

Andrzej Cegielski

Programowanie matematyczne
Część I: Programowanie liniowe

Wydawnictwa Uniwersytetu Zielonogórskiego

Zielona Góra 2002

Spis treści

| | |
|---|------------|
| Przedmowa | vii |
| 1. Wiadomości wstępne | 1 |
| 1.1. Zadania programowania matematycznego | 1 |
| 1.2. Oznaczenia i proste fakty | 4 |
| 2. Podstawy programowania liniowego | 7 |
| 2.1. Przykłady wprowadzające | 7 |
| 2.2. Podstawowe pojęcia | 11 |
| 2.3. Podstawy teoretyczne | 22 |
| 2.3.1. Zbiory wypukłe | 22 |
| 2.3.2. Punkty ekstremalne | 25 |
| 2.3.3. Projekcja metryczna na zbiór wypukły | 28 |
| 2.3.4. Twierdzenia o oddzielaniu i ich konsekwencje | 30 |
| 2.3.5. Funkcje wypukłe | 34 |
| 2.3.6. Stożki | 36 |
| 2.3.7. Lemat Farkasa | 39 |
| 2.3.8. Dalsze konsekwencje twierdzeń o oddzielaniu | 40 |
| 3. Metoda sympleksowa | 45 |
| 3.1. Tablica sympleksowa | 45 |
| 3.2. Opis metody sympleksowej | 52 |
| 3.2.1. Wymiana zmiennej bazowej (piwotyzacja) | 53 |
| 3.2.2. Krawędzie zbioru rozwiązań dopuszczalnych | 55 |
| 3.2.3. Piwotyzacja przy znanym dopuszczalnym rozwiązaniu bazowym (II faza metody sympleksowej) | 56 |
| 3.2.4. Przypadki szczególne | 60 |
| 3.2.5. Piwotyzacja przy nieznanym dopuszczalnym rozwiązaniu bazowym (I faza metody sympleksowej) | 68 |
| 3.3. Degeneracja | 75 |

| | |
|--|------------|
| 3.4. Zrewidowana metoda sympleksowa | 79 |
| 4. Dualizm w programowaniu liniowym | 83 |
| 4.1. Definicja i przykłady | 83 |
| 4.2. Twierdzenia o dualności | 87 |
| 4.3. Dualny algorytm sympleksowy | 97 |
| 5. O złożoności obliczeniowej ZPL | 101 |
| 5.1. Złożoność obliczeniowa algorytmu sympleksowego | 101 |
| 5.2. Metoda Karmarkara punktów wewnętrznych | 104 |
| 5.2.1. Transformacja ZPL do postaci normalnej | 105 |
| 5.2.2. Idea metody Karmarkara | 108 |
| 5.2.3. Przekształcenie rzutowe | 110 |
| 5.2.4. Relaksacja ograniczeń | 112 |
| 5.2.5. Opis metody Karmarkara punktów wewnętrznych | 116 |
| 5.2.6. Zbieżność i złożoność obliczeniowa metody punktów wewnętrznych | 118 |
| 6. Zadanie transportowe | 123 |
| 6.1. Podstawowe pojęcia | 123 |
| 6.1.1. Zbilansowane zadanie transportowe | 123 |
| 6.1.2. Niezbilansowane zadanie transportowe | 124 |
| 6.1.3. Macierz transportowa | 126 |
| 6.1.4. Dualne zadanie transportowe | 131 |
| 6.2. Wyznaczanie dopuszczalnego rozwiązania bazowego | 132 |
| 6.2.1. Opis ogólny | 132 |
| 6.2.2. Metoda kąta północno-zachodniego | 135 |
| 6.2.3. Metoda elementu minimalnego macierzy kosztów | 136 |
| 6.2.4. Metoda aproksymacyjna Vogla | 137 |
| 6.3. Algorytm transportowy | 138 |
| 6.3.1. Podstawowe własności | 138 |
| 6.3.2. Ogólny opis algorytmu transportowego | 142 |
| 6.3.3. Algorytm transportowy | 143 |

| | |
|--|------------|
| 7. Programowanie całkowitoliczbowe | 149 |
| 7.1. Definicje i przykłady wprowadzające | 149 |
| 7.2. Metoda Gomory’ego | 155 |
| 7.3. Metody podziału i ograniczeń | 160 |
| 7.3.1. Ogólny opis metod podziału i ograniczeń | 163 |

