Método Simplex

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

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 prátoco 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.

Método Simplex

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

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

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 $y_j = \theta^* > 0$.

Método Simplex

Seja ℓ um índice que minimiza (1), ou seja

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

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 1.

(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* $\bar{B}$.

Prova do Teorema 1

Prova: Primeiramente, vamos mostrar (a), ou seja, que as colunas $A_{B(i)}$, $i \neq \ell$, e A_j são linearmente independentes.

Suponha, por absurdo, que os vetores $A_{\bar{B}(i)}$, $i = 1, ..., m$, com $\bar{B}(i)$ definido como anteriormente, sejam linearmente dependentes. Neste caso, existem coeficientes $\lambda_1, ..., \lambda_m$, não todos nulos, tais que

$$\sum_{i=1}^{m} \lambda_i A_{\bar{B}(i)} = 0.$$

Prova do Teorema 1

Ou seja,

$$B^{-1}\sum_{i=1}^{m}\lambda_i A_{\bar{B}(i)} = \sum_{i=1}^{m}\lambda_i B^{-1}A_{\bar{B}(i)} = 0,$$

o que mostra que os vetores $B^{-1}A_{\bar{B}(i)}$ também são linearmente dependentes.

Como $B^{-1}B = I$ e $A_{B(i)}$ é a i-ésima coluna de B, segue que os vetores $B^{-1}A_{\bar{B}(i)} = B^{-1}A_{B(i)}$, para todo $i \neq \ell$, são todas as colunas da matriz identidade, com exceção da coluna ℓ. Em particular, eles vetores são linearmente independentes e suas ℓ-ésimas componentes são todas nulas.

Prova do Teorema 1

Por outro lado, $B^{-1}A_{\bar{B}(\ell)} = B^{-1}A_j = -d_B$. Sua ℓ-ésima componente, $-d_{B(\ell)}$, é não-nula pela definição de ℓ.

Portanto, $B^{-1}A_{\bar{B}(\ell)}$ é linearmente independente de todos os outros vetores $B^{-1}A_{\bar{B}(i)}$, o que contraria o fato destes vetores serem linearmente dependentes que concluímos a partir da suposição que os vetores $A_{\bar{B}(i)}$ são linearmente dependentes.

Prova do Teorema 1

Vamos agora mostrar (b), ou seja, o vetor $y = x + \theta^* d$ é uma solução básica viável associada à matriz base $\bar{B}$.

Note que, por construção, $y \geq 0$, $Ay = b$ e $y_i = 0$ para todo $i \neq \bar{B}(1), ..., \bar{B}(m)$. Além disso, como mostrado no item (a) deste teorema, as colunas $A_{\bar{B}(1)}, ..., A_{\bar{B}(m)}$ são linearmente independentes.

Portanto, y é uma solução básica viável associada à matriz $\bar{B}$. $\square$

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

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

Uma iteração do Método Simplex

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

Teorema 2. *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 $-\infty$.*

Prova do Teorema 2

Prova: Se o algoritmo pára pelo critério de parada do Passo P2, então as condições de otimalidade do Teorema 1 da aula sobre "condições de otimalidade" são satisfeitas, B é uma base ótima e a solução básica viável atual é ótima.

Se o algoritmo pára pelo critério de parada do Passo P3, temos uma solução básica viável x e encontramos uma variável não-básica x_j tal que $\bar{c}_j < 0$. A direção d correspondente satisfaz $Ad = 0$ e $d \geq 0$.

Note que, neste caso, $x + \theta d \in P$ para todo $\theta > 0$. Como $c^T y = \bar{c}_j < 0$, quanto maior o valor de θ, menor será o custo. Ou seja, o custo ótimo é $-\infty$.

Prova do Teorema 2

Em cada iteração, o algoritmo move de um ponto por uma quantidade positiva θ^* em uma direção d que satisfaz $c^T d < 0$. Portanto, o custo de uma solução básica viável calculada em uma iteração é sempre estritamente menor do que o custo da solução básica viável da iteração anterior. Desta forma, nenhuma solução básica viável é visitada mais de uma vez.

Como o número de soluções básicas viáveis é finito, o algoritmo deve terminar sua execução em um número finito de passos. $\square$

Método Simplex para problemas degenerados

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 1 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.

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.