Introdução à Programação Linear

J. P. S. de Sousa

Contents

Chapter 1

Introdução

1.1 Um Breve Histórico da Programação Linear

A história dos métodos de otimização de forma mais geral pode ser tracejada pelo menos até a invenção do cálculo diferencial de forma independente por Isaac Newton e Gotfried Leibiniz entre os séculos XVII e XVIII, cujos métodos poderiam ser usados, apesar da pouca viabilidade, para otimizar o valor de uma função contínua. No decorrer do tempo, o estudo do cálculo e, consequentemente, dos métodos de otimização se expadiram com os trabalhos de outros matemáticos como Leonard Euler, Adrien-Marie Legendre e Friedrich Gauss. Diversos métodos foram desenvolvidos ao longo dos séculos por esses matemáticos, porém muitos com a propriedade de serem inviáveis de serem aplicados em larga escala, como os Multiplicadores de Lagrange.

Os problemas de programação matemática, e mais ainda os de programação linear, surgem naturalmente em uma série de aplicações, porém, durante e após a Segunda Guerra Mundial, a resolução desses problemas tornou-se crítica. A razão é que num conflito daquela escala problemas da natureza de "gerir e alocar recursos escassos de maneira eficiente" tornaram-se uma das principais preocupações dos militares da época.

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 significava escrever um programa de computador, estando na verdade atrelada ao sentido de "planejamento" de uma solução.

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 Notações e Terminologias

Hello World

1.3 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

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

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 $f(\mathbf{x})$ de **função objetivo**, as variáveis $x_1, \ldots, x_2$ de **variáveis de decisão** e o conjunto de inequações de **restrições**. A verdade é 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 $\mathbf{x} \in \mathbb{R}^n$ que são solução para o Problema de Programação Linear é chamado de **conjunto viável**, ou factível.

1.3.1 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

$$A\mathbf{x} \leq \mathbf{b}$$

tal que

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

Notação de desigualdades para matrizes pode parecer estranha, mas em contextos aplicados ela é muito comum ao descrever problemas como os quais estamos lidando. Nesta notação, temos que: se $\mathbf{x}$ e $\mathbf{y}$ são vetores com n componentes e $\mathbf{x} < \mathbf{y}$, então para todo $i = 1, \ldots, n$ é verdade que $x_i < y_i$. A notação é análoga para os demais tipos de desigualdades.

1.3.2 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

$$\mathbf{x} \geq 0$$

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.

1.3.3 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 temos 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: $7x_1 + 2x_2 + x_3 = 9$.

1.3.4 Problemas de Maximização e Minimização

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

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

1.3.5 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

$$\begin{aligned} \text{Maximizar:} \quad & z = \mathbf{c}^t\mathbf{x} \\ \text{sujeito a:} \quad & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

e quando de minimização, então sua forma canônica é

$$\begin{aligned} \text{Minimizar:} \quad & z = \mathbf{c}^t\mathbf{x} \\ \text{sujeito a:} \quad & A\mathbf{x} \geq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

1.4 Modelos de Programação Linear

Chapter 2

Teoria da Programação Linear

2.1 Conjuntos Convexos

2.1.1 Definição de Conjunto Convexo

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 X no $\mathbb{R}^n$ é chamado convexo se para quaisquer vetores $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{V}$ é verdade que

$$\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in \mathbb{V}$$

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

Aqueles com algum conhecimento de geometria analítica devem se lembrar que uma combinação linear de vetores na forma $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ 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 scales são positivos e sua soma é igual a 1. Então supondo um conjunto de pontos formado por $\mathbf{x}_1, \ldots, \mathbf{x}_n$ e escalares $\lambda_1, \ldots, \lambda_n$, então a combinação linear

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

é convexa se, e somente 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 e menores que 1.

2.1.2 Meio-espaços e Hiperplanos

Meio-espaços 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 meio-espaços.

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

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

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

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

ou por

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

Os primeiros são **meio-espaços fechados**, isto é, eles incluem sua fronteira, que consiste do hiperplano H. Por sua vez, os segundos são **meio-espaços abertos**, pois não incluem sua fronteira. Tanto hiperplanos quanto meio-espaços 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 meio-espaços fechados. Portanto, se R é a região factível do PPL cujo conjunto de restrições é dado por $A\mathbf{x} \leq \mathbf{b}$, por exemplo, então

$$X = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$$

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, denominamos como poliedros 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 dos conceitos de poliedro e politopo.

Definição 2.2. Um poliedro convexo é a interseção de um número finito de meio-espaços. Um politopo é um poliedro convexo limitado.

Para aqueles não familiarizados com o conceito de limitado, esse na matemática não é equivalente ao de finito. De modo geral, um conjunto limitado é aquele que pode ser contido dentro de uma região finita no mesmo espaço que o contém.