Przedmowa

Skrypt ten zawiera materiał przygotowany na podstawie moich dwusemestralnych wykładów z programowania matematycznego, które prowadzę od kilku lat na kierunku matematyka najpierw w Wyższej Szkole Inżynierskiej w Zielonej Górze, potem na Politechnice Zielonogórskiej, a w końcu na Uniwersytecie Zielonogórskim. Pierwsza część skryptu, obejmująca programowanie liniowe odpowiada w dużej części pierwszemu semestrowi wykładów. Natomiast druga część skryptu dotyczy programowania nieliniowego, w szczególności zadań minimalizacji różniczkowalnej bez ograniczeń i z ograniczeniami. Skrypt ten przeznaczony jest głównie dla studentów kierunku matematyka, jednak studenci innych kierunków, jak na przykład informatyka, informatyka i ekonometria, czy zarządzanie i marketing powinni bez większych trudności zrozumieć znakomitą większość przedstawionego w skrypcie materiału.

W skrypcie tym zajmujemy się zagadnieniami matematycznymi sformułowanymi najczęściej jako problemy minimalizacji bądź maksymalizacji funkcji wielu zmiennych przy zadanych ograniczeniach. Rozpatrujemy więc warunki konieczne i wystarczające na to, aby funkcja osiągała minimum względnie maksimum oraz badamy pewne własności zbiorów rozwiązań tych problemów. Ponieważ programowanie matematyczne jest działem matematyki blisko związanym z naukami stosowanymi, takimi, jak nauki techniczne czy ekonomia, więc umieściłem w skrypcie liczne przykłady, które powinny dać czytelnikowi możliwość zapoznania się z licznymi zastosowaniami.

Z uwagi na te zastosowania nie ograniczamy się więc do warunków istnienia rozwiązań, ale zajmujemy się również metodami ich wyznaczania. Metody te podane są – najogólniej rzecz ujmując – jako schematy iteracyjne bądź algorytmy, które prowadzą albo do dokładnych rozwiązań, albo ich odpowiednich przybliżeń. Przedstawiamy również dowody zbieżności tych algorytmów do rozwiązania. Dla części tych metod analizujemy również szybkość tej zbieżności. Pozwala to z jednej strony lepiej zrozumieć czytelnikowi istotę tych metod z drugiej zaś strony – skuteczniejsze ich stosowanie poprzez dobór odpowiedniej metody do rozpatrywanego zagadnienia.

Programowanie liniowe wyodrębniło się jako przedmiot głównie za sprawą G. B. Dantziga, który – w związku z zagadnieniami planowania i ich zastosowań do celów wojskowych – stworzył w 1947 roku metodę sympleksową. Jednak już wcześniej, bo w roku 1939 L. Kantorowicz wydzielił pewną klasę zadań programowania liniowego i podał sposób ich rozwiązania. W roku 1947 ukazała się także ważna praca T. C. Koopmansa wskazująca na znaczenie metod programowania liniowego w ekonomii. Kantorowicz i Koopmans za swoje osiągnięcia zostali uhonorowani w 1975 roku Nagrodą Nobla w dziedzinie ekonomii. Pochodząca od Dantziga metoda sympleksowa jest do dzisiaj powszechnie stosowana. Jednak ostatnie dwadzieścia lat przyniosło przełomowe odkrycia, dzięki którym dysponujemy nowymi metodami rozwiązywania zadań programowania liniowego.

Pierwsza część skryptu dotycząca programowania liniowego podzielona jest na 7 rozdziałów.

We wstępie przedstawiamy klasyfikacje zadań programowania matematycznego.

Rozdział 1. poświęcony jest podstawom programowania liniowego. Po kilku krótkich przykładach podajemy różne postaci zadań programowania liniowego oraz pojęcia i fakty używane w dalszej części skryptu. Następnie przedstawiamy elementy analizy wypukłej. Na ich podstawie podajemy w szczególności warunki konieczne i dostateczne istnienia rozwiązań zadań programowania liniowego.

