \text{se } m + 1 \leq i \leq n. \end{cases}$$

O coeficiente de custo relativo r_j indica o quanto mudará na função objetivo se um vetor x_j entrar na base. Quando todos os coeficientes r_j forem maiores ou iguais a zero, então a solução encontrada é a solução ótima factível. Porém, se existir r_j negativo, então a solução pode ser melhorada colocando x_j na solução. Caso houver mais de um r_j menor do que zero, então podemos escolher qualquer um desses elementos para entrar na base e criar uma nova solução, entretanto o mais comum a ser escolhido é o menor valor negativo.

Para melhor justificar a última linha do novo *Tableau*, iremos voltar a discussão da função objetivo. Sabemos que

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = z \Leftrightarrow c_1x_1 + c_2x_2 + \cdots + c_nx_n - z = 0.$$

Assim, em um novo sistema teremos $m + 1$ variáveis básicas, em que podemos exigir que z seja uma delas. Por esse motivo que não implementamos uma coluna exclusivamente para z , pois ela seria da forma $(0, 0, \dots, 1)$.

Inicialmente, uma linha contendo os valores c_i pode ser adicionada na matriz padrão, em que teremos uma entrada na direita sendo zero para representar a equação adicional. Usando as operações de pivoteamento, conseguiremos anular os elementos desta nova linha que estão relacionadas com as variáveis básicas. Obteremos que:

$$r_{m+1}x_{m+1} + r_{m+2}x_{m+2} + \cdots + r_nx_n - z = -z_0,$$

com os r_i sendo os coeficientes de custos relativos. Dessa forma, poderemos trabalhar com a nova linha adicionada na matriz como qualquer outra linha. Quando é feito o pivoteamento, utilizando o Método Simplex, o valor da função avaliada na nova solução básica factível aparecerá no último elemento da matriz, por exemplo, seja o PPL

$$\begin{array}{ll} \text{minimizar} & -3x_1 - x_2 - 3x_3 \\ \text{sujeito a} & 2x_1 + x_2 + x_3 \leq 2 \\ & x_1 + 2x_2 + 3x_3 \leq 5 \\ & 2x_1 + 2x_2 + x_3 \leq 6 \\ & x_1, x_2, x_3 \geq 0 \end{array}$$

para transformar esse PPL no padrão do Método Simplex, adicionaremos as variáveis básicas x_4, x_5 e x_6 .

$$\begin{aligned}
 &\text{minimizar} && -3x_1 - x_2 - 3x_3 + x_4 + x_5 + x_6 \\
 &\text{sujeito a} && 2x_1 + x_2 + x_3 + x_4 = 2 \\
 &&& x_1 + 2x_2 + 3x_3 + x_5 = 5 \\
 &&& 2x_1 + 2x_2 + x_3 + x_6 = 6 \\
 &&& x_1, x_2, x_3, x_4, x_5 \geq 0
 \end{aligned}$$

colocando em um *Tableau*, obteremos

	x_1	x_2	x_3	x_4	x_5	x_6	b
	2	1	1	1	0	0	2
	1	2	3	0	1	0	5
	2	2	1	0	0	1	6
r_j	-3	-1	-3	0	0	0	0

Agora, poderemos fazer o pivoteamento. Para tal, temos que os coeficientes de redução $r_1, r_2, r_3 < 0$. Portanto, podemos escolher qualquer uma dessas colunas para fazer o pivoteamento. Escolhendo a coluna relacionada a r_2 , teremos que o pivô será o primeiro elemento dessa coluna, ou seja, o elemento circulado

	x_1	x_2	x_3	x_4	x_5	x_6	b
	2	①	1	1	0	0	2
	1	2	3	0	1	0	5
	2	2	1	0	0	1	6
r_j	-3	-1	-3	0	0	0	0

podemos somar a segunda linha pela primeira multiplicada por -2 . Resultando em

	x_1	x_2	x_3	x_4	x_5	x_6	b
	2	①	1	1	0	0	2
	-3	0	1	-2	1	0	1
	2	2	1	0	0	1	6
r_j	-3	-1	-3	0	0	0	0

com o mesmo processo, somando a última linha pela primeira multiplicada por -2 . Obtendo então

	x_1	x_2	x_3	x_4	x_5	x_6	b
	2	1	1	1	0	0	2
	-3	0	1	-2	1	0	1
	-2	0	-1	-2	0	1	2
r_j	-1	0	-2	1	0	0	2

continuamos com $r_1, r_3 < 0$. Assim, podemos escolher uma dessas duas colunas relacionadas. Escolhendo a coluna relacionada a r_3 , temos que o pivô será o segundo elemento da terceira coluna. Ou seja, o elemento circulado no *Tableau*

	x_1	x_2	x_3	x_4	x_5	x_6	b
	2	1	1	1	0	0	2
	-3	0	①	-2	1	0	1
	-2	0	-1	-2	0	1	2
r_j	-1	0	-2	1	0	0	2

fazendo o pivôteamento, o mesmo processo anterior, teremos

	x_1	x_2	x_3	x_4	x_5	x_6	b
	5	1	0	3	-1	0	1
	-3	0	1	-2	1	0	1
	-5	0	0	-4	1	1	3
r_j	-7	0	0	-3	2	0	4

Agora, só resta o elemento $r_1 < 0$. Com isso, o pivô será o primeiro elemento da primeira coluna. E assim,

	x_1	x_2	x_3	x_4	x_5	x_6	b
	1	1/5	0	3/5	-1/5	0	1/5
	0	3/5	1	-1/5	2/5	0	8/5
	0	1	0	-1	0	1	4
r_j	0	7/5	0	6/5	3/5	0	27/5

Como vimos anteriormente, as colunas relacionadas as variáveis x_1 , x_3 e x_6 , tem uma entrada com valor 1 e as outras 0. Ou seja, podemos tomar $x_2 = x_4 = x_5 = 0$ e igualar $x_1 = 1/5$, $x_3 = 8/5$ e $x_6 = 4$. Concluindo que a solução ótima factível será $(1/5, 0, 8/5, 0, 0, 4)$. Além disso, pela construção do *Tableau*, temos que a última entrada da última coluna é o valor da função objetivo avaliado na solução básica factível gerada pelo procedimento de pivoteamento. Ou seja, o valor mínimo da função objetivo será $-27/5$. Perceba que $r_1 = r_3 = 0$, e se caso quiséssemos continuar o procedimento do Método Simplex escolhendo a primeira ou a terceira coluna para trabalharmos, entraríamos em um ciclo até chegarmos novamente nesse valor mínimo da função objetivo. Isso ocorre pois a solução ótima é degenerada. Por esse motivo, o método parou quando $r_1, r_2, r_3 \geq 0$.

3.4.4 Método Simplex de Duas Fases

Até então, estudamos Problemas de Programação Lineares nos quais tínhamos $A = [B \ I]$, com B sendo uma matriz $m \times m$ e I a matriz identidade. Com isso, obtivemos

facilmente uma solução básica inicial. Entretanto, se essa solução inicial não for óbvia, teremos que encontrar uma solução básica inicial, para assim aplicar o Método Simplex.

Veremos agora que podemos resolver um PPL auxiliar, utilizando o Método Simplex, a fim de encontrar uma solução básica fictível não trivial. Por meio das operações simples e elementares vistas anteriormente, as restrições de um PPL podem ser vistas como:

$$\begin{aligned} \text{sujeito a } \quad Ax &= b \\ \text{e } x &\geq 0 \end{aligned}$$

Para encontrar uma solução factível para esse problema, vamos considerar um Problema Artificial de Programação Linear da seguinte forma:

$$\begin{aligned} \text{minimizar } \quad & \sum_{i=1}^m y_i \\ \text{sujeito a } \quad & Ax + y = b \\ & x \geq 0 \\ & y \geq 0 \end{aligned} \tag{14}$$

com $y = (y_1, y_2, \dots, y_m)$ o vetor de variáveis artificiais. Se o problema original tiver uma solução factível então o problema artificial também terá ($y = 0$). Se o problema inicial não tiver solução factível então o problema artificial terá o valor mínimo maior que zero.

Perceba que o problema artificial já está na forma canônica com variáveis x e y . Além disso, com esse novo problema, temos que uma solução inicial factível é dada por $y = b$. Podemos aplicar o Método Simplex para obter soluções básicas factíveis em cada etapa do método. Quando o mínimo encontrado em (14) teremos que as variáveis artificiais serão nulas, impedindo a degeneração, uma vez que não haverá variável básica y . Caso houver algum y_i que seja zero e básico, isso indica uma solução básica degenerada. Com isso, utilizaremos as variáveis não básicas x_i para substituir y_i e assim criar uma solução que dependa apenas de x .

Se tivermos um PPL com as seguintes restrições

$$\begin{aligned} 2x_1 + x_2 + 2x_3 &= 4 \\ 3x_1 + 3x_2 + x_3 &= 3 \\ x_1, x_2, x_3 &\geq 0 \end{aligned} \tag{15}$$

então adicionaremos as variáveis artificiais x_4 e x_5 e a função objetivo será dado por $x_4 + x_5$, teremos um PPL artificial da seguinte maneira

$$\begin{aligned} \text{minimizar } \quad & x_4 + x_5 \\ \text{sujeito a } \quad & 2x_1 + x_2 + 2x_3 + x_4 = 4 \\ & 3x_1 + 3x_2 + x_3 + x_5 = 3 \\ & x_1, x_2, x_3, x_4, x_5 \geq 0. \end{aligned}$$

Agora, para resolvermos o PPL artificial, vamos escrevê-lo em um *Tableau*

$$\begin{array}{c|cccccc}
 & x_1 & x_2 & x_3 & x_4 & x_5 & b \\
 \hline
 & 2 & 1 & 2 & 1 & 0 & 4 \\
 & 3 & 3 & 1 & 0 & 1 & 3 \\
 \hline
 c_j & 0 & 0 & 0 & 1 & 1 & 0
 \end{array}$$

Perceba que na última linha teremos os coeficientes de custo c_j , para aplicarmos o Método Simplex, precisamos do coeficiente de custo relativo r_j . Para isso, iremos zerar as entradas de c_j relacionadas as variáveis artificiais x_4 e x_5 . E assim, é dado que

$$\begin{array}{c|cccccc}
 & 2 & 1 & 2 & 1 & 0 & 4 \\
 & 3 & 3 & 1 & 0 & 1 & 3 \\
 \hline
 r_j & -5 & -4 & -3 & 0 & 0 & -7
 \end{array}$$

com o coeficiente de custo relativo explícito, podemos aplicar o Método Simplex como vimos anteriormente. Utilizando o processo de pivoteamento, temos o seguinte *Tableau*

$$\begin{array}{c|cccccc}
 & 0 & -3/4 & 1 & 3/4 & -1/2 & 3/2 \\
 & 1 & 5/4 & 0 & -1/4 & 1/2 & 1/2 \\
 \hline
 r_j & 0 & 0 & 0 & 1 & 1 & 0
 \end{array}$$

logo, a solução encontrada é dada por $x_1 = 1/2$, $x_2 = 0$ e $x_3 = 3/2$. Agora com essa solução básica factível, podemos trabalhar com o problema inicial.

Esse procedimento será chamado de Método Simplex de Duas Fases para Problemas de Programação Linear. A fase I consiste em introduzir as variáveis artificiais para encontrar uma solução básica factível. Já a fase II consiste em minimizar a função objetivo, utilizando a solução básica factível encontrada na fase I. Na fase II, as variáveis artificiais e a função artificial da fase I são omitidas. E na fase I as variáveis artificiais são adicionadas apenas nas equações que não tem variáveis de folga.

