

Capítulo 2

Álgebra Linear, Análise Convexa e Conjuntos Convexos Poliedrais

2.1. Introdução

Neste Capítulo, apresenta-se dois tópicos altamente relevantes e relacionados com a teoria e aplicação de resolução de problemas de programação linear. Inicialmente, apresenta-se uma revisão de tópicos de álgebra linear básica. Dentre esses tópicos, tem especial importância as informações relativas à resolução de um sistema linear de equações algébricas. Entretanto, o tópico mais relevante está relacionado com a teoria de conjuntos convexos. As restrições de um problema de programação linear podem ser caracterizadas como sendo um conjunto convexo. Na verdade podem ser caracterizadas como um tipo especial de conjunto convexo, isto é, um conjunto convexo poliedral. Portanto, conceitos relacionados com conjuntos convexos tais como subespaço, raio, hiperplano, ponto extremo de um conjunto convexo e direção extrema de um conjunto convexo são muito importantes para entender a metodologia de resolução de problemas de programação linear.

Adicionalmente, deve-se observar que a caracterização de pontos extremos e direções extremas e, especialmente, a teoria e aplicação relacionadas com o Teorema Geral da Representação (TGR) de um conjunto convexo representam os tópicos mais relevantes apresentados neste Capítulo. Esses tópicos são usados no próximo Capítulo para entender o método simplex para resolução de problemas de programação linear. Finalmente, deve-se observar que os conceitos de ponto extremo e de direção extrema apresentados neste Capítulo podem ser interpretados como relacionados a uma visão geométrica do problema de programação linear e esses conceitos são analisados de forma independente da visão algébrica apresentada no próximo Capítulo. Além disso, ambos conceitos são equivalentes, embora em muitas aplicações, um tipo de visão geométrica ou algébrica possa ser mais adequada para interpretar problemas e provar teoremas relacionados com a teoria e aplicação de problemas de programação linear.

2.2. Revisão de Tópicos de Álgebra Linear

2.2.1. Vetores

Um vetor é um arranjo de n números arranjados em linha ou coluna.

Exemplos de vetores são os seguintes:

$$a = \begin{bmatrix} 1 & 3 & 6 \end{bmatrix} \qquad b = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Vetores Especiais:

O chamado vetor i -ésimo unitário é um vetor especial em que todos seus elementos são iguais a zero, exceto a posição i , que vale 1. Assim, esse vetor assume a seguinte forma:

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i$$

Combinação Linear de Vetores no Espaço E^n :

Um vetor b de dimensão n é uma combinação linear dos vetores $a_1, a_2, \dots, a_k$ em E^n se satisfaz o seguinte:

$$b = \sum_{j=1}^k \lambda_j a_j \tag{2.1}$$

onde os λ_j são números reais. Se, adicionalmente, a seguinte relação for verdadeira:

$$\sum_{j=1}^k \lambda_j = 1 \tag{2.2}$$

então o vetor b é uma combinação afim dos vetores $a_1, a_2, \dots, a_k$. Adicionalmente, se $\lambda_j \geq 0$, então a combinação é convexa. Nesse contexto, E^n é chamado de espaço euclidiano de dimensão n e constituído por todos os vetores de dimensão n .

Independência Linear de Vetores no Espaço E^n :

Os vetores de dimensão n , $a_1, a_2, \dots, a_k$ são chamados de linearmente independentes (LI) se cumprem a seguinte relação:

$$\sum_{j=1}^k \lambda_j a_j = 0 \implies \forall j = 1, 2, \dots, k. \tag{2.3}$$

Adicionalmente, se os vetores não são linearmente independentes, então eles são linearmente dependentes.

Exemplo 2.1: Mostrar que os seguintes vetores são linearmente independentes:

$$a_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad a_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Se a_1 e a_2 são linearmente independentes então temos o seguinte:

$$\lambda_1 a_1 + \lambda_2 a_2 = 0 \implies \begin{bmatrix} 1 \\ 2 \end{bmatrix} \lambda_1 + \begin{bmatrix} -1 \\ 1 \end{bmatrix} \lambda_2 = 0 \implies \begin{cases} \lambda_1 - \lambda_2 = 0 \\ 2\lambda_1 + \lambda_2 = 0 \end{cases} \implies \begin{cases} \lambda_1 = \lambda_2 \\ 3\lambda_1 = 0 \end{cases} \implies \begin{cases} \lambda_1 = 0 \\ \lambda_2 = 0 \end{cases}$$

Portanto, como foi provado que $\lambda_1 = \lambda_2 = 0$, então a_1 e a_2 são linearmente independentes.

Exemplo 2.2: Mostrar que os seguintes vetores são linearmente dependentes:

$$a_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad a_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \qquad a_3 = \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}$$

Inicialmente, tentamos provar que os vetores a_1 , a_2 e a_3 são linearmente independentes. Assim temos o seguinte:

$$\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 = 0 \implies \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \lambda_1 + \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \lambda_2 + \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} \lambda_3 = 0$$

$$\begin{cases} \lambda_1 - \lambda_2 = 0 \\ 2\lambda_1 + \lambda_2 + 3\lambda_3 = 0 \\ 3\lambda_1 - \lambda_2 + 2\lambda_3 = 0 \end{cases} \implies \begin{cases} \lambda_1 = \lambda_2 \\ 3\lambda_2 + 3\lambda_3 = 0 \\ 2\lambda_2 + 2\lambda_3 = 0 \end{cases} \implies \begin{cases} \lambda_1 = \lambda_2 \\ \lambda_2 + \lambda_3 = 0 \\ \lambda_2 + \lambda_3 = 0 \end{cases} \implies \begin{cases} \lambda_1 = \lambda_2 \\ \lambda_2 = -\lambda_3 \end{cases}$$

$$\lambda_1 = \lambda_2 = \alpha \qquad \lambda_3 = -\alpha$$

para qualquer α real. Assim, por exemplo, para $\alpha = 1$ temos $\lambda_1 = \lambda_2 = 1$ e $\lambda_3 = -1$ e diferentes de zero. Portanto, concluímos que os vetores a_1 , a_2 e a_3 não são linearmente independentes, isto é, são linearmente dependentes.

Vetores que Geram o Espaço Euclidiano E^n

Um conjunto de vetores $a_1, a_2, \dots, a_k$ em E^n gera o espaço E^n se qualquer vetor em E^n pode ser representado como uma combinação linear dos vetores $a_1, a_2, \dots, a_k$. Assim, dado um vetor $b \in E^n$, então devem existir escalares $\lambda_1, \lambda_2, \dots, \lambda_k$ de forma que:

$$b = \sum_{j=1}^k \lambda_j a_j \quad (2.4)$$

Exemplo 2.3: Mostrar que os seguintes vetores geram o espaço E^2 :

$$a_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad a_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad a_3 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

De (2.4) temos o seguinte:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \lambda_1 + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \lambda_2 + \begin{bmatrix} -1 \\ 3 \end{bmatrix} \lambda_3 \implies \begin{cases} b_1 = \lambda_1 + \lambda_2 - \lambda_3 \\ b_2 = \lambda_2 + 3\lambda_3 \end{cases}$$

Fazemos $\lambda_3 = \alpha$ e temos o seguinte:

$$\lambda_2 = b_2 - 3\alpha$$

$$\lambda_1 = b_1 - \lambda_2 + \lambda_3 = b_1 - (b_2 - 3\alpha) + \alpha = b_1 - b_2 + 4\alpha$$

Portanto, uma solução geral assume a seguinte forma:

$$\lambda_1 = b_1 - b_2 + 4\alpha \quad \lambda_2 = b_2 - 3\alpha \quad \lambda_3 = \alpha$$

para qualquer valor de α real. Assim, duas soluções particulares são as seguintes:

1. Quando $\alpha = 0$:

$$\lambda_1 = b_1 - b_2 \quad \lambda_2 = b_2 \quad \lambda_3 = 0$$

2. Quando $\alpha = b_2$:

$$\lambda_1 = b_1 + 3b_2 \quad \lambda_2 = -2b_2 \quad \lambda_3 = b_2$$

Exemplo 2.4: No exemplo anterior, mostrar que os vetores a_1 , a_2 e a_3 geram o seguinte vetor:

$$b = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

Existem muitas formas de encontrar a representação solicitada, uma representação para cada valor de α escolhido. Assim, escolhemos arbitrariamente um $\alpha = 2$. Substituindo este valor na solução geral do exemplo anterior temos o seguinte:

$$\lambda_1 = 4 - 8 + 4(2) = 4 \quad \lambda_2 = 8 - 3(2) = 2 \quad \lambda_3 = 2$$

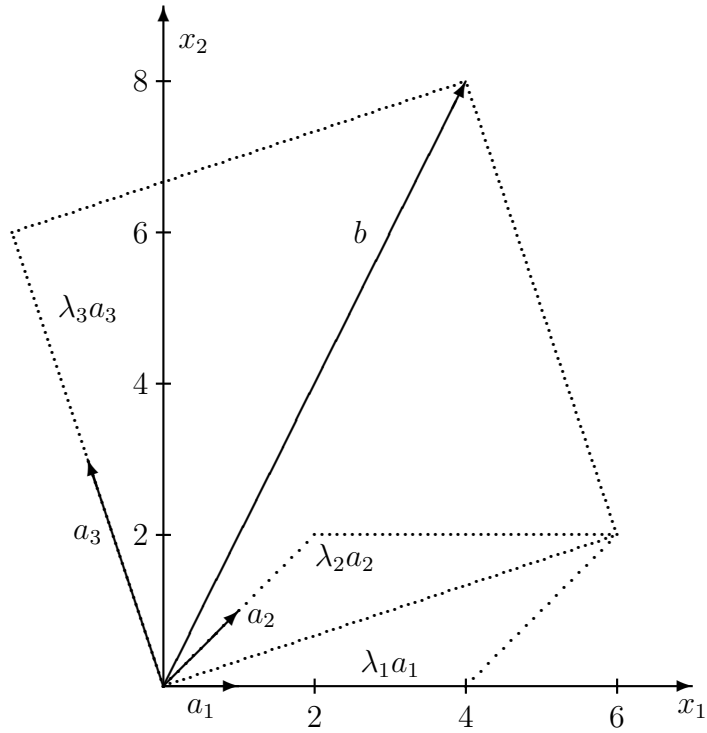


Figura 2.1: Os vetores a_1 , a_2 e a_3 geram o vetor b .

Assim, foi encontrada a seguinte representação:

$$\begin{bmatrix} 4 \\ 8 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

cuja representação gráfica é mostrada na Figura 2.1.

Vetores que Formam uma Base no Espaço Euclidiano E^n

Um conjunto de vetores $a_1, a_2, \dots, a_k$ em E^n forma uma base no espaço E^n se satisfazem as seguintes condições:

1. Se $a_1, a_2, \dots, a_k$ gera o espaço E^n .
2. Se qualquer um dos vetores for removido, então o conjunto restante de vetores não gera o espaço E^n .

Forma Equivalente de Definir uma Base no Espaço Euclidiano E^n

Um conjunto de vetores $a_1, a_2, \dots, a_k$ em E^n forma uma base no espaço E^n se $k = n$ e esses vetores são linearmente independentes.

Propriedade Importante Relacionada com uma Base no Espaço Euclidiano E^n

Dado um vetor $b \in E^n$, então uma base de E^n representada pelos vetores $a_1, a_2, \dots, a_n$ representa o vetor b de maneira única. Assim, a dimensão da base é única, formada por n vetores no espaço E^n , mas o conjunto de vetores que formam uma base não é único.

Exemplo 2.5: Mostrar que os seguintes vetores:

$$a_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad a_3 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

que formam uma base em E^2 , representam de maneira única o vetor:

$$b = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

De (2.4) temos o seguinte:

$$\begin{bmatrix} 4 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \lambda_1 + \begin{bmatrix} -1 \\ 3 \end{bmatrix} \lambda_3 \Rightarrow \begin{cases} \lambda_1 - \lambda_3 = 4 \\ 3\lambda_3 = 8 \end{cases} \Rightarrow \begin{cases} \lambda_1 = \frac{20}{3} \\ \lambda_3 = \frac{8}{3} \end{cases}$$

Portanto, existe uma única combinação possível para representar o vetor b combinando os vetores a_1 e a_3 . Na Figura 2.2 é mostrada essa representação de forma gráfica.

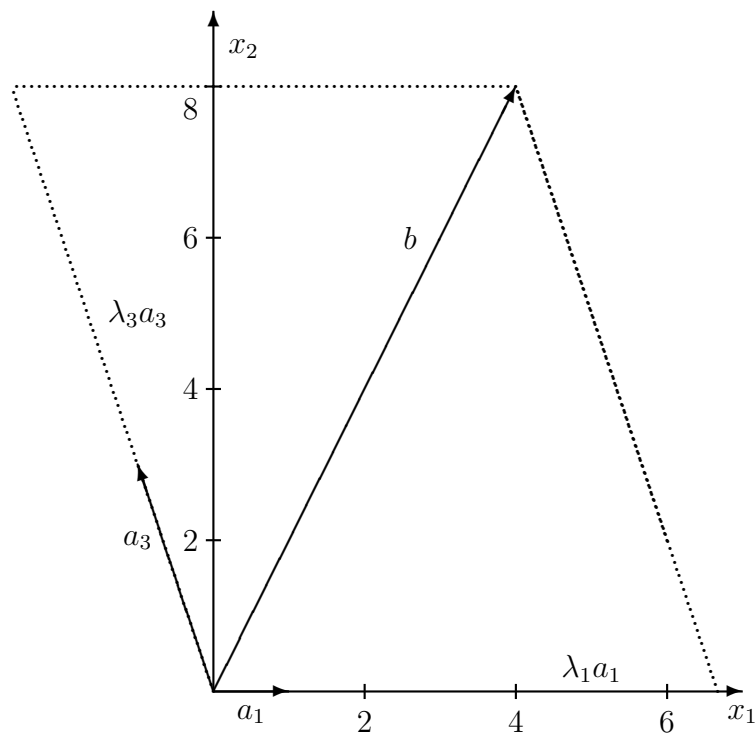


Figura 2.2: O vetor b é representado de forma única pelos vetores a_1 e a_3 .

2.3. Matrizes

Uma matriz é um arranjo retangular de números. Assim, por exemplo, uma matriz A assume a seguinte forma:

$$A = \begin{bmatrix} 1 & -4 \\ 6 & 8 \end{bmatrix}$$

Nesse contexto, é comum representar uma matriz por linhas e colunas. Assim, a matriz A mostrada anteriormente pode ser representada por colunas da seguinte forma:

$$A = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \implies a_1 = \begin{bmatrix} 1 \\ 6 \end{bmatrix} \quad \text{e} \quad a_2 = \begin{bmatrix} -4 \\ 8 \end{bmatrix}$$

A mesma matriz pode ser representada por linhas da seguinte forma:

$$A = \begin{bmatrix} a^1 & a^2 \end{bmatrix} \implies a^1 = \begin{bmatrix} 1 & -4 \end{bmatrix} \quad \text{e} \quad a^2 = \begin{bmatrix} 6 & 8 \end{bmatrix}$$

2.3.1. Solução de um Sistema Linear de Equações Usando Operações Elementares de Matrizes

Em uma matriz podem ser realizadas operações elementares de matrizes por linhas ou colunas. As operações elementares por linhas são as seguintes:

1. A linha i e a linha j podem ser trocadas.
2. A linha i pode ser multiplicada por um escalar $k \neq 0$.
3. A linha i pode ser substituída pela linha i adicionada de k vezes a linha j .

Essas propriedades de operações elementares em uma matriz são usadas para mostrar a forma de resolver um sistema algébrico linear da forma $Ax = b$ formado por m equações e n variáveis. Assim, a matriz A é de dimensão $m \times n$, b é um vetor de dimensão m e x é um vetor de dimensão n .

Propriedade Fundamental para Resolver um Sistema Algébrico Linear

Supor que a matriz $[A, b]$ foi transformada para a matriz $[A', b']$ usando apenas um número finito de operações elementares de matrizes. Nesse contexto, a seguinte relação é verdadeira:

$$Ax = b \iff A'x = b'$$

Em outras palavras, uma solução x que resolve um sistema também é solução do outro sistema algébrico linear. Essa propriedade fundamental é usada como base para justificar as várias formas de resolver um sistema algébrico linear. Portanto, usamos essa propriedade para mostrar os métodos de Gauss-Jordan e de Gauss.

Método de Gauss-Jordan

Neste método a estratégia fundamental consiste em transformar a matriz A até assumir a forma da matriz identidade, isto é, $A' = I$. Nesse caso a solução x é encontrada de forma trivial, isto é, $I x = b' \Rightarrow x = b'$.