Głównym celem rozdziału 2. jest przedstawienie metody sympleksowej, która służy rozwiązaniu zadania programowania liniowego. Pokazujemy skończoną zbieżność tej metody dla zadania niezdegenerowanego, a następnie pokazujemy, jak z postaci tablicy sympleksowej można wywnioskować, że brak jest rozwiązań optymalnych bądź, że istnieje ich wiele, że zbiór tych rozwiązań jest nieograniczony oraz, że zadanie programowania liniowego jest sprzeczne (nie ma tzw. rozwiązań dopuszczalnych). Przedstawiamy również metodę postępowania w przypadku zadań zdegenerowanych. Ostatni ustęp tego rozdziału poświęcony jest zrewidowanej metodzie sympleksowej.

Z każdym zadaniem programowania liniowego można wiązać tzw. zadanie dualne. Związki zachodzące między tymi zadaniami omawiamy w rozdziale 3. Wskazujemy także na ich interpretację ekonomiczną. Na koniec przedstawiamy dualny algorytm sympleksowy.

W rozdziale 4. czytelnik znajdzie podstawowe informacje o wykładniczej złożoności obliczeniowej metody sympleksowej. Ponadto w rozdziale tym przedstawiamy metodę Karmarkara rozwiązania zadań programowania liniowego i podajemy dowód wielomianowej złożoności obliczeniowej tej metody. Rozdział ten nie jest niezbędny do zrozumienia pozostałej części skryptu i może być przez czytelnika pominięty przy pierwszym czytaniu.

Ostatnie dwa rozdziały dotyczą specjalnych zadań programowania liniowego.

W rozdziale 5. omawiamy zadanie transportowe i metody jego rozwiązania. W szczególności podajemy, jak wyznaczyć rozwiązanie dopuszczalne tego zadania, a następnie przedstawiamy algorytm transportowy.

W rozdziale 6. zawarte są elementy programowania całkowitoliczbowego. Skupiamy się głównie na metodach rozwiązywania zadań programowania liniowego całkowitoliczbowego. Podajemy dwie metody: metodę Gomory'ego oraz metodę podziału i ograniczeń.

Na podstawie doświadczeń nabytych przy pisaniu tego skryptu mogę powiedzieć, że jego przygotowanie przypomina nieco działanie algorytmu zbieżnego do rozwiązania optymalnego w nieskończenie wielu krokach.

Z każdym krokiem jesteśmy bliżej celu, ale trzeba wybrać moment, w którym otrzymane przybliżenie uznamy za zadowalające. Mając więc świadomość, że przedstawiona postać skryptu nie jest wolna od niedociągnięć, zdecydowałem się przekazać ją do rąk czytelnika licząc na krytyczne uwagi z jego strony.

Zielona Góra, wrzesień 2002

Andrzej Cegielski

ROZDZIAŁ 1

Wiadomości wstępne

Pomiędzy duchem i materią pośredniczy matematyka.

[H. Steinhaus]

1.1. Zadania programowania matematycznego

Niech dane będą: podzbiór $\mathcal{X} \subset \mathbb{R}^n$, funkcja $f : \mathcal{X} \rightarrow \mathbb{R}$ oraz podzbiór $D \subset \mathcal{X}$. Zadanie programowania matematycznego polega na wyznaczeniu punktu x^* , w którym funkcja f ograniczona do zbioru D osiąga minimum (lokalne bądź globalne), o ile takie minimum istnieje. Zbiór D , zwany *zbiorem rozwiązań dopuszczalnych*, podawany jest najczęściej w postaci

$$D = \{x \in \mathcal{X} : c_i(x) = 0 \text{ dla } i \in E \text{ oraz } c_i(x) \leq 0 \text{ dla } i \in I\},$$

gdzie $\mathcal{X} \subset \mathbb{R}^n$, $E = \{1, \dots, p\}$, $I = \{p+1, \dots, m\}$, $c_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i \in E \cup I$. Funkcja f nazywa się *funkcją celu*, zaś funkcje c_i , $i \in E \cup I$, nazywają się *funkcjami ograniczeń* lub *ograniczeniami*. Ograniczenia ponumerowane wskaźnikami $i \in E$ nazywamy *równościowymi*, zaś ponumerowane wskaźnikami $i \in I$ – *nierównościowymi*. Zbiór \mathcal{X} jest najczęściej podawany w postaci: $\mathcal{X} = \mathbb{R}^n$, $\mathcal{X} = \mathbb{R}_+^n$, $\mathcal{X} = \mathbb{Z}^n$, $\mathcal{X} = \mathbb{Z}_+^n$. W ostatnich dwóch przypadkach zadanie nazywa się zadaniem *programowania całkowitoliczbowego*.