Das definições percebemos que poliedros convexos podem ser dados como um sistema e inequações ou equações lineares. Por conseguinte, o espaço das soluções viáveis 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 2.3. 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 2.4. Seja P um conjunto de pontos em $\mathbb{R}^n$. 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 2.5. 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 Conjuntos Poliédricos

Essa seção se dedicará a descrever conceitos importantes relacionados aos conjutnos poliédricos, tais como os pontos extremos, raios, direções e faces.

2.2.1 Pontos Extremos

O conceito de ponto extremo de um conjunto convexo X é uma especialização do conceito de "independência linear" para combinações lineares convexas.

Definição 2.6. Um vetor $\mathbf{x}$ num conjunto convexo X é chamado ponto extremo de X se ele não pode ser expresso como combinação linear convexa de nenhum outro subconjunto de X.

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 meio-espaço, 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 meio-espaço 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 meio-espaços e hiperplanos. Para viés de síntese, admita a seguinte definição

Definição 2.7. Seja X um conjunto convexo definidor pelos vetores $\mathbf{x}$ que satisfazem $A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{p}$. Se $\mathbf{a}_i^\intercal$ é uma linha da matriz A, então o hiperplano dado por $\mathbf{a}^\intercal\mathbf{x} = b_i$ é chamado de hiperplano definidor de X.

Em sequência, provaremos o teorema a seguir que caracteriza pontos extremos de um conjunto poliédrico.

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, e somente se, $\bar{\mathbf{x}}$ pertence a pelo menos n hiperplanos linearmente independentes definidores 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 $-\mathbf{d}$ espectivamente. 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.

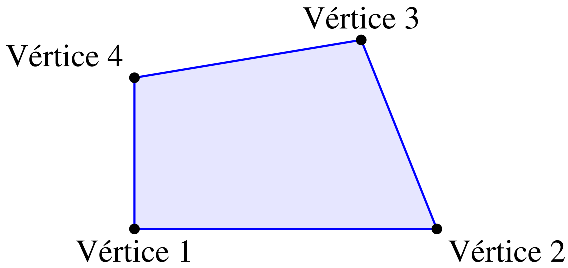


Figure 2.1: Interseção de hiperplanos em $\mathbb{R}^2$

- 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.

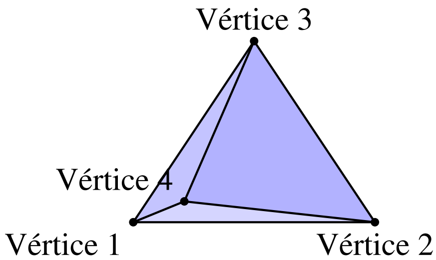


Figure 2.2: Interseção de Hiperplanos em $\mathbb{R}^3$

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, sendo definido 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 vértice do raio, e que se estende indefinitivamente 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 $X = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{p}\}$. Se temos que para todo $\mathbf{x} \in X$, $\mathbf{x} + \lambda\mathbf{d} \in X$, então é verdade que

$$A(\mathbf{x} + \lambda\mathbf{d}) \leq \mathbf{b}$$
$$\mathbf{x} + \lambda\mathbf{d} \geq \mathbf{p}$$

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

Proposição 2.3. Seja $X = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq p\}$ 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

$$A(\mathbf{x} + \lambda\mathbf{d}) \leq \mathbf{b}$$
$$A\mathbf{x} + A(\lambda\mathbf{d}) \leq \mathbf{b}$$

mas como $A\mathbf{x} \leq \mathbf{b}$ e $\lambda > 0$, então é preciso que

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

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

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

[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 $\lim_{\lambda\to\infty} \frac{1}{\lambda}(\mathbf{p}-\mathbf{x}) = 0$, e com isso

$$\mathbf{d} > 0$$

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 $A\mathbf{x} \leq \mathbf{b}$, 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 $\mathbf{d}$ 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.

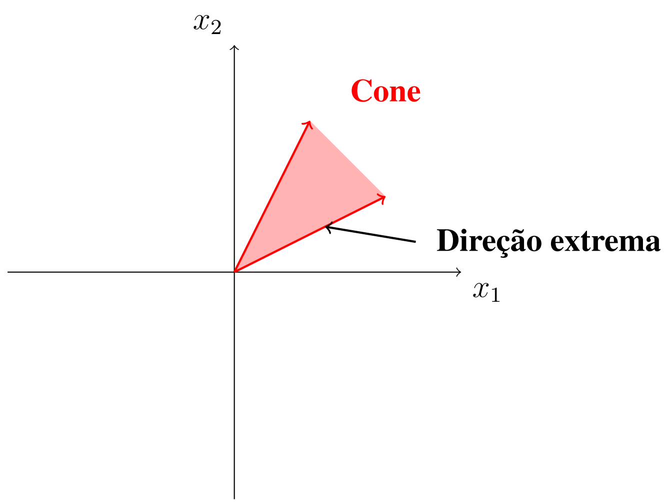


Figure 2.3: Representação de um Cone e seus raios extremos em $\mathbb{R}^2$

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 das equações 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 definidores 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 fa ce imprópria[2].