Se considerarmos o problema com as restrições conforme 3.4.4, ou seja,

$$\begin{aligned}
 \text{minimizar} \quad & 4x_1 + x_2 + x_3 \\
 \text{sujeito a} \quad & 2x_1 + x_2 + 2x_3 = 4 \\
 & 3x_1 + 3x_2 + x_3 = 3 \\
 & x_1, x_2, x_3 \geq 0
 \end{aligned}$$

vemos que não temos uma solução básica factível de imediato. Para encontrar tal solução, iremos utilizar o Método Simplex de Duas Fases. Como a fase I foi feita no exemplo anterior, vamos apenas continuar a fase II.

O *Tableau* inicial será montado com as informações da fase I e adicionaremos inicialmente os coeficientes da função objetivo na última linha. Logo, o *Tableau* será dado por

	x_1	x_2	x_3	b
	0	$-3/4$	1	$3/2$
	1	$5/4$	0	$1/2$
c_j	4	1	1	0.

Zerando os valores de custo referentes as colunas x_1 e x_3 , isto é, para que os zeros apareçam nas colunas básicas, teremos que:

	x_1	x_2	x_3	b
	0	$-3/4$	1	$3/2$
	1	$5/4$	0	$1/2$
r_j	0	$-13/4$	0	$-7/2$

Assim, temos que o coeficiente de custo relativo r_2 é negativo, isto é, $r_2 < 0$. Com isso, faremos a regra de pivoteamento nesta coluna, resultando no *Tableau* final:

	x_1	x_2	x_3	b
	$3/5$	0	1	$9/5$
	$4/5$	1	0	$2/5$
r_j	$13/5$	0	0	$-11/5$

Obtemos, dessa forma, a solução mínima da função objetivo com os valores $x_1 = 0$, $x_2 = 2/5$ e $x_3 = 9/5$.

Portanto, com esse processo desenvolvido, é possível resolver qualquer Problema de Programação Linear, tendo solução básica trivial ou não.

3.4.5 Introdução ao Algoritmo do Método Simplex Revisado

O Método Simplex Revisado é uma pequena alteração do Método Simplex clássico. Tal alteração, consiste em apenas resolvermos as contas necessárias na iteração atual. Caso m seja muito menor que n , percebemos que muitas operações que fazíamos no *Tableau* seriam quase que desnecessárias. Pensando nisso, o Método Simplex Revisado é baseado em uma interpretação matricial do Método Simplex. Sendo assim, transformaremos toda a discussão feita anteriormente com uma representação matricial.

3.4.5.1 Representação Matricial do Método Simplex

Sabemos que podemos representar um PPL da seguinte forma:

$$\begin{aligned} &\text{minimizar} && c^T x \\ &\text{sujeito a} && Ax = b \\ &&& e \quad x \geq 0 \end{aligned}$$

sendo $x, c^T \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$.

Vimos anteriormente que a matriz A tem m colunas linearmente independentes, ou seja, A tem uma submatriz $B \in \mathbb{R}^{m \times m}$ sendo uma base do espaço \mathbb{E}^m . É usual utilizarmos a matriz B como sendo as m primeiras colunas de A . Então, podemos dizer que $A = [B \ D]$, $x = (x_B, x_D)$ e $c^T = (c_B^T, c_D^T)$. Com isso, obtemos o seguinte problema:

$$\begin{aligned} &\text{minimizar} && c_B^T x_B + c_D^T x_D \\ &\text{sujeito a} && Bx_B + Dx_D = b \\ &\text{com } x_B \geq 0 && \text{ e } x_D \geq 0 \end{aligned}$$

Podemos então utilizar as equações das restrições e encontrar uma solução básica factível da seguinte maneira:

$$Bx_B + Dx_D = b \Rightarrow x_B = B^{-1}b - B^{-1}Dx_D.$$

Avaliando a função objetivo na solução básica encontrada, obtemos:

$$\begin{aligned} z &= c_B^T(B^{-1}b - B^{-1}Dx_D) + c_D^T x_D = c_B^T B^{-1}b - c_B^T B^{-1}Dx_D + c_D^T x_D \\ z &= c_B^T B^{-1}b + (c_D^T - c_B^T B^{-1}D)x_D. \end{aligned}$$

Notemos que essa função expressa o custo de qualquer solução em termos de x_D . Com isso, definimos:

$$r_D^T = c_D^T - c_B^T B^{-1}D,$$

em que r_D^T é o vetor de custo relativo para as variáveis não básicas. Da mesma forma que utilizamos o custo relativo anteriormente, este vetor servirá para decidir qual vetor deixará a base.

Se colocar as matrizes em um *Tableau* como fizemos anteriormente, teremos que:

$$\left[\begin{array}{cc|c} A & b \\ \hline c^T & 0 \end{array} \right] = \left[\begin{array}{cc|c} B & D & b \\ \hline c_B^T & c_D^T & 0 \end{array} \right]$$

Além disso, se a matriz B for utilizada como base, teremos que a forma matricial que buscamos será dada como se segue:

$$T = \left[\begin{array}{c|cc} I & B^{-1}D & B^{-1}b \\ \hline 0 & c_D^T - c_B^T B^{-1}D & -c_B^T B^{-1}b \end{array} \right]$$

Ou seja, temos que as soluções básicas factíveis são dadas da forma $x_b = B'^{-1}b$, com B' uma matriz $m \times m$ igual a matriz B a menos de operações matriciais básicas.

Analisando o pivoteamento com a representação matricial em partes, conseguimos desconsiderar os alguns cálculos para encontrar a solução ótima. Portanto, reduzimos a quantidade de operações realizadas, tornando o algoritmo menos custoso computacionalmente.

3.4.5.2 Algoritmo do Método Simplex Revisado

O algoritmo do Método Simplex Revisado precisará de duas informações básicas, sendo elas: a matriz B^{-1} a inversa de uma base atual e uma solução básica factível $x_B = b = B^{-1}b$. Vejamos o algoritmo:

Passo 0 Definir B^{-1} , D , c_B^T e c_D^T ;

Passo 1: Calcular $r_D^T = c_D^T - c_B^T B^{-1} D$. Para isso, calcule $\lambda^T = c_B^T B^{-1}$, e então calcular $r^T = c_D^T - \lambda^T D$. Se $r_D \geq 0$ então o programa será finalizado e retornará $x_b = B^{-1}b$;

Passo 2: Determinar qual o vetor A_q deve entrar na base escolhendo o menor coeficiente de custo negativo e calcule $b_q = B^{-1}A_q$;

Passo 3: Se $a_{iq} \leq 0$ então o algoritmo para, pois tem o conjunto solução K é infinito. Caso contrário calcule a razão b_i/a_{iq} com $a_{iq} > 0$;

Passo 4: Selecionar o pivô a_{jq} com j o índice da menor razão b_i/a_{iq} não negativa e pivotar em relação ao pivô a_{jq} ;

Passo 5: Atualizar B^{-1} e a solução atual é $B^{-1}b$ e Voltar ao **Passo 1**;

No **Passo 2** dividimos as contas em duas partes. Isso é feito, pois não precisaremos inverter a matriz B a cada iteração, mas teremos de resolver o sistema $B^T \lambda = c_B$, diminuindo consideravelmente o custo computacional. Veremos um exemplo de como será resolvido um PPL, utilizando os passos apresentados anteriormente. Consideremos o problema a seguir.

$$\begin{aligned} \text{minimizar} \quad & -3x_1 - x_2 - 3x_3 \\ \text{sujeito a} \quad & 2x_1 + x_2 + x_3 \leq 2 \\ & x_1 + 2x_2 + 3x_3 \leq 5 \\ & 2x_1 + 2x_2 + x_3 \leq 6 \\ & \text{e } x_1, x_2, x_3 \geq 0. \end{aligned}$$

Teremos então que, se adicionarmos as variáveis de folga, o *Tableau* dado será dado por

D	B	b
2	1 1	2
1	2 3	5
2	2 1	6

Perceba que separamos o *Tableau* de maneira que a matriz D seja formada pelas três primeiras colunas, a matriz $B = B^{-1}$ seja formada pela terceira, quarta e quinta coluna do *Tableau* e o vetor b sendo a última coluna do problema. Neste caso, temos uma solução básica inicial dada por $x_B = (0,0,0,2,5,6)$. Assim, consideraremos um novo *Tableau* utilizando apenas a matriz B^{-1} e x_B , conforme a seguir.

Variável	B^{-1}			x_B
4	1	0	0	2
5	0	1	0	5
6	0	0	1	6.

Sabemos que $c^T = [-3 \ -1 \ -3 \ 0 \ 0 \ 0]$. Com isso, $c_D^T = [-3 \ -1 \ -3]$ e $c_B^T = [0 \ 0 \ 0]$. Para o primeiro passo do algoritmo, teremos que calcular o vetor de custo relativo $r_D^T = c_D - c_B^T B^{-1} D$. Calculando $\lambda^T = c_B B^{-1}$ e $r_D^T = c_D^T - \lambda^T D$. Teremos que $\lambda^T = [0 \ 0 \ 0] B^{-1} = [0 \ 0 \ 0]$ e assim

$$r_D^T = c_D^T - \lambda^T B^{-1} D = [-3 \ -1 \ -3] - [0 \ 0 \ 0] \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 2 & 1. \end{bmatrix}$$

Obtemos então que $r_D^T = [-3 \ -1 \ -3]$. Escolhendo a segunda coluna para entrar na base, contrariando os critérios anteriores demonstrados, a fim de simplificar as operações, teremos que:

Variável	B^{-1}			x_B	$B^{-1} A_2$
4	1	0	0	2	1
5	0	1	0	5	2
6	0	0	1	6	2.

Notemos que a coluna A_2 , referente a matriz D , foi multiplicada por B^{-1} para descobrir sua representação nesta base. Neste caso, $B^{-1} = I_3$. Por isso, não modificou as entradas de A_2 . Entretanto, veremos que nas próximas iterações isso não ocorrerá.

Fazendo a divisão entre os elementos da coluna de A_2 com b , teremos que o pivô será o primeiro elemento. Com isso, realizando o pivôteamento, obtemos então

Variável	B^{-1}			x_B
2	1	0	0	2
5	-2	1	0	1
6	-2	0	1	2

calculando $\lambda^T = [-1 \ 0 \ 0] B^{-1} = [-1 \ 0 \ 0]$, teremos que $r_D^T = [-1 \ 0 \ -2]$. Portanto, escolheremos a coluna A_3 para sair da base, obtendo o seguinte *Tableau*

Variável	B^{-1}			x_B	$B^{-1} A_3$
2	1	0	0	2	1
5	-2	1	0	1	1
6	-2	0	1	2	-1.

Fazendo o pivôteamento, teremos então que

<i>Variável</i>	B^{-1}			x_B
2	1	-1	0	1
3	-2	1	0	1
6	-2	1	1	3

calculando novamente o coeficiente de custo, obteremos que $r_D^T = [-7 \ 0 \ 0]$. Então, a coluna A_1 entrará na base, conforme se segue

<i>Variável</i>	B^{-1}			x_B	$B^{-1}A_1$
2	1	-1	0	1	5
3	-2	1	0	1	-3
6	-2	1	1	3	-5.

Com o pivôteamento,

<i>Variável</i>	B^{-1}			x_B
1	3/5	-1/5	0	1/5
3	-1/5	2/5	0	8/5
6	-1	0	1	4

Assim, teremos o coeficiente de custo igual a $r_D^T = [0 \ 7/5 \ 0]$. Portanto, a solução ótima será dada por $x^T = [1/5 \ 0 \ 4 \ 0 \ 0 \ 4]$.

Perceba que a cada iteração do processo, é reduzido significativamente a quantidade de contas a serem feitas, se comparado com o método do *Tableau* feito anteriormente. Um problema a ser considerado nesse método, é que a matriz inversa de B tem que ser dada. Contudo, em alguns casos teremos que $B = I$ e a inversa é fácil de fornecer. Porém, quando precisarmos trabalhar com o Método Simplex de Duas Fases, teremos que na segunda fase do processo a matriz B^{-1} não é uma matriz trivial. Dessa forma, precisaríamos de outros métodos numéricos para criar um algoritmo mais completo.