Czasem zbiór \mathcal{X} dany jest w postaci $\mathcal{X} = \{0,1\}^n$. Wówczas mówimy o zadaniu *programowania binarnego*.

Dokładniej rzecz ujmując, programowanie matematyczne zajmuje się:

- warunkami istnienia rozwiązań dopuszczalnych (warunkami niesprzeczności problemu),
- warunkami (koniecznymi i wystarczającymi) istnienia minimum (lub przynajmniej skończonego kresu dolnego),
- metodami wyznaczenia tego minimum i punktu x^* realizującego to minimum (w sposób dokładny lub przybliżony),

Czasem do zadań programowania matematycznego zalicza się również:

- zbudowanie – dla konkretnego zagadnienia praktycznego – odpowiedniego modelu matematycznego w postaci zadania minimalizacji funkcji wielu zmiennych,
- interpretację rozwiązania takiego zadania.

Na ogół zakłada się, że zbiór \mathcal{X} jest domknięty i funkcje ograniczeń c_i , $i \in E \cup I$ są ciągłe. W konsekwencji zbiór D jest domknięty. Jeśli ponadto jest on ograniczony i funkcja f jest również ciągła, to na mocy twierdzenia Weierstrassa osiąga ona minimum na D .

Zadania programowania matematycznego dzielimy na:

- *minimalizację różniczkowalną* (lub inaczej gładką), gdy wszystkie funkcje f, c_i , $i \in E \cup I$, są różniczkowalne.
- *minimalizację nieróżniczkowalną* (lub inaczej niegładką), gdy przynajmniej jedna z funkcji f, c_i , $i \in E \cup I$, nie jest różniczkowalna.

Minimalizację różniczkowalną dzielimy na:

- *minimalizację bez ograniczeń*, gdy $D = \mathbb{R}^n$ (lub inaczej, funkcje ograniczeń nie występują)

- *minimalizację z ograniczeniami*, gdy $D \subset \mathbb{R}^n$

Wśród zadań minimalizacji z ograniczeniami wyróżniamy:

- *zadanie programowania liniowego*, gdy wszystkie funkcje f, c_i , $i \in E \cup I$, są liniowe,
- *zadanie programowania kwadratowego*, gdy funkcja f jest kwadratowa, zaś wszystkie funkcje c_i , $i \in E \cup I$, są liniowe,
- *zadanie programowania wypukłego*, gdy $E = \emptyset$ i wszystkie funkcje f, c_i , $i \in I$, są wypukłe.

W zadaniu minimalizacji z ograniczeniami:

- ograniczenie równościowe $c_i(x) = 0$ można zastąpić dwoma ograniczeniami nierównościowymi $c_i(x) \leq 0$ i $-c_i(x) \leq 0$,
- ograniczenie nierównościowe $c_i(x) \leq 0$ można zastąpić ograniczeniem równościowym $c_i(x) + u_i = 0$, wprowadzając tak zwaną *zmienną uzupełniającą* $u_i \geq 0$,
- *zmienną wolną* ($x_j \in \mathbb{R}$) można przedstawić jako różnicę dwóch zmiennych nieujemnych $x_j = x_j^+ - x_j^-$, gdzie $x_j^+, x_j^- \geq 0$.

W konsekwencji, postacie zadania programowania matematycznego z ograniczeniami wyłącznie równościowymi, bądź z ograniczeniami wyłącznie nierównościowymi, ze zmiennymi wyłącznie wolnymi, bądź to ze zmiennymi wyłącznie nieujemnymi są sobie w pewnym sensie równoważne, tzn. z każdej z nich można (przynajmniej teoretycznie) przejść do dowolnej innej. Inna rzecz, że takie przejście może okazać się mało efektywne z punktu widzenia metod programowania matematycznego.

