Método Simplex

Marina Andretta

ICMC-USP

10 de setembro de 2019

Baseado no livro Introduction to Linear Optimization, de D. Bertsimas e J. N. Tsitsiklis.

Método Simplex

Já vimos que se um problema de programação linear na forma padrão tem uma solução ótima, existe uma solução básica viável que é ótima.

O método Simplex é baseado neste fato e busca por uma solução ótima movendo de uma solução básica viável a outra, pelas arestas do conjunto viável, sempre por uma direção que reduz o custo.

Em algum momento, uma solução básica viável é atingida, na qual nenhuma das arestas leva a uma redução do custo. Esta solução básica viável é ótima e o algoritmo termina sua execução.

A partir de agora, vamos sempre considerar que o problema de programação linear está escrito na forma padrão

$$\begin{array}{rl} \text{minimizar} & c^T x \\ \text{sujeita a} & Ax = b, \\ & x \geq 0, \end{array}$$

com P o conjunto viável correspondente. Supomos que a matriz $A \in \mathbb{R}^{m \times n}$ tem linhas linearmente independentes.

Usaremos a notação A_i para a i-ésima coluna de A e a_i^T sua i-ésima linha.

Método Simplex

Muitos algoritmos de otimização são estruturados da seguinte maneira: dada uma solução viável, procura-se em sua vizinhança um ponto viável próximo com custo menor. Se nenhum solução viável próxima leva a uma redução de custo, temos uma solução ótima local.

Para problemas de otimização gerais, uma solução ótima local não necessariamente é uma solução ótima global. Felizmente, no caso de problemas de programação linear, otimalidade local implica em otimalidade global, porque estamos minimizando uma função convexa em um conjunto convexo.

Método Simplex

Vamos agora nos concentrar em como encontrar uma direção na vizinhança de uma solução básica viável que proporcione redução do custo.

Vamos também desenvolver condições de otimalidade para verificar se uma dada solução básica viável é ótima.

Direção viável

Suponha que estamos em um ponto $x \in P$ e queremos sair de x, andando em uma direção $d \in \boldsymbol{R}^n$. Claramente, estamos interessados em direções d que não leve diretamente a um ponto inviável.

Isso leva a seguinte definição.

Definição 1. *Seja x um elemento de um poliedro P. Um vetor $d \in \boldsymbol{R}^n$ é dito uma direção viável a partir de x se existe um escalar positivo θ para o qual $x + \theta d \in P$.*

Direção viável

Sejam x uma solução básica viável do problema na forma padrão, $B(1), ..., B(m)$ os índices das variáveis básicas e $B = (A_{B(1)} ... A_{B(m)})$ a matriz base correspondente.

Em particular, temos que $x_i = 0$ para toda variável não-básica e o vetor das variáveis básicas $x_B = (x_{B(1)}, ..., x_{B(m)})$ é dado por

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

Direção viável

Considere a possibilidade de mover de x para um novo vetor $x + \theta d$ selecionando uma variável não-básica x_j (que começa valendo 0) e a aumentando para um valor positivo θ, mantendo as demais variáveis não-básicas em 0.

Algebricamente, $d_j = 1$ e $d_i = 0$ para todo índice não-básico i diferente de j.

Ao mesmo tempo, o vetor x_B de variáveis básicas muda para $x_B + \theta d_B$, com $d_B = (d_{B(1)}, ..., d_{B(m)})$ o vetor com as componentes de d correspondentes às variáveis básicas.

Direção viável

Dado que estamos interessados somente em soluções viáveis, queremos que $A(x + \theta d) = b$.

Como x é viável, temos $Ax = b$. Assim, para que $A(x + \theta d) = b$ com $\theta > 0$, precisamos que $Ad = 0$.

Direção viável

Lembre-se que $d_j = 1$ e $d_i = 0$ para todo índice não-básico diferente de j.

Então,

$$0 = Ad = \sum_{i=1}^{n} A_i d_i = \sum_{i=1}^{m} A_{B(i)} d_{B(i)} + A_j = B d_B + A_j.$$

Como a matriz base B é inversível, temos

$$d_B = -B^{-1} A_j. \tag{1}$$

Direção viável

A direção d que construímos será chamada de j-ésima direção básica.

Por enquanto garantimos que as restrições de igualdade são satisfeitas quando nos movemos de x na direção d. Vejamos o que acontece com as restrições de não-negatividade.

Lembre-se que a variável x_j é aumentada e todas as outras variáveis não-básicas se mantêm em 0. Portanto, precisamos apenas nos preocupar com o sinal das variáveis básicas.

Direção viável

Vamos analisar dois casos:

(a) Suponha que x seja uma solução básica não-degenerada.