Exemplo 2.6: Use o método de Gauss-Jordan para resolver o seguinte sistema:

$$\left. \begin{array}{l} x_1 + 2x_2 + x_3 = 10 \\ -x_1 + 2x_2 - x_3 = 6 \\ x_2 + x_3 = 2 \end{array} \right\} \Rightarrow \begin{bmatrix} 1 & 2 & 1 \\ -1 & 2 & -1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 6 \\ 2 \end{bmatrix}$$

Realizando operações elementares por linha das matrizes encontramos o seguinte:

$$\left[\begin{array}{ccc|c} A & : & b \end{array} \right] \Rightarrow \left[\begin{array}{cccc|c} 1 & 2 & 1 & : & 10 \\ -1 & 2 & -1 & : & 6 \\ 0 & 1 & 1 & : & 2 \end{array} \right] \Rightarrow \left[\begin{array}{cccc|c} 1 & 2 & 1 & : & 10 \\ 0 & 4 & 0 & : & 16 \\ 0 & 1 & 1 & : & 2 \end{array} \right]$$

$$\left[\begin{array}{cccc|c} 1 & 2 & 1 & : & 10 \\ 0 & 1 & 0 & : & 4 \\ 0 & 0 & 1 & : & -2 \end{array} \right] \Rightarrow \left[\begin{array}{cccc|c} 1 & 2 & 0 & : & 12 \\ 0 & 1 & 0 & : & 4 \\ 0 & 0 & 1 & : & -2 \end{array} \right] \Rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & : & 4 \\ 0 & 1 & 0 & : & 4 \\ 0 & 0 & 1 & : & -2 \end{array} \right]$$

Portanto, a solução do sistema algébrico linear é a seguinte:

$$A' x = b' \Rightarrow I x = b' \Rightarrow x = b'$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ -2 \end{bmatrix} \Rightarrow x_1 = 4 \quad x_2 = 4 \quad x_3 = -2$$

Método de Eliminação de Gauss

No método de Gauss, o processo de transformação da matriz A termina quando a matriz A' se transforma em uma matriz triangular superior. Finalmente, para encontrar os valores das variáveis x realizamos um processo de substituição a partir do valor do último elemento de x .

Exemplo 2.7: Use o método de Gauss para resolver o sistema do exemplo anterior.

Neste caso temos o seguinte sistema:

$$\begin{bmatrix} 1 & 2 & 1 \\ -1 & 2 & -1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 6 \\ 2 \end{bmatrix}$$

Realizando operações elementares por linha das matrizes encontramos o seguinte:

$$\left[\begin{array}{ccc|c} A & : & b \end{array} \right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 1 & 10 \\ -1 & 2 & -1 & 6 \\ 0 & 1 & 1 & 2 \end{array} \right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 1 & 10 \\ 0 & 4 & 0 & 16 \\ 0 & 1 & 1 & 2 \end{array} \right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 1 & 10 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -2 \end{array} \right]$$

Assim, a matriz A' assume uma forma triangular superior e temos o seguinte sistema equivalente:

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 4 \\ -2 \end{bmatrix}$$

Portanto, o sistema algébrico anterior pode ser facilmente resolvido por substituição reversa a partir da solução de x_3 da seguinte forma:

$$x_3 = -2 \quad x_2 = 4 \Rightarrow x_1 + 2x_2 + x_3 = 10 \Rightarrow x_1 + 2(4) + (-2) = 10 \Rightarrow x_1 = 4$$

2.3.2. Inversa de uma Matriz

Se A é uma matriz quadrada de dimensão $n \times n$, então sua inversa A^{-1} deve satisfazer a seguinte relação:

$$A A^{-1} = A^{-1} A = I$$

A inversa de uma matriz pode ser encontrada usando operações elementares de linhas de matrizes. Assim, a estratégia é a seguinte:

$$\left[\begin{array}{cc} A & : & I \end{array} \right] \Rightarrow [\text{operações elementares de linhas}] \Rightarrow \left[\begin{array}{cc} I & : & A^{-1} \end{array} \right]$$

Se não for possível encontrar uma matriz identidade usando a estratégia anterior, então a matriz A não tem inversa. Nesse caso aparece pelo menos uma linha com todos os elementos iguais a zero, mostrando que a matriz A não tem posto completo.

Propriedade Importante:

Uma matriz quadrada A de dimensão $n \times n$ tem inversa $\iff$ suas linhas (ou equivalentemente suas colunas) são linearmente independentes ou, em outras palavras, a matriz A tem posto completo.

Exemplo 2.8: Encontrar a inversa da seguinte matriz:

$$A = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 2 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

Realizando operações elementares por linha da matrizes encontramos o seguinte:

$$\left[\begin{array}{ccc|ccc} A & : & I \end{array} \right] \Rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & : & 1 & 0 & 0 \\ -1 & 2 & -1 & : & 0 & 1 & 0 \\ 0 & 1 & 1 & : & 0 & 0 & 1 \end{array} \right] \Rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & 4 & 0 & : & 1 & 1 & 0 \\ 0 & 1 & 1 & : & 0 & 0 & 1 \end{array} \right] \Rightarrow$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 0 & : & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 1 & : & -\frac{1}{4} & -\frac{1}{4} & 1 \end{array} \right] \Rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & : & \frac{5}{4} & \frac{1}{4} & -1 \\ 0 & 1 & 0 & : & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 1 & : & -\frac{1}{4} & -\frac{1}{4} & 1 \end{array} \right] \Rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & : & \frac{3}{4} & -\frac{1}{4} & -1 \\ 0 & 1 & 0 & : & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 1 & : & -\frac{1}{4} & -\frac{1}{4} & 1 \end{array} \right]$$

Observação importante: É oportuno lembrar a seguinte operação de matrizes:

$$\left[\begin{array}{ccc|ccc} A & : & I \end{array} \right] \Rightarrow A^{-1} \left[\begin{array}{ccc|ccc} A & : & I \end{array} \right] \Rightarrow \left[\begin{array}{ccc|ccc} A^{-1}A & : & A^{-1}I \end{array} \right] \Rightarrow \left[\begin{array}{ccc|ccc} I & : & A^{-1} \end{array} \right]$$

2.3.3. Posto de uma Matriz

O posto de uma matriz A de dimensão $m \times n$ é igual o número máximo de linhas (ou colunas) linearmente independentes da matriz A . Assim, temos o seguinte:

$$Posto(A) \leq \min(m, n)$$

Nesse contexto, se $Posto(A) = \min(m, n)$, então a matriz A é de posto completo.

Propriedade importante: Pode ser demonstrado que se o posto de uma matriz A é k , $Posto(A) = k$, então a matriz A pode ser reduzida da seguinte forma:

$$A \Rightarrow \left[\begin{array}{c} \text{Operações} \\ \text{elementares} \\ \text{de matrizes} \end{array} \right] \Rightarrow A' = \left[\begin{array}{ccc|ccc} I_k & : & Q \\ 0 & : & 0 \end{array} \right]$$

Portanto, a dimensão de k determina o posto da matriz. Se $k = \min(m, n)$, então a matriz A tem posto completo e se $k < \min(m, n)$, então a matriz A não tem posto completo. Adicionalmente, o número de linhas com valores iguais a zero permite identificar o número de linhas redundantes. Deve-se observar que a matriz que se encontra abaixo da matriz Q deve ter todos seus elementos iguais a zero. Em caso contrário, seria possível incrementar o posto da matriz A , isto é, a dimensão de I_k , transferindo para a diagonal um valor diferente de zero.

2.3.4. Análise do Processo de Solução de um Sistema de Equações Lineares

Consideremos que pretendamos resolver o seguinte sistema de equações algébricas lineares:

$$A x = b$$

em que A é uma matriz de dimensão $m \times n$, b é um vetor de dimensão m e x é o vetor de incógnitas de dimensão n .

Nesse contexto podem acontecer as seguintes possibilidades:

1. O Sistema algébrico linear é inconsistente.

Se $\text{Posto}[A, b] > \text{posto}[A]$, então o vetor b não pode ser representado como uma combinação linear dos vetores $a_1, a_2, \dots, a_n$ que são as colunas da matriz A . Em outras palavras, não é possível representar o vetor b como uma combinação linear dos vetores coluna da matriz A e, portanto, está caracterizado um sistema inconsistente. Veja que estamos pretendendo encontrar escalares $x_1, x_2, \dots, x_n$ (os elementos do vetor x) que permitam representar o vetor b como uma combinação linear dos vetores coluna $a_1, a_2, \dots, a_n$ (as colunas da matriz A). Nesse contexto, se o vetor b não for representável então temos um sistema inconsistente.

Exemplo 2.9: Consideremos o seguinte sistema algébrico linear:

$$\left. \begin{array}{l} x_1 + 2x_2 + x_3 - 2x_4 = 10 \\ -x_1 + 2x_2 - x_3 + x_4 = 6 \\ x_2 + x_3 = 2 \\ -x_1 + 9x_2 + 2x_3 = 29 \end{array} \right\} \Rightarrow \begin{bmatrix} 1 & 2 & 1 & -2 \\ -1 & 2 & -1 & 1 \\ 0 & 1 & 1 & 0 \\ -1 & 9 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 10 \\ 6 \\ 2 \\ 29 \end{bmatrix}$$

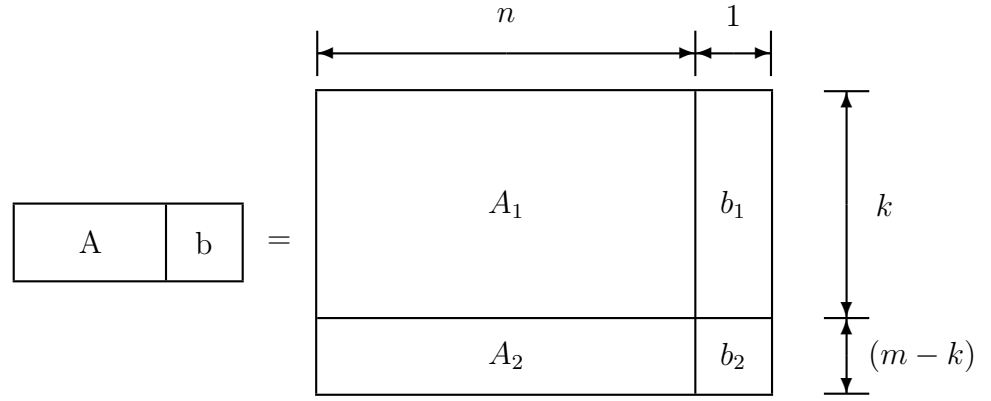
Para o sistema de equações mostrado anteriormente, pode ser provado que $\text{Posto}[A, b] = 4 > \text{Posto}[A] = 3$. Assim, como o $\text{Posto}[A, b] > \text{Posto}[A]$, então não é possível representar o vetor b como uma combinação linear dos vetores coluna da matriz A , isto é, não existem valores escalares para x_1, x_2, x_3 e x_4 que permitam de que o seguinte sistema seja verdadeiro:

$$\begin{bmatrix} 10 \\ 6 \\ 2 \\ 29 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ -1 \end{bmatrix} x_1 + \begin{bmatrix} 2 \\ 2 \\ 1 \\ 9 \end{bmatrix} x_2 + \begin{bmatrix} 1 \\ -1 \\ 1 \\ 2 \end{bmatrix} x_3 + \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} x_4$$

Na verdade, os quatro vetores coluna não são linearmente independentes e não geram o espaço E^4 e, particularmente, não podem gerar o vetor b indicado. Assim, neste caso, dizemos que o sistema algébrico linear é inconsistente. Na prática não se encontra o $\text{Posto}[A, b]$ e o $\text{Posto}[A]$ porque representam tarefas trabalhosas. Assim, adota-se uma estratégia geral de operações elementares de matrizes para qualquer sistema algébrico linear e, no final, deve-se identificar o tipo de solução como é mostrado adiante.

2. O Sistema algébrico linear tem solução factível.

Se $\text{Posto}[A] = \text{Posto}[A, b] = k$, então existe pelo menos uma solução factível. Neste caso o sistema algébrico linear ou melhor a matriz $[A, b]$ pode ser ordenada da seguinte forma:



onde a matriz A_1 é de dimensão $k \times n$, o vetor b_1 é de dimensão k , a matriz A_2 é de dimensão $(m - k) \times n$ e o vetor b_2 é de dimensão $(m - k)$. Assim, de forma geral, podemos supor que:

$$posto[A_1] = Posto[A_1, b_1] = k$$

Exemplo 2.10: Consideremos o seguinte sistema algébrico linear:

$$\left. \begin{array}{l} x_1 + 2x_2 + x_3 - 2x_4 = 10 \\ -x_1 + 2x_2 - x_3 + x_4 = 6 \\ x_2 + x_3 = 2 \\ -x_1 + 9x_2 + 2x_3 = 28 \end{array} \right\} \Rightarrow \begin{bmatrix} 1 & 2 & 1 & -2 \\ -1 & 2 & -1 & 1 \\ 0 & 1 & 1 & 0 \\ -1 & 9 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 10 \\ 6 \\ 2 \\ 28 \end{bmatrix}$$

Para o sistema de equações mostrado anteriormente, pode ser provado que $Posto[A, b] = Posto[A] = 3 < \min(m, n)$. Assim, podemos concluir que os vetores coluna da matriz A não são linearmente independentes. Nesse contexto, podemos ordenar o sistema anterior da seguinte forma:

$$\begin{array}{|c|c|} \hline A & b \\ \hline \end{array} = \begin{array}{|cccc|c|} \hline 1 & 2 & 1 & -2 & 10 \\ -1 & 2 & -1 & 1 & 6 \\ 0 & 1 & 1 & 0 & 2 \\ -1 & 9 & 2 & 0 & 28 \\ \hline \end{array}$$

Neste caso, se o vetor x resolve o sistema $A_1x = b_1$, então também resolve automaticamente o sistema $A_2x = b_2$. Portanto, as equações $A_2x = b_2$ são chamadas de redundantes ou dependentes e devem ser eliminadas do processo de solução. Entretanto, não é trivial identificar as equações redundantes. Também neste caso é usada uma estratégia geral mostrada adiante em que no final do processo podem ser identificadas as equações redundantes.

Após eliminar as equações redundantes temos o sistema reduzido $A_1 x = b_1$ com $\text{Posto}[A] = k$. Neste contexto, é possível separar A_1 em duas parcelas da seguinte forma:

$$A_1 = \begin{array}{c} \begin{array}{cc} \xleftarrow{k} & \xrightarrow{(n-k)} \\ \hline \begin{array}{|c|c|} \hline B & N \\ \hline \end{array} & \begin{array}{c} \updownarrow k \\ \hline \end{array} \end{array} \end{array}$$

onde B é uma matriz quadrada de dimensão $k \times k$ e $\text{Posto}[B] = k$. Nesse contexto, a matriz B é chamada de **matriz básica** porque suas colunas formam uma base para o espaço E^k e N é chamada de **matriz não básica**.

Para o Exemplo 2.10 podemos fazer a seguinte separação para obter A_1 :

$$A_1 = \begin{bmatrix} 1 & 2 & 1 & -2 \\ -1 & 2 & -1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \Rightarrow B = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 2 & -1 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{e} \quad N = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

onde B é a matriz básica e N é a matriz não básica.

Também, podemos fazer a seguinte separação de variáveis:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ \dots\dots\dots \\ x_{k+1} \\ \vdots \\ x_n \end{bmatrix} \Rightarrow x_B = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \quad \text{e} \quad x_N = \begin{bmatrix} x_{k+1} \\ \vdots \\ x_n \end{bmatrix}$$

Assim, podemos fazer a seguinte transformação:

$$A_1 x = b_1 \Rightarrow \begin{bmatrix} B & N \end{bmatrix} \begin{bmatrix} x_B \\ x_N \end{bmatrix} = b_1 \Rightarrow B x_B + N x_N = b_1$$

Como B é de posto completo, então existe a inversa B^{-1} e, portanto, a relação anterior pode ser modificada da seguinte forma:

$$B^{-1} B x_B + B^{-1} N x_N = B^{-1} b_1 \Rightarrow x_B = B^{-1} b_1 - B^{-1} N x_N \quad (2.5)$$

Para o exemplo que estamos resolvendo temos o seguinte:

$$B^{-1} = \begin{bmatrix} \frac{3}{4} & -\frac{1}{4} & -1 \\ \frac{1}{4} & \frac{1}{4} & 0 \\ -\frac{1}{4} & -\frac{1}{4} & 1 \end{bmatrix}$$

e substituindo em (2.5) temos o seguinte:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & -\frac{1}{4} & -1 \\ \frac{1}{4} & \frac{1}{4} & 0 \\ -\frac{1}{4} & -\frac{1}{4} & 1 \end{bmatrix} \begin{bmatrix} 10 \\ 6 \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{3}{4} & -\frac{1}{4} & -1 \\ \frac{1}{4} & \frac{1}{4} & 0 \\ -\frac{1}{4} & -\frac{1}{4} & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} [x_4]$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ -2 \end{bmatrix} + \begin{bmatrix} \frac{7}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \end{bmatrix} [x_4] \Rightarrow \begin{bmatrix} x_1 = 4 + \frac{7}{4}\lambda \\ x_2 = 4 + \frac{1}{4}\lambda \\ x_3 = -2 - \frac{1}{4}\lambda \\ x_4 = \lambda \end{bmatrix}$$

onde λ é um número real.

