Método Simplex

Prof. Alexandre Salles da Cunha

Universidade Federal de Minas Gerais
Departamento de Ciência da Computação
Belo Horizonte, Brasil

acunha@dcc.ufmg.br

Julho 2020

UFMG
UNIVERSIDADE FEDERAL
DE MINAS GERAIS

UFMG - ICEx
DEPARTAMENTO DE CIÊNCIA DA
COMPUTAÇÃO

Referências bibliográficas

Relativas à esta apresentação

1. D. Bertsimas e J. N. Tsitsiklis. Introduction to Linear Optimization, Athena Scientific, 1997. [Cap. 3]
2. R. J. Vanderbei. Linear Programming: Foundations and Extensions, Springer, 4a. Edição, 2014. [Caps. 2,3 e 4]
3. V. Chvátal. Linear Programming. Freeman, 1983. [Caps. 2 e 4].

1. Se um PPL possui uma solução ótima, então existe uma solução ótima que é básica viável para o poliedro que define a região de viabilidade do PPL.

2. Iniciando em uma solução básica viável, move-se para uma solução básica vizinha aprimorante (menor custo, no caso de minimização, maior lucro, no caso de maximização). Esta operação é chamada de pivoteamento.

3. Devidamente excluindo a possibilidade de ciclagem como consequência da degeneração, o processo se repete por um número finito de vezes até que uma de duas condições seja satisfeita:
   - Encontra-se uma direção (raio extremo) que permita melhorar indefinidamente a função objetivo, caracterizando um PPL ilimitado
   - Nenhuma solução básica vizinha é aprimorante. Logo, a otimalidade da solução básica atual é demonstrada.

PPL

$$
\begin{array}{lllll}
\max & 5x_1 & +4x_2 & +3x_3 & \\
 & 2x_1 & +3x_2 & +x_3 & \leq 5 \\
 & 4x_1 & +1x_2 & +2x_3 & \leq 11 \\
 & 3x_1 & +4x_2 & +2x_3 & \leq 8 \\
 & x_1, & x_2, & x_3 & \geq 0
\end{array}
$$

PPL no formato padrão

$$
\begin{array}{llllllll}
\max & 5x_1 & +4x_2 & +3x_3 & & & & \\
 & 2x_1 & +3x_2 & +x_3 & +x_4 & & & =5 \\
 & 4x_1 & +1x_2 & +2x_3 & & +x_5 & & =11 \\
 & 3x_1 & +4x_2 & +2x_3 & & & +x_6 & =8 \\
 & x_i & \geq 0 & & i & =1,\ldots,6 & &
\end{array}
$$

Solução (básica viável) inicial, corresp. ao dicionário $\{x_4, x_5, x_6\}$

"Zerando" os valores de x_1, x_2, x_3, mantendo no máximo $m = 3$ entradas não nulas para as demais variáveis (folgas).

$$
\begin{array}{llllll}
\max & v & = 0 & +5x_1 + 4x_2 & +3x_3 \\
 & x_4 & = 5 & -2x_1 - 3x_2 & -x_3 \\
 & x_5 & = 11 & -4x_1 - 1x_2 & -2x_3 \\
 & x_6 & = 8 & -3x_1 - 4x_2 & -2x_3
\end{array}
$$

- Observe que pela primeira linha do quadro acima, obtemos uma informação importante:

$$v \geq 0, \text{ pois } x_1, x_2, x_3 \geq 0.$$

- É possível melhorar esta solução ? Haveria vantagem em aumentar x_1, por exemplo, mantendo $x_2 = x_3 = 0$?
- O que limitaria o crescimento de x_1 ao manter $x_2 = x_3 = 0$?

Solução (básica viável) inicial

"Zerando" os valores de x_1, x_2, x_3, mantendo no máximo $m = 3$ entradas não nulas para as demais variáveis (folgas).

$$
\begin{array}{llll}
\max\ v & = 0 & +5x_1 + 4x_2 & +3x_3 \\
x_4 & = 5 & -2x_1 - 3x_2 & -x_3 \\
x_5 & = 11 & -4x_1 - 1x_2 & -2x_3 \\
x_6 & = 8 & -3x_1 - 4x_2 & -2x_3
\end{array}
$$

- Passo 1 (otimalidade): identificação da variável independente (não básica) que deve ser tornar dependente (básica): x_1.
- Passo 2 (teste da razão): diante do crescimento de x_1 e da manutenção das demais variáveis independentes no nível zero, qual a variável dependente que mais limita x_1 crescer ? x_4.
- Passo 3 (pivoteamento): x_1 entra no dicionário, x_4 sai. O sistema é reescrito.

Após a operação de pivoteamento: dicionário para $\{x_1, x_5, x_6\}$

$$
\begin{array}{lllll}
\max & v & = 12.5 & -2.5x_4 - 3.5x_2 & +0.5x_3 \\
 & x_1 & = 2.5 & -0.5x_4 - 1.5x_2 & -0.5x_3 \\
 & x_5 & = 1 & +2x_4 + 5x_2 & \\
 & x_6 & = 0.5 & +1.5x_4 + 0.5x_2 & -0.5x_3
\end{array}
$$

- Incrementar x_3 mantendo as demais variáveis independentes (não básicas) no nível atual ($x_4 = x_2 = 0$) é atrativo.
- Crescimento de x_3 é limitado pelo decrescimento de x_1, x_6.
- Teste da razão: a variável dependente x_6 se tornará independente, pois é a mais restritiva ao crescimento de x_3.