Então $x_B > 0$. Assim, $x_B + \theta d_B \geq 0$ para um valor de θ suficientemente pequeno de θ e a viabilidade é mantida.

Em particular, d é uma direção viável.

Direção viável

(b) Suponha que x seja degenerada.

Então d nem sempre é uma direção viável.

Note que é possível que uma variável básica $x_{B(i)}$ seja 0 e a componente correspondente $d_{B(i)}$ de $d_B = -B^{-1}A_j$ seja negativa.

Neste caso, se seguimos a j-ésima direção básica, a restrição de não-negatividade de $x_{B(i)}$ é imediatamente violada.

Custo reduzido

Vamos ver agora os efeitos no custo de andar em uma direção básica.

Se d é a j-ésima direção básica, então a taxa de mudança do custo $c^T d$ ao longo da direção d é dada por $c_B^T d_B + c_j$, com $c_B = (c_{B(1)}, \ldots, c_{B(m)})$.

Como $d_B = -B^{-1}A_j$, isso é o mesmo que $c_j - c_B^T B^{-1} A_j$. Esta quantidade é importante o suficiente para ter uma definição.

Custo reduzido

Definição 2. *Sejam* x *uma solução básica,* B *uma matriz base associada a* x *e* c_B *o vetor de custos das variáveis básicas. Para cada* j*, definimos o* *custo reduzido* $\bar{c}_j$ *da variável* x_j *pela fórmula*

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

Intuitivamente, c_j é o custo por unidade de crescimento da variável x_j e o termo $-c_B^T B^{-1} A_j$ é o custo das mudanças nas variáveis básicas necessárias para satisfazer as restrições $Ax = b$.

Exemplo

Considere o problema de programação linear

$$
\begin{array}{lccccccccc}
\text{minimizar} & c_1x_1 & + & c_2x_2 & + & c_3x_3 & + & c_4x_4 & & \\
\text{sujeita a} & x_1 & + & x_2 & + & x_3 & + & x_4 & = & 2, \\
 & 2x_1 & & & + & 3x_3 & + & 4x_4 & = & 2, \\
 & & & & x_1, & x_2, & x_3, & x_4 & \geq & 0,
\end{array}
$$

As duas primeiras colunas da matriz A são $A_1 = (1, 2)$ e $A_2 = (1, 0)$. Como elas são linearmente independentes, podemos escolher x_1 e x_2 como variáveis básicas.

Exemplo

A matriz base correspondente é

$$B = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}.$$

Definimos $x_3 = x_4 = 0$ e, resolvendo o sistema $B \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, temos $x_1 = x_2 = 1$.

Assim, $x = (1, 1, 0, 0)$ é uma solução básica viável não-degenerada.

Exemplo

Uma direção básica correspondente a aumentar a variável não-básica x_3 é contruída da seguinte forma: primeiramente, definimos $d_3 = 1$ e $d_4 = 0$. A direção de mudança das variáveis básicas é calculada usando a fórmula (1):

$$\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} d_{B(1)} \\ d_{B(2)} \end{pmatrix} = d_B = -B^{-1}A_3 =$$

$$-\begin{pmatrix} 0 & 0.5 \\ 1 & -0.5 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -1.5 \\ 0.5 \end{pmatrix}.$$

Portanto, a terceira direção básica é dada por $d = (-1.5, 0.5, 1, 0)$.

Exemplo

O custo de mover de x na direção básica d é $c^T d = -1.5c_1 + 0.5c_2 + c_3$.

Este é o custo reduzido da variável x_3.

Condições de otimalidade

Considere agora o custo reduzido (Definição 2) para o caso de uma variável básica.

Como B é a matriz $(A_{B(1)} \dots A_{B(m)})$, temos que $B^{-1}(A_{B(1)} \dots A_{B(m)}) = I$, com I a matriz identidade de dimensão $m \times m$.

Em particular, $B^{-1}A_{B(i)}$ é a i-ésima coluna da matriz identidade, denotada por e_i.

Condições de otimalidade

Portanto, para toda variável básica $x_{B(i)}$, temos

$$\bar{c}_{B(i)} = c_{B(i)} - c_B^T B^{-1} A_{B(i)} = c_{B(i)} - c_B^T e_i = c_{B(i)} - c_{B(i)} = 0.$$

Então, o custo reduzido de toda variável básica é zero.

O teorema a seguir diz respeito a condições de otimalidade de um ponto. Dada nossa interpretação de custos reduzidos como a taxa de mudança do custo ao longo de certa direção, o teorema é intuitivo.

