MAC-315 - Programação Linear

Prof. Marcelo Queiroz

http://www.ime.usp.br/~mqz

Notas de Aula[1]

3 O Método Simplex

Neste capítulo consideraremos sempre o problema de programação linear na forma canônica

$$(PLC)\begin{cases} \min & c'x \\ \text{s.a} & x \in P \end{cases}$$

onde

$$P = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\},$$

$b \in \mathbb{R}^m$ e $A \in \mathbb{R}^{m\times n}$ possui linhas linearmente independentes.

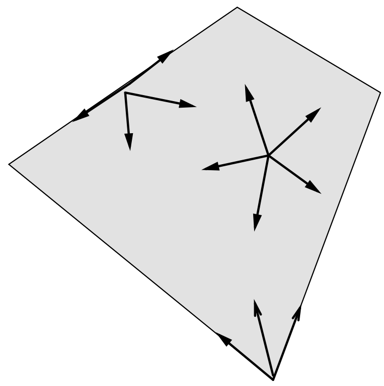
3.1 Condições de otimalidade

Uma estratégia comum a vários algoritmos de otimização consiste em construir uma sequência de pontos viáveis $x^0, x^1, \ldots, x^k$ que correspondem a melhorias sucessivas do valor da função ojetivo, com a propriedade de que x^{i+1} está "próximo" de x^i; se nenhum ponto nas proximidades de x^k melhoram o valor da função objetivo, o método pára e a solução obtida é *localmente ótima*.

Exercício 3.1 *Mostre que se a solução x^* é localmente ótima para o PLC, ou seja, se existe um $\varepsilon > 0$ tal que $c'x^* \leq c'(x^* + d)$ para qualquer $d \in \{d \in \mathbb{R}^n \mid \|d\| < \varepsilon$ e $x^* + d \in P\}$, então x^* é solução ótima do (PLC).*

A partir de um ponto $x \in P$, estamos interessados em considerar direções que não saem imediatamente do conjunto viável, como na figura a seguir:

[1] Baseadas no livro de Bertsimas & Tsitsiklis: *Introduction to Linear Optimization.*



Definição 3.1 *$d \in \mathbb{R}^n$ é chamada de* **direção viável** *a partir de $x \in P$ se existe $\theta > 0$ tal que $x + \theta d \in P$.*

Observe que na definição acima $x+\alpha d \in P$, $\forall \alpha \in [0, \theta]$, por convexidade. Se x é uma solução básica viável e $\{B_1, \ldots, B_m\}$ são os índices de uma base associada, temos

$$x_B = (x_{B_1}, \ldots, x_{B_m})' = B^{-1}b = \left[\begin{array}{c|c|c} A^{B_1} & \cdots & A^{B_m} \end{array}\right]^{-1} b$$

e $x_j = 0$ para todas as variáveis não básicas. Estamos interessados em restringir a busca ao conjunto das soluções básicas, pois este é um conjunto finito (corolário 2.1) que contém a solução ótima do PLC sempre que uma solução ótima existir (teorema 2.8).

Usaremos como noção de "proximidade" a definição de adjacência e buscaremos uma solução básica adjacente restringindo as direções d consideradas a manter ativas $n-1$ dentre as restrições ativas em x: além das m restrições que formam $Ax = b$, manteremos ativas $n - m - 1$ das restrições $x_i \geq 0$ ativas ($x_i = 0$); como são $n - m$ as variáveis não-básicas, podemos atingir este objetivo exigindo que $d_j = 1$ para alguma variável x_j não-básica, e $d_i = 0$ para todas as demais variáveis não-básicas. O novo ponto $y = x + \theta d$ terá que satisfazer

$$\begin{aligned} y_B &= x_B + \theta d_B \\ y_j &= x_j + \theta d_j = 0 + \theta = \theta \\ y_i &= x_i + \theta d_i = 0 + \theta 0 = 0, \forall i \neq j, B_1, \ldots, B_m. \end{aligned}$$

Como estamos interessados em soluções viáveis, exigiremos que $y \in P$ para algum valor de $\theta > 0$. Por um lado, isso força y a satisfazer $Ay = b$, ou seja

$$b = Ay = A(x + \theta d) = Ax + \theta Ad = b + \theta Ad \Longrightarrow Ad = 0;$$

lembrando que $d_j = 1$ e $d_i = 0, \forall i \neq j, B_1, \ldots, B_m$, temos $0 = Ad = \left(\sum_{i=B_1,\ldots,B_m} A^i d_i\right) + d_j A^j + \left(\sum_{i\neq j,B_1,\ldots,B_m} A^i d_i\right) = B d_B + A^j$ e portanto

$$d_B = -B^{-1}A^j.$$

Por outro lado, precisamos garantir que $y \geq 0$, ou seja, que $x + \theta d \geq 0$. Esta análise será feita na próxima seção.

A direção d acima construída, obtida fixando-se $d_j = 1$ e $d_i = 0, \forall i \neq j, B_1, \ldots, B_m$, é chamada de j-ésima direção básica. Com respeito a esta direção podemos distinguir duas situações:

1. x é uma solução não-degenerada: neste caso $x_B > 0$, e consequentemente para $\theta > 0$ pequeno teremos $x + \theta d \geq 0$. Em particular, d é uma direção viável de acordo com a definição 3.1;

2. x é degenerada: então existe um índice B_i tal que $x_{B_i} = 0$, e neste caso a direção construída pode ser viável ou não, dependendo do sinal de d_{B_i}.

Queremos observar o comportamento da função objetivo nas direções básicas. A variação do valor da função objetivo andando-se na j-ésima direção básica d será

$$c'(x + \theta d) - c'x = \theta c'd = \theta(c_j + c'_B d_B) = \theta(c_j - c'_B B^{-1} A^j).$$

Definição 3.2 *Seja x uma solução básica e B a matriz básica associada. Para cada índice j definimos o* **j-ésimo custo reduzido** *como*

$$\bar{c}_j = c_j - c'_B B^{-1} A^j.$$

Exemplo 3.1 *Considere o PLC*

$$\begin{array}{lccccccccc}
\min & c_1x_1 & + & c_2x_2 & + & c_3x_3 & + & c_4x_4 & & \\
s.a & x_1 & + & x_2 & + & x_3 & + & x_4 & = & 2 \\
 & 2x_1 & & & + & 3x_3 & + & 4x_4 & = & 2 \\
 & & & & x_1, \ldots, x_4 \geq 0 & & & & & .
\end{array}$$

Considere a base associada às variáveis $\{x_1, x_2\}$, com $B = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ e a solução correspondente $x = (1, 1, 0, 0)'$, que é não-degenerada. A direção básica associada a $j = 3$ é obtida fazendo-se $d_3 = 1$, $d_4 = 0$ e calculando-se

$$\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = -B^{-1}A^3 = -\begin{bmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ \frac{1}{2} \end{bmatrix};$$

O custo reduzido $\bar{c}_3$ corresponde à variação da função objetivo entre os pontos x e $x + d$, cujo valor é

$$\bar{c}_3 = c_3 - c'_B B^{-1} A^3 = c_3 + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}' \begin{bmatrix} -\frac{3}{2} \\ \frac{1}{2} \end{bmatrix} = -\frac{3}{2}c_1 + \frac{1}{2}c_2 + c_3.$$

Note que na definição de custo reduzido, se j é um índice básico, então $\bar{c}_j = c_j - c'_B B^{-1} A^j = c_j - c'_B e^j = c_j - c_j = 0$. Se existe uma variável não-básica x_j tal que $\bar{c}_j < 0$ e é possível andar na j-ésima direção básica sem sair do poliedro, então é possível passar para um ponto melhor. Do contrário, a solução considerada é ótima, como veremos a seguir.

Teorema 3.1 *Seja x uma solução básica viável e $\bar{c}$ o vetor de custos reduzidos. Então*

1. *Se $\bar{c} \geq 0$, x é solução ótima do PLC;*

2. *Se x é ótima e não-degenerada, então $\bar{c} \geq 0$.*

Prova.

Seja B a matriz básica associada a x. Seja $y \in P$ uma solução viável qualquer, e considere $d = y - x$. Temos $Ax = Ay = b \Longrightarrow Ad = 0$, e consequentemente $\sum_{i=1}^{n} A^i d_i = B d_B + \sum_{i \in N} A^i d_i = 0$, onde N é o conjunto de índices das variáveis não-básicas. Como B é inversível, podemos escrever

$$d_B = -\sum_{i \in N} B^{-1} A^i d_i$$

e

$$c'd = c'_B d_B + \sum_{i \in N} c_i d_i = \sum_{i \in N} (c_i - c'_B B^{-1} A^i) d_i = \sum_{i \in N} \bar{c}_i d_i.$$

1. Note que se $i \in N$, então $d_i = y_i - x_i \geq 0$, pois $x_i = 0$ (variável não-básica) e $y_i \geq 0$ (viabilidade); além disso $\bar{c}_i \geq 0$ por hipótese. Assim $c'd = \sum_{i \in N} \bar{c}_i d_i \geq 0$, ou seja $c'y \geq c'x$, o que mostra que x é ótimo.

2. Suponha por contradição que existe um índice (não-básico) j tal que $\bar{c}_j < 0$ e seja d a j-ésima direção básica a partir de x. Como x é não-degenerada, $x_B > 0$ e portanto existe um $\theta > 0$ tal que $x + \theta d \in P$. Mas

   $$c'(x + \theta d) = c'x + \theta c'd = c'x + \theta \Big(\sum_{i \in N} \bar{c}_i d_i\Big) = c'x + \theta \bar{c}_j < c'x,$$

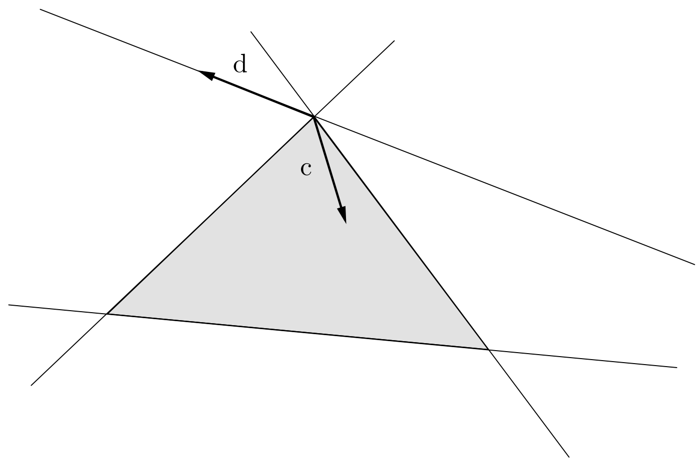
   onde a última igualdade é obtida lembrando-se que $d_j = 1$ e $d_i = 0, \forall i \in N \backslash \{j\}$; mas a desigualdade obtida contradiz a otimalidade de x. Concluimos, portanto, que $\bar{c} \geq 0$.