No exemplo que está sendo analisado existem duas possibilidades:

- a) Quando $k = n$: Neste caso N é vazio e o sistema tem apenas uma solução única que assume a seguinte forma:

$$x_B = B^{-1}b_1 = \begin{bmatrix} 4 \\ 4 \\ -2 \end{bmatrix}$$

- b) Quando $k < n$: Neste caso podemos dar valores arbitrários aos elementos de x_N (x_4 no exemplo) e, uma vez escolhidos esses valores, pode-se encontrar uma solução para os elementos de x_B usando a seguinte relação:

$$x_B = B^{-1}b_1 - B^{-1}N x_N$$

Neste caso, o sistema aceita infinitas soluções e esse tipo de problema é o que **interessa em programação linear**.

Observação importante: Em um sistema como o mostrado anteriormente, se especificamos valores arbitrários para os elementos de x_N , então encontramos uma solução única para os elementos de x_B já que a matriz B tem inversa. Entretanto, um tipo de **solução particularmente especial** é aquela encontrada dando valores iguais a zero para todos os elementos de x_N . Nesse contexto temos o seguinte:

$$x_N = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \implies x_B = B^{-1}b_1$$

Esse tipo de solução é conhecida como **solução básica** do sistema $A_1x = b_1$. Finalmente, se nessa estratégia de busca de uma solução para o sistema algébrico linear acontece que encontramos uma solução com $x_N = 0$ e $x_B = B^{-1}b_1 \geq 0$, então esse tipo de solução é conhecida como **solução básica factível (SBF)**. O conceito de SBF é crucial ao analisar os tópicos de resolução de problemas de programação linear (PL).

Resumindo, ao resolver um sistema de equações algébricas lineares pode acontecer uma das seguintes alternativas:

1. Se $\text{Posto}[A, b] > \text{Posto}[A] \implies Ax = b$ é um sistema inconsistente e não existe solução.
2. Se $\text{Posto}[A, b] = \text{Posto}[A] = k = n \implies Ax = b$ tem uma única solução.
3. Se $\text{Posto}[A, b] = \text{Posto}[A] = k < n \implies Ax = b$ tem infinitas soluções.

Observação importante: Todas as características mencionadas anteriormente podem ser verificadas implementando a eliminação de Gauss da seguinte forma:

$$Ax = b \implies \left[\begin{array}{l} \text{Eliminação de Gauss} \\ \text{com operações elementares} \\ \text{de linhas de matrizes} \end{array} \right] \implies A'x = b'$$

Após realizar o processo de transformação mostrado anteriormente, encontramos um sistema com a seguinte estrutura:

$$\left[\begin{array}{ccc} I_k & \vdots & Q \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 0 \end{array} \right] x = \left[\begin{array}{c} b'_1 \\ \cdots \\ b'_2 \end{array} \right]$$

Nesse contexto, aparecem as seguintes possibilidades:

1. Se $\text{Posto}[A] = k < m$ e $b'_2 \neq 0$, então o sistema é inconsistente e não existe solução. Veja que nesse caso existe uma equação inconsistente por cada elemento de $b'_2 \neq 0$.
2. Se $\text{Posto}[A] = k$ e $k = n \leq m$ e $b'_2 = 0$, então o sistema tem solução única.
3. Se $\text{Posto}[A] = k$, $k < n$ e $b'_2 = 0$ (quando $k < m$), então o sistema tem infinitas soluções.

2.4. Conjuntos Convexos e Funções Convexas

2.4.1. Conjunto Convexo

Um conjunto X em E^n é chamado de conjunto convexo se dados dois elementos x_1 e x_2 que pertencem a X então o elemento genérico $x = \lambda x_1 + (1 - \lambda)x_2 \in X$ para cada $\lambda \in [0, 1]$. A Figura 2.3 mostra um conjunto convexo e um conjunto não convexo.

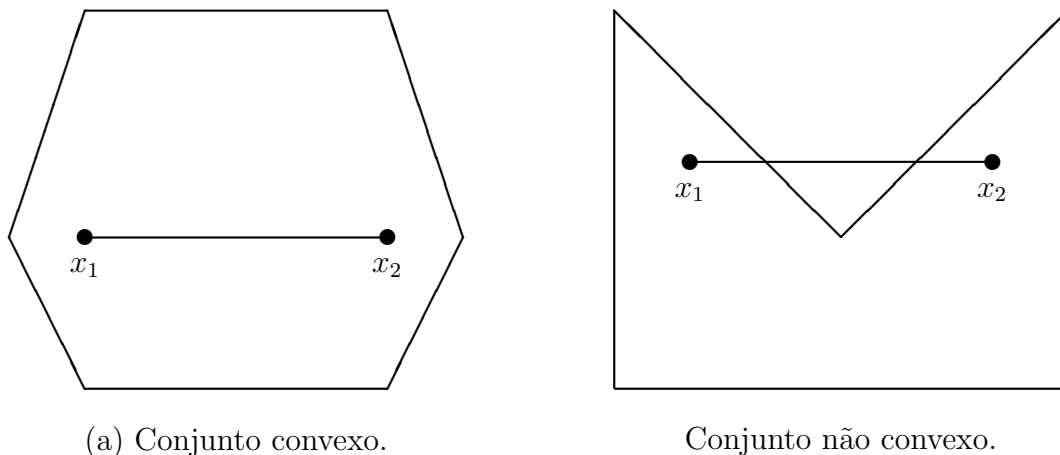
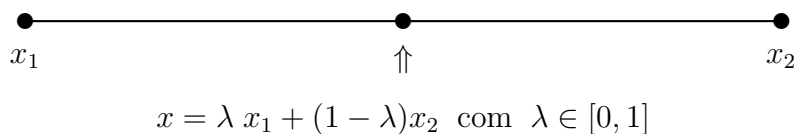


Figura 2.3: Conjuntos convexo e não convexo.

Deve-se observar que em E^2 , os pontos que satisfazem a condição $x = \lambda x_1 + (1 - \lambda)x_2$ com $\lambda \in [0, 1]$, são pontos do segmento de reta que une x_1 e x_2 e pode ser observado graficamente da seguinte forma:



Deve-se observar que para cada $\lambda \in [0, 1]$ é gerado um ponto no segmento de reta que une os pontos x_1 e x_2 . Assim, todo ponto no segmento de reta que une x_1 e x_2 é chamado de combinação convexa de x_1 e x_2 .

Exemplo 2.11: Exemplos de conjuntos convexos. Nos exemplos mostrados A é uma matriz de dimensão $m \times n$, b é um vetor de dimensão m e x é um vetor de dimensão n . Nesse contexto, os seguintes conjuntos são conjuntos convexos:

- $X = \{x : A x = b\}.$
- $X = \{x : A x = b; \quad x \geq 0\}.$
- $X = \{x : A x \leq b; \quad x \geq 0\}.$

Como ilustração, provamos apenas que o primeiro conjunto é um conjunto convexo.

Sejam x_1 e x_2 dois pontos que pertencem ao conjunto X . Então, devemos provar que X é um conjunto convexo, isto é, devemos provar que o ponto $x' = \lambda x_1 + (1 - \lambda)x_2 \in X$ para $\lambda \in [0, 1]$.

Se $x_1 \in X \implies A x_1 = b$.

Se $x_2 \in X \implies A x_2 = b$.

Multiplicando a primeira relação por λ e a segunda relação por $(1 - \lambda)$ e somado ambas relações temos o seguinte:

$$\lambda A x_1 = \lambda b \quad \text{e} \quad (1 - \lambda) A x_2 = (1 - \lambda) b$$

$$\lambda A x_1 + (1 - \lambda) A x_2 = \lambda b + (1 - \lambda) b$$

$$A[\lambda x_1 + (1 - \lambda) x_2] = b \implies A x' = b$$

e, portanto, $x' = \lambda x_1 + (1 - \lambda)x_2 \in X$.

Observação importante: Na prova anterior x_1 é um vetor de dimensão n que pertence ao conjunto X . Em outros casos, x_1 também representa o primeiro elemento do vetor x e, portanto, é um escalar. Entretanto, o contexto identifica facilmente se x_1 é um escalar ou um vetor de dimensão n . Também sempre que usamos o sistema $A x = b$ deve estar implícito que x tem dimensão n e, portanto, estamos trabalhando no espaço E^n .

2.4.2. Tipos Especiais de Conjuntos Convexos: Hiperplano, Subespaço e Raio

A seguir apresentamos alguns conjuntos convexos muito importantes na resolução de problemas de programação linear.

Hiperplano

Um hiperplano H no espaço E^n é um conjunto convexo definido da seguinte forma:

$$H = \{x : p x = k\}$$

onde p é um vetor conhecido de dimensão n e diferente de zero e k é um escalar também conhecido. A Figura 2.4 mostra um hiperplano. Devemos observar que, por exemplo, a relação $3x_1 - x_2 + 4x_3 = 9$ é um hiperplano em E^3 . Assim, o sistema algébrico $A x = b$ está formado por m hiperplanos em E^n e o conjunto $X = \{x : A x = b\}$ está definido por pontos x que se encontram dentro da interseção de m hiperplanos.

Existe uma definição alternativa de hiperplano. Seja $x_o \in H$, então temos o seguinte:

$$p x_o = k$$

$$p x = k$$

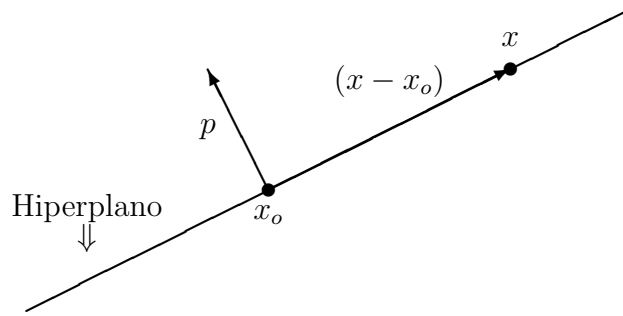


Figura 2.4: Hiperplano.

Das duas relações anteriores encontramos uma forma alternativa de caracterizar um hiperplano:

$$p(x - x_o) = 0$$

O hiperplano em E^2 representa a equação de uma reta, em E^3 representa a equação de um plano e, assim sucessivamente.

Subespaço

Um subespaço no espaço E^n é um conjunto convexo definido da seguinte forma:

$$H^+ = \{x : p x \geq k\}$$

ou também o subespaço complementar:

$$H^- = \{x : p x \leq k\}$$

Assim, claramente, a união dos dois subespaços representa o espaço completo E^n , isto é, $E^n = H^+ \cup H^-$. Portanto, um hiperplano separa o espaço E^n em duas regiões ou subespaços como é mostrado na Figura 2.5. Deve-se observar que um subespaço é um conjunto convexo.

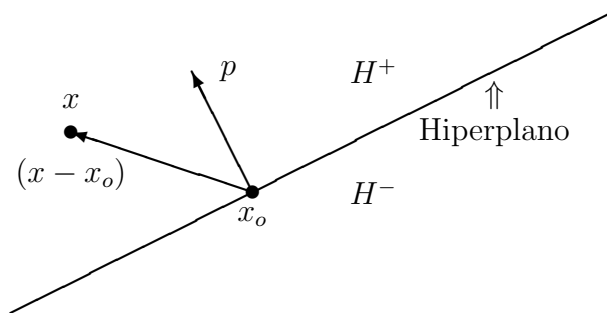


Figura 2.5: Subespaços.

Raio

O raio é um conjunto convexo definido da seguinte forma:

$$R = \{x : x = x_o + \lambda d; \quad \lambda \geq 0\}$$

onde d é um vetor conhecido e diferente de zero e x_o é um ponto conhecido e chamado de vértice de um raio. A Figura 2.6 mostra um raio. Um raio também é um conjunto convexo e essa propriedade é demonstrada através de um teorema.

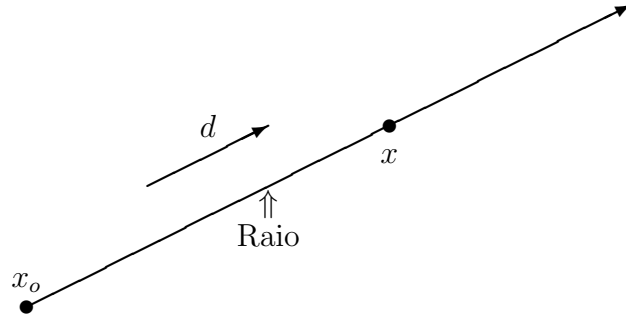


Figura 2.6: Raio.

Teorema 2.1: Um raio é um conjunto convexo.

Prova: Provaremos que o raio:

$$R = \{x : x = x_o + \lambda d; \quad \lambda \geq 0\}$$

é um conjunto convexo.

Sejam x_1 e x_2 dois pontos que pertencem ao raio R . Assim, devemos provar que o ponto genérico $x' = \lambda x_1 + (1 - \lambda)x_2 \in R$ para $\lambda \in [0, 1]$.

$$\text{Se } x_1 \in R \implies x_1 = x_o + \lambda_1 d \quad \text{e} \quad \lambda_1 \geq 0.$$

$$\text{Se } x_2 \in R \implies x_2 = x_o + \lambda_2 d \quad \text{e} \quad \lambda_2 \geq 0.$$

Multiplicando a primeira relação por λ e a segunda relação por $(1 - \lambda)$ e somado ambas relações temos o seguinte:

$$\lambda x_1 = \lambda x_o + \lambda \lambda_1 d$$

$$(1 - \lambda) x_2 = (1 - \lambda) x_o + (1 - \lambda) \lambda_2 d$$

$$\lambda x_1 + (1 - \lambda) x_2 = x_o + [\lambda \lambda_1 + (1 - \lambda) \lambda_2] d$$

com $\lambda_1 \geq 0$; $\lambda_2 \geq 0$ e $\lambda \in [0, 1]$. Agora fazemos as seguintes substituições:

$$x' = \lambda x_1 + (1 - \lambda) x_2 \quad \text{e} \quad \lambda' = \lambda \lambda_1 + (1 - \lambda) \lambda_2$$

e podemos verificar facilmente que $\lambda' \geq 0$ porque $\lambda \geq 0$; $(1 - \lambda) \geq 0$; $\lambda_1 \geq 0$ e $\lambda_2 \geq 0$. Portanto, foi demonstrado que o ponto genérico x' :

$$x' = \lambda x_1 + (1 - \lambda)x_2 = x_o + \lambda' d; \quad \text{com } \lambda' \geq 0 \in R$$

e, portanto, foi provado que R é um conjunto convexo.

2.4.3. Pontos extremos, Direções e Direções Extremas de um Conjunto Convexo

Este tópico é fundamental para entender alguns tópicos relacionados com a resolução de problemas de programação linear. Deve-se mencionar que o conjunto de restrições de um problema de PL é um conjunto convexo e, adicionalmente, pode ter direções e direções extremas.

Ponto Extremo de um Conjunto Convexo

Um ponto $x \in X$ é chamado de ponto extremo se esse ponto não pode ser representado como uma combinação estrita e convexa de dois pontos distintos do conjunto X .

Assim, se $x = \lambda x_1 + (1 - \lambda)x_2$ com $\lambda \in [0, 1]$ e x_1 e $x_2 \in X \rightarrow$ se $x = x_1 = x_2$, então x é um ponto extremo de X . Em outras palavras, um ponto extremo $x \in X$ é um ponto que pode ser representado por uma combinação convexa apenas dele próprio. A Figura 2.7 mostra o ponto extremo x_1 e os pontos x_2 e x_3 que não são pontos extremos. Deve-se observar que a definição de ponto extremo mostrada anteriormente tem pouca utilidade prática, apenas é usada para encontrar formas alternativas para caracterizar um ponto extremo.

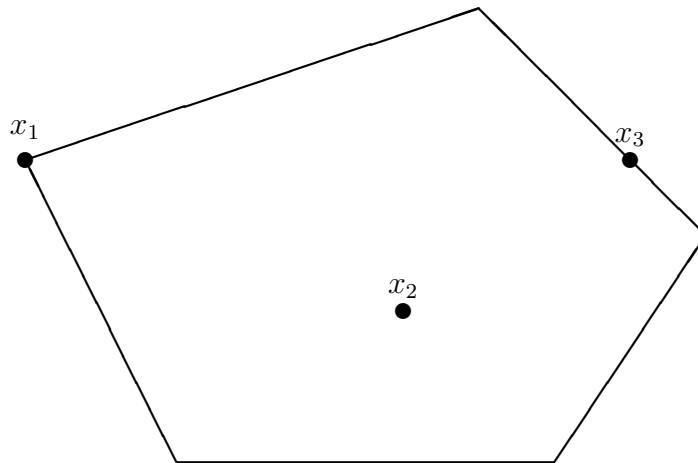


Figura 2.7: Ponto extremo de um conjunto convexo.

Direção de um Conjunto Convexo

Dado um conjunto convexo X , então um vetor $d \neq 0$ é chamado de direção desse conjunto convexo se para cada ponto x_o desse conjunto X o raio:

$$R = \{x : x = x_o + \lambda d; \quad \lambda \geq 0\}$$

também pertence ao conjunto convexo X . Nesse contexto, existem raios em conjuntos convexos ilimitados e, portanto, nesse tipo de conjuntos existem direções. A Figura 2.8 mostra um conjunto convexo limitado e outro conjunto convexo ilimitado.