Teorema 1. *Considere uma solução básica viável x associada a uma matriz base B. Seja $\bar{c}$ o vetor correspondente de custos reduzidos.*

(a) *Se $\bar{c} \geq 0$, então x é ótima.*

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

Note que o Teorema 1 admite que x seja uma solução básica viável degenerada, mas $\bar{c}_j < 0$ para algum índice não-básico j.

Existe um teorema análogo ao Teorema 1 que dá condições sob as quais uma solução básica viável é uma solução ótima única.

Condições de otimalidade

De acordo com o Teorema 1, para decidir se uma solução básica viável não-degenerada é ótima, precisamos apenas verificar se todos os custos reduzidos são não-negativos, o que é o mesmo que examinar as $n - m$ direções básicas.

Se x é uma solução básica viável degenerada, não existe um teste computacionalmente simples como este para ser feito. Felizmente, o método Simplex, que desenvolveremos a seguir, consegue lidar com isso de uma maneira eficiente.

Note que para usar o Teorema 1 e verificar se uma dada solução básica é ótima, precisamos satisfazer duas condições: viabilidade e não-negatividade dos custos reduzidos (otimalidade). Isso nos leva à seguinte definição:

Definição 3. *A matriz base B é chamada de ótima se:*

(a) $B^{-1}b \geq 0$ *e*

(b) $\bar{c}^T = c^T - c_B^T B^{-1} A \geq 0^T$.

Condições de otimalidade

Claramente, se uma base ótima é encontrada, a solução básica correspondente é viável e satisfaz as condições de otimalidade. Portanto, é ótima.

No entanto, no caso degenerado, ter uma solução básica viável ótima não quer necessariamente dizer que os custos sejam não-negativos.

Método Simplex

Agora estamos prontos para desenvolver o Método Simplex, detalhando como passar de uma solução básica viável a outra melhor, sempre que uma direção básica de descida é descoberta.

O Método Simplex foi proposto por George B. Dantzig (1914–2005) em 1947 e foi o primeiro método prático para resolver problemas de programação linear. Nesta época, computadores estavam começando a surgir e a resolução deste tipo de problemas se tornava importante na prática, devido a aplicações militares, econômicas e em transportes.

Vamos supor que toda solução básica viável é não-degenerada. Essa suposição será relaxada (e explicitada) mais à frente.

Suponha que estamos em uma solução básica viável x e que calculamos os custos reduzidos $\bar{c}_j$ das variáveis não-básicas.

Se todos eles são não-negativos, o Teorema 1 da aula sobre "condições de otimalidade" garante que temos uma solução ótima e podemos parar.

Por outro lado, se o custo reduzido $\bar{c}_j$ de alguma variável não-básica é negativo, a j-ésima direção básica d é uma direção viável na qual o custo decresce (ou direção de descida).

Método Simplex

Ao mover ao longo desta direção d, a variável não-básica j se torna positiva e todas as outras variáveis não-básicas permanecem em 0.

Neste caso, dizemos que x_j (ou A_j) entra na base.

Quando começamos a mover a partir de x ao longo da direção d, estamos em pontos da forma $x + \theta d$, com $\theta \geq 0$.

Como o custo decresce ao longo da direção d, queremos andar o máximo possível nesta direção.

Método Simplex

Isso nos leva ao ponto $x + \theta^* d$, com

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

A mudança no custo resultante é $\theta^* c^T d$, que é o mesmo que $\theta^* \bar{c}_j$.

Método Simplex

Vamos agora desenvolver uma fórmula para calcular θ^*.

Dado que $Ad = 0$, temos que $A(x + \theta d) = Ax + \theta Ad = b$ para todo θ. Ou seja, as restrições de igualdade nunca são violadas.

Então, $x + \theta d$ pode se tornar inviável somente se alguma de suas componentes se tornar negativa.

Vamos analisar dois casos:

(a) Se $d \geq 0$, então $x + \theta d \geq 0$ para todo $\theta \geq 0$.

Como o vetor $x + \theta d$ nunca se torna inviável, definimos $\theta^* = \infty$.

(b) Se $d_i < 0$ para algum i, a restrição $x_i + \theta d_i \geq 0$ se torna $\theta \leq -\frac{x_i}{d_i}$.

Esta restrição em θ deve ser satisfeita para todo i tal que $d_i < 0$. Portanto, o maior valor possível para θ é

$$\theta^* = \min_{\{i \mid d_i < 0\}} \left(-\frac{x_i}{d_i} \right).$$

(b) (cont.) Lembre-se que se x_i é uma variável não-básica, então ou x_i é uma variável que vai entrar na base (neste caso, $d_i = 1$), ou $d_i = 0$.

Em ambos os casos, d_i é não-negativo.

