

Ruana Máira Schneider

Método simplex para programação linear

Florianópolis

2013

Ruana Maíra Schneider

Método simplex para programação linear

Orientador:
Melissa Weber Mendonça

UNIVERSIDADE FEDERAL DE SANTA CATARINA

Florianópolis

2013

Esta monografia foi julgada adequada como **TRABALHO DE CONCLUSÃO DE CURSO** no Curso de Matemática — Habilitação Licenciatura e aprovada em sua forma final pela Banca Examinadora designada pela Portaria no 20CC 210

Prof. Nereu Estanislau Burin
Professor da disciplina

Banca Examinadora:

Prof. Melissa Weber Mendonça
Orientador

Prof. Clóvis Caesar Gonzaga

Prof. Juliano de Bem Francisco

Sumário

Resumo

Introdução p. 5

1 Resultados da álgebra linear p. 6

2 Descrição dos Problemas de Programação Linear p. 11

2.1 Soluções básicas p. 15

2.2 Relações de convexidade p. 19

3 Método Simplex p. 25

3.1 Procedimento computacional do método simplex p. 31

3.2 As duas fases do método simplex p. 36

3.3 Número de iterações p. 38

3.4 Método simplex revisado p. 44

4 Dualidade e problema dual p. 49

4.1 Método simplex dual p. 53

Conclusões e perspectivas p. 57

Referências Bibliográficas p. 59

Resumo

Palavras-chave: programação linear, método simplex, dualidade.

Um problema de programação linear (PL) é um problema de otimização cuja função a ser minimizada ou maximizada é linear bem como o seu conjunto de restrições (relações de interdependência entre as variáveis). As técnicas de resolução de problemas de PL são amplamente utilizadas principalmente na Física, Engenharia e Economia.

É possível mostrar que se o conjunto das soluções viáveis de um problema de PL, ou seja, o conjunto de soluções em que as variáveis assumem valores positivos e satisfazem todas as restrições, for não vazio e a função objetivo for limitada inferiormente neste conjunto, ele será um conjunto convexo e fechado. Esse conjunto possuirá um número finito de pontos extremos e a solução ótima do problema corresponderá a um ponto extremo deste conjunto. Desse modo, o método simplex, proposto por Georges Dantzig [3] em 1947, é um procedimento matricial que percorre esses pontos extremos em busca da solução ótima. O algoritmo utiliza um critério de busca de forma que a solução seguinte seja sempre “melhor” do que a anterior. Como o conjunto possui um número finito de pontos extremos isso nos garante que o algoritmo termina em algum ponto, que será a solução ótima do problema. Existem “versões melhoradas” desse método que obtêm uma solução em menos tempo computacional. O método simplex revisado, por exemplo, é um esquema que ordena os cálculos evitando operações desnecessárias, de forma a minimizar o tempo da computação.

Introdução

A *Programação linear*, utilizada para solucionar uma classe de problemas de Programação Matemática, é um dos ramos mais desenvolvidos e mais utilizados da *Pesquisa Operacional*. As técnicas de programação linear, apesar de possuírem uma estrutura relativamente simples, são amplamente utilizadas em diversas áreas, como a Física, Engenharia e Economia.

A apresentação do Método Simplex, proposto por Georges Dantzig [3], em 1947, durante a Segunda Guerra Mundial, foi um marco na história da Programação Matemática e simboliza o início do estudo da programação linear. O Método Simplex é um procedimento matricial que percorre os pontos extremos do conjunto das soluções viáveis em busca de uma solução ótima. O algoritmo utiliza um critério na busca de forma que a solução seguinte seja sempre “melhor” do que a anterior. Como o conjunto possui um número finito de pontos extremos isso garante que o algoritmo termina em algum ponto, que será uma solução ótima do problema.

Apesar do bom desempenho do simplex, para problemas cuja dimensão é muito grande, esse algoritmo tem um custo operacional alto, pois levará muito tempo para a obtenção de uma solução ótima, o que, na prática, o torna inviável. Desse modo, é possível estudarmos novas “versões” do método de forma a obter uma solução em menos tempo computacional. O método simplex revisado, por exemplo, é um esquema que ordena os cálculos evitando operações desnecessárias, de forma a minimizar o tempo da computação.

Desse modo, tendo em vista a sua vasta aplicabilidade, o propósito desta pesquisa é o estudo destes métodos numéricos para resolução de problemas de programação linear.

1 Resultados da álgebra linear

Neste capítulo listaremos alguns resultados importantes da álgebra linear que foram necessários no desenvolvimento teórico da pesquisa. As demonstrações dos teoremas aqui apresentados podem ser consultadas em [2].

Definição 1.1. Um conjunto $V \neq \emptyset$ é um espaço vetorial sobre um corpo K se satisfaz para todos os elementos, denominados vetores, as seguintes operações:

Adição: A cada par u, v de vetores de V corresponde um vetor $u + v \in V$, denominado de soma de u e v , de modo que:

$$Ai) \quad u + v = v + u, \quad \forall u, v \in V \text{ (propriedade comutativa).}$$

$$Aii) \quad (u + v) + w = u + (v + w) \quad \forall u, v, w \in V \text{ (propriedade associativa).}$$

$$Aiii) \quad \text{Existe em } V \text{ um vetor, denominado nulo e denotado por } 0, \text{ tal que } 0 + v = 0 \quad \forall v \in V.$$

$$Aiv) \quad \text{Para cada vetor } v \in V \text{ existe um vetor em } V \text{ denotado por } -v \text{ tal que } v + (-v) = 0.$$

Multiplicação : A cada par $\alpha \in K$ e $v \in V$, corresponde um vetor $\alpha v \in V$ denominado produto por escalar de α por v , de modo que:

$$Mi) \quad (\alpha\beta)v = \alpha(\beta v) \quad \forall \alpha, \beta \in K \text{ e } \forall v \in V \text{ (propriedade associativa).}$$

$$Mii) \quad 1 \cdot v = v, \quad \forall v \in V \text{ (onde } 1 \text{ é o elemento identidade de } K).$$

$$Miii) \quad \alpha(v + u) = \alpha v + \alpha u \quad \forall \alpha \in K \text{ e } \forall u, v \in V.$$

$$Miv) \quad (\alpha + \beta)v = \alpha v + \beta v \quad \forall \alpha, \beta \in K \text{ e } \forall v \in V.$$

Definição 1.2. Um vetor $v \in V$, com V um espaço vetorial sobre um corpo K , é uma combinação linear dos vetores $v_1, v_2, \dots, v_n \in V$ se existirem escalares $\alpha_1, \alpha_2, \dots, \alpha_n \in K$ tais que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n.$$

Definição 1.3. Seja V um espaço vetorial sobre um corpo K e B um subconjunto de V . Dizemos que B é um conjunto gerador de V , se todo elemento de V for uma combinação linear de um número finito de elementos de B .

Definição 1.4. Dizemos um conjunto de vetores v_1, v_2, \dots, v_n em um corpo K é linearmente independente quando $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$ implica que $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$ com $\alpha_i \in K, i = 1, 2, \dots, n$. Um conjunto de vetores será linearmente dependente se não for linearmente independente.

Definição 1.5. Seja V um espaço vetorial sobre um corpo K e B um subconjunto de V . Dizemos que B é uma base de V se:

- i) B for conjunto gerador de V e
- ii) B for linearmente independente.

Definição 1.6. Um espaço vetorial V sobre um corpo K tem dimensão m se:

- i) Existirem m vetores v_1, v_2, \dots, v_m em V linearmente independentes e
- ii) Não existirem $(m + 1)$ vetores linearmente independentes em V .

Denotamos o espaço vetorial V sobre um corpo K , com dimensão m por $V = K^m$.

Definição 1.7. Uma matriz de ordem $m \times n$ sobre um corpo K é um conjunto de elementos distribuídos em m linhas e n colunas. Geralmente é representada na forma:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}$$

ou, de forma abreviada: $A_{m \times n} = [a_{ij}]$ com $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$.

Definição 1.8. Uma matriz $A_{m \times n}$ é dita ser quadrada quando $m = n$.

Definição 1.9. Uma matriz identidade de dimensão $n \times n$, denotada por I_n é tal que $a_{ij} = 1$ sempre que $i = j$ e $a_{ij} = 0$ sempre que $i \neq j$.

Definição 1.10. Dada $A_{m \times n}$ uma matriz, a matriz $B_{n \times m}$, transposta de A , é tal que $b_{ij} = a_{ji}$ para todo $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$. Representamos a transposta de A por A^T .

Definição 1.11. Um vetor de dimensão m é uma matriz de apenas uma coluna e m linhas (vetor coluna), ou de apenas uma linha e m colunas (vetor linha).

Observação 1.1. Vamos tratar todos os vetores como vetores coluna. Quando quisermos representar um vetor linha, tomaremos v como um vetor coluna e o vetor linha será v^T .

Definição 1.12. A multiplicação de uma matriz $A_{m \times n}$ por uma matriz $B_{n \times p}$ resulta em uma matriz $C_{m \times p}$ obtida da seguinte forma:

$$c_{ij} = \sum_{k=1}^n a_{ik}b_{kj} \quad i = 1, 2, \dots, m \quad e \quad j = 1, 2, \dots, p.$$

Observação 1.2. O produto de duas matrizes $A_{m \times n}$ e $B_{p \times q}$ só está bem definido quando $n = p$.

Definição 1.13. Um sistema de equações lineares com m equações e n incógnitas é representado algebricamente da seguinte forma:

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

onde os a_{ij} 's são coeficientes conhecidos, os b_i 's são constantes dadas e os x_i 's são as incógnitas do sistema.

O sistema pode ser representado matricialmente da seguinte forma:

$$Ax = b$$

onde $A_{m \times n}$ é a matriz dos coeficientes das equações lineares, x é o vetor das incógnitas e b é o vetor das constantes dadas.

Definição 1.14. Dado um sistema de equações lineares na forma matricial $Ax = b$, dizemos que a matriz aumentada do sistema é a matriz

$$\left[\begin{array}{c|cccc} A & & & & b \end{array} \right] = \begin{bmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{bmatrix}$$

Definição 1.15. *Seja A uma matriz quadrada de ordem n . Dizemos que a matriz inversa de A , denotada por A^{-1} é tal que:*

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

Se A possui inversa, então dizemos que A é não-singular.

Observação 1.3. *Dado o sistema de equações lineares representado matricialmente por $Ax = b$, se multiplicarmos ambos os lados por A^{-1} pela esquerda, temos $x = A^{-1}b$. Ou seja, se soubermos qual é a matriz inversa de A podemos encontrar a solução do sistema.*

Teorema 1.1. *O sistema de equações lineares representado matricialmente por $Ax = b$ onde A é uma matriz quadrada tem solução única se e somente se A for não-singular.*

Observação 1.4. *Nos resultados deste trabalho, sempre utilizaremos o corpo \mathbb{R} dos números reais.*

Observação 1.5. *Note que podemos tomar algumas, ou todas as colunas de uma matriz como vetores isolados e verificar se este conjunto de vetores é ou não linearmente independente.*

Definição 1.16. *O posto de uma matriz A , ou “rank” é o número de linhas ou colunas linearmente independentes de A .*

Teorema 1.2. *Uma matriz A , quadrada de ordem n , é não-singular (admite inversa) se, e somente se, A tem posto n .*

Observação 1.6. *As seguintes operações elementares sobre as linhas de uma matriz não alteram o seu posto:*

- i) Troca de linhas.*
- ii) Multiplicação de uma linha por um escalar não nulo.*
- iii) Adição de uma linha a outra linha.*

Definição 1.17. *Dizemos que uma matriz está na forma escalonada quando o número de elementos nulos que precedem o primeiro elemento não nulo de cada linha cresce de uma linha para outra.*

Exemplo 1.1. *A matriz:*

$$\begin{bmatrix} 1 & 0 & 4 & 8 \\ 0 & 4 & 9 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

está na forma escalonada.

Definição 1.18. *O primeiro elemento não nulo de uma linha de uma matriz na forma escalonada é chamado pivô.*

Definição 1.19. *Dizemos que uma matriz está na forma canônica quando está escalonada e:*

i) *todos os pivôs valem 1.*

ii) *toda coluna que possui um pivô de uma linha tem todos os outros elementos iguais a zero.*

Definição 1.20. *Dizemos que dois sistemas de equações lineares são equivalentes se ambos possuem a mesma solução.*

Teorema 1.3. *Se realizarmos operações elementares sobre as linhas de uma matriz aumentada de um sistema de equações lineares, o sistema resultante terá a mesma solução que o sistema original.*

Definição 1.21. *Dado o sistema de equações lineares representado matricialmente por $Ax = b$, o método da eliminação gaussiana, utilizado para encontrarmos a solução do sistema, consiste em realizar operações elementares sobre as linhas da matriz aumentada do sistema de forma que a matriz A , dos coeficientes, esteja na forma escalonada.*

Exemplo 1.2. *Resolver o sistema $Ax=b$ com*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 3 \\ 1 & 2 & 4 \end{bmatrix}$$

e $b^T = (0 \quad 1 \quad 2)$.