Ponieważ $\max_{x \in D} f(x) = -\min_{x \in D} -f(x)$, więc zadanie maksymalizacji można sprowadzić do zadania minimalizacji i odwrotnie.

1.2. Oznaczenia i proste fakty

W dalszej części używać będziemy następujących oznaczeń i konwencji:

- $x = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$ oznacza element przestrzeni \mathbb{R}^n w postaci wektora kolumnowego (czasem będziemy zapisywać współrzędne wektora w postaci ξ_j , $j = 1, \dots, n$, szczególnie wówczas, gdy symbolem x_k oznaczać będziemy k -ty wyraz ciągu elementów przestrzeni \mathbb{R}^n),
- $x \geq 0$ oznacza, że wszystkie współrzędne wektora x są nieujemne,
- $\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x \geq 0\}$
- $x > 0$ oznacza, że wszystkie współrzędne wektora x są nieujemne, przy czym przynajmniej jedna z nich jest dodatnia,
- $x \gg 0$ oznacza, że wszystkie współrzędne wektora x są dodatnie,
- $\mathbb{R}_{++}^n = \{x \in \mathbb{R}^n : x \gg 0\}$,
- $\langle x, y \rangle$ oznacza iloczyn skalarny wektorów $x, y \in \mathbb{R}^n$,
- $x^\top y$ jest standardowym iloczynem skalarnym wektorów $x, y \in \mathbb{R}^n$ zapisanym w konwencji mnożenia macierzy, czyli

$$x^\top y = \sum_{j=1}^n x_j y_j,$$

- $\|x\| = \sqrt{x^\top x}$ oznacza normę euklidesową wektora $x \in \mathbb{R}^n$,
- $e_j = (0, \dots, 0, 1, 0, \dots, 0)^\top$ oznacza j -ty wersor, tzn. element przestrzeni euklidesowej odpowiedniego wymiaru, którego j -ta współrzędna jest równa 1 zaś pozostałe są równe 0),

- I oznacza macierz jednostkową odpowiedniego wymiaru.

Jeśli A jest macierzą typu $m \times n$, to:

- A_i oznacza i -tą kolumnę macierzy A ,
- A_J oznacza podmacierz utworzoną z kolumn A_j macierzy A , $j \in J \subset \{1, \dots, n\}$,
- A^i oznacza i -ty wiersz macierzy A ,
- $r(A)$ oznacza rząd macierzy A ,
- Ax jest kombinacją liniową kolumn macierzy A , gdzie $x \in \mathbb{R}^n$ jest wektorem współczynników tej kombinacji,
- $u^\top A$ jest kombinacją liniową wierszy macierzy A , gdzie $u \in \mathbb{R}^m$ jest wektorem współczynników tej kombinacji.

Niech

$$A = \begin{bmatrix} A_{11} & \dots & A_{1r} \\ \dots & \dots & \dots \\ A_{p1} & \dots & A_{pr} \end{bmatrix}$$

oznacza macierz w zapisie blokowym z wyróżnionymi podmacierzami A_{ij} , $i = 1, \dots, p$, $j = 1, \dots, r$, takimi, że podmacierze we wspólnej kolumnie macierzy blokowej A mają tę samą liczbę kolumn i podmacierze we wspólnym wierszu macierzy blokowej A mają tę samą liczbę wierszy. Podobne oznaczenia dotyczą macierzy

$$B = \begin{bmatrix} B_{11} & \dots & B_{1s} \\ \dots & \dots & \dots \\ B_{r1} & \dots & B_{rs} \end{bmatrix}.$$

Założmy ponadto, że istnieją iloczyny $A_{ik}B_{kj}$, $i = 1, \dots, p$, $j = 1, \dots, s$, $k = 1, \dots, r$. Wówczas

$$AB = [C_{ij}] = \begin{bmatrix} C_{11} & \dots & C_{1s} \\ \dots & \dots & \dots \\ C_{p1} & \dots & C_{ps} \end{bmatrix},$$

gdzie $C_{ij} = \sum_k A_{ik} B_{kj}$. Innymi słowy, reguły mnożenia macierzy w zapisie blokowym są takie same, jak dla zwykłych macierzy.

