Introdução à Programação Linear

J. P. S. de Sousa

Contents

Chapter 1

Introdução

1.1 Um Breve Histórico da Programação Linear

O problema de programação matemática, em sentido mais amplo, pode ser entendido como aquele que pretende otimizar uma função - seja maximizar ou minimizar - dado um conjunto de restrições. Esse problema surge naturalmente em uma série de aplicações, porém ele se tornou especialmente interessante durante e após a Segunda Guerra Mundial. A razão é que num conflito daquela escala problemas de natureza de "gerir e alocar recursos escassos de maneira eficiente" se tornaram uma das principais preocupações dos militares.

Provavelmente o primeiro a formular explicitamente a teoria e problemas de programação linear foi o matemático soviético Leonid Kantorovich (1912-1986) enquanto trabalhava em problemas de planejamento e gestão de recursos para os militares da URSS na década de 1930. Contudo, os trabalhos de Kantorovich permaneceram desconhecidos no Ocidente até o final dos anos de 1950 devido à Cortina de Ferro da Guerra Fria.

A historiografia da Programação Linear geralmente atribui como marco de seu início o desenvolvimento do método simplex em 1947 pelo matemático norte-americano George B. Dantzig (1914-2005), que o criou como parte dos esforços do *Scientific Computation Optimum Program* (SCOOP) da força aérea dos Estados Unidos. O termo "programação linear" é atribuído ao matemático americano T. C. Koopmans, em que a palavra programação não tinha o sentido que possui hoje, estando mais atrelada ao sentido de "planejamento".

Desde a invenção do simplex, a teoria e os algoritmos, não só de programação linear, mas da otimização matemática de forma geral, foram desenvolvidos e expandidos por uma série de autores ao longo do século XX. Suas aplicações espalharam-se para diversos campos, como a própria matemática, a pesquisa operacional, a economia, estatística e diversas outras ciências. Os problemas de PL também impulsionaram o desenvolvimento de métodos para problemas de maior complexidade, como a programação discreta, a programação não linear, a programação combinatória, programação estocástica, problemas de controle ótimo, etc.

1.2 O Problema de Programação Linear

Como dito anteriormente, o problema geral da programação matemática é otimizar uma função - seja maximizá-la ou minimizá-la - dada um conjunto de restrições sobre suas variáveis. Quando tanto a função quanto suas restrições são lineares, chamamos o problema de programação matemática de **Problema de Programação Linear** (PPL).

Dada uma função $f : \mathbb{R}^n \to \mathbb{R}$ definida como

$$f(x) = c^t x, \ c \in \mathbb{R}^n$$

e um conjunto de inequações

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

chamamos a função a ser otimizada $f(x)$ de **função objetivo**, as variáveis $x_1, \ldots, x_2$ de **variáveis de decisão** e o conjunto de inequações de **restrições**. A verdade que as restrições não necessariamente precisam ser da forma $a_i^t x \leq b_i$, podendo ser da forma $a_i^t x \geq b_i$ ou $a_i^t x = b_i$. O conjunto dos vetores $x \in \mathbb{R}$ que são solução para o Problema de Programação Linear é chamado de **conjunto viável**, ou factível.

1.2.1 Notação Matricial

Para facilitar a notação, podemos usar a representação matricial para o conjunto de restrições. Se a_{ij} é a entrada da i-ésima linha e j-ésima coluna da matriz A, de modo que

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

então reduzimos as restrições simplesmente a

$$Ax \leq b$$

, tal que

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Notação de desigualdades para matrizes pode parecer estranha, e realmente é. Mas ela é muito comum ao descrever problemas como os quais estamos lidando. Nesta notação, temos que: se a_i é o vetor cujas componentes são as entradas da i-ésima linha de A, então a desigualdade acima diz que, para todo $i = 1, 2, \ldots, m$ é verdade que

$$a_i^t x \leq b_i$$

.

1.2.2 Restrições de Não Negatividade

Por se tratar de um ramo que se desenvolveu em boa parte a partir de problemas concretos, a programação linear lida com variáveis de decisão que geralmente representam quantidades de alguma grandeza do mundo real, e por isso são não negativas. Logo, é comum a adição ao conjunto de restrições a expressão

$$x \geq 0$$

que significa que todas as componentes do vetor das variáveis de decisão são não negativas.

Mas isso não anula a existência de variáveis que possam assumir valores negativos, que geralmente são chamadas de **variáveis irrestritas**. Contudo, se x_j for uma variável irrestrita, então podemos expressá-la como $x_j = x_j^+ - x_j^-$, tal que $x_j^+ \geq 0$ e $x_j^- \geq 0$. Portanto, podemos expressar uma variável irrestrita a partir de outras duas não negativas. Ademais, se temos k variáveis irrestritas, necessitaremos de k variáveis x^+ e apenas uma x^-, que será a mais negativa das expressões $x_j = x_j^+ - x_j^-$.

Se $x_j > l_j$ então podemos trocá-la por $x_j' = x_j - l_j$, que será não negativa. De forma análoga, se tivermos que $x_j < u_j$, então podemos trocar essa variável por $x_j' = u_j - x_j$, que também será não negativa.

1.2.3 Inequações

Podemos transitar facilmente entre os tipos de inequações apenas multiplicando-as por -1. Por exemplo, uma inequação da forma "menor ou igual" como $2x_1 + 3x_2 \leq 5$ é equivalente a $-2x_1 - 3x_2 \geq -5$, que é da forma "maior ou igual".

Também é importante sabermos com transitar entre equações e inequações, pois o método simplex, por exemplo, aceita a entrada da restrições apenas como igualdades. Desse modo, se toemos uma inequação como $4x_1 + x_2 \geq 10$, para a transformar numa igualdade, iremos subtrair do primeiro membro uma nova variável não negativa x_3, que é chamada de **variável de folga**, obtendo assim a igualdade $4x_1 + x_2 - x_3 = 10$. Da mesma forma, se tivéssemos a inequação $7x_1 + 2x_2 \leq 9$, poderíamos somar a variável de folga x_3 a inequação para obter uma igualdade: $7x_1 + 2x_2 + x_3 = 9$.

1.2.4 Problemas de Maximização e Minimização

A otimização da função objetivo $f(x) = z$ pode se referir a sua maximização, ou sua minimização. Contudo, ambos os problemas são equivalentes. Se o problema é $\max z = c^t x$, temos que ele é equivalente ao $\min z' = -c^t x$, tal que se $\hat{z}$ é o valor ótimo do primeiro e $\hat{z}'$ o do segundo, então $\hat{z} = -\hat{z}'$.