De acordo com o teorema acima, para concluir que uma solução básica é ótima basta calcular o custo reduzido para as $n - m$ direções básicas e verificar o sinal do mesmo. Isso motiva a seguinte

Definição 3.3 *Uma matriz básica é dita* **ótima** *se* $B^{-1}b \geq 0$ *e* $\bar{c}' = c' - c'_B B^{-1} A \geq 0'$.

No entanto, no caso de uma solução básica degenerada, não é possível concluir a partir do teorema que se existe $\bar{c}_j < 0$ então a solução necessariamente não é ótima: a j-ésima direção básica poderia não ser viável, e uma outra base associada ao mesmo vértice degenerado poderia estar associada a um (outro) vetor de custos reduzidos não-negativo. Em outras palavras, custos reduzidos não-negativos são uma condição *suficiente* de otimalidade, mas não *necessária.* O caso degenerado será tratado detalhadamente na seção seguinte.

Exercício 3.2 *Construa um exemplo de uma solução ótima degenerada e uma base que não satisfaz* $\bar{c} \geq 0$, *traduzindo o desenho a seguir para uma formulação como PLC;*



3.2 Desenvolvimento do método simplex

Consideraremos inicialmente o caso em que todas as soluções básicas do poliedro são não-degeneradas. Seja x uma solução básica e d a j-ésima direção básica a partir de x, satisfazendo $d_B = -B^{-1}A^j$, $d_j = 1$ e $d_i = 0$, $\forall i \neq j, B_1, \ldots, B_m$. Estamos interessados em analisar a viabilidade da solução $x + \theta d$: lembrando que $Ad = 0$ sabemos que $A(x + \theta d) = b$, $\forall \theta \in \mathbb{R}$; restaria verificar que $x + \theta d \geq 0$. Temos duas situações possíveis:

1. $d \geq 0$: neste caso $x + \theta d \geq 0$, $\forall \theta \geq 0$;

2. d possui alguma coordenada negativa: neste caso, existe um índice B_i tal que $d_{B_i} < 0$, e o valor máximo de θ que satisfaz $x + \theta d \geq 0$

será
$$\theta^* = \min_{B_i | d_{B_i} < 0} \left(-\frac{x_{B_i}}{d_{B_i}} \right);$$

Exemplo 3.2 *Voltando ao exemplo 3.1, tínhamos a solução básica $x = (1,1,0,0)'$ associada à base $\{x_1, x_2\}$ e a direção básica correspondente a $j = 3$ era $d = (-\frac{3}{2}, \frac{1}{2}, 1, 0)'$. Supondo que $c = (2,0,0,0)'$ teremos $\bar{c}_3 = -3 < 0$ e consequentemente a função objetivo decresce ao longo da direção d. Teremos*
$$\theta^* = \min_{B_i | d_{B_i} < 0} \left(-\frac{x_{B_i}}{d_{B_i}} \right) = -\frac{x_1}{d_1} = \frac{2}{3},$$
e consequentemente
$$y = x + \theta^* d = x + \frac{2}{3} d = (0, \frac{4}{3}, \frac{2}{3}, 0)'.$$
Note que as colunas A^2 e A^3 associadas a $y_i \neq 0$ são linearmente independentes, e y é a solução básica associada à base $\{x_2, x_3\}$. Em relação à base associada a x temos a entrada de x_3 na base e a saída de x_1: a entrada determinada pela escolha da direção básica associada a $j = 3$ e a saída determinada pelo índice que define θ^*, ou seja, pela primeira restrição a se tornar ativa no caminho de x para $x + \theta d$ (a restrição $x_1 \geq 0$). Note que y e x são adjacentes, pois possuem $m - 1 = 1$ variáveis básicas em comum (a variável x_2).

Supondo que $\theta^* = -\frac{x_{B_k}}{d_{B_k}} > 0$, teremos o novo ponto $y = x + \theta^* d$ que satisfaz
$$\begin{aligned}
&y_B = x_B + \theta^* d_B = x_B - \tfrac{x_{B_k}}{d_{B_k}} d_B \geq 0 \\
&y_{B_k} = x_{B_k} + \theta^* d_{B_k} = x_{B_k} - \tfrac{x_{B_k}}{d_{B_k}} d_{B_k} = 0 \\
&y_j = x_j + \theta^* d_j = -\tfrac{x_{B_k}}{d_{B_k}} > 0 \\
&y_i = x_i + \theta^* d_i = 0 + \theta^* 0 = 0, \forall i \neq j, B_1, \ldots, B_m.
\end{aligned}$$

Assim, substituindo a variável x_{B_k} pela variável x_j na base, obtemos a nova base $\{B_1, \ldots, B_{k-1}, j, B_{k+1}, \ldots, B_m\}$ associada a y, com a matriz básica
$$\bar{B} = \left[\begin{array}{c|c|c|c|c|c|c} A^{B_1} & \cdots & A^{B_{k-1}} & A^j & A^{B_{k+1}} & \cdots & A^{B_m} \end{array} \right].$$

Teorema 3.2

1. *As colunas A^{B_i} e A^j são linearmente independentes e $\bar{B}$ é uma matriz básica;*

2. *O vetor $y = x + \theta^* d$ é uma solução básica viável associada à matriz básica $\bar{B}$.*

Prova.

1. Suponha por contradição que $A^{B_i},\ i = 1, \ldots, k-1, k+1, \ldots, m$ e A^j sejam linearmente dependentes, ou seja, que exista $\lambda \neq 0$ tal que $\bar{B}\lambda = 0$. Então $B^{-1}\bar{B}\lambda = 0$, ou seja, os vetores $B^{-1}A^{B_i} = e^i,\ i = 1, \ldots, k-1, k+1, \ldots, m$ e $B^{-1}A^j = -d_B$ são linearmente dependentes. Mas $(B^{-1}A^j)_k = -d_{B_k} > 0$ por construção, e portanto λ satisfaz

$$\left\{\begin{array}{cccccc} \lambda_1 & & -\lambda_k d_{B_1} & & & = & 0 \\ & \ddots & \vdots & & & & \vdots \\ & \lambda_{k-1} & -\lambda_k d_{B_{k-1}} & & & = & 0 \\ & & -\lambda_k d_{B_k} & & & = & 0 \\ & & -\lambda_k d_{B_{k+1}} & \lambda_{k+1} & & = & 0 \\ & & \vdots & & \ddots & & \vdots \\ & & -\lambda_k d_{B_m} & & \lambda_m & = & 0. \end{array}\right.$$

   Da linha k concluimos que $\lambda_k = 0$ e das outras $\lambda = 0$. Esta contradição estabelece o item 1.

2. Por $Ad = 0$ temos $Ay = b$ e pela definição de θ^* temos $y \geq 0$. Além disso $y_i = 0,\ \forall i \neq \bar{B}_1, \ldots, \bar{B}_m$ e as colunas $A^{\bar{B}_1}, \ldots, A^{\bar{B}_m}$ são linearmente independentes pela parte 1. Isso estabelece a afirmação de 2.

Além disto, para o caso não-degenerado que estamos considerando temos que $y \neq x$ (pois $d \neq 0$ e $\theta^* > 0$), y é adjacente a x (pois possui $m-1$ variáveis básicas em comum) e $c'y < c'x$ (pois $c'd < 0$).

Podemos resumir **uma iteração típica do simplex** como segue:

1. No início da iteração conhecemos uma base formada pelas colunas $A^{B_1}, \ldots, A^{B_m}$ e o vértice associado x.

2. Calcule os custos reduzidos $\bar{c}_j = c_j - c_B' B^{-1}A^j$ para todos os índices não-básicos j. Se $\bar{c} \geq 0$, x é uma solução ótima e o algoritmo pára. Do contrário, escolha algum j tal que $\bar{c}_j < 0$.

3. Calcule $u = B^{-1}A^j$. Se $u \leq 0$ tem-se $\theta^* = \infty$, logo o custo ótimo é $-\infty$ e o algoritmo pára.

4. Se existe algum componente $u_i > 0$, define-se $\theta^* = \min_{u_i > 0} \dfrac{x_{B_i}}{u_i}$.

5. Seja k um índice tal que $\theta^* = \frac{x_{B_k}}{u_k}$. Forme uma nova base substituindo A^{B_k} por A^j. Os valores da nova solução básica y são dados por $y_j = \theta^*$ e $y_{B_i} = x_{B_i} - \theta^* u_i,\ \forall i \neq k$ (coordenadas básicas), além de $y_{B_k} = 0$ (não-básica).

A inicialização do simplex será discutida na seção 3.5. Por enquanto consideramos dada uma base viável inicial (lembre-se que para problemas canônicos viáveis em que A possui linhas l.i. sempre existe alguma base viável).

Teorema 3.3 *Suponha que o conjunto viável é não-vazio e que toda solução básica é não-degenerada. Nestas condições o método simplex termina em um número finito de iterações, com uma das seguintes propriedades:*

1. *B é uma base ótima e x é uma solução ótima; ou*

2. *Existe d tal que $Ad = 0$, $d \geq 0$ e $c'd < 0$ e o valor ótimo do problema é $-\infty$.*

Prova.

> A cada iteração intermediária o método passa de uma solução básica viável x para uma solução básica viável y que satisfaz $c'y < c'x$. Como o número de soluções básicas é finito (corolário 2.1) o método pára obrigatoriamente depois de um número finito de passos. Há dois critérios de parada: No passo 2, se todos os custos reduzidos são não-negativos, então B é uma base ótima (teorema 3.1 (a)). No passo 3, se $u \leq 0$ a j-ésima direção básica satisfaz $Ad = 0$ (por definição), $d \geq 0$ (porque $u \leq 0$) e $c'd = \bar{c}_j < 0$ (pelo passo 2); assim $x + \theta d \in P$ para qualquer $\theta \geq 0$ e $\lim_{\theta \to \infty} c'(x + \theta d) = -\infty$, o que mostra que o problema é ilimitado.