A matriz aumentada do sistema é:

$$[A \quad b] = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 1 & 3 & 3 & 1 \\ 1 & 2 & 4 & 2 \end{array} \right]$$

Subtraindo a linha 1 das linhas 2 e 3 temos:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array} \right]$$

Assim, chegamos à solução $x_3 = 2, x_2 = 1$ e como $x_1 + 2x_2 + 3x_3 = 0$, temos que $x_1 = -8$.

2 *Descrição dos Problemas de Programação Linear*

Um problema de programação linear é um problema de otimização onde a função que queremos otimizar (minimizar ou maximizar), chamada de função objetivo, é uma função linear e o conjunto das restrições (relações de interdependência entre as variáveis) são igualdades ou desigualdades lineares.

Exemplo 2.1. *Uma empresa fabrica dois produtos P_1 e P_2 . O produto P_1 leva 5 horas na linha de montagem, 2 horas para a finalização e é vendido por R\$120,00, já o produto P_2 , leva 3 horas na linha de montagem, 4 horas para a finalização e é vendido por R\$210,00. Numa semana, a empresa dispõe de 108 horas para a linha de montagem e 60 horas para a finalização. Qual a quantidade, x_1 e x_2 , de cada produto, que deve ser produzida por semana de modo que o lucro seja máximo?*

Podemos representar o problema por:

$$\begin{aligned} \text{maximize} \quad & 120x_1 + 210x_2 \\ \text{sujeito a} \quad & 5x_1 + 3x_2 \leq 108 \\ & 2x_1 + 4x_2 \leq 60 \\ & x_1, x_2 \geq 0. \end{aligned}$$

As variáveis de um problema de programação linear devem ser todas não negativas, (na maioria dos casos as variáveis do problema representam quantidade). Apesar de as restrições poderem ser de desigualdade e, em alguns casos encontrarmos variáveis livres, todos os problemas de programação linear podem ser transformados na forma padrão, onde todas as restrições

são igualdades e todas as variáveis devem ser não negativas:

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

onde b_i, c_j e a_{ij} com $i = 1, \dots, m$ e $j = 1, \dots, n$, são constantes reais fixadas e os x_j com $j = 1, \dots, n$ são as variáveis. Na forma padrão, precisamos que todos os b_i 's sejam não negativos. Caso algum b_i seja negativo, multiplicamos a restrição correspondente por -1 .

Em notação matricial, podemos escrever:

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

onde x é um vetor coluna de dimensão n , c é um vetor de dimensão n , A é uma matriz $m \times n$ e b é um vetor de dimensão m . A inequação vetorial $x \geq 0$ significa que cada componente do vetor x deve ser não negativa.

Nem sempre os problemas de programação linear já estão na forma padrão. Quando temos um problema onde as restrições não são desigualdades, usamos variáveis de folga ou variáveis de excesso para transformar o problema para a forma padrão. Por exemplo, quando temos um problema de programação linear da forma:

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

Podemos transformar o problema para a forma padrão adicionando variáveis y_i 's não negativas, chamadas *variáveis de folga*, da seguinte forma:

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

Se quisermos escrever esse problema em notação matricial, a matriz que representa as equações de restrição terá dimensão $m \times (n + m)$ e será da forma $[A, I]$ onde A é a matriz original do problema e I é a matriz identidade de dimensão m .

Note que maximizar uma função é o mesmo que minimizarmos esta função multiplicada por -1 . Assim, o problema inicial na forma padrão se torna:

$$\begin{aligned}
\text{minimize} \quad & -120x_1 - 210x_2 \\
\text{sujeito a} \quad & 5x_1 + 3x_2 + x_3 = 108 \\
& 2x_1 + 4x_2 + x_4 = 60 \\
& x_1, x_2, x_3, x_4 \geq 0.
\end{aligned}$$

Da mesma forma, se tivermos um problema na forma:

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

Subtraímos variáveis y_i não negativas, neste caso chamadas *variáveis de excesso*, de modo

a transformar o problema para:

$$\begin{aligned}
 \text{minimize} \quad & c_1x_1 + c_2x_2 + \dots + c_nx_n, \\
 \text{sujeito a} \quad & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n - y_1 = b_1 \\
 & a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n - y_2 = b_2 \\
 & \vdots \\
 & a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n - y_m = b_m \\
 & x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0 \\
 & y_1 \geq 0, y_2 \geq 0, \dots, y_m \geq 0.
 \end{aligned}$$

É claro que podemos utilizar o mesmo processo quando temos um conjunto de restrições misto, com desigualdades de menor ou igual e de maior ou igual. Nesse caso, adicionamos variáveis de folga ou de excesso adequadas para cada restrição.

Se tivermos um problema de programação linear na forma padrão, no qual uma ou mais variáveis não precisam ser necessariamente não negativas, ou seja, se não tivermos a restrição $x_i \geq 0$ para algum i , então temos variáveis livres. Nesse caso, o problema também pode ser transformado para a forma padrão. Para isso, podemos utilizar duas técnicas.

Uma das técnicas que podemos utilizar é “particionar” a variável em questão. Por exemplo, se tivermos o problema:

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

Como não temos $x_1 \geq 0$, escrevemos

$$x_1 = u_1 - v_1$$

com $u_1 \geq 0$ e $v_1 \geq 0$. Se substituirmos x_1 por $u_1 - v_1$ no problema, continuamos com equações lineares e agora todas as variáveis devem ser não negativas. O problema passa a ter $n + 1$ variáveis: $u_1, v_1, x_2, \dots, x_n$.

Outra técnica para transformar o problema acima para a forma padrão é eliminar a variável

x_1 bem como uma das equações de restrição.

Passo 1: Tomamos uma das m equações, cujo coeficiente da variável x_1 seja não nulo, digamos a i -ésima equação, e escrevemos x_1 como combinação linear das variáveis x_2, x_3, \dots, x_n mais uma constante, nesse caso b_i (ou seja, isolamos a variável x_1).

Passo 2: Substituímos x_1 no problema geral.

Dessa forma, temos um novo problema onde as variáveis devem ser todas não negativas. Uma vez encontrada a solução para esse novo problema, basta substituir os valores encontrado para as variáveis x_2, x_3, \dots, x_n na combinação linear que representa x_1 que havíamos obtido no passo 1, para encontrar o valor de x_1 .

Exemplo 2.2.

$$\begin{aligned} \text{minimize} \quad & x_1 + 2x_2 + 2x_3 \\ \text{sujeito a} \quad & 2x_1 + x_2 + x_3 = 7 \\ & x_1 + 3x_2 + x_3 = 8 \\ & x_2, x_3 \geq 0. \end{aligned}$$

Da segunda restrição temos: $x_1 = 8 - 3x_2 - x_3$. Substituindo essa igualdade na função objetivo e na primeira restrição temos:

$$\begin{aligned} \text{minimize} \quad & 8 - x_2 + x_3 \\ \text{sujeito a} \quad & -5x_2 - x_3 = -9 \\ & x_2, x_3 \geq 0 \end{aligned}$$

$$\begin{aligned} \text{minimize} \quad & -x_2 + x_3 \\ \text{sujeito a} \quad & 5x_2 + x_3 = 9 \\ & x_2, x_3 \geq 0. \end{aligned}$$

Resolvemos o problema reduzido, cuja solução será $x_2 = 9/2$ e $x_3 = 0$. Então encontramos o valor de x_1 , que será $x_1 = -11/2$.

2.1 Soluções básicas

Em um sistema de equações lineares da forma

$$Ax = b$$

onde x é um vetor de dimensão n , b é um vetor de dimensão m e A é uma matriz $m \times n$, com $m < n$ de posto m , é possível selecionar um conjunto de m colunas linearmente independentes de A . Suponha, por exemplo, que as m primeiras colunas de A sejam linearmente independentes. Chamamos de B a matriz composta por essas m colunas, que será não singular. Se tomarmos x_B como sendo o vetor com as variáveis de mesmo índice das colunas escolhidas para B , nesse caso, x_1, \dots, x_m podemos resolver a equação

$$Bx_B = b$$

que terá solução única dada por:

$$x_B = B^{-1}b.$$

Se tomarmos $x = (x_B, 0)$ obtemos uma solução para $Ax = b$, pois teremos

$$Ax = Bx_B = b.$$

Definição 2.1. *Dado um conjunto de m equações lineares com n variáveis, $Ax = b$, seja B uma submatriz não singular de dimensão $m \times m$ composta por colunas de A . Então, se todas as $n - m$ componentes de x , não associadas com as colunas de B , forem nulas, a solução que resulta deste conjunto de equações é dita ser uma **solução básica** para o problema em relação à base B . As componentes de x de mesmo índice das colunas de B são chamadas de **variáveis básicas**.*

A matriz B é chamada de base, pois B é composta por m colunas linearmente independentes que podem ser consideradas uma base para o espaço \mathbb{R}^m . A solução básica é uma expressão do vetor b como combinação linear dos vetores dessa base B .

Se $m < n$ e as equações forem linearmente independentes, então existem várias soluções para o sistema. Para que A possua ao menos uma solução básica precisamos supor primeiramente que $n > m$. Além disso, as linhas de A devem ser linearmente independentes, pois se houver uma linha que seja combinação linear das outras, iremos obter uma restrição contraditória (e portanto, nenhuma solução) ou uma redundância que poderia ser eliminada. Portanto é necessário que A tenha posto m para que haja ao menos uma solução básica para o sistema.

Definição 2.2. *Se uma ou mais variáveis básicas em uma solução básica for nula, essa solução é dita ser uma solução básica degenerada.*

Consideramos agora o sistema:

$$Ax = b,$$

$$x \geq 0$$

que representa as restrições de um problema de programação linear na forma padrão.

Definição 2.3. Um vetor x que satisfaz esse sistema é dito ser uma **solução viável** para essas restrições. Uma solução viável que também é básica é dita ser uma **solução viável básica**. Se a solução também for degenerada, é chamada de **solução viável básica degenerada**.

Observação 2.1. Um problema de programação linear pode ter várias soluções viáveis.

Exemplo 2.3. Dado o problema

$$\begin{aligned} \text{minimize} \quad & x_1 + 4x_2 + x_3 \\ \text{sujeito a} \quad & 2x_1 - 2x_2 + x_3 \leq 4 \\ & x_1 - x_3 \leq 1 \\ & x_1, x_2, x_3 \geq 0. \end{aligned}$$

temos que as soluções $x_1 = 3, x_2 = 2, x_3 = 2$ e $x_1 = 5, x_2 = 5, x_3 = 4$ são ambas viáveis para as restrições.

Definição 2.4. Uma solução viável para as restrições que atinge o valor mínimo para a função objetivo sujeita a essas restrições é dita ser uma **solução viável ótima**. Se essa solução for também básica, é chamada de **solução básica viável ótima**.

Teorema 2.1 (Fundamental da Programação Linear). Dado um problema de programação linear na forma padrão, onde A é uma matriz de dimensão $m \times n$ e posto m , temos que:

- (i) Se existe uma solução viável, existe uma solução viável básica.
- (ii) Se existe uma solução viável ótima, existe uma solução viável básica ótima.

Demonstração:

(i) Denotando as colunas de A por a_1, a_2, \dots, a_n e supondo que $x = (x_1, x_2, \dots, x_n)$ seja uma solução viável para o problema, temos:

$$x_1 a_1 + x_2 a_2 + \dots + x_n a_n = b.$$

Suponha que p componentes do vetor x (solução viável) sejam positivas. Para facilitar a demonstração, suponha que sejam exatamente as p primeiras componentes (com $p \leq n$). Como

as componentes restantes deste vetor serão nulas, temos:

$$x_1 a_1 + x_2 a_2 + \dots + x_p a_p = b. \quad (2.1)$$

Então temos duas possibilidades, os vetores a_1, a_2, \dots, a_p são linearmente independentes ou linearmente dependentes.

Suponha primeiramente que a_1, a_2, \dots, a_p sejam linearmente independentes. Então, $p \leq m$ pois A tem posto m . Se $p = m$ a solução será básica e a demonstração termina. Se $p < m$, como A tem posto m , podemos encontrar $m - p$ vetores, dos $n - p$ vetores restantes, tais que o conjunto desses m vetores ($p + (m - p) = m$) seja linearmente independente. Atribuindo zero às outras $m - p$ variáveis, teremos uma solução básica degenerada.

