

Metody optymalizacji

Wykład nr 8

Paweł Zieliński

Katedra Podstaw Informatyki,
Wydział Informatyki i Telekomunikacji,
Politechnika Wroclawska

Zadanie LP

Rozważmy zadanie programowania liniowego w postaci standardowej

$$\mathbb{X} = \begin{cases} \min \mathbf{c}^T \mathbf{x} \\ \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

gdzie $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $m \leq n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{b} \geq \mathbf{0}$ a $\mathbf{x} \in \mathbb{R}^n$ wektorem zmiennych decyzyjnych.



Zadanie LP

Rozważmy zadanie programowania liniowego w postaci standardowej

$$\mathbb{X} = \begin{cases} \min \mathbf{c}^T \mathbf{x} \\ \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

gdzie $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $m \leq n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{b} \geq \mathbf{0}$ a $\mathbf{x} \in \mathbb{R}^n$ wektorem zmiennych decyzyjnych.

- Niech \mathbf{B} będzie bazą, $\mathbf{x}^B = \mathbf{B}^{-1} \mathbf{b}$ a $\mathbf{x}^T = [\mathbf{x}^B, \mathbf{0}]$ będzie rozwiązaniem bazowym dopuszczalnym. Takie rozwiązanie może być, np. wyznaczone za pomocą metody dwóch faz.



Zadanie LP

Rozważmy zadanie programowania liniowego w postaci standardowej

$$\mathbb{X} = \begin{cases} \min \mathbf{c}^T \mathbf{x} \\ \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

gdzie $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $m \leq n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{b} \geq \mathbf{0}$ a $\mathbf{x} \in \mathbb{R}^n$ wektorem zmiennych decyzyjnych.

- Niech \mathbf{B} będzie bazą, $\mathbf{x}^B = \mathbf{B}^{-1} \mathbf{b}$ a $\mathbf{x}^T = [\mathbf{x}^B, \mathbf{0}]$ będzie rozwiązaniem bazowym dopuszczalnym. Takie rozwiązanie może być, np. wyznaczone za pomocą metody dwóch faz.
- Niech J_B będzie zbiorem indeksów kolumn (zmiennych) bazowych. Załóżmy, że J_B jest uporządkowany (zapewnienie, że algorytm sympleks skończy działanie - nie wpadnie w cykl).



Algorytm sympleks

Krok 0 Oblicz \mathbf{y}^j przedstawienia kolumn \mathbf{A}_j w bazie \mathbf{B} , $\mathbf{y}^j = \mathbf{B}^{-1}\mathbf{A}_j$.

Krok 1 Sprawdź kryterium optymalności, tj. $(\forall j \notin J_{\mathbf{B}})(c_j - (\mathbf{c}^{\mathbf{B}})^T \mathbf{y}^j \geq 0)$. Jeśli kryterium jest spełnione, to **STOP** \mathbf{x} jest rozwiązaniem optymalnym.

W przeciwnym przypadku wybierz kolumnę \mathbf{A}_j taką, że $k \notin J_{\mathbf{B}}$ i $c_k - (\mathbf{c}^{\mathbf{B}})^T \mathbf{y}^k < 0$, która wchodzi i do bazy (lub zmienna x_k).

Krok 2 Wybierz kolumnę \mathbf{B}_{i^*} (zmienną x_{i^*}) wychodzącą z bazy \mathbf{B} ,

$$\frac{x_{i^*}^{\mathbf{B}}}{y_{i^*}^k} = \min \left\{ \frac{x_i^{\mathbf{B}}}{y_i^k} : y_i^k > 0, i \in J_{\mathbf{B}} \right\}.$$

Jeśli $\mathbf{y}^k \leq \mathbf{0}$, wówczas nie ma rozwiązania optymalnego skończonego, **STOP**.

Krok 3 Uaktualnij: $J_{\mathbf{B}}$, rozwiązanie bazowe $\bar{\mathbf{x}}$, bazę $\bar{\mathbf{B}}$ i $\bar{\mathbf{y}}^j$ przedstawienia kolumn macierzy w bazie $\bar{\mathbf{B}}$

$\mathbf{B} := \bar{\mathbf{B}}$, $\mathbf{x} := \bar{\mathbf{x}}$, $\mathbf{y}^j := \bar{\mathbf{y}}^j$. Idź do Krok 1.



Algorytm sympleks...

Uaktualnienie $(m + 1) \times (n + 1)$ elementów tablicy sympleksowej!

Algorytm sympleks...

Uaktualnienie $(m + 1) \times (n + 1)$ elementów tablicy sympleksowej!