Logo, para transformar passar de um problema para outro, basta multiplicarmos os coeficientes da função por -1.

$$\text{maximizar } c^t x = \text{minimizar } -c^t x$$

1.2.5 Forma Padrão e Canônica

Um PPL está na forma **padrão** se todas as restrições são igualdades e todas as variáveis são não negativas. O método simplex, por exemplo, resolve programas lineares apenas quando estão na forma padrão.

A forma canônica de um PPL depende do tipo de otimização que estamos buscando. Se o problema é de maximização, então sua forma canônica é dada por

$$\begin{array}{rl} \text{Maximizar:} & z = c^t x \\ \text{sujeito a:} & Ax \leq b \\ & x \geq 0 \end{array}$$

e quando de minimização, então sua forma canônica é

$$\begin{array}{rl} \text{Minimizar:} & z = c^t x \\ \text{sujeito a:} & Ax \geq b \\ & x \geq 0 \end{array}$$

1.3 Modelos de Programação Linear

Chapter 2

Teoria da Programação Linear

2.1 Espaços Convexos

2.1.1 Conjuntos Convexos

Um dos conceitos mais fundamentais para a Teoria de PL é a de **convexidade**. Na Geometria, convexidade é definida como a propriedade de uma figura em que, para quaisquer pontos no seu interior, o segmento de reta entre esses pontos está inteiramente contido na figura. Mas podemos definir convexidade algebricamente da seguinte forma:

Definição 2.1. Um conjunto $\mathbb{V}$ no $\mathbb{R}^n$ é chamado convexo se para quaisquer vetores $x_1, x_2 \in \mathbb{V}$ é verdade que

$$\lambda x_1 + (1 - \lambda)x_2 \in \mathbb{V}$$

para todo $\lambda \in [0, 1]$

Aqueles com algum conhecimento de geometria analítica devem se lembrar que uma combinação linear de vetores na forma $\lambda x_1 + (1 - \lambda)x_2$ com $\lambda \in [0, 1]$ define o segmento de reta que une as extremidades desses vetores. Esse tipo de combinação é chamada de convexa.

De forma mais geral, uma combinação linear convexa é aquela em que todos os escales são positivos e sua soma é igual a 1. Então supondo um conjunto de geradores formado por $x_1, \ldots, x_n$ e escalares $\lambda_1, \ldots, \lambda_n$, então a combinação linear

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

é convexa se

$$\sum_{i=1}^{n} \lambda_i = 1,\ 0 \leq \lambda_i \leq 1$$

Além disso, chamamos de combinação estritamente convexa aquelas em que os escales são estritamente maiores que zero ou menores que 1.

2.1.2 Semiespaços e Hiperplanos

Semiespaços e hiperplanos generalizam o conceito geométrico do plano que é dividido por um reta, gerando-se dois semiplanos. Um hiperplano no $\mathbb{R}^n$ é um espaço de dimensão $n - 1$. Um hiperplano divide o $\mathbb{R}^n$ em dois semiespaços.

Algebricamente, um hiperplano H é o conjunto dos vetores $x \in \mathbb{R}^n$ que solucionam uma equação linear na forma $\mathbf{a}^\intercal x = k$, onde $k \in \mathbb{R}$ é uma constante e $a \in \mathbb{R}^n$ é um vetor não nulo, geralmente conhecido como vetor normal ou gradiente.

$$H = \{x \mid a^\intercal x = k\}$$

Por sua vez, semiespaços são definidos algebricamente como inequações lineares. Se H é um hiperplano no $\mathbb{R}^n$ definido pela equação $a^\intercal x = k$, então esse hiperplano divide o $\mathbb{R}^n$ em hiperespaços H^+ e H^- que podem ser dados por

$$H^+ = \{x \mid a^\intercal x \geq k\} \quad H^- = \{x \mid a^\intercal x \leq k\}$$

ou por

$$H^+ = \{x \mid a^\intercal x > k\} \quad H^- = \{x \mid a^\intercal x < k\}$$

Os primeiros são **semiespaços fechados**, isto é, eles incluem sua fronteira, que consiste do hiperplano H. Por sua vez, os segundos são **semiespaços abertos**, pois não incluem sua fronteira. Tanto hiperplanos quanto semiespaços são conjuntos convexos. A demonstração disso fica como exercício para o leitor.

Caso se lembres da forma como expressamos as restrições de um PPL, perceberás que o conjunto viável é na verdade a interseção de vários semiespaços fechados. Portanto, se R é a região factível do PPL cujo conjunto de restrições é dado por $Ax \leq b$, por exemplo, então

$$R = \{x \mid Ax \leq b\}$$

e como a interseção de conjuntos convexos é também convexo, então as restrições de um problema de programação linear formam um conjunto convexo.

2.1.3 Poliedros e Politopos

Normalmente, aprendemos no ensino básico que poliedros são figuras tridimensionais formadas por faces poligonais. Por sua vez, polígonos seriam figuras bidimensionais limitadas, ou seja possuem uma área finita, formadas por segmentos de retas que se encontram em pontos chamados de vértices.

Politopos surgem como uma espécie generalização dos conceitos clássicos de polígonos e poliedros **convexos** para o caso n-dimensional. Dessa forma, poliedros convexos, por exemplo, seriam politopos tridimensionais, enquanto que os polígonos convexos são politopos bidimensionais.

Apesar da noção clássica de poliedros se limitar ao caso tridimensional, a literatura matemática estende este conceito para n dimensões. Contudo, a definição formal de um poliedro de n-dimensões não é um consenso, e não é difícil encontrar definições diferentes, e por vezes não equivalentes, do que se trata esse objeto matemático.

Dito isso, e pelo fato da Teoria de PL se concentrar mais especificamente na classe dos poliedros **convexos**, que são melhores definidos, não iremos tratar das outras classes. Dessa forma, segue uma definição conveniente para ao Teoria de PL do que é um poliedro convexo.

Definição 2.2. Um poliedro convexo é a interseção de um número finito de semiespaços.

Por sua vez, politopos são dados como

Definição 2.3. Um politopo é um poliedro convexo limitado

Para aqueles que não lembrarem, o conceito de limitado na matemática não é equivalente ao de finito. Formalmente, um conjunto limitado é aquele que pode ser contido dentro de uma bola de raio finito.