3.4.6 Aplicação do Método Simplex

Um problema clássico resolvido no artigo Dantzig (1990) é o problema da dieta. Tal problema visa encontrar quais são os alimentos mais baratos que suprem a necessidade diária de um ser humano. Esse problema clássico foi resolvido utilizando Método Simplex. Com isso, outros trabalhos foram criados baseados nesse problema clássico, como o artigo Kripka e Peccati (2014), que, ao invés de buscar o menor custo monetário em uma dieta, buscou minimizar o consumo de calorias diárias, considerando os nutrientes básicos diários de uma pessoa.

Para a modelagem do problema, foram considerados doze micro e macro nutrientes, tais que completam a necessidade diária. Os autores buscaram diminuir o consumo de

calorias semanais. Para isso, foram modeladas 6 refeições diárias como um Problema de Programação Linear. No total, foram considerados 154 tipos de alimentos.

Aplicando o Método Simplex para a resolução dessa dieta, foi encontrada uma solução apontando a seguinte dieta, conforme a Quadro 1 abaixo.

Quadro 1 – Quadro nutricional com os alimentos em cada refeição.

Café da manhã	Lanche matinal	Almoço	Lanche da tarde	Janta	Lanche
0.7 fatia de queijo prato	0.3 fatia de mamão	4.3 banana prata	0.3 fatia de mamão	4.2 banana	0.2 fatia de mamão
1.8 fatia de presunto	0.4 fatia de queijo prato	6.4 colheres de óleo de soja	0.4 fatia de queijo prato	5.5 colheres de óleo de soja	0.2 fatia de mamão
1 colher de café solúvel	0.5 fatia de presunto	11 fatias de chuchu	0.5 fatia de presunto	8.4 fatias de chuchu	0.7 laranja
0.1 fatia de pão de centeio	0.6 fatia de abacaxi	1.1 sobrecoxa de frango	0.6 fatia de abacaxi	1 sobrecoxa de frango	0.2 fatia de abacaxi
0.7 colher de margarina	0.4 fatia de melão	2.2 colheres de couve	0.4 fatia de melão	2 colheres de couve	0.4 fatia de melão
1.8 colher de geleia de frutas		3.5 colheres de abóbora		2.8 colheres de abóbora	2.9 biscoitos salgados
3.3 fatias de abacaxi		4.4 colheres de pepino		3.2 colheres de pepino	0.2 iogurte
0.2 fatia de melão					

Fonte: Kripka e Peccati (2014).

Assim, o total resultante de calorias utilizando essa dieta foi de 12.239,28 por semana. Já uma dieta buscando os mesmos nutrientes feito por uma nutricionista, resultaria em 17.561,62 calorias semanais. Ou seja, com o modelo minimizado pelo Método Simplex, teve uma redução de quase 30%.

Por fim, o Método Simplex é muito utilizado em diversas empresas para diferentes objetivos, como minimizar despesas ou maximizar lucros. A importância desse exemplo é dado pela sua aplicação no cotidiano de uma pessoa. Mostrando a importância, e a diferença, na utilização da matemática em diferentes contextos.

4 PROBLEMAS IRRESTRITOS DE PROGRAMAÇÃO NÃO LINEAR

Os problemas estudados anteriormente buscavam encontrar o mínimo de uma função linear em um conjunto determinado por equações e inequações lineares. Essas características determinavam o chamado Problema de Programação Linear. Veremos agora quais são as características dos Problemas Irrestritos de Programação Não Linear. De uma forma geral, o Problema de Programação Não Linear pode ser definido por

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & x \in \mathbb{R}^n \end{array}$$

em que $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Como $x \in \mathbb{R}^n$, temos que não há restrições no conjunto estudado, diferente do estudo abordado anteriormente. Esta é a característica dos Problemas Irrestritos de Programação Não Linear.

Nesses problemas, temos que podem existir dois tipos de mínimo da função objetivo. Para melhor compreender a diferença entre esses mínimos, vejamos as próximas definições.

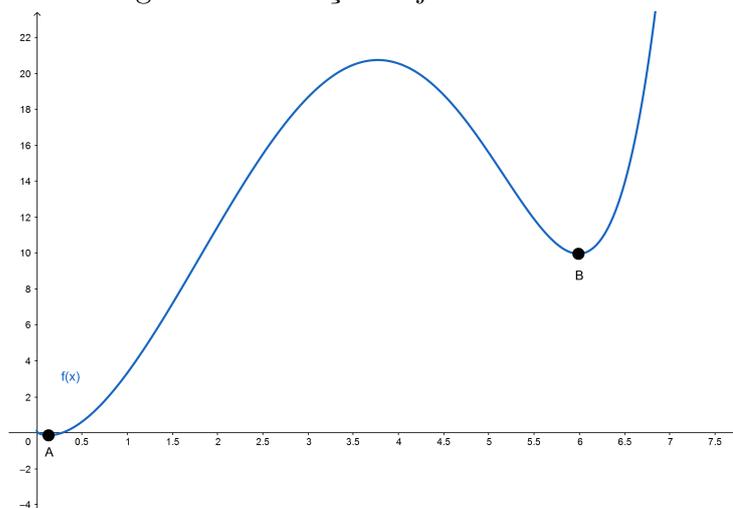
Definição 38. *Mínimo Global* Dizemos que $a \in D(f)$ é um minimizador global quando, para todo $x \in D(f)$ temos $f(x) \leq f(a)$. E o valor obtido por $f(a)$ é chamado de mínimo global.

Agora definiremos o conceito de mínimo local, o qual é mais restrito do que a definição feita anteriormente.

Definição 39. *Mínimo Local* Dizemos que $a \in D(f)$ é um minimizador local quando, existe $\varepsilon > 0$ tal que para todo x tal que $\|a - x\| < \varepsilon$ temos $f(x) \geq f(a)$. E o valor obtido por $f(a)$ é chamado de mínimo local.

O minimizador local, minimiza a função objetivo em uma determinada região. Já o minimizador global, minimiza a função objetivo em todo seu domínio. Vejamos o problema de minimizar $f(x) = -x^3 + 5x^2 - \sqrt{x} + 0.12e^x$. Não é trivial encontrarmos os pontos de mínimo local e mínimo global. Mesmo utilizando a função derivada, a minimização não poderá ser explorada analiticamente. Podemos ilustrar o gráfico da função na Figura 15.

Figura 15 – Função objetivo não linear



Fonte: Autoria própria.

Neste caso, temos um mínimo local e um mínimo global. Os pontos $A = (a, f(a))$ e $B = (b, f(b))$ relacionam o minimizador ao seu mínimo. O ponto $x = a$ é um minimizador global e o ponto $x = b$ é um minimizador local.

Utilizaremos o método de Nelder e Mead (1965), o qual também é baseado na Geometria Simplética, para encontrar uma solução mínima de forma iterativa. O Método Nelder-Mead é um método livre de derivadas, sendo capaz de encontrar mínimos em funções que não possuem derivadas contínuas. Ou ainda, é utilizado quando o cálculo dessas derivadas exige um alto custo computacional.

Segundo (CONN; SCHEINBERG; VICENTE, 2009), a otimização é, muitas vezes, relacionada com a derivada da função igualando a zero. Porém, existem várias funções sem derivadas, ou com derivadas complicadas, que também podem ser otimizadas. Portanto, existe uma grande quantidade de problemas que podem ser resolvidos utilizando métodos livres de derivação.

4.1 Método de Nelder-Mead

O algoritmo de Nelder-Mead é um dos mais populares métodos livre de derivadas. Sendo muito utilizado na engenharia e, provavelmente, o mais citado dos métodos. Dentre as razões para este sucesso, destacam-se a simplicidade e a facilidade de adaptação para minimizar ou maximizar funções. Para este estudo, será utilizado o livro *Introduction to derivative-free optimization* de Conn, Scheinberg e Vicente (2009).

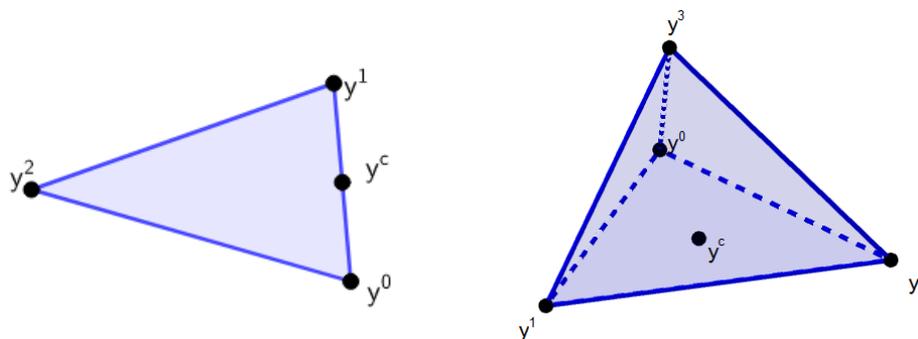
Este algoritmo é um método de pesquisa direta, o qual utiliza apenas o valor da função objetivo em uma quantidade finita de pontos por iteração, para assim, decidir qual a próxima ação tomar.

Todo o processo será baseado no espaço \mathbb{R}^n . Como definimos anteriormente,

teremos um simplex de dimensão n neste espaço. Isto é, teremos $n + 1$ vértices que iremos denotar por $Y = \{y^0, y^1, \dots, y^n\}$. Tais vértices serão ordenados em um vetor auxiliar, com valores da função objetivo de forma crescente. Além dos vértices do simplex, teremos que definir um novo ponto que chamaremos de centróide.

Definição 40. O ponto $y^c = \sum_{i=0}^{n-1} \frac{y^i}{n}$ é chamado centróide do conjunto $Y = \{y^0, y^1, \dots, y^{n-1}\}$.

Figura 16 – Exemplos de centróides em \mathbb{R}^2 e \mathbb{R}^3 .



Fonte: Autoria própria.

Após determinado o centróide, o método irá determinar um novo ponto que estará na reta que contém os pontos y^n e y^c . Para isso, o método seleciona um $\delta \in \mathbb{R}$, de maneira que o novo ponto será da forma:

$$y = y^c + \delta(y^c - y^n), \tag{16}$$

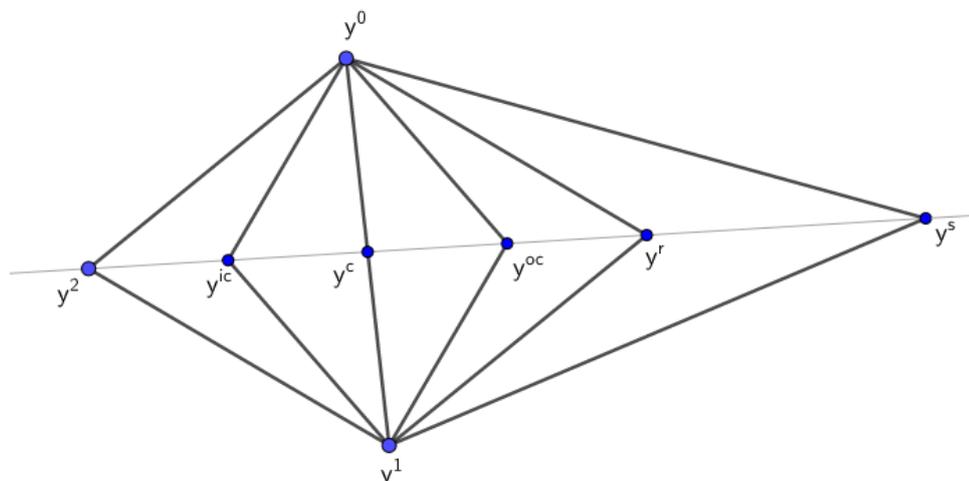
utilizando os procedimentos de reflexão, expansão ou contração para substituir o pior vértice (y^n).