Após a operação de pivoteamento: dicionário para $\{x_1, x_5, x_3\}$

$$\begin{array}{llll} \max\ v & = 13 & -x_4 - 3x_2 & -x_6 \\ x_1 & = 2 & -2x_4 - 2x_2 & +x_6 \\ x_5 & = 1 & +2x_4 + 5x_2 & \\ x_3 & = 1 & +3x_4 + x_2 & -2x_6 \end{array}$$

- Passo 1 (teste de otimalidade):
  - A primeira linha, em conjunto com a não negatividade das variáveis, informa:

$$v \leq v + x_4 + 3x_2 + x_6 = 13 \rightarrow v \leq 13.$$

  - Para $x_1 = 2, x_5 = 1, x_3 = 1$ e $x_2 = x_4 = x_6 = 0$, $v = 13$.
  - Logo, provamos a otimalidade da solução atual.

O Problema a ser resolvido

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.:} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

- $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$
- a j-ésima coluna de A é designada por A_j
- a i-ésima linha de A é designada por a_i.

Observação: minimizar $c^T x$ equivale a maximizar $-c^T x$.
As funções objetivo são simétricas e o conjunto de otimizadores, caso existam, é o mesmo.

Uma idéia muito comum nos algoritmos de otimização

1. Dada uma solução viável, procure uma solução viável mais barata nas vizinhanças da solução atual.
2. Esta procura de soluções viáveis normalmente é feita tentando movimentar sobre uma *direção viável* aprimorante.
3. Se uma direção aprimorante não existe, a solução corrente é um ótimo local para o problema (não necessáriamente global).
4. O PPL é um problema de otimização convexa e todo ótimo local é global.
5. O Método Simplex, ilustrado no exemplo, explora exatamente esta idéia.

Seja x um ponto em um poliedro P. Um vetor $d \in \mathbb{R}^n$ é dito ser uma direção viável em x, se existir um escalar positivo θ para o qual $(x + \theta d) \in P$.

Obtendo uma direção viável conveniente

1. Vamos considerar a possibilidade de mover a partir de x para uma solução básica vizinha (adjacente).

2. Por definição, uma solução básica adjacente a x deve compartilhar $n-1$ restrições justas, cujos vetores a_i são linearmente independentes.

3. Como as restrições $a_i^T x = b_i, \forall i = 1, \ldots m$ $(Ax = b)$ precisam ser satisfeitas em qualquer solução viável, escolhemos uma direção d que permite alterar uma única variável não básica, mantendo as demais no nível atual.

4. Logo, para obter uma solução básica adjacente a x precisamos nos movimentar ao longo de uma direção viável que permita que apenas uma componente não básica de x torne-se básica.

A direção definida abaixo é denominada j-ésima direção básica

1. A base (e o dicionário) é definida pelo índice das variáveis básicas: $\{B(1), \ldots, B(m)\}$.

2. d_N: componentes não básicas $(i \notin \{B(1), \ldots, B(m)\})$
   - Vamos designar por j o índice da coluna não básica liberado para se tornar básica.
   - $d_j = 1$ para uma única variável não básica.
   - $d_i = 0$ para todas as demais variáveis não básicas, $i \neq j$.

3. d_B: componentes básicas $(i \in \{B(1), \ldots, B(m)\})$
   - Uma vez que $A(x + \theta d) \in P$ para algum $\theta > 0$, temos que $Ax + \theta Ad = b$ e então $Ad = 0$.
   - Então: $0 = Ad = \sum_{i=1}^{n} A_i d_i = \sum_{i=1}^{m} A_{B(i)} d_i + A_j = A_j + B d_B$.
   - Portanto $d_B = -B^{-1} A_j$ uma vez que B admite inversa.

Temos que nos preocupar exclusivamente com a variação das variáveis básicas, pois apenas uma variável não básica, x_j, será acrescida.

Quando x_j cresce, temos que garantir que $x_B + \theta d_B \geq 0$. Dois casos precisam ser considerados:

- x_B é não degenerada.
  Então $x_B > 0$ e então existe um $\theta > 0$ suficientemente pequeno que garante a viabilidade para $x_B + \theta d_B$.
- x_B é degenerada. Logo, existe $x_{B(i)} = 0$ para algum $i \in \{B(1), \ldots, B(m)\}$.
  Neste caso, se $d_{B(i)} < 0$, qualquer movimento no sentido de d_B faz com que a restrição $x_{B(i)} \geq 0$ seja violada. Assim sendo, a direção d neste caso não é viável, e consequentemente $\theta = 0$.

Efeito do deslocamento ao longo da direção básica d na função objetivo

- Variação da função objetivo: $\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} d_i = \sum_{i=1}^{n} c_i d_i$.
- Uma vez que $d_j = 1$, $d_i = 0$ para i não básica $i \neq j$ e $d_B = -B^{-1}A_j$:
  $\sum_{i=1}^{n} c_i d_i = c_j - c_B^T B^{-1} A_j$, onde $c_B^T = (c_{B(1)}, \ldots, c_{B(m)})$.
  - A parcela c_j é a taxa unitária de variação por permitir j aumentar.
  - A parcela $-c_B^T B^{-1} A_j$ é a compensação das mudanças nas variáveis básicas, ao garantir $A(x + \theta d) = b$.