Note que este teorema é uma versão construtiva do teorema 2.8.

O método simplex para problemas degenerados

Considere a aplicação do método simplex descrito acima em problemas degenerados. Duas situações novas podem aparecer durante a execução do algoritmo:

1. Se o vértice x do início da iteração é degenerado, pode existir algum $u_i > 0$ com correspondente $x_{B_i} = 0$, e portanto $\theta^* = 0$; isto acontece quando a j-ésima direção básica é inviável, e neste caso a "nova" solução básica é $y = x$. A base associada a y no entanto é diferente da de x, pois foi obtida desta última através da substituição da coluna A^{B_k} por A^j; em particular a conclusão do teorema 3.2 ainda é válida.

2. Mesmo que tenhamos $\theta^* > 0$, pode acontecer de uma ou mais componentes básicas se anularem no novo ponto $x + \theta^* d$; isso acontece quando há empates do tipo $\frac{x_{B_l}}{u_l} = \frac{x_{B_k}}{u_k} = \theta^*$. Como apenas uma variável pode sair da base, as outras permanecerão na base com valor 0, e portanto o novo vértice y será degenerado.

Mudanças de bases associadas ao mesmo vértice podem levar o algoritmo a descobrir eventualmente uma direção básica viável com custo reduzido negativo. Por outro lado, pode-se obter uma repetição cíclica de bases associadas a um mesmo vértice, e o algoritmo não terminar nunca. Este fenômeno é chamado de *ciclagem* e pode ser evitado através de escolhas cuidadosas das variáveis que entram e saem da base nos casos de empate.

Seleção do pivô

Nos passos 2 e 5 do simplex não estão determinadas as escolhas da variável que entra na base (índice j tal que $\bar{c}_j < 0$) e da variável que sai da base (índice k tal que $\theta^* = \frac{x_{B_k}}{u_k}$). Regras determinísticas de escolha destas variáveis são chamadas de *regras de pivotação*.

Com relação à variável que *entra* na base, algumas escolhas são naturais:

1. Escolha o índice j associado ao custo reduzido $\bar{c}_j$ mais negativo. Esta escolha está associada à maior *taxa de decréscimo* da função objetivo; o decréscimo real depende do tamanho de passo θ^* que daremos naquela direção.

2. Escolha o índice j com o valor de $\theta_j^* \bar{c}_j$ mais negativo (aqui θ_j^* é o valor de θ^* associado à j-ésima direção básica). Esta escolha nos levará ao vizinho y de menor custo.

Estas duas regras são computacionalmente caras, pois é necessário calcular o custo reduzido associado a todas as variáveis não-básicas. Uma regra muito simples e que também é muito eficiente na prática é a **regra de Bland**, ou regra do menor índice, que escolhe o primeiro j (na ordem natural dentre os índices não-básicos) que satisfaz $\bar{c}_j < 0$.

A regra de Bland também pode ser aplicada na escolha da variável que sai da base: basta escolher o menor índice B_k que satisfaz $\theta^* = \frac{x_{B_k}}{u_k}$.

Veremos na seção 3.4 que o uso da regra de Bland garante que não ocorre ciclagem em problemas degenerados.

3.3 Implementações do método simplex

Vamos discutir algumas alternativas de implementação do método simplex que diferem essencialmente no modo como as contas (por exemplo custos reduzidos e direções básicas) são realizadas e na quantidade de informação que é aproveitada de uma iteração para a iteração seguinte.

Vamos contabilizar o número aproximado de operações em cada iteração usando a notação $\mathcal{O}(f(n))$, que pode ser entendida – *informalmente* – como a ordem ou taxa de crescimento da função; assim $\mathcal{O}(n)$ representa uma função qualquer que cresce (no máximo) linearmente e $\mathcal{O}(n^2)$ representa uma função com taxa de crescimento (no máximo) quadrático. Por exemplo, para calcular o produto interno $p'b$ onde $p, b \in \mathbb{R}^m$ são necessárias $\mathcal{O}(m)$ operações aritméticas; o produto matriz-vetor Bb requer $\mathcal{O}(m^2)$ operações; calcular a inversa de uma matriz B ou resolver um sistema linear da forma $Bx = b$ requer $\mathcal{O}(m^3)$ operações aritméticas.

Implementação ingênua

A implementação mais direta do método simplex como formulado na seção anterior leva em consideração que apenas os índices $B_1, \ldots, B_m$ das variáveis básicas são conhecidos no início de cada iteração. A matriz B é formada a partir das colunas $A^{B_1}, \ldots, A^{B_m}$ e calcula-se o vetor auxiliar

$$p' = c'_B B^{-1}$$

resolvendo-se o sistema linear $p'B = c'_B$; este vetor é chamado de vetor de *multiplicadores* associados à base B. O custo reduzido $\bar{c}_j = c_j - c'_B B^{-1} A^j$ associado à variável x_j é então calculado como

$$\bar{c}_j = c_j - p' A^j.$$

Dependendo da regra de pivotação utilizada pode ser necessário calcular todos os custos reduzidos, ou então calcular um por vez até encontrar o primeiro $\bar{c}_j$ negativo. Supondo que a variável x^j seja escolhida para entrar na base, calculamos $u = B^{-1}A^j$ resolvendo o sistema $Bu = A^j$, e determinamos a j-ésima direção básica. Finalmente calculamos θ^* e a variável a sair da base, para determinar a nova solução básica viável.

Note que para calcular p e u são necessárias $\mathcal{O}(m^3)$ operações aritméticas. Calcular o custo reduzido de uma variável custa $\mathcal{O}(m)$ operações, e como no pior caso teremos que calcular todos os custos reduzidos, teríamos um gasto de $\mathcal{O}(m(n - m)) = \mathcal{O}(mn)$. No total, teremos $\mathcal{O}(m^3 + mn)$ operações aritméticas.

Método simplex revisado

A parte computacionalmente mais custosa da implementação ingênua do simplex consiste em resolver dois sistemas de equações. Alternativamente, poderíamos considerar que a matriz B^{-1} é conhecida no início de uma iteração qualquer: isso tornaria os cálculos de p e u simples produtos matriz-vetor com um custo de $\mathcal{O}(m^2)$ operações aritméticas. Para esta alternativa ser viável, precisamos de um método eficiente de atualização de B^{-1} nas mudanças de base ao final de cada iteração.

Seja $B = \left[\begin{array}{c|c|c} A^{B_1} & \cdots & A^{B_m} \end{array}\right]$ a matriz básica disponível no início de uma iteração e $\bar{B} = \left[\begin{array}{c|c|c|c|c|c|c} A^{B_1} & \cdots & A^{B_{k-1}} & A^j & A^{B_{k+1}} & \cdots & A^{B_m} \end{array}\right]$ a matriz que queremos calcular para a próxima iteração. Estas duas matrizes possuem $n-1$ colunas em comum; é de se esperar que as matrizes B^{-1} e $\bar{B}^{-1}$ também possuam algum tipo de informação em comum. Em particular,

$$B^{-1}\bar{B} = \begin{bmatrix} 1 & 0 & \cdots & u_1 & \cdots & 0 \\ 0 & 1 & \cdots & u_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & & u_k & & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_m & \cdots & 1 \end{bmatrix};$$

portanto se conseguirmos descobrir uma matriz Q tal que $QB^{-1}\bar{B} = I$, então poderemos calcular $\bar{B}^{-1} = QB^{-1}$. Porém gostaríamos de fazer isso usando apenas $\mathcal{O}(m^2)$ operações, e isso só será possível se Q possuir uma estrutura muito particular, como veremos a seguir.

Definição 3.4 *Dada uma matriz qualquer, a operação que consiste em somar um múltiplo de uma linha qualquer a uma outra linha (ou à mesma linha) é chamada de* **operação elementar de linhas**.

Exemplo 3.3 *Seja* $Q \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ *e* $C = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 5 & 6 \end{bmatrix}$ *e considere o produto* $QC = \begin{bmatrix} 11 & 14 \\ 2 & 4 \\ 5 & 6 \end{bmatrix}$. *Note que a matriz Q representa a operação elementar de linhas que correponde a somar à primeira linha de C o dobro da terceira linha.*

Em geral, somar à i-ésima linha de uma matriz C o produto da j-ésima linha da mesma matriz por β é equivalente a multiplicar C à esquerda pela matriz $Q = I + D_{ij}$, onde D_{ij} é uma matriz com β na posição (i,j) e zero em todas as outras posições. Com exceção do caso (degenerado) $i = j$ e $\beta = -1$, que corresponde a apagar completamente a informação de uma das linhas não será utilizado, temos em todos os outros casos $\det(Q) \neq 0$ e portanto Q é inversível. Note ainda que para calcular o produto QC são necessárias apenas $\mathcal{O}(m)$ operações.

Considere agora que aplicamos uma sequência de operações elementares de linhas representadas pelas matrizes $Q_1, \ldots, Q_K$; o resultado final será equivalente a multiplicar à esquerda pela matriz $Q = Q_K \cdots Q_1$.

Examinando a estrutura da matriz $B^{-1}\bar{B}$ não é difícil encontrar uma sequência de operações elementares de linhas que transformam a matriz $B^{-1}\bar{B}$ na identidade:

1. Divida a k-ésima linha por u_k (lembre-se que $u_k = -d_{B_k} > 0$);

2. para cada $i \neq k$ subtraia a k-ésima linha multiplicada por u_i.

A sequência destas operações equivale a multiplicar $B^{-1}\bar{B}$ à esquerda por uma matriz Q e isso garante, como vimos antes, que $\bar{B}^{-1} = QB^{-1}$; isso mostra que para calcular $\bar{B}^{-1}$ basta aplicar a sequência de operações acima à matriz B^{-1}.