Suponha agora que a_1, a_2, \dots, a_p sejam linearmente dependentes. Portanto, existem constantes y_1, y_2, \dots, y_p não nulas tais que:

$$y_1 a_1 + y_2 a_2 + \dots + y_p a_p = 0.$$

Multiplicando esta equação por um escalar ε e subtraindo-a da equação (2.1), temos:

$$(x_1 - \varepsilon y_1) a_1 + (x_2 - \varepsilon y_2) a_2 + \dots + (x_p - \varepsilon y_p) a_p = b.$$

Para cada ε , as componentes $x_i - \varepsilon y_i$ correspondem a uma solução para o sistema de equações, embora possamos ter $x_i - \varepsilon y_i \geq 0$. Denotando $y = (y_1, y_2, \dots, y_p, 0, 0, \dots, 0)$, para cada ε ,

$$x - \varepsilon y$$

é uma solução para o sistema de equações. Para $\varepsilon = 0$, temos a solução viável inicial. Conforme ε aumenta, as outras componentes crescem, decrescem ou permanecem as mesmas, o que depende do valor do y_i correspondente (positivo, negativo ou nulo). Assumindo que ao menos um y_i seja positivo, ao menos uma componente irá decrescer conforme ε aumenta. Se aumentarmos ε até o menor valor tal que uma ou mais componentes se tornem nulas, ou seja,

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

Para esse valor de ε a solução será viável e possuirá no máximo $p - 1$ componentes positivas. Repetindo este processo, podemos eliminar variáveis positivas até que tenhamos uma solução viável cujas colunas sejam linearmente independentes. Assim, voltamos ao caso anterior e a demonstração termina.

(ii) Seja $x = (x_1, x_2, \dots, x_n)$ uma solução ótima. Como no item anterior, temos dois ca-

sos: a_1, a_2, \dots, a_p são linearmente independentes ou linearmente dependentes. No primeiro caso a demonstração é a mesma do item anterior. No segundo caso também teremos a mesma demonstração, só nos resta mostrar que $x - \varepsilon y$ é ótima. Para tanto, note que o valor correspondente à solução $x - \varepsilon y$ é

$$c^T x - \varepsilon c^T y.$$

Para ε suficientemente pequeno, $x - \varepsilon y$ é uma solução viável para valores positivos ou negativos de ε . Assim, concluímos que $c^T y = 0$. Quando $c^T y \neq 0$, um ε pequeno e com sinal apropriado pode ser determinado de modo a tornar a expressão anterior menor do que $c^T x$ mantendo a sua viabilidade. Isso violaria a suposição de otimalidade de x e, portanto, teríamos $c^T y = 0$. Estabelecendo que a nova solução viável com componentes positivas menores também seja ótima, o restante da prova é análogo ao item (i). ■

Esse teorema nos diz que sempre que o problema admitir uma solução viável ótima, admitirá uma solução básica viável ótima, ou seja, precisamos encontrar soluções básicas viáveis para depois decidir uma solução entre elas que seja ótima. Como temos n variáveis e m restrições, teremos no máximo:

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}$$

soluções básicas, que corresponde a todas as maneiras de tomarmos m colunas de n e assim, temos um número finito de possibilidades (assumindo que todas as soluções básicas são soluções viáveis não degeneradas). Se expandirmos a técnica utilizada na demonstração deste teorema, podemos deduzir o método simplex, que será apresentado posteriormente.

2.2 Relações de convexidade

Definição 2.5. Um conjunto X em K^n é dito ser convexo se para todos os pontos $x_1, x_2 \in X$ e para todo α real, $0 < \alpha < 1$, o ponto $\alpha x_1 + (1 - \alpha)x_2$ pertence a X .

Exemplo 2.4. No \mathbb{R}^2 temos:



conjunto convexo



conjunto não convexo

Definição 2.6. Um ponto x em um conjunto convexo X é dito ser um ponto extremo de X se não existirem dois pontos x_1 e x_2 em X distintos tais que $x = \alpha x_1 + (1 - \alpha)x_2$ para algum $\alpha, 0 < \alpha < 1$.

Definição 2.7. Um ponto $x \in X$ é uma combinação convexa de x_1, x_2, \dots, x_p , pontos de X se e somente se tivermos $x = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_p x_p$, com $\alpha_i \geq 0, i = 1, 2, \dots, p$ e $\alpha_1 + \alpha_2 + \dots + \alpha_p = 1$.

Definição 2.8. Um conjunto X em K^n é dito ser uma variedade linear se para todos os pontos $x_1, x_2 \in X$ e para todo λ real, o ponto $\lambda x_1 + (1 - \lambda)x_2$ pertence a X .

Definição 2.9. Um hiperplano em K^n é uma variedade linear de dimensão $n - 1$.

Proposição 2.1. Seja a um vetor não nulo de dimensão n e c um número real. O conjunto

$$H = \{x \in K^n : a^T x = c\}$$

é um hiperplano em K^n .

Definição 2.10. Seja a um vetor não nulo em K^n e c um número real. Correspondente ao hiperplano $H = \{x : a^T x = c\}$ existem os semi-espacos fechados, positivo e negativo:

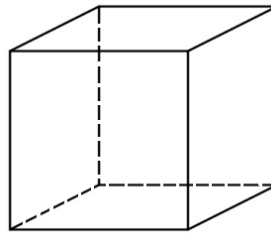
$$H_+ = \{x : a^T x \geq c\}$$

$$H_- = \{x : a^T x \leq c\}$$

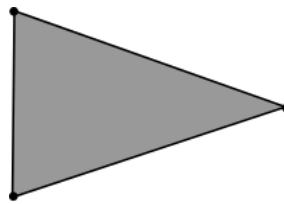
Definição 2.11. Um conjunto que pode ser expresso como interseção de um número finito de semi-espacos fechados é dito politopo convexo.

Corolário 2.1. Se o politopo convexo X é limitado, então é um poliedro convexo, ou seja, X é composto por pontos que são combinação convexa de um número finito de pontos.

Exemplo 2.5. No \mathbb{R}^3 o cubo é um poliedro:



Observação 2.2. *Os pontos extremos de um poliedro convexo correspondem aos seus vértices.*



Teorema 2.2 (Equivalência entre pontos extremos e soluções básicas.). *Seja A uma matriz $m \times n$, de posto m e b um vetor de dimensão m . Seja X o poliedro convexo composto por todos os vetores x de dimensão n que satisfazem,*

$$Ax = b$$

$$x \geq 0$$

Um vetor x é um ponto extremo de X se e somente se x é uma solução básica para o problema.

Demonstração:

Seja $x = (x_1, x_2, \dots, x_m, 0, 0, \dots, 0)$ uma solução viável básica (suponhamos sem perda de generalidade que as colunas da solução básica são as m primeiras). Então temos:

$$x_1 a_1 + x_2 a_2 + \dots + x_m a_m = b$$

onde $a_i, i = 1, \dots, m$ representam as colunas de A correspondentes à solução básica. Essas m colunas são linearmente independentes. Suponhamos que x possa ser expresso como combinação convexa de dois outros pontos de X ; digamos $x = \alpha y + (1 - \alpha)z$, com $0 < \alpha < 1$, e $y \neq z$. Como todas as componentes de x, y , e z são não negativas e $0 < \alpha < 1$ segue que todas as $n - m$ componentes restantes de y e z são nulas. Assim, temos

$$y_1 a_1 + y_2 a_2 + \dots + y_m a_m = b$$

e

$$z_1 a_1 + z_2 a_2 + \dots + z_m a_m = b.$$

Como os vetores a_1, a_2, \dots, a_m são linearmente independentes, segue que $x = y = z$ e portanto, x é um ponto extremo de X .

Suponha agora que x é um ponto extremo de X . Sem perda de generalidade assumimos que as componentes não nulas de x são as k primeiras. Então:

$$x_1 a_1 + x_2 a_2 + \dots + x_k a_k = b$$

com $x_i > 0$, $i = 1, 2, \dots, k$. Para mostrar que x é uma solução básica, basta mostrar que a_1, a_2, \dots, a_k são linearmente independentes. Vamos supor por contradição que não sejam. Então, existe uma combinação linear não trivial que será nula:

$$y_1 a_1 + y_2 a_2 + \dots + y_k a_k = 0.$$

Definimos o vetor $y = (y_1, y_2, \dots, y_k, 0, \dots, 0)$ de dimensão n . Como $x_i > 0$, $1 \leq i \leq k$, podemos escolher ε tal que

$$x + \varepsilon y \geq 0, \quad x - \varepsilon y \geq 0.$$

Então, temos

$$x = \frac{1}{2}(x + \varepsilon y) + \frac{1}{2}(x - \varepsilon y)$$

expressando x como combinação convexa de dois vetores distintos em X , o que não pode ocorrer. Logo, a_1, a_2, \dots, a_k são L.I. e a solução é básica. ■

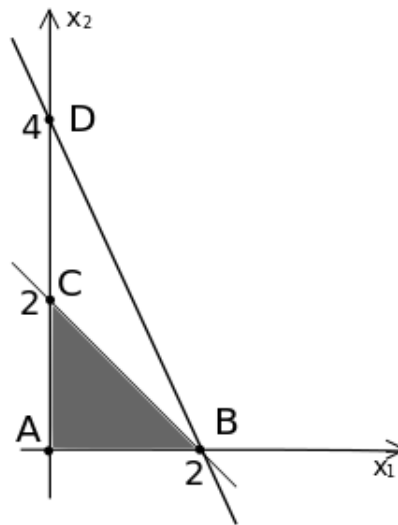
Exemplo 2.6. *Dado o conjunto de restrições de um problema de PL:*

$$x_1 + x_2 \leq 2$$

$$2x_1 + x_2 \leq 4$$

$$x_1, x_2 \geq 0.$$

O conjunto viável pode ser representado graficamente (em cinza):



Introduzindo as variáveis de folga, temos:

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

Note que o vértice A corresponde à solução $x = [0 \ 0 \ 2 \ 4]$, que corresponde à base composta pelas colunas 3 e 4 da matriz do simplex (solução trivial correspondente às variáveis de folga).

O vértice C corresponde à solução $x = [0 \ 2 \ 0 \ 2]$, relativa à base composta pelas colunas 2 e 4.

O vértice B, por sua vez, corresponde à solução $x = [2 \ 0 \ 0 \ 0]$, que pode ser obtida tomando 3 bases diferentes: as colunas 1 e 2, 1 e 3 ou 1 e 4.

O vértice D, corresponde à solução básica $x = [0 \ 4 \ -2 \ 0]$, relativa à base composta pelas colunas 2 e 3, porém não é viável, pois temos $x_3 < 0$.

Portanto, cada solução básica viável corresponde a um vértice do conjunto viável, porém, cada vértice pode corresponder a mais de uma solução básica viável.

Corolário 2.2. Se o conjunto convexo X , como no teorema 2.2, é não vazio, possui ao menos

um ponto extremo.

Corolário 2.3. *Se existe uma solução finita ótima para um problema de programação linear, existe uma solução finita ótima que é um ponto extremo do conjunto X , de restrições.*

Corolário 2.4. *O conjunto X , de restrições, possui no máximo um número finito de pontos extremos.*

Proposição 2.2. *Uma função objetivo linear $c^T x$ atinge seu mínimo ao longo de um poliedro convexo X em um ponto extremo de X .*

Demonstração: Sejam x_1, x_2, \dots, x_k pontos extremos de X . Então, qualquer ponto x de X pode ser expresso por uma combinação convexa dos pontos extremos:

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k$$

onde $\alpha_i \geq 0$, $i = 1, 2, \dots, k$ com $\alpha_1 + \alpha_2 + \dots + \alpha_k = 1$.

Então

$$c^T x = \alpha_1 (c^T x_1) + \alpha_2 (c^T x_2) + \dots + \alpha_k (c^T x_k).$$

Seja $z_0 = \min \{c^T x_i : i = 1, 2, \dots, k\}$, então temos:

$$c^T x \geq (\alpha_1 + \alpha_2 + \dots + \alpha_k) z_0 = z_0.$$

Logo, o mínimo de $c^T x$ em X é z_0 .

3 *Método Simplex*

Neste capítulo vamos utilizar o desenvolvimento apresentado por Luenberger em [6], para entender o método simplex. A eliminação gaussiana é utilizada amplamente durante os procedimentos do método simplex, portanto, vamos ressaltar como é feito o pivotamento em um sistema de equações lineares. Considere o conjunto de equações lineares:

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

onde $m \leq n$.