- Możemy uaktualniać tylko $(m + 1) \times (m + 1)$ elementów tablicy sympleksowej - kluczem jest posiadanie informacji o B^{-1} . Mając informacje o B^{-1} możemy wyznaczyć x^B , y^j przedstawienia kolumn macierzy A w bazie B , wskaźniki optymalności:

$$c_j - z_j = c_j - (c^B)^T y^j = c_j - (c^B)^T B^{-1} A_j = c_j - \pi^T A_j.$$



Algorytm sympleks...

Uaktualnienie $(m + 1) \times (n + 1)$ elementów tablicy sympleksowej!

- Możemy uaktualniać tylko $(m + 1) \times (m + 1)$ elementów tablicy sympleksowej - kluczem jest posiadanie informacji o B^{-1} . Mając informacje o B^{-1} możemy wyznaczyć x^B , y^j przedstawienia kolumn macierzy A w bazie B , wskaźniki optymalności:

$$c_j - z_j = c_j - (\mathbf{c}^B)^T \mathbf{y}^j = c_j - (\mathbf{c}^B)^T B^{-1} \mathbf{A}_j = c_j - \boldsymbol{\pi}^T \mathbf{A}_j.$$

- Możemy wyznaczać wskaźniki optymalności (Krok 1) tak długo aż znajdziemy pierwszą kolumnę $k \notin J_B$, dla której $c_j - (\mathbf{c}^B)^T \mathbf{y}^j < 0$ i przerwać przeszukiwanie - wiemy wtedy, że aktualne rozwiązanie nie jest optymalne.



Algorytm sympleks...

Uaktualnienie $(m + 1) \times (n + 1)$ elementów tablicy sympleksowej!

- Możemy uaktualniać tylko $(m + 1) \times (m + 1)$ elementów tablicy sympleksowej - kluczem jest posiadanie informacji o B^{-1} . Mając informacje o B^{-1} możemy wyznaczyć x^B , y^j przedstawienia kolumn macierzy A w bazie B , wskaźniki optymalności:

$$c_j - z_j = c_j - (\mathbf{c}^B)^T \mathbf{y}^j = c_j - (\mathbf{c}^B)^T B^{-1} \mathbf{A}_j = c_j - \boldsymbol{\pi}^T \mathbf{A}_j.$$

- Możemy wyznaczać wskaźniki optymalności (Krok 1) tak długo aż znajdziemy pierwszą kolumnę $k \notin J_B$, dla której $c_j - (\mathbf{c}^B)^T \mathbf{y}^j < 0$ i przerwać przeszukiwanie - wiemy wtedy, że aktualne rozwiązanie nie jest optymalne.

Biorąc pod uwagę powyższe spostrzeżenia nowy, zrewidowany, algorytm sympleks w każdej iteracji będzie tylko pamiętał B^{-1} , x^B , $\boldsymbol{\pi}$ i wartość funkcji celu $(\mathbf{c}^B)^T x^B$.



Zrewidowany algorytm sympleks

Założmy, że



Zrewidowany algorytm sympleks

Założmy, że

- pierwsza tablica simpleksowa zawiera macierz jednostkową \mathbf{I} , która znajduje się po lewej stronie (w pierwszych m kolumnach). Zatem mamy początkową bazę $\mathbf{B} = \mathbf{I}$ i $\mathbf{x}^{\mathbf{B}} = \mathbf{b}$.



Zrewidowany algorytm sympleks

Założmy, że

- pierwsza tablica simpleksowa zawiera macierz jednostkową \mathbf{I} , która znajduje się po lewej stronie (w pierwszych m kolumnach). Zatem mamy początkową bazę $\mathbf{B} = \mathbf{I}$ i $\mathbf{x}^{\mathbf{B}} = \mathbf{b}$.
- współczynniki funkcji celu zmiennych odpowiadającym kolumnom macierzy \mathbf{I} są równe zero.