Exemplo 3.4 *Seja* $B^{-1} = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 3 & 1 \\ 4 & -3 & -2 \end{bmatrix}$, $u = \begin{bmatrix} -4 \\ 2 \\ 2 \end{bmatrix}$ *e suponha que* $k = 3$. *Queremos então transformar o vetor* u *no vetor* $e^3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. *Para isso, dividimos a terceira linha por 2, somamos à primeira linha o produto da terceira linha por 4 e subtraimos da segunda linha o produto da terceira linha por 2. Com isso, obtemos*

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

Atualizando a inversa da matriz básica desta maneira obtemos uma implementação do simplex conhecida como *método simplex revisado*:

0. No início da iteração conhecemos uma base formada pelas colunas $A^{B_1}, \ldots, A^{B_m}$ e a matriz inversa de base B^{-1}, além do vértice associado x.

1. Calcule o vetor $p' = c'_B B^{-1}$ os custos reduzidos $\bar{c}_j = c_j - p'A^j$ para todos os índices não-básicos j. Se $\bar{c} \geq 0$, x é uma solução

ótima e o algoritmo pára. Do contrário, escolha algum j tal que $\bar{c}_j < 0$.

2. Calcule $u = B^{-1}A^j$. Se $u \leq 0$ tem-se $\theta^* = \infty$, logo o custo ótimo é $-\infty$ e o algoritmo pára.

3. Se existe algum componente $u_i > 0$, define-se
$$\theta^* = \min_{u_i>0} \frac{x_{B_i}}{u_i}.$$

4. Seja k um índice tal que $\theta^* = \frac{x_{B_k}}{u_k}$. Forme uma nova base substituindo A^{B_k} por A^j. Os valores da nova solução básica y são dados por $y_j = \theta^*$ e $y_{B_i} = x_{B_i} - \theta^* u_i$, $\forall i \neq k$ (coordenadas básicas), além de $y_{B_k} = 0$ (não-básica).

5. Monte a matriz $[\ B^{-1} \mid u\ ] \in \mathbb{R}^{m\times(m+1)}$. Aplique a sequência de m operações elementares de linhas de modo a tornar a última coluna igual ao vetor e^k (k-ésimo vetor da base canônica); as primeiras m colunas da matriz correspondem à proxima inversa de base $\bar{B}^{-1}$.

Note que nesta versão o cálculo de p e u é feito por produtos matriz-vetor, que custam $\mathcal{O}(m^2)$ operações aritméticas. A atualização de B^{-1} corresponde a uma sequência de m operações elementares de linhas, cada uma das quais custando $\mathcal{O}(m)$, num total de $\mathcal{O}(m^2)$. Finalmente o cálculo dos custos reduzidos pode levar (no pior caso) $\mathcal{O}(mn)$. Cada iteração do simplex revisado custa portanto $\mathcal{O}(m^2 + mn)$ operações aritméticas.

Exemplo *Considere o problema*

$$\begin{cases} \min & -x_1 & -x_2 & (+0x_3 & +0x_4) & \\ s.a & x_1 & & +x_3 & & = 1 \\ & x_1 & +x_2 & & +x_4 & = 2 \\ & & & x \geq 0 & & \end{cases}$$

e considere a base inicial formada pelas variáveis x_3 e x_4.

1. *iteração:* $B^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $x = (0,0,1,2)'$. *Calculamos*

$$p' = c_B' B^{-1} = (0,0) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = (0,0),$$

$$\bar{c}_1 = c_1 - p'A^1 = -1 - (0,0) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -1$$

$$e\ \bar{c}_2 = c_2 - p'A^2 = -1 - (0,0) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -1.$$

Por Bland, escolhemos x_1 *para entrar na base e calculamos*

$u = B^{-1}A^1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

e $\theta^* = \min\{\frac{x_{B_1}}{u_1}, \frac{x_{B_2}}{u_2}\} = \min\{\frac{1}{1}, \frac{2}{1}\} = 1$*,*

de onde $k = 1$ *e* x_3 *sai da base. Fazemos*

$x \longleftarrow x + \theta^*(1, 0, -1, -1)' = \ (1, 0, 0, 1)'$

e calculamos a nova inversa da base $\{x_1, x_4\}$ *aplicando as operações elementares de linhas que transformam u em* $e^k = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$*:*

$$\begin{bmatrix} B^{-1} | u \end{bmatrix} = \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right] \rightsquigarrow \begin{bmatrix} \bar{B}^{-1} | e^k \end{bmatrix} = \left[\begin{array}{cc|c} 1 & 0 & 1 \\ -1 & 1 & 0 \end{array}\right].$$

2. iteração: $B^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$, $x = (1, 0, 0, 1)'$. *Calculamos*

$p' = c'_B B^{-1} = (-1, 0) \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = (-1, 0)$,

$\bar{c}_2 = c_2 - p'A^2 = -1 - (-1, 0) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -1$

e $\bar{c}_3 = c_3 - p'A^3 = 0 - (-1, 0) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1.$

Como $\bar{c}_2 < 0$*,* x_2 *entra na base; calculamos*

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

e $\theta^* = \frac{x_{B_2}}{u_2} = 1$*,*

de onde $k = 2$ *e* x_4 *sai da base. Fazemos*

$x \longleftarrow x + \theta^*(0, 1, 0, -1)' = (1, 1, 0, 0)'$

e calculamos a nova inversa da base $\{x_1, x_2\}$ *aplicando as operações elementares de linhas que transformam u em* $e^k = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Como $u = e^k$ *neste caso teremos* $\bar{B}^{-1} = B^{-1}$*: a base mudou (porque os índices mudaram), mas a matriz básica permaneceu igual (porque as colunas* A^2 *e* A^4 *são iguais).*

3. iteração: $B^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$, $x = (1, 1, 0, 0)'$. *Calculamos*

$p' = c'_B B^{-1} = (-1, -1) \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = (0, -1)$,

$$\bar{c}_3 = c_3 - p'A^3 = 0 - (0,-1)\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0$$

$$\text{e } \bar{c}_4 = c_4 - p'A^4 = 0 - (0,-1)\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1.$$

Como $\bar{c} \geq 0$, esta base é ótima. A solução ótima do problema é $x^* = (1,1,0,0)'$ com valor ótimo $c'x^* = -2$. Observe que como $\bar{c}_3 = 0$ e a direção básica associada $(-1,1,1,0)'$ é viável, todos os pontos do segmento que une $(1,1,0,0)'$ até $(0,2,1,0)'$ são ótimos.

Desenhe o poliedro em $\mathbb{R}^2$ associado a este problema (considerando x^3 e x^4 variáveis residuais) bem como o trajeto do simplex (sempre considerando as duas primeiras coordenadas de pontos e direções).

Implementação do simplex tabular

A próxima implementação a ser discutida é conhecida por *simplex tabular*. Ao invés de supor que se conhece no início de uma iteração qualquer apenas a matriz B^{-1}, esta implementação mantém atualizados os vetores $B^{-1}b$ (valores das variáveis básicas) e $B^{-1}A^1, \ldots, B^{-1}A^n$, armazenados no formato matricial

$$B^{-1}\,[b \mid A]$$

conhecido como *tableau* do simplex. A coluna $B^{-1}A^j = u$ associada à variável que entra na base e a k-ésima linha associada à variável x_{B_k} que sai da base são chamadas de *coluna* e *linha pivô*, respectivamente; o elemento $(B^{-1}A^j)_k = u_k$ é o *elemento pivô*, e é sempre positivo (a menos que o algoritmo pare por causa do critério de ilimitação $u \leq 0$).

A informação contida no tableau do simplex pode ser interpretada lembrando-se que, se B é uma matriz inversível qualquer, então

$$b = Ax \iff B^{-1}b = B^{-1}Ax,$$

para qualquer $x \in \mathbb{R}^n$. Os coeficientes que aparecem no tableau correspondem à representação do sistema linear acima.

Como no caso do simplex revisado, precisamos garantir que a informação transmitida para o proxima iteração, neste caso o tableau $B^{-1}\,[b \mid A]$, seja atualizada ao final de cada iteração. O tableau ao final de uma iteração qualquer deverá conter $\bar{B}^{-1}\,[b \mid A]$, o que pode ser obtido multiplicando-se o tableau original à esquerda pela matriz Q que satisfaz $QB^{-1} = \bar{B}^{-1}$. Ainda como no simplex revisado, isso é feito aplicando-se operações elementares de linhas que visam transformar a coluna $B^{-1}A^j$ na k-ésima coluna da identidade, ou ainda, transformando o elemento pivô em 1 e as demais coordenadas daquela coluna em 0.

A escolha da variável que sai da base, assim como o cálculo de θ^*, corresponde a considerar as frações $\frac{(B^{-1}b)_k}{(B^{-1}A^j)_k}\left(=\frac{x_{B_k}}{u_k}\right)$ para $(B^{-1}A^j)_k > 0$. A determinação de ilimitação está associada à condição $B^{-1}A^j \leq 0$.

Para prover o tableau do simplex com a informação necessária para a decisão da variável que entra na base, acrescentamos uma linha extra (linha 0) contendo as informações $-c'x = -c'_B x_B = -c'_B B^{-1}b$ (onde x é a solução associada à base atual) e o vetor $\bar{c}' = c' - c'_B B^{-1}A$:

$-c'_B x_B$	$\bar{c}_1$	$\cdots$	$\bar{c}_n$
x_{B_1}	$\vert$		$\vert$
$\vdots$	$B^{-1}A^1$	$\cdots$	$B^{-1}A^n$
x_{B_m}	$\vert$		$\vert$

e com isso temos no tableau toda a informação que o método simplex precisa. Veremos a seguir que a regra de atualização da linha 0 é idêntica à das outras linhas: adicionar à linha 0 um múltiplo da linha pivô de modo a zerar a entrada $\bar{c}_j$, que corresponde à variável que entra na base e fica justamente acima da coluna $B^{-1}A^j = u$ no tableau.