Das definições percebemos que poliedros convexos podem ser dados como um sistema de inequações ou equações lineares. Por conseguinte, o espaço das soluções viáveis R de um problema de PL é um poliedro convexo. Contudo, nem sempre esse espaço será um politopo, pois nem sempre temos um conjunto limitado de soluções factíveis.

Visto isso, passaremos a nos referir a conjuntos convexos definidos como $X = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{p}\}$, onde A é uma matriz $m \times n$, como **conjunto poliédrico** ou **poliedro**.

2.1.4 Cone Convexo

Cones lineares, chamados geralmente apenas de cones, são um subconjunto de um espaço vetorial $\mathbb{V}$ fechado sobre multiplicação por escalar positivo. Algebricamente, um cone é definido como um conjunto C em que para todo $\mathbf{x} \in C$ e todo $\lambda \in \mathbb{R}$ não negativo é verdade que $\lambda\mathbf{x} \in C$. Dessa definição segue que a origem sempre é um membro do cone, caso em que $\lambda = 0$. Uma classe especial dos cones são os convexos, que são definidos como

Definição 2.4. Um cone C é convexo se para quaisquer $\mathbf{x_1}, \mathbf{x_2} \in C$ temos que $\mathbf{x_1}+\mathbf{x_2} \in C$

Do que foi dito até então, é fácil ver que, se um conjunto é fechado por multiplicação por escalares positivos e sobre a adição, então ele é um conjunto convexo.

Devido a essas definições, combinações lineares entre pontos tais que os escales são estritamente positivos são chamadas de combinações cônicas. Dito isso, iremos enunciar uma definição que será mais importante no futuro.

Definição 2.5. Seja P um conjunto de pontos em um espaço vetorial $\mathbb{V}$. Chama-se envoltória cônica o conjunto $\mathrm{Cone}(P)$ o conjunto de todas as combinações cônicas dos elementos de P

Uma classe ainda mais especial dos cones, são os cones poliédricos convexos

Definição 2.6. Se A é um matriz $m \times n$, então um cone poliédrico convexo C é aquele definido por

$$C = \{\mathbf{x} \mid A\mathbf{x} \leq 0\}$$

Uma propriedade interessante dessa classe de cones é que a iésima linha da matriz A é o vetor normal ao íésimo hiperplano dado por

$$\mathbf{a}^{\mathsf{T}}\mathbf{x} = 0$$

Os cones poliédricos convexos desempenham um importante papel na Teoria dos Problemas Duais de PL, que abordaremos nos capítulos mais adiante.

2.2 Geometria dos Espaços Convexos

2.2.1 Pontos Extremos

O conceito de ponto extremo de um conjunto convexo $\mathbb{V}$ é uma especialização do conceito de "independência linear" para combinações lineares convexas.

Definição 2.7. Um vetor x num conjunto convexo $\mathbb{V}$ é chamado ponto extremo de $\mathbb{V}$ se ele não pode ser expresso como combinação linear convexa de nenhum outro subconjunto de $\mathbb{V}$.

De grosso modo, podemos dizer que os pontos extremos de um conjunto convexo é um conjunto linearmente independente se considerarmos apenas suas combinações convexas. Além disso, é possível ver também que os pontos extremos formam uma "base" para conjuntos convexos, ou seja, qualquer ponto de um conjunto convexo pode ser dado como combinação convexa dos pontos extremos, mas não se preocupe com isso agora, pois iremos explorar essa ideia mais adiante.

Agora, iremos explorar mais o significado geométrico de um ponto extremo. Mas para isso, iremos demonstrar um resultado que irá nos auxiliar nesta tarefa.

Proposição 2.1. Se α e β são números reais e γ é uma combinação linear estritamente convexa desses números, tal que $\alpha \leq \gamma$ e $\beta \leq \gamma$, então

$$\alpha = \beta = \gamma$$

Demonstração. Primeiros vamos explicitar a ideia geométrica de uma combinação linear convexa para números reais. Tradicionalmente, o conjunto dos reais é compreendido como uma reta cujos pontos estão, respectivamente, associados a um número real. Dessa forma, o conjunto das combinações lineares convexas entre dois pontos da reta real é o segmento de reta que une esses dois pontos.

Dito isso, se γ é uma combinação convexa de α e β, então é verdade que

$$\alpha \leq \gamma \leq \beta \tag{2.1}$$

uma vez que o ponto associado a γ deve estar disposto no segmento unindo os pontos associados a α e β. Visto que, por hipótese

$$\alpha \leq \gamma \quad \beta \leq \gamma$$

então, da expressão à direita e de (2.1) segue que $\beta = \gamma$, ao passo que

$$\begin{aligned} \lambda\alpha + (1-\lambda)\gamma &= \gamma \\ \lambda\alpha &= \lambda\gamma \\ \alpha &= \gamma \end{aligned}$$

■

Essa proposição, num primeiro momento, parece estar desconectada da ideia dos pontos extremos, mas apesar de simples, ela é fundamental para construirmos a geometria daquele conceito. Vamos enunciar mais uma proposição para seguirmos com o nosso objetivo. Em resumo, ela diz que, se um ponto em um hiperplano é combinação convexa de outros dois que compartilham o mesmo semiespaço, então esses dois últimos também estão no hiperplano em questão.

Proposição 2.2. Sejam os pontos $\mathbf{x}'$ e $\mathbf{x}''$ pertencentes a um mesmo semiespaço H^-. Se uma combinação estritamente convexa desses pontos está no hiperplano H, então ambos os pontos pertencem a esse hiperplano.

