L'algoritmo del simplesso rivisto

Ricerca operativa

Giovanni Righini

Università degli Studi di Milano

L’algoritmo *revised simplex*

Per eseguire i test di ammissibilità e ottimalità e scegliere il pivot su cui eseguire la prossima iterazione non è necessario conoscere tutti i coefficienti del tableau.

L’idea quindi è di rappresentare il tableau in un modo alternativo, ma equivalente, risparmiando alcune operazioni.

A questo scopo si sfrutta la dualità.

Coppia duale

Consideriamo un problema di PL in forma standard

$$\begin{aligned} \text{P) minimize } z &= c^T x \\ \text{subject to } Ax &= b \\ x &\geq 0 \end{aligned}$$

con $n + m$ variabili non-negative e m vincoli di uguaglianza.

Il suo duale è

$$\begin{aligned} \text{D) maximize } w &= b^T y \\ \text{subject to } A^T y &\leq c \end{aligned}$$

con m variabili libere e $m + n$ vincoli di disuguaglianza.

Base primale

Scelta una base di m colonne, il primale si può riscrivere come segue:

$$
\begin{aligned}
\text{P) minimize } z = & \, c_B^T x_B + c_N^T x_N \\
\text{subject to } & \, B x_B + N x_N = b \\
& \, x_B, x_N \geq 0.
\end{aligned}
$$

dove

$$
\begin{aligned}
A &= [B|N] \\
x^T &= [x_B|x_N] \\
c^T &= [c_B|c_N]
\end{aligned}
$$

Base duale

Il duale si può mettere a sua volta in forma standard, inserendo variabili non-negative di slack:

$$\begin{aligned} \text{D) maximize } w &= b^T y \\ \text{subject to } A^T y + s &= c \end{aligned}$$

con m variabili y libere, $m + n$ variabili $s \geq 0$ e $m + n$ equazioni.

Dato che $A^T = \begin{bmatrix} B^T \\ N^T \end{bmatrix}$, il duale si può riscrivere come segue:

$$\begin{aligned} \text{D) maximize } w &= b^T y \\ \text{subject to } B^T y + s_B &= c_B \\ N^T y + s_N &= c_N. \end{aligned}$$

Per il terorema degli scarti complementari, per ogni coppia di basi primale/duale che si corrispondono si ha

$$x_i s_i = 0.$$

Quindi $s_B = 0$.

Base duale

Da $B^T y = c_B$ si ottiene

$$y = (B^T)^{-1} c_B.$$

Da $N^T y + s_N = c_N$, si ottiene

$$\begin{aligned} s_N &= c_N - N^T (B^T)^{-1} c_B = \\ &= c_N - N^T (B^{-1})^T c_B = \\ &= c_N - (B^{-1} N)^T c_B. \end{aligned} \tag{1}$$

Nel primale invece si ha

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

$$x_N = 0.$$

Così le soluzioni primale e duale si possono calcolare a partire da B, N e i dati **iniziali**.

Cambio di base

Per eseguire un passo di pivot, si sceglie una variabile primale fuori base con costo ridotto negativo, cioè con colonna $q \in N$ tale che $s_q < 0$.

Indichiamo con $T = B^{-1}A$ il tableau corrente (che non vogliamo calcolare esplicitamente). Sia T_q la sua colonna (l'unica che calcoliamo esplicitamente) corrispondente alla variabile x_q:

$$T_q = B^{-1}A_q,$$

dove A_q indica la colonna q della matrice A.

Indichiamo con $p \in B$ l'indice della variabile primale uscente:

$$p = \text{argmin}_{i:T_{iq}>0} \left\{ \frac{(x_B)_i}{T_{iq}} \right\},$$

dove $(x_B)_i$ indica il valore della variabile che è in base sulla riga i.

Il nuovo valore di x_q, entrando in base, è $x'_q = \dfrac{(x_B)_p}{T_{pq}}$.

Cambio di base

Diamo una descrizione algebrica dell'operazione di pivot dalla soluzione di base x alla soluzione di base x', sapendo che:

$$Ax = b \quad Ax' = b$$

perchè entrambe le soluzioni sono ammissibili,

$$x_N = 0 \quad x'_i = 0 \ \forall i \in N \backslash \{q\}.$$

Poiché

$$b = Ax = Bx_B$$

$$b = Ax' = Bx'_B + A_q x'_q,$$

si ha

$$x'_B = x_B - B^{-1} A_q x'_q.$$

Nel duale

$$y^T = c_B^T B^{-1}$$

$$A_q^T y + s_q = c_q \quad \text{ossia} \quad s_q = c_q - y^T A_q.$$

Obiettivo

Il valore dell'obiettivo z dopo il pivot è

$$\begin{aligned} z' = c^T x' = c_B^T x_B' + c_q x_q' &= \\ = c_B^T x_B - c_B^T B^{-1} A_q x_q' + c_q x_q' &= \\ = c_B^T x_B - y^T A_q x_q' + c_q x_q' &= \\ = c_B^T x_B - (c_q - s_q) x_q' + c_q x_q' &= \\ = c_B^T x_B + s_q x_q' = z + s_q x_q'&. \end{aligned}$$

Se l'iterazione non è degenere, $s_q < 0$ e $x_q' > 0$ implicano $z' < z$.

Quindi tutto ciò che serve per procedere con le iterazioni dell'algoritmo del simplesso (x, y, s, z) può essere calcolato a partire dai dati iniziali eseguendo solo prodotti tra vettore e matrice e non tra matrici, pur di conoscere B^{-1}.