No início de uma iteração qualquer do simplex a linha 0 contém

$$[-c'_B x_B \mid \bar{c}'] = [0 \mid c'] - p'\,[b \mid A]$$

onde $p' = c'_B B^{-1}$. Suponha que apliquemos a atualização proposta à linha 0, que corresponde a somar à linha 0 o produto da linha pivô por um múltiplo positivo (de modo a zerar o próximo $\bar{c}_j$). A linha 0 teria então a nova forma

$$[0 \mid c'] - \bar{p}'\,[b \mid A]$$

onde $\bar{p}'$ corresponde a uma nova combinação linear das linhas de $[b \mid A]$. Lembrando que a variável j ocupa o lugar da variável B_k na próxima base, teremos

$$c_j - \bar{p}'A^j = c_{\bar{B}_k} - \bar{p}'A^{\bar{B}_k} = 0.$$

Considere os outros índices básicos B_i, $i \neq k$. A B_i-ésima coluna do tableau é $B^{-1}A^{B_i} = e^i$, assim o elemento da linha k e coluna B_i é 0 e portanto a atualização proposta não vai alterar o elemento $\bar{c}_{B_i}$, que vai continuar sendo 0:

$$c_{\bar{B}_i} - \bar{p}'A^{\bar{B}_i} = 0,\ \ i \neq k.$$

Com isso teremos $c_{\bar{B}_i} - \bar{p}'A^{\bar{B}_i} = 0,\ i = 1, \ldots m$ ou ainda $c'_{\bar{B}} = \bar{p}'\bar{B}$. Multiplicando os dois lados desta equação por $\bar{B}^{-1}$ concluimos que $\bar{p}' = c'_{\bar{B}}\bar{B}^{-1}$, ou seja, a nova linha 0 corresponde exatamente ao tableau associado à nova base $\bar{B}$.

Uma iteração do simplex tabular fica assim:

0. É conhecido o tableau associado a uma base B.

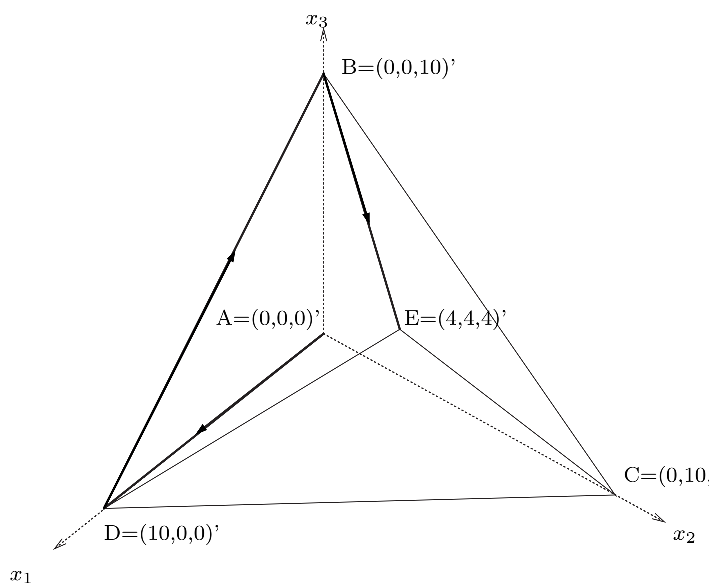
1. Examinamos a linha 0 buscando um custo reduzido negativo. Se tal custo não existir, temos uma base ótima; do contrário, escolhemos j tal que $\bar{c}_j < 0$.

2. Se a j-ésima coluna $B^{-1}A^j = u$ é não-positiva, o problema é ilimitado; do contrário, escolha k tal que $\frac{x_{B_k}}{u_k} = \min_{u_i>0}\{\frac{x_{B_i}}{u_i}\}$. A variável x_{B_k} sai da base e a variável x_j entra em seu lugar.

3. Some a cada linha do tableau um múltiplo positivo da linha pivô (linha k) de modo a tornar o elemento pivô (u_k) igual a 1 e todo o resto da coluna pivô (coluna j) igual a 0.

Exemplo 3.5 *Considere o problema*

$$\begin{cases} \min & -10x_1 - 12x_2 - 12x_3 \\ s.a & x_1 + 2x_2 + 2x_3 \leq 20 \\ & 2x_1 + x_2 + 2x_3 \leq 20 \\ & 2x_1 + 2x_2 + x_3 \leq 20 \\ & x \geq 0 \end{cases}$$



que na forma canônica fica

$$\begin{cases} \min & (-10, -12, -12, 0, 0, 0)x \\ s.a & \begin{bmatrix} 1 & 2 & 2 & 1 & 0 & 0 \\ 2 & 1 & 2 & 0 & 1 & 0 \\ 2 & 2 & 1 & 0 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 20 \\ 20 \\ 20 \end{bmatrix} \\ & x \geq 0 \end{cases}$$

Usaremos como solução básica viável inicial $x = (0, 0, 0, 20, 20, 20)'$ *associada à base* $\{x_4, x_5, x_6\}$. *Como* $c_B = 0$ *teremos* $[-c_B x_B \mid c -$

$c'_B B^{-1}A] = [0 \mid c]$, *e como* $B^{-1} = I$ *teremos o tableau inicial*

		$\downarrow$ x_1	x_2	x_3	x_4	x_5	x_6
	0	-10	-12	-12	0	0	0
$x_4 =$	20	1	2	2	1	0	0
$\leftarrow$ $x_5 =$	20	$\boxed{2}$	1	2	0	1	0
$x_6 =$	20	2	2	1	0	0	1

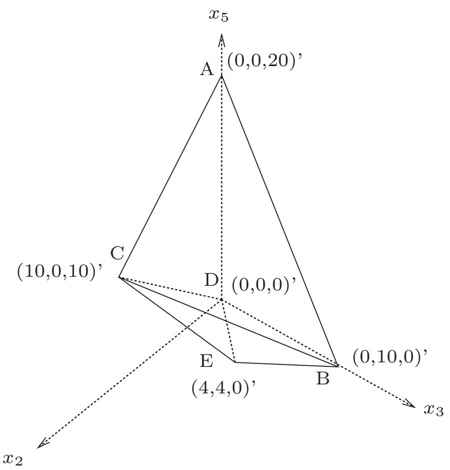
1. iteração: *Escolhemos* x_1 *para entrar na base* ($\bar{c}_1 = -10 < 0$), *e notamos que o* $\min\{\frac{x_4}{u_1}, \frac{x_5}{u_2}, \frac{x_6}{u_3}\}$ *é atingido para* $k = 2$ *e* $k = 3$. *Escolhendo* $k = 2$ (x_5 *sai da base), teremos que pivotar sobre o elemento da linha 2, coluna 1 do tableau (indicado pelas setas). A nova base* $\{x_4, x_1, x_6\}$ *está associada ao tableau*

		x_1	x_2	$\downarrow$ x_3	x_4	x_5	x_6
	100	0	-7	-2	0	5	0
$\leftarrow$ $x_4 =$	10	0	1.5	$\boxed{1}$	1	-0.5	0
$x_1 =$	10	1	0.5	1	0	0.5	0
$x_6 =$	0	0	1	-1	0	-1	1

e à solução $x = (10, 0, 0, 10, 0, 0)'$. *Note que os números do tableau correspondem ao problema* $\min -7x_2 - 2x_3 + 5x_5$ *sujeito a*

$$\begin{array}{rcccccccccc}
10 = & & & 1.5x_2 & + & x_3 & + & x_4 & - & 0.5x_5 & & \\
10 = & x_1 & + & 0.5x_2 & + & x_3 & & & + & 0.5x_5 & & \\
0 = & & & x_2 & - & x_3 & & & - & x_5 & + & x_6
\end{array}$$

que pode ser visualizado considerando-se as variáveis x_1, x_4 *e* x_6 *como residuais:*



2. iteração: *Voltando ao simplex, vamos escolher* $\bar{c}_3 = -2 < 0$ *para entrar na base (por Bland escolheríamos* $\bar{c}_2 = -7 < 0$). *Há um* *empate entre $\frac{x_4}{u_1}$ e $\frac{x_1}{u_2}$; escolhamos $k = 1$ como o índice que realiza o mínimo (x_4 sai da base). Pivotando sobre o elemento da linha 1, coluna 3, temos*

			$\downarrow$				
		x_1	x_2	x_3	x_4	x_5	x_6
	120	0	-4	0	2	4	0
$x_3 =$	10	0	1.5	1	1	-0.5	0
$x_1 =$	0	1	-1	0	-1	1	0
$\leftarrow$ $x_6 =$	10	0	2.5	0	1	-1.5	1

associado à solução $x = (0, 0, 10, 0, 0, 10)'$.

3. iteração: *Neste último tableau teremos $j = 2$ (o único índice tal que $\bar{c}_j < 0$) e $k = 3$ (pois $\frac{10}{2.5} < \frac{10}{1.5}$); assim x_2 entra na base e x_6 sai da base. O tableau final é*

		x_1	x_2	x_3	x_4	x_5	x_6
	136	0	0	0	3.6	1.6	1.6
$x_3 =$	4	0	0	1	0.4	0.4	-0.6
$x_1 =$	4	1	0	0	-0.6	0.4	0.4
$x_2 =$	4	0	1	0	0.4	-0.6	0.4

que corresponde à solução ótima $x^* = (4, 4, 4, 0, 0, 0)'$ (pois $\bar{c} \geq 0$). O caminho traçado pelo simplex pode ser acompanhado na figura do poliedro original como a sequência de pontos A-D-B-E. Esta sequência depende das escolhas das variáveis que entram e saem da base: se na segunda iteração tivéssemos usado Bland e escolhido x_2 para entrar na base, por exemplo, teríamos uma pivotação degenerada (linha 3, coluna 2) e passaríamos da base $\{x_4, x_1, x_6\}$ para a base $\{x_4, x_1, x_2\}$, associadas ao mesmo ponto D=(10,0,0,10,0,0)'. O caminho seria A-D-E, mas o número de iterações continuaria igual. Como o simplex só modifica um índice básico por iteração, é impossível passar da base inicial $\{x_4, x_5, x_6\}$ para a base ótima $\{x_1, x_2, x_3\}$ em menos de 3 iterações.

Complexidade do simplex tabular e comparação com o revisado

O simplex tabular mantém atualizado o tableau que tem tamanho $\mathcal{O}(mn)$. A atualização de cada entrada do tableau usa um número constante de operações (uma soma/subtração e um produto), portanto a atualização do tableau inteiro custa $\mathcal{O}(mn)$ operações aritméticas. A escolha da variável que entra ($\mathcal{O}(n)$) e da variável que sai ($\mathcal{O}(m)$) não afeta a ordem de grandeza da expressão mn, assim uma iteração do simplex tabular tem complexidade $\mathcal{O}(mn)$.