Então, precisamos considerar apenas as variáveis básicas e temos a fórmula equivalente

$$\theta^* = \min_{\{i=1,\ldots,m \mid d_{B(i)}<0\}} \left(-\frac{x_{B(i)}}{d_{B(i)}} \right). \tag{2}$$

Note que $\theta^* > 0$, já que, por causa da não-degenerescência, $x_{B(i)} > 0$ para todo i.

Exemplo

Considere o problema de programação linear

$$\begin{array}{lccccccccc} \text{minimizar} & c_1x_1 & + & c_2x_2 & + & c_3x_3 & + & c_4x_4 & & \\ \text{sujeita a} & x_1 & + & x_2 & + & x_3 & + & x_4 & = & 2, \\ & 2x_1 & & & + & 3x_3 & + & 4x_4 & = & 2, \\ & & & & x_1, & x_2, & x_3, & x_4 & \geq & 0, \end{array}$$

já usado no exemplo da aula sobre "condições de otimalidade".

Como já vimos, $x = (1, 1, 0, 0)$ é uma solução básica viável não-degenerada e a terceira direção básica é dada por $d = (-1.5, 0.5, 1, 0)$.

Exemplo

O custo de mover de x na direção básica d é $c^T d = -1.5c_1 + 0.5c_2 + c_3$, que é o custo reduzido $\bar{c}_3$ da variável x_3.

Suponha que $c = (2, 0, 0, 0)$. Neste caso, temos $\bar{c}_3 = -3$.

Como $\bar{c}_3$ é negativo, consideramos vetores da forma $x + \theta d$, com $\theta \geq 0$.

Exemplo

Conforme θ cresce, a única componente de $x + \theta d$ que decresce é a primeira, já que $d_1 < 0$.

O maior valor possível para θ é dado por $\theta^* = -\frac{x_1}{d_1} = \frac{2}{3}$.

Isso nos leva ao ponto $y = x + \frac{2}{3}d = \left(0, \frac{4}{3}, \frac{2}{3}, 0\right)$.

Exemplo

Note que as colunas A_2 e A_3 correspondentes às variáveis não-nulas de y são $(1, 0)$ e $(1, 3)$, respectivamente.

Como são linearmente independentes, elas formam uma base e o vetor y é uma nova solução básica viável.

Em particular, a variável x_3 entrou na base e a variável x_1 saiu.

Método Simplex

Uma vez que θ^* é escolhido, supondo que ele seja finito, movemos para a nova solução viável $y = x + \theta^* d$.

Como $x_j = 0$ e $d_j = 1$, temos que $y_j = \theta^* > 0$.

Seja ℓ um índice que minimiza (2), ou seja

$$-\frac{x_{B(\ell)}}{d_{B(\ell)}} = \min_{\{i=1,\dots,m \mid d_{B(i)}<0\}} \left(-\frac{x_{B(i)}}{d_{B(i)}}\right) = \theta^*.$$

Note que $d_{B(\ell)} < 0$ e $x_{B(\ell)} + \theta^* d_{B(\ell)} = 0$.

Assim, a variável básica $x_{B(\ell)}$ se torna 0, enquanto a variável não-básica x_j se torna positiva, o que sugere que x_j deve substituir $x_{B(\ell)}$ na base.

Então, subtituímos na base antiga B a coluna $A_{B(\ell)}$ pela coluna A_j e obtemos a matriz

$$\bar{B} = \begin{pmatrix} | & & | & | & | & & | \\ A_{B(1)} & \cdots & A_{B(\ell-1)} & A_j & A_{B(\ell+1)} & \cdots & A_{B(m)} \\ | & & | & | & | & & | \end{pmatrix}.$$

Equivalentemente, estamos trocando o conjunto de índices de variáveis básicas $\{B(1), ..., B(m)\}$ pelo novo conjunto $\{\bar{B}(1), ..., \bar{B}(m)\}$, cujos índices são dados por

$$\bar{B}(i) = \begin{cases} B(i), & i \neq \ell, \\ j, & i = \ell. \end{cases}$$

Teorema 2.

(a) *As colunas* $A_{B(i)}$, $i \neq \ell$, *e* A_j *são linearmente independentes e, portanto,* $\bar{B}$ *é uma base.*

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

Método Simplex

Como θ^* é positivo, a nova solução básica viável $x + \theta^* d$ é diferente de x.

Como d é um direção de descida, o custo da nova solução básica viável é estritamente menor.

Atingimos, então, o objetivo de mover para uma nova solução básica viável com custo menor.

Método Simplex

Podemos agora definir uma iteração típica do Método Simplex, também conhecida como um pivô.

Para facilitar a notação, vamos definir um vetor $u = (u_1, ..., u_m)$ como

$$u = -d_B = B^{-1}A_j,$$

com A_j a coluna que entra na base. Em particular, $u_i = -d_{B(i)}$, para $i = 1, ..., m$.