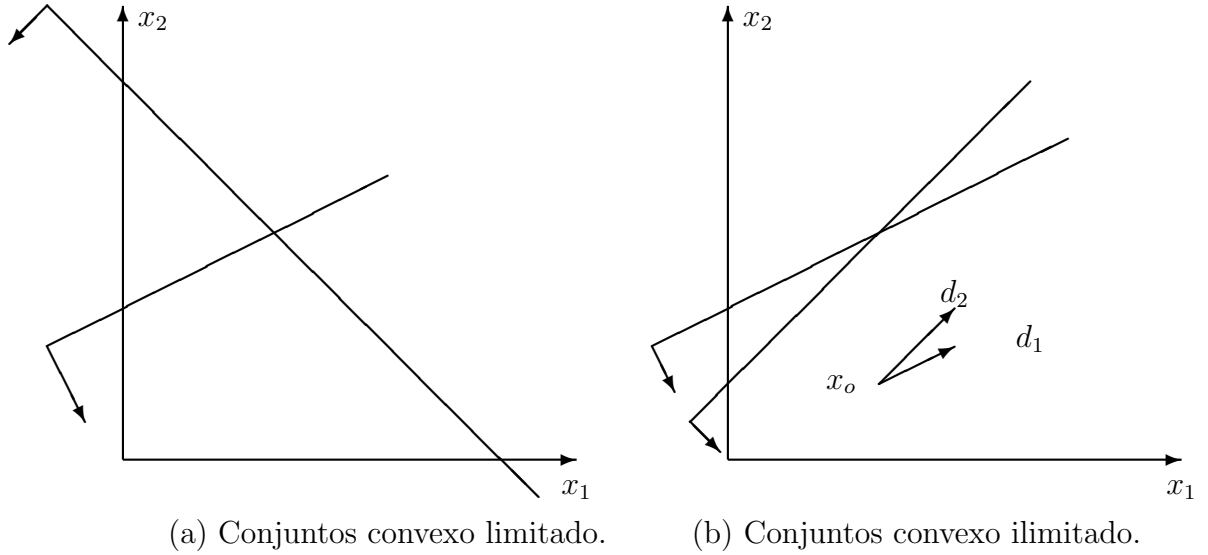


Figura 2.8: Conjuntos convexos.

Direção Extrema de um Conjunto Convexo

Uma direção de um conjunto convexo X é chamada de direção extrema desse conjunto convexo se não pode ser representada como uma combinação linear positiva de duas direções distintas de X . O conceito de direção extrema de um conjunto convexo X é muito importante para entender a técnica de resolução de problemas de programação linear. Adiante, deve ser melhor formulado o conceito de direção extrema no contexto de um problema de PL.

2.4.4. Pontos extremos, Direções e Direções Extremas de um Conjunto Convexo Poliedral

Neste caso limitamos a análise de pontos extremos, direções e direções extremas de um tipo especial de conjunto convexo, isto é, do chamado conjunto convexo poliedral. As chamadas restrições de um problema de programação linear na verdade são um conjunto convexo poliedral e, por esse motivo, esse é o tipo de conjunto convexo que interessa analisar.

Um conjunto convexo poliedral é a interseção de um número finito de subespaços. Se o conjunto poliedral é limitado, então esse tipo de conjunto é chamado de politopo. Assim, um conjunto convexo poliedral é um tipo especial de conjunto convexo. Um conjunto convexo poliedral pode ser definido da seguinte forma:

$$X = \{x : Ax \leq b; \quad x \geq 0\} \quad (2.6)$$

onde cada linha ou restrição do sistema algébrico linear mostrado anteriormente representa um subespaço. A Figura 2.8 mostra conjuntos poliedrais típicos.

Deve-se observar que o conjunto poliedral apresentado anteriormente tem a restrição na forma $\leq$ apenas para facilidade de análise posterior. Devemos lembrar que uma restrição do tipo $\geq$ pode ser facilmente transformada na forma $\leq$ e uma restrição de igualdade (que identifica um hiperplano) pode ser transformada em duas restrições de desigualdade e ambas restrições podem assumir a forma $\leq$ (um hiperplano pode ser substituído por dois subespaços). Portanto, a forma $Ax \leq b$ representa uma forma totalmente geral de representação de restrições de igualdade e de desigualdade. Na análise realizada nesta Subseção e no restante do Capítulo, se não for mencionado explicitamente, assumimos que o conjunto convexo X é o definido em (2.6) e no espaço E^n .

Ponto Extremo de um Conjunto Convexo Poliedral

Neste caso pretendemos encontrar uma forma alternativa de caracterizar um ponto extremo que seja útil para analisar de forma mais adequada as características de um problema de programação linear. Essa forma alternativa é encontrada através de um teorema a partir da definição de ponto extremo mostrado anteriormente.

Teorema 2.2: Forma alternativa de caracterizar um ponto extremo

Um ponto $\bar{x} \in X$ pertence a n hiperplanos linearmente independentes $\iff \bar{x}$ não pode ser escrito como uma combinação convexa de dois pontos do conjunto X (que é equivalente a afirmar que $\bar{x}$ é um ponto extremo).

Prova: $\implies$

Assumimos que no ponto $\bar{x} \in X$ estão ativos n hiperplanos linearmente independentes. Assim, pretendemos demonstrar que $\bar{x}$ não pode ser representado como uma combinação convexa de dois pontos diferentes de X , o que equivale a dizer que $\bar{x}$ é um ponto extremo.

Em geral, $\bar{x} \in X$ pode ser representado através de combinação convexa de dois pontos de X na forma:

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

em que $0 < \lambda < 1$ e com x' e $x'' \in X$. Inicialmente, supor que em $\bar{x} \in X$ existe uma restrição ativa definida pela restrição $\alpha x \leq \beta$, então a seguinte relação é verdadeira:

$$\alpha \bar{x} = \beta$$

que substituindo na relação anterior produz a seguinte relação:

$$\alpha \bar{x} = \beta \implies \alpha[\lambda x' + (1 - \lambda)x''] = \beta \implies \lambda(\alpha x') + (1 - \lambda)(\alpha x'') = \beta$$

A única forma de garantir de que a relação anterior seja sempre verdadeira é que as seguintes relações sejam verdadeiras:

$$\alpha x' = \beta \quad \text{e} \quad \alpha x'' = \beta$$

Portanto, foi provado que se em $\bar{x}$ existe um hiperplano ativo, então esse mesmo hiperplano está também ativo em x' e x'' . Generalizando, podemos afirmar que se em $\bar{x}$ existem n hiperplanos ativos linearmente independentes, então esses mesmos hiperplanos estão também ativos em x' e x'' .

Em resumo, se em $\bar{x}$ existem n hiperplanos ativos linearmente independentes (LI), então esses mesmos n hiperplanos estão ativos em x' e x'' . Entretanto, dois pontos diferentes no espaço E^n não podem ter os mesmos n hiperplanos ativos e, portanto, a única possibilidade é que $\bar{x} = x' = x''$, o que prova que $\bar{x}$ não pode ser representado por uma combinação convexa de dois pontos diferentes de X . Assim, foi provado que $\bar{x}$ é um ponto extremo.

Prova: $\Leftarrow$ Prova pelo absurdo.

Supor que o número máximo de hiperplanos LI ativos em $\bar{x} \in X$ é $r < n$. Então provaremos que a interseção desses hiperplanos não representa um ponto extremo.

Os hiperplanos ativos podem ser representados por $G \bar{x} = g$ onde G é uma matriz de dimensão $r \times n$. Também, seja $d \neq 0$ uma solução de $G d = 0$. Esse d existe porque $n > r$ e assim é possível escolher valores adequados para alguns elementos de d (veja que na verdade $G d = 0$ é um sistema com infinitas soluções). Assim, existem dois pontos muito próximos de $\bar{x}$ obtidos a partir de um $\epsilon > 0$ de forma que $x' = \bar{x} + \epsilon d$ e $x'' = \bar{x} - \epsilon d$ que ainda pertencem a X porque $G x' = g$ e $G x'' = g$ de acordo com o seguinte raciocínio:

$$G x' = G (\bar{x} + \epsilon d) = G \bar{x} + \epsilon G d = G \bar{x} + \epsilon(0) = g$$

$$G x'' = G (\bar{x} - \epsilon d) = G \bar{x} - \epsilon G d = G \bar{x} - \epsilon(0) = g$$

e, adicionalmente, as restrições de X que não estavam ativas em $\bar{x}$ permanecem factíveis para um $\epsilon > 0$ suficientemente pequeno. Assim, por exemplo, podemos escolher um $\bar{x} = 0,7x' + 0,3x''$ e com x' e $x'' \in X$, o que mostra, claramente, que $\bar{x}$ não é um ponto extremo, já que foi representado por uma combinação convexa de dois pontos x' e $x'' \in X$ e, portanto, foi provado que $\bar{x}$ não é um ponto extremo.

Observação importante: A caracterização de ponto extremo como sendo todo ponto $\bar{x} \in X$ em que n hiperplanos LI se encontram ativos é fundamental para entender a técnica de resolução de problemas de programação linear. Essa forma de caracterização chamamos de interpretação geométrica de um ponto extremo em contraposição com a interpretação algébrica que é abordada adiante.

Forma Sistemática para Encontrar os Pontos Extremos de um Conjunto Convexo Poliedral

Uma vez caracterizado um conjunto convexo poliedral X , então é possível encontrar todos os pontos extremos do conjunto X . Para isso devemos montar uma estratégia usando o conceito de ponto extremo encontrado no Teorema 2.2, isto é, um ponto extremo de um conjunto convexo poliedral X é qualquer ponto $x \in X$ (é um ponto factível) e, adicionalmente, nesse ponto devem existir n hiperplanos ativos linearmente independentes. A estratégia é muito simples, mas

extremamente demorada. Assim, essa estratégia assume a seguinte forma para um conjunto X conhecido no espaço E^n :

1. Supor que estamos trabalhando no espaço E^n e o conjunto convexo poliedral X tem k restrições de igualdade (são hiperplanos) e p restrições de desigualdade (são subespaços). Assim, encontrar os valores de n , k e p . Com essa informação, encontrar o número máximo de candidatos a ponto extremo usando a seguinte relação:

$$N_{max} = \binom{p}{n-k} = \frac{p!}{(n-k)!(p+k-n)!}$$

que é apenas a relação muito conhecida de encontrar o número de combinações existentes em p elementos tomados em grupos de $(n-k)$. Assim, queremos identificar um ponto $\bar{x}$ em que existam n hiperplanos ativos linearmente independentes (HALI). Deve-se observar que sempre existem k HALI relacionados com as restrições de igualdade (hiperplano) e, portanto, devemos escolher $(n-k)$ HALI do grupo de p subespaços.

2. Para cada candidato a ponto extremo fazer o seguinte:

- a) Identificar os subespaços que devem ter os hiperplanos ativos, isto é, $(n-k)$ subespaços devem ter seus HA. Juntar esses $(n-k)$ HA com as k restrições de igualdade (hiperplanos que sempre se encontram ativos) para formar um sistema algébrico linear de n equações e n incógnitas.
- b) Resolver o sistema algébrico linear e encontrar a solução $\bar{x}$.
- c) Verificar se a solução $\bar{x}$ encontrada em (b) não viola alguma das $p-(n-k)$ restrições de desigualdade não consideradas para montar o conjunto de $(n-k)$ hiperplanos ativos. Em outras palavras, verificar se $\bar{x}$ pertence à interseção dos $p-(n-k)$ subespaços não considerados na montagem do sistema algébrico linear. Se $\bar{x}$ for factível, isto é, não viola nenhuma restrição de desigualdade (subespaço), então $\bar{x}$ é um ponto extremo e, em caso contrário, $\bar{x}$ não é ponto extremo (é um ponto em que existem n HALI, mas é infactível).

Observação importante: A estratégia anterior pode ser muito demorada já que, para analisar cada candidato, a ponto extremo devemos resolver um sistema algébrico linear de n equações com n incógnitas. Assim, essa estratégia é usada apenas em caráter ilustrativo e para problemas com um número reduzido de variáveis.

Exemplo 2.12: Encontre o conjunto de pontos extremos do seguinte conjunto convexo poliedral não vazio:

$$X = \left\{ (x_1, x_2) : \begin{array}{l} -x_1 + 2x_2 \leq 6 \\ -x_1 + x_2 \leq 2 \\ x_2 \geq 1 \\ x_1 \geq 0 \end{array} \right\}$$

Neste caso estamos em E^2 e, portanto, $n = 2$. Adicionalmente, $p = 4$ e $k = 0$. Assim, o número máximo da candidatos a ponto extremo é seguinte:

$$N_{max} = \binom{4}{2} = \frac{4!}{2! 2!} = 6$$

Existem quatro subespaços que identificamos de forma trivial. Os seis candidatos são os seguintes:

1. Quando os subespaços 1 e 2 estão ativos:

Neste caso temos que resolver o sistema algébrico linear (SAL) seguinte:

$$-x_1 + 2x_2 = 6 \quad \text{e} \quad -x_1 + x_2 = 2$$

cuja solução é seguinte: $x_1 = 2$ e $x_2 = 4$.

O ponto $(2, 4)$ não viola as restrições 3 e 4 e, portanto, representa um ponto extremo.

2. Quando os subespaços 1 e 3 estão ativos:

Neste caso temos que resolver o SAL seguinte:

$$-x_1 + 2x_2 = 6 \quad \text{e} \quad x_2 = 1$$

cuja solução é seguinte: $x_1 = -4$ e $x_2 = 1$.

O ponto $(-4, 1)$ viola a restrição 2, $(-(-4) + 1 = 5 > 2)$, e a restrição 4, $(x_1 = -4 < 0)$ e, portanto, não é ponto extremo. Lembremos que se a solução encontrada viola apenas uma restrição, então ela não representa um ponto extremo.

3. Quando os subespaços 1 e 4 estão ativos:

Neste caso temos que resolver o SAL seguinte:

$$-x_1 + 2x_2 = 6 \quad \text{e} \quad x_1 = 0$$

cuja solução é seguinte: $x_1 = 0$ e $x_2 = 3$.

O ponto $(0, 3)$ viola a restrição 2, $(-(0) + 3 = 3 > 2)$ e, portanto, não é ponto extremo.

4. Quando os subespaços 2 e 3 estão ativos:

Neste caso temos que resolver o SAL seguinte:

$$-x_1 + x_2 = 2 \quad \text{e} \quad x_2 = 1$$

cuja solução é seguinte: $x_1 = -1$ e $x_2 = 1$.

O ponto $(-1, 1)$ viola a restrição 4 e, portanto, não é um ponto extremo.

5. Quando os subespaços 2 e 4 estão ativos:

Neste caso temos que resolver o SAL seguinte:

$$-x_1 + x_2 = 2 \quad \text{e} \quad x_1 = 0$$

cuja solução é seguinte: $x_1 = 0$ e $x_2 = 2$.

O ponto $(0, 2)$ não viola a restrição 1 nem a restrição 3 e, portanto, é ponto extremo.

6. Quando os subespaços 3 e 4 estão ativos:

Neste caso temos que resolver o SAL seguinte:

$$x_2 = 1 \quad \text{e} \quad x_1 = 0$$

cuja solução é seguinte: $x_1 = 0$ e $x_2 = 1$.

O ponto $(0, 1)$ não viola a restrição 1 nem a restrição 2 e, portanto, é um ponto extremo.

A Figura 2.9 mostra o conjunto X com os seis pontos, candidatos a pontos extremos, onde podemos observar que em todos esses pontos existem $n = 2$ HALI, mas apenas 3 deles são factíveis. Assim, nesse conjunto X existem 3 pontos extremos: $(2, 4)$, $(0, 2)$ e $(0, 1)$.