Fattorizzazione LU

Per calcolare rapidamente l'inversa di B, si rappresenta $B = LU$ come prodotto tra

- una matrice L *unit lower triangular* (gli elementi sulla diagonale hanno valore 1 e sopra la diagonale hanno valore 0);
- una matrice U *upper triangular* (gli elementi sotto la diagonale hanno valore 0).

Il calcolo di $T_q = B^{-1}A_q$, cioè tale che $BT_q = A_q$, si divide in due step:

- Trovare $d : Ld = A_q$
- Trovare $T_q : UT_q = d$.

Analogamente, il calcolo di $y = (B^{-1})^T c_B$, cioè tale che $B^T y = c_B$, si divide in due step:

- Trovare $d : U^T d = c_B$
- Trovare $y : L^T y = d$.

Tutti questi step sono calcolabili efficientemente per eliminazione Gaussiana.

Aggiornamento di L e U

Nel passaggio da B a B', quando x_q entra in base e x_p esce, bisogna aggiornare efficientemente L e U.

$L^{-1}B = U$ è triangolare superiore.
Sia j l'indice della colonna di U che corriponde a x_p.
$L^{-1}B'$ è ancora triangolare superiore tranne che nella colonna j.

Aggiornamento di L e U

Con

- una permutazione ciclica delle colonne che sposta j in fondo e tutte le colonne da $j+1$ a n a sinistra di una posizione,
- una permutazione ciclica delle righe che sposta la riga j in fondo e tutte le righe da $j+1$ a n in alto di una posizione,

si ottiene una nuova matrice triangolare superiore con in aggiunta alcuni elementi non-zero sull'ultima riga.

Essa si può esprimere come prodotto tra una matrice L' identica a I tranne che nell'ultima riga ed una matrice U' identica alla matrice permutata tranne che nell'ultimo elemento in basso a destra.

Anche questi nuovi coefficienti si ricavano in modo efficiente per eliminazione Gaussiana.

Esempio

$$L^{-1}B = U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} & u_{15} \\ & u_{22} & u_{23} & u_{24} & u_{25} \\ & & u_{33} & u_{34} & u_{35} \\ & & & u_{44} & u_{45} \\ & & & & u_{55} \end{bmatrix} \qquad L^{-1}A_q = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix}$$

Supponiamo che la colonna modificata sia $j = 2$.

$$L^{-1}B' = \begin{bmatrix} u_{11} & w_1 & u_{13} & u_{14} & u_{15} \\ & w_2 & u_{23} & u_{24} & u_{25} \\ & w_3 & u_{33} & u_{34} & u_{35} \\ & w_4 & & u_{44} & u_{45} \\ & w_5 & & & u_{55} \end{bmatrix}$$

Indichiamo con P_j la matrice della permutazione ciclica da j a n.

$$P_j L^{-1} B' P_j^T = \begin{bmatrix} u_{11} & u_{13} & u_{14} & u_{15} & w_1 \\ & u_{33} & u_{34} & u_{35} & w_3 \\ & & u_{44} & u_{45} & w_4 \\ & & & u_{55} & w_5 \\ & u_{23} & u_{24} & u_{25} & w_2 \end{bmatrix}$$

Esempio

La matrice

$$P_j L^{-1} B' P_j^T = \begin{bmatrix} u_{11} & u_{13} & u_{14} & u_{15} & w_1 \\ & u_{33} & u_{34} & u_{35} & w_3 \\ & & u_{44} & u_{45} & w_4 \\ & & & u_{55} & w_5 \\ & u_{23} & u_{24} & u_{25} & w_2 \end{bmatrix}$$

è il prodotto di due matrici triangolari L_1 e U_1:

$$L_1 = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ 0 & l_{52} & l_{53} & l_{54} & 1 \end{bmatrix} \quad U_1 = \begin{bmatrix} u_{11} & u_{13} & u_{14} & u_{15} & w_1 \\ & u_{33} & u_{34} & u_{35} & w_3 \\ & & u_{44} & u_{45} & w_4 \\ & & & u_{55} & w_5 \\ & & & & \hat{w}_2 \end{bmatrix}$$

Esempio

I valori dei nuovi coefficienti si calcolano per eliminazione Gaussiana.

$$\begin{bmatrix} u_{11} & u_{13} & u_{14} & u_{15} & w_1 \\ & u_{33} & u_{34} & u_{35} & w_3 \\ & & u_{44} & u_{45} & w_4 \\ & & & u_{55} & w_5 \\ & u_{23} & u_{24} & u_{25} & w_2 \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ 0 & l_{52} & l_{53} & l_{54} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{13} & u_{14} & u_{15} & w_1 \\ & u_{33} & u_{34} & u_{35} & w_3 \\ & & u_{44} & u_{45} & w_4 \\ & & & u_{55} & w_5 \\ & & & & \hat{w}_2 \end{bmatrix}$$

$$\begin{aligned} l_{52}u_{33} &= u_{23} \\ l_{52}u_{34} + l_{53}u_{44} &= u_{24} \\ l_{52}u_{35} + l_{53}u_{45} + l_{54}u_{55} &= u_{25} \\ l_{52}w_3 + l_{53}w_4 + l_{54}w_5 + \hat{w}_2 &= w_2 \end{aligned}$$

Quindi la nuova fattorizzazione $B' = L'U'$ si ottiene così:

$$L_1U_1 = P_jL^{-1}B'P_j^T \quad \Leftrightarrow \quad L_1U_1 = (P_jL^{-1}L')(U'P_j^T)$$

$L_1 = P_jL^{-1}L'$ e quindi $L' = LP_j^TL_1$
$U_1 = U'P_j^T$ e quindi $U' = U_1P_j$.