Definição: Custo reduzido de Programação Linear

Seja x uma solução básica, B a base associada a x e c_B o vetor de custos das variáveis básicas. Para todo $j = 1, \ldots, n$, definimos o custo reduzido $\bar{c}_j$ da variável x_j como:

$$\bar{c}_j = c_j - c_B^T B^{-1} A_j$$

$$
\begin{aligned}
\min \quad & c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 \\
& x_1 + x_2 + x_3 + x_4 = 2 \\
& 2x_1 + 3x_3 + 4x_4 = 2 \\
& x \geq 0
\end{aligned}
$$

- Base escolhida $B = [A_1 \;\; A_2]$.
- Fixando $x_3 = x_4 = 0$ e resolvendo o sistema $Bx_B = b$, $x_B = (x_1, x_2)^T = (1, 1)^T$.
- Assumindo que d seja a direção básica associada à variável não básica x_3, temos:
  $$d_B = (d_1, d_2)^T = -B^{-1}A_3 = -\begin{bmatrix} 0 & 1/2 \\ 1 & -1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ \frac{1}{2} \end{bmatrix}.$$
- $\overline{c}_3 = c_3 - 3c_1/2 + c_2/2$

- Aplicando a definição para a variável básica $x_{B(i)}$, para algum $i \in \{1, \ldots, m\}$, associada à $i-$ésima coluna de B:

$$\overline{c}_{B(i)} = c_{B(i)} - c_B^T B^{-1} A_{B(i)} = c_{B(i)} - c_B^T e_i,$$

  onde e_i é um vetor m dimensional com todas as entradas nulas, exceto a i-ésima entrada que é 1.

- Logo, $\overline{c}_{B(i)} = c_{B(i)} - c_B^T e_i = c_{B(i)} - c_{B(i)} = 0$

Teorema

Considere uma solução básica x associada a uma base B. Assuma que $\overline{c}$ seja o correspondente vetor de custos reduzidos.

1. *Se $\overline{c} \geq 0$, então x é uma solução ótima.*
2. *Se x é ótima e não degenerada, então $\overline{c} \geq 0$.*

Obs: O resultado é intuitivo, dada nossa interpretação de custos reduzidos como a taxa de variação da função objetivo como função do deslocamento ao longo de uma direção.

Prova: (1)

- Vamos assumir $\bar{c} \geq 0$ e que y denote uma solução viável qualquer para o problema, com a qual obtemos uma direção viável $d = y - x$ (d é viável por convexidade).
- Viabilidade de x, y: $Ax = Ay = b$ e portanto $Ad = 0$.
- Então $0 = Ad = Bd_B + \sum_{i \in N} A_i d_i$ (N é o conj. dos índices das variáveis não básicas em x).
- Logo $d_B = -\sum_{i \in N} B^{-1} A_i d_i$
- E então $c^T d = c_B^T d_B + \sum_{i \in N} c_i d_i = \sum_{i \in N} (c_i - c_B^T B^{-1} A_i) d_i = \sum_{i \in N} \bar{c}_i d_i$.
- Para qualquer $i \in N$, $x_i = 0$. Logo, $d_i = y_i - x_i \geq 0$ e $c^T d = \sum_{i \in N} \bar{c}_i d_i \geq 0$

(2): por contradição

- Suponha que x seja uma solução básica viável ótima, não degenerada e que exista $j : \overline{c}_j < 0$.
- Então j necessariamente deve ser índice de uma variável não básica, uma vez que $\overline{c}_i = 0, \forall i \in \{B(1), \dots, B(m)\}$.
- Considere a $j-$ ésima direção básica, d.
  Esta direção é necessariamente viável dado que x é não degenerada. Ou seja, existe $\theta > 0$ tal que $x + \theta d$ é uma solução viável.
- Assim temos:

$$c^T(x+\theta d) = c^T x + \theta c^T d = c^T x + \theta \sum_{j \in N} \overline{c}_i d_i = c^T x + \theta \overline{c}_j < c^T x,$$

contrariando a otimalidade de x.

1. Uma solução ótima x pode ser tal que o custo reduzido de alguma variável não básica é negativo. Ainda não dispomos do certificado de otimalidade, que virá quando a *base certa*, associada à mesma solução básica viável ótima, for escolhida.

2. De acordo com o teorema para decidir se uma solução básica viável não degenerada x é ótima, precisamos verificar os custos reduzidos de suas $n - m$ componentes não básicas.

3. No caso de uma solução básica degenerada, um teste igualmente simples não é disponível.

Definição

Uma base B é ótima se as duas condições seguintes são satisfeitas:

1. $B^{-1}b \geq 0$
2. $\overline{c}^T = c^T - c_B^T B^{-1} A \geq 0.$

Observe que, diante desta definição:

1. Se uma base ótima B é encontrada, a solução básica associada é viável, satisfaz as condições de otimalidade dadas pelo Teorema anterior (independentemente de ser degenerada ou não) e portanto é ótima.
2. Por outro lado, uma solução básica viável ótima degenerada não necessariamente possui custos reduzidos não negativos, pois mais de uma base podem ser associadas à mesma solução básica degenerada. Mediante troca de bases, o mesmo ponto no espaço é representado por outra base, ótima.

Sobre a degeneração de P

De início, vamos assumir a hipótese de que todos as soluções básicas de $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ são não degeneradas. Em seguida relaxaremos esta hipótese.