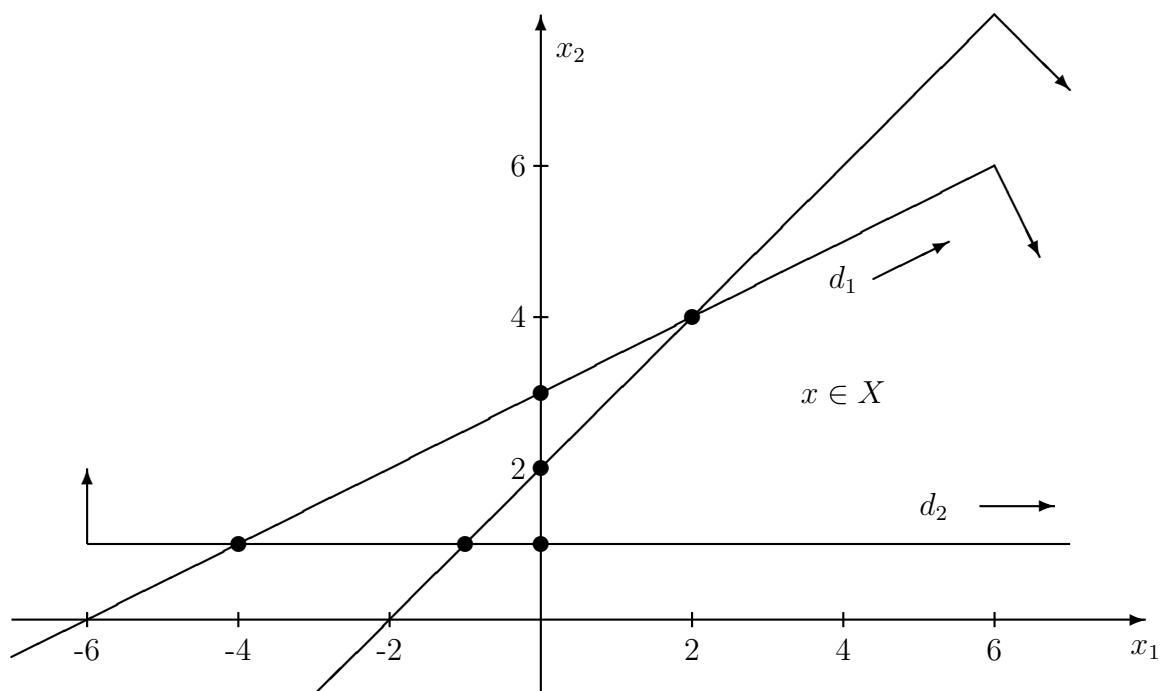


Figura 2.9: Pontos extremos do conjunto X do Exemplo 2.12.

Direções de um Conjunto Convexo Poliedral

Usando a definição de direção de um conjunto convexo, apresentada anteriormente, podemos encontrar facilmente o conjunto de direções de um conjunto convexo poliedral como o definido em (2.6).

Exemplo 2.13: Encontre o conjunto de direções do seguinte conjunto convexo poliedral não vazio:

$$X = \{x : Ax \leq b; \quad x \geq 0\}$$

Supor que d é uma direção de X e, portanto, qualquer ponto x' com a seguinte propriedade:

$$x' = x_o + \lambda d; \quad \lambda \geq 0$$

com $d \neq 0$ e $x_o \in X$ deve ser um ponto de X , isto é, $x' = x_o + \lambda d \in X$.

Se $x' \in X \implies Ax' \leq b$ e $x' \geq 0$. Assim, temos o seguinte:

$$A(x_o + \lambda d) \leq b \implies Ax_o + \lambda Ad \leq b \quad (2.7)$$

$$x_o + \lambda d \geq 0 \quad (2.8)$$

Como $x_o \in X$, então $Ax_o \leq b$ e de (2.7) temos o seguinte;

$$\lambda Ad \leq 0 \implies Ad \leq 0$$

porque $\lambda \geq 0$. Também, de (2.8) e levando em conta que $x_o \geq 0$ temos o seguinte:

$$\lambda d \geq 0 \implies d \geq 0$$

Portanto, o vetor d é uma direção do conjunto convexo poliedral X se d pertence ao seguinte conjunto D :

$$D = \{d : Ad \leq 0; \quad d \geq 0 \quad d \neq 0\}$$

Observação importante: Pode ser facilmente provado que o conjunto de direções D é um conjunto convexo.

Exemplo 2.14: Encontre o conjunto de direções do seguinte conjunto convexo poliedral não vazio:

$$X = \{x : Ax = b; \quad x \geq 0\}$$

Supor que d é uma direção de X e, portanto, qualquer ponto x' com a seguinte propriedade:

$$x' = x_o + \lambda d; \quad \lambda \geq 0$$

com $d \neq 0$ e $x_o \in X$ deve ser um ponto de X , isto é, $x' = x_o + \lambda d \in X$.

Se $x' \in X \implies A x' = b$ e $x' \geq 0$. Assim, temos o seguinte:

$$A(x_o + \lambda d) = b \implies A x_o + \lambda A d = b \quad (2.9)$$

$$x_o + \lambda d \geq 0 \quad (2.10)$$

Como $x_o \in X$, então $A x_o = b$ e de (2.9) temos o seguinte;

$$\lambda A d = 0 \implies A d = 0$$

porque $\lambda \geq 0$. Também, de (2.10) e levando em conta que $x_o \geq 0$ temos o seguinte:

$$\lambda d \geq 0 \implies d \geq 0$$

Portanto, o vetor d é uma direção do conjunto convexo poliedral X se d pertence ao seguinte conjunto D :

$$D = \{d : A d = 0; \quad d \geq 0 \quad d \neq 0\}$$

Como pode ser observado, a forma de encontrar as direções do conjunto X nos dois exemplos anteriores é praticamente a mesma. A mesma coisa acontece com todos os outros conjuntos convexos poliedrais. Assim, o conjunto D de direções do conjunto convexo $X = \{x : Ax \geq 0; \quad x \geq 0\}$ é dado por $D = \{d : A d \geq 0; \quad d \geq 0 \quad d \neq 0\}$. Adicionalmente, nem todas as variáveis de decisão estão obrigadas a cumprir a restrição $x \geq 0$ e, portanto, apenas aquelas variáveis $x_i \geq 0$ devem gerar a correspondente restrição $d_i \geq 0$.

Exemplo 2.15: Encontre o conjunto de direções do seguinte conjunto convexo poliedral não vazio:

$$X = \left\{ (x_1, x_2) : \begin{array}{l} -x_1 + 2x_2 \leq 6 \\ -x_1 + x_2 \leq 2 \\ x_2 \geq 1 \\ x_1 \geq 0 \end{array} \right\}$$

Devemos observar que cada restrição $a^i x (=)(\leq)(\geq) b_i$ de uma linha de $A x (=)(\leq)(\geq) b$ em X gera uma restrição do tipo $a^i d (=)(\leq)(\geq) 0$ em D e cada restrição do tipo $x_j (=)(\leq)(\geq) 0$ em X gera uma restrição do tipo $d_j (=)(\leq)(\geq) 0$ em D . Assim temos o seguinte:

Pela restrição $-x_1 + 2x_2 \leq 6 \implies -d_1 + 2d_2 \leq 0$.

Pela restrição $-x_1 + x_2 \leq 2 \implies -d_1 + d_2 \leq 0$.

Pela restrição $x_2 \geq 1 \implies d_2 \geq 0$.

Pela restrição $x_1 \geq 0 \implies d_1 \geq 0$.

Portanto, o conjunto de direções de X está dado pelo conjunto:

$$D = \left\{ d : d = (d_1, d_2); \quad d \neq 0; \quad \begin{array}{l} -d_1 + 2d_2 \leq 0 \\ -d_1 + d_2 \leq 0 \\ d_1 \geq 0 \\ d_2 \geq 0 \end{array} \right\}$$

A Figura 2.10 mostra o conjunto de direções D .

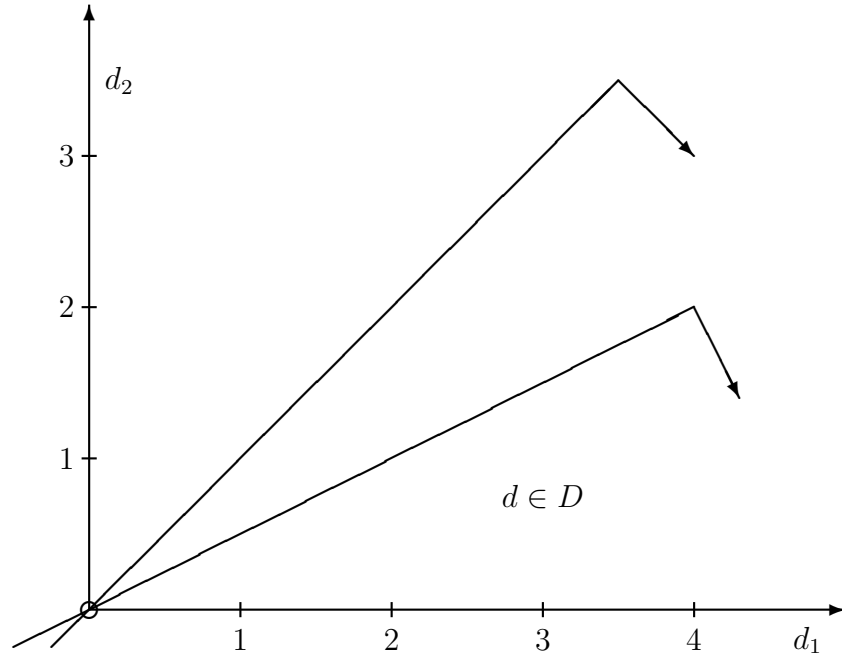


Figura 2.10: Região factível do conjunto D que são as direções do conjunto X .

Direções Extremas de um Conjunto Convexo Poliedral

As direções do conjunto convexo poliedral X :

$$X = \{x : Ax \leq b; \quad x \geq 0\}$$

são dadas pelos elementos que fazem parte do conjunto convexo D :

$$D = \{d : Ad \leq 0; \quad d \geq 0; \quad d \neq 0\}$$

Geometricamente, este conjunto pode ser encontrado simplesmente deslocando todos os hiperplanos que definem X , paralelamente aos mesmos até encontrar a origem. Lembremos a semelhança das seguintes relações:

$$Ax \leq b \iff Ad \leq 0$$

Para evitar duplicações na representação de direções, pode-se normalizar os vetores d usando a seguinte norma:

$$|d_1| + |d_2| + \dots + |d_n| = 1$$

mas como todos os elementos de $d \geq 0$, então a norma pode ser simplificada da seguinte forma:

$$d_1 + d_2 + \dots + d_n = 1$$

Assim, o conjunto de direções normalizadas de X formam o conjunto D que assume a seguinte forma:

$$D = \{d : \begin{array}{l} d \leq 0; \quad d \geq 0; \quad d \neq 0; \quad d_1 + d_2 + \dots + d_n = 1 \end{array}\}$$

em que d assume a seguinte forma:

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$$

Propriedade importante: Os pontos extremos de D normalizado são as direções extremas de X .

Essa propriedade importante permite encontrar as direções extremas de X usando a mesma estratégia usada para encontrar os pontos extremos de X . Assim, para conhecer as direções extremas de X precisamos apenas encontrar os pontos extremos do conjunto D normalizado.

Exemplo 2.16: Encontre o conjunto de direções normalizadas D e as direções extremas do seguinte conjunto convexo poliedral não vazio:

$$X = \left\{ (x_1, x_2) : \begin{array}{l} -x_1 + 2x_2 \leq 6 \\ -x_1 + x_2 \leq 2 \\ x_2 \geq 1 \\ x_1 \geq 0 \end{array} \right\}$$

No Exemplo 2.15 já foi encontrado o conjunto de direções D do conjunto convexo X . Assim, para encontrar o conjunto de direções normalizadas precisamos apenas adicionar a restrição de normalização. Assim, de forma trivial, podemos encontrar o conjunto D normalizado que assume a seguinte forma:

$$D = \left\{ d : d = (d_1, d_2); \quad d \neq 0; \quad \begin{array}{l} -d_1 + 2d_2 \leq 0 \\ -d_1 + d_2 \leq 0 \\ d_1 + d_2 = 1 \\ d_1 \geq 0 \\ d_2 \geq 0 \end{array} \right\}$$

A Figura 2.11 mostra o conjunto D normalizado. Veja que esse conjunto está representado apenas por um segmento de reta e, portanto, trivialmente verificamos que tem dois pontos extremos.

Pode-se verificar facilmente que a segunda restrição é irrelevante (não participa na caracterização da região factível de D) e, portanto, pode ser eliminada. Assim, o conjunto D normalizado apresenta as seguintes restrições:

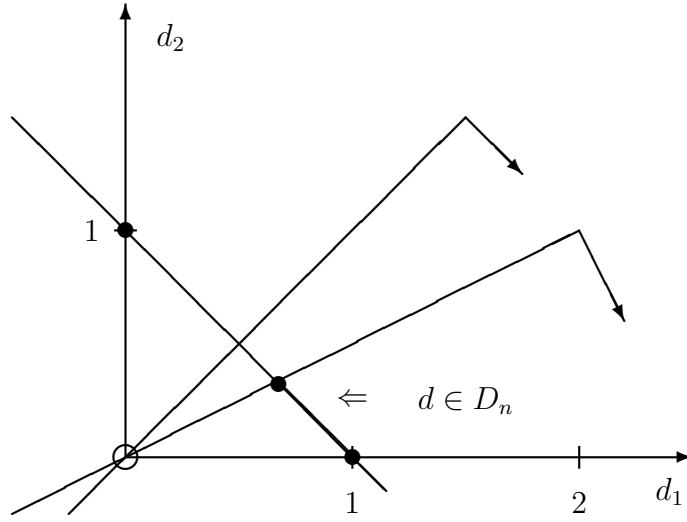


Figura 2.11: Região factível do conjunto D normalizado que são as direções do conjunto X .

$$\begin{aligned} -d_1 + 2d_2 &\leq 0 \\ d_1 + d_2 &= 1 \\ d_1 &\geq 0 \\ d_2 &\geq 0 \end{aligned}$$

sendo que adicionalmente existe a restrição trivial $d \neq 0$.

Para encontrar os pontos extremos do conjunto D normalizado, usamos a mesma estratégia usada para resolver o Exemplo 2.12.

Neste caso temos que $n = 2$, $k = 1$ e $p = 3$. Assim, o número de candidatos a ponto extremo é seguinte:

$$N_{max} = \binom{3}{1} = \frac{3!}{1! 2!} = 3$$

Existem três subespaços candidatos a fornecer um hiperplano, enumerados implicitamente, e um hiperplano que sempre se encontra ativo, isto é, a segunda relação matemática. Os três candidatos a ponto extremo são os seguintes:

1. Quando o subespaço 1 está ativo:

Neste caso temos que resolver o sistema algébrico linear (SAL) seguinte:

$$-d_1 + 2d_2 = 0 \quad \text{e} \quad d_1 + d_2 = 1$$

cuja solução é seguinte: $d_1 = \frac{2}{3}$ e $d_2 = \frac{1}{3}$.

O ponto $(2/3, 1/3)$ não viola os subespaços 2 e 3, veja que $d_1 = \frac{2}{3} > 0$ e $d_2 = \frac{1}{3} > 0$ e, portanto, representa um ponto extremo.

2. Quando o subespaço 2 está ativo:

Neste caso temos que resolver o sistema algébrico linear (SAL) seguinte:

$$d_1 = 0 \quad \text{e} \quad d_1 + d_2 = 1$$

cujas soluções são seguintes: $d_1 = 0$ e $d_2 = 1$.

O ponto $(0, 1)$ viola o subespaço 1, veja que $-1(0) + 2(1) = 2 > 0$ e, portanto, não é um ponto extremo.

3. Quando o subespaço 3 está ativo:

Neste caso temos que resolver o sistema algébrico linear (SAL) seguinte:

$$d_2 = 0 \quad \text{e} \quad d_1 + d_2 = 1$$

cujas soluções são seguintes: $d_1 = 1$ e $d_2 = 0$.

O ponto $(1, 0)$ não viola os subespaços 1 e 2, veja que $-1(1) + 2(0) = -1 < 0$ e $d_1 = 1 > 0$ e, portanto, representa um ponto extremo.

A Figura 2.11 mostra o conjunto D com os 3 pontos, candidatos a pontos extremos, onde podemos observar que em todos esses pontos existem $n = 2$ HALI, mas apenas dois são factíveis (veja que a restrição $-d_1 + d_2 \leq 0$ não foi considerada na análise). Assim, nesse conjunto D existem 2 pontos extremos: $(2/3, 1/3)$ e $(1, 0)$. Como esses pontos extremos de D normalizado são também direções extremas do conjunto X , então concluímos que $(2/3, 1/3)$ e $(1, 0)$ são direções extremas de X . Finalmente, concluímos que o conjunto convexo poliedral X tem 3 pontos extremos e duas direções extremas como pode ser visto na Figura 2.9.

2.4.5. Faces, Arestas e Pontos extremos Adjacentes de um Conjunto Convexo Poliedral

Existem ainda alguns conceitos relacionados com um conjunto convexo poliedral X que eventualmente são usados na formulação de técnicas de solução de problemas de programação linear.

Face Própria de X

A face própria de um conjunto convexo poliedral X está formada por todos os pontos $x \in X$ que pertencem a um conjunto não vazio de hiperplanos ativos, isto é, não são pontos interiores de X . Seja $r(F)$ o número máximo de hiperplanos ativos linearmente independentes em todos os pontos da face F , então se define a dimensão de F da seguinte forma:

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

Em outras palavras, cada hiperplano ativo linearmente independente produz a perda de um grau de liberdade. Esse fato produz as seguintes consequências:

1. Um ponto extremo é uma face própria de dimensão zero porque tem n hiperplanos ativos linearmente independentes.
2. Uma aresta é uma face própria de dimensão 1 porque tem $(n - 1)$ hiperplanos ativos linearmente independentes. Assim, **uma aresta é uma face própria de um grau de liberdade porque tem somente $(n - 1)$ hiperplanos ativos linearmente independentes** (um hiperplano a menos que um ponto extremo). Portanto, uma aresta é um conjunto de pontos que apresentam $(n - 1)$ hiperplanos ativos linearmente independentes, isto é, em todos os pontos de uma aresta existem $(n - 1)$ hiperplanos ativos linearmente independentes.
3. A $\dim(X) = n - r(F) = n - 0 = n \rightarrow$ apresenta dimensão completa porque neste caso temos o conjunto completo X e deve existir pelo menos um ponto em que nenhum hiperplano se encontra ativo (um ponto interior de X). Assim, o conjunto X e o conjunto vazio são chamados de faces impróprias de X .
4. A face própria de maior dimensão que assume a seguinte forma:

$$\dim(F) = \dim(X) - 1 = n - 1$$

é chamada de faceta de X e representa um tópico muito importante em problemas de programação linear inteira.