O valor de δ é importante para definir o próximo simplex. Ou seja, o próximo conjunto Y_1 . Para cada valor diferente de δ , teremos um novo ponto. Por exemplo, se tomarmos os seguintes valores:

- $\delta = 1$ teremos uma reflexão;
- $\delta = 2$ teremos uma expansão;
- $\delta = \frac{1}{2}$ teremos uma contração externa;
- $\delta = -\frac{1}{2}$ teremos uma contração interna.

Aplicando o método, teremos a seguinte relação dos novos vértices em \mathbb{R}^2 :

Figura 17 – Exemplos do processo de nós no \mathbb{R}^2 .

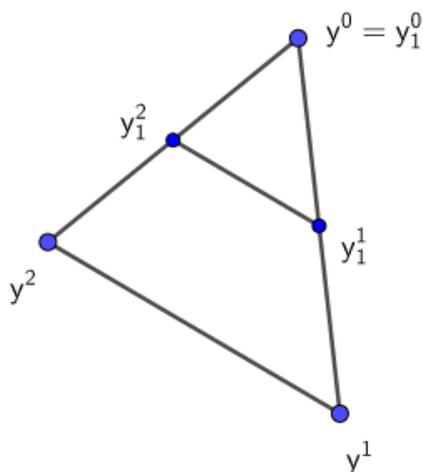


Fonte: Autoria própria.

Na Figura 17, os pontos y^0 , y^1 e y^2 são dados e o ponto y^2 é o pior ponto (aqui o pior ponto corresponde àquele em que a função objetivo tem o maior valor). Obtemos o centróide y^c como o ponto equidistante de y^0 e y^1 . Baseado na reta que passa por y^2 e y^c , criaremos novos pontos y^{ic} , y^{oc} , y^r e y^s que são chamados de pontos de contração interna, contração externa, reflexão e expansão, respectivamente. Esses novos pontos substituirão o ponto y^2 em caso de sucesso do método.

Além desses casos, existe um outro tipo de nó, nomeado de encolhimento. No entanto, neste nó, nós substituiremos todos os vértices do simplex com exceção do y_0 (o vértice relacionado ao menor valor da função).

Figura 18 – Exemplo de encolhimento em \mathbb{R}^2 .



Fonte: Autoria própria.

Então, para todo $i = 1, 2, \dots, n$, teremos que y_i deve ser substituído por $y^0 + \frac{1}{2}(y^i - y^0)$. A Figura 18 mostra como seria feita essa substituição se o encolhimento for aceito em \mathbb{R}^2 . Sendo assim, y^0 , y^1 e y^2 são os pontos iniciais do simplex e y_1^0 , y_1^1 e y_1^2 os vértices obtidos no encolhimento. Chamaremos de nós esses pontos gerados pelos diferentes valores de δ , entre os quais iremos escolher aquele que irá prosseguir para a próxima iteração no lugar do pior ponto do simplex. A fim de facilitar a aplicação do método, na prática, trabalhamos com os valores de δ definidos inicialmente. Sendo assim, podemos utilizar diferentes valores de δ respeitando as seguintes desigualdades:

$$0 < \delta^s < 1, \quad -1 < \delta^{ic} < 0 < \delta^{oc} < \delta^r < \delta^e,$$

em que δ^s , δ^{ic} , δ^{oc} , δ^r e δ^e estão relacionados às iterações de encolhimento, contração interna, contração externa, reflexão e expansão, respectivamente.

Introduziremos o algoritmo do Método de Nelder-Mead, utilizando uma aplicação dos nós, assim como mostrado anteriormente.

4.1.1 Minimizando Funções com o Algoritmo do Método de Nelder-Mead

Para buscarmos soluções para Problemas Irrestritos de Programação Linear, estudaremos o Método de Nelder-Mead. O método abordado neste trabalho tem uma pequena diferença do método original, pois no método clássico se a expansão for aceita então obteríamos $f^e < f^0$, ignorando a possibilidade de $f^r < f^e$. Ou seja, o Método Moderno de Nelder-Mead prevê a possibilidade de melhorar o ponto escolhido caso tenha sucesso na etapa de reflexão e, daí, escolhe a melhor opção entre f^r e f^e , otimizando o processo.

Passo 0: Determine o simplex inicial $Y_0 = \{y_0^0, y_0^1, \dots, y_0^n\}$. Defina $k = 0$ e as constantes

$$\delta^s = \frac{1}{2}, \delta^{ic} = -\frac{1}{2}, \delta^{oc} = \frac{1}{2}, \delta^r = 1 \text{ e } \delta^e = 2.$$

Passo 1: Para $k = 1, 2, \dots$. Defina o conjunto Y_k e avalie f nos pontos de Y_k ;

Passo 2: Reordene o conjunto Y_k de maneira que

$$f^0 = f(y^0) \leq f^1 = f(y^1) \leq \dots \leq f^n = f(y^n).$$

Passo 3: Defina o centróide $y^c = \sum_{i=0}^{n-1} \frac{y^i}{n}$.

Passo 4: (Reflexão) Faça uma reflexão do pior vértice (y^n) sobre o centróide:

$$y^r = y^c + \delta^r(y^c - y^n)$$

e avalie $f^r = f(y^r)$. Se $f^0 \leq f^r \leq f^{n-1}$ defina $Y_{k+1} = \{y_k^0, y_k^1, \dots, y^r, \dots, y_k^{n-1}, y^r\}$. Vá para o **Passo 2**.

Passo 5: (Expansão) Se $f^r < f^0$ calcule o ponto de expansão

$$y^e = y^c + \delta^e(y^c - y^n)$$

e avalie $f^e = f(y^e)$. Se $f^e \leq f^r$ defina $Y_{k+1} = \{y_k^0, y_k^1, \dots, y_k^{n-1}, y^e\}$, caso contrário, defina $Y_{k+1} = \{y_k^0, y_k^1, \dots, y_k^{n-1}, y^r\}$. Vá para o **Passo 2**.

Passo 6: (Contração) Se $f^r \geq f^{n-1}$ é feito uma contração entre y^r (contração interna) e y^n (contração externa).

Passo 6.1: (Contração externa) Se $f^r < f^n$, execute uma contração externa com

$$y^{oc} = y^c + \delta^{oc}(y^c - y^n)$$

e avalie $f^{oc} = f(y^{oc})$. Se $f^{oc} \leq f^r$, defina $Y_{k+1} = \{y_k^0, y_k^1, \dots, y_k^{n-1}, y^{oc}\}$ e vá para o **Passo 2**. Caso contrário, vá para o **Passo 7**;

Passo 6.2: (Contração interna) Se $f^r \geq f^n$, execute uma contração interna

$$y^{ic} = y^c + \delta^{ic}(y^c - y^n)$$

e avalie $f^{ic} = f(y^{ic})$. Se $f^{ic} < f^n$, defina $Y_{k+1} = \{y_k^0, y_k^1, \dots, y_k^{n-1}, y^{ic}\}$ e vá para o **Passo 2**. Caso contrário, vá para o **Passo 7**;

Passo 7: (Encolhimento) Defina os pontos $y_s^i = y^0 + \delta^s(y^i - y^0)$, com $i = 1, 2, \dots, n$, e o conjunto $Y_{k+1} = \{y^0, y_s^1, y_s^2, \dots, y_s^n\}$. Volte para o **Passo 2**.

Um critério de parada pode consistir no “tamanho” do simplex. Neste sentido, vamos definir o diâmetro de um simplex da seguinte maneira:

Definição 41. O diâmetro do simplex $Y = \{y^0, \dots, y^n\}$ é dado por

$$diam(Y) = \max_{0 \leq i < j \leq n} \|y^i - y^j\|.$$

Por exemplo, se tivermos um simplex definido por $Y = \{e_1, \dots, e_n\}$, sendo e_i os vetores da base canônica com $i = 1, \dots, n$, teremos que $diam(Y) = \sqrt{2}$.

Para reduzir o custo computacional, podemos aproximar o diâmetro pelo valor

$$\Delta(Y) = \max_{0 \leq i \leq n} \|y^i - y^0\|.$$

Note que teremos $\Delta(Y) \leq diam(Y) \leq 2\Delta(Y)$. Com isto, podemos definir como critério de parada $\Delta(Y_k) < \Delta_{tol}$, em que Δ_{tol} seria um valor suficientemente pequeno para garantir a precisão desejada para a solução.

Tal algoritmo tem um número exato de nós por iteração, sendo 1 se o nó é de reflexão, 2 se o nó é de expansão ou contração e $n + 1$ se o nó é de encolhimento, em que n é a dimensão do simplex.

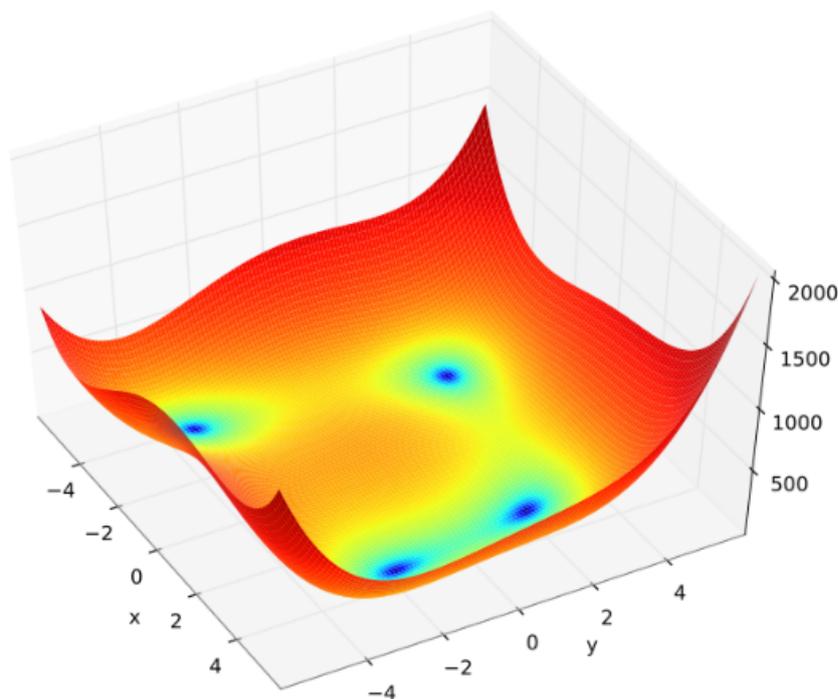
Um exemplo da aplicação do Método de Nelder-Mead seria na busca do ponto de mínimo da Função de Himmelblau. Utilizaremos a tese da autora Gonçalves (2013) para observar geometricamente como ocorre a convergência do método.

A Função de Himmelblau é dada por

$$f(x,y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2,$$

com o seguinte gráfico da função

Figura 19 – Função de Himmeblau.



Fonte: Gonçalves (2013).