1. Partimos de $x = (x_B, x_n)$, solução básica viável inicial não degenerada, associada à base $B = [A_{B(1)}, \ldots, A_{B(m)}]$, $x_N = 0, x_B = B^{-1}b > 0$.
2. Teste de otimalidade: Se $\overline{c}_j \geq 0 : \forall j \notin \{B(1), \ldots, B(m)\}$, x é uma solução básica ótima e resolve o PPL.
3. Caso contrário, existe $j \notin \{B(1), \ldots, B(m)\} : \overline{c}_j < 0$, x não é ótima.
4. A $j-$ésima direção básica $d = (d_B, d_N)^T : d_B = -B^{-1}A_j$, $d_j = 1$ e $d_k = 0, \forall k \notin \{j, B(1), \ldots, B(m)\}$ é uma direção viável aprimorante.
5. Ao longo de $\theta d, \theta > 0$, a componente x_j torna-se positiva, as componentes não básicas $x_k : k \neq j$ permanecem em 0 e as variáveis básicas variam em $-\theta B^{-1}A_j$.
6. Teste da razão: θ deve ser tão grande quanto possível.

$$x(\theta) = x + \theta d, \ \ \theta \geq 0$$

- Máximo deslocamento ao longo de d:

$$\theta^* = \max\{\theta \geq 0 : (x + \theta d) \in P\}.$$

- Como $Ad = 0$, $Ax(\theta) = b$, apenas as restrições $x_B \geq 0$ precisam ser verificadas:

  1. Se $d \geq 0$, $x(\theta) \geq 0$ para qualquer $\theta \geq 0$. Logo, $\theta^* = \infty$, o PPL é ilimitado, d é um raio extremo de $P : c^T d < 0$.
  2. Caso contrário, existe $i \in \{1, \ldots, m\} : d_{B(i)} < 0$. Para todo i tal que $d_{B(i)} < 0$, devemos garantir que $x_{B(i)} + \theta d_{B(i)} \geq 0$. Logo, $\theta \leq -\frac{x_{B(i)}}{d_{B(i)}}$.

- Então $\theta^* = \min\{-\frac{x_{B(i)}}{d_{B(i)}} : i = 1, \ldots, m, \ d_{B(i)} < 0\}$

- Observe que $\theta^* > 0$ uma vez que $x_B > 0$.

Assumindo que θ^* é finito, $y = x(\theta^*) = x + \theta^* d$.

- $y_j = x_j + \theta^* = -\frac{x_{B(l)}}{d_{B(l)}} > 0$,
  onde $B(l) \in \{B(1), \ldots, B(m)\}$ denota o índice da variável básica que definiu θ^*.
- $y_{B(i)} \geq 0$ para todo $i = 1, \ldots, m$.
- Em particular, para $B(l)$ temos:

$$y_{B(l)} = x_{B(l)} + \theta d_{B(l)} = x_{B(l)} - \frac{x_{B(l)}}{d_{B(l)}} d_{B(l)} = 0$$

o que sugere que a coluna A_j substitua a coluna $A_{B(l)}$ na base, uma vez que houve uma troca das restrições de não negatividade que ficaram justas.

Teorema

Considere os m índices $\{B(1), \ldots, B(l-1), j, B(l+1), \ldots, B(m)\}$ *e a correspondente matriz*

$$\overline{B} = [A_{B(1)}, \ldots, A_{B(l-1)}, A_j, A_{B(l+1)}, \ldots, A_{B(m)}]$$

de colunas de A. Veja que $\overline{B}$ *difere da base B associada à solução básica viável x apenas pela* $l-$*ésima coluna:* $A_{B(l)}$ *foi substituída por* A_j*. Então:*

1. *As colunas* $A_{B(i)} : i \neq l$ *e* A_j *são linearmente independentes e, assim,* $\overline{B}$ *é uma matriz básica.*
2. *O vetor* $y = x + \theta^* d$ *é a solução básica associada a* $\overline{B}$*.*

(1): por contradição

- Vamos supor que os vetores $A_{\overline{B}(i)}$ sejam l.d. Então, existem coeficientes $\lambda_1, \ldots, \lambda_m$, não todos nulos, tais que $\sum_{i=1}^{m} \lambda_i A_{\overline{B}(i)} = 0$. Multiplicando por B^{-1} temos $\sum_{i=1}^{m} \lambda_i B^{-1} A_{\overline{B}(i)} = 0$ e os vetores $B^{-1} A_{\overline{B}(i)}$ também são l.d. Vamos mostrar que este não é o caso.
- Observe que $B^{-1}B = I$. Logo $B^{-1}A_{B(i)} = e_i, i \neq l$ (e_i denota o vetor de zeros, exceto pela i-ésima entrada, que é 1). Todos os vetores $e_i = B^{-1}A_{B(i)} : i \neq l$ são l.i. e possuem um 0 na $l-$ésima entrada.
- Por outro lado $B^{-1}A_{\overline{B}(l)} = B^{-1}A_j = -d_B$.
- Por definição, $d_{B(l)} < 0$. Logo $B^{-1}A_j$ e $\{B^{-1}A_{B(i)} : i \neq l\}$ são l.i..
- Logo, $\overline{B}$ é uma base.

(2):

- Temos que $y \geq 0, Ay = b$.
- Além disto $y_i = 0, \forall i \notin \{\overline{B}(1), \ldots, \overline{B}(m)\}$.
- Acabamos de mostrar que $A_{\overline{B}(1)}, \ldots, A_{\overline{B}(m)}$ são l.i..
- Portanto, segue que $y = (y_{\overline{B}}, y_{\overline{N}})$ é uma solução básica viável associada a $\overline{B}$, y_B é solução única para $\overline{B}z = b$ e $y_{\overline{N}} = 0$.

$$\begin{aligned} \min \quad & c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 \\ & x_1 + x_2 + x_3 + x_4 = 2 \\ & 2x_1 + 3x_3 + 4x_4 = 2 \\ & x \geq 0 \end{aligned}$$

- Considerando a solução básica $x = (1, 1, 0, 0)$, $\overline{c}_3 = -3c_1 + \frac{c_2}{2} + c_3$.
- Se $c^T = (2, 0, 0, 0)^T$ temos que $\overline{c}_3 = -3$. A correspondente direção básica é $(-\frac{3}{2}, \frac{1}{2}, 1, 0)$.
- Considerando pontos da forma $x + \theta d, \theta > 0$, $\theta^* = -\frac{x_1}{d_1} = \frac{2}{3}$.
- O novo ponto obtivo é $y = x + \frac{2}{3}d = (0, \frac{4}{3}, \frac{2}{3}, 0)$.
- As colunas associadas às novas variáveis básicas, x_2, x_3, são l.i..