Definição 2.13. Seja X um conjunto poliédrico e F uma face desse conjunto Denotemos por $r(F)$ o número mínimo de hiperplanos definidores de X linearmente independentes necessários para deifnir F.

Se a face F do poliedro X é definida pelo sistema $B\mathbf{x} = \mathbf{y}$, então temos que $r(F)$ deve ser igual ao posto de B, já que esse é o número de hiperplanos LI que estão sendo usados para definir F.

Definição 2.14. Seja X um poliedro convexo em $\mathbb{R}^n$ definido por m hiperplanos linearmente independentes, e F uma face de X. A dimensão de F é dada por:

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

Como já observamos, $r(F)$ corresponde ao posto do sistema que define F, de modo que a dimensão da face, pelo Teorema Núcleo-Posto da Álgebra Linear, corresponde ao núcleo desse sistema. Por exemplo,

- São necessários ao menos n hiperplanos definidores para definir um ponto extremo, e, portanto, pontos extremos são faces de dimensão zero, uma vez que o sistema que os define tem posto cheio.

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

- São necessários dois pontos extremos para definir uma aresta, que é uma face de dimensão 1, pois são necessários $n - 1$ hiperplanos definidores LI para definir a reta que contém a aresta, e o sistema terá portanto núcleo com dimensão 1, como esperado.

O racioncínio pode ser generalizado para qualquer face de dimensão até a dimensão do prório poliedro convexo, que corresponde ao número máximo, limitado pela dimensão do espaço vetorial, de hiperplanos definidores linearmente independentes.

2.3 Envoltórias Convexas

Nesta seção, iremos agora nos concentrar numa outra forma de enxergar os conjuntos poliédricos que consiste em observá-los como o conjunto das combinações convexas de seus pontos extremos. Para isso, iremos investigar um pouco sobre os espaços afim e usar propriedades desses espaços para definir e demonstrar alguns resultados sobre essa nova forma de ver os poliedros, inclusive a sua equivalência com a definição inicialmente dada.

2.3.1 Espaços Afim

Antes de ir mais afundo na natureza geométrica dos conjuntos poliédricos, iremos recuar um pouco e observar alguns conceitos algébricos. Relembrando, uma das noções primárias da Álgebra Linear é a de independência linear, em que 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 2.15. Sejam $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vetores do $\mathbb{R}^n$. $\mathbf{u}$ é uma combinação afim dos vetores citados se existem $\lambda_1, \ldots, \lambda_k$ tal que

$$\mathbf{u} = \sum_{i=1}^{k} \lambda_i \mathbf{v}_i \qquad \sum_{i=1}^{k} \lambda_i = 1$$

Combinações afim são uma especialização das combinações lineares, visto a restrição de que a soma dos escalares que acompanham os vetores deve ser igual a 1. Observemos também que as combinaçãoes convexas, por sua vez, são uma especialização das combinações afim, em que adiciona-se a restrição de que os pesos devem ser positivos. As combinações lineares afim de um conjunto de vetores formam um conjunto fechado chamado de **subsespaço afim**

Os hiperplanos afins, de forma análoga ao lineares, podem ser definidos a partir de um sistema linear na forma $A\mathbf{x} = \mathbf{b}$. 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 afirmarmos com certeza que eles são afimmente dependentes? É isso que mostraremos no resultado a seguir.

Lema 2.1. Se $P = \{\mathbf{p}_1, \ldots, \mathbf{p}_k\} \subset \mathbb{R}^n$ é um conjunto finito de pontos com $k > n+1$, então existem $\mu_1, \ldots, \mu_n \in \mathbb{R}$ não todos nulos tal que

$$\sum_{i=1}^{k} \mu_i \mathbf{p}_i = 0 \qquad \sum_{i=1}^{k} \mu_i = 0$$

e P é um conjunto afimmente dependente.