5. **Pontos extremos adjacentes:** Dois pontos extremos de X são chamados de adjacentes se o segmento de reta que os une é uma aresta. Nesse contexto, dois pontos extremos adjacentes apresentam $(n - 1)$ hiperplanos ativos linearmente independentes comuns.

A Figura 2.12 mostra as diferentes faces próprias do conjunto convexo poliedral X .

2.4.6. Representação de Conjuntos Convexos Poliedrais

Nesta Subseção é analisada a representação de conjuntos convexos poliedrais. Assim, um tópico a ser abordado é o chamado Teorema Geral da Representação (TGR) que é o tópico mais importante a ser analisado neste Capítulo. Assim, é analisada uma propriedade que afirma que um conjunto convexo poliedral X pode ser completamente representado e, portanto, substituído por uma combinação convexa dos pontos extremos de X mais uma combinação linear não negativa das direções extremas de X . Assim, inicialmente definimos o que significa combinação convexa e combinação linear positiva. Um ponto $x \in X$ no espaço E^n é representado por uma combinação convexa de k pontos $x_1, x_2, \dots, x_k$ se são verdadeiras as seguintes relações:

$$x = \sum_{j=1}^k \lambda_j x_j \quad \sum_{j=1}^k \lambda_j = 1 \quad \lambda_j \geq 0$$

Também, um ponto $x \in X$ no espaço E^n é representado por uma combinação linear não negativa de p pontos $d_1, d_2, \dots, d_p$ se são verdadeiras as seguintes relações:

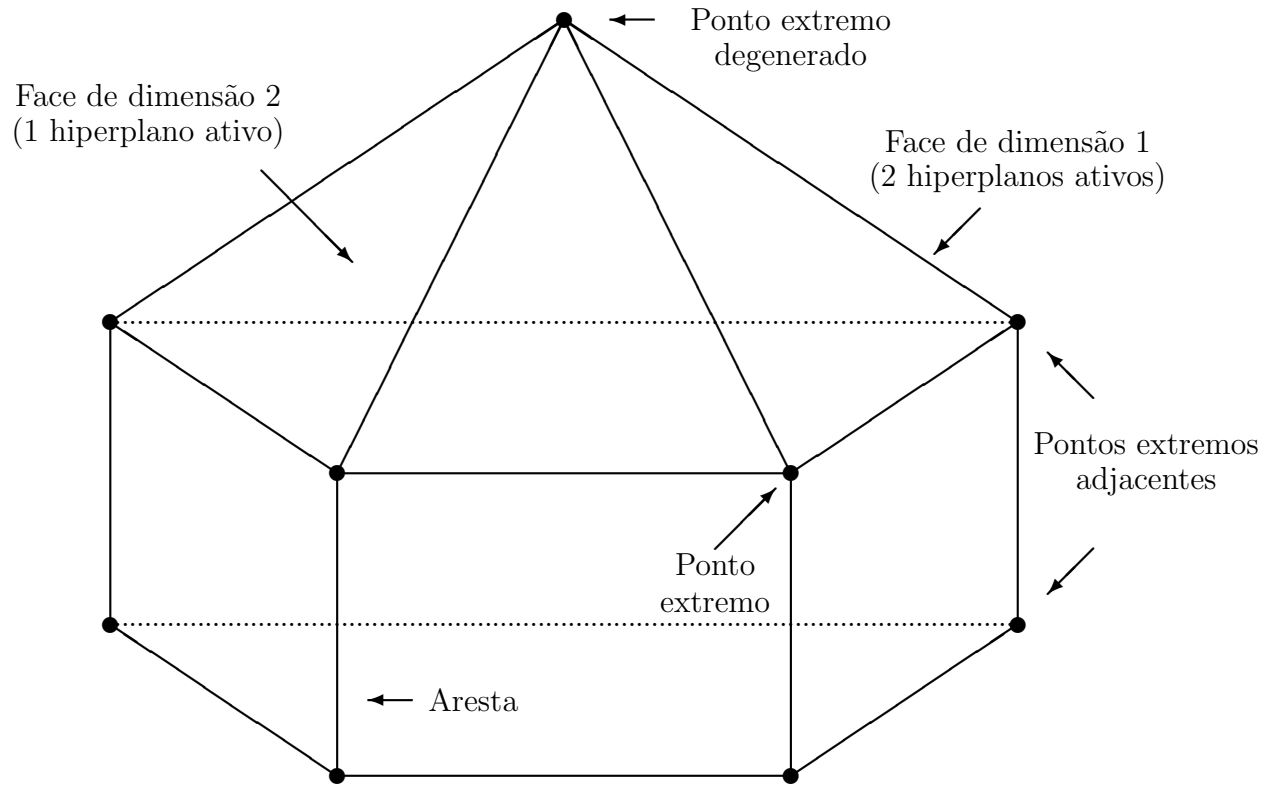


Figura 2.12: Faces próprias do conjunto convexo poliedral X .

$$x = \sum_{j=1}^k \mu_j d_j \quad \mu_j \geq 0$$

Caso Especial: Conjunto Poliedral Limitado ou Politopo

Existe uma importante propriedade, que pode ser provada como um caso particular do TGR, que afirma que qualquer ponto $x \in X$ de um conjunto convexo poliedral limitado pode ser **completamente representado** por uma **combinação convexa** de seus pontos extremos. Essa representação convexa não é única. Essa propriedade é ilustrada na Figura 2.13.

Na Figura 2.13 vamos tentar representar o ponto genérico x como uma combinação convexa dos pontos extremos x_1, x_2, x_3, x_4 e x_5 (não necessariamente usaremos todos eles).

Representamos inicialmente o ponto x como uma combinação convexa dos pontos x_4 e y :

$$x = \lambda y + (1 - \lambda)x_4 \quad \lambda \in [0, 1]$$

Entretanto, também temos o seguinte:

$$y = \mu x_1 + (1 - \mu)x_2 \quad \mu \in [0, 1]$$

Substituindo y na primeira relação temos o seguinte:

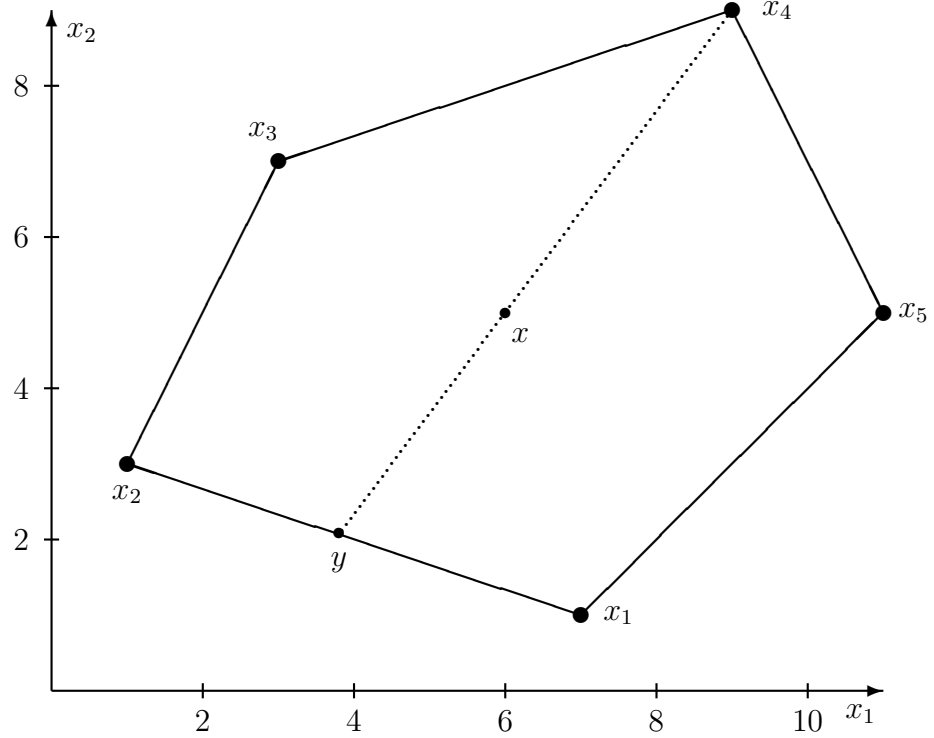


Figura 2.13: Representação de um conjunto convexo poliedral limitado.

$$x = \lambda[\mu x_1 + (1 - \mu)x_2] + (1 - \lambda)x_4 \quad \mu \in [0, 1]; \quad \lambda \in [0, 1]$$

$$x = (\lambda \mu) x_1 + [\lambda(1 - \mu)]x_2 + (1 - \lambda)x_4 \quad \mu \in [0, 1]; \quad \lambda \in [0, 1]$$

que é uma representação de x através de uma combinação convexa dos pontos extremos x_1 , x_2 e x_4 . Os coeficientes são valores entre 0 e 1 e a soma de coeficiente é igual a 1. Assim verificamos facilmente que $(\lambda \mu) \in [0, 1]$; $[\lambda(1 - \mu)] \in [0, 1]$ e $(1 - \lambda) \in [0, 1]$. Adicionalmente, temos o seguinte:

$$\lambda \mu + \lambda(1 - \mu) + (1 - \lambda) = 1$$

o que caracteriza uma combinação convexa. Devemos observar que essa representação não é única. Assim, por exemplo, se escolhermos um y no segmento de reta que une x_1 e x_5 , então x pode ser representado como uma combinação convexa dos pontos extremos x_1 , x_3 e x_5 .

Caso Especial: Conjunto Poliedral Ilimitado

Existe também uma propriedade, que pode ser provada como um caso particular do TGR, que afirma que qualquer ponto $x \in X$ de um conjunto convexo poliedral ilimitado pode ser **completamente representado** por uma **combinação convexa** de seus pontos extremos e mais uma combinação linear não negativa de suas direções extremas. Essa representação convexa não é única. Essa propriedade é ilustrada na Figura 2.14.

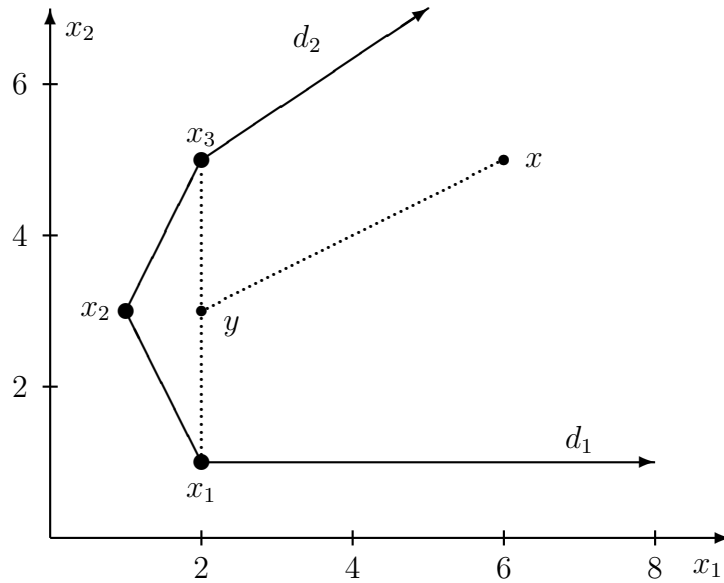


Figura 2.14: Representação de um ponto em um conjunto convexo poliedral ilimitado.

Na Figura 2.14 existem 3 pontos extremos, x_1 , x_2 , x_3 , e duas direções extremas, d_1 e d_2 (não necessariamente usaremos todos eles). Assim, o ponto x pode ser representado apenas por uma direção paralela a d_2 e pelo ponto y que é uma combinação convexa dos pontos extremos x_1 e x_3 . Assim, temos o seguinte:

$$x = y + \mu d_2 = \lambda x_1 + (1 - \lambda)x_3 + \mu d_2$$

onde $\lambda \in [0, 1]$ e $\mu \geq 0$.

Também, essa representação de x não é única. Assim, pode-se encontrar outra representação combinando as direções d_1 e d_2 da seguinte forma:

$$x = \lambda x_1 + (1 - \lambda)x_3 + \mu_1 d_1 + \mu_2 d_2 \quad \lambda \in [0, 1]; \quad \mu_1 > 0; \quad \mu_2 > 0$$

Caso Geral: O Teorema Geral da representação - TGR

O Teorema Geral da Representação é um dos tópicos mais importantes relacionados com técnicas de solução de problemas de programação linear. Entre suas aplicações mais importantes podem ser citadas as seguintes: (1) permite provar as condições nas quais um problema de programação linear é ilimitado, (2) permite provar de que se existe um ótimo finito para um PL, então um ponto extremo é ótimo de um PL e, (3) permite encontrar uma grande família de problemas lineares equivalentes em que o modelo de PL usado na decomposição matemática de Dantzig-Wolfe é um dos mais importantes e conhecidos.

Teorema 2.3: O Teorema Geral da Representação - TGR

Seja $X = \{x : A x \leq b; \quad x \geq 0\}$ um conjunto convexo poliedral não vazio. Nesse contexto, o conjunto de pontos extremos definido por $\{x_1, x_2, \dots, x_k\}$ é não vazio e finito. Além disso, o conjunto de direções extremas é vazio se e somente se X é limitado. Se X não é limitado, então o conjunto de direções extremas $\{d_1, d_2, \dots, d_l\}$ é não vazio e finito. Então, nesse contexto, um

ponto $\bar{x} \in X \iff \bar{x}$ pode ser representado como uma combinação convexa dos pontos extremos $\{x_1, x_2, \dots, x_k\}$ mais uma combinação linear não negativa das direções extremas $\{d_1, d_2, \dots, d_l\}$, isto é,

$$\bar{x} \in X \implies \begin{cases} \bar{x} = \sum_{j=1}^k \lambda_j x_j + \sum_{j=1}^l \mu_j d_j \\ \sum_{j=1}^k \lambda_j = 1 \\ \lambda_j \geq 0 \quad j = 1, 2, \dots, k \\ \mu_j \geq 0 \quad j = 1, 2, \dots, l \end{cases} \quad (2.11)$$

Especificamos o conjunto de pontos extremos: $S_p = \{x_1, x_2, \dots, x_k\}$.

Especificamos também o conjunto de direções extremas: $S_d = \{d_1, d_2, \dots, d_l\}$.

1. Provamos que o número de pontos extremos é finito e não vazio.

Provamos que o número de pontos extremos é finito e não vazio, isto é, $1 \leq k < \infty$ onde k é o número de pontos extremos (os elementos de $S_p \implies k = |S_p|$).

- a) Provamos que existe pelo menos um ponto extremo: $1 \leq k \implies k \geq 1$. Seja $\bar{x} \in X$. Nesse contexto, se $\bar{x} \in S_p$, então existe pelo menos um ponto extremo, isto é, o próprio $\bar{x}$, então $k \geq 1$.

Agora supor que $\bar{x} \notin S_p$.

Seja r o número máximo de hiperplanos ativos linearmente independentes (HALI) em $\bar{x}$. Esse ponto pode ser representado na forma $\bar{x} = \lambda y_1 + (1 - \lambda)y_2$ onde $0 < \lambda < 1$ e $y_1, y_2 \in X$ com $y_1 \neq y_2$. Veja que apenas estamos representando $\bar{x}$ como uma combinação convexa de dois pontos diferentes de X . Observe, adicionalmente, que $0 \leq r < n$ (zero para o caso em que $\bar{x}$ é um ponto interior de X onde não existe HALI e $r < n$ já que $r = n$ significaria que $\bar{x}$ é um ponto extremo).

Seja $d = y_2 - y_1 \neq 0$ e, nesse contexto, é possível provar que as seguintes relações são verdadeiras:

$$y_1 = \bar{x} - (1 - \lambda)d \quad \text{e} \quad y_2 = \bar{x} + \lambda d \quad (2.12)$$

As relações anteriores podem ser provadas da seguinte forma:

$$\bar{x} = \lambda y_1 + (1 - \lambda)y_2 = \lambda y_1 - y_1 + y_1 + (1 - \lambda)y_2 = -(1 - \lambda)y_1 + (1 - \lambda)y_2 + y_1$$

$$\bar{x} = (y_2 - y_1)(1 - \lambda) + y_1 \implies y_1 = \bar{x} - (1 - \lambda)d$$

$$\bar{x} = \lambda y_1 + (1 - \lambda)y_2 = \lambda y_1 - \lambda y_2 + y_2 = -\lambda(y_2 - y_1) + y_2 = -\lambda d + y_2 \implies y_2 = \bar{x} + \lambda d$$

onde em ambos os casos foi usada a relação $d = y_2 - y_1$.