Utilizando o procedimento da eliminação gaussiana, descrito no Capítulo 2, se as primeiras m colunas de A forem linearmente independentes, o sistema pode ser transformado para a forma canônica:

$$\begin{array}{rcccccl} x_1 & & +y_{1,m+1}x_{m+1} & y_{1,m+2}x_{m+2} & \dots & y_{1,n}x_n & = & y_{10} \\ x_2 & & +y_{2,m+1}x_{m+1} & y_{2,m+2}x_{m+2} & \dots & y_{2,n}x_n & = & y_{20} \\ & \ddots & & & & & & \\ x_m & & +y_{m,m+1}x_{m+1} & y_{m,m+2}x_{m+2} & \dots & y_{m,n}x_n & = & y_{m0} \end{array}$$

cuja solução básica correspondente é

$$\begin{aligned} x_1 = y_{10}, x_2 = y_{20}, \dots, x_m = y_{m0} \\ x_{m+1} = 0, \dots, x_n = 0. \end{aligned}$$

Como vimos no Capítulo 2, o vetor $(x_1, x_2, \dots, x_m, 0, 0, \dots, 0)$ é uma solução básica para o sistema $Ax = b$. Dessa forma, as variáveis x_1, x_2, \dots, x_m são chamadas *básicas* e as variáveis

restantes, *não básicas*.

Dessa forma, um sistema de equações está na forma canônica se dentre as n variáveis existirem m básicas de tal forma que cada uma figure em apenas uma equação, seu coeficiente seja unitário e não haja outra dessas m variáveis em outra equação, ou seja, se a matriz que representa os coeficientes estiver na forma canônica. Podemos representar esse sistema apenas pelos coeficientes de cada variável:

$$\begin{array}{cccccccc} 1 & 0 & \dots & 0 & y_{1,m+1} & y_{1,m+2} & \dots & y_{1,n} & = & y_{10} \\ 0 & 1 & \dots & 0 & y_{2,m+1} & y_{2,m+2} & \dots & y_{2,n} & = & y_{20} \\ \vdots & & & & & & & & & \\ 0 & 0 & \dots & 1 & y_{m,m+1} & y_{m,m+2} & \dots & y_{m,n} & = & y_{m0}. \end{array}$$

Vimos que uma solução básica corresponde a um vértice do poliedro formado pelas restrições do problema. O método simplex parte de uma solução básica inicial e, seguindo um critério, escolhe uma nova solução básica de forma que o valor da função objetivo diminua. Para partir de uma solução básica para outra, basta “trocar” uma de suas variáveis básicas por uma não básica, seguindo um critério que veremos posteriormente. Para “trocar” esta variável básica por uma não básica, basta trocarmos uma coluna da base B por outra que não pertence à base B , de forma que as colunas continuem sendo linearmente independentes. Para isso, utilizamos a eliminação gaussiana.

Por exemplo, suponha que queiramos substituir a variável x_p , $1 \leq p \leq m$, pela variável não básica, x_q (isso só pode ser feito se $y_{pq} \neq 0$). Dividimos a linha p por y_{pq} para que o coeficiente de x_q se torne unitário na p -ésima equação e subtraímos múltiplos da linha p de cada uma das outras linhas de forma que tenhamos coeficiente nulo para x_q em todas as outras equações. Assim, a q -ésima coluna do sistema é feita de zeros exceto na p -ésima linha. Os coeficientes da nova forma canônica y'_{ij} , se tornam:

$$\begin{cases} y'_{ij} = y_{ij} - \frac{y_{pj}}{y_{pq}} y_{iq}, & i \neq p \\ y'_{pj} = \frac{y_{pj}}{y_{pq}}. \end{cases} \quad (3.1)$$

Apesar de utilizarmos o processo de pivotamento para encontrar novas variáveis básicas, existem algumas condições necessárias que devem ser satisfeitas para que a viabilidade da nova solução básica seja preservada (não negatividade).

Hipótese de não degenerabilidade: Dado um problema de programação linear

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

todas as soluções básicas são soluções viáveis não degeneradas.

Essa hipótese é utilizada em todo o desenvolvimento do método simplex, já que o método falha se a mesma não se aplica. Veremos um exemplo na próxima seção.

Assim, para determinar um vetor que deve deixar a base, utilizamos o seguinte processo: suponha que tenhamos uma solução básica viável

$$x = (x_1, x_2, \dots, x_m, 0, 0, \dots, 0),$$

ou seja,

$$x_1 a_1 + x_2 a_2 + \dots + x_m a_m = b$$

onde a_i , $i = 1, 2, \dots, m$, representa a i -ésima coluna de A .

Utilizando a hipótese de não degenerabilidade, temos $x_i > 0$, $i = 1, 2, \dots, m$. Vamos adicionar à base o vetor a_q , $q > m$, cuja representação em termos da base atual é

$$a_q = y_{1q} a_1 + y_{2q} a_2 + \dots + y_{mq} a_m.$$

Multiplicando a_q por uma variável $\varepsilon \geq 0$ e subtraindo da equação anterior, temos:

$$(x_1 - \varepsilon y_{1q}) a_1 + (x_2 - \varepsilon y_{2q}) a_2 + \dots + (x_m - \varepsilon y_{mq}) a_m + \varepsilon a_q = b.$$

Então, para qualquer $\varepsilon \geq 0$ a última equação nos fornece b como combinação linear de no máximo $m + 1$ vetores. Mas para termos uma nova solução básica precisamos eliminar um vetor (uma coluna de A) dessa combinação linear. Se $\varepsilon = 0$, temos a antiga solução básica viável. Conforme aumentamos ε , a última equação nos dá uma solução não básica. Os coeficientes dos outros vetores irão crescer ou decrescer linearmente de acordo com o crescimento de ε . Como vimos na demonstração do teorema fundamental da programação linear, o valor de ε correspondente ao primeiro valor que anula um ou mais coeficientes, será:

$$\varepsilon = \min \left\{ \frac{x_i}{y_{iq}} : y_{iq} > 0 \quad i = 1, 2, \dots, m \right\}$$

onde $x_i = y_{i0}$.

Neste caso, temos uma nova solução básica viável, na qual o vetor a_q foi substituído pelo vetor a_p , onde p corresponde ao índice da mínima razão a qual foi tomada como ε . Se o mínimo é alcançado por mais de um índice i então a nova solução é degenerada e qualquer um dos vetores com componente zero pode ser escolhido para deixar a base.

Se nenhum dos y_{iq} 's for positivo, então todos os coeficientes aumentam conforme ε aumenta, e portanto, nenhuma solução básica viável pode ser obtida. Observamos porém que nesse caso, existem soluções viáveis com coeficientes grandes (arbitrários). Isso significa que o conjunto X das soluções viáveis é ilimitado. Assim, dada uma solução viável para um problema de programação linear temos duas opções: ou existe uma nova solução viável que pode ser encontrada, ou o conjunto de soluções viáveis é ilimitado.

A operação completa é dada por:

$$x'_i = x_i - \frac{y_{iq}}{y_{pq}} x_p.$$

A ideia do método simplex é selecionar colunas para sair e entrar na solução básica de tal forma que a nova solução obtenha um valor menor para a função objetivo do que a solução anterior. Portanto, precisamos escolher uma coluna para entrar na base de tal forma que o valor da função objetivo diminua e uma coluna para sair da base de modo que a viabilidade seja mantida.

Dada uma solução básica viável:

$$(x_B, 0) = (y_{10}, y_{20}, \dots, y_{m0}, 0, 0, \dots, 0)$$

onde a matriz dos coeficientes possui uma matriz identidade nas primeiras m colunas:

$$\begin{array}{cccccccc} a_1 & a_2 & \dots & a_m & a_{m+1} & a_{m+1} & \dots & a_n & b \\ 1 & 0 & \dots & 0 & y_{1,m+1} & y_{1,m+2} & \dots & y_{1,n} & y_{10} \\ 0 & 1 & \dots & 0 & y_{2,m+1} & y_{2,m+2} & \dots & y_{2,n} & y_{20} \\ \vdots & & & & & & & & \\ 0 & 0 & \dots & 1 & y_{m,m+1} & y_{m,m+2} & \dots & y_{m,n} & y_{m0} \end{array}$$

Vamos representar o valor da função objetivo correspondente a qualquer solução x por z :

$$z = c_1 x_1 + c_2 x_2 + \dots + c_n x_n,$$

e portanto, para a solução básica acima, o valor correspondente é:

$$z_0 = c_B^T x_B,$$

onde $c_B = [c_1, c_2, \dots, c_m]$.

Se valores arbitrários forem assumidos para $x_{m+1}, x_{m+2}, \dots, x_n$ podemos resolver para as variáveis restantes da forma

$$\begin{aligned} x_1 &= y_{10} - \sum_{j=m+1}^n y_{1j} x_j \\ x_2 &= y_{20} - \sum_{j=m+1}^n y_{2j} x_j \\ &\vdots \\ x_m &= y_{m0} - \sum_{j=m+1}^n y_{mj} x_j. \end{aligned}$$

Usando essas equações, podemos eliminar x_1, x_2, \dots, x_m da forma geral. Assim, obtemos

$$z = c^T x = z_0 + (c_{m+1} - z_{m+1})x_{m+1} + (c_{m+2} - z_{m+2})x_{m+2} + \dots + (c_n - z_n)x_n$$

onde

$$z_j = y_{1j}c_1 + y_{2j}c_2 + \dots + y_{mj}c_m,$$

$$m + 1 \leq j \leq n.$$

Essa equação nos dá o valor da função objetivo z para qualquer solução de $Ax = b$ em termos das variáveis x_{m+1}, \dots, x_n . Dessa forma, podemos determinar se existe vantagem em alterar a solução básica introduzindo uma das variáveis não básicas. Note que se $c_j - z_j$ for negativo para algum j , $m + 1 \leq j \leq n$ então um aumento do x_j de zero para algum valor positivo diminuiria o custo total, e portanto, seria uma solução melhor. Ou seja, a solução poderá ser melhorada sempre que houver $c_j - z_j$ negativo para algum j , $m + 1 \leq j \leq n$.

Se tomarmos z como uma variável adicional, e

$$c_1 x_1 + c_2 x_2 + \dots + c_n x_n - z = 0$$

como mais uma equação linear do sistema (na última linha), então temos $(m + 1)$ equações. Se pivotarmos o sistema de forma que as variáveis básicas sejam x_1, x_2, \dots, x_m, z , a última linha nos dará o valor de z em termos das variáveis não básicas, como queríamos, pois dessa forma

podemos observar os valores de $c_i - z_1$'s que estão na última linha.

Teorema 3.1 (Melhoria da solução básica viável). *Dada uma solução básica viável não degenerada com valor objetivo correspondente a z_0 , suponha que para algum j ocorra que $c_j - z_j \leq 0$. Então, existe uma solução viável com valor objetivo $z \leq z_0$. Se a coluna a_j puder ser substituída por algum vetor na base original para produzir uma nova solução básica viável, essa nova solução terá $z \leq z_0$. Se a_j não pode ser substituída para produzir uma nova solução básica viável, então o conjunto X das soluções básicas viáveis é ilimitado e a função objetivo pode ser tomada arbitrariamente pequena (para menos infinito).*

Demonstração: O resultado é uma consequência imediata do que encontramos anteriormente. Seja $(x_1, x_2, \dots, x_m, 0, 0, \dots, 0)$ uma solução básica viável cujo valor objetivo é z_0 e suponha que $c_{m+1} - z_{m+1} < 0$. Então, em qualquer caso, uma nova solução viável pode ser construída da forma

$$(x'_1, x'_2, \dots, x'_m, x'_{m+1}, 0, \dots, 0) \quad \text{com} \quad x'_{m+1} > 0.$$

Substituindo essa solução na relação:

$$z = c^T x = z_0 + (c_{m+1} - z_{m+1})x_{m+1} + (c_{m+2} - z_{m+2})x_{m+2} + \dots + (c_n - z_n)x_n$$

obtemos

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

e portanto, $z < z_0$ para qualquer tal solução. Queremos tornar x'_{m+1} tão grande quanto seja possível. Na medida em que x'_{m+1} cresce, as outras componentes crescem, permanecem constantes ou decrescem. Assim, x'_{m+1} pode ser aumentada até um $x'_i = 0$, $i \leq m$, de forma que obtenhamos uma nova solução básica viável, ou que nenhum dos x'_i 's decresçam e então x'_{m+1} pode ser aumentado sem limites, indicando que o conjunto das soluções é ilimitado e que a função objetivo não possui um valor mínimo. ■

Teorema 3.2 (Condição de otimalidade). *Se para alguma solução básica viável $c_j - z_j \geq 0$ para todo j , então esta solução é ótima.*

Demonstração: Isso segue imediatamente da relação

$$z = c^T x = z_0 + (c_{m+1} - z_{m+1})x_{m+1} + (c_{m+2} - z_{m+2})x_{m+2} + \dots + (c_n - z_n)x_n$$

uma vez que qualquer outra solução viável deve ter $x_i \geq 0$ para todo i e, portanto, o valor z da função objetivo satisfará $z - z_0 \geq 0$. ■

Como os valores $c_j - z_j$ são importante no desenvolvimento do método simplex, os representaremos por:

$$r_j = c_j - z_j.$$

3.1 Procedimento computacional do método simplex

Vamos utilizar agora os conceitos apresentados anteriormente para deduzir o método simplex. Para utilizar o método, precisamos de um ponto de partida, ou seja, é necessário que já tenhamos uma solução básica viável. Seja então o problema de programação linear

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

onde A tem dimensão $m \times n$, x e c têm dimensão n e b tem dimensão m .