Uma iteração do Método Simplex

P1. Em uma iteração típica, começamos com uma base formada pelas colunas básicas $A_{B(1)}, ..., A_{B(m)}$ e uma solução básica viável x.

P2. Calcule os custos reduzidos $\bar{c}_j = c_j - c_B^T B^{-1} A_j$ para todos os índices não-básicos j.

Se todos os custos reduzidos forem não-negativos, a solução básica viável correspondente é ótima e o algoritmo pára.

Caso contrário, escolha um j para o qual $\bar{c}_j < 0$.

P3. Calcule $u = B^{-1}A_j$.

Se nenhuma componente de u for positiva, temos $\theta^* = \infty$, custo ótimo é $-\infty$ e o algoritmo pára.

P4. Se alguma componente de u é positiva, calcule

$$\theta^* = \min_{\{i=1,\dots,m \mid u_i>0\}} \left(\frac{x_{B(i)}}{u_i} \right).$$

P5. Seja ℓ um índice tal que $\theta^* = \frac{x_{B(\ell)}}{u_\ell}$.

Monte uma nova base, substituindo $A_{B(\ell)}$ por A_j.

Se y é a nova solução básica viável, os valores das novas variáveis básicas são $y_j = \theta^*$ e $y_{B(i)} = x_{B(i)} - \theta^* u_i$, para todo $i \neq \ell$.

O Método Simplex começa com uma solução básica viável arbitrária, o que, para problemas de programação linear na forma padrão viáveis, sempre existe.

O Teorema 3 garante que, no caso não-degenerado, o Método Simplex funciona corretamente e termina em um número finito de iterações.

Teorema 3. *Suponha que o conjunto viável seja não-vazio e que toda solução básica viável seja não-degenerada. Então, o Método Simplex termina em um número finito de iterações. Ao terminar, há duas possibilidades:*

(a) *Temos uma base ótima B e uma solução básica viável associada que é ótima.*

(b) *Encontramos um vetor d que satisfaz $Ad = 0$, $d \geq 0$, $c^T d < 0$ e o custo ótimo é $-\infty$.*

Até este momento estávamos supondo que todas as soluções básicas são não-degeneradas.

Vamos ver o que acontece se o mesmo algoritmo é usado na presença de degenerescência.

Método Simplex para problemas degenerados

Neste caso, as seguintes novas possibilidades podem ser encontradas na execução do algoritmo:

(a) Se a solução básica viável atual x é degenerada, θ^* pode ser igual a 0, o que faz com que a nova solução básica viável y seja igual a x.

Isso acontece se alguma variável $x_{B(\ell)}$ é igual a 0 e a componente corresponde na direção d, $d_{B(\ell)}$, é negativa.

Mesmo assim, podemos definir uma nova base $\bar{B}$, substituindo $A_{B(\ell)}$ por A_j e o Teorema 2 ainda é válido.

Método Simplex para problemas degenerados

(b) Mesmo se θ^* é positivo, pode acontecer de mais de uma das variáveis básicas originais se tornarem 0 em $x + \theta^* d$.

Como somente uma delas deixa a base, as outras permanecem valendo 0 e a nova solução básica viável é degenerada.

Método Simplex para problemas degenerados

Mudanças de base mesmo permanecendo na mesma solução básica viável não é inútil, já que uma sequência de mudanças de tais bases pode levar à descoberta de uma direção viável com redução de custo.

Por outro lado, uma sequência de mudanças de base pode fazer com que o algoritmo volte à base inicial e permaneça em um laço infinito. Este fenômeno é chamado de ciclagem.

Método Simplex para problemas degenerados

Pode-se pensar que ciclagem é um evento raro, mas em vários casos de problemas de programação linear que tem uma estrutura muito bem definida, a maioria das soluções básicas viáveis é degenerada e a ciclagem é uma possibilidade real.

Felizmente, a ciclagem pode ser evitada escolhendo com rigor as variáveis que entram e saem da base.

Escolha do pivô

O Método Simplex, como apresentado aqui, possui alguns graus de liberdade.

No Passo P2, podemos escolher qualquer j com custo reduzido $\bar{c}_j$ negativo.

No Passo P5, pode haver diversos índices ℓ que dão o mínimo usado para definir θ^* e podemos escolher qualquer um deles.

Regras para fazer essas escolhas são chamadas de regras de pivotamento.

Escolha do pivô

Para escolher a coluna que entra na base, as regras a seguir são candidatas naturais:

(a) Escolha a coluna A_j, com $\bar{c}_j < 0$, cujo custo é o mais negativo.