Agora consideremos deslocamentos de $\bar{x}$ nas direções e sentidos de d e $-d$ com a pretensão de encontrar um hiperplano ativo sem sair de X . Ambos deslocamentos são possíveis e de tamanho finito sem sair de X . Escolhendo a direção d ou a direção $-d$ sempre é possível encontrar um hiperplano ativo e como $X \subseteq \{x : x \geq 0\}$, então no pior caso, o tamanho do passo é limitado pelos hiperplanos definidos por $x \geq 0$. Assim, sem perda de generalidade, supor que esse deslocamento para encontrar um hiperplano ativo acontece na direção $-d$, isto é, o deslocamento é na direção onde se encontra y_1 . Assim, temos o seguinte:

$$\bar{\gamma} = \max\{\gamma : y_1 = \bar{x} - \gamma d \in X < \infty\}$$

e calculamos $\bar{y}_1 = \bar{x} - \bar{\gamma}d$. Veja que $\bar{y}_1$ é um ponto encontrado com o máximo deslocamento possível sem sair de X o que significa que esse deslocamento para encontrar $\bar{y}_1$ termina quando se atinge um ou vários hiperplanos. Assim, podemos afirmar que o número máximo de hiperplanos ativos LI em $\bar{y}_1$ é $\bar{r} \geq r + 1$ porque todos os hiperplanos ativos em $\bar{x}$ continuam ativos em $\bar{y}_1$ (lembramos que $\bar{x}$ é uma combinação convexa estrita de $\bar{y}_1$ e y_2 e, além disso, pelo menos um hiperplano adicional linearmente independente está ativo em $\bar{y}_1$, isto é, aquele hiperplano que limita o incremento de γ na direção $-d$ ao encontrar o valor de $\bar{\gamma}$).

Portanto, se $\bar{r} = n \implies \bar{y}_1 \in S_p$ e, dessa forma, encontramos um ponto extremo e $k \geq 1$. Por outro lado, se $\bar{r} < n$, então substituir $\bar{x}$ por $\bar{y}_1$, isto é $\bar{x} = \bar{y}_1$, e repetir o processo até que $\bar{r} = n$ o que deve acontecer em um número finito de passos porque $\bar{r}$ aumenta em cada passo pelo menos em uma unidade e porque n é finito.

b) Provamos que $k < \infty$:

Como existem $(n + m)$ candidatos a hiperplanos ativos em X e um ponto extremo é definido por n hiperplanos ativos LI, então o número máximo de pontos extremos é calculado usando a seguinte relação:

$$N_{max} = \binom{m+n}{n} = \frac{(m+n)!}{n! m!} < \infty$$

Portanto, como N_{max} é finito, então $k \leq N_{max} < \infty$ o que prova esta parte do teorema.

2. Agora provamos que se X é não limitado, então o conjunto de direções extremas é não vazio e finito, isto é, $1 \leq l < \infty$.

Se X é limitado então, por definição, o conjunto de direções extremas ou, melhor ainda, o conjunto de direções de X é vazio. Assim, temos o seguinte:

$$D = \phi$$

Se X não é limitado, então $D \neq \phi \implies 1 \leq l$.

Observemos também que D tem a mesma estrutura de X e que os pontos extremos de D normalizado são também as direções extremas de X e como o número de pontos extremos de D é finito (já que tem a mesma estrutura de X), então o número de direções extremas de X também deve ser finito e, portanto, temos o seguinte:

$$l < \infty \implies 1 \leq l < \infty$$

3. Agora provamos que se $\bar{x} \in X$, então $\bar{x}$ pode ser escrito na forma apresentada em (2.11).

Inicialmente definimos um novo conjunto $\bar{X}$ da seguinte forma:

$$\bar{X} = X \cap \{x : \sum_{j=1}^n x_j \leq M\}$$

onde M é um número suficientemente grande de forma que não exclua cada ponto extremo x_j , $j = 1, 2, \dots, k$ e também não exclua o ponto $\bar{x}$ em análise.

Observe que, nesse contexto, $\bar{X}$ é limitado e os pontos extremos de X são também pontos extremos de $\bar{X}$. Seja $\bar{S}_p = \{x_1, x_2, \dots, x_k, x_{k+1}, \dots, x_{k+u}\}$ os pontos extremos de $\bar{X}$ onde $0 \leq u < \infty$.

a) Primeiro provaremos que $\bar{x}$ pode ser escrito como uma combinação convexa dos pontos extremos de $\bar{X}$, isto é, por pontos de $\bar{S}_p$.

Se $\bar{x} \in S_p \implies \bar{x}$ é um ponto extremo de $\bar{X}$, já que continuamos no espaço E^n .

Se $\bar{x} \notin S_p$, então representamos os hiperplanos ativos de X em $\bar{x}$ da seguinte forma:

$$G x = g$$

Neste caso, $Posto[G] \leq (n - 1)$ porque em caso contrário $\bar{x}$ seria um ponto extremo de $\bar{X}$. Agora, encontramos uma solução para o sistema $G d = 0$, $d \neq 0$ e, a partir desse resultado, encontramos a seguinte relação:

$$\bar{\gamma}_1 = \max\{\gamma : y_1 = \bar{x} + \gamma d \in \bar{X}\}$$

veja que, no contexto anterior, d é uma direção que tem os mesmos hiperplanos ativos existentes em $G x = g$. Assim, y_1 é um ponto a partir de $\bar{x}$ e que preserva esses hiperplanos ativos. Também, observe que $0 < \bar{\gamma}_1 < \infty$ porque $\bar{X}$ é limitado. Também, devemos observar que γ aumenta de valor até que encontramos um hiperplano ativo. Assim, para $\bar{\gamma}_1$ temos $\bar{y}_1 = \bar{x} + \bar{\gamma}_1 d$. Portanto, em $\bar{y}_1 \in \bar{X}$ temos pelo menos um hiperplano ativo adicional linearmente independente de $\bar{X}$ em relação a $\bar{x}$. Se este novo hiperplano ativo em $\bar{y}_1$ junto com o sistema $G x = g$ produzem um sistema de $Posto = n$, então $\bar{y}_1$ é um ponto extremo de $\bar{X}$. Em caso contrário, se o posto ainda não é igual a n , então repetir o passo usado para encontrar $\bar{y}_1$ e, no pior caso, após $[n - Posto[G]]$ iterações encontramos um ponto extremo de $\bar{X}$, que seria um novo ponto $\bar{y}_1 \in \bar{X}$ em que existem n HALI e o posto de G no sistema $G \bar{y}_1 = g$ é igual a n .

Por outro lado, definimos um parâmetro $\bar{\gamma}_2$ da seguinte forma:

$$\bar{\gamma}_2 = \max\{\gamma : y_2 = \bar{x} + \gamma(\bar{x} - \bar{y}_1)\}$$

que deve gerar o seguinte ponto:

$$\bar{y}_2 = \bar{x} + \bar{\gamma}_2(\bar{x} - \bar{y}_1)$$

Observemos que $\bar{\gamma}_2 < \infty$ porque $\bar{X}$ é limitado e $\bar{\gamma}_2$ é o máximo permitido para encontrar um hiperplano ativo. Também, $\bar{\gamma}_2 > 0$ já que propositalmente foi escolhida uma direção de forma que se acompanhe a direção $(\bar{x} - \bar{y}_1)$ com um γ positivo e também não pode ser igual a zero já que $G(\bar{x} + \gamma(\bar{x} - \bar{y}_1)) = g$ para $\gamma \neq 0$ foi resolvido de forma a encontrar um $\bar{\gamma}_2$ que foi limitado por um novo hiperplano que não se encontrava ativo em $\bar{x}$. Assim, a nova solução $G(\bar{y}_2) = g$ tem pelo menos um hiperplano ativo LI adicional em $\bar{y}_2$ em relação a $\bar{x}$.

Resumindo, $\bar{y}_1$ é um ponto extremo de $\bar{X}$ e $\bar{y}_2$ tem pelo menos um hiperplano ativo LI adicional em relação a $\bar{x}$. Além disso, $\bar{x}$ é uma combinação convexa de $\bar{y}_1$ e $\bar{y}_2$ como pode ser provado da seguinte forma:

$$\bar{y}_2 = \bar{x} + \bar{\gamma}_2(\bar{x} - \bar{y}_1) \implies \bar{x} = \left(\frac{\bar{\gamma}_2}{1 + \bar{\gamma}_2} \right) \bar{y}_1 + \left(\frac{1}{1 + \bar{\gamma}_2} \right) \bar{y}_2$$

que pode ser facilmente verificado que é uma combinação convexa já que $\bar{\gamma}_2 > 0$ e a soma dos coeficientes de $\bar{y}_1$ e $\bar{y}_2$ é igual a 1.

Portanto, o ponto $\bar{x}$ em que existem r HALI foi representado por uma combinação convexa dos pontos $\bar{y}_1$ e $\bar{y}_2$, sendo que $\bar{y}_1$ é um ponto extremo de $\bar{X}$ e em $\bar{y}_2$ existem pelo menos $(r + 1)$ HALI. Nesse contexto, podem acontecer dois casos:

- Se $\bar{y}_2 \in \bar{S}_p$, então conseguimos representar $\bar{x}$ como uma combinação convexa de dois pontos extremos, isto é, $\bar{y}_1$ e $\bar{y}_2$.
- Se $\bar{y}_2 \notin \bar{S}_p$:

Nesse caso $\bar{y}_2$ pode ser representado como uma combinação convexa estrita dos pontos $\bar{y}_3$ e $\bar{y}_4$ (repetindo o método para calcular $\bar{y}_1$ e $\bar{y}_2$), onde um desses pontos, por exemplo, $\bar{y}_3$ é um ponto extremo de $\bar{X}$, $\bar{y}_3 \in \bar{S}_p$ e o outro, $\bar{y}_4$ tem pelo menos um HALI adicional em relação a $\bar{y}_2$.

Neste ponto da demonstração $\bar{y}_4$ tem pelo menos 2 HALI adicionais em relação com $\bar{x}$. Continuando este processo, em cada passo substituindo um ponto por outros dois pontos, em que um desses pontos é um ponto extremo e o outro tem um HALI adicional, então em um número determinado e finito de iterações este segundo ponto deve atingir n HALI e, portanto, deve representar um ponto extremo.

Portanto, usando um máximo de $[n - \text{Posto}[G] + 1]$ pontos extremos de $\bar{X}$ terminamos o processo de representação. Essa representação assume a seguinte forma:

$$\bar{x} = \sum_{j=1}^{k+u} \delta_j x_j \quad \text{onde} \quad \sum_{j=1}^{k+u} \delta_j = 1 \quad \text{e} \quad \delta_j \geq 0 \quad (2.13)$$

- b) Agora completamos a prova eliminando o efeito da restrição $x \leq M$ adicionado. isto é, agora devemos provar que (2.13) é verdadeiro para X .

Se $\delta_j = 0$ para $j > k$, então (2.13) é equivalente a (2.11) e, portanto, X é limitado.

Em caso contrário, consideremos que existe algum ponto extremo de (2.13), x_v com $v > k$ e $\delta_v > 0$. Observe que, nesse contexto, x_v é um ponto extremo novo gerado pela restrição $\sum_{j=1}^n x_j \leq M$, isto é, $\sum_{j=1}^n x_j = M$ é um dos n hiperplanos LI que identificam x_v como ponto extremo de $\bar{X}$. Os outros, $(n - 1)$ hiperplanos ativos pertencem ao conjunto X original e identificam uma aresta de X . Consequentemente, existe um

ponto extremo $x_{i(v)}$ de X , $1 \leq i(v) \leq k$ que é um ponto extremo adjacente de x_v em $\bar{X}$ através dessa aresta. Além disso, $(x_v - x_{i(v)})$ é uma direção de X porque não existe outro hiperplano de X que pare o incremento de $x(v)$ nessa direção. Adicionalmente, seja $\bar{d} = \frac{1}{\theta_v}(x_v - x_{i(v)})$ onde θ_v assume um valor tal que $\bar{d}$ seja um vetor unitário. Nesse contexto, $\bar{d} \in D$, isto é, $\bar{d}$ é uma direção extrema. Adicionalmente, $(n-1)$ hiperplanos LI de X que se encontram ativos na caracterização de x_v também se encontram ativos no sistema $A d \leq 0; \quad d \geq 0$ na caracterização de $\bar{d}$. Também, esses $(n-1)$ hiperplanos ativos que caracterizam x_v , junto com o hiperplano $\sum_{j=1}^n d_j = 1$ produzem o conjunto de n hiperplanos ativos LI de D em $\bar{d}$. Portanto, a direção $\bar{d}$ é um ponto extremo de D que chamaremos $d_{j(v)}$ que por sua vez é uma direção extrema de X . Consequentemente, temos que $x_v = x_{i(v)} + \theta_v d_{j(v)}$. Substituindo esta relação em (2.13) temos o seguinte:

$$\bar{x} = \sum_{j=1}^k \delta_j x_j + \sum_{v=k+1}^{k+u} \delta_v x_{i(v)} + \sum_{v=k+1}^{k+u} \delta_v \theta_v d_{j(v)} \quad (2.14)$$

que tem a forma de (2.11).

4. Finalmente, demonstraremos que se $\bar{x}$ pode ser escrito na forma apresentada em (2.11), então $\bar{x} \in X$.

Devemos demonstrar que se $\bar{x}$ pode ser escrito como em (2.11), então $\bar{x}$ cumpre com as seguintes relações:

$$\bar{x} \in X \implies \{\bar{x} : A \bar{x} \leq b; \quad \bar{x} \geq 0\}$$

Se $\bar{x}$ cumpre com (2.11), então temos o seguinte:

$$\left\{ \begin{array}{l} \bar{x} = \sum_{j=1}^k \lambda_j x_j + \sum_{j=1}^l \mu_j d_j \\ \sum_{j=1}^k \lambda_j = 1 \\ \lambda_j \geq 0 \quad j = 1, 2, \dots, k \\ \mu_j \geq 0 \quad j = 1, 2, \dots, l \end{array} \right. \quad (2.15)$$

Sabemos que os x_j são pontos extremos de X e, portanto, cumprem com as seguintes relações:

$$A x_j \leq b \quad x_j \geq 0; \quad j = 1, \dots, k \quad (2.16)$$

Sabemos que os d_j são direções extremas de X (ou também equivalentemente pontos extremos de D) e, portanto, cumprem com as seguintes relações:

$$A d_j \leq 0 \quad d_j \geq 0 \text{ e } d_j \neq 0; \quad j = 1, \dots, l \quad (2.17)$$

a) Provamos que $\bar{x} \geq 0$:

$\bar{x} \geq 0$ é verdadeiro porque analisando (2.15) ou (2.11) podemos verificar o seguinte:

$$\lambda_j \geq 0; \quad x_j \geq 0; \quad \mu_j \geq 0 \quad \text{e} \quad d_j \geq 0 \implies \bar{x} \geq 0$$

b) Provamos que $A \bar{x} \leq b$:

Multiplicando $\bar{x}$ em (2.15) por A pela esquerda temos o seguinte:

$$A \bar{x} = A \left\{ \sum_{j=1}^k \lambda_j x_j + \sum_{j=1}^l \mu_j d_j \right\}$$

$$A \bar{x} = \lambda_1 A x_1 + \lambda_2 A x_2 + \dots + \lambda_k A x_k + \mu_1 A d_1 + \mu_2 A d_2 + \dots + \mu_l A d_l \quad (2.18)$$

mas na relação anterior temos o seguinte:

$$\mu_1 A d_1 + \mu_2 A d_2 + \dots + \mu_l A d_l \leq 0$$

já que $\mu_j \geq 0$ e cada $A d_j \leq 0$; $j = 1, 2, \dots, l$ [ver (2.17)]. Assim, em (2.18) temos o seguinte:

$$A \bar{x} \leq \lambda_1 A x_1 + \lambda_2 A x_2 + \dots + \lambda_k A x_k \quad (2.19)$$

mas em (2.19) cada x_j é ponto extremo de X e, portanto, $x_j \in X$ e para cada x_j é verdadeira a seguinte relação:

$$A x_j \leq b \implies \lambda_j A x_j \leq \lambda_j b; \quad j = 1, 2, \dots, k \quad (2.20)$$

De (2.19) e (2.20) temos o seguinte:

$$A \bar{x} \leq \lambda_1 A x_1 + \lambda_2 A x_2 + \dots + \lambda_k A x_k \leq \lambda_1 b + \lambda_2 b + \dots + \lambda_k b$$

$$A \bar{x} \leq b(\lambda_1 + \lambda_2 + \dots + \lambda_k) = b \sum_{j=1}^k \lambda_j = b[1] = b \implies A \bar{x} \leq b$$

Portanto, como $\bar{x} \geq 0$ e $A \bar{x} \leq b \implies \bar{x} \in X$.

Corolário 2.1: Qualquer $\bar{x} \in X$ pode ser representado na forma apresentada em (2.11) usando no máximo $\min\{(n+1), (k+l)\}$ variáveis positivas de tipo λ_j e μ_j .