Tomamos a matriz dos coeficientes das equações na forma canônica para a solução básica viável já conhecida. Calculamos os r_j 's, correspondentes a essa solução inicial e adicionamos uma linha, formada por esses r_j 's. Essa linha terá apenas n componentes e a matriz aumentada do sistema possui $n + 1$ colunas, então, na $(n + 1)$ -ésima componente da linha adicionada, representamos o respectivo valor da função objetivo multiplicado por menos um. Temos assim, a *matriz do simplex* que possuirá $m + 1$ linhas e $n + 1$ colunas. Suponha que as variáveis básicas sejam x_1, x_2, \dots, x_m . Então a matriz do simplex terá a forma:

$$\begin{bmatrix} 1 & 0 & \dots & 0 & y_{1,m+1} & y_{1,m+2} & \dots & y_{1j} & \dots & y_{1,n} & y_{10} \\ 0 & 1 & & & & & & & & & \\ \vdots & & & & & & & & & & \\ 0 & 0 & \dots & 0 & y_{i,m+1} & y_{i,m+2} & \dots & y_{ij} & \dots & y_{i,n} & y_{i0} \\ \vdots & & & & & & & & & & \\ 0 & 0 & \dots & 1 & y_{m,m+1} & y_{m,m+2} & \dots & y_{mj} & \dots & y_{m,n} & y_{m0} \\ 0 & 0 & \dots & 0 & r_{m+1} & r_{m+2} & \dots & r_j & \dots & r_n & -z_0 \end{bmatrix}$$

A solução básica correspondente a esse sistema é:

$$x_i = \begin{cases} y_{i0} & 0 \leq i \leq m \\ 0 & m+1 \leq i \leq n \end{cases}$$

onde $y_{i0} \geq 0, i = 1, 2, \dots, m$. O valor correspondente da função objetivo é z_0 .

Os valores dos r_j 's, indicam se o valor da função objetivo irá crescer ou decrescer se x_j for trazido para a solução, ou seja, se pivotarmos nesta coluna. Se todos esses coeficientes forem não negativos, então a solução indicada é ótima. Se existir algum que seja negativo, é possível melhorar a solução (desde que o problema não seja degenerado) trazendo o vetor correspondente para a base, como vimos no capítulo anterior. Quando mais de um desses coeficientes forem negativos, qualquer um deles pode ser trazido à base, para facilitar o algoritmo do método, escolhemos o mais negativo dentre eles.

Podemos considerar:

$$c_1x_1 + c_2x_2 + \dots + c_nx_n - z = 0$$

como uma nova equação. A solução básica do sistema aumentado teria $m+1$ variáveis básicas mas podemos exigir que z seja uma delas. Por isso, não é necessário que adicionemos uma coluna correspondente a z já que esta seria sempre $(0, 0, \dots, 0, 1)$. Adicionamos à matriz inicial uma última linha composta pelos c_i 's e completada por zeros, que representará essa equação adicional. Usando as operações padrão de pivoteamento, podemos reduzir a zero os elementos nesta linha correspondentes a variáveis básicas. Como já vimos, isto é equivalente a transformar a equação adicional para a forma

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

o que é equivalente a

$$c^T x = z_0 + \sum_{j=m+1}^n (c_j - z_j)x_j.$$

Dessa forma, a última linha pode ser tratada como qualquer outra no processo de pivotamento, iniciando com os c_j 's e reduzindo os termos correspondentes às variáveis básicas a zero através de operações entre linhas.

Uma vez que uma coluna q tenha sido selecionada para ser pivotada, calculamos as razões y_{i0}/y_{iq} para os elementos positivos de $y_{iq}, i = 1, 2, \dots, m$, da q -ésima coluna, para determinar qual elemento será o pivô, que nesse caso será o elemento y_{pq} , onde p é o índice correspondente

à menor razão encontrada. Quando pivotamos no elemento y_{pq} a solução continua sendo viável e o valor da função objetivo sofre um decréscimo (se o problema não for degenerado é claro). Se houver mais de um elemento correspondente à razão mínima, qualquer um pode ser utilizado e se não houver elementos não negativos na coluna, o problema é ilimitado. Atualizamos então toda a matriz com y_{pq} sendo o pivô da coluna correspondente à nova variável básica, e transformamos a última linha da mesma maneira que as outras (exceto a linha q que é a linha do pivô escolhido) de forma a obter uma nova matriz na forma canônica. Como anteriormente, o valor da função objetivo de correspondente à essa nova solução básica aparece no último elemento da última linha da matriz.

O algoritmo simplex pode ser resumido nos seguintes passos:

Passo 0 Formar a matriz correspondente a uma solução básica viável.

Passo 1 Se todos $r_j \geq 0$, pare. A solução básica viável atual já é ótima.

Passo 2 Selecionar q tal que $r_q < 0$ para determinar qual variável não básica irá se tornar básica.

Passo 3 Calcular as razões y_{i0}/y_{iq} para $y_{iq} \geq 0, i = 1, 2, \dots, m$. Se nenhum $y_{iq} \geq 0$, pare, o problema é ilimitado. Caso contrário, selecionar p como sendo o índice i correspondente à menor razão.

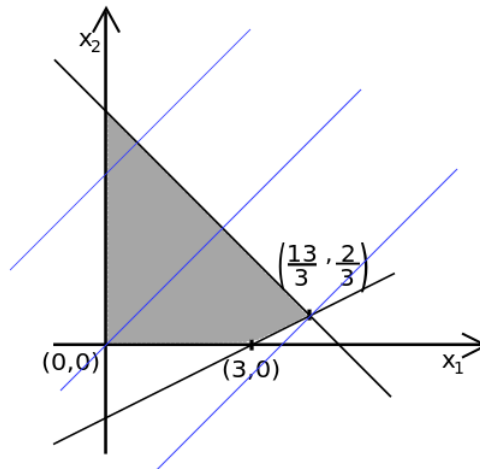
Passo 4 Pivotar no pq -ésimo elemento, atualizando todas as linhas incluindo a última. Retornar ao passo 1.

A prova de que o algoritmo resolve o problema, quando não há degeneração, é baseada na teoria que foi apresentada. O processo termina somente se for encontrada uma solução ou se o problema for degenerado. Como vimos, o número de soluções básicas viáveis é finito e nenhuma base é repetida por causa do decréscimo estrito da função objetivo, portanto, o algoritmo deve terminar, encontrando uma base que satisfaz uma das duas condições.

Exemplo 3.1.

$$\begin{aligned} \text{minimize} \quad & -x_1 + x_2 \\ \text{sujeito a} \quad & x_1 + x_2 \leq 5 \\ & x_1 - 2x_2 \leq 3 \\ & x \geq 0. \end{aligned}$$

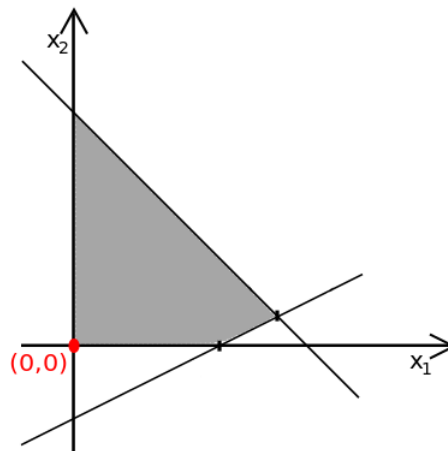
Podemos representar graficamente o conjunto viável para as restrições acima bem como as curvas de nível da função objetivo:



Primeiramente, escrevemos o problema na matriz do simplex, adicionando variáveis de folga, x_3 e x_4 :

$$\left[\begin{array}{cccc|c} a_1 & a_2 & a_3 & a_4 & b \\ \hline 1 & 1 & 1 & 0 & 5 \\ 1 & -2 & 0 & 1 & 3 \\ \hline -1 & 1 & 0 & 0 & 0 \end{array} \right]$$

Note que em relação às variáveis x_1 e x_2 , essa solução básica (correspondente às variáveis de folga) pode ser representada graficamente por:

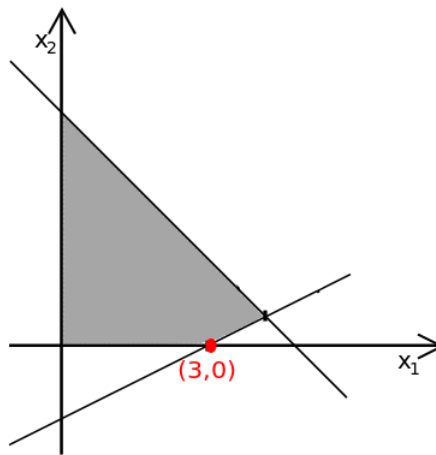


Escolhemos o menor valor dentre as 4 componentes da última linha. A coluna correspondente a este elemento será utilizada para pivotamento. Escolhemos a menor razão entre os

elementos das duas primeiras linhas da última coluna e os elementos das duas primeiras linhas da coluna escolhida, neste caso, a primeira (5/1 e 3/1). Portanto, a matriz atualizada fica:

$$\left[\begin{array}{cccc|c} a_1 & a_2 & a_3 & a_4 & b \\ \hline 0 & 3 & 1 & -1 & 2 \\ 1 & -2 & 0 & 1 & 3 \\ \hline 0 & -1 & 0 & 1 & 3 \end{array} \right]$$

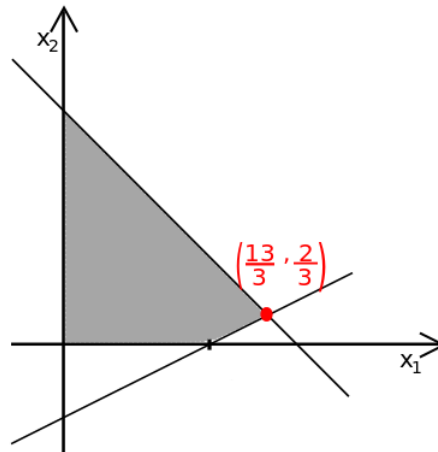
Agora, relação às variáveis x_1 e x_2 , essa solução básica corresponde a:



Utilizando o mesmo procedimento, temos mais um passo:

$$\left[\begin{array}{cccc|c} a_1 & a_2 & a_3 & a_4 & b \\ \hline 0 & 1 & 1/3 & -1/3 & 2/3 \\ 1 & 0 & 2/3 & 1/3 & 13/3 \\ \hline 0 & 0 & 1/3 & 2/3 & 11/3 \end{array} \right]$$

Cuja representação em relação às variáveis x_1 e x_2 é:



Como não temos mais elementos negativos nos 4 primeiros elementos da última coluna, o método termina. Desse modo, a solução será: $x_1 = 13/3$ e $x_2 = 2/3$ e o valor da função objetivo será $-11/3$.

3.2 As duas fases do método simplex

Para problemas como o apresentado no Exemplo (3.1), onde as restrições são todas de menor ou igual, uma solução básica inicial pode ser obtida imediatamente a partir das variáveis de folga adicionadas ao problema. Entretanto, isso nem sempre acontece e para iniciarmos o método simplex precisamos ter uma solução básica inicial. Para tanto, aplicamos o método simplex para o problema auxiliar de encontrar uma solução básica inicial. Seja o problema de PL na forma canônica:

$$\begin{aligned} Ax &= b \\ x &\geq 0. \end{aligned} \tag{3.2}$$

Para encontrar uma solução para o problema (3.2), construímos um problema auxiliar adicionando variáveis artificiais:

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^m y_i, \\ \text{sujeito a} \quad & Ax + y = b \\ & x \geq 0 \\ & y \geq 0 \end{aligned} \tag{3.3}$$

onde $y = (y_1, y_2, \dots, y_m)$ é um vetor com variáveis artificiais. Note que se o problema (3.3) possui valor mínimo igual a zero, com $y = 0$, podemos encontrar uma solução viável para (3.2).

Se o menor valor que o problema (3.3) atinge for positivo, então o problema (3.2) não possuirá solução viável.

Assim, resolvemos o problema (3.3) utilizando o método simplex. Uma solução viável inicial para este problema será $y = b$. Note que se o menor valor do problema (3.3) for 0, nenhuma das variáveis y_i será básica, isso significa que teremos uma solução básica apenas envolvendo as variáveis de x e assim, podemos retornar ao problema original e resolvê-lo a partir dessa solução básica inicial que obtivemos.