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

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

Além disso, observemos que o conjunto $\{\mathbf{p}_2 - \mathbf{p}_1, \ldots, \mathbf{p}_k - \mathbf{p}_1\}$ também é linearmente dependente, pois contém mais de n vetores, já que $k > n+1$ por hipótese, e também que

$$\begin{gathered} \sum_{i=2}^{k} \mu_i (\mathbf{p}_i - \mathbf{p}_1) \\ \sum_{i=2}^{k} \mu_i \mathbf{p}_i - \mathbf{p}_1 \sum_{i=2}^{k} \mu_i \\ -\mu_1 \mathbf{p}_1 - \mathbf{p}_1 \sum_{i=2}^{k} \mu_i \end{gathered}$$

Dessa forma, se tivermos que

$$\mu_1 = -\sum_{i=2}^{k} \mu_i$$

então é verdade que

$$\sum_{i=2}^{k} \mu_i (\mathbf{p}_i - \mathbf{p}_1) = 0$$

e também

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

Resta mostrar que o conjunto P é afimmente dependente. Para isso, suponha que P esteja ordenado de forma que para todo $i \in I = \{1, \ldots, k\}$ é verdade que $\mu_1 \geq \mu_i$. Obeservamos então que

$$\mu_1 \mathbf{p}_1 + \mu_2 \mathbf{p}_2 + \ldots + \mu_k \mathbf{p}_k = 0$$
$$\mu_2 \mathbf{p}_2 + \ldots + \mu_k \mathbf{p}_k = -\mu_1 \mathbf{p}_1$$

Dividindo a equação acima toda por $-\mu_1$ otbemos

$$\sum_{i=2}^{k} -\frac{\mu_i}{\mu_1} \mathbf{p}_i = \mathbf{p}_1$$

Mas também

$$\sum_{i=2}^{k} \mu_i = -\mu_1$$
$$\sum_{i=2}^{k} -\frac{\mu_i}{\mu_1} = 1$$

Portanto $\mathbf{p}_1$ pode ser dado como combinação afim dos demais pontos de P, e o conjunto dos pontos é afimmente dependente. ■

2.3.2 Envoltória Convexa

Definição 2.16. Seja P um conjunto de pontos em $\mathbb{R}^n$. Chamamos de envoltória convexa de P, ou fecho convexo, o conjunto $\mathrm{Conv}(P)$ dado por todas as combinações convexas dos pontos de P

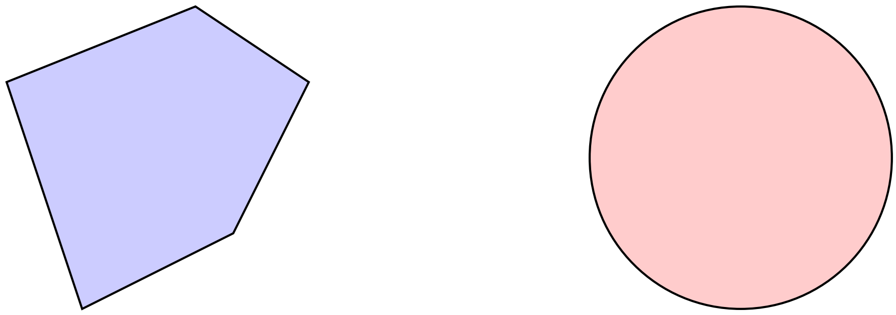
Em outras referências, pode-se encontrar uma definição equivalente de que a envoltória convexa, é o menor conjunto convexo que contém todos os pontos de P. Neste texto, usaremos a definição destacada por ela ser mais algebricamente fecunda.

Definição 2.17. 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.



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

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

Definição 2.18. Seja X um conjunto convexo em $\mathbb{R}^d$. Se P é o conjunto dos pontos de extremos de X, então P é chamado de perfil de X.

Agora provemos o resultado a seguir, que diz a envoltória convexa de um conjunto de pontos é igual a envoltória do seu perfil.

Proposição 2.4. Seja P um conjunto convexo e P' o perfil de P. Então é correto dizer que

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

Demonstração. Suponha que $\mathbf{p}_k \in P$ seja um combinação convexa dos demais pontos de P, o que implica, pela própria definição de envoltória convexa, que

$$\mathrm{Conv}(P) = \mathrm{Conv}(P \setminus \{\mathbf{p}_k\})$$

Logo, ao obter a envoltória de P menos os pontos não extremos, ou seja, a envoltória de seu perfil, o conjunto gerado será o mesmo que $\mathrm{Conv}(P)$. ■

Corolário 2.1. Se P' é perfil de P, então P' é perfil de $\mathrm{Conv}(P)$

Demonstração. Como $X = \text{Conv}(P) = \text{Conv}(P')$, então todos os pontos de X são gerados por pontos de P', que é convexo-independente, logo P' trata-se do perfil de X. ■

Uma classe especial de envoltórias convexas são os **simplex** A ideia geométrica por detrás do conceito de simplex é do "mais simples conjunto convexo num 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 2.19. Chama-se simplex a envoltória convexa de um conjunto de $d+1$ pontos em $\mathbb{R}^d$ afimmente independetes.