Demonstração. Seja $\bar{\mathbf{x}}$ uma combinação estritamente convexa de $\mathbf{x}'$ e $\mathbf{x}''$ tal que

$$\lambda\mathbf{x}' + (1-\lambda)\mathbf{x}'' = \bar{\mathbf{x}}$$

Se $H = \{\mathbf{x} \mid \mathbf{a}^\intercal\mathbf{x} = b\}$ e $\bar{\mathbf{x}} \in H$, então

$$\mathbf{a}^\intercal\bar{\mathbf{x}} = b$$

e disso segue que

$$\lambda(\mathbf{a}^\intercal\mathbf{x}') + (1-\lambda)(\mathbf{a}^\intercal\mathbf{x}'') = b \tag{2.2}$$

Digamos que H^- seja definido como

$$H^- = \{\mathbf{x} \mid \mathbf{a}^\intercal\mathbf{x} \leq b\}$$

Com efeito, se $\mathbf{x}'$ e $\mathbf{x}''$ estão em H^-, então

$$\mathbf{a}^\intercal\mathbf{x}' \leq b \quad \mathbf{a}^\intercal\mathbf{x}'' \leq b$$

Como bem sabemos, o resultado do produto escalar entre vetores é um número real, e o que a expressão 2.2 está nos mostrando é que b é um combinação convexa dos números reais $\mathbf{a}^\intercal\mathbf{x}'$ e $\mathbf{a}^\intercal\mathbf{x}''$. Ora, mas também temos que $\mathbf{x}', \mathbf{x}'' \in H^-$, o que implica pela Proposição 2.1 que

$$\mathbf{a}^\intercal\mathbf{x}' = b \quad \mathbf{a}^\intercal\mathbf{x}'' = b$$

ou seja, $\mathbf{x}', \mathbf{x}'' \in H$ ■

Com esse último resultado em mãos, agora iremos caracterizar os pontos extremos a partir das noções já vistas de semiespaços e hiperplanos. Para viés de síntese, sendo $X = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0\}$, com A sendo $m \times n$, iremos chamar os $n + m$ hiperplanos associados aos semiespaços definidores de X (as n restrições agrupadas em A mais a m restrições de não negatividade) de **hiperplanos definidores** de X.

Teorema 2.1. Um vetor $\bar{\mathbf{x}}$ é ponto extremo de um conjunto poliédrico $X = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{p}\}$, com $A \in \mathbb{R}^{m\times n}$, se $\bar{\mathbf{x}}$ pertence a pelo menos n hiperplanos linearmente independentes que definem X.

Demonstração. Suponha por contradição que $\bar{\mathbf{x}}$ pertence a pelo menos n hiperplanos linearmente independentes definidores de X, mas não é um ponto extremo. Com efeito, $\bar{\mathbf{x}}$ pode ser dado como combinação estritamente convexa de vetores $\mathbf{x}'$ e $\mathbf{x}''$ em X. Pela Proposição 2.2, tanto $\bar{\mathbf{x}}'$ quanto $\bar{\mathbf{x}}''$ devem pertencer a esses n hiperplanos definidores.

Entretanto, o sistema linear dado pelas equações desses hiperplanos é uma sistema quadrado $n \times n$ cujas linhas são linearmente independentes, o que implica que ele tem solução única. Portanto

$$\mathbf{x}' = \mathbf{x}''$$

logo $\bar{\mathbf{x}}$ não pode ser uma combinação estritamente convexa daqueles vetores, contrariando a afirmação de que ele poderia.

Por outro lado, suponhamos agora a contrapositiva de que $\bar{\mathbf{x}}$ pertença a $r < n$ hiperplanos linearmente independentes, e que o sistema linear dado por eles é

$$B\mathbf{x} = \mathbf{c}$$

onde B é uma matriz $r \times n$. O que faremos agora será construir vetores $\mathbf{x}'$ e $\mathbf{x}''$ dos quais $\bar{\mathbf{x}}$ é uma combinação estritamente convexa, e não poderia, pois, ser um ponto extremo.

Por ser uma matriz retangular larga (mais colunas do que linhas), o posto de B é no máximo r, logo ela não pode ter posto cheio, o que implica que existe um vetor $\mathbf{d} \neq 0$ tal que $B\mathbf{d} = 0$. Construímos o vetor $\mathbf{d}$ porque se temos um sistema com infinitas soluções e conhecemos pelos uma, podemos encontrar outras somando essa solução com uma vinda do sistema homogêneo, e usaremos essa ideia para construir $\mathbf{x}'$ e $\mathbf{x}''$.

Digamos que

$$\mathbf{x}' = \bar{\mathbf{x}} + \epsilon\mathbf{d} \quad \mathbf{x}'' = \bar{\mathbf{x}} - \epsilon\mathbf{d}$$

Por conseguinte temos que tanto $\mathbf{x}'$ quanto $\mathbf{x}''$ satisfazem o sistema $B\mathbf{x} = \mathbf{c}$, pertencendo aos r hiperplanos nos quais $\bar{\mathbf{x}}$ também está.

Para que $\mathbf{x}'$ e $\mathbf{x}''$ satisfaçam as demais $n - r$ restrições, basta tomarmos um $\epsilon > 0$ suficientemente pequeno. Isso porque, se pensarmos geometricamente, $\mathbf{x}'$ e $\mathbf{x}''$ estão ambos nas semirretas com vértice em $\bar{\mathbf{x}}$ e cujas direções são dadas por $\mathbf{d}$ e $-\mathbf{d}$ respectivamente. Desse modo, para que não avancemos para além da fronteira de X, basta que ϵ seja pequeno o suficiente para tal.

Por fim, escolhendo $\lambda = 0.5$, temos que

$$
\begin{gathered}
0.5 \cdot \mathbf{x}' + 0.5 \cdot \mathbf{x}'' \\
0.5 \cdot (\bar{\mathbf{x}} + \epsilon \mathbf{d}) + 0.5 \cdot (\bar{\mathbf{x}} - \epsilon \mathbf{d}) \\
0.5 \cdot \bar{\mathbf{x}} + 0.5 \cdot \bar{\mathbf{x}}
\end{gathered}
$$

concluindo então que

$$0.5 \cdot \mathbf{x}' + (1 - 0.5) \cdot \mathbf{x}'' = \bar{\mathbf{x}}$$

e encontramos $\mathbf{x}'$ e $\mathbf{x}''$ dos quais uma combinação estritamente convexa é $\bar{\mathbf{x}}$. Portanto se $\bar{\mathbf{x}}$ não pertence ao menos n dos hiperplanos que definem X, então $\bar{\mathbf{x}}$ não pode ser um ponto extremo. ■

Você talvez deve estar se perguntando agora sobre os pontos extremos que pertencem a mais do que n dos hiperplanos definidores. Para esses pontos usamos uma nomenclatura especial, que iremos formalizar na definição a seguir.