1. Uma vez que $\theta^* > 0$, y é uma solução básica distinta de x.
2. Uma vez que d é uma direção de redução de custos e então $c^T y < c^T x$.
3. Em parte, cumprimos nosso propósito inicial de mover para uma solução básica adjacente, de menor custo.
4. Vamos enunciar uma iteração típica do Método Simplex, definindo um vetor $u^T = (u_1, \ldots, u_m)^T$, de forma que $u = -d_B = B^{-1}A_j$, onde A_j é a coluna que entra na base. Ou seja, $u_i = -d_{B(i)} : i = 1, \ldots, m$.

1. Inicie a iteração com as colunas básicas $B(1), \ldots, B(m)$ e a solução básica viável associada x. Seja N o conjunto dos índices das variáveis não básicas.

2. Calcule os custos reduzidos $\overline{c}_j = c_j - c_B^T B^{-1} A_j, \forall j \in N$. Se $\overline{c} \geq 0$, páre, a solução x é ótima. Caso contrário, escolha $j \in N : \overline{c}_j < 0$.

3. Calcule $u = B^{-1} A_j$. Se $u \leq 0$, temos que $\theta^* = \infty$ e o custo ótimo é $-\infty$. Páre, o problema é ilimitado.

4. Se alguma componente de u é positiva, faça $\theta^* = \min_{i \in \{1,\ldots,m\} : u_i > 0} \frac{x_{B(i)}}{u_i}$.
   Seja l tal que $\theta^* = \frac{x_{B(l)}}{u_l}$.

5. Forme uma nova base substituindo $A_{B(l)}$ por A_j. Denote por y a nova solução básica viável. Então $y_j = \theta^*$ e $y_{B(i)} = x_{B(i)} - \theta^* u_i, \; i = 1, \ldots, m, i \neq l$.

Teorema

Assuma que o conjunto de soluções viáveis do PL é não vazio e que toda solução básica é não degenerada. Então, o Método Simplex termina após um número finito de iterações. Ao final, temos duas possibilidades:

1. *Dispomos de uma base ótima B e a correspondente solução básica ótima.*
2. *Encontramos um vetor $d : Ad = 0, d \geq 0$ e $c^T d < 0$. Portanto o PPL é ilimitado e o custo ótimo é $-\infty$.*

- Se o algoritmo termina no Passo 2, as condições de otimalidade propostas previamente foram satisfeitas e B é uma base ótima.
- Se o algoritmo termina no Passo 3, x é uma solução básica viável, $\overline{c}_j < 0$ e a direção básica associada a x_j, d, é um raio extremo satisfazendo $Ad = 0, d \geq 0$, $c^T d = \overline{c}_j < 0$, de forma que $x + \theta d \in P$ para qualquer $\theta > 0$ e o custo pode ser tão negativo quando desejemos.
- A cada iteração, $\theta^* > 0$ e $c^T d < 0$, o custo diminui e não podemos então visitar uma solução básica mais de uma vez.
- Como o número de soluções básicas é finito, o algoritmo então termina.

1. Se a solução básica corrente x for degenerada, θ^* pode ser nulo, caso $x_{B(l)} = 0$ e $d_{B(l)} < 0$. Neste caso, a nova solução obtida y é idêntica a x. Ainda assim, podemos definir uma nova matriz $\overline{B}$ com a entrada da coluna A_j e a saída de $A_{B(l)}$ de forma que $\overline{B}$ será uma base.

2. Mesmo se $\theta^* > 0$, pode acontecer de mais de uma variável básica se tornar zero no novo ponto $x + \theta^* d$. Uma vez que apenas uma destas variáveis básicas sairá da base para a entrada de x_j (e de A_j na nova base), a nova solução encontrada $y = x + \theta^* d$ será degenerada.

- As mudanças de base associadas à mesma solução básica podem, em alguns casos, não ser inúteis, levando a descobrir uma direção básica de redução de custo.
- Por outro lado, uma sequência de mudanças de base associadas ao mesmo ponto x pode levar à repetição das colunas associadas à representação da base, ocasionando **a ciclagem**.
- Para problemas altamente estruturados (fluxo em redes, etc) a possibilidade não é remota.
- A ciclagem pode ser evitada através do uso de regras adequadas de quais variáveis entrarão e/ou sairão da base.

Ilustração dos dois casos em uma solução degenerada.

Exemplo a seguir (Chvátal, p. 30)

1. PPL na forma de maximização
2. Regra pivoteamento determinística empregada:
    - Entrada na base: variável com custo reduzido mais positivo.
    - Em caso de empate no teste da razão: sai da base a variável com menor índice.

O Método Simplex pode não terminar ?

Dicionário 1, $B = \{5, 6, 7\}$, entra x_1, sai x_5

$$\begin{array}{rcrrrrrrrr}
x_5 & = & & - & 0.5x_1 & + & 5.5x_2 & + & 2.5x_3 & - & 9x_4 \\
x_6 & = & & - & 0.5x_1 & + & 1.5x_2 & + & 0.5x_3 & - & x_4 \\
x_7 & = & 1 & - & x_1 & & & & & & \\
\hline
z & = & & + & 10x_1 & - & 57x_2 & - & 9x_3 & - & 24x_4
\end{array}$$

Dicionário 2, $B = \{1, 6, 7\}$, entra x_2, sai x_6

$$\begin{array}{rcrrrrrrrr}
x_1 & = & & + & 11x_2 & + & 5x_3 & - & 18x_4 & - & 2x_5 \\
x_6 & = & & - & 4x_2 & - & 2x_3 & + & 8x_4 & + & x_5 \\
x_7 & = & 1 & - & 11x_2 & - & 5x_3 & + & 18x_4 & + & 2x_5 \\
\hline
z & = & & + & 53x_2 & + & 41x_3 & - & 204x_4 & - & 20x_5
\end{array}$$

O Método Simplex pode não terminar ?