A razão por detrás da restrição de que os pontos devem ser afimmente independentes é porque eles não podem estar todos contidos num mesmo hiperplano afim, de modo que o conjunto convexo gerado por eles tenha dimensão igual ao do $\mathbb{R}^d$.

(a) Simplex em 2D

(b) Simplex em 3D

Figure 2.5: Comparação entre simplex de dimensões diferentes.

2.3.3 Teorema de Carathéodory

O Teorema de Carathéodory é um dos resultados mais fundamentais da geometria convexa, pois ele nos permite dizer que um ponto qualquer em um conjunto convexo pode ser gerado por um número finito de outros pontos nesse conjunto, mas não somente isso como também o limitante do tamanho desse conjunto de geradores.

Teorema 2.1 (Carathéodory)**.** Seja $P \subset \mathbb{R}^n$ um conjunto finito de pontos. Se $\mathbf{x} \in \text{Conv}(P)$, então $\mathbf{x} \in \text{Conv}(P')$ para algum $P' \subset P$ com cardinalidade igual a $d+1$.

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

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

$$\begin{aligned} \mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{p}_i \\ \sum_{i=1}^{k} \mu_i = 1 \end{aligned} \tag{2.4}$$

Provemos este resultado por indução em k

Base: trivial

Hipótse de Indução: se $\mathbf{x} \in P$, então existe $P' \subset P$ com $\#P = k - 1$ e $\#P' \leq d + 1$ tal que $\mathbf{x} \in P'$. Porvemos para o caso em que $\#P = k$.

Se $k \leq d + 1$, então nada há a demonstrar, do contrário, o lema 2.1 garante que existem $\mu_1, \ldots, \mu_k \in \mathbb{R}$ não todos nulos tal que

$$\sum_{i=1}^{k} \mu_i \mathbf{p}_i = 0 \qquad \sum_{i=1}^{k} \mu_i = 0$$

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

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{p}_i - \alpha \sum_{i=1}^{k} \mu_i \mathbf{p}_i$$

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

donde temos que

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

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

$$\alpha = \min_{i \in \{1,\ldots,k\}} \left\{ \frac{\lambda_i}{\mu_i} \mid \mu_i > 0 \right\}$$

Logo, $\alpha > 0$, e para $i \in \{1, \ldots, k\}$, 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 $\lambda_{j^*} - \alpha\mu_{j^*} = 0$, e podemos expressar $\mathbf{x}$ como

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

donde concluímos que $\mathbf{x}$ pode ser escrito como combinação convexa de $k-1$ pontos de $\mathrm{Conv}(P)$, caso para o qual vale a hipótese de indução. Logo, $\mathbf{x}$ pode ser escrito como combinação convexa de até $d+1$ pontos de $\mathrm{Conv}(P)$. ■

Pela proposição 2.4, temos que $\mathbf{x} \in \mathrm{Conv}(P)$ está na envoltória convexa do perfil de P. Portanto, $\mathbf{x}$ pode ser expresso como combinação convexa de até $d+1$ pontos extremos de P. Formalizemos essas conclusão no corolário a seguir.

Corolário 2.1. Seja $P \subset \mathbb{R}^d$ finito. Se $\mathbf{x} \in \mathrm{Conv}(P)$, então $\mathbf{x}$ pode ser dado como combinação convexa de até $d+1$ pontos extremos de P.

2.3.4 Topologia de Evoltórias Convexas

Nesta seção iremos analisar algumas propriedades topológicas básicas de interesse sobre a envoltória convexa de um conjunto de pontos. Mas não se preocupe, apesar de usarmos conceitos de topologia, iremos apresentá-los e usá-los de forma bem elementar a fim de não desviar muito dos nossos objetios e complicar em demasiado nossas construções, já que está fora do escopo desse texto as noções mais aprofundadas desse tema.

Proposição 2.5. Seja P um conjunto convexo. Se $\bar{\mathbf{x}}$ é um ponto extremo de P, então $\bar{\mathbf{x}}$ é um ponto na fronteira de P.