A função tem quatro minimizadores globais distintos, sendo eles:

$$f(3.0,2.0) = 0.0$$

$$f(-2.805118,3.131312) = 0.0$$

$$f(-3.779310, -3.283186) = 0.0$$

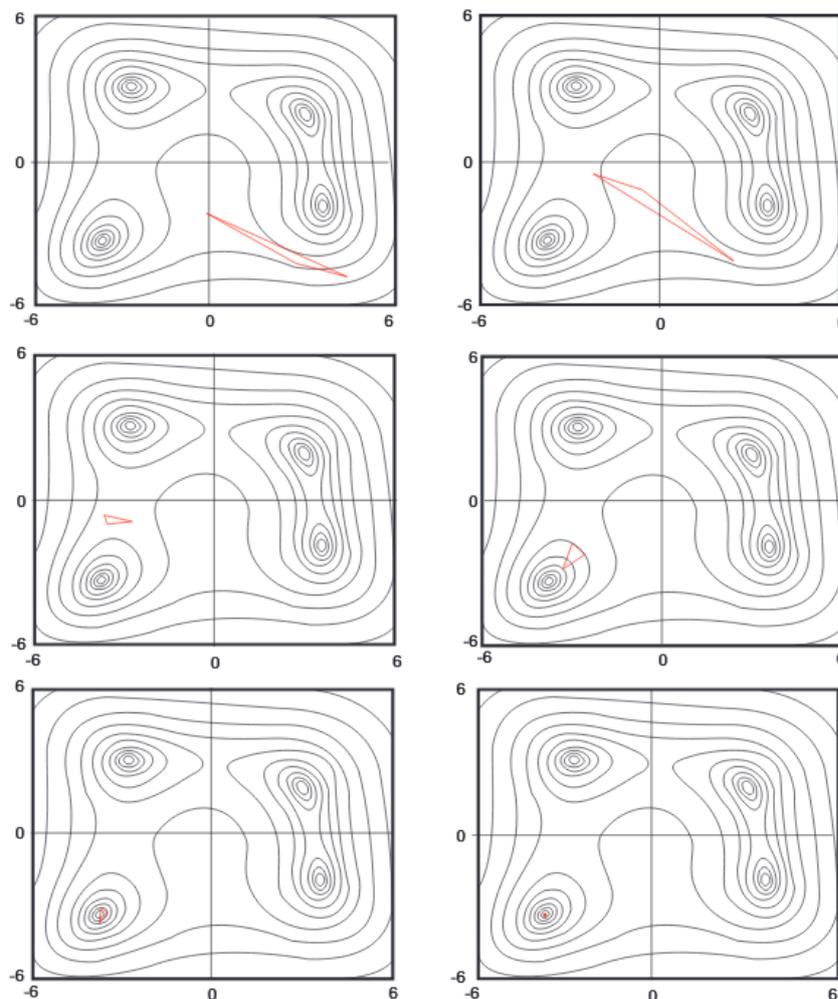
$$f(3.584428, -1.848126) = 0.0.$$

Aplicando o Método de Nelder-Mead, podemos observar geometricamente como ocorre a convergência¹ na Figura 20 do método.

Um ponto importante é que o minimizador que o método encontro depende, em grande parte dos casos, diretamente do simplex definido inicialmente. Além disso, esse exemplo é muito importante pois mostra que, mesmo utilizando um simplex próximo de um minimizador, o Método de Nelder-Mead não necessariamente converge para o minimizador mais próximo.

¹Para melhorar a compreensão, a visualização completa do método pode ser acessada em: <https://upload.wikimedia.org/wikipedia/commons/9/96/Nelder_Mead2.gif>

Figura 20 – Convergência do Método de Nelder-Mead na função de Himmelblau.



Fonte: Gonçalves (2013).

Esse fato ocorre em diversas funções, pois o método busca encontrar um minimizador global ou local. Mas não conseguimos escolher para qual minimizador o método convergirá.

4.1.2 Nelder-Mead e Simplex

O algoritmo de Nelder-Mead modifica o simplex para que se adapte a curva no local. Isto é, o simplex se adapta a paisagem local (CONN; SCHEINBERG; VICENTE, 2009). Porém o desempenho do algoritmo de Nelder-Mead, quando funciona, está diretamente relacionado com o ajuste da curvatura da função.

Um problema decorrente do método, é que o simplex pode não estabelecer uma convergência global, pois é necessário que os vértices não sejam pontos estacionários². Para resolver esse problema, é preciso reiniciar o método sempre que a geometria ou a

²Pontos estacionários de uma função são os pontos que avaliados na derivada dessa função resultam em zero.

estabilidade dos vértices do simplex derem problema. Como estamos trabalhando com métodos livres de derivadas, esse problema é encontrado utilizando Λ -posicionamento na maioria dos casos. Porém, em conjuntos simplex, é mais comum a utilização do cálculo de volume do simplex. Para essa discussão, utilizaremos o trabalho complementar de Gonçalves (2013).

Sabemos que no espaço \mathbb{R}^n , um simplex não pode ter dimensão maior que n . Neste caso, suponhamos que o conjunto $Y = \{y^0, \dots, y^n\}$ forme um simplex de dimensão n . Então temos a seguinte definição.

Definição 42. *A matriz referente as faces do simplex será dada por*

$$L(Y) = [y^1 - y^0 \ \dots \ y^n - y^0]^T = \begin{bmatrix} y_1^1 - y_1^0 & \dots & y_n^1 - y_n^0 \\ \vdots & \ddots & \vdots \\ y_1^n - y_1^0 & \dots & y_n^n - y_n^0 \end{bmatrix} \quad (17)$$

com $y^j = [y_1^j \ \dots \ y_n^j]^T$.

Uma observação interessante, é que se tomarmos a matriz $[L(Y) \ (-L(Y))]$ teremos que essa matriz será a máxima base positiva³ de \mathbb{R}^n .

Agora iremos generalizar o conceito de volume, que geralmente é empregado a conjuntos tridimensionais.

Definição 43. *Seja o conjunto $Y = \{y^0, \dots, y^n\}$. O volume de um simplex relacionado a Y será dado por*

$$\text{vol}(Y) = \frac{|\det(L(Y))|}{n!},$$

com a matriz $L(Y)$ da Definição 17.

Perceba que, da Geometria Analítica, temos que a área de um triângulo é dado por $\frac{|\det(M)|}{2} = \frac{|\det(M)|}{2!}$, sendo M a matriz relacionada com os lados do triângulo. O volume do tetraedro relacionando a uma matriz M é dado por $\frac{|\det(M)|}{6} = \frac{|\det(M)|}{3!}$. Ou seja, coincide com a generalização feita.

Em métodos simpléticos, como o de Nelder-Mead, é utilizado o diâmetro do conjunto, em que neste caso é o diâmetro do simplex. Essa escolha é feita pois não há a necessidade de definir o ponto⁴ y^c para tal análise.

Agora introduziremos o volume normalizado. Essa definição substituirá o Λ -posicionamento na verificação do simplex.

Definição 44. *Seja $Y = \{y^0, \dots, y^n\}$ o conjunto de $n + 1$ pontos. O volume normalizado de Y será dado por*

³ M é base positiva de \mathbb{R}^n se para todo $x \in \mathbb{R}^n$ temos que $\alpha_1 m_1 + \dots + \alpha_p m_p = x$, com $\alpha_i \geq 0$ para $i = 1, \dots, p$.

⁴No livro de Andrew R. Conn (2009), o ponto y^c é chamado de *centering point*.

$$von(Y) = vol\left(\frac{1}{diam(Y)}Y\right) = \frac{|det(L(Y))|}{n! diam(Y)^n}.$$

Observe que para chegar a última igualdade na Definição 44 foi utilizada, além da definição de volume, a seguinte propriedade de determinante $det(\alpha A) = \alpha^n det(A)$, se $A \in \mathbb{R}^{n \times n}$.

As definições de diâmetro e volume serão matematicamente convenientes para a manipulação de um simplex, utilizando reflexão ou encolhimento. Com tais definições, veremos alguns fatos importantes do simplex utilizando Nelder-Mead.

Para sabermos se o simplex escolhido poderá ser utilizado, precisaremos saber se o simplex é não degenerado. Para isso, faremos a seguinte definição.

Definição 45. *Seja o conjunto $Y = \{y^0, \dots, y^n\}$, um simplex será não degenerado quando $det(L(Y)) \neq 0$. Ou seja, $vol(Y) > 0$.*

Um simplex degenerado, seguindo a definição, tem a matriz $L(Y)$ com as colunas linearmente dependentes. Logo, o volume do simplex será igual a zero. Para obter êxito no Método de Nelder-Mead, é necessário que o simplex seja não degenerado. Dessa forma, nesse trabalho, consideraremos que o simplex inicial sempre será não degenerado.

Sabemos que o Método de Nelder-Mead trabalha, basicamente, modificando o conjunto inicial $Y = \{y^0, \dots, y^n\}$. Portanto, utilizaremos a notação Y_k para descrever a k – ésimas iteração do método sobre o conjunto Y .

Teorema 46. *Caso a iteração k seja de reflexão, expansão ou contração, então temos que*

$$vol(Y_{k+1}) = |\delta| vol(Y_k). \quad (18)$$

Se a iteração for de encolhimento, então

$$vol(Y_{k+1}) = \delta^n vol(Y_k). \quad (19)$$

sendo δ referente a cada tipo de nó.

Demonstração. Sem perda de generalidade, suponhamos que o pior vértice y^n é tal que $y^n = 0$. Isto é, o pior vértice da k – ésimas iteração está na origem. Sabemos que o novo vértice, se não houver encolhimento, é dado da seguinte forma:

$$v_k = y_k^c + \delta(y_k^c - y_k^n) = (1 + \delta)y_k^c, \quad (20)$$

com

$\delta = \delta^r$ Se o nó de reflexão for aceito

$\delta = \delta^e$ Se o nó de expansão for aceito

$\delta = \delta^{oc}$ Se o nó de contração externa for aceito

$\delta = \delta^{ic}$ Se o nó de contração interna for aceito

Logo, pela Equação (20), o novo vértice também pode ser escrito da seguinte forma:

$$v_k = L_k w \quad \text{com} \quad w = \left[\frac{1 + \delta}{n}, \dots, \frac{1 + \delta}{n} \right].$$

Sendo que L_k é a matriz relacionada ao conjunto Y_k dos vértices do simplex. Como o volume de um simplex não depende da ordem de seus vértices, também podemos supor que v_k seja o pior vértice de Y_k . Então, teremos que

$$|\det(L_{k+1})| = |\det(L_k - L_k w e^T)| = |\det(L_k)| |\det(I - w e^T)|,$$

o qual $e^T = (1, \dots, 1)$. Temos então que a matriz $I - w e^T$ tem $n - 1$ autovalores iguais a 1 e, pela construção de w , temos um auto valor igual a $-\delta$. Como o determinante de uma matriz é o produto de seus autovalores, então $\det(I - w e^T) = -\delta$. Sendo assim,

$$|\det(L_{k+1})| = |\delta| |\det(L_k)|.$$

Portanto,

$$\text{vol}(Y_{k+1}) = |\delta| \text{vol}(Y_k).$$

Agora faremos a demonstração da Equação (19). Seja Y_k o conjunto referente ao simplex da k - ésima iteração. Se a iteração $k + 1$ for de encolhimento, sabemos que $Y_{k+1} = \{y_k^i = y^0 + \delta^s(y^i - y^0), i = 1, \dots, n\}$. Então, por definição, teremos que,

$$\begin{aligned} L(Y_{k+1}) &= [y_k^0 + \delta^s(y_k^1 - y_k^0) - y_k^0 \quad \dots \quad y_k^0 + \delta^s(y_k^1 - y_k^0) - y_k^0] \\ &= [\delta^s(y_k^1 - y_k^0) \quad \dots \quad \delta^s(y_k^1 - y_k^0)] \\ &= \delta^s[(y_k^1 - y_k^0) \quad \dots \quad (y_k^1 - y_k^0)] \\ L(Y_{k+1}) &= \delta^s L(Y_k) \end{aligned}$$

Portanto, podemos calcular o volume do simplex

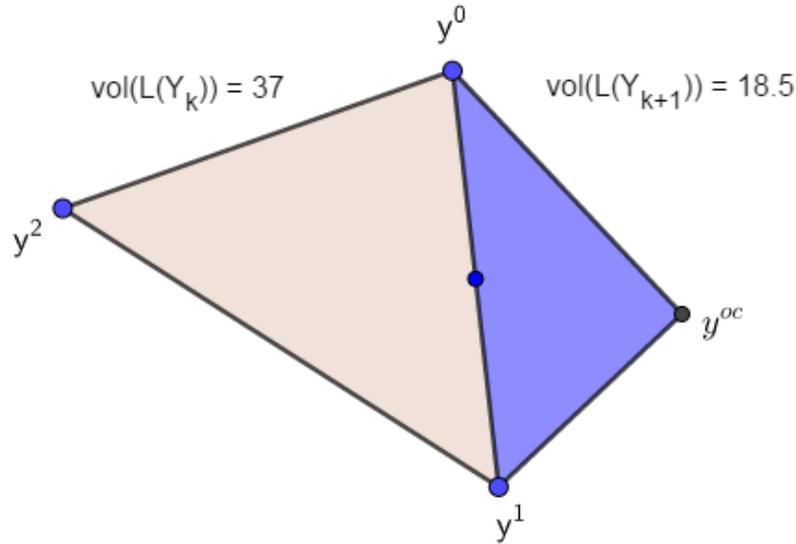
$$\text{vol}(Y_{k+1}) = \frac{|\det(L(Y_{k+1}))|}{n!} = \frac{|\det(\delta L(Y_k))|}{n!} = \frac{(\delta^s)^n |\det(L(Y_k))|}{n!}$$

$$\Leftrightarrow \text{vol}(Y_{k+1}) = (\delta^s)^n \text{vol}(Y_k).$$

sendo δ referente a iteração de encolhimento. □

A Figura 21 ilustra a contração externa no \mathbb{R}^2 .