Dicionário 3, $B = \{1, 2, 7\}$, entra x_3, sai x_1

$$
\begin{array}{rcrrrrrrrrr}
x_2 & = & & - & 0.5x_3 & + & 2x_4 & + & 0.25x_5 & - & 0.25x_6 \\
x_1 & = & & - & 0.5x_3 & + & 4x_4 & + & 0.75x_5 & - & 2.75x_6 \\
x_7 & = & 1 & + & 0.5x_3 & - & 4x_4 & - & 0.75x_5 & - & 13.25x_6 \\
\hline
z & = & & + & 14.5x_3 & - & 98x_4 & - & 6.75x_5 & - & 13.25x_6
\end{array}
$$

Dicionário 4, $B = \{2, 3, 7\}$, entra x_4, sai x_2

$$
\begin{array}{rcrrrrrrrrr}
x_3 & = & & + & 8x_4 & + & 1.5x_5 & - & 5.5x_6 & - & 2x_1 \\
x_2 & = & & - & 2x_4 & - & 0.5x_5 & + & 2.5x_6 & + & x_1 \\
x_7 & = & 1 & & & & & & & - & x_1 \\
\hline
z & = & & + & 18x_4 & + & 15x_5 & - & 93x_6 & - & 29x_1
\end{array}
$$

O Método Simplex pode não terminar ?

Dicionário 5, $B = \{3, 4, 7\}$, entra x_5, sai x_3

$$
\begin{array}{rcrrrrrrrrr}
x_4 & = & & - & 0.25x_5 & + & 1.25x_6 & + & 0.5x_1 & - & 0.5x_2 \\
x_3 & = & & - & 0.5x_5 & + & 4.5x_6 & + & 2x_1 & - & 4x_2 \\
x_7 & = & 1 & & & & & - & x_1 & & \\
\hline
z & = & & + & 10.5x_5 & - & 70.5x_6 & - & 20x_1 & - & 9x_2
\end{array}
$$

Dicionário 6, $B = \{4, 5, 7\}$, entra x_6, sai x_4

$$
\begin{array}{rcrrrrrrrrr}
x_5 & = & & + & 9x_6 & + & 4x_1 & - & 8x_2 & - & 2x_3 \\
x_4 & = & & - & x_6 & - & 0.5x_1 & + & 1.5x_2 & + & 0.5x_3 \\
x_7 & = & 1 & & & - & x_1 & & & & \\
\hline
z & = & & + & 24x_6 & + & 22x_1 & - & 93x_2 & - & 21x_3
\end{array}
$$

Este dicionário é idêntico ao inicial.

Dicionário 7, $B = \{5, 6, 7\}$, entra x_1, sai x_5

$$
\begin{array}{rcccccccccc}
x_6 & = & & - & 0.5x_1 & + & 1.5x_2 & + & 0.5x_3 & - & x_4 \\
x_5 & = & & - & 0.5x_1 & + & 5.5x_2 & + & 2.5x_3 & - & 9x_4 \\
x_7 & = & 1 & - & x_1 & & & & & & \\
\hline
z & = & & + & 10x_1 & - & 57x_2 & - & 9x_3 & - & 24x_4
\end{array}
$$

Até o momento, não particularizamos regra alguma para escolher as **regras de pivoteamento**, isto é, escolher qual variável entra e sai da base. Em particular, na descrição que apresentamos do Método Simplex, temos flexibilidade quanto à escolha de:

1. Qual coluna não básica com custo reduzido atrativo deve ser eleita para entrar na base;
2. Diante de mais de uma variável que definam θ^*, qual escolher para deixar a base.

Algumas prováveis candidatas a regras de pivotemento:

- Escolher para entrar na base, a coluna não básica A_j que possui o custo reduzido $\overline{c}_j$ mais negativo (min).
- Escoher para entrar na base, a coluna não básica A_j que promova o maior decréscimo na função objetivo, isto é $\theta^* \overline{c}_j$. Esta regra pode acelerar a queda da função objetivo, mais paga-se maior preço computacional, para se avaliar o θ^* associado a entrada de cada coluna que possua $\overline{c}_j < 0$. A experiência computacional existente sugere que o tempo global de cpu não decresce com a adoção desta estratégia.
- Mesmo a primeira estratégia acima pode ser muito cara para problemas de larga escala (requer que todas as colunas tenham seus custos reduzidos avaliados). Assim sendo, uma regra simples consiste na regra de escolher a coluna com custo reduzido negativo, de menor índice.

1. Os vetores $B^{-1}A_j$ são de fundamental importância para se obter:
   - Custos reduzidos ($\overline{c}_j = c_j - c_B^T B^{-1} A_j$)
   - A direção de busca d ($Bd_B = -A_j$)
   - O valor de θ^*, pelo teste da razão (depende de d_B)

2. A forma como esta informação é calculada e quais informações são preservadas de iteração para iteração diferenciam as implementações do método.