Demonstração. Para provarmos esse resultado, basta mostrarmos que para qualquer bola aberta B centrada em $\bar{\mathbf{x}}$, B intercepta pontos tanto de P quanto de $\mathbb{R}^d \setminus P$. Suponha por contradição que haja uma bola aberta centrada em $\bar{\mathbf{x}}$ tal que $B \subsetneq P$. Isso implica que $\bar{\mathbf{x}}$ está no interior de um segmento de reta inteiramente contido em P, e não é, pois, ponto extremo. ■

Antes de partirmos para o próximo resultado, elucidemos um fato. Suponha que X seja um conjunto convexo no $\mathbb{R}^d$ gerado por pontos de P. Para todo ponto $\mathbf{x}$ somente uma das afirmações pode ser verdadeira:

1. $\mathbf{x}$ está em X.
2. Existe um hiperplano com equação $\mathbf{a}^\intercal\mathbf{y} = b$ tal que

   $$\mathbf{a}^\intercal\mathbf{x} > b \text{ e } \mathbf{a}^\intercal\mathbf{q} < b$$

   para todo $\mathbf{q} \in X$

A segunda afirmação é mais forte e ela é consequência do Teorema do Hiperplano Separador, um importante resultado em geometria convexa, mas o qual não provaremos aqui. Dessa forma, temos duas opções, dado um ponto e um conjunto convexo, então ou o ponto está no conjunto, ou existe um hiperplano separando os dois. Com isso em mente, iremos agora provar que conjuntos convexos finitamente gerados são fechados.

Proposição 2.6. Se C é a envoltória convexa de P, então C é fechado.

Demonstração. O que faremos na verdade será mostrar que se $\mathbf{x}$ não é um ponto de X, então ele não poderia ser um ponto de fronteira.

Se $\mathbf{x} \notin X$, então existe um hiperplano H com equação $\mathbf{a}^\intercal\mathbf{y} = b$ que divide o $\mathbb{R}^d$ em dois meio-espaços abertos que separam $\mathbf{x}$ e X. Digamos que $\mathbf{x} \in H^+$ e que $X \subset H^-$. Supnha que $\mathbf{z}$ seja um ponto no hiperplano H tal que $||\mathbf{z} - \mathbf{x}||$ seja a mínima distância entre $\mathbf{x}$ e o hiperplano H.

$$\mathbf{z} = \underset{\mathbf{v}\in H}{\operatorname{argmin}}\{||\mathbf{v} - \mathbf{x}||\}$$

Se $\alpha = ||\mathbf{z} - \mathbf{x}||$, então a bola aberta centrada em $\mathbf{x}$ e com raio menor que α estará totalmente contida em H^+. Como meios-espaços são conjuntos convexos, e $X \subset H^-$, então existe uma bola aberta centrada em $\mathbf{x}$ que não contém nenhum ponto de X, e $\mathbf{x}$ não poderia ser um ponto de fronteira de X. ■

Proposição 2.7. Se $C = \mathrm{Conv}(P)$ é a envoltória convexa de um conjunto finito P de pontos, então C é limitado.

Demonstração. Com efeito, uma vez que P é finito, então deve have uma bola aberta que contém todos os seus pontos e, por conseguinte, a envoltória convexa desse conjunto, portanto $C = \mathrm{Conv}(P)$ é limitado. ■

Para fins de síntese, reunamos as proposiçõe anteriores no teorema a seguir

Teorema 2.2. Se $C = \mathrm{Conv}(P)$ é a envoltória convexa de um conjunto finito de pontos, então C é compacto.[3]

[3]Chama-se de compactos os conjuntos que são simultâneamente fechados e limitados.

2.4 Envoltória Cônica

Na seção anterior, estudamos sobre a envoltória convexa de um conjunto de pontos. Nesta, estudaremos o cone gerado a partir de um conjunto de direções em $\mathbb{R}^d$, chamado de envoltória cônica ou fecho cônico. Muitos dos resultados provados aqui serão análogos também àqueles da seção anterior.

2.4.1 Envoltória e Independência Cônica

Definição 2.20. Seja D um conjunto de direções em $\mathbb{R}^n$. A envoltória cônica de D, denotado por $\mathrm{Cone}(D)$, é o conjunto de todas as combinações cônicas dos elementos de D.

Definição 2.21. Um conjunto D é cônico-independente se nenhum de seus elementos poder ser dado como combinação cônica dos demais

A seguir, um fato semelhante ao levantado também na seção anterior sobre redundâncias no conjunto de geradores de um cone convexo.

Proposição 2.8. Se $D = \{d_1, \ldots, d_m\}$ é um conjunto de direções, e $D' = \{d_1, \ldots, d_k\}$ é um subconjunto conico-independente de D com $k < m$, então

$$\mathrm{Cone}(D') = \mathrm{Cone}(D)$$

Demonstração. O argumento é análogo ao mostrado na prova da proposição 2.4 ∎

2.4.2 Teorema de Carathéodory para Cones

Nesta parte, iremos demonstrar uma versão do Teorema de Carathéodory para os cones convexos

Teorema 2.3. Seja $D = \{\mathbf{d}_1, \ldots, \mathbf{d}_k\}$ um conjunto de direções em $\mathbb{R}^n$. Se $\mathbf{d} \in D$, então $\mathbf{d}$ pode ser dado como combinação cônica de até n direções de D

Demonstração. Vamos provar por indução em k.