Lembramos que a complexidade do simplex revisado era $\mathcal{O}(m^2+mn)$ que representa exatamente a mesma ordem de grandeza que $\mathcal{O}(mn)$ (pois $m \leq n \Longrightarrow m^2 \leq mn$). Esta complexidade do simplex revisado foi calculada supondo-se que todos os custos reduzidos eram computados: utilizando-se Bland poderíamos ter uma situação em que o primeiro custo reduzido já é negativo e aquela iteração do simplex teria complexidade $\mathcal{O}(m^2)$ (correspondente a atualizar B^{-1} e calcular p' e u). Já no simplex tabular não é possível deixar de atualizar todo o tableau e melhorar a complexidade $\mathcal{O}(mn)$. Assim concluimos que uma iteração do simplex revisado não pode ser mais mais lenta que uma iteração do simplex tabular, e deve inclusive ser muito mais rápida na maioria das iterações.

Em relação ao uso da memória, as iterações do simplex revisado só armazenam a inversa da base, que ocupa espaço $\mathcal{O}(m^2)$. Já o tableau ocupa $\mathcal{O}(mn)$. Apesar da descrição do problema em princípio ter tamanho $\mathcal{O}(mn)$ (por causa da matriz A), na maioria dos problemas práticos (e grandes) a estrutura de A é esparsa e permite seu armazenamento eficiente; em contrapartida o tableau não será esparso em geral, mesmo que A e B o sejam, e o uso da memória no simplex tabular será sempre $\mathcal{O}(mn)$.

Resumindo esta discussão, temos as seguintes avaliações da complexidade destas duas implementações do simplex:
Revisado: Memória $\mathcal{O}(m^2)$, Tempo (melhor caso) $\mathcal{O}(m^2)$, Tempo (pior caso) $\mathcal{O}(mn)$;
Tabular: Memória $\mathcal{O}(mn)$, Tempo (melhor caso) $\mathcal{O}(mn)$, Tempo (pior caso) $\mathcal{O}(mn)$

Aspectos práticos de implementação

Algumas idéias específicas de melhoria da implementação do simplex estão associadas com aumentar a estabilidade numérica do método. Tanto o simplex revisado quanto o tabular partem de uma informação inicial obtida diretamente da descrição do problema, e passam a carregar informação de uma iteração para a outra acumulando erros, devido aos arredondamentos e truncamentos em função da representação finita dos valores de ponto flutuante.

Uma idéia neste sentido é a *reinversão*. Depois de um certo número de iterações do simplex revisado em que a inversa de base já acumulou uma certa quantidade de erros, a matriz B^{-1} é recalculada a partir dos valores originais das colunas $A^{B_1}, \ldots, A^{B_m}$. No caso do simplex tabular pode-se recomputar o tableau a partir da nova B^{-1} pela definição do tableau $(-c'_B x_B, \bar{c}, B^{-1}b, B^{-1}A^1, \ldots, B^{-1}A^n)$.

Outra idéia relacionada no caso do simplex revisado é, ao invés de atualizar a matriz B^{-1}, armazenar as m operações elementares de linhas que transformam B^{-1} em $\bar{B}^{-1}$, ou seja, os múltiplos da linha pivô de B^{-1} que foram somados a cada linha (de B^{-1}). Para calcular $u = \bar{B}^{-1}A^j$, por exemplo, poderíamos aplicar a mesma sequência de operações elementares ao vetor $B^{-1}A^j$, a um custo de $\mathcal{O}(m)$ operações aritméticas (supondo $B^{-1}A^j$ conhecido). Esta idéia pode ser utilizada ao longo de várias iterações, guardando a cada iteração o índice da linha pivô e os m coeficientes correspondentes à atualização da inversa de base. Supondo que se armazene todos os vetores $B^{-1}A^j$ onde B é a base inicial, o custo de se calcular p' e u após l iterações será $\mathcal{O}(lm)$; utilizando-se esta estratégia em sequências de L iterações (onde L é constante), depois das quais a matriz B^{-1} é computada explicitamente, os custos de computar p' e u serão $\mathcal{O}(m)$ em todas as iterações do método, o que representa uma grande economia de tempo.

Ciclagem e a regra de Bland

Exemplo 3.6 *Neste exemplo veremos que o simplex pode realmente ciclar dependendo das escolhas das variáveis que entram e saem da base em cada iteração. Considere o tableau inicial*

		$\downarrow$ x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-\frac{3}{4}$	20	$-\frac{1}{2}$	6	0	0	0
$\leftarrow$ $x_5 =$	0	$\boxed{\frac{1}{4}}$	-8	-1	9	1	0	0
$x_6 =$	0	$\frac{1}{2}$	-12	$-\frac{1}{2}$	3	0	1	0
$x_7 =$	1	0	0	1	0	0	0	1

e suponha que adotemos as seguintes regras: ***1)*** *a variável a entrar da base é sempre aquela com o menor custo reduzido (mais negativo);* ***2)*** *dentre as variáveis que podem sair da base escolhemos sempre aquela de menor índice. Então o primeiro pivô será o elemento na posição (1,1) e a sequência de tableaux obtida é*

		x_1	$\downarrow$ x_2	x_3	x_4	x_5	x_6	x_7
	3	0	-4	$-\frac{7}{2}$	33	4	0	0
$x_1 =$	0	1	-32	-4	36	3	0	0
$\leftarrow$ $x_6 =$	0	0	$\boxed{4}$	$\frac{3}{2}$	-15	-2	1	0
$x_7 =$	1	0	0	1	0	0	0	1

				$\downarrow$				
		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	0	0	-2	18	1	1	0
$\leftarrow$ $x_1 =$	0	1	0	8	-84	-12	8	0
$x_2 =$	0	0	1	$\frac{3}{8}$	$-\frac{15}{4}$	$-\frac{1}{2}$	$\frac{1}{4}$	0
$x_7 =$	1	0	0	1	0	0	0	1

					$\downarrow$			
		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$\frac{1}{4}$	0	0	-3	-2	3	0
$x_3 =$	0	$\frac{1}{8}$	0	1	$-\frac{21}{2}$	$-\frac{3}{2}$	1	0
$\leftarrow$ $x_2 =$	0	$-\frac{3}{64}$	1	0	$\frac{3}{16}$	$\frac{1}{16}$	$-\frac{1}{8}$	0
$x_7 =$	1	$-\frac{1}{8}$	0	0	$\frac{21}{2}$	$\frac{3}{2}$	-1	1

						$\downarrow$		
		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-\frac{1}{2}$	16	0	0	-1	1	0
$\leftarrow$ $x_3 =$	0	$-\frac{5}{2}$	56	1	0	2	-6	0
$x_4 =$	0	$-\frac{1}{4}$	$\frac{16}{3}$	0	1	$\frac{1}{3}$	$-\frac{2}{3}$	0
$x_7 =$	1	$\frac{5}{2}$	-56	0	0	-2	6	1

							$\downarrow$	
		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-\frac{7}{4}$	44	$\frac{1}{2}$	0	0	-2	0
$x_5 =$	0	$-\frac{5}{4}$	28	$\frac{1}{2}$	0	1	-3	0
$\leftarrow$ $x_4 =$	0	$\frac{1}{6}$	-4	$-\frac{1}{6}$	1	0	$\frac{1}{3}$	0
$x_7 =$	1	0	0	1	0	0	0	1

		$\downarrow$						
		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-\frac{3}{4}$	20	$-\frac{1}{2}$	6	0	0	0
$\leftarrow$ $x_5 =$	0	$\frac{1}{4}$	-8	-1	9	1	0	0
$x_6 =$	0	$\frac{1}{2}$	-12	$-\frac{1}{2}$	3	0	1	0
$x_7 =$	1	0	0	1	0	0	0	1

onde o último tableau é idêntico ao inicial e o simplex repetirá a mesma sequência de tableaux infinitamente. Observe que $\theta^* = 0$ *em todas as iterações e que a solução viável foi sempre* $x = (0, 0, 0, 0, 0, 0, 1)'$ *com custo 3. Apenas as bases que representam esta solução básica degenerada foram alternando-se ao longo da execução do algoritmo.*

Veremos a seguir que a regra de Bland garante que o método simplex não apresenta ciclagem.

Teorema *O método simplex utilizando a **regra de Bland:***

1. *dentre as variáveis candidatas a entrar na base, escolha a de índice* $j = \min\{i \mid \bar{c}_i < 0\}$*;*

2. *dentre as variáveis candidatas a sair da base, escolha a de índice* $B_k = \min\{B_l \mid u_l > 0 \text{ e } \frac{x_{B_l}}{u_l} = \theta^*\}$*;*

não apresenta ciclagem.

Prova.

A prova é feita por contradição: suponha que possa existir um ciclo usando a regra de Bland, formado pelas bases

$$B^{(1)}, B^{(2)}, \dots, B^{(l)}, B^{(1)}, B^{(2)}, \dots, B^{(l)}, \dots$$

onde todas as bases estão associadas a uma mesma solução degenerada $\bar{x}$.

constatação: toda variável que entra na base durante a ciclagem terá que sair antes de l iterações (altura em que entrará de novo na base).

consequência: dentre as variáveis que entram e saem de alguma base durante a ciclagem existe uma variável x_q de índice máximo.

Vamos explorar a estrutura do problema no momento em que x_q entrará na base. Seja B a matriz básica nesta iteração. Se x_q entrará na base pela regra de Bland, é porque

$$\bar{c}_q < 0 \quad \text{e} \quad \bar{c}_j \geq 0, \ \forall j < q.$$

Considere o problema equivalente

$$\begin{cases} \min & \bar{c}' x \\ \text{s.a} & \bar{A} x = \bar{b} \\ & x \geq 0 \end{cases}$$

onde $\bar{A} = B^{-1}A$, $\bar{b} = B^{-1}b$ e $\bar{c}' = c' - c'_B B^{-1} A$. Note que

$$\bar{c}' x \ = \overset{=0}{\bar{c}'_B}\, x_B + \sum_{\substack{j \in N \\ j < q}} \overset{\geq 0}{\bar{c}'_j}\, x_j + \overset{<0}{\bar{c}'_q}\, x_q + \sum_{\substack{j \in N \\ j > q}} \bar{c}'_j x_j,$$

e que as variáveis $\{x_j \mid j \in N \text{ e } j > q\}$ nunca entram nas bases do ciclo (porque x_q é a variável de índice máximo que entra e sai de alguma base).