Definição 2.8. Seja $X = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$ um conjunto poliédrico em que A é $m \times n$. Se $\mathbf{x}$ é um ponto extremo de X tal que $\mathbf{x}$ pertença a mais do que n dos $n + m$ hiperplanos definidores de X, então $\mathbf{x}$ é chamado de **ponto extremo degenerado**. O número excedente de hiperplanos que possuem $\mathbf{x}$ é chamado de **ordem de degeneração**.

Em resumo, podemos definir um ponto extremo com n hiperplanos definidores linearmente independentes do nosso conjunto convexo X. Se existe mais de uma maneira de definir esse ponto extremos com n hiperplanos, então esse ponto é um ponto degenerado.

Pontos extremos de um PPL são soluções de um sistema formado por um subconjunto linearmente independente de restrições na forma padrão. Da Álgebra Linear, sabemos que, em $\mathbb{R}^n$, a interseção de n hiperplanos linearmente independentes é um ponto, o qual também pode ser chamado de **vértice**.

Para visualizar isso, considere exemplos em dimensões menores:

- No plano ($\mathbb{R}^2$): Os hiperplanos são retas, e ao desenharmos um polígono, seus vértices correspondem às interseções de pares de retas.

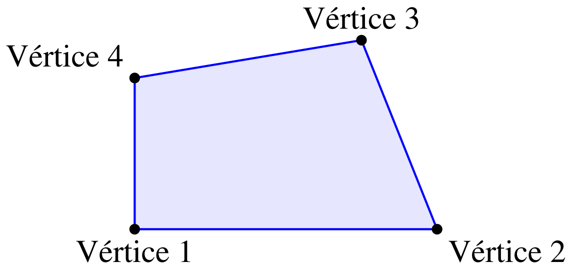


Figure 2.1: Interseção de hiperplanos em $\mathbb{R}^2$

- No espaço tridimensional ($\mathbb{R}^3$): Um vértice de um poliedro é o ponto de encontro de três faces, que correspondem a hiperplanos em um espaço tridimensional.

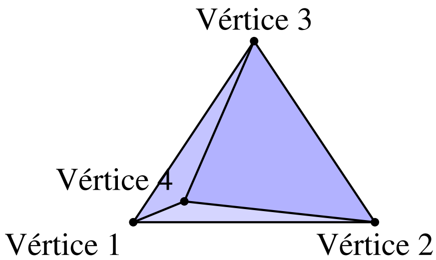


Figure 2.2: Interseção de Hiperplanos em $\mathbb{R}^3$

Portanto, o conceito algébrico de ponto extremo está diretamente relacionado ao conceito geométrico de vértice. Concluímos que os pontos extremos de um conjunto viável de um PPL são, geometricamente, os vértices do poliedro formado por esse conjunto.

Pontos extremos desempenham um papel importante na Programação Linear, pois como veremos adiante, a solução ótima de um PPL sempre está em um ponto extremo do conjunto viável. Para o método simplex, a existência de pontos extremos degenerados requer precauções, pois esses podem afetar o desempenho do método, inclusive fazendo rodar indefinitivamente.

2.2.2 Raios e Direções

Um raio também é uma classe de conjuntos convexos, que são definidos a seguir.

Definição 2.9. Um raio r é uma coleção de pontos dados na forma

$$r = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} + \lambda \mathbf{d}, \lambda > 0\}$$

onde $\mathbf{d}$ é um vetor em $\mathbb{R}^n$ não nulo e λ é um real positivo

Podemos também interpretar raios geometricamente como sendo uma semirreta com origem em $\mathbf{x}$, dito como o vértice do raio, e que se estende na direção do vetor $\mathbf{d}$, chamado de diretor ou direção do raio.

Definição 2.10. Seja X um conjunto convexo contido no $\mathbb{R}^n$. O vetor não nulo $\mathbf{d} \in \mathbb{R}^n$ é uma direção de X se para qualquer ponto $\mathbf{x_0} \in X$ e qualquer $\lambda \geq 0$ é verdade que $\mathbf{x_0} + \lambda\mathbf{d} \in X$

A Definição expande a noção de direção para qualquer conjunto convexo além dos raios. Podemos dizer, então, que um conjunto X possui uma direção se ele contém um raio. Observe também que um conjunto convexo X possui uma direção se, e somente se, ele não é limitado. A demonstração disso fica como exercício para o leitor[1].

Agora iremos nos aprofundar um pouco mais nas características algébricas de uma direção para um conjunto poliédrico $X = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$. Se temos que para todo $\mathbf{x} \in X$, $\mathbf{x} + \lambda\mathbf{d} \in X$, então é verdade que

$$\begin{aligned} A(\mathbf{x} + \lambda\mathbf{d}) &\leq \mathbf{b} \\ \mathbf{x} + \lambda\mathbf{d} &\geq 0 \end{aligned}$$

Porém, para que $\mathbf{d}$ seja uma direção de X, temos que ele deve satisfazer algumas condições algébricas a mais, que serão formalizadas no seguinte enunciado.

Proposição 2.3. Seja $X = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$ um conjunto poliédrico. O vetor $\mathbf{d}$ é uma diereção de X se, e somente se,

$$\mathbf{d} > 0, \quad A\mathbf{d} \leq 0$$

Demonstração. Com efeito, se $\mathbf{d}$ é uma direção de X, então

$$\begin{aligned} A(\mathbf{x} + \lambda\mathbf{d}) &\leq \mathbf{b} \\ A\mathbf{x} + A(\lambda\mathbf{d}) &\leq b \end{aligned}$$

mas como $A\mathbf{x} \leq \mathbf{b}$, então é preciso que

$$A\mathbf{d} \leq 0$$

Ademais, se $\mathbf{x} + \lambda\mathbf{d} \geq 0$, então

$$\mathbf{d} \geq \frac{1}{\lambda} - \mathbf{x} \tag{2.3}$$

[1]Tente argumentar o porquê de X possuir uma direção implicar em ele não poder ter uma bola o contendo. Para a recíproca, argumente o porquê de um conjunto ser contido por uma bola não poder ter uma direção na qual ele se estende indefinitivamente.

como $\lambda \geq 0$ pode ser arbitrariamente grande, então $\lim_{\lambda\to\infty} -\mathbf{x} = 0$, e com isso

$$d \geq 0$$

.