Base: Para $k = 1$ é trivial.

Hipótese de Indução: Suponha que o teorema vale para todo k. Disso decorre dois casos

I - Se $k \leq n$, não há nada a ser demosntrado

II - Se $k > n$, então D é linearmente dependente, o que implica

$$\exists \mu_1, \ldots, \mu_k \in \mathbb{R}^n, \left(\sum_{i=1}^{k} \mu_i \mathbf{d}_i = 0 \right)$$

Seja $\theta \in (0, 1)$ tal que

$$\theta \sum_{i=1}^{k} \mu_i \mathbf{d}_i = 0$$

ao passo que

$$\mathbf{d} = \sum_{i=1}^{k} \lambda_i \mathbf{d}_i$$

o que implica

$$\mathbf{d} = \sum_{i=1}^{k} (\lambda_i - \theta\mu_i) \mathbf{d}_i$$

A estratégia aqui será a mesma daquela usada para provar o Teorema de Carathéodory: escolha θ como o menor $\frac{\lambda_j}{\mu_j}$ tal que $j \in J$ onde J é o conjunto dos índices de 1 até k para os quais $\mu_k > 0$. Dessa forma, $\theta > 0$, e se seguem dois casos:

I - $\mu_i < 0$

Então $-\theta\mu_i > 0 \Rightarrow \lambda_i - \theta\mu_i > 0$

II - $\mu_i > 0$

Então $-\theta\mu_i < 0$, mas uma vez que

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

e pela nossa escolha de θ, teremos que $\lambda_i - \theta\mu_i > 0$.

Se j^* for o índice tal que $\theta = \frac{\lambda_{j^*}}{\mu_{j^*}}$, então para quando $i = j^*$ temos $\lambda_i - \theta\mu_i = 0$, e

$$\mathbf{x} = \sum_{i=1}^{j^*-1} (\lambda_i - \theta\mu_i)\mathbf{d}_i + \sum_{i=j^*+1}^{k} (\lambda_i - \theta\mu_i)\mathbf{d}_i$$

Logo $\mathbf{d}$ pode ser escrito como combinação cônica de $k - 1$ direções e, pela hipotese de indução, pode ser então escrito como combinação de até n direções. Portanto, pelo Princípio de Indução, o teorema vale para qualquer conjunto de direções. ■

Corolário 2.2. Seja D um conjunto de direções em $\mathbb{R}^n$. Se $\mathbf{x} \in \mathrm{Cone}(D)$, então $\mathbf{x}$ pode ser dado como combinação cônica de até no máximo n direções extremas de $\mathrm{Cone}(D)$.

O corolário anterior é ainda mais interessante no caso em que o cone convexo é finitamente gerado. Se D é finito, então qualquer direção de $\mathrm{Cone}(D)$ pode ser gerada por até n direções extremas de $\mathrm{Cone}(D)$, o que é equivalente a dizer que pode ser gerado por até n direções de D. Se D for cônico-independente, então ele é o conjunto das direções extremas de $\mathrm{Cone}(D)$.

2.5 Equivalência de Representações de Poliedros

Esta seção se dedicará a provar a equivalência entre conjuntos poliédricos e conjuntos convexos dados como fecho de um conjunto de pontos. Para isso, iremos apresentar o conceito de projeção de um poliedro e um algoritmo capaz de obter essa projeção.

2.5.1 Projeção de Poliedros e Eliminição de Fourier-Motzkin

Definição 2.22. Seja $X = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \mid A\mathbf{x} + B\mathbf{y} \leq \mathbf{b}\}$ um conjunto poliédrico. Chamamos P de projeção de X nas variáveis de $\mathbf{y}$, se para todo $\mathbf{y}$ tamos que

$$\mathbf{y} \in P \Leftrightarrow \exists \mathbf{x} \in \mathbb{R}^{d_1}, ((\mathbf{x}, \mathbf{y}) \in X)$$

A definição 2.22 é a geralmente dada para caracterizar projeções de poliedros. Porém, apresentaremos aqui uma forma alternativa de definir esse conceito que nos será mais fácil de aplicar mais adiante. Antes de a apresentarmos, introduzemos o conceito de uma **matriz de seleção de variáveis**.

Suponha que temos um vetor $\mathbf{x} = (x_1, x_2, x_3)$, e queremos selecionar a suas duas últimas componentes. Para isso, podemos aplicar a seguinte matriz

$$\pi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

de modo que $\pi\mathbf{x} = (x_2, x_3)$. Logo, matrizes de seleção representam operadores lineares capazes de realizar uma projeção de um vetor em suas componentes. Repare que não estamos utilizando necessariamente o nome "matriz de projeção", pois nas referências de Álgebra Linear, esse norme destina-se aos operadores que realizam a projeção ortogonal de um vetor num subespaço de dimensão menor.