Considere agora o momento em que x_q sairá da base, e seja $\bar{B}$ a matriz básica nesta iteração e x_p a variável que entrará no lugar de x_q. Podemos calcular os custos reduzidos com relação à formulação anterior: se x_p entrará na base, é porque

$$\bar{\bar{c}}_p = \bar{c}_p - \bar{c}'_{\bar{B}}\bar{B}^{-1}\bar{A}^p = \bar{c}_p - \bar{c}'_{\bar{B}}\bar{u} < 0,$$

onde $\bar{u} = \bar{B}^{-1}\bar{A}^p$. Como x_p é uma das variáveis que entram e saem da base, $p < q$ e consequentemente $\bar{c}_p \geq 0$; isso mostra que

$$\bar{c}'_{\bar{B}}\bar{u} = \bar{c}_{\bar{B}_1}\bar{u}_1 + \bar{c}_{\bar{B}_2}\bar{u}_2 + \cdots + \bar{c}_{\bar{B}_m}\bar{u}_m > 0.$$

Como x_q está na base $\bar{B}$ temos $q = \bar{B}_s$ (para algum s) e um dos termos da somatória acima é $\bar{c}_q\bar{u}_s$; lembrando que $\bar{c}_q < 0$ e que $\bar{u}_s > 0$ (pois este é o pivô), temos $\bar{c}_q\bar{u}_s < 0$. Como $\bar{c}_{\bar{B}_1}\bar{u}_1 + \bar{c}_{\bar{B}_2}\bar{u}_2 + \cdots + \bar{c}_{\bar{B}_m}\bar{u}_m > 0$ precisa existir um outro índice $\bar{B}_r$ tal que $\bar{c}_{\bar{B}_r}\bar{u}_r > 0$; como $x_{\bar{B}_r}$ participa de uma base do ciclo, $\bar{B}_r \notin \{j \in N \mid j > q\}$ e consequentemente $\bar{c}_{\bar{B}_r} \geq 0$, de onde

$$\bar{c}_{\bar{B}_r} > 0 \quad \text{e} \quad \bar{u}_r > 0.$$

Em particular, $x_{\bar{B}_r}$ não pertencia à base original B (pois $\bar{c}_B = 0$) e assim $\bar{x}_{\bar{B}_r} = 0$ na solução degenerada $\bar{x}$ associada ao ciclo. Mais ainda, $\bar{B}_r < q$, porque $x_{\bar{B}_r}$ entrou na base em algum momento do ciclo. Nos termos do simplex tabular teríamos a seguinte situação:

		$x_{\bar{B}_r}$	$\downarrow$ x_p	x_q
		0	$\bar{\bar{c}}_p$ <0	0
		0		0
$\leftarrow$ $x_q =$	0	0	$\bar{u}_s$ >0	1
		0		0
$\leftarrow$ $x_{\bar{B}_r} =$	0	1	$\bar{u}_r$ >0	0
		0		0

Mas as condições

$$\bar{u}_r > 0, \quad \frac{\bar{x}_{\bar{B}_r}}{\bar{u}_r} = 0 \quad \text{e} \quad \bar{B}_r < q$$

contradizem o fato de x_q ter sido escolhida pela regra de Bland para sair da base. Esta contradição mostra que a hipótese da existência do ciclo sob a regra de Bland não pode ser verdadeira.

conclusão: este argumento mostra que a regra de Bland garante a seguinte propriedade: *uma variável x_q só pode sair* *da base depois que entre na base alguma variável de índice maior do que q, que era não-básica quando x_q entrou na base.* Esta propriedade não é compatível com a existência de um ciclo, onde existe uma variável x_q com índice máximo dentre as variáveis que entram e saem da base.

3.4 Inicialização do simplex

Até agora consideramos a existência de uma base inicial dada. Em alguns casos uma tal base pode ser facilmente construída; se o problema foi obtido a partir de um problema da forma $Ax \leq b$ com $b \geq 0$, na forma canônica as colunas associadas às variáveis residuais formam uma matriz identidade, e a solução básica associada será $x_B = b \geq 0$, $x_N = 0$ que é viável.

No caso geral a matriz A não possui necessariamente uma base tal que $B = I$, e encontrar uma base viável inicial para o simplex é tão difícil quanto resolver um problema de programação linear. Considere o problema

$$\begin{cases} \min & c'x \\ \text{s.a} & Ax = b \\ & x \geq 0, \end{cases}$$

e suponha que $b \geq 0$ (como as restrições são de igualdade, podemos multiplicar algumas das linhas por -1 se necessário).

A técnica usual para produzir uma base viável inicial consiste em acrescentar uma matriz identidade à matriz original, identidade esta associada a m *variáveis artificiais*, e resolver o problema auxiliar:

$$\begin{cases} \min & y_1 + y_2 + \cdots + y_m \\ \text{s.a} & Ax + Iy = b \\ & x, y \geq 0. \end{cases}$$

Note que x é solução viável do problema original se e somente se $\begin{bmatrix} x \\ y=0 \end{bmatrix}$ é solução viável do problema auxiliar. A base associada às variáveis $y_1, \ldots, y_m$ produz a solução $x = 0$ e $y = b \geq 0$, que é viável no problema auxiliar. Note ainda que $y \geq 0$ garante que

$$y_1 + y_2 + \cdots + y_m = 0 \iff y_1 = y_2 = \cdots = y_m = 0,$$

de onde podemos afirmar que

o problema original é viável se e somente se o valor ótimo do problema auxiliar é zero.

Podemos então aplicar o simplex ao problema auxiliar a partir da base inicial $\{y_1, \ldots, y_m\}$ e observar o resultado: se a solução ótima $\begin{bmatrix} x^* \\ y^* \end{bmatrix}$

tem valor ótimo zero então $y^* = 0$ e x^* é viável no problema original; se o valor ótimo é positivo, o problema original é inviável (o que não deixa de ser uma resposta ao problema).

Neste ponto ainda não resolvemos *necessariamente* o problema de produzir uma base para o problema original. Suponha que o problema original é viável e que o método simplex produz uma solução ótima $\begin{bmatrix} x^* \\ y^*=0 \end{bmatrix}$. Se a base ótima produzida contiver apenas variáveis x_i originais, então ela será uma base viável para o problema original (pois será formada por m colunas de A linearmente independentes, associadas à solução x^* viável). Por outro lado se a base ótima contiver variáveis artificiais y_i (com $y_i^* = 0$), as colunas correspondentes na matriz básica associada não pertencem à matriz A e sim à identidade acrescentada ao problema. Se não pudermos forçar a saida das variáveis artificiais da base não poderemos eliminá-las do problema e voltar ao contexto do problema original.

Retirando as variáveis artificiais da base

Suponha que a solução ótima do problema auxiliar seja $\begin{bmatrix} x^* \\ y^*=0 \end{bmatrix}$ e que a base ótima obtida é formada pelas variáveis $x_{B_1}, \ldots, x_{B_k}$ e $y_{B_{k+1}}, \ldots, y_{B_m}$, onde $k < m$ (se necessário pode-se reordenar os índices básicos). Temos duas possibilidades:

A matriz A possui linhas L.I. Neste caso as colunas $A^{B_1}, \ldots, A^{B_k}$ podem ser completadas com outras colunas da matriz A até formar uma base; as variáveis associadas às outras colunas possuem valor 0 na solução x^* porque são não-básicas, e entrarão na base também com valor 0;

A matriz A possui linhas L.D. Neste caso não existem bases, mas é possível identificar linhas redundantes (combinações lineares de outras linhas) e reduzir o número de linhas até obter um conjunto linearmente independente.

O procedimento utilizado para obter uma base inicial do simplex após resolver o problema auxiliar consiste em tentar retirar as variáveis artificiais da base. A cada passo ocorrerá OU a substituição de uma variável artificial por uma variável original OU a eliminação de uma variável artificial juntamente com uma linha redundante. Esse procedimento será explicado fazendo-se referência ao tableau do simplex.

Seja $l > k$ e considere a l-ésima variável básica, que é uma variável artificial. Suponha que a l-ésima linha de $B^{-1}A$ possui um elemento diferente de zero, na coluna $B^{-1}A^j$.

		x_{B_i}	$\downarrow$ x_j	y_{B_l}
		0	?	0
		0	?	0
$x_{B_i} =$	?	1	?	0
		0	?	0
$\leftarrow \quad y_{B_l} =$	0	0	$\neq 0$	1
		0	?	0

Então a coluna A^j é não-básica e além disso é linearmente independente de $A^{B_1}, \ldots, A^{B_k}$: basta ver que os vetores $B^{-1}A^{B_i} = e^i$, $i < k < l$ possuem a l-ésima coordenada igual a zero, e portanto $B^{-1}A^j$ não pode ser combinação linear de $B^{-1}A^{B_i}$, $i < k < l$ (equivalentemente A^j não pode ser combinação linear de A^{B_i}, $i < k < l$). Podemos introduzir x_j na base e retirar y_{B_l} fazendo uma pivotação sobre o elemento $B_l^{-1}A^j \neq 0$. Como $y_{B_l} = 0$, a pivotação será degenerada, isso é, não modificará nem a solução x^* nem o valor ótimo do problema auxiliar, e poderá ser efetuada mesmo que o pivô seja negativo.

Suponha por outro lado que a l-ésima linha de $B^{-1}A$ é nula. Então $B_l^{-1}A = \sum_{i=1}^m B_{li}^{-1}A_i = 0'$, o que mostra que as linhas de A são linearmente dependentes. Como $B_l^{-1}b = y_{B_l} = 0$, temos que a l-ésima linha do tableau $B_l^{-1}A = B_l^{-1}b$ é redundante e pode ser eliminada. Se estivermos usando o simplex revisado, podemos concluir que a linha de índice B_l do problema original é redundante: como $B^{-1}I^{B_l} = e^l \Longrightarrow B_l^{-1}I^{B_l} = 1 \Longrightarrow B_{lB_l}^{-1} = 1$, teremos $\sum_{i=1}^{m} B_{li}^{-1}A_i = 0' \Longrightarrow A_{B_l} = \sum_{\substack{i=1 \\ i \neq B_l}}^{m} B_{li}^{-1}A_i$.