Por outro lado, se $\mathbf{d} > 0$ e $A\mathbf{d} \leq 0$, então para $\lambda \geq 0$ temos que

$$A(\lambda\mathbf{d}) \leq 0$$

e se $A\mathbf{x} \leq b$, então

$$\begin{aligned} A\mathbf{x} + A(\lambda\mathbf{d}) &\leq \mathbf{b} \\ A(\mathbf{x} + \lambda\mathbf{d}) &\leq \mathbf{b} \end{aligned}$$

e como $\mathbf{x} \in X$ é arbitrário, então d é uma direção de X. ■

Definição 2.11. Seja X um conjunto convexo e $\mathbf{d}$ uma direção desse conjunto. O vetor $\mathbf{d}$ é chamado de direção extrema se ele não pode ser dado como combinação linear positiva de outras direções desse conjunto.

A ideia de direções extremas é análoga a de pontos extremos. O vetor $\mathbf{d}$ é direção extrema se não existe outras duas direções $\mathbf{d}_1$ e $\mathbf{d}_2$ tal que

$$\begin{aligned} \lambda_1\mathbf{d}_1 + \lambda_2\mathbf{d}_2 &= \mathbf{d} \\ \lambda_1, \lambda_2 &\geq 0 \end{aligned}$$

Os raios cuja a direção é dada por uma direção extrema são chamados de raios extremos.

Os conceitos de raios e direções podem ser usados para definir os cones convexos. Seja $C \subset \mathbb{R}^n$ um cone definido como

$$C = \{\mathbf{x} \mid \lambda\mathbf{x}, \lambda \geq 0\}$$

Podemos perceber que o subconjunto de C dado pelos múltiplos escalares positivos de $\mathbf{x} \in C$ formam um raio com vértice na origem do $\mathbb{R}^n$ com direção dada por $\mathbf{x}$. Dessa forma, um cone pode ser definido a partir de suas direções, mas nem todas são necessárias para tal, já que podemos usar somente as suas direções extremas, que formam um conjunto minimal gerador das demais.

Portanto, se C é um cone e $D = \{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_n\}$ o conjunto das suas direções extremas, então o cone pode ser dado como

$$C = \{\mathbf{x} \mid \mathbf{x} = \sum_{j=1}^{n} \lambda_j\mathbf{d}_j, \lambda_j \geq 0\}$$

Porém iremos formalizar isso num resultado futuro.

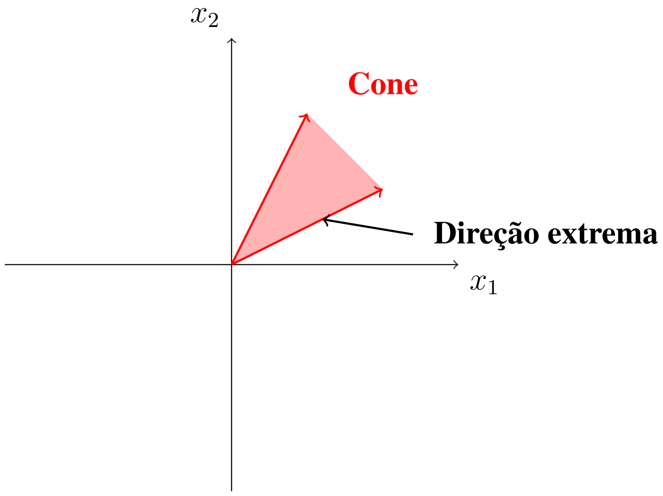


Figure 2.3: Representação de um Cone e seus raios extremos em $\mathbb{R}^2$

2.2.3 Faces e Arestas

Nesta seção apenas abordaremos algumas definições algébricas para conceitos historicamente geométricos. A primeira delas será a **face**.

Definição 2.12. Seja X um conjunto convexo e $F = \{\mathbf{x} \in X \mid B\mathbf{x} = \mathbf{c}\}$ em que $B\mathbf{x} = \mathbf{c}$ é o sistema linear cujas equações é um subconjunto não vazio dos hiperplanos definidores de X. O conjunto F é então chamado de face de X.

Em outras palavras uma face é um subconjunto de X cujos elementos são solução de um sistema linear dado por 1 ou mais hiperplanos de X. Em dimensão três, quando pensamos em faces de um poliedro, pensamos naquelas que tem dimensão dois, mas a nossa definição permite que outros elementos de dimensão menores, como as arestas e os vértices desse poliedro também sejam chamados de "faces".

Na verdade, em dimensão n podemos ter faces de qualquer dimensão entre 1 e n. As faces de dimensão $n-1$, como as faces propriamente ditas de um poliedro em dimensão 3, ou as arestas de um polígono bidimensional, são chamadas de **facetas**. Facetas também podem ser entendidas como a parte viável (que está em X) de um hiperplano definidor. As faces com dimensão entre $n-1$ e 0 são chamadas de faces próprias de X, enquanto aquela de dimensão n, que é o próprio X é chamada de face imprópria[2].

[2]Se nossa definição considerasse o conjunto vazio também como face de um conjunto convexo, então ele também seria uma face imprópria

Vamos agora dizer que $r(F)$ da face F de um conjunto convexo X é o número mínimo de hiperplanos necessários para definir F como o conjunto solução de um sistema linear. Esse número, para pontos extremos, que são faces de dimensão 0, é n, como vimos no teorema 2.1. Para uma aresta, que é um segmento de reta, e possui, pois, dimensão 1, esse número é $n - 1$. Podemos concluir que no geral a dimensão $\dim(F)$ de uma face F e o número $r(F)$ estão relacionadas por

$$r(F) = n - \dim(F)$$

Ainda sobre arestas, sabemos da geometria clássica que elas sempre conectam dois vértices, par esse chamado de adjacentes. Como vimos, pontos extremos equivalem aos vértices de um poliedro convexo n dimensional, e, analogamente, os pontos extremos que são conectados por uma aresta (o que nem sempre ocorre, visto que os poliedros podem ser ilimitados) são também chamados de **adjacentes**. O método simplex, ao caminhar pelos vértices da região viável em busca da solução ótima do Programa Linear, sempre o faz caminhando pelas arestas da região viável, indo de um vértice para outro que seja adjacente.

[section]

2.3 Envoltória Convexa

2.3.1 Envoltória e Independência Convexa

Definição 2.3.1. Seja P um conjunto de pontos em $\mathbb{R}^n$. Chamamos de envoltória convexa P, ou fecho convexo, o conjunto $\text{Conv}(P)$ dado por todas as combinações convexas dos pontos de P

Em outras referências, pode-se encontrar uma definição equivalente de que a envoltória convexa, popularmente conhecida como *covex hull*, é o menor conjunto convexo que contém todos os pontos de P. Neste texto, usaremos a definição destacada por ela ser mais algebricamente fecunda.