Prova: Dado um $\bar{x} \in X$ então, pelo Teorema Geral da Representação, existe uma solução para o sistema (2.11). Seja r o posto da matriz de coeficientes associada com as restrições de igualdade em (2.11) (observe que em (2.11) existem $(n+1)$ restrições de igualdade, $(k+l)$ restrições de desigualdade e $(k+l)$ variáveis). Portanto, o valor de r obedece a seguinte relação:

$$r = \min\{(n+1), (k+l)\} \quad (2.21)$$

Adicionalmente, observemos que o conjunto de variáveis (λ, μ) que satisfazem (2.11) formam um conjunto poliedral no espaço E^{k+l} e, pelo Teorema Geral da Representação, existe pelo menos um ponto extremo. Assim, esse ponto extremo deve ser caracterizado por $(k+l)$ hiperplanos ativos LI e as restrições de igualdade sempre fornecem r hiperplanos ativos LI e, portanto, estão faltando $\{(k+l)-r\}$ hiperplanos ativos LI adicionais e esses hiperplanos devem ser fornecidos pelas variáveis λ_j e μ_j (quando os subespaços tipo $\lambda_j \geq 0$ tem o hiperplano ativo, isto é $\lambda_j = 0$) quando essas variáveis assumem valores iguais a zero. Assim, na caracterização desse ponto extremo $\bar{x}$ de (2.11), pelo menos $(k+l) - r$ das variáveis λ_j e μ_j devem ser iguais a zero. Portanto, o número máximo de variáveis de tipo λ_j e μ_j que podem ser positivas (e portanto participam na representação de $\bar{x} \in X$) é encontrado através da seguinte relação:

$$\{(k+l) - [(k+l) - r]\} = r$$

Como $r = \min\{(n+1), (k+l)\}$ então o Corolário 2.1 está provado.

Observações: As seguintes observações são muito importantes:

1. Se X é um conjunto convexo limitado então $l = 0$ e de (2.21) temos o seguinte:

$$r = \min\{(n+1), k\}$$

que indica que um ponto $\bar{x} \in X$ pode ser representado por não mais de que $\min\{(n+1), k\}$ de seus pontos extremos. Isso significa que, por exemplo, no espaço E^2 com $n = 2$, um ponto $\bar{x} \in X$ pode ser representado por no máximo 3 pontos extremos. Assim, se 3 pontos extremos formam os vértices de um triângulo então esses 3 pontos extremos podem representar todos os pontos dentro desse triângulo.

Como ilustração adicional, supor que em E^2 o conjunto X tem 5 pontos extremos. Nesse contexto, qualquer $\bar{x} \in X$ pode ser representado por não mais de que $\min\{3, 5\} = 3$ de seus pontos extremos.

2. Se X é um conjunto convexo ilimitado, então um $\bar{x} \in X$ pode ser representado por não mais que $(n+1)$ pontos extremos e n direções extremas. Em E^n podem existir mais de n direções extremas, mas para representar um $\bar{x} \in X$ são necessárias apenas n direções extremas já que na representação devemos usar pelo menos um ponto extremo.

O Teorema Geral da Representação é o tópico mais importante para provar o funcionamento do método simplex, analisado em detalhe no próximo capítulo, para resolver problemas de programação linear. Entretanto, os conceitos usados na prova podem ser usados para outras provas e outras aplicações como, por exemplo, para montar uma estratégia que permite encontrar um ponto extremo a partir de um ponto interior.

2.4.7. Estratégia para Encontrar um Ponto Extremo a Partir de um Ponto Interior

A partir de um ponto interior de $\bar{x} \in X$ é possível encontrar um ponto extremo de X usando o seguinte algoritmo:

1. Seja $\bar{x}$ um ponto interior de X . Adicionar ao conjunto X a restrição $\sum_{j=1}^n x_j \leq M$ onde M é um número suficientemente grande para evitar problemas com conjuntos ilimitados. Assim, esse novo conjunto é chamado de $\bar{X}$.
2. Verificar o número de hiperplanos ativos no ponto $\bar{x} \in \bar{X}$:

a) Se nenhum hiperplano está ativo então escolher uma direção arbitrária:

$$d = [d_1, d_2, \dots, d_n]$$

b) Se existem hiperplanos ativos então encontrar a direção d usando a relação:

$$G d = 0$$

onde G é uma matriz constituída pelos coeficientes das restrições do problema que se encontram com os hiperplanos ativos, isto é, uma parcela das restrições $A x \leq b$ e $x \geq 0$ cujos hiperplanos se encontram ativos.

3. Encontrar o parâmetro:

$$\bar{\gamma}_1 = \max\{\gamma_1 : x = \bar{x} + \gamma_1 d \in \bar{X}; \quad \gamma_1 > 0\}$$

Se todos os $\gamma_1 \leq 0$ então encontrar $\bar{\gamma}_1$ da seguinte forma:

$$\bar{\gamma}_1 = \max\{\gamma_1 : x = \bar{x} - \gamma_1 d \in \bar{X}; \quad \gamma_1 > 0\}$$

4. O ponto $\bar{x}' = \bar{x} + \gamma_1 d$ (ou $\bar{x}' = \bar{x} - \gamma_1 d$ com a segunda possibilidade) tem um hiperplano ativo adicional em relação a $\bar{x}$.
5. Fazer $\bar{x} = \bar{x}'$ e repetir os passos 2, 3 e 4 até que o ponto $\bar{x}$ tenha n hiperplanos ativos (sem levar em conta $x_1 + x_2 + \dots + x_n \leq M$), isto é, sem levar em conta o subespaço artificial adicionado.

É muito importante observar que no Passo 3 estamos procurando o γ_1 que permita que o ponto x encontre um hiperplano ativo, isto é, o avanço de x termina quando encontra um novo hiperplano ativo.

Exemplo 3.16: No seguinte conjunto convexo:

$$X = \{(x_1, x_2) : \begin{array}{l} 3x_1 - x_2 \geq 2 \\ -x_1 + x_2 \leq 2 \\ -x_1 + 2x_2 \leq 8 \\ x_2 \geq 2 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{array}\}$$

encontrar um ponto extremo a partir do ponto interior inicial $\bar{x} = (4, 3)$.

Usamos o algoritmo, passo a passo, para encontrar um ponto extremo de X a partir do ponto inicial conhecido.

1. O conjunto $\bar{X}$ assume a seguinte forma:

$$\bar{X} = \{(x_1, x_2) : x \in X; \quad x_1 + x_2 \leq M\}$$

2. O ponto $\bar{x} = (4, 3)$ é um ponto interior de $\bar{X}$ porque não possui nenhum hiperplano ativo. Assim, escolhemos a seguinte direção arbitrária:

$$d = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

3. Encontrando $\bar{\gamma}_1$:

O novo ponto procurado apresenta a seguinte estrutura:

$$x = \bar{x} + \gamma_1 d = \begin{bmatrix} 4 \\ 3 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix} \gamma_1 = \begin{bmatrix} 4 - \gamma_1 \\ 3 + 2\gamma_1 \end{bmatrix}$$

Substituindo esse valor de x nas sete restrições temos o seguinte:

Restrição 1	$3(4 - \gamma_1) - (3 + 2\gamma_1) = 2 \implies \gamma_1 = \frac{7}{5}$	
Restrição 2	$-(4 - \gamma_1) + (3 + 2\gamma_1) = 2 \implies \gamma_1 = 1$	
Restrição 3	$-(4 - \gamma_1) + 2(3 + 2\gamma_1) = 8 \implies \gamma_1 = \frac{6}{5}$	
Restrição 4	$(3 + 2\gamma_1) = 2 \implies \gamma_1 = -\frac{1}{2}$	Não interessa
Restrição 5	$(4 - \gamma_1) = 0 \implies \gamma_1 = 4$	
Restrição 6	$(3 + 2\gamma_1) = 0 \implies \gamma_1 = -\frac{3}{2}$	Não interessa
Restrição 7	$(4 - \gamma_1) + (3 + 2\gamma_1) = M \implies \gamma_1 = M - 7$	É muito grande

Assim, o máximo γ_1 permitido é $\bar{\gamma}_1 = 1$.

4. O seguinte ponto:

$$\bar{x}' = \bar{x} + \bar{\gamma}_1 d = \begin{bmatrix} 4 - \bar{\gamma}_1 \\ 3 + 2\bar{\gamma}_1 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

apresenta um hiperplano ativo (o hiperplano da segunda restrição).

5. O ponto:

$$\bar{x} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

apresenta um hiperplano ativo.

2. Encontramos a direção d usando a relação:

$$G d = 0 \implies \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0 \implies -d_1 + d_2 = 0$$

$$d_1 = d_2 \implies d = \begin{bmatrix} d_1 \\ d_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{foi escolhido qualquer } d_1$$

3. Encontrando $\bar{\gamma}_1$:

O novo ponto procurado apresenta a seguinte estrutura:

$$x = \bar{x} + \gamma_1 d = \begin{bmatrix} 3 \\ 5 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \gamma_1 = \begin{bmatrix} 3 + \gamma_1 \\ 5 + \gamma_1 \end{bmatrix}$$

Substituindo esse valor de x nas sete restrições temos o seguinte:

Restrição 1	$3(3 + \gamma_1) - (5 + \gamma_1) = 2 \implies \gamma_1 = -1$	Não interessa
Restrição 2	$-(3 + \gamma_1) + (5 + \gamma_1) = 2 \implies \gamma_1 = 0$	Está ativa
Restrição 3	$-(3 + \gamma_1) + 2(5 + \gamma_1) = 8 \implies \gamma_1 = 1$	
Restrição 4	$(5 + \gamma_1) = 2 \implies \gamma_1 = -3$	Não interessa
Restrição 5	$(3 + \gamma_1) = 0 \implies \gamma_1 = -3$	Não interessa
Restrição 6	$(5 + \gamma_1) = 0 \implies \gamma_1 = -5$	Não interessa
Restrição 7	$(3 + \gamma_1) + (5 + \gamma_1) = M \implies \gamma_1 = \frac{1}{2}M - 7$	É muito grande

Assim, o máximo γ_1 permitido é $\bar{\gamma}_1 = 1$.

4. O seguinte ponto:

$$\bar{x}' = \bar{x} + \bar{\gamma}_1 d = \begin{bmatrix} 3 + \bar{\gamma}_1 \\ 5 + \bar{\gamma}_1 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

apresenta dois hiperplanos ativos (os hiperplanos da segunda e da terceira restrição).

5. O ponto:

$$\bar{x} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

apresenta dois hiperplanos ativos e, portanto, representa um ponto extremo de X .

A Figura 2.15 mostra a representação gráfica do exemplo resolvido. O Leitor está convidado a repetir a solução do exemplo anterior seguindo a seguinte estratégia: (1) a partir do ponto (4, 3), mas escolhendo uma direção inicial (0, 1) e, (2) a partir do ponto (6, 5).

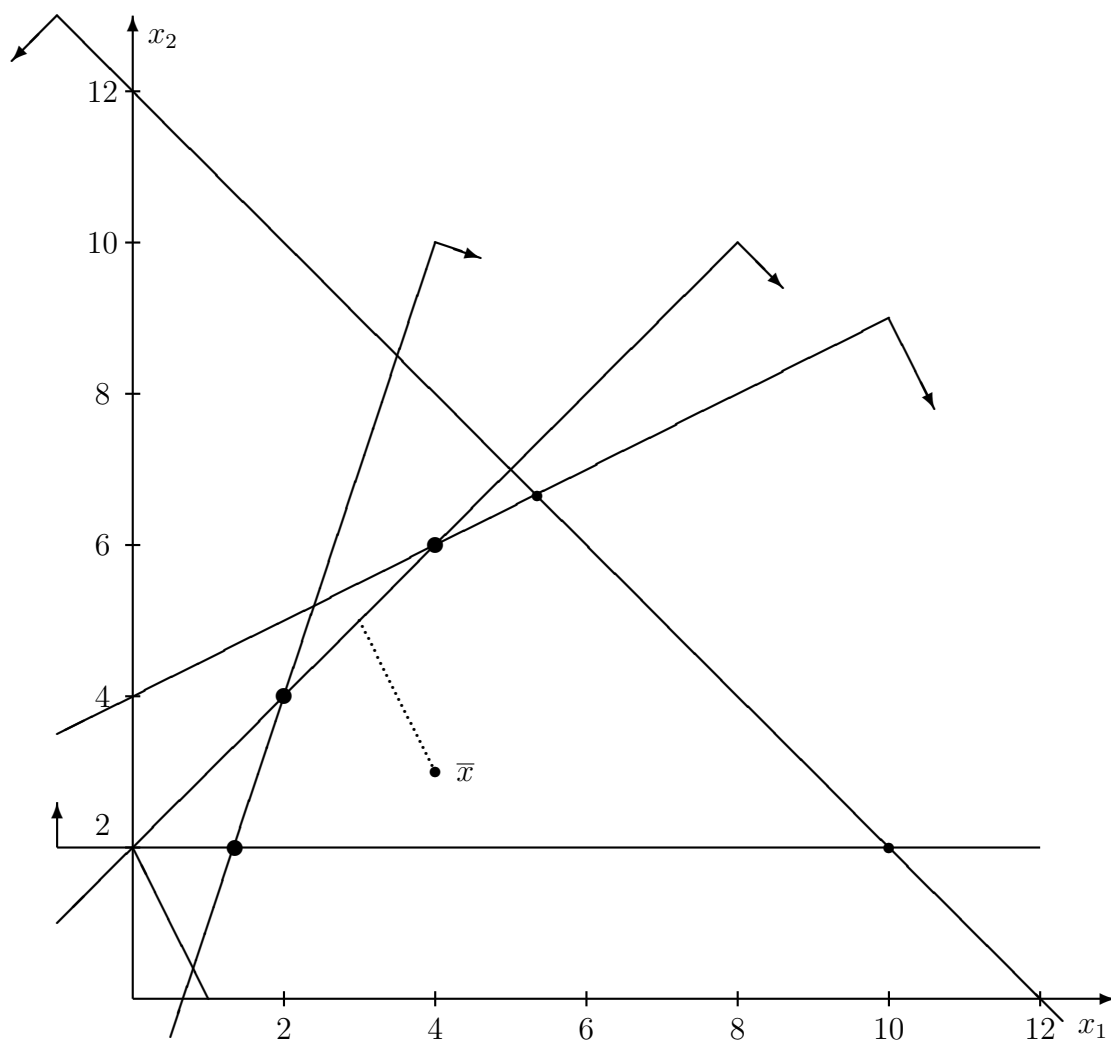


Figura 2.15: Encontrando um ponto extremo a partir de um ponto interior.

2.4.8. Funções Convexas e Côncavas

As propriedades das funções convexas não são fundamentais na teoria de programação linear. Entretanto, esses conceitos são fundamentais na teoria de programação não linear. Em programação linear a função objetivo é linear e cada restrição pode ser representada como uma função linear e, portanto, nesse contexto as propriedades das funções convexas ou côncavas se tornam pouco relevantes.

Uma função $f(\cdot)$ no ponto $x = (x_1, x_2, x_3, \dots, x_n)$ é convexa se é verdadeira a seguinte relação para dois pontos genéricos x_1 e x_2 :

$$f[\lambda x_1 + (1 - \lambda)x_2] \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \quad \forall \lambda \in [0, 1] \quad (2.22)$$

Assim, o valor da função $f(\cdot)$ no ponto $x = \lambda x_1 + (1 - \lambda)x_2$ que é uma combinação convexa de dois pontos x_1 e x_2 é menor ou igual que a combinação convexa entre $f(x_1)$ e $f(x_2)$.

Quando $f(\cdot)$ é uma função de uma única variável x então pode ser demonstrado que $\lambda f(x_1) + (1 - \lambda)f(x_2)$; $\forall \lambda \in [0, 1]$ corresponde aos pontos de uma corda que une $f(x_1)$ e $f(x_2)$. Assim, em uma função convexa, a corda tem uma altura maior ou igual ao valor de $f(x)$ (se encontra sempre acima de $f(x)$).

Uma função $f(\cdot)$ é côncava se e somente se $-f(x)$ é convexa. Neste caso, é válida a seguinte relação:

$$f[\lambda x_1 + (1 - \lambda)x_2] \geq \lambda f(x_1) + (1 - \lambda)f(x_2) \quad \forall \lambda \in [0, 1] \quad (2.23)$$

para dois pontos genéricos x_1 e x_2 .

Existem várias formas para demonstrar se uma função $f(\cdot)$ é convexa. Entre as principais podem ser citadas as seguintes:

1. Usando a definição apresentada em (2.22). Assim, deve-se escolher dois pontos genéricos x_1 e x_2 e demonstrar que a relação (2.22) é verdadeira.
2. Demonstrando que o epígrafe de $f(x)$ é um conjunto convexo.
3. Demonstrando que o Hessiano de $f(x)$, $H(x)$, é uma matriz semi-definida positiva.

Finalmente, deve-se observar que todas as funções existentes em programação linear são lineares e, pode-se demonstrar facilmente que uma função linear é convexa e côncava ao mesmo tempo.