Exemplo *Considere o problema*

$$\begin{cases} \min & x_1 - x_2 - x_3 \\ s.a & x_1 + x_2 + x_3 = 2 \\ & x_1 - 2x_2 + x_3 = 2 \\ & 2x_1 + 2x_2 + 2x_3 = 4 \\ & x_1, x_2, x_3 \geq 0 \end{cases}$$

e o problema auxiliar

$$\begin{cases} \min & y_1 + y_2 + y_3 \\ s.a & x_1 + x_2 + x_3 + y_1 = 2 \\ & x_1 - 2x_2 + x_3 + y_2 = 2 \\ & 2x_1 + 2x_2 + 2x_3 + y_3 = 4 \\ & x_1, x_2, x_3, y_1, y_2, y_3 \geq 0 \end{cases}$$

Temos um vértice inicial fazendo $x_1 = x_2 = x_3 = 0$ e $(y_1, y_2, y_3)' = b = (2, 2, 4)'$. A matriz básica correspondente é a identidade e $c_B =$

$(1,1,1)'$. *Os custos reduzidos são* $\bar{c}'_N = c'_N - c'_B B^{-1} A^N = (0,0,0) - (1,1,1)\,[A^1 \mid A^2 \mid A^3] = (-4,-1,-4)$. *A aplicação do simplex tabular para o problema auxiliar usando a regra de Bland gera os seguintes tableaux:*

		$\downarrow$ x_1	x_2	x_3	y_1	y_2	y_3
	-8	-4	-1	-4	0	0	0
$\leftarrow$ $y_1 =$	2	1	1	1	1	0	0
$y_2 =$	2	1	-2	1	0	1	0
$y_3 =$	4	2	2	2	0	0	1

		x_1	x_2	x_3	y_1	y_2	y_3
	0	0	3	0	4	0	0
$x_1 =$	2	1	1	1	1	0	0
$y_2 =$	0	0	-3	0	-1	1	0
$y_3 =$	0	0	0	0	-2	0	1

Este último corresponde a uma solução ótima do problema auxiliar com valor ótimo 0. Assim sabemos que $(x_1, x_2, x_3)' = (2,0,0)'$ *é viável no problema original. Porém a base obtida contém variáveis artificiais.*

Observando a segunda linha do tableau (associada à variável y_2*) nas posições corespondentes às colunas* x_1, x_2, x_3 *encontramos um elemento não-nulo (associado à variável* x_2*). Podemos então introduzir* x_2 *na base, eliminando* y_2 *através de uma pivotação degenerada:*

		x_1	$\downarrow$ x_2	x_3	y_1	y_2	y_3
$x_1 =$	2	1	1	1	1	0	0
$\leftarrow$ $y_2 =$	0	0	-3	0	-1	1	0
$y_3 =$	0	0	0	0	-2	0	1

		x_1	x_2	x_3	y_1	y_2	y_3
$x_1 =$	2	1	0	1	$\frac{2}{3}$	$\frac{1}{3}$	0
$x_2 =$	0	0	1	0	$\frac{1}{3}$	$-\frac{1}{3}$	0
$y_3 =$	0	0	0	0	-2	0	1

Note que a linha zero já não tem importância nesta pivotação: apenas trocamos um índice da base, sem sair do ponto $(2,0,0,0,0,0)'$. *Observando a terceira linha do tableau (associada à variável* y_3*) nas posições corespondentes às colunas* x_1, x_2, x_3 *temos apenas zeros, o que permite concluir que esta linha do tableau é redundante no contexto do problema* *original. Eliminamos a linha e a coluna associadas a y_3 e chegamos ao tableau:*

		x_1	x_2	x_3	y_1	y_2
$x_1 =$	2	1	0	1	$\frac{2}{3}$	$\frac{1}{3}$
$x_2 =$	0	0	1	0	$\frac{1}{3}$	$-\frac{1}{3}$

Agora temos apenas variáveis originais na base, e com isso podemos voltar para o problema original. Note que nas colunas artificiais do tableau a informação contida é $B^{-1}I = B^{-1}$ (a identidade corresponde às colunas associadas às variáveis y_1 e y_2). Para preencher a linha zero do problema original associada a esta base, calculamos

$$-c_B'B^{-1}b = -(1,-1)\begin{bmatrix}2\\0\end{bmatrix} = -2$$

$$\bar{c}_3 = c_3 - c_B'B^{-1}A^3 = -1 - (1,-1)\begin{bmatrix}1\\0\end{bmatrix} = -2$$

e lembramos que $\bar{c}_1 = \bar{c}_2 = 0$ (porque x_1 e x_2 são básicas). Teremos então o tableau a seguir

				$\downarrow$
		x_1	x_2	x_3
	-2	0	0	-2
$\leftarrow$ $x_1 =$	2	1	0	1
$x_2 =$	0	0	1	0

que nos conduz a solução ótima do problema original:

		x_1	x_2	x_3
	2	2	0	0
$x_3 =$	2	1	0	1
$x_2 =$	0	0	1	0

A solução ótima é $(0,0,2)'$ com valor ótimo -2.

O método simplex de duas fases

Resumimos então o comportamento do método simplex completo. Note que não são mais necessários nem o conhecimento de uma base inicial nem a hipótese de que as linhas de A sejam L.I.:

Fase 1: inicialização

1. Se necessário, multiplique as linhas originais por -1 de modo a obter $b \geq 0$;

2. Crie variáveis artificiais $y_1, \dots, y_m$, se necessário, e aplique o método simplex ao problema auxiliar com o objetivo de minimizar $\sum_{i=1}^{m} y_i$;

3. Se o valor ótimo do problema auxiliar é positivo, o problema original é inviável e o método pára;

4. Se o valor ótimo do problema auxiliar é zero e todas as variáveis básicas são originais, elimine as variáveis artificiais e vá para a fase 2;

5. Se a l-ésima variável básica é artificial, observe a linha $B_l^{-1}A$ do tableau. Se esta linha é nula, elimine a l-ésima linha do tableau (no caso do simplex revisado, elimine a B_l-ésima linha da matriz A e a l-ésima linha e B_l-ésima coluna da matriz inversa de base); elimine também a variável artificial y_{B_l}. Se ao contrário existir j tal que $B_l^{-1}A^j \neq 0$, faça uma mudança de base usando $B_l^{-1}A^j$ como pivô, e elimine a variável artificial y_{B_l}. Repita esta operação até que a base possua apenas variáveis originais.

Fase 2: solução

1. Considere a inversa de base ou o tableau do final da fase 1;

2. Calcule os custos reduzidos das variáveis usando esta base inicial e os coeficientes da função objetivo original;

3. Aplique o método simplex para o problema original.

Note que, utilizando-se uma regra anti-ciclagem qualquer, o método simplex de duas fases terminará em tempo finito para qualquer problema de programação linear canônico, cobrindo todas as possibilidades a seguir:

1. Se o problema original é inviável, isso é detectado no passo 3 da fase 1;

2. Se o problema é viável mas a matriz A possui linhas redundantes, isso é detectado e corrigido no final da fase 1;

3. Se o valor ótimo é $-\infty$, isso é detectado durante o passo 3 da fase 2;

4. Se existe uma solução ótima, o simplex produzirá uma base ótima no final da fase 2.

3.5 Eficiência computacional do método simplex

Na seção 3.3 fizemos uma análise preliminar do custo computacional de UMA iteração do método simplex, para cada uma das implementações propostas. Estamos interessados em analisar o número de iterações necessárias para resolver um problema de programação linear.

O número de iterações no pior caso

Embora possamos exprimir a complexidade computacional de uma iteração do simplex como uma função polinomial das dimensões m e n, o número de iterações do simplex no pior caso não admite uma tal limitação a priori. Um exemplo de comportamento não polinomial do simplex pode ser construído como segue.

Teorema 3.5 *Seja $\varepsilon \in (0, \frac{1}{2})$ e considere o problema de programação linear*

$$\begin{cases} \min & -x_n \\ s.a & \varepsilon \leq x_1 \leq 1 \\ & \varepsilon x_{i-1} \leq x_i \leq 1 - \varepsilon x_{i-1}, \quad i = 2, \ldots, n. \end{cases}$$

Então são verdadeiras as seguintes afirmações:

1. *O conjunto viável possui 2^n vértices;*

2. *Os vértices podem ser ordenados numa sequência $v^1, \ldots, v^{2^n}$ de modo que v^{i+1} é adjacente a v^i e $c'v^{i+1} < c'v^i$;*

3. *Existe uma regra de pivotação sob a qual o simplex percorre a sequência de vértices $v^1, \ldots, v^{2^n}$ nesta ordem, levando $2^n - 1$ iterações.*

Este exemplo não responde a seguinte pergunta (em aberto): *é verdade que exemplos como o anterior existem para quaisquer regras de pivotação possíveis?*

O número de iterações no caso médio

Apesar da existência comprovada de exemplos especiais cuja solução exige um número exponencial de iterações do método simplex, a observação do comportamento do método simplex em situações típicas sugere que o número de iterações é em média $\mathcal{O}(m)$. A dificuldade de produzir resultados gerais no caso médio reside na dificuldade de estabelecer uma distribuição de probabilidade razoável sobre o universo de problemas possíveis.

Um exemplo de resultado particular para o caso médio do simplex segue. Suponha que são dados m vetores-linha $A_1, \ldots, A_m \in \mathbb{R}^n$, um vetor $c \in \mathbb{R}^n$ e um vetor $b \in \mathbb{R}^m$ e considere que para cada índice $i = 1, \ldots, m$ podemos escolher uma das desigualdades $A_i x \leq b_i$ ou $A_i x \geq b_i$ com probabilidade igual. Poderão ser gerados deste modo 2^m problemas lineares diferentes, dos quais um certo número L de problemas serão viáveis. Pode-se mostrar então que, usando uma regra especial de pivotação, no caso médio dentre os L problemas viáveis o método simplex terminará em no máximo $\frac{n}{2}$ iterações (vide Schrijver: "Theory of Linear and Integer Programming"; o resultado é de M. Haimovich). Esta estimativa linear está em concordância com uma "intuição empírica" do comportamento do simplex.

Exercícios sugeridos para o capítulo 3: 3.2-3.7, 3.9, 3.12, 3.17-3.20, 3.22-3.25, 3.28 e 3.33.