Exemplo 3.2. Dado o conjunto de restrições, encontre uma solução viável:

$$2x_1 + x_2 + 2x_3 = 4$$

$$3x_1 + 3x_2 + x_3 = 3$$

$$x_1, x_2, x_3 \geq 0.$$

Introduzimos variáveis artificiais $x_4, x_5 \geq 0$. Assim, a matriz inicial fica:

x_1	x_2	x_3	x_4	x_5	b
2	1	2	1	0	4
3	3	1	0	1	3
0	0	0	1	1	0

Transformando as variáveis x_4 e x_5 em variáveis básicas, ou seja, zerando o elemento da última linha da coluna 4 e da coluna 5, temos:

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

Escolhemos a primeira coluna para entrar na base, realizando as operações obtemos:

x_1	x_2	x_3	x_4	x_5	b
0	-1	4/3	1	-2/3	2
1	1	1/3	0	1/3	1
0	1	-4/3	0	5/3	-2

Escolhemos agora a terceira coluna para entrar na base, realizando as operações obtemos:

x_1	x_2	x_3	x_4	x_5	b
0	$-3/4$	1	$3/4$	$-1/2$	$3/2$
1	$5/4$	0	$-1/4$	$1/2$	$1/2$
0	0	0	1	1	0

Como as duas variáveis artificiais não fazem mais parte da solução básica, obtemos uma solução básica inicial para o problema, a saber: $x_1 = 1/2, x_2 = 0, x_3 = 3/2$.

Existem outras maneiras de se obter uma solução inicial. Chamamos esse problema de problema de viabilidade convexa. Uma opção é o método de projeções ortogonais sucessivas, que segundo Dounq, em [4] introduzido por Von Neumann, em 1933. Mais informações sobre este método podem ser obtidas em [4].

3.3 Número de iterações

Para um problema prático da forma:

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

com $m < 50$ e $m + n < 200$, segundo Chvatal, em [1], Dantzig, em 1963, registrou que o número de iterações é, normalmente da ordem de $3m/2$ e raramente atinge $3m$. Essa observação se confirma com os resultados obtidos experimentalmente recentemente. Na verdade, o número de iterações é proporcional a m . Até então acreditava-se que o método simplex e suas variações necessitavam de um número de iterações limitado por uma expressão polinomial do tamanho do problema.

De fato, segundo Chvatal, em [1], em 1963, realizou-se um estudo, liderado por H. W. Kuhn e R. E. Quandt, no qual eles resolveram uma quantidade de problemas onde a função objetivo a ser minimizada $c^T x$ era tal que $c = (1, 1, \dots, 1)$, $b_i = 10000$ para todo i e cada a_{ij} era obtido de maneira aleatória dentro de um conjunto de números inteiros entre 1 e 1000. Um dos resultados obtidos por exemplo, foi o de que em 100 problemas resolvidos com $m = n = 50$ foram necessárias em média 95.2 iterações.

Entretanto, ainda segundo Chvatal em [1], em 1972, V. Klee e G. J. Mint mostraram que para alguns problemas o método simplex percorreria exatamente $2^n - 1$ iterações até encontrar

uma solução ótima. Esses problemas são da forma:

$$\begin{aligned} &\text{maximize} \quad \sum_{j=1}^n 10^{n-j} x_j \\ &\text{sujeito a} \quad (2 \sum_{j=1}^{i-1} 10^{i-j} x_j) + x_i \leq 100^{i-1} \\ & \quad \quad \quad x_j \geq 0 \end{aligned}$$

com $i = (1, 2, \dots, n)$ e $j = (1, 2, \dots, n)$. Assim, se tivermos um problema desse tipo com $n = 50$, teremos $2^{50} - 1 \approx 10^{15}$ iterações. Se considerarmos que um computador realiza um milhão de iterações por segundo e que em 365 dias aproximadamente 3×10^7 segundos, se dividirmos a quantidade total de iterações pela quantidade de iterações realizadas em um ano ($3 \times 10^7 \times 10^6$), esse problema levaria

$$\frac{10^{15}}{3 \times 10^7 \times 10^6} \approx 33$$

33 anos para ser resolvido.

Exemplo 3.3. Para um problema da forma que Klee e Minty apresentaram com $n = 3$ esperamos encontrar o resultado após $2^3 - 1 = 7$ iterações.

$$\begin{aligned} &\text{maximize} \quad 100x_1 + 10x_2 + x_3 \\ &\text{sujeito a} \quad x_1 \leq 1 \\ & \quad \quad \quad 20x_1 + x_2 \leq 100 \\ & \quad \quad \quad 200x_1 + 20x_2 + x_3 \leq 10000 \\ & \quad \quad \quad x_1, x_2, x_3 \geq 0. \end{aligned}$$

Reescrevemos o problema transformando-o para a forma padrão:

$$\begin{aligned} &\text{minimize} \quad -100x_1 - 10x_2 - x_3 \\ &\text{sujeito a} \quad x_1 + x_4 = 1 \\ & \quad \quad \quad 20x_1 + x_2 + x_5 = 100 \\ & \quad \quad \quad 200x_1 + 20x_2 + x_3 + x_6 = 10000 \\ & \quad \quad \quad x_1, \dots, x_6 \geq 0. \end{aligned}$$

Assim, a matriz inicial do simplex fica:

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 20 & 1 & 0 & 0 & 1 & 0 & 100 \\ 200 & 20 & 1 & 0 & 0 & 1 & 10000 \\ \hline -100 & -10 & -1 & 0 & 0 & 0 & 0 \end{array} \right]$$

Após a primeira iteração temos:

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -20 & 1 & 0 & 80 \\ 0 & 20 & 1 & 200 & 0 & 1 & 9800 \\ \hline 0 & -10 & -1 & 100 & 0 & 0 & 100 \end{array} \right]$$

Após a segunda iteração temos:

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -20 & 1 & 0 & 80 \\ 0 & 0 & 1 & 200 & -20 & 1 & 8200 \\ \hline 0 & 0 & -1 & -100 & 10 & 0 & 900 \end{array} \right]$$

Após a terceira iteração temos:

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 20 & 1 & 0 & 0 & 1 & 0 & 80 \\ -200 & 0 & 1 & 0 & -20 & 1 & 8000 \\ \hline 100 & 0 & -1 & 0 & 10 & 0 & 1000 \end{array} \right]$$

Após a quarta iteração temos:

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 20 & 1 & 0 & 0 & 1 & 0 & 100 \\ -200 & 0 & 1 & 0 & -20 & 1 & 8000 \\ \hline -100 & 0 & 0 & 0 & -10 & 1 & 9000 \end{array} \right]$$

Após a quinta iteração temos:

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -20 & 1 & 0 & 80 \\ 0 & 0 & 1 & 200 & -20 & 1 & 8200 \\ \hline 0 & 0 & 0 & 100 & -10 & 1 & 9100 \end{array} \right]$$

Após a sexta iteração temos:

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -20 & 1 & 0 & 80 \\ 0 & 20 & 1 & -200 & 0 & 1 & 9800 \\ \hline 0 & 10 & 0 & -100 & 0 & 1 & 9900 \end{array} \right]$$

Após a sétima iteração temos:

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 20 & 1 & 0 & 0 & 1 & 0 & 100 \\ 200 & 20 & 1 & 0 & 0 & 1 & 10000 \\ \hline 100 & 10 & 0 & 0 & 0 & 1 & 10000 \end{array} \right]$$

E o algoritmo termina encontrando a solução $x_3 = 10000$, $x_4 = 1$, $x_5 = 100$ e $z = -10000$.

No exemplo anterior, utilizamos como critério para escolher a coluna que entra na solução básica como sendo o valor mais negativo da última coluna. Note que se no mesmo exemplo escolhêssemos na primeira iteração a coluna 3 para entrar na base, obteríamos a solução em apenas uma iteração. Dessa forma, poderíamos escolher outro critério para selecionar a coluna que deve entrar na solução básica. De fato, poderíamos utilizar o critério chamado “regra do maior crescimento”, ou seja, escolhemos a coluna cuja entrada na base nos daria o maior crescimento para a função objetivo. No caso dos problemas de Klee e Minty, com esse critério, teríamos uma solução em apenas uma iteração. Entretanto, em 1973, R. G. Jeroslow construiu problemas de PL nos quais se utilizarmos o critério de “maior crescimento” o número de iteração seria exponencial em relação a m e n . É claro que não basta olharmos apenas para o número de iterações. A “regra do maior crescimento” requer mais tempo computacional do que o critério do coeficiente mais negativo.

Problemas do tipo apresentados por Klee e Minty e outros que requerem um grande número de iterações, são chamados de “patológicos” pois raramente são encontrados na prática. Apesar desses casos apresentados, o método simplex funciona muito bem na prática, para problemas

de até 1 milhão de variáveis.

Da mesma forma, é necessário analisarmos se é possível que o método simplex entre um ciclo e não encontre uma solução ótima, o que de fato pode ocorrer, se não exigirmos a hipótese de não degenerabilidade.

Exemplo 3.4.

$$\begin{aligned}
 & \text{minimize} && -10x_1 + 57x_2 + 9x_3 + 24x_4 \\
 & \text{sujeito a} && 0.5x_1 - 5.5x_2 - 2.5x_3 + 9x_4 \leq 0 \\
 & && 0.5x_1 - 1.5x_2 - 0.5x_3 + x_4 \leq 0 \\
 & && x_1 \leq 1 \\
 & && x_1, x_2, x_3, x_4 \geq 0.
 \end{aligned}$$

Transformando para a forma padrão temos:

$$\begin{aligned}
 & \text{minimize} && -10x_1 + 57x_2 + 9x_3 + 24x_4 \\
 & \text{sujeito a} && 0.5x_1 - 5.5x_2 - 2.5x_3 + 9x_4 + x_5 = 0 \\
 & && 0.5x_1 - 1.5x_2 - 0.5x_3 + x_4 + x_6 = 0 \\
 & && x_1 + x_7 = 1 \\
 & && x_1, \dots, x_7 \geq 0.
 \end{aligned}$$

Assim, a matriz inicial do simplex fica:

$$\left[\begin{array}{cccccccc|c}
 0.5 & -5.5 & -2.5 & 9 & 1 & 0 & 0 & 0 & 0 \\
 0.5 & -1.5 & -0.5 & 1 & 0 & 1 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 \hline
 -10 & 57 & 9 & 24 & 0 & 0 & 0 & 0 & 0
 \end{array} \right]$$

Após a primeira iteração temos:

$$\left[\begin{array}{cccccccc|c}
 1 & -11 & -5 & 18 & 2 & 0 & 0 & 0 & 0 \\
 0 & 4 & 2 & -8 & -1 & 1 & 0 & 0 & 0 \\
 0 & 11 & 5 & -18 & -2 & 0 & 1 & 1 & 1 \\
 \hline
 0 & -53 & -41 & 204 & 20 & 0 & 0 & 0 & 0
 \end{array} \right]$$

Após a segunda iteração temos:

$$\left[\begin{array}{cccccc|cc} 1 & 0 & 0.5 & -4 & -0.75 & 2.75 & 0 & 0 \\ 0 & 1 & 0.5 & -2 & -0.25 & 0.25 & 0 & 0 \\ 0 & 0 & -0.5 & 4 & 0.75 & -2.75 & 1 & 1 \\ \hline 0 & 0 & -14.5 & 98 & 6.75 & 13.25 & 0 & 0 \end{array} \right]$$

Após a terceira iteração temos:

$$\left[\begin{array}{cccccc|cc} 2 & 0 & 1 & -8 & -1.5 & 5.5 & 0 & 0 \\ -1 & 1 & 0 & 2 & 0.5 & -2.5 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ \hline 29 & 0 & 0 & -18 & -15 & 93 & 0 & 0 \end{array} \right]$$

Após a quarta iteração temos:

$$\left[\begin{array}{cccccc|cc} -2 & 4 & 1 & 0 & 0.5 & -4.5 & 0 & 0 \\ -0.5 & 0.5 & 0 & 1 & 0.25 & -1.25 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ \hline 20 & 9 & 0 & 0 & -10.5 & 70.5 & 0 & 0 \end{array} \right]$$

Após a quinta iteração temos:

$$\left[\begin{array}{cccccc|cc} -4 & 8 & 2 & 0 & 1 & -9 & 0 & 0 \\ 0.5 & -1.5 & 0.5 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ \hline -22 & 93 & 21 & 0 & 0 & -24 & 0 & 0 \end{array} \right]$$

Após a sexta iteração temos:

$$\left[\begin{array}{cccccc|cc} 0.5 & -5.5 & -2.5 & 9 & 1 & 0 & 0 & 0 \\ 0.5 & -1.5 & -0.5 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ \hline -10 & 57 & 9 & 24 & 0 & 0 & 0 & 0 \end{array} \right]$$

Note que após a sexta iteração a matriz do simplex se torna igual a inicial e se o método prosseguir, fará as mesmas operações descritas acima.