Figura 21 – Comparação do volume de um simplex usando contração externa no \mathbb{R}^2 .



Fonte: Autoria própria.

Assim, temos uma relação entre os volumes do simplex gerados pelo Método de Nelder-Mead. Dessa maneira, poderemos analisar algumas propriedades do método.

Teorema 47. *Se o simplex inicial é não degenerado, então todos os outros simplex gerados pelo Método de Nelder-Mead também serão não degenerados.*

Demonstração. Sabemos que um simplex é não degenerado se $\text{vol}(L(Y)) \neq 0$ e também sabemos que $\delta \neq 0$ para todos os tipos de nós. Portanto, como $\text{vol}(Y_{k+1}) = |\delta| \text{vol}(Y_k)$ se o nó não for de encolhimento e $\text{vol}(Y_{k+1}) = \delta^n \text{vol}(Y_k)$ se for de encolhimento, teremos que $\text{vol}(L(Y_k)) \neq 0$, para toda toda iteração k do Método de Nelder-Mead. \square

Definiremos $f(y_k^0)$ como sendo f avaliada no melhor ponto da k – ésima iteração do Método de Nelder-Mead. Para analisarmos a convergência do método, definiremos $\{f(y_k^0)\}_{k \in \mathbb{N}}$ como uma sequência da função objetivo avaliada no melhor ponto de $Y_k = \{y_k^0, \dots, y_k^n\}$. Para provar a convergência da sequência $\{f(y_k^0)\}_{k \in \mathbb{N}}$, precisamos antes mostrar um lema.

Lema 48. *Se não há encolhimento então $f(y_{k+1}^i) \leq f(y_k^i)$, para todo $k \in \mathbb{N}$, se $i \in \{0, \dots, n\}$. Além disso, a desigualdade é estrita para algum i .*

Demonstração. Se a iteração $k + 1$ não é de encolhimento, então temos que o único vértice modificado do simplex anterior será y_k^n . Pelo algoritmo é dado que

$$\begin{aligned} f(y_{k+1}^i) &= f(y_k^i), \quad \text{para } i \neq n \\ f(y_{k+1}^i) &< f(y_k^i), \quad \text{para } i = n, \end{aligned}$$

como queríamos demonstrar. \square

Com isso, é dado que $\sum_{j=0}^n f(y_{k+1}^j) < \sum_{j=0}^n f(y_k^j)$. Essa é uma propriedade explorada por Kelley (1999) no estudo de problemas encontrados no Método de Nelder-Mead.

Teorema 49. *Seja f uma função limitada inferiormente em \mathbb{R}^n , e seja o Método de Nelder-Mead. Então*

$$\{f(y_k^0)\}_{k \in \mathbb{N}}$$

é convergente.

Demonstração. Se não há encolhimento nas iterações do método, sabemos que $f(y_{k+1}^0) \leq f(y_k^0)$. Caso houvesse iteração de encolhimento, sabemos que o vértice y_k^0 é o único mantido no novo simplex e portanto teríamos que $f(y_{k+1}^0) = f(y_k^0)$. Ou seja, a sequência é monótona decrescente, e como f é limitada inferiormente, teremos que a sequência é monótona decrescente e limitada. Portanto, convergente. \square

Dessa forma, sabemos agora que dado um simplex não degenerado, teremos que

$$\{f(y_k^0)\}_{k \in \mathbb{N}}$$

é convergente se $k \rightarrow \infty$. Entretanto, perceba que essa sequência não demonstra que converge para o ponto de mínimo da função.

Teorema 50. *Seja f uma função limitada inferiormente. Se o número de encolhimento é finita então*

$$\{f(y_k^0)\}_{k \in \mathbb{N}}$$

converge, para todo $i \in \{0, \dots, n\}$. Além disso, $f^(y^0) \leq \dots \leq f^*(y^n)$, com $f^*(y^i) = \lim_{k \rightarrow \infty} f(y_k^i)$.*

Demonstração. Os vértices do simplex são ordenados, a cada iteração, dessa maneira $f(y^0) \leq \dots \leq f(y^n)$, como f é limitada inferiormente temos que as sequências

$$\{f(y_k^0)\}_{k \in \mathbb{N}}$$

são limitadas, monotonas e decrescentes. Portanto são convergentes e o limite respeita $f^*(y^0) \leq \dots \leq f^*(y^n)$. \square

Posteriormente, trabalharemos com a ordem dos elementos obtidos pelo Método de Nelder-Mead e com o teorema anterior poderemos demonstrar novos resultados relacionados com a ordem dos elementos.

Agora analisaremos o caso que não ocorre o encolhimento em pelo menos um número finito de iterações, uma vez que já sabemos que o nó de encolhimento modifica $n - 1$ vértices, aumentando a quantidade de contagens significativamente em que a convergência das sequências também é afetada.

Teorema 51. *Dado uma função limitada inferiormente f . Se há um número finito de iterações sem encolhimento, então todos os vértices do simplex convergem para um único ponto.*

Demonstração. Sabemos que se há um número finito de iterações sem encolhimento, então se $k > K$ é fato que a k – ésima iteração é de encolhimento para algum $K \in \mathbb{N}$. Logo, temos que

$$\text{vol}(Y_{k+1}) = \delta^n \text{vol}(Y_k).$$

Como $0 < \delta < 1$, então

$$\text{vol}(Y_k) \rightarrow 0 \text{ quando } k \rightarrow \infty.$$

Logo,

$$|\det(L(Y_k))| \rightarrow 0 \text{ se } k \rightarrow \infty.$$

Ou seja, o ângulo entre vetores referentes as colunas de L_k tenderão a zero quando k tender ao infinito. Portanto, todos os vértices do simplex irão convergir para um único ponto.

□

Perceba que, novamente, tal demonstração não implica que a convergência seja em um ponto estacionário.

Podemos mostrar outros resultados se considerarmos apenas as funções estritamente convexas. Tais funções são uma extensão da definição de conjuntos convexos trabalhados anteriormente.

Definição 52. *Uma função f é dita estritamente convexa em \mathbb{R}^n quando, para todo $x, y \in \mathbb{R}^n$ com $x \neq y$ e $0 < \lambda < 1$ temos que*

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y).$$

Teorema 53. *Se f é estritamente convexa, então para qualquer simplex não degenerado o Método de Nelder-Mead não terá iteração de encolhimento.*

Demonstração. Se a função é estritamente convexa e $c = \sum_{i=0}^m \lambda_i y_i$ então

$$f(c) = f\left(\sum_{i=0}^n \lambda_i y_i\right) < \sum_{i=0}^{n-1} \lambda_i f(y^i) \leq \max\{f(y^1), \dots, f(y^n)\}$$

se $\lambda \in (0,1)$. A iteração de encolhimento ocorre apenas se a contração interna ou externa falham. Logo, para esse demonstração, vamos focar na contração externa. Para a contração

externa ser aceita temos que $f(y_k^{n-1}) \leq f(y_k^r) < f(y_k^n)$. Portanto, para algum $\lambda \in (0,1)$, teríamos que

$$f(y_k^{oc}) = f(\lambda y_k^c + (1 - \lambda)y_k^r) < \lambda f(y_k^c) + (1 - \lambda)f(y_k^r) \leq \max\{f(y_k^c), f(y_k^r)\}$$

sendo $y_k^c = \sum_{i=0}^{n-1} \frac{y_k^i}{n}$ o centroide e $y_k^r = y_k^c + \delta^r(y_k^c - y_k^c)$ o ponto do nó de reflexão do método. Sabemos que $f(y_k^n) \leq f(y_k^r)$ e também $f(y_k^c) \leq f(y_k^{n-1})$, pois

$$f(y_k^c) = f\left(\sum_{i=0}^{n-1} \frac{y_k^i}{n}\right) < \sum_{i=0}^{n-1} \frac{1}{n} f(y_k^i) \leq \max\{f(y_k^1, \dots, y_k^{n-1})\}.$$

Assim, $\max\{f(y_k^c), f(y_k^r)\} = f(y_k^r)$. Dessa maneira, $f(y_k^{oc}) < f(y_k^r)$ e logo a contração externa é aceita. A demonstração para a contração interna é similar. Portanto, sempre que a iteração entrar em um nó do tipo de contração, se a função for extritamente convexa, então a iteração acabará neste nó. □

4.1.3 Ordem Lexicográfica Decrescente

Agora abordaremos a maneira de corrigir um problema na organização dos elementos no início de uma iteração do Método de Nelder-Mead. Se uma iteração anterior acabar com reflexão, expansão ou contração e o valor avaliado de tal vértice for igual a um ou mais valores avaliados nos vértices anteriores, precisamos então organizar o conjunto dos vértices de uma única maneira. Para isso, teremos que definir uma regra para que esse novo vértice seja posicionado.

Definição 54. *O índice k^* é chamado de índice de mudança da k – ésima iteração sendo*

$$k^* = \min\{i \in (0, \dots, n) \mid y_k^i \neq y_{k+1}^i\}.$$

Por exemplo, se tivermos o conjunto $(f_k^0, f_k^1, f_k^2, f_k^3, f_k^4) = (1, 2, 2, 3, 3)$ e a iteração $k+1$ nos fornecer o valor 2, temos que o novo conjunto será dado por $(f_{k+1}^0, f_{k+1}^1, f_{k+1}^2, f_{k+1}^3, f_{k+1}^4) = (1, 2, 2, 2, 3)$. Logo, nesse exemplo, $k^* = 3$, pois $f_k^3 \neq f_{k+1}^3$.

Em vista disso, temos que $k^* \in \{1, 2, \dots, n-1\}$ se a iteração for de reflexão, $k^* = 0$ se a iteração for de expansão, $k^* \in \{0, 2, \dots, n\}$ se a iteração for de contração e $k^* = 0$ ou $k^* = 1$ se a iteração for de encolhimento. Isso é dado observando as possibilidades do fim de iteração do algoritmo trabalhado anteriormente.

Teorema 55. *Seja f uma função limitada inferiormente em \mathbb{R}^n . Se o Método de Nelder-Mead parte de um simplex inicial não degenerado, se não há encolhimento, e se existe um inteiro $j \in \{0, 1, \dots, n-1\}$ para o qual*

$$f(y_*^j) < f(y_*^{j+1}) \text{ em que } f(y_*^j) = \lim_{k \rightarrow \infty} f(y_k^j), \quad (21)$$

então existe uma iteração com índice $T \in \mathbb{N}$ tal que para todo $t > T$, o índice da mudança satisfaz:

$$k^* > j. \quad (22)$$

Isto é, os j primeiros vértices de todos os simplex permanecem fixos após a iteração T .