Formalmente, uma selção é um operador linear $T : \mathbb{R}^d \to \mathbb{R}^n$, com $d > n$ tal que, se $I = \{1, \dots, d\}$ são os índices das componentes de $\mathbf{x}$ e $J \subset I$ os índices das componentes selecionadas por T, então a matriz π de T é dada como

$$\pi_{i,j} = \begin{cases} 1, & \text{se } i = j \text{ e } j \in J \\ 0, & \text{caso constrário} \end{cases}$$

Repare que π é a matriz identidade do $\mathbb{R}^d$ com apenas as linhas cujos índices estão J. Observe também que estamos somente considerando projeções que preservem a ordem das componentes do vetor original. Dessa forma, daremos uma nova definição para projeções de poliedros.

Definição 2.23. Seja $X = \{\mathbf{x} \in \mathbb{R}^d \mid A\mathbf{x} \leq \mathbf{b}\}$ um conjunto poliédrico e $\Pi : \mathbb{R}^d \to \mathbb{R}^n$ um operador de seleção. Chamos P de projeção de X nas variáveis selecionadas por Π se para todo $y \in \mathbb{R}^n$ é satisfeito que

$$\mathbf{y} \in P \iff \exists \mathbf{x} \in X, (\Pi(\mathbf{x}) = \mathbf{y})$$

ou seja, P é a imagem de Π aplicado a X.

Uma característica de poliedros é que a suas "sombras" são também poliedros, ou seja, se tomarmos um poliedro tridimensional e o projetarmos em uma superfície bidimensional, essa "sombra" terá a forma de um polígono. O que faremos agora será provar esse fato, isto é, que projeções de poliedros são poliedros.

Para cumprir nossa tarefa, iremos utilizar um clássico algoritmo de resolução de sistemas de iniquações linares análogo à Eliminação Gaussiana: o Algoritmo de Eliminação de Fourier-Motzkin, criado pelo famoso matemático francês Jean-Baptiste Fourier e redescoberto pelo matemático israelo-americano Theodore Motzkin. Esse algoritmo recebe um sistema de inequações lineares e elimina uma de suas variáveis a fim de facilitar a verificação de existência de uma solução ou se o sistema é limitado ou não. O algortimo 1 apresenta um passo para obter um novo sistema $B\mathbf{y} \leq \mathbf{c}$ a partir de $A\mathbf{x} \leq \mathbf{b}$ com uma variável eliminada x_k de $\mathbf{x}$ eliminada, isto é, a k-ésima coluna de C zerada.

2.5.2 Teorema da Equivalência

Algoritmo 1: Passo de eliminação da variável x_k no método de Fourier–Motzkin

Entrada: Um sistema $A\mathbf{x} \leq \mathbf{b}$ e uma variável x_k de $\mathbf{x}$

Saída: Um sistema $B\mathbf{y} \leq \mathbf{c}$ com a variável x_k eliminada

início
- // Classificação das inequações segundo o sinal de $a_{i,k}$
- $I_+ \leftarrow \{\, i \mid a_{i,k} > 0 \,\}$;
- $I_- \leftarrow \{\, i \mid a_{i,k} < 0 \,\}$;
- $I_0 \leftarrow \{\, i \mid a_{i,k} = 0 \,\}$;
- // Inicializa o novo sistema $B\mathbf{y} \leq \mathbf{c}$
- $B \leftarrow [\,]$; $\quad \mathbf{c} \leftarrow [\,]$;
- // Combina pares de restrições para eliminar x_k
- **para cada** $(i, j) \in I_+ \times I_-$ **faça**
  - // Calcula combinação linear das linhas i e j
  - // de A para eliminar x_k
  - $\alpha \leftarrow \frac{1}{a_{i,k}}$; $\quad \beta \leftarrow -\frac{1}{a_{j,k}}$;
  - $\mathbf{d} \leftarrow \alpha\mathbf{a}_i^\intercal + \beta\mathbf{a}_j^\intercal$;
  - $c \leftarrow \alpha b_i + \beta b_j$;
  - adicionar linha $(\mathbf{d},\, c)$ a $(B, \mathbf{c})$;
- **fim**
- // Mantém as inequações que não dependem de x_k
- **para cada** $i \in I_0$ **faça**
  - adicionar linha $(\mathbf{a}_i^\intercal,\, b_i)$ a $(B, \mathbf{c})$;
- **fim**
- **retorna** $B\mathbf{y} \leq \mathbf{c}$;

fim

Chapter 3

Método Simplex

Chapter 4

Solução Inicial e Convergência

Chapter 5

Dualidade

Chapter 6

Programação Linear Inteira