Como o custo reduzido é a taxa de mudança da função custo, esta regra escolhe a direção ao longo da qual o custo tem redução mais rápida.

No entanto, a redução de fato do custo depende de quanto vai ser andado nesta direção.

Por isso, temos a próxima regra.

Escolha do pivô

(b) Escolha a coluna A_j, com $\bar{c}_j < 0$, para a qual o decréscimo do custo $\theta^*|\bar{c}_j|$ é o maior.

Esta regra oferece a possibilidade de atingir otimalidade depois de um número menor de iterações.

Por outro lado, o custo computacional em cada iteração é maior, já que é necessário calcular θ^* para cada coluna com $\bar{c}_j < 0$.

Evidências empíricas sugerem que o tempo do algoritmo no geral, usando esta regra, não melhora.

Escolha do pivô

Para problemas de grande porte, mesmo a regra que escolhe o $\bar{c}_j$ mais negativo pode ser computacionalmente custosa, porque requer o cálculo do custo reduzido para toda variável.

Na prática, regras mais simples às vezes são usadas, como a regra do menor índice, que escolhe o menor j tal que $\bar{c}_j < 0$.

Usando esta regra, ao descobrir um custo reduzido negativo não é necessário calcular os demais.

Outras regras já foram testadas, cada uma com algumas vantagens e desvantagens.

Escolha do pivô

Já para escolher a coluna que sai da base, a opção mais simples é novamente a regra do menor índice: de todas as variáveis que podem sair da base, escolha a que tem menor índice.

Uma constatação importante é que usando esta regra tanto para a entrada como para a saída da base, a ciclagem pode ser evitada.

Implementações do Método Simplex

Vamos agora discutir diferentes implementações possíveis para o Método Simplex.

Note que os vetores $B^{-1}A_j$ são essenciais, já que são usados para calcular os custos reduzidos, as direções de descida e o tamanho de passo θ^*.

A maior diferença entre implementações do Método Simplex está em como são calculados $B^{-1}A_j$ e como as informações são aproveitadas de uma iteração para outra.

Implementações do Método Simplex

Para comparar diferentes implementações, precisamos das seguintes informações: se são dados uma matriz $B \in \boldsymbol{R}^{m \times m}$ e vetores $b, p \in \boldsymbol{R}^m$,

- calcular a inversa de B ou resolver um sistema linear na forma $Bx = b$ custa $O(m^3)$ operações aritméticas;
- calcular o produto Bb gasta $O(m^2)$ operações aritméticas;
- calcular o produto interno $p^T b$ gasta $O(m)$ operações aritméticas.

Implementação ingênua

Primeiramente vamos discutir a implementação mais direta, na qual nenhuma informação auxiliar é passada de uma iteração a outra.

No início de uma iteração qualquer, temos os índices $B(1), ..., B(m)$ das variáveis básicas.

Montamos a matriz B e calculamos $p^T = c_B^T B^{-1}$ resolvendo o sistema linear $p^T B = c_B^T$, com p desconhecido.

O vetor p é chamado de vetor de multiplicadores do simplex associados à base B.

Implementação ingênua

O custo reduzido $\bar{c}_j = c_j - c_B^T B^{-1} A_j$ para qualquer variável x_j é calculado usando a fórmula $\bar{c}_j = c_j - p^T A_j$.

Dependendo da regra de pivotamento usada, precisamos calcular todos os custos reduzidos ou apenas um por vez, até encontrar algum com valor negativo.

Implementação ingênua

Quando a coluna A_j é selecionada para entrar na base, resolvemos o sistema linear $Bu = A_j$ para determinar o vetor $u = B^{-1}A_j$.

Neste ponto, podemos calcular a direção pela qual iremos sair da solução básica viável atual.

Finalmente, calculamos θ^*, a variável que sairá da base e calculamos a nova solução básica viável.

Implementação ingênua

Note que precisamos de $O(m^3)$ operações aritméticas para resolver cada um dos sistemas $p^T B = c_B^T$ e $Bu = A_j$.

Além disso, calcular os custos reduzidos de todas as variáveis custa $O(mn)$ operações, porque precisamos calcular o produto interno do vetor p com cada coluna não-básica A_j.

Implementação ingênua

Portanto, o custo computacional total por iteração é $O(m^3 + mn)$. Veremos a seguir que podemos ter implementações alternativas com custo de $O(m^2 + mn)$ operações por iteração.

Isso mostra que esta implementação ingênua é bastante ineficiente em geral.

Por outro lado, para certos problemas com estrutura especial os sistemas lineares $p^T B = c_B^T$ e $Bu = A_j$ podem ser resolvidos muito eficientemente, e, para estes problemas, esta implementação tem interesse prático.