Demonstração. Provaremos este teorema por contradição. Pela Hipótese dada em (21), temos que $f(y_*^j) + \delta = f(y_*^{j+1})$, para algum $\delta > 0$. Seja $\varepsilon > 0$ tal que $\delta > \varepsilon \Leftrightarrow \delta - \varepsilon > 0$. Como $f(y_*^j) = \lim_{t \rightarrow \infty} f(y_t^j)$, existe $T \in \mathbb{N}$ tal que para todo $t \geq T$ temos $f(y_t^j) - \varepsilon \leq f(y_*^j)$. Teremos então,

$$f(y_k^j) < f(y_k^j) + \delta - \varepsilon \leq f(y_*^j) + \delta = f(y_*^{j+1}).$$

Perceba que a última igualdade é dada pela suposição feita anteriormente. Como buscamos encontrar o mínimo da função, teremos que $f(y_*^{j+1}) \leq f(y_s^{j+1})$, para algum $s \in \mathbb{N}$, teremos que para todo $t \geq T$

$$f(y_t^j) < f(y_s^{j+1}).$$

Agora, se $k^* \leq j$, para algum $t \geq T$, então, como não há encolhimento

$$f(y_{t+1}^{j+1}) = f(y_t^j) \text{ e } y_{t+1}^j = y_t^{j-1}, \quad j > k^*,$$

contradizendo a Inequação (22). Portanto, $k^* > j$, para todo $k \geq K$. □

Além do auxílio para encontrar o índice de mudança, esse teorema também é importante para outro resultado que auxilia na convergência do método.

Corolário 56. *Seja f uma função limitada inferiormente em \mathbb{R}^n . Se não houver encolhimento e o índice de mudança for $k_t^* = 0$, para toda iteração t , então temos que $f_*^0 = f_*^1 = \dots = f_*^n$.*

Demonstração. Notemos que $k_t^* = 0$ para todas as iterações. Além disso, como por hipótese não há encolhimento, então segue que

$$\begin{aligned} f_{t+1}^i &< f_t^i & \text{ e } & y_{t+1}^i \neq y_t^i & \text{ quando } i = 0 \\ f_{t+1}^i &= f_t^i & \text{ e } & y_{t+1}^{i+1} = y_t^i & \text{ quando } i \geq 1 \end{aligned}$$

Daí, como $f_*^i = \lim_{t \rightarrow \infty} f_t^i$, então

$$f_*^i = f_*^{i+1} \text{ para } i \geq 1.$$

Portanto, $f_*^0 = f_*^1 = \dots = f_*^n$. □

Observemos que, mesmo com esses teoremas demonstrados, ainda não conseguimos demonstrar a convergência do método para um ponto de mínimo.

4.1.4 Convergência do Método de Nelder-Mead

No artigo de Nelder e Mead (1965), o qual foi proposto o método, não contém uma demonstração da convergência para o mínimo da função objetivo. Dessa forma, muitos trabalhos buscaram essa demonstração. O artigo de Lagarias et al. (1998) mostrou a convergência para problemas unidimensionais e comentou sobre a dificuldade em afirmar a convergência do método para espaços bidimensionais.

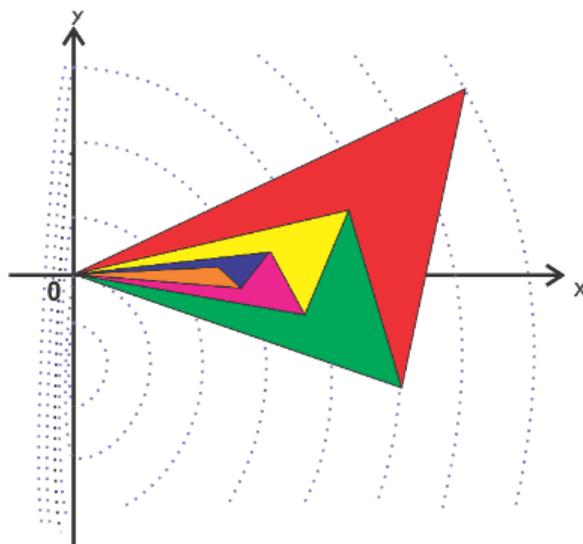
Podemos ver na verdade que existem funções tais que o método não converge para o mínimo da função. Conforme Conn, Scheinberg e Vicente (2009 apud WOODS, 1985, p. 149), foi encontrado uma função diferenciável e não convexa com duas variáveis para a qual o método falha. O motivo da falha do método é causado por uma sequência de nós de encolhimento convergindo para um ponto diferente do mínimo da função.

McKinnon (1998) consegue demonstrar a convergência para um ponto não estacionário em uma família de funções de duas variáveis estritamente convexas com suas terceiras, ou mais, derivadas contínuas. Um caso abordado no artigo é dado por

$$f(x,y) = \begin{cases} 360x^2 + y + y^2 & \text{se } x \leq 0 \\ 6x^2 + y + y^2 & \text{se } x \geq 0 \end{cases} \quad (23)$$

com os valores iniciais $x_1 = (0,0)$, $x_2 = (1,1)$ e $x_3 = ((1 + \sqrt{33})/8, 1 - \sqrt{33})/8$. Aplicando o Método de Nelder-Mead, teremos que o algoritmo converge para o ponto $(0,0)$ com $f(0,0) = 0$. Vejamos a figura abaixo.

Figura 22 – Convergência do Método de Nelder-Mead em f dada em (23).



Notemos que a função é minimizada no ponto $(0,0.5)$ e $f(0,0.5) = -0.25$. Além disso, é importante observar que mesmo sem definir um critério de parada, o método convergiu em 1336 iterações. Em princípio, a justificativa desse fato, se dá pela propagação de erros causados pelo computador, e também, pela limitação computacional na representação de números. Ou seja, em algum momento, as operações feitas pelo Método de Nelder-Mead, resultaram em um número tão pequeno que não pode ser considerado por um computador.

Ainda, nos trabalhos de Torczon de 1989 e 1991, o Método de Nelder-Mead é modificado para que haja convergência para funções continuamente diferenciáveis e limitadas inferiormente (TORCZON, 1989; TORCZON, 1991).

Existem diversas modificações criadas a partir do método clássico, o qual foi apresentado nesse trabalho, com garantia de convergência em um conjunto específico de funções. Mas o Método de Nelder-Mead é um dos métodos de busca direta mais utilizado pois na maioria dos problemas da literatura o método é capaz de encontrar um minimizador. Uma técnica, utilizada por Gonçalves (2013) para diminuir a chance de erro da convergência do método, é recomençar o Método de Nelder-Mead em diferentes simplex e depois comparar os valores obtidos. Além disso, na prática, o algoritmo converge para a grande maioria dos problemas testados na literatura.

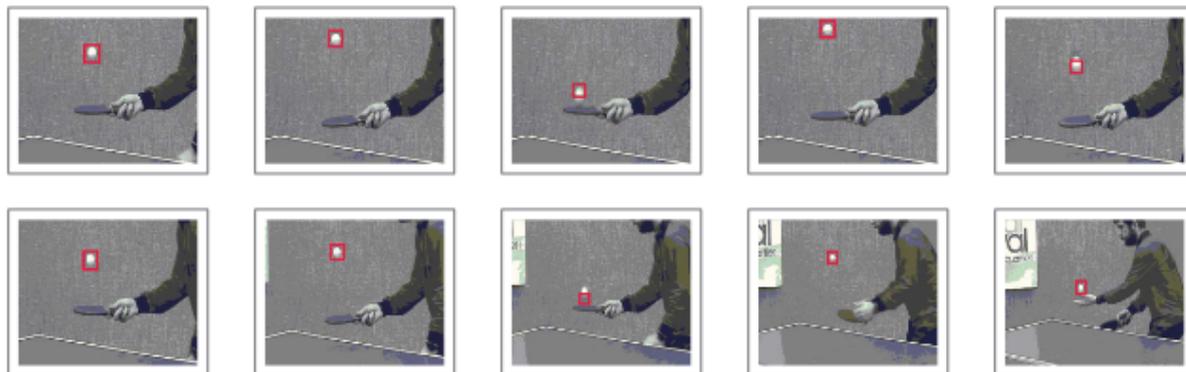
4.1.5 Aplicações do Método de Nelder-Mead

É comum encontrarmos na literatura aplicações utilizando o Método de Nelder-Mead para encontrar parâmetros e minimizar o erro causado na solução de um problema. O artigo de Ouria e Toufigh (2009) aproxima um polinômio de grau 4 utilizando o Método de Nelder-Mead a uma linha freática⁵, buscando modelar um problema de infiltração não confinado. Tal problema raramente resulta em uma linha freática contínua e suave.

Já no artigo de Victor, Bandeira e Lima (2012), o Método de Nelder-Mead foi utilizado para encontrar objetos, em tempo real, em vídeos. Lembrando que um vídeo é formado por uma sequência de imagens, em que cada imagem é chamada de *frame*. Então, para cada *frame*, o algoritmo localiza um objeto pré definido. O vídeo utilizado foi de um jogador de tênis de mesa, no qual o objetivo era localizar a bolinha na imagem. O resultado obtido está ilustrado na Figura 23 abaixo.

⁵Linha freática é a linha que separa em uma superfície a parte seca da parte úmida.

Figura 23 – Identificação em imagem utilizando o Método de Nelder-Mead.



Fonte: Victor, Bandeira e Lima (2012).

O retângulo ao redor da bolinha foi gerado pelo Método de Nelder-Mead. Um ponto bem importante, é que com apenas 100 iterações por frame, o método conseguiu localizar a bolinha mesmo com um afastamento da imagem. O processo não ultrapassou de 0,01 segundos por *frame*, sendo uma quantidade de tempo razoável até para processos em tempo real.

5 CONSIDERAÇÕES FINAIS

Ao longo desse trabalho, abordamos os problemas de otimização, que podem ser generalizados da seguinte forma

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && x \in D \end{aligned} \tag{24}$$

sendo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ chamada de função objetivo e $D \subseteq \mathbb{R}^n$ de conjunto factível. Para a resolução de uma família de problemas da forma (24), estudamos dois métodos iterativos livres de derivação: o Método Simplex, criado para solucionar Problemas de Programação Linear (PPL) e o Método de Nelder-Mead, o qual é utilizado para solucionar Problemas Irrestrito de Programação Não Linear em que $D = \mathbb{R}^n$.

No Capítulo 2, introduzimos a Geometria Simplética, utilizando conceitos de conjuntos convexos e conjuntos afim, com a finalidade de definir um $n - simplex$. Vimos que um $n - simplex$ pode ser entendido como a generalização do conceito de triângulo em \mathbb{R}^2 . Determinamos, dessa maneira, que um $2 - simplex$ é um triângulo no plano e um $3 - simplex$ é um tetraedro no espaço, formados por 3 e 4 vértices, respectivamente. Para isso, foi necessário definir que se $Y = \{y_0, y_1, \dots, y_n\}$ é independentemente afim então o fecho convexo de Y é um $n - simplex$, em que os pontos y_i são chamados de vértices do $simplex$.

Iniciamos o estudo de otimização no Capítulo 3, definindo os Problema de Programação Linear, que foram generalizados pelo seguinte esquema:

$$\begin{aligned} &\text{minimizar} && c^T x \\ &\text{sujeito a} && Ax = b \\ &&& e \quad x \geq 0 \end{aligned} \tag{25}$$

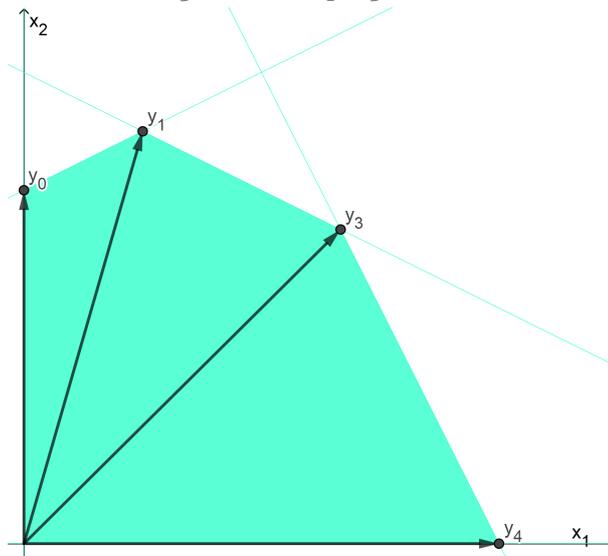
com $x, c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ e $A \in \mathbb{R}^{m \times n}$. Como a matriz A não é invertível, existem diversas soluções para o sistema $Ax = b$ encontrado em (25). Essas soluções, embora não sejam todas soluções para o problema, são importantes para o funcionamento dos métodos iterativos que buscam a solução do problema, isto é, a solução ótima. Diante disso, foram chamadas de soluções viáveis (ou factíveis).