Definição 2.3.2. Seja $P \subset \mathbb{R}^n$ é um conjunto de pontos. P é convexo-independente se nenhum de seus pontos pode ser dado como combinação convexa de outros dois.

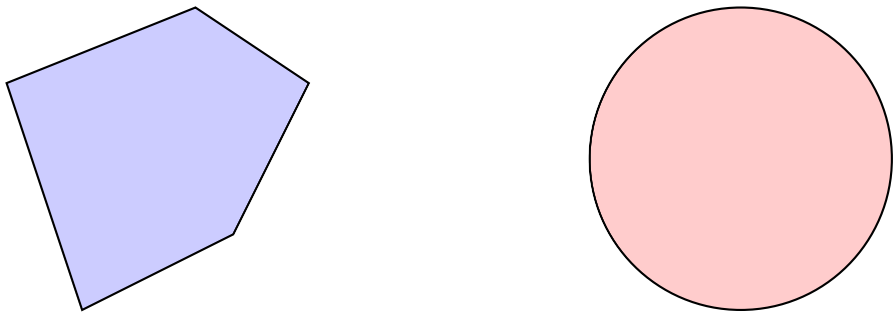
A definição acima especializa o conceito de independência linear para conjuntos convexos, que nos será uma definição útil. Dessa forma, um conjunto convexo-dependente será aquele no qual há pelo menos um ponto que pode ser dado como combinação convexa de outros dois. Outra especialização que faremos será do conceito de geradores: se um conjunto convexo X é a envoltória convexa de um conjunto P, então P é um conjunto gerador de X. Ademais, se P é convexo-independente, então os pontos de P são os pontos extremos de X, o que especializa o conceito de base da álgebra linear para os pontos extremos.

Na Teoria de PL, nosso interesse irá se concentrar nos conjuntos convexo que são finitamente gerados, isto é, que podem ser determinados por um conjunto finito de pontos geradores. Exemplo de conjuntos convexos finitamente gerados são os poliedros tridimensionais e os polígonos, enquanto que o círculo, ou a esfera, são exemplos de conjuntos convexos que não são finitamente gerados.

Proposição 2.3.3. Seja $P = \{p_1, \ldots, p_m\}$ um conjunto convexo dependente e $P' = \{p_1, \ldots, p_k\}$ um subconjunto convexo-independente de P com $k < m$. Então é correto dizer que

$$\text{Conv}(P') = \text{Conv}(P)$$

Dado um conjunto de geradores, é fácil observar que, ao obtermos um novo conjunto sem as redundâncias, então a envoltória convexa, como foi definida em 2.3.1, será exatamente a mesmas para ambos os conjuntos. Além disso, se P for um conjunto convexo independente, então seus elementos são os pontos extremos de $\text{Conv}(P)$, como fora observado mais previamente.



(a) Conjunto convexo finitamente gerado (b) Conjunto convexo não finitamente gerado

Figure 2.4: Comparação entre um polígono convexo e um círculo.

Uma classe especial de envoltórias convexas são os **simples** A ideia geométrica por detrás do conceito de simplex é do "mais simples conjunto convexo em um espaço euclideano de dimensão finita". Por exemplo: o menor número de pontos necessários para delimitar uma região no plano é 3, e a região delimitada por eles é um triângulo. Já em dimensão 3, para delimitar uma região no espaço são necessários ao menos quatro pontos, que definirão um tetraedro. Generalizando essa primitiva, obtemos a definição.

Definição 2.3.4. Chama-se simplex a envoltória convexa de um conjunto de $n+1$ pontos em $\mathbb{R}^n$ convexo-independente.

(a) Simplex em 2D (b) Simplex em 3D

Figure 2.5: Comparação entre simplex de dimensões diferentes.

Os simplex nos permitem a mais facilmente descrever a natureza geométrica por deterás de conjuntos convexos finitamente gerados.

2.3.2 Espaços Afim

Antes de ir mais afundo na natureza geométrica dos conjuntos poliédricos, iremos retornar um pouco a alguns conceitos algébricos. Relembrando, uma das noções primárias da Álgebra Linear é a de independência linear, donde dizemos que um conjunto de vetores é linearmente independente se nenhum vetor poder ser dado como combinação linear dos demais. Uma classe de combinações lineares que nos é últil são as chamadas combinações afim

Definição 1. Sejam $v_1, \ldots, v_k$ vetores do $\mathbb{R}$. u é uma combinação afim dos vetores citados se existem $\lambda_1, \ldots, \lambda_k$ tal que

$$u = \sum_{i=1}^{k} \lambda_i v_i \qquad \sum_{i=1}^{k} \lambda_i = 1$$

Portanto, combinações afim estão em um nível intermediário de restrições sobre os pesos entre as combinações lineares tradicionais e as combinações convexas.

As combinações lineares afim formam um conjunto fechado chamado de **subsespaço afim**, que são análogos aos hiperplanos lineares a menos da restrição de conterem a origem do $\mathbb{R}^n$.

Os hiperplanos afim, de forma análoga ao lineares, podem ser definidos a partir de um sistema linear na forma $Ax = b$ ou através de um conjunto de no mínimo n geradores. Um conjunto de vetores em que pelo menos um dos membros é combinação afim dos demais é chamado de afimmente dependente. Geometricamente, isso significa que pelo menos um dos vetores estará no espaço afim gerado pelos demais.

É sabido que qualquer conjunto com mais de n vetores em $\mathbb{R}^n$ é linearmente dependente, mas qual será o número mínimo de vetores em um conjunto para afirmamos com certeza que eles são afimmente dependentes?

A envoltória convexa de dois pontos é um segmento de reta, e, por sua vez, o subsespaço afim gerado por esses mesmos pontos é a a própria reta que eles definem, ou seja, obtemos o subespaço afim gerado ao extender infinitamente o segmento de reta em ambos os sentido. Já três pontos não colineares definem um triângulo, que é uma figura plana, e o subsespaço afim gerado por esses três pontos é justamnte o plano que contém esse triângulo; é como se tivessemos expandido infinitamente a área do triângulo assim como fizemos com o comprimento do segmento de reta anteriormente. Por conseguinte, para obte o *span* afim a partir de uma envoltória convexa, basta extendemos infinitamente a envoltória nas direções possíveis.