Esse fenômeno só ocorre se possuímos solução básica degenerada, ou seja, quando uma solução básica assume o valor zero para alguma variável básica, pois para que o método se

prenda em um ciclo é necessário que o valor da função objetivo não se altere, o que só ocorre quando trazemos uma solução degenerada para a base.

3.4 Método simplex revisado

O método simplex revisado consiste em uma maneira de aplicarmos o método a um problema de programação linear, de forma a reduzir a quantidade de operações realizadas. Se um problema possuir muito mais variáveis do que restrições (se n for muito maior do que m) os pivôs irão aparecer em poucas colunas da matriz, como as outras colunas não são usadas explicitamente, não é necessário que atualizemos toda a matriz a cada iteração.

Para entendermos a teoria do método simplex revisado, precisamos reformular o método simplex, já apresentado, na forma matricial.

Considere o problema:

$$\begin{aligned} &\text{minimize} && c^T x \\ &\text{sujeito a} && Ax \geq b \\ &&& x \geq 0. \end{aligned} \tag{P}$$

Dada uma solução básica inicial para o problema primal, chamamos de B a submatriz da matriz original A composta pelas m colunas de A correspondentes às variáveis básicas dessa solução e de D a submatriz composta pelas colunas restantes. Vamos assumir que B é composta pelas m primeiras colunas de A (se não for, basta permutar as colunas de A). Então, separamos A , x e c :

$$\begin{aligned} A &= [B, D] \\ x &= (x_B, x_D), \quad c = [c_B, c_D] \end{aligned}$$

o problema na forma padrão se torna:

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

A solução básica inicial correspondente à base B é $x = (x_B, 0)$, onde $x_B = B^{-1}b$, como vimos no capítulo anterior. Nesse caso, obtemos essa solução fixando $x_D = 0$. Entretanto, para qualquer valor de x_D , o valor de x_B é calculado por:

$$x_B = B^{-1}b - B^{-1}Dx_D$$

se substituirmos essa expressão na função objetivo, temos:

$$z = c_B^T(B^{-1}b - B^{-1}Dx_D) + c_D^T x_D$$

$$z = c_B^T B^{-1}b + (c_D^T - c_B^T B^{-1}D)x_D$$

que nos dá o valor da função objetivo para qualquer solução do problema em termos de x_D . Então, temos

$$r = c_D^T = c_B^T B^{-1}D.$$

Escrevendo na forma matricial, temos inicialmente:

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

Que não está necessariamente na forma canônica. Usando a matriz B como base e realizando as operações para transformar a matriz na forma canônica obtemos:

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

O método simplex revisado consiste nos seguintes passos: dada uma solução básica inicial, sua base B , sua inversa B^{-1} , $x_B = y_0 = B^{-1}b$:

Passo 1 Calcular o valor do vetor r correspondente à solução atual, $r = c_D - \lambda B^{-1}D$, onde $\lambda = c_B B^{-1}$. Se $r \geq 0$ a solução atual é ótima.

Passo 2 Selecionar a coluna a_j para entrar na base de forma que j seja tal que r_j seja o mais negativo. Calcular $y_j = B^{-1}a_j$ que nos dá o vetor a_j em termos da base B .

Passo3 Calcular as razões y_{i0}/y_{ij} para determinar qual coluna sairá da base.

Passo 4 Atualizar B^{-1} e a nova solução $B^{-1}b$. Retornar ao primeiro passo.

Atualizamos B^{-1} com as operações usuais de pivotamento aplicadas a uma matriz composta de B^{-1} e y_j onde o elemento que se tornará pivô foi escolhido no passo 3. Podemos atualizar B^{-1} ao mesmo tempo adicionando-o como uma coluna desse matriz. Normalmente a matriz da solução inicial é uma matriz identidade e, conseqüentemente, sua inversa também será. Entretanto, se tivermos outra matriz correspondente à solução inicial, calculamos sua inversa explicitamente para iniciarmos o método simplex revisado.

Exemplo 3.5.

$$\begin{aligned}
 & \text{maximize} && 3x_1 + x_2 + x_3 \\
 & \text{sujeito a} && 2x_1 + x_2 + x_3 \leq 2 \\
 & && x_1 + 2x_2 + 3x_3 \leq 5 \\
 & && 2x_1 + 2x_2 + x_3 \leq 6 \\
 & && x_1, x_2, x_3, x_4 \geq 0.
 \end{aligned}$$

Claramente uma solução inicial é dada pelas variáveis de folga. Temos a matriz inicial:

a_1	a_2	a_3	a_4	a_5	a_6	b
2	1	1	1	0	0	2
1	2	3	0	1	0	5
2	2	1	0	0	1	6

com $c^T = [-3 \ -1 \ -3000]$. Iniciando o método simplex revisado temos:

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

Calculamos $\lambda^T = [0, 0, 0]B^{-1} = [0, 0, 0]$ e

$$(c_D - \lambda D)^T = [-3, -1, -3].$$

Para facilitar os cálculos, vamos escolher a coluna a_2 para entrar na base. Podemos representá-la multiplicando a_2 da matriz inicial por B^{-1} :

Variável	B^{-1}	x_B	y_2
4	1 0 0	2	1
5	0 1 0	5	2
6	0 0 1	6	2

Escolhendo a menor razão, pivotamos e atualizamos a matriz:

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

Assim,

$$\lambda^T = [-1, 0, 0]B^{-1} = [-1, 0, 0]$$

$$c_1 - z_1 = -1, c_3 - z_3 = -2 \quad e \quad c_4 - z_4 = 1.$$

Selecioneamos a_3 para entrar na base:

Variável	B^{-1}			x_B	y_2
2	1	0	0	2	1
5	-1	1	0	1	1
6	-2	0	1	2	-1

Após calcular a menor razão e pivotarmos, temos:

Variável	B^{-1}			x_B
2	3	-1	0	1
3	-2	1	0	1
6	-4	1	1	3

Assim,

$$\lambda^T = [-1, -3, 0]B^{-1} = [3, -2, 0]$$

$$c_1 - z_1 = -7, c_4 - z_4 = -3 \quad e \quad c_5 - z_5 = 2.$$

Agora escolhemos a_1 para entrar na base:

Variável	B^{-1}			x_B	y_2
2	1	-1	0	1	5
3	-2	1	0	1	-3
6	-4	1	1	3	-5

Após calcular a menor razão e pivotarmos, temos:

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

Assim,

$$\lambda^T = [-3, -3, 0]B^{-1} = [-6/5, -3/5, 0]$$

$$c_2 - z_2 = 7/5, c_4 - z_4 = 6/5 \quad e \quad c_5 - z_5 = 3/5.$$

Como todos os c_i 's são não negativos, temos a solução ótima: $x^T = (1/5, 0, 8/5, 0, 0, 4)$

Observe que para encontrar uma solução ótima utilizando o método simplex revisado não precisamos atualizar toda a matriz do simplex a cada iteração (como fazíamos anteriormente). Assim, podemos poupar alguns cálculos e encontrar a solução de maneira mais rápida.

4 Dualidade e problema dual

Para cada problema de programação linear, existe um problema dual correspondente. Esse par de problemas (primal e dual) se relaciona da seguinte forma: cada solução viável em um deles, limita uma solução ótima do outro. A *forma simétrica* da dualidade é dada pelo par de problemas definidos a seguir.

Definição 4.1. Dado o problema de programação linear, chamado de problema primal,

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{sujeito a} && Ax \geq b \\ & && x \geq 0 \end{aligned} \tag{P}$$

o respectivo problema dual é

$$\begin{aligned} & \text{maximize} && b^T \lambda \\ & \text{sujeito a} && A^T \lambda \leq c \\ & && \lambda \geq 0. \end{aligned} \tag{D}$$

Se A é uma matriz $m \times n$ então x é um vetor de dimensão n , b é um vetor de dimensão m , c é um vetor de dimensão n e λ é um vetor de dimensão m . O vetor x é a variável do problema primal e o vetor λ é a variável do dual.

Podemos obter o dual de qualquer problema de programação linear da seguinte forma:

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

que pode ser escrito da forma:

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{sujeito a} && Ax \geq b \\ & && -Ax \geq -b \\ & && x \geq 0. \end{aligned}$$

Usando um vetor dual particionado da forma $[u, v]$ o dual correspondente é :

$$\begin{aligned} & \text{maximize} && u^T b - v^T b \\ & \text{sujeito a} && u^T A - v^T A \leq c \\ & && u \geq 0 \\ & && v \geq 0 \end{aligned}$$

Tomando $\lambda = u - v$, obtemos o par de problemas na *forma antissimétrica* da relação de dualidade:

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{sujeito a} && Ax = b \\ & && x \geq 0 \\ \\ & \text{maximize} && b^T \lambda \\ & \text{sujeito a} && A^T \lambda \leq c. \end{aligned}$$

Note que aqui o vetor λ não está restrito a valores não negativos.

Em geral, se alguma das inequações no problema primal for transformada em equação, as componentes correspondentes de λ no problema dual se tornam variáveis livres. Se algum dos componentes de x no problema primal for uma variável livre, então a correspondente desigualdade em $\lambda^T A \leq c$ é transformada em igualdade no dual.

Considere o par de problemas acima na forma antissimétrica, podemos enunciar o seguinte lema:

Lema 4.1 (Teorema fraco da dualidade). *Se x e λ são viáveis para os problemas acima, respectivamente, então, $c^T x \geq b^T \lambda$.*

Demonstração: Temos:

$$\lambda^T b = \lambda^T Ax \geq c^T x$$

que vale desde que $x \geq 0$ e $\lambda^T A \geq c^T$. ■

Corolário 4.1. *Se x_0 e λ_0 forem viáveis para os problemas acima, respectivamente, e se $c^T x_0 = y_0^T b$, então x_0 e λ_0 são ótimos para seus respectivos problemas.*

Teorema 4.1. Teorema da dualidade da programação linear *Se um dos problemas acima possuir uma solução finita ótima, então o outro também possuirá, e os valores correspondentes das funções objetivos são iguais. Se um dos problemas for ilimitado, então o outro não possuirá solução viável.*

A segunda parte desse teorema é uma consequência imediata do lema 4.1. Note que se o problema primal for ilimitado e λ for viável para o dual, teremos $\lambda b \leq -M$ para qualquer M , o que é impossível. A demonstração da primeira parte deste teorema utiliza o Teorema do Hiperplano Separador, um resultado que não será mencionado neste trabalho. Daremos uma prova deste resultado, de forma menos geral, baseada na técnica do método simplex mostrando o teorema a seguir.

Teorema 4.2. *Se o problema de programação linear tem uma solução básica viável ótima correspondente à base B , então, o vetor $\lambda = c_B^T B^{-1}$ é uma solução ótima para o problema dual. Os valores ótimos para os dois problemas são iguais.*

Demonstração: É possível encontrarmos uma solução ótima para o problema dual através da solução obtida para o primal pelo método simplex. Para isso, vamos utilizar a formulação matricial do método simplex descrita no capítulo anterior.

Considere o problema primal e seu respectivo dual:

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{sujeito a} && Ax \geq b \\ & && x \geq 0 \end{aligned} \tag{P}$$

$$\begin{aligned} & \text{maximize} && b^T \lambda \\ & \text{sujeito a} && A^T \lambda \leq c. \end{aligned} \tag{D}$$

Conhecendo a solução básica viável ótima $(x_B, 0)$ correspondente à base B , vamos determinar uma solução para o problema dual em termos dessa base B .

Inicialmente, reescrevemos a matriz A como $A = [B, D]$. Como a solução básica viável $x_B = B^{-1}b$ é ótima, o vetor dos coeficientes r_j 's será não negativo em cada componente.

Da matriz do simplex temos que:

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

e como r_D é não negativo em cada componente, temos $c_B^T B^{-1} D \leq c_D^T$.

Tome $\lambda^T = c_B^T B^{-1}$. Essa escolha para λ resolve o problema dual, pois

$$\lambda^T A = [\lambda^T B, \lambda^T D] = [c_B^T, c_B^T B^{-1} D] \leq [c_B^T, c_D^T] = c^T$$

Como $\lambda^T A \leq c^T$, λ é viável para o dual. Por outro lado,

$$\lambda^T b = c_B^T B^{-1} b = c_B^T x_B$$

e portanto o valor da função objetivo dual para esse λ é igual ao valor do problema primal. Pelo lema 1, λ é ótimo. ■

Para obtermos uma solução ótima para o dual a partir de uma solução ótima encontrada pelo simplex para o primal, basta olharmos para B^{-1} , base dessa solução ótima. Se a solução viável básica inicial foi obtida através da adição das variáveis de folga, por exemplo, então na matriz final, B^{-1} aparece onde no início havia uma matriz identidade. Na última linha da matriz, as componentes correspondentes à essa matriz identidade serão $c_I^T - c_B^T B^{-1}$. Assim, subtraindo c_I^T dos elementos correspondentes na última linha, encontramos o valor da solução $\lambda^T = c_B^T B^{-1}$, do dual, multiplicada por -1 . No caso das variáveis de folga, citado acima, $c_I = 0$, então basta olhar para os elementos na última linha correspondentes a B^{-1} .