Método Simplex revisado

O grande esforço computacional da implementação ingênua do Método Simplex está na necessidade de resolver dois sistemas lineares.

Para uma implementação alternativa, a matriz B^{-1} pode ser disponibilizada no início de cada iteração e os vetores $c_B^T B^{-1}$ e $B^{-1} A_j$ são calculados usando apenas produtos matriz-vetor.

Para que esta abordagem seja prática, precisamos de uma maneira eficiente de atualizar a matriz B^{-1} toda vez que uma mudança é feita na base.

Sejam

$$B = \begin{pmatrix} A_{B(1)} & \dots & A_{B(m)} \end{pmatrix}$$

a matriz base do início de uma iteração e

$$\bar{B} = \begin{pmatrix} A_{B(1)} & \dots & A_{B(\ell-1)} & A_j & A_{B(\ell+1)} & \dots & A_{B(m)} \end{pmatrix}$$

a matriz base do início da próxima iteração.

Ambas as matrizes têm as mesmas colunas, com exceção da ℓ-ésima, já que a coluna $A_{B(\ell)}$ é substituída por A_j.

Assim, podemos explorar informações de B^{-1} para calular $\bar{B}^{-1}$.

Definição 4. *Dada uma matriz, não necessariamente quadrada, a operação de adicionar um múltiplo constante de uma linha à mesma linha ou a outra é chamada de operação elementar de linha.*

Operações elementares em linhas em uma matriz C podem ser escritas, de forma equivalente, multiplicando uma matriz adequada Q por C.

Exemplo

Considere as matrizes

$$Q = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ e } C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}.$$

Calculando QC, temos

$$QC = \begin{pmatrix} 11 & 14 \\ 3 & 4 \\ 5 & 6 \end{pmatrix},$$

que é o mesmo que somar à primeira linha de C sua terceira linha multiplicada por 2.

Método Simplex revisado

Generalizando o exemplo anterior, podemos ver que multiplicar a j-ésima linha por β e somá-la à i-ésima linha de uma matriz é o mesmo que multiplicar esta matriz, pela esquerda, pela matriz $Q = I + D_{ij}$, com D_{ij} uma matriz com todas as componentes nulas, a menos do elemento (i, j), que é igual a β.

Como o determinante desta matriz Q é 1, ela é inversível.

Método Simplex revisado

Suponha agora que aplicamos uma sequência de K operações elementares em linhas e que a k-ésima operação desta sequência corresponde à multiplicação pela esquerda por uma certa matriz inversível Q_k.

Então, a sequência destas operações elementares é o mesmo que uma multiplicação pela esquerda da matriz inversível $Q_K Q_{K-1} \ldots Q_2 Q_1$.

Portanto, realizar uma sequência de operações elementares em linhas em uma matriz é equivalente a multiplicar esta matriz pela esquerda por uma certa matriz inversível.

Como $B^{-1}B = I$, temos que $B^{-1}A_{B(i)}$ é a i-ésima coluna da identidade e_i.

Usando isto, temos que

$$B^{-1}\bar{B} = \left(\begin{array}{ccccccc} | & & | & | & | & & | \\ e_1 & \dots & e_{\ell-1} & u & e_{\ell+1} & \dots & e_m \\ | & & | & | & | & & | \end{array} \right) =$$

$$\left(\begin{array}{ccccc} 1 & & u_1 & & \\ & \ddots & \vdots & & \\ & & u_\ell & & \\ & & \vdots & \ddots & \\ & & u_m & & 1 \end{array} \right),$$

com $u = B^{-1}A_j$.

Vamos então aplicar uma sequência de operações elementares em linhas para transformar a matriz $B^{-1}\bar{B}$ na matriz identidade.

Considere a seguinte sequência:

(a) Para cada $i \neq \ell$, some à linha i a ℓ-ésima linha multiplicada por $-\frac{u_i}{u_\ell}$.

Isso sempre pode ser feito, já que $u_\ell > 0$. Com isso, o elemento u_i será trocado pelo elemento 0.

(b) Divida a ℓ-ésima linha por u_ℓ.

Com isso, o elemento u_i será trocado pelo elemento 1.

Método Simplex revisado

O que esta sequência de operações elementares está fazendo é somar a cada linha um múltiplo da ℓ-ésima linha para substituir a coluna u pela ℓ-ésima coluna da identidade e_ℓ.

Esta sequência de operações elementares é equivalente a multiplicar pela esqueda a matriz $B^{-1}\bar{B}$ por uma certa matriz inversível Q.

Como o resultado desta multiplicação é a matriz identidade, temos que $QB^{-1}\bar{B} = I$. Ou seja, $QB^{-1} = \bar{B}^{-1}$.