Dentre as soluções viáveis, definimos ainda as soluções básicas do sistema, as quais são as soluções que têm $n - m$ valores iguais a zero. Vimos que se justifica a sua definição pelo fato de que são importantes para o desenvolvimento do Método Simplex, o qual foi utilizado para obter a solução ótima.

Vimos ainda que a ideia central do Método Simplex é reduzir a busca de soluções ótimas pelos vértices do politopo gerado pelas restrições $Ax = b$ e $x \geq 0$. Dois resultados teóricos sustentam o Método Simplex: o primeiro deles é o Teorema Fundamental da

Programação Linear (Teorema 31 do Capítulo 3), o qual garante que se há uma solução ótima então há uma solução básica ótima e o segundo resultado (Teorema 34 do Capítulo 3) indica que os vértices do politopo são soluções básicas factíveis. Neste sentido, o Método Simplex pode ser visto, de maneira geométrica, como um método que gera uma sequência de vértices de politopos até obter uma solução ótima. Repetindo a Figura (14) para melhor compreender, podemos ter uma interpretação geométrica do Método Simplex.

Figura 24 – Intersecção de Inequações com vetores do PPL.



Fonte: Autoria própria

Lembremos que o polígono é a representação das restrições em um PPL no \mathbb{R}^2 e os vetores são as soluções básicas do sistema. Temos, portanto, que cada iteração do Método Simplex determina um vetor y_i da Figura 24.

Um problema encontrado no Método Simplex é que precisamos estabelecer uma solução básica inicial. Nos casos em que todas as restrições são dadas por inequações, definiremos facilmente a solução básica utilizando as variáveis de folga e excesso. Caso houver, uma ou mais, equações, a solução básica não é trivial. Para lidar com esse problema, é desenvolvido o Método Simplex de Duas Fases. Tal método consiste em resolver um PPL artificial (Fase I), encontrando assim uma solução básica do PPL original. E por fim, utilizarmos a solução básica encontrada, para encontrar a solução ótima (Fase II). Utilizando os dois métodos, em teoria, é possível solucionar qualquer PPL definido anteriormente.

Buscando minimizar o custo computacional, foi introduzido ainda o algoritmo do Método Simplex Revisado. Esse é um método baseado em matrizes e é capaz de desconsiderar algumas operações que seriam feitas no Método Simplex clássico. Se o PPL tem a matriz das restrições $A \in \mathbb{R}^{m \times n}$, com m muito menor que n , então o custo computacional do novo método reduzirá significativamente.

Para comprovar a eficácia da aplicação do Método Simplex, abordamos o artigo Kripka e Peccati (2014) que modifica o problema clássico da dieta. O problema modificado busca minimizar a quantidade de calorias em uma dieta, considerando os nutrientes mínimos a serem consumidos. A dieta resultante obteve quase que 30% menos calorias do que uma dieta feita por uma nutricionista.

Utilizando o Método Simplex podemos organizar o PPL em um tableau, de maneira a encontrar a solução ótima apenas realizando operações básicas da Álgebra Linear.

No Capítulo 4, abordamos os Problemas Irrestritos de Programação Não Linear e sua resolução através do Método de Nelder-Mead. O método funciona da seguinte maneira: em uma função objetivo $f : \mathbb{R}^n \rightarrow \mathbb{R}$, fornecemos inicialmente $n + 1$ pontos iniciais construindo um conjunto $Y = \{y^0, y^1, \dots, y^n\}$, que por sua vez, gera um $n - simplex$. Baseado nesses pontos, o método gera um novo simplex. Para isso, é substituído, a cada iteração, o pior ponto y^n . Essa substituição é feita baseada na reta que passa por y^n e y^c , sendo y^c o ponto equidistante de y^0, y^1, \dots, y^{n-1} . Utilizando as propriedades dos conjuntos simpléticos, conseguimos demonstrar algumas propriedades do Método de Nelder-Mead. Vimos um exemplo numérico, em que o método converge rapidamente para um dos pontos de mínimo da função. Contudo, a desvantagem desse método é que não há um resultado geral que garanta a convergência para um valor de mínimo da função objetivo. Inclusive, apresentamos um exemplo no qual o Método de Nelder-Mead não converge para o minimizador da função. Trazemos também uma aplicação do método na localização de objetos em vídeos, feita em tempo real. Essa aplicação utilizou 100 iterações do método, e ainda obteve um resultado muito satisfatório.

O presente trabalho tem sua importância pelo fato de que podemos solucionar tanto os Problemas de Programação Linear, utilizando o Método Simplex, assim como podemos utilizar o Método de Nelder-Mead para solucionar Problemas Irrestritos de Programação Não Linear. Ou seja, grande parte dos problemas podem ser solucionados utilizando os métodos estudados aqui, os quais são métodos clássicos e amplamente utilizados pelas mais diversas áreas de conhecimento. Além disso, a discussão sobre os conceitos da Geometria Convexa e dos Simpléticos mostrou-se contributiva, uma vez que dificilmente as literaturas na área de otimização realizam uma abordagem ampla como a que foi feita neste trabalho. Dessa forma, a bagagem teórica desenvolvida aqui permite trabalhos futuros tanto na área de otimização quanto na área de Análise Convexa, desenvolvendo novos métodos ou variações dos métodos aqui estudados.

Referências

- BAZARAA, M. S.; SHETTY, C. M.; SHERALI, H. D. **Nonlinear Programming-Theory and algorithms**. 3. ed. [S.l.: s.n.], 1979. Citado na página 7.
- CONN, A. R.; SCHEINBERG, K.; VICENTE, L. N. **Introduction to Derivative-Free Optimization**. [S.l.]: the Society for Industrial and Applied Mathematics and the Mathematical Programming Society, 2009. ISBN 9780898716689. Citado 5 vezes nas páginas 8, 23, 57, 63 e 72.
- DANTZIG, G. B. Origins of the simplex method. 1987. Disponível em: <<https://apps.dtic.mil/dtic/tr/fulltext/u2/a182708.pdf>>. Acesso em: 01 de Novembro de 2020. Citado 2 vezes nas páginas 8 e 10.
- DANTZIG, G. B. The diet problem. **The Computer Journal Institute for Operations Research and the Management Sciences**, v. 20, p. 43–47, 1990. Disponível em: <<https://www.mpi-inf.mpg.de/fileadmin/inf/d1/teaching/winter18/Ideen/Materialien/Dantzig-Diet.pdf>>. Acesso em: 01 de Novembro de 2020. Citado 2 vezes nas páginas 9 e 54.
- FRIEDLANDER, A. **Elementos de Programação Não-Linear**. [s.n.], 1994. Disponível em: <<https://www.ime.unicamp.br/~friedlan/livro.htm>>. Acesso em: 28 de Outubro de 2020. Citado na página 7.
- GIVENS, G. H.; HOETING, J. A. **Computational Statistics**. 2. ed. John Wiley and Sons, Inc., 2013. ISBN 9780470533314. Disponível em: <[http://home.usc.edu.cn/~liweiyu/documents/Geof%20H.%20Givens%20%20Jennifer%20A.%20Hoeting\(auth.\)%20-%20Comp.pdf](http://home.usc.edu.cn/~liweiyu/documents/Geof%20H.%20Givens%20%20Jennifer%20A.%20Hoeting(auth.)%20-%20Comp.pdf)>. Acesso em: 01 de Novembro de 2020. Citado na página 9.
- GONÇALVES, A. M. S. **O Problema de min-max-min com restrições pelo Método de Nelder-Mead**. Tese (Doutorado) — UNIVERSIDADE FEDERAL DO RIO DE JANEIRO, 2013. Disponível em: <<https://www.cos.ufrj.br/uploadfile/1378146249.pdf>>. Acesso em: 27 de Outubro de 2020. Citado 7 vezes nas páginas 10, 61, 62, 63, 64, 72 e 73.
- IMPA. **Geometria Simplética**. 2020. Disponível em: <<https://impa.br/pesquisa/geometria-simpletica/>>. Acesso em: 01 de Novembro de 2020. Citado na página 8.
- KELLEY, C. T. Detection and remediation of stagnation in the Nelder-Mead algorithm using a sufficient decrease condition. **SIAM Journal on Optimization**, v. 10, n. 1, p. 43–55, 1999. Disponível em: <<https://doi.org/10.1137/S1052623497315203>>. Citado na página 68.
- KRIPKA, R.; PECCATI, C. Uma variação para o problema clássico da dieta: A minimização do consumo de calorias. **Revista CIATEC-UPF**, v. 6, p. 13, 05 2014. Citado 3 vezes nas páginas 54, 55 e 77.
- LAGARIAS, J. et al. Convergence properties of the Nelder–Mead simplex method in low dimensions. **SIAM Journal on Optimization**, v. 9, p. 112–147, 12 1998. Citado na página 72.

LUENBERGER, D. G.; YE, Y. **Linear and Nonlinear Programming**. 3. ed. [S.l.]: Springer Science+Business Media, LLC, 2008. ISBN 9780387745022, 9780387745039. Citado 2 vezes nas páginas 25 e 36.

MCKINNON, K. I. M. Convergence of the Nelder–Mead simplex method to a nonstationary point. **Society for Industrial and Applied Mathematics**, 1998. ISSN 10957189. Disponível em: <<https://doi.org/10.1137/S1052623496303482>>. Acesso em: 27 de Outubro de 2020. Citado na página 72.

NELDER, J. A.; MEAD, R. A Simplex Method for Function Minimization. **The Computer Journal**, v. 7, n. 4, p. 308–313, 01 1965. ISSN 0010-4620. Disponível em: <<https://doi.org/10.1093/comjnl/7.4.308>>. Citado 3 vezes nas páginas 10, 57 e 72.

OURIA, A.; TOUFIGH, M. M. Application of Nelder-Mead simplex method for unconfined seepage problems. **Applied Mathematical Modelling**, v. 33, n. 9, p. 3589 – 3598, 2009. ISSN 0307-904X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0307904X08003296>>. Citado na página 73.

ROCKAFELLAR, T. R. **Convex Analysis**. 2. ed. [S.l.]: Princeton University Press, 1972. ISBN 0691080690. Citado na página 11.

TORCZON, V. On the convergence of the multidirectional search algorithm. **SIAM Journal on Optimization**, v. 1, n. 1, p. 123–145, 1991. Disponível em: <<https://doi.org/10.1137/0801010>>. Citado na página 73.

TORCZON, V. J. **Multidirectional search: A direct search algorithm for parallel machines**. Tese (Doutorado) — Rice University, 1989. Disponível em: <<https://scholarship.rice.edu/handle/1911/16304>>. Acesso em: 11 de Novembro de 2020. Citado na página 73.

VICTOR, S. E. C. B. de H.; BANDEIRA, S. S.; LIMA, E. A. de O. Métodos de otimização direta aplicados ao rastreamento de imagens em vídeo. **Congresso de Matemática Aplicada e Computacional**, 2012. Disponível em: <<http://sbmac.locaweb.com.br/cmacts/cmact-ne/2012/trabalhos/PDF/156.pdf>>. Acesso em: 17 de Novembro de 2020. Citado 2 vezes nas páginas 73 e 74.

WOODS, D. J. **An Interactive Approach for Solving Multi-objective Optimization Problems**. Tese (Doutorado) — Rice University, 1985. Citado na página 72.