1. Nenhuma informação adicional será transportada de iteração para iteração no método.
2. Em cada iteração típica, conhecemos os índices $B(1), \ldots, B(m)$ das variáveis básicas.
3. Montamos a matriz B e resolvemos o sistema linear $p^T B = c_B^T$.
4. O custo reduzido para uma coluna j é calculado via $\bar{c}_j = c_j - p^T A_j$.
5. Dependendo da regra de pivoteamento adotado, podemos evitar calcular todos os custos reduzidos.
6. Uma vez que escolhemos a coluna A_j para entrar na base, resolvemos $Bu = A_j$ $(d_B = -u)$.
7. A partir de então, podemos calcular θ^* e definir qual variável sairá da base.

Implementação inocente - complexidade computacional

- Resolver os sistemas lineares ($p^T B = c_B^T$ e $Bu = A_j$): $O(m^3)$. Dependendo da estrutura da matriz B, este custo pode ser reduzido sensivelmente (exemplo: Problemas de Fluxo em Redes)
- Cálculo do custo reduzido para todas as colunas: $O(mn)$
- Custo total por iteração: $O(nm + m^3)$.
- Veremos como fazer a iteração a um custo de $O(m^2 + mn)$

- Boa parte do custo computacional por iteração é relacionado à resolução de dois sistemas lineares.
- Neste método, a matriz B^{-1} é explicitamente mantida em cada iteração e os vetores $c_B^T B^{-1}$ e $B^{-1}A_j$ são calculados via multiplicação matricial, ao custo de $O(m^2)$.
- Para ser eficiente, o método requer um esquema bastante eficiente de atualização da inversa da base B^{-1} (que custe menos de $O(m^3)$, caso contrário não haveria ganho) quando implementarmos uma operação de pivoteamento.

- Base no início da iteração:

$$B = [A_{B(1)}, \dots, A_{B(m)}]$$

- Base ao final da iteração:

$$\overline{B} = [A_{B(1)}, \dots, A_{B(l-1)}, A_j, A_{B(l+1)}, \dots, A_{B(m)}]$$

- Todas, exceto uma coluna de B e $\overline{B}$, são compartilhadas.
- Em função disto, é razoável esperar que B^{-1} contenha informação que possa ser usada na determinação de $\overline{B}^{-1}$.

Definição

Dada uma matriz, não necessariamente quadrada, a operação de adicionar um múltiplo de uma linha à mesma linha ou a outra linha da matriz é chamada de operação linha elementar (não muda as informações dos sistemas lineares representados pela matriz).

Exemplo

Digamos que queiramos multiplicar a terceira linha da matriz C abaixo por 2 para somar à primeira linha da matriz. Isto equivale a multiplicar C pela matriz Q dada abaixo.

$$C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} Q = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} QC = \begin{bmatrix} 11 & 14 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

- Multiplicar a $j-$ésima linha de uma matriz por β e adicionar este produto à $i-$ésima linha da matriz, equivale a pré mutiplicar a matriz por $Q = I + D_{ij}$, onde D_{ij} é uma matriz de zeros, exceto em sua entrada na $i-$ésima linha e $j-$ésima coluna que é β.
- Observe que a matriz Q possui determinante 1 e então admite inversa.
- Suponha que queiramos fazer uma sequência de K operações linha elementares, cada qual correspondendo à pré-multiplicação por uma matriz Q_k correspondente. Esta sequência de operações consiste em pré multiplicar a matriz de entrada por: $Q_K Q_{K-1} \dots Q_2 Q_1$.

- Uma vez que $B^{-1}B = I$, temos que $B^{-1}A_{B(i)} = e_i$.
- Usando esta observação temos $B^{-1}\overline{B} = \begin{bmatrix} 1 & & u_1 & & \\ & \ddots & \vdots & & \\ & & u_l & & \\ & & \vdots & \ddots & \\ & & u_m & & 1 \end{bmatrix}$
- Para transformar $B^{-1}\overline{B}$ na identidade:
  - Para cada $i \neq l$, adicionamos o produto da $l-$ ésima linha (linha pivot) por $\frac{-u_i}{u_l}$ e somamos à $i-$ ésima linha ($u_l \neq 0$, logo a operação substitui u_i por 0 na $i-$ésima linha).
  - Dividimos a $l-$ ésima linha por u_l.
- Seja Q a matriz que representa o produto destas m operações.
- Então $QB^{-1}\overline{B} = I \rightarrow \overline{B}^{-1} = QB^{-1}$.

Manutenção explícita da inversa da base

Vamos adicionar o seguinte passo ao final da iteração típica do Método Simplex:

- Forme a matriz $m \times (m+1)$ dada por: $[B^{-1} \mid u]$.
  Reproduza em B^{-1} as operações linha elementares necessárias para transformar u em e_l.
  As primeiras m colunas da matriz resultante fornece $\overline{B}^{-1}$.

- Observe que recalcular a inversa da base agora custa $O(m^2)$.

- Desta forma, a primeira inversão da base custa $O(m^3)$. As demais atualizações da base ocorrem ao custo de $O(m^2)$. O custo por iteração foi reduzido para $O(nm + m^2)$, se todos os custos reduzidos forem calculados.

- Vamos considerar $l = 3$ (a terceira coluna de B será alterada para compor $\overline{B}$) e

$$B^{-1} = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 3 & 1 \\ 4 & -3 & -2 \end{bmatrix} u = \begin{bmatrix} -4 \\ 2 \\ 2 \end{bmatrix}$$

- Então temos: $Q = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$ e

$$\overline{B}^{-1} = \begin{bmatrix} 9 & -4 & -1 \\ -6 & 6 & 3 \\ 2 & -1.5 & -1 \end{bmatrix}$$

- Regra lexicográfica. É adequada à implementação da versão que emprega todas as variáveis explícitas no quadro Simplex ou no Simplex Revisado, mantendo-se a inversa da base explícita. Eem algumas implementações bastante sofisticadas do Simplex Revisado, isto é abolido, e o uso da regra lexicográfica não é tão simples.