Zrewidowany algorytm sympleks

Założmy, że

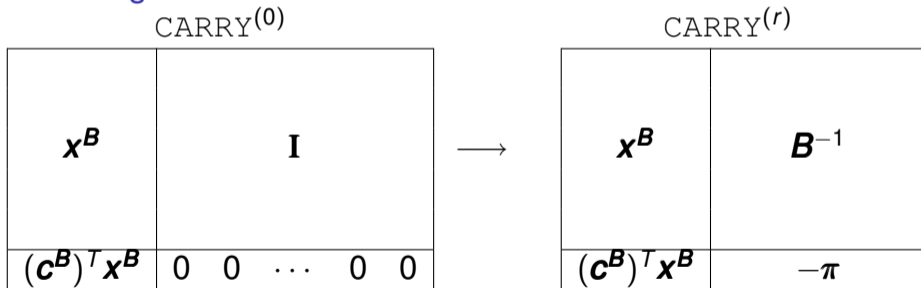
- pierwsza tablica sympleksowa zawiera macierz jednostkową \mathbf{I} , która znajduje się po lewej stronie (w pierwszych m kolumnach). Zatem mamy początkową bazę $\mathbf{B} = \mathbf{I}$ i $\mathbf{x}^{\mathbf{B}} = \mathbf{b}$.
- współczynniki funkcji celu zmiennych odpowiadającym kolumnom macierzy \mathbf{I} są równe zero.

Tę lewą część tablicy sympleksowej, z macierzą \mathbf{I} , nazwiemy **CARRY**, która zawiera $(m + 1) \times (m + 1)$ elementów.



Zrewidowany algorytm sympleks...

Początkowa tablica $\text{CARRY}^{(0)}$ i tablica $\text{CARRY}^{(r)}$ w iteracji r algorytmu zrewidowanego.



Zrewidowany algorytm sympleks - szkic

Krok 1 (Pricing operation) Obliczaj wskaźniki optymalności $c_j - \pi^T \mathbf{A}_j$ dla kolumn niebazowych $j \notin J_B$ aż znajdzie się kolumna $k \notin J_B$, dla której $c_k - \pi^T \mathbf{A}_k < 0$ (kolumna \mathbf{A}_k wchodzi do bazy). Jeśli $(\forall j \notin J_B)(c_j - (\mathbf{c}^B)^T \mathbf{y}^j \geq 0)$, to aktualne rozwiązanie jest optymalne, STOP.

Krok 2 (Column generation) Oblicz przestawienie $\mathbf{y}^k = \mathbf{B}^{-1} \mathbf{A}_k$ kolumny \mathbf{A}_k w bazie. Wybierz kolumnę \mathbf{B}_{i^*} (zmienną x_{i^*}) wychodzącą z bazy \mathbf{B} ,

$$\frac{x_{i^*}^B}{y_{i^*}^k} = \min \left\{ \frac{x_i^B}{y_i^k} : y_i^k > 0, i \in J_B \right\}.$$

Jeśli $\mathbf{y}^k \leq \mathbf{0}$, to nie ma rozwiązania optymalnego skończonego, STOP.

Krok 3 Uaktualnij $\text{CARRY}^{(r)}$ aby otrzymać $\text{CARRY}^{(r+1)}$. Uaktualnienie dokonujemy tak jak dla algorytmu sympleks korzystając z policzonego \mathbf{y}^k .

Krok 4 (Update basis) Uaktualnij indeksy kolumn należących do bazy J_B .



Zrewidowany algorytm sympleks...

Algorytm potrzebuje następujących informacji:

Zrewidowany algorytm sympleks...

Algorytm potrzebuje następujących informacji:

- tablicę zawierającą oryginalną macierz \mathbf{A} ,

Zrewidowany algorytm sympleks...

Algorytm potrzebuje następujących informacji:

- tablicę zawierającą oryginalną macierz \mathbf{A} ,
- informację o kolumnach bazowych J_B ,

Zrewidowany algorytm sympleks...

Algorytm potrzebuje następujących informacji:

- tablicę zawierającą oryginalną macierz \mathbf{A} ,
- informację o kolumnach bazowych J_B ,
- aktualną tablicę $\text{CARRY}^{(r)}$.

Zrewidowany algorytm sympleks...

Algorytm potrzebuje następujących informacji:

- tablicę zawierającą oryginalną macierz \mathbf{A} ,
- informację o kolumnach bazowych J_B ,
- aktualną tablicę $\text{CARRY}^{(r)}$.

Kluczowym krokiem jest Krok 2: mając \mathbf{B}^{-1} mamy wystarczającą ilość informacji, aby wygenerować każdy element tablicy sympleksowej.



Główne zalety zrewidowanej wersji algorytmu

- Nie obliczamy wszystkich wskaźników optymalności odpowiadającym kolumnom niebazowym. Wybieramy pierwszą, dla której wskaźnik jest ujemny.

Główne zalety zrewidowanej wersji algorytmu

- Nie obliczamy wszystkich wskaźników optymalności odpowiadającym kolumnom niebazowym. Wybieramy pierwszą, dla której wskaźnik jest ujemny.
- Ponieważ J_B jest uporządkowany spełniamy założenie twierdzenia 3 (zob. wykłady nr 4 i 5), które zapewnia, że algorytm sympleks wykonuje skończoną liczbę kroków.



