Método Simplex revisado

Marina Andretta

ICMC-USP

26 de setembro de 2016

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

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.

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

- 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.

Método Simplex revisado

Sejam

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

a matriz base do início de uma iteração e

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

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 1. *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.

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.

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

$$B^{-1}\bar{B} = \begin{pmatrix} | & & | & | & | & & | \\ e_1 & \dots & e_{\ell-1} & u & e_{\ell+1} & \dots & e_m \\ | & & | & | & | & & | \end{pmatrix} =$$

$$\begin{pmatrix} 1 & & u_1 & & \\ & \ddots & \vdots & & \\ & & u_\ell & & \\ & & \vdots & \ddots & \\ & & u_m & & 1 \end{pmatrix},$$

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

Método Simplex revisado

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.

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

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

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

$$B^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 3 & 1 \\ 4 & -3 & -2 \end{pmatrix} \xrightarrow{l_1+2l_3} \begin{pmatrix} 9 & -4 & -1 \\ -2 & 3 & 1 \\ 4 & -3 & -2 \end{pmatrix} \xrightarrow{l_2-l_3}$$

$$\begin{pmatrix} 9 & -4 & -1 \\ -6 & 6 & 3 \\ 4 & -3 & -2 \end{pmatrix} \xrightarrow{l_3/2} \begin{pmatrix} 9 & -4 & -1 \\ -6 & 6 & 3 \\ 2 & -\frac{3}{2} & -1 \end{pmatrix} = \bar{B}^{-1}.$$

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 $A_{B(1)}, \ldots, A_{B(m)}$, 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

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

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 $i \neq \ell$.

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 $\bar{B}^{-1}$.