Exemplo 4.1. No exemplo 2, do capítulo 4 temos o problema primal e seu respectivo dual:

$$\begin{aligned} \text{minimize} \quad & -x_1 + x_2 \\ \text{sujeito a} \quad & x_1 + x_2 \leq 5 \\ & x_1 - 2x_2 \leq 3 \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned} \tag{P}$$

$$\begin{aligned} \text{maximize} \quad & 5\lambda_1 + 3\lambda_2 \\ \text{sujeito a} \quad & \lambda_1 + \lambda_2 \leq -1 \\ & \lambda_1 - 2\lambda_2 \leq 1. \end{aligned} \tag{D}$$

Matriz do simplex (inicial):

$$\left[\begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & b \\ \hline 1 & 1 & 1 & 0 & 5 \\ 1 & -2 & 0 & 1 & 3 \\ \hline -1 & 1 & 0 & 0 & 0 \end{array} \right]$$

Matriz do simplex (final):

$$\left[\begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & b \\ \hline 0 & 1 & 1/3 & -1/3 & 2/3 \\ 1 & 0 & 2/3 & 1/3 & 13/3 \\ \hline 0 & 0 & 1/3 & 2/3 & 11/3 \end{array} \right]$$

Logo, uma solução ótima para o problema dual é $\lambda_1 = -1/3$ e $\lambda_2 = -2/3$.

Note que $5(-1/3) + 3(-2/3) = -11/3$.

Uma solução ótima dos problemas primal e dual também satisfaz uma outra relação:

Teorema 4.3 (Folgas complementares - forma antissimétrica). *Seja x e λ soluções viáveis para os problemas primal e dual respectivamente, na forma antissimétrica. Uma condição suficiente e necessária para que ambos sejam soluções ótimas é que tenhamos, para todo i :*

$$i) \ x_i > 0 \implies \lambda^T a_i = c_i.$$

$$ii) \ x_i = 0 \iff \lambda^T a_i < c_i.$$

Demonstração: Se as condições impostas são satisfeitas, $(\lambda^T A - c^T)x = 0$. Então, $\lambda^T b = c^T x$ e pelo corolário do lema anterior, as duas soluções são ótimas. Reciprocamente, se duas soluções forem ótimas, pelo teorema da dualidade, temos que $\lambda^T b = c^T x$ e assim, $(\lambda^T A - c^T)x = 0$. Se cada componente de x for não negativa e cada componente de, $(\lambda^T A - c^T)$ for não positiva, as condições (i) e (ii) são satisfeitas. ■

Podemos reescrever o Teorema 4.3 para a forma simétrica do problema dual.

Teorema 4.4 (Folgas complementares - forma simétrica). *Sejam x e λ soluções viáveis para os problemas primal e dual respectivamente, na forma simétrica. Uma condição suficiente e necessária para que ambos sejam soluções ótimas é que para todo i e j*

$$i) \ x_i > 0 \implies \lambda^T a_i = c_i.$$

$$ii) \ x_i = 0 \iff \lambda^T a_i < c_i.$$

$$iii) \ \lambda_j > 0 \implies a^j x = b_j.$$

$$iv) \ \lambda_j = 0 \iff a^j x > b_j.$$

Onde a^j é a j -ésima linha de A .

4.1 Método simplex dual

Não é necessário construirmos uma nova matriz para aplicarmos o método simplex no problema dual. Utilizamos a mesma matriz construída no caso primal, apenas as operações serão alteradas, seguindo os critérios do algoritmo que apresentaremos a seguir. Essa técnica é denominada *método simplex dual*.

Dado o problema de programação linear

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

tome uma base B tal que $\lambda^T = c_B^T B^{-1}$ é viável para o dual. A solução básica no primal, $x_B = B^{-1}b$, é *dual viável* pois é viável para o dual. Se $x_B \geq 0$, então essa solução também será primal viável e, portanto, ótima.

Como o vetor λ é viável para o dual, satisfaz $\lambda^T a_j \leq c_j$, para $j = 1, 2, \dots, n$, permutando essa base para as m primeiras colunas de A , temos

$$\lambda^T a_j = c_j$$

para $j = 1, \dots, m$

e, exceto em caso de degeneração do dual, temos

$$\lambda^T a_j < c_j$$

para $j = m + 1, \dots, n$.

Encontramos então um vetor $\bar{\lambda}$ de modo que uma das equações se torne uma inequações e uma das inequações se torne uma equação e ao mesmo tempo o valor da função objetivo dual cresça. As m igualdades na nova solução determinam uma nova base. Chamando a i -ésima linha de B^{-1} de u^i , temos:

$$\bar{\lambda}^T = \lambda^T - \varepsilon u^i$$

temos $\bar{\lambda}^T a_j = \lambda^T a_j - \varepsilon u^i a_j$. Lembrando que $z_j = \lambda^T a_j$, e como $u^i a_j = \lambda_{ij}$ o ij -ésimo elemento da matriz, temos

$$\bar{\lambda}^T a_j = c_j, \quad \text{com } j = 1, 2, \dots, m \text{ e } i \neq j$$

$$\bar{\lambda}^T a_j = c_j - \varepsilon$$

$$\bar{\lambda}^T a_j = z_j - \varepsilon \lambda_{ij}, \quad \text{com } j = m + 1, \dots, n$$

e

$$\bar{\lambda}^T b = \lambda^T b - \varepsilon x_{B_i}.$$

Assim, podemos descrever o algoritmo:

Passo 1 Dada uma solução básica viável dual x_B , se $x_B \geq 0$ a solução é ótima. Se x_B é não negativo,

selecione um índice i tal que a i -ésima componente de $x_B, x_{B_i} < 0$.

Passo 2 Se todos os $\lambda_{ij} \geq 0, j = 1, 2, \dots, n$, então o dual não possui máximo. Se $\lambda_{ij} < 0$ para algum j , então seja

$$\epsilon_0 = \frac{z_k - c_k}{\lambda_{ik}} = \min \left\{ \frac{z_j - c_j}{\lambda_{ij}} : \lambda_{ij} < 0 \right\}$$

Passo 3 Formar uma nova base B substituindo a_i por a_k . Usando esta base, determine a solução básica dual viável correspondente x_B e retorne ao passo 1.

Exemplo 4.2.

$$\begin{aligned} \text{minimize} \quad & 3x_1 + 4x_2 + 5x_3 \\ \text{sujeito a} \quad & x_1 + x_2 + 3x_3 \geq 5 \\ & 2x_1 + 2x_2 + x_3 \geq 6 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{aligned}$$

Transformando o problema para a forma padrão, temos:

$$\left[\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & b \\ \hline -1 & -2 & -3 & 1 & 0 & -5 \\ -2 & -2 & -1 & 0 & 1 & -6 \\ \hline 3 & 4 & 5 & 0 & 0 & 0 \end{array} \right]$$

A solução básica é viável para o dual pois todos os elementos da última linha são positivos. Escolhemos uma variável para sair da solução básica, digamos $x_5 = -6$ e calculamos as razões: $-(3/-2)$, $-(4/-2)$ e $-(5/-1)$, e escolhemos a menor. Após a primeira iteração temos:

$$\left[\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & b \\ \hline 0 & -1 & -5/2 & 1 & -1/2 & -2 \\ 1 & 1 & 1/2 & 0 & -1/2 & 3 \\ \hline 0 & 1 & 7/2 & 0 & 3/2 & 9 \end{array} \right]$$

Após a segunda iteração temos:

$$\left[\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & b \\ \hline 0 & 1 & 5/2 & -1 & 1/2 & 2 \\ 1 & 0 & -2 & 1 & -1 & 1 \\ \hline 0 & 0 & 1 & 1 & 1 & 11 \end{array} \right]$$

$x_1 = 1, x_2 = 2, x_3 = 0$ e o valor da função será -11.

Note que dado um problema primal, o seu problema dual pode ter menos variáveis (ou vice e versa). Geralmente, um problema com menos variáveis se torna mais rápido de ser resolvido e pelo Teorema da Dualidade, sabemos que o valor da função objetivo do problema dual para uma solução ótima para o problema dual será igual ao valor da função objetivo para uma solução ótima para o primal. Assim dependendo do problema, esse resultado nos permite resolvê-lo de maneira mais rápida (utilizando o problema dual) obtendo uma solução ótima.

Conclusões e perspectivas

O desenvolvimento do método simplex trouxe grande avanço para a programação linear. Como apresentamos, o método simplex percorre as soluções básicas de um problema de PL, indo de um vértice a outro, definindo uma sequência de vértices com valores que se aproximam do ótimo. Esse método ainda é muito utilizado atualmente sendo sobrepujado por algoritmos de pontos interiores quando o problema é de grande porte.

Apesar da boa performance do simplex, nem sempre ele pode ser aplicado, como mostramos no exemplo proposto por Klee e Minty, em que simplex precisa percorrer todos os vértices para encontrar uma solução ótima. Nesses casos o número de passos do simplex é exponencial em relação ao tamanho do problema. A partir daí, acreditou-se que um bom algoritmo deveria conter um número “polinomial” de passos em relação ao tamanho do problema, ou seja, o tempo para computação seria limitado por uma função polinomial. Mais informações sobre complexidade de tempo podem ser encontradas em [9].

Gonzaga cita em [5] que Khachiyan anunciou em 1979 o método dos elipsóides, de tempo polinomial, que constrói uma sequência de elipsóides na qual cada novo elipsóide possui o volume menor do que o anterior, cada um contendo um conjunto de soluções, que converge para uma solução ótima. Assim, uma solução pode ser encontrada para qualquer grau de aproximação desejado continuando o processo. Entretanto, depois de alguns testes constatou-se que em quase todos os casos o simplex é muito mais rápido do que o método dos elipsóides. Mesmo assim, a descoberta não foi em vão uma vez que mostrou que existem algoritmos de tempo polinomial para a programação linear.

Mais tarde, segundo Gonzaga, em [5], em 1984, foi anunciado o método de Karmarkar, de tempo polinomial. Esse método é classificado como sendo um método de pontos interiores, pois diferente do simplex, que percorre apenas os vértices do conjunto viável, ele percorre o conjunto viável pelo seu interior convergindo para uma solução ótima. A partir daí, vários pesquisadores buscaram um melhoramento desse método, o que permitiu que o método afim escala, publicado por Dikin, fosse redescoberto. Uma significativa melhora também foi obtida com o estudo de algoritmos de trajetória central.

A realização deste trabalho de conclusão de curso teve grande importância na minha formação

acadêmica uma vez que possibilitou o contato com uma área da Matemática que não consta no currículo do curso de graduação. Além disso, este trabalho impulsionou o contato com um ambiente de programação, no qual pude implementar os métodos apresentados, o que é essencial quando se trabalha com Matemática Aplicada. Um próximo passo, a partir daqui, seria dar continuidade ao estudo destes métodos, partindo para o estudo de métodos de pontos interiores.

Referências Bibliográficas

- [1] CHVATAL, Vasek. **Linear programming**. New York: W. H. Freeman, c1983. 478p.
- [2] COELHO, Flávio Ulhoa; LOURENÇO, Mary Lilian. **Um curso de álgebra linear**. São Paulo: Edusp, 2005. 261p.
- [3] DANTZIG, George Bernard. **Linear programming and extensions**. Princeton, N.J.: Princeton University Press, 1963.
- [4] DOUNG, Phan Canh; HUE, Le Thanh. **An alternating projections algorithm for solving linear programs**. Vietnã: Acta Mathematica Vietnamica, 2009. Volume 34, número 3, pp. 335-343.
- [5] GONZAGA, Clóvis Caesar. **Algoritmos de pontos interiores para programação linear**. Rio de Janeiro: Instituto de Matematica Pura e Aplicada, 1989.
- [6] LUENBERGER, David G. **Introduction to linear and nonlinear programming**. Reading: Addison-Wesley Publishing, c1973. xii, 356p
- [7] PUCCINI, Abelardo de Lima. **Introdução à programação linear**. Rio de Janeiro: Livros Técnicos e Científicos, 1987. 248p
- [8] RAMALHETE, Manuel; GUERREIRO, Jorge; MAGALHÃES, Alipio. **Programação linear**. Lisboa: McGraw-Hill de Portugal, c1984- 2v.
- [9] http://pt.wikipedia.org/wiki/Complexidade_de_tempo