Główne zalety zrewidowanej wersji algorytmu

- Nie obliczamy wszystkich wskaźników optymalności odpowiadającym kolumnom niebazowym. Wybieramy pierwszą, dla której wskaźnik jest ujemny.
- Ponieważ J_B jest uporządkowany spełniamy założenie twierdzenia 3 (zob. wykłady nr 4 i 5), które zapewnia, że algorytm sympleks wykonuje skończoną liczbę kroków.
- Krok 1, czyli obliczanie wskaźników, korzysta z oryginalnych kolumn macierzy A .

W wielu przypadkach macierz ograniczeń jest macierzą rzadką (zawiera dużo elementów zerowych), np. w zagadnieniach sieciowych (problem najkrótszej ścieżki, problem maksymalnego przepływu). W tym przypadku możliwa jest redukcja złożoności obliczeniowej i pamięciowej.



Metoda dwóch faz dla zrewidowanego algorytmu

$$\begin{aligned} \min \quad & \sum_{i \in [m]} \mathbf{1} x_i^A \\ \left\{ \begin{array}{l} \mathbf{I} \mathbf{x}^A + \mathbf{A} \mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \\ \mathbf{x}^A \geq \mathbf{0}, \end{array} \right. \end{aligned} \quad (1)$$

gdzie $\mathbf{x}^A \in \mathbb{R}^m$ jest wektorem sztucznych zmiennych.



Metoda dwóch faz dla zrewidowanego algorytmu

$$\begin{aligned} \min \quad & \sum_{i \in [m]} 1 x_i^A \\ \left\{ \begin{array}{l} \mathbf{I} \mathbf{x}^A + \mathbf{A} \mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \\ \mathbf{x}^A \geq \mathbf{0}, \end{array} \right. \end{aligned} \quad (1)$$

gdzie $\mathbf{x}^A \in \mathbb{R}^m$ jest wektorem sztucznych zmiennych.

Aby zastosować zrewidowany algorytm sympleks do zadania (1) musimy zmodyfikować to zadanie i skonstruować równoważne.

Zauważ, że współczynniki funkcji celu odpowiadające kolumnom macierzy jednostkowej (przy zmiennych sztucznych) są równe 1, a powinny być równe 0.



Metoda dwóch faz dla zrewidowanego algorytmu...

Od funkcji celu (wiersza zerowego) odjąć wszystkie wiersze macierzy $\mathbf{I}x^A + \mathbf{A}x$. Otrzymujemy następujące zadanie programowania liniowego:

$$\begin{aligned} \min \quad & \sum_{i \in [m]} 0 x_i^A - \sum_{i \in [m]} a_{i1} x_1 - \cdots - \sum_{i \in [m]} a_{in} x_n \\ \left\{ \begin{array}{l} \mathbf{I}x^A + \mathbf{A}x = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \\ \mathbf{x}^A \geq \mathbf{0}, \end{array} \right. \end{aligned} \quad (2)$$

gdzie $d_j = -\sum_{i \in [m]} a_{ij}$ są współczynnikami funkcji celu oryginalnych zmiennych decyzyjnych x_j , $j \in [n]$, rozwiązywanego w **pierwszej fazie** zadania (2) za pomocą zrewidowanego algorytmu sympleks a wskaźniki optymalności odpowiadające kolumnom oryginalnych zmiennych są postaci: $d_j - \pi^T \mathbf{A}_j$, $j \in [n]$.



Metoda dwóch faz dla zrewidowanego algorytmu...

W drugiej fazie

- wracamy do oryginalnych współczynników c_j zmiennych x_j ,
- generujemy $-\pi^T = -(\mathbf{c}^B)^T \mathbf{B}^{-1}$ (współczynniki funkcji celu zmiennych sztucznych są zerowe).

Zauważ, że $-\pi^T$ jest wskaźnikiem optymalności kolumn odpowiadających macierzy jednostkowej (zmiennych sztucznych),
 $-\pi_j = 0 - (\mathbf{c}^B)^T \mathbf{B}^{-1} \mathbf{e}_j$.



Postać iloczynowa macierzy odwrotnej \mathbf{B}^{-1}

W jaki sposób pamiętać macierz \mathbf{B}^{-1} w kolejnych iteracjach?

Macierz bazowa \mathbf{B} różni się tylko jedną kolumną od macierzy bazowej w kolejnej iteracji $\overline{\mathbf{B}}$.

Zatem niewielkie modyfikacje macierzy odwrotnych w kolejnych iteracjach algorytmu wykonuje się za pomocą przekształceń elementarnych zastosowanych do \mathbf{B}^{-1} .