- Regra de Bland: muito simples de ser implementada.

Pivoteamento de Bland

1. Encontre o menor índice j para o qual o custo reduzido $\overline{c}_j$ é negativo para entrar na base.
2. Casa haja empate entre as variáveis básicas x_i que se anulam primeiro com o incremento da variável x_j, escolha aquela com o menor índice i para sair da base.

Observações:

- Esta regra não nos obriga a calcular todos os custos reduzidos.
- É compatível com a implementação do Simplex Revisado, na qual os custos reduzidos são calculados na ordem natural das variáveis.
- Provou-se não haver ciclagem com esta regra de pivotemento. Assim, o Método Simplex converge após um número finito de mudanças de base.

Encontrando uma base inicial viável

Assumimos dispor de uma solução inicial básica viável. Não descrevemos ainda como obter esta solução.

- Em alguns casos, encontrar esta solução é trivial. Exemplo: $P' = \{x \in \mathbb{R}^n_+ : Ax \leq b\}$ e o vetor $b \geq 0$. Neste caso, introduzindo as variáveis de folga temos $P = \{(x, s) \in (\mathbb{R}^n_+, \mathbb{R}^m_+) : Ax + Is = b\}$ e o vetor $(x, s)^T = (0, b)^T$ é uma solução básica viável.

- De um modo geral, encontrar esta solução não é tão trivial, requerendo a solução de um PPL auxiliar (Simplex Fase 1).

PPL Original

$$\begin{aligned} \min \qquad & c^T x \\ & Ax = b \\ & x \geq 0 \end{aligned}$$

Assumimos sem perda de generalidade que $b \geq 0$ (caso contrário, multiplicamos alguma linha do sistema linear por -1)

Problema Auxiliar (PA)

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} y_i \\ & Ax + y = b \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

- Introduzimos as variáveis artificiais $y \in \mathbb{R}^m_+$.
- Uma solução básica inicial para o Prob. Auxiliar é dada por $(x, y) = (0, b)$.
- Resolvemos PA. Se a função objetivo ótima for 0 sabemos que há solução viável para o problem original. Caso contrário, o problema original é inviável.
- Como obter a solução básica inicial para o problema original ?

Para iniciar o Método Simplex, necessitamos de uma solução básica inicial e a matriz básica associada. Vamos assumir que no ótimo de PA tenhamos $\sum_{j=1}^{m} y_j = 0$ (prob. original é viável).

Analisando a solução do PA:

1. Se ao final da resolução de PA as colunas básicas na solução ótima forem exclusivamente as variáveis x, basta tomarmos esta solução como uma sol. básica viável inicial (e as correspondentes colunas compondo a base) para iniciar o problema original.
2. Se existirem variáveis y_j básicas (necessariamente degeneradas) na solução ótima de PA, devemos fazer operações de pivoteamento forçado para removê-las da base.

- Vamos supor (como de costume) que $posto(A) = m$.
- Vamos assumir que $k < m$ colunas de A aparecem na base ótima de PA. Os índices das variáveis básicas associadas a estas colunas são $B(1), \ldots, B(k)$
- Vale observar que as colunas $A_{B(1)}, \ldots, A_{B(k)}$ são linearmente independentes, uma vez que estão na base.
- Como a matriz $A \in \mathbb{R}^{m \times n}$ possui posto m, suas colunas geram o $\mathbb{R}^m$. Logo, é possível escolher outras $m - k$ colunas de A, que sejam linearmente independentes a $A_{B(1)}, \ldots, A_{B(k)}$.
- Vamos identificar estas colunas, eliminando as colunas associadas a y da base.

- Vamos supor que a $l-$ésima variável básica ($l > k$) seja artificial. Dentre as variáveis $x_j : j \notin \{B(1), \ldots, B(k)\}$, procuramos uma coluna j (e variável x_j) tal que a $l-$ésima entrada do vetor $B^{-1}A_j$ seja distinto de 0.

- A_j é l.i. com $A_{B(1)}, \ldots, A_{B(k)}$.

- Então fazemos um pivoteamento forçado, no qual trazemos a coluna A_j para a base e eliminados a coluna da variável artificial da base.

- Repetimos o processo enquanto houver colunas básicas associadas a variáveis artificiais.

A complexidade do método Simplex depende:

1. da complexidade computacional em cada iteração ($O(nm)$, por exemplo, no caso do Simplex Revisado)
2. do número de iterações, ou operações de pivoteamento.

- Observação experimental média: ao longo dos anos observou-se que o método normalmente executa $O(m)$ operações de pivoteamento na média.
- Pior caso: Para todas as regras de pivoteamento conhecidas até o momento, há exemplos de problemas não degenerados nos quais o método percorre todos os vértices do poliedro até encontrar a solução ótima. O primeiro resultado patológico deste tipo deve-se a Klee e Minty (1972).

Questão em aberto

- Para várias regras populares de pivoteamento, há exemplos de PPLs e solução básica inicial que fazem o método percorrer todos os vértices.
- Entretanto, estes exemplos não excluem a possibilidade de existir uma regra de pivotemento que não apresente o comportamento de pior caso exponencial.
- Esta é uma das questões mais importantes em Teoria de Programação Linear, ainda em aberto.
- Então atenção: não se conhece uma regra de pivoteamento imune a esta patologia. Isto não significa que se provou que qualquer implementação do Método Simplex será exponencial.
- O Problema de Programação Linear é polinomialmente solúvel (Algoritmo do Elipsóide).