Condições de otimalidade

Marina Andretta

ICMC-USP

19 de outubro de 2016

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.

Método Simplex

Nas próximas aulas vamos ver com detalhes o desenvolvimento do Simplex e algumas de suas diferentes implementações.

Também discutiremos algumas dificuldades que surgem na presença de degenerescência.

Método Simplex

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 $A \in \boldsymbol{R}^{m \times n}$ tem linhas linearmente independentes.

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

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

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

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 $(A_{B(1)} \ldots A_{B(m)})$, temos que $B^{-1}(A_{B(1)} \ldots A_{B(m)}) = I$, 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.

Condições de otimalidade

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* $\bar{c} \geq 0$*.*

Prova do Teorema 1

Prova: Primeiramente, vamos mostrar (a), ou seja, se $\bar{c} \geq 0$, então x é ótima.

Suponha que $\bar{c} \geq 0$ e seja y uma solução viável arbitrária. Defina $d = y - x$. Como x e y são viáveis, temos que $Ax = Ay = b$ e, portanto, $Ad = 0$.

Podemos reescrever a última igualdade como

$$Bd_B + \sum_{i \in N} A_i d_i = 0,$$

com N o conjunto de índices das variáveis não-básicas (definidas pela base B).

Prova do Teorema 1

Como B é inversível, temos

$$d_B + B^{-1}\sum_{i \in N} A_i d_i = 0 \Rightarrow d_B = -B^{-1}\sum_{i \in N} A_i d_i$$

e

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

Para todo índice não-básico $i \in N$, devemos ter $x_i = 0$. Como y é viável, $y_i \geq 0$. Então, $d_i = y_i - x_i \geq 0$ e $\bar{c}_i d_i \geq 0$ para todo $i \in N$.

Portanto, $c^T(y - x) = c^T d \geq 0 \Rightarrow c^T y \geq c^T x$. Como y é uma solução viável arbitrária, x é ótimo.

Prova do Teorema 1

Vamos agora mostrar (b), ou seja, se x é ótima e não-degenerada, então $\bar{c} \geq 0$.

Suponha que x seja uma solução básica viável não-degenerada ótima. Suponha, por absurdo, que $\bar{c}_j < 0$ para algum j. Como o custo reduzido de variáveis básicas é sempre zero, x_j deve ser uma variável não-básica e $\bar{c}_j$ é a taxa de mudança do custo ao longo da j-ésima direção básica.

Como x é não-degenerada, a j-ésima direção básica é uma direção viável que decresce o custo. Ao mover nesta direção, obtemos soluções viáveis com custos menores do que x. Portanto, x não é ótima, o que contraria nossa hipótese. $\square$

Condições de otimalidade

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.