Niech \mathbf{B}^{-1} będzie aktualną odwrotnością macierzy bazowej, to odwrotność $\overline{\mathbf{B}}^{-1}$ w kolejnej iteracji wyznaczana jest następująco:

$$\overline{\mathbf{B}}^{-1} = \mathbf{E}\mathbf{B}^{-1},$$

gdzie \mathbf{E} jest macierzą przekształcenia.



Postać iloczynowa macierzy odwrotnej B^{-1} ...

Kolumna A_k wchodzi do bazy, kolumna A_{i^*} wychodzi z bazy), macierz przekształcenia E jest postaci:

$$E = \begin{bmatrix} 1 & & \eta_1 & 0 & \cdots & 0 \\ & \ddots & & & & \\ & & \eta_{i^*} & 0 & & \\ & & \eta_{i^*+1} & 1 & & \\ & & \vdots & & \ddots & \\ 0 & \cdots & \eta_m & 0 & & 1 \end{bmatrix}$$

gdzie

$$\eta^T = \left[-\frac{y_1^k}{y_{i^*}^k}, \dots, -\frac{y_{i^*-1}^k}{y_{i^*}^k}, \frac{1}{y_{i^*}^k}, -\frac{y_{i^*+1}^k}{y_{i^*}^k}, \dots, -\frac{y_m^k}{y_{i^*}^k} \right],$$

y^k jest przedstawieniem A_k w bazie B .



Postać iloczynowa macierzy odwrotnej \mathbf{B}^{-1} ...

Informacja o przejściu z \mathbf{B}^{-1} do $\bar{\mathbf{B}}^{-1}$ pamiętana jest za pomocą \mathbf{E} lub oszczędniej za pomocą wektora η . Jeśli założymy, że macierz bazowa w pierwszej iteracji jest równa \mathbf{I} , wówczas macierz \mathbf{B}^{-1} po r -tej iteracji możemy obliczyć

$$\mathbf{B}^{-1} = \mathbf{E}_r \mathbf{E}_{r-1} \cdots \mathbf{E}_1.$$

Jest to **postać iloczynowa macierzy odwrotnej \mathbf{B}^{-1}** .



Postać iloczynowa macierzy odwrotnej \mathbf{B}^{-1} ...

Informacja o przejściu z \mathbf{B}^{-1} do $\bar{\mathbf{B}}^{-1}$ pamiętana jest za pomocą \mathbf{E} lub oszczędniej za pomocą wektora η . Jeśli założymy, że macierz bazowa w pierwszej iteracji jest równa \mathbf{I} , wówczas macierz \mathbf{B}^{-1} po r -tej iteracji możemy obliczyć

$$\mathbf{B}^{-1} = \mathbf{E}_r \mathbf{E}_{r-1} \cdots \mathbf{E}_1.$$

Jest to **postać iloczynowa macierzy odwrotnej \mathbf{B}^{-1}** .

Podobnie obliczamy

$$\pi^T = (\cdots (((\mathbf{c}^B)^T \mathbf{E}_r) \mathbf{E}_{r-1}) \cdots) \mathbf{E}_1,$$

$$\mathbf{y}^k = \mathbf{E}_r (\cdots (\mathbf{E}_2 (\mathbf{E}_1 \mathbf{A}_k)) \cdots).$$

Powyższa metoda ma zastosowanie do dużych problemów z macierzami rzadkimi.



Postać iloczynowa macierzy odwrotnej \mathbf{B}^{-1} ...



Od czasu do czasu ciąg przekształceń elementarnych może być zbyt długi, wtedy dokonuje się **reinwersji**, jest to podyktowane względami pamięciowymi i własnościami numerycznymi (błędy numeryczne), czyli oblicza się macierz \mathbf{B}^{-1} (również $\mathbf{x}^{\mathbf{B}}$) korzystając z oryginalnych kolumn macierzy \mathbf{A} .

Następnie zastępuje się aktualny dłuższy ciąg przekształceń elementarnych krótszym, który prowadzi do macierzy \mathbf{B}^{-1} .



Uwagi na temat treści wykładu

Treść wykładu w całości została przygotowana na podstawie książek

-  Christos H. Papadimitriou, Kenneth Steiglitz.
Combinatorial optimization: algorithms and complexity.
Dover Publications Inc., 1998.
-  Narsingh Deo, Janusz S. Kowalik, Maciej M. Sysło.
Algorytmy optymalizacji dyskretnej z programami w języku PASCAL.
Wydawnictwa Naukowego PWN, 1999.