ROZDZIAŁ 2

Podstawy programowania liniowego

Każda wiedza zawiera tyle prawdy, ile jest w niej zawartej matematyki.

[I. Kant]

2.1. Przykłady wprowadzające

Przykład 2.1.1. Przedstawimy matematyczny model następującego zagadnienia:

Ogrodnik chce obsadzić ogród wielkości 100 m^2 różami i goździkami. Może on zainwestować maksymalnie 720 zł, zaś z powodu wysokiego nakładu pracy może on zarezerwować najwyżej 60 m^2 na goździki. Ile m^2 powinien on obsadzić każdym rodzajem kwiatów, aby osiągnąć maksymalny zysk. Robocizna i koszty materiałowe wynoszą dla róż 6 zł/m^2 i dla goździków 9 zł/m^2 . Zysk wynosi dla róż 1 zł/m^2 , zaś dla goździków 2 zł/m^2 .

Wprowadzając oznaczenia:

- x_1 – pole powierzchni (w m^2) obsadzonej różami,
- x_2 – pole powierzchni (w m^2) obsadzonej goździkami,

model matematyczny będzie miał postać:

$$\begin{array}{llll}
 \text{maksymalizować} & x_1 + 2x_2 & & \\
 \text{przy ograniczeniach} & x_1 + x_2 & \leq & 100 \\
 & 6x_1 + 9x_2 & \leq & 720 \\
 & x_2 & \leq & 60 \\
 & x_1, x_2 & \geq & 0
 \end{array}$$

Powyższy przykład podpada pod następujący schemat.

Przykład 2.1.2 (*zagadnienie analizy działalności gospodarczej*). Producent wytwarza n towarów P_1, \dots, P_n wykorzystując m surowców lub ogólniej, czynników produkcyjnych R_1, \dots, R_m . Producent dysponuje b_i jednostkami surowca R_i , $i = 1, \dots, m$. Do wyprodukowania jednostki towaru P_j potrzeba a_{ij} jednostek surowca R_i , $i = 1, \dots, m$, $j = 1, \dots, n$. Jednostka wyprodukowanego towaru P_j przynosi zysk c_j jednostek pieniężnych, $j = 1, \dots, n$. Ile jednostek każdego z towarów powinien wytwarzać producent, aby zapewnić sobie maksymalny zysk?

Oznaczając przez x_j ilość jednostek wyprodukowanego towaru P_j , $j = 1, \dots, n$, możemy powyższe zagadnienie sformułować następująco:

$$\begin{array}{llll}
 \text{maksymalizować} & c_1x_1 + c_2x_2 + \dots + c_nx_n & & \\
 \text{przy ograniczeniach} & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & \leq & b_1 \\
 & a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n & \leq & b_2 \\
 & \dots & & \\
 & a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n & \leq & b_m \\
 & x_1, \dots, x_n & \geq & 0.
 \end{array}$$

Przykład 2.1.3 (*zagadnienie diety*). Przedstawimy model matematyczny następującego zagadnienia:

Hodowca karmi krowy dwoma rodzajami karmy, nazwijmy je A i B. Dzienna racja żywnościowa krowy musi zawierać czynniki odżywcze I, II i III w ilościach co najmniej 6, 12 i 4 gramy. Zawartość czynników odżywczych w poszczególnych rodzajach karmy i ich ceny podane są w tabeli:

| | karma A | karma B | minimalna zawartość (g) |
|-----------------------------|---------|---------|-------------------------|
| czynnik odżywczy I (g/kg) | 2 | 1 | 6 |
| czynnik odżywczy II (g/kg) | 2 | 4 | 12 |
| czynnik odżywczy III (g/kg) | 0 | 7 | 4 |
| cena (zł/kg) | 5 | 7 | |

Ile kilogramów karmy A i B powinna zawierać dzienna racja żywnościowa krowy, aby przy zachowaniu podanego minimalnego zapotrzebowania na czynniki odżywcze, powstałe koszty były minimalne?