Isso nos permite agora responder a questão feita mais para trás: expandir infinitamente um simplex de dimensão n que é gerado por $n + 1$ vetores do $\mathbb{R}^n$, , é suficiente para obtermos todo o $\mathbb{R}^n$ e qualquer outro vetor do $\mathbb{R}^n$ pode ser dado como combinação afim dos pontos extremos do simplex que fora expandido. Portanto, a conclusão que chegamos é que qualquer conjunto com mais de $n + 1$ vetores é necessariamnte afimmente independente. Esse resultado é provado formalmente de forma mais algébrica no lema a seguir

Lema 2.3.1. Se $P = \{p_1, \dots, p_k\} \subset \mathbb{R}^n$ é um conjunto finito de pontos com $k > n+1$, então existem $\mu_1, \dots, \mu_n \in \mathbb{R}$ não todos nulos tal que

$$\sum_{i=1}^{k} \mu_i p_i = 0 \qquad \sum_{i=1}^{k} \mu_i = 0$$

e P é um conjunto afimmente dependente.

Demonstração. Da álgebra linear, sabemos que, se $\#P = k > n$, então P é um conjunto linearmente dependente, o que implica dizer que há $\mu_1, \dots, \mu_n \in \mathbb{R}$ não todos nulos tal que

$$\sum_{i=1}^{k} \mu_i p_i = 0$$

Além disso, observemos que o conjunto $\{p_2 - p_1, \dots, p_k - p_1\}$ também é linearmente dependente, portanto

$$\sum_{i=2}^{k} \mu_i (p_i - p_1) = 0$$

e fazendo $\mu_1 = -\sum_{i=2}^{k} \mu_i$ obtemos que

$$\sum_{i=1}^{k} \mu_i = 0$$

■

2.3.3 Simplex e o Teorema de Carathéodory

Antes de prosseguirmos, iremos demonstra um lema importante para um resultado fundamental da geometria convexa: o Teorema de Carathéodory. Enunciemos o lema

Teorema 2.3.1 (Carathéodory)**.** Seja $P \subset \mathbb{R}^n$ um conjunto finito de pontos. Se $x \in \mathrm{Conv}(P)$, então $x \in \mathrm{Conv}(P')$ para algum $P' \subset P$ com cardinalidade igual a $n + 1$.

Demonstração. Com efeito, se $x \in \mathrm{Conv}(P)$, e $\#P = k$, então existem $\lambda_1, \dots, \lambda_k \in \mathbb{R}$

com $\lambda_i \in [0, 1]$ tal que

$$\begin{aligned} x &= \sum_{i=1}^{k} \lambda_i p_i \\ \sum_{i=1}^{k} \mu_i &= 1 \end{aligned} \tag{2.4}$$

Se $k \leq n+1$, então nada há a demonstrar, do contrário, o lema 2.3.1 garante que existem $\mu_1, \dots, \mu_n \in \mathbb{R}$ não todos nulos tal que

$$\sum_{i=1}^{k} \mu_i p_i = 0$$

e como $\alpha \sum_{i=1}^{k} \mu_i p_i = 0$ para todo $\alpha \in \mathbb{R}$ então podemos rescrever a equação 2.4 como

$$x = \sum_{i=1}^{k} \lambda_i p_i - \alpha \sum_{i=1}^{k} \mu_i p_i$$

$$x = \sum_{i=1}^{k} (\lambda_i - \alpha \mu_i) p_i$$

donde temos que

$$\sum_{i=1}^{k} (\lambda_i - \alpha \mu_i) = 1$$

Agora iremos tentar escrever x como uma combinação convexa de até $n + 1$ pontos. Para atingir esse objetivo, escolhemos α tal que

$$\alpha = \min \left\{ \frac{\lambda_i}{\mu_i} \mid i \in \{1, \dots, k\} \text{ e } \mu_i > 0 \right\}$$

Logo, $\alpha > 0$, e para $i \in \{1, \dots, k\}$, temos dois casos.

I - $\mu_i \geq 0$

Temos que

$$\lambda_i - \alpha \mu_i = \mu_i \left(\frac{\lambda_i}{\mu_i} - \alpha \right)$$

e pela nossa escolha de α, então $\lambda_i - \alpha \mu_i \geq 0$

II - $\mu_i < 0$

Uma vez que $\alpha > 0$, então $\alpha\mu_i > 0$ e, portanto, $\lambda_i - \alpha\mu_i \geq 0$

digamos que $j*$ seja tal que $\alpha = \frac{\lambda_{j*}}{\mu_{j*}}$. Observe que $\lambda_{j^*} - \alpha\mu_{j^*}$, e podemos expressar x como

$$x = \sum_{i=1}^{j^*-1}(\lambda_i - \alpha\mu_i)p_i + \sum_{i=j^*+1}^{k}(\lambda_i - \alpha\mu_i)p_i$$

donde concluímos que x pode ser escrito como combinação convexa de $k-1$ pontos de $convP$

Como podemos repetir esse processo enquanto $k > n+1$, ou seja, enquanto o lema 2.3.1 pode ser aplicado, então x pode ser escrito como combinação convexa de até $n+1$ pontos de $\mathrm{Conv}(P)$. ■

Para capturar o significado geométrico do Teorema de Carathéodory, poderíamos enunciá-lo como "um ponto qualquer num fecho convexo sempre estará no interior de um simplex gerado por pontos desse fecho". Ademais, se $P \subset \mathbb{R}^n$ é um conjunto convexo-independente, então para qualquer $x \in \mathrm{Conv}(P)$, x pode ser dado como combinação convexa de até $n+1$ pontos de P, ou seja, de até $n+1$ pontos extremos. Formalizemos essa conclusão como corolário do Teorema de Carathéodory.

Corolário 2.3.1. Seja $P \subset \mathbb{R}^n$ um conjunto convexo-independente. Se $x \in \mathrm{Conv}(P)$ então x pode ser dado como combinação convexa de até $n+1$ pontos de P

Em outras palavras, o corolário afirma que um ponto qualquer num fecho convexo pode ser dado como combinação convexa de até $n+1$ pontos extremos.

Chapter 3

Método Simplex

Chapter 4

Solução Inicial e Convergência

Chapter 5

Dualidade

Chapter 6

Programação Linear Inteira