Portanto, se aplicarmos a mesma sequência de operações elementares na matriz B^{-1}, obtemos $\bar{B}^{-1}$.

Exemplo

Sejam

$$B = \begin{pmatrix} 1/7 & 5/21 & 1/3 \\ 0 & 2/3 & 1/3 \\ 2/7 & -11/21 & -1/3 \end{pmatrix}, \quad A_j = \begin{pmatrix} 4/7 \\ 2 \\ -20/7 \end{pmatrix} \text{ e } \ell = 3.$$

Queremos construir $\bar{B}$ substituindo a terceira coluna de B por A_j e, então, calcular $\bar{B}^{-1}$.

Exemplo

Fazendo as contas explicitamente, temos que

$$\bar{B} = \begin{pmatrix} 1/7 & 5/21 & 4/7 \\ 0 & 2/3 & 2 \\ 2/7 & -11/21 & -20/7 \end{pmatrix} \text{ e } \bar{B}^{-1} = \begin{pmatrix} 9 & -4 & -1 \\ -6 & 6 & 3 \\ 2 & -\frac{3}{2} & -1 \end{pmatrix}.$$

Vamos agora usar operações elementares em linhas para obter $\bar{B}^{-1}$.

Exemplo

Temos que

$$B^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 3 & 1 \\ 4 & -3 & -2 \end{pmatrix} \text{ e } u = B^{-1}A_j = \begin{pmatrix} -4 \\ 2 \\ 2 \end{pmatrix}.$$

Nosso objetivo é transformar o vetor u em $e_3 = (0, 0, 1)$ somando múltiplos da linha 3 às demais linhas.

Para isso, multiplicamos a terceira linha por 2 e somamos à primeira. Depois, substraímos a terceira linha da segunda. Finalmente, dividimos a terceira linha por 2.

Exemplo

Executando estas mesmas operações em B^{-1}, temos

$$B^{-1} = \left(\begin{array}{rrr} 1 & 2 & 3 \\ -2 & 3 & 1 \\ 4 & -3 & -2 \end{array}\right) \xrightarrow{l_1+2l_3} \left(\begin{array}{rrr} 9 & -4 & -1 \\ -2 & 3 & 1 \\ 4 & -3 & -2 \end{array}\right) \xrightarrow{l_2-l_3}$$

$$\left(\begin{array}{rrr} 9 & -4 & -1 \\ -6 & 6 & 3 \\ 4 & -3 & -2 \end{array}\right) \xrightarrow{l_3/2} \left(\begin{array}{rrr} 9 & -4 & -1 \\ -6 & 6 & 3 \\ 2 & -\frac{3}{2} & -1 \end{array}\right) = \bar{B}^{-1}.$$

Uma iteração do Método Simplex revisado

Implementando a atualização de B^{-1} como descrito, temos o Método Simplex revisado. Descrevemos aqui uma iteração deste método.

P1. Em uma iteração típica, começamos com uma base formada pelas colunas básicas $A_{B(1)}, ..., A_{B(m)}$, uma solução básica viável x e a inversa B^{-1} da matriz base B.

P2. Calcule o vetor linha $p^T = c_B^T B^{-1}$ e, então, os custos reduzidos $\bar{c}_j = c_j - p^T A_j$ para todos os índices não-básicos j.

Se todos os custos reduzidos forem não-negativos, a solução básica viável correspondente é ótima e o algoritmo pára.

Caso contrário, escolha um j para o qual $\bar{c}_j < 0$.

P3. Calcule $u = B^{-1}A_j$.

Se nenhuma componente de u for positiva, temos $\theta^* = \infty$, custo ótimo é $-\infty$ e o algoritmo pára.

P4. Se alguma componente de u é positiva, calcule

$$\theta^* = \min_{\{i=1,\ldots,m \mid u_i>0\}} \left(\frac{x_{B(i)}}{u_i} \right).$$

Uma iteração do Método Simplex revisado

P5. Seja ℓ um índice tal que $\theta^* = \frac{x_{B(\ell)}}{u_\ell}$.

Monte uma nova base, substituindo $A_{B(\ell)}$ por A_j.

Se y é a nova solução básica viável, os valores das novas variáveis básicas são $y_j = \theta^*$ e $y_{B(i)} = x_{B(i)} - \theta^* u_i$, para todo $i \neq \ell$.

Uma iteração do Método Simplex revisado

P6. Monte a matriz $(B^{-1} \mid u) \in \boldsymbol{R}^{m \times (m+1)}$.

Some a cada uma de suas linhas um múltiplo da ℓ-ésima linha de forma a transformar a última coluna na ℓ-ésima coluna da identidade e_ℓ.

As primeiras m colunas desta matriz é $\bar{B}^{-1}$.