Oznaczając przez:

- x_1 – ilość (w kg) karmy A w dziennej racji żywnościowej,
- x_2 – ilość (w kg) karmy B w dziennej racji żywnościowej,

powyższe zagadnienie będzie miało następujący model matematyczny:

$$\begin{array}{ll}
 \text{minimalizować} & 5x_1 + 7x_2 \\
 \text{przy ograniczeniach} & \begin{array}{ll} 2x_1 + x_2 & \geq 6 \\ 2x_1 + 4x_2 & \geq 12 \\ 7x_2 & \geq 4 \\ x_1, x_2 & \geq 0. \end{array}
 \end{array}$$

Przykład 2.1.4 (*zagadnienie produkcyjne*). Producent ma za zadanie wyprodukować m towarów P_1, \dots, P_m w ilościach b_1, \dots, b_m wykorzystując n surowców R_1, \dots, R_n . Z jednostki surowca R_j można wyprodukować a_{ij} jednostek towaru P_i , $i = 1, \dots, m$, $j = 1, \dots, n$. Jednostka surowca R_j

kosztuje c_j jednostek pieniężnych, $j = 1, \dots, n$. Ile powinno wynosić zużycie poszczególnych surowców do wyznaczonego profilu produkcyjnego, aby łączne koszty surowców były minimalne?

Oznaczając przez x_j ilość jednostek surowca R_j użytego do produkcji, $j = 1, \dots, n$, powyższe zagadnienie możemy sformułować następująco:

$$\begin{array}{ll} \text{minimalizować} & c_1x_1 + c_2x_2 + \dots + c_nx_n \\ \text{przy ograniczeniach} & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ & a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ & \dots \\ & a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \\ & x_1, \dots, x_n \geq 0. \end{array}$$

Przykład 2.1.5 (*zagadnienie transportowe*). W sieci m magazynów A_1, \dots, A_m składa się pewien towar. Należy go dostarczyć do sieci n sklepów B_1, \dots, B_n . Zapas magazynu A_i wynosi a_i jednostek towaru, $i = 1, \dots, m$. Sklep B_j potrzebuje b_j jednostek towaru, $j = 1, \dots, n$. Zakładamy, że

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j \quad (2.1)$$

(łączna podaż jest równa łącznemu popytowi). Koszty transportu jednostki towaru z magazynu A_i do sklepu B_j wynoszą c_{ij} jednostek pieniężnych, $i = 1, \dots, m$, $j = 1, \dots, n$. Należy określić plan transportowy o minimalnych kosztach zaspokajający zapotrzebowanie wszystkich sklepów (czyli, przy podanym założeniu, wyczerpujący łączne zapasy magazynów).

Jeśli przez x_{ij} oznaczymy ilość towaru transportowanego z magazynu A_i do sklepu B_j , $i = 1, \dots, m$, $j = 1, \dots, n$, to powyższe zagadnienie możemy sformułować następująco:

$$\begin{array}{ll} \text{minimalizować} & \sum_{i=1}^m \sum_{j=1}^n c_{ij}x_{ij} \\ \text{przy ograniczeniach} & \sum_{j=1}^n x_{ij} = a_i, \quad i = 1, \dots, m \\ & \sum_{i=1}^m x_{ij} = b_j, \quad j = 1, \dots, n \\ & x_{ij} \geq 0, \quad i = 1, \dots, m, \\ & \quad j = 1, \dots, n. \end{array} \quad (2.2)$$

2.2. Podstawowe pojęcia

Definicja 2.2.1. Pod pojęciem *zadania programowania liniowego* (ZPL) rozumie się zadanie maksymalizacji lub minimalizacji funkcji liniowej przy ograniczeniach w postaci równości lub nierówności liniowych.

Postać klasyczna ZPL:

$$\begin{array}{ll}
 \text{maksymalizować} & c_1x_1 + \dots + c_nx_n \\
 \text{przy ograniczeniach} & a_{11}x_1 + \dots + a_{1n}x_n \leq b_1 \\
 & \dots \\
 & a_{m1}x_1 + \dots + a_{mn}x_n \leq b_m \\
 & x_j \geq 0, \quad j = 1, \dots, n.
 \end{array}$$

Postać standardowa ZPL:

$$\begin{array}{ll}
 \text{maksymalizować} & c_1x_1 + \dots + c_nx_n \\
 \text{przy ograniczeniach} & a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\
 & \dots \\
 & a_{m1}x_1 + \dots + a_{mn}x_n = b_m \\
 & x_j \geq 0, \quad j = 1, \dots, n.
 \end{array}$$

Ponieważ $\min_x f(x) = -\max_x -f(x)$, więc, podobnie jak w sytuacji ogólnej, problem minimalizacji można sprowadzić do problemu maksymalizacji.

Jeśli jedna ze zmiennych x_j może przyjmować dowolne wartości rzeczywiste, to również, podobnie jak w przypadku ogólnym, można ją wówczas zastąpić dwiema zmiennymi nieujemnymi x_j^+ i x_j^- , przy czym $x_j = x_j^+ - x_j^-$.

Postać klasyczną ZPL można sprowadzić do postaci standardowej wprowadzając tzw. *zmiennne uzupełniające* u_1, \dots, u_m :

$$\begin{array}{ll}
\text{maksymalizować} & c_1x_1 + \dots + c_nx_n \\
\text{przy ograniczeniach} & a_{11}x_1 + \dots + a_{1n}x_n + u_1 = b_1 \\
& \dots \\
& a_{m1}x_1 + \dots + a_{mn}x_n + u_m = b_m \\
& x_j \geq 0, \quad j = 1, \dots, n, \\
& u_i \geq 0, \quad i = 1, \dots, m.
\end{array}$$

Ostatnia postać nosi nazwę *postaci kanonicznej* ZPL.

Postać standardową ZPL można sprowadzić do postaci klasycznej przedstawiając każde równanie $a_{i1}x_1 + \dots + a_{in}x_n = b_i$ jako dwie nierówności $a_{i1}x_1 + \dots + a_{in}x_n \leq b_i$ i $-a_{i1}x_1 - \dots - a_{in}x_n \leq -b_i$, $i = 1, \dots, m$.

Niech $x = (x_1, \dots, x_n)^\top$, $c = (c_1, \dots, c_n)^\top$, $b = (b_1, \dots, b_m)^\top$ i

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}.$$

Wówczas ZPL można przedstawić w zapisie macierzowym:

- w postaci klasycznej:

$$\begin{array}{ll}
\text{maksymalizować} & c^\top x \\
\text{przy ograniczeniach} & Ax \leq b \\
& x \geq 0,
\end{array}$$

- w postaci standardowej:

$$\begin{array}{ll}
\text{maksymalizować} & c^\top x \\
\text{przy ograniczeniach} & Ax = b \\
& x \geq 0.
\end{array}$$

Po sprowadzeniu postaci klasycznej do standardowej, ZPL ma następujący zapis macierzowy:

$$\begin{array}{ll} \text{maksymalizować} & c^\top x \\ \text{przy ograniczeniach} & \begin{bmatrix} A & I \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = b \\ & x, u \geq 0. \end{array}$$

Jest to zapis macierzowy postaci kanonicznej ZPL.

Jeśli postać standardową sprowadzi się do postaci klasycznej w opisany uprzednio sposób, to jej zapis macierzowy wygląda następująco:

$$\begin{array}{ll} \text{maksymalizować} & c^\top x \\ \text{przy ograniczeniach} & \begin{bmatrix} A \\ -A \end{bmatrix} x \leq \begin{bmatrix} b \\ -b \end{bmatrix} \\ & x \geq 0. \end{array}$$

Jeśli rząd $r(A)$ macierzy A jest równy m , to postać standardową ZPL można sprowadzić do postaci kanonicznej wyznaczając z układu równań $Ax = b$, m pewnych zmiennych jako funkcje pozostałych $n - m$ zmiennych. Można to uczynić na przykład metodą eliminacji Gaussa. Metodą tą można również wyznaczyć rzędy macierzy A i $[A, b]$, a także usunąć liniowo zależne równania jeśli $r(A) = r([A, b]) < m$. Zilustrujemy to następującym przykładem.

Przykład 2.2.2. Przy pomocy metody eliminacji Gaussa przedstawimy układ równań

$$\begin{aligned} 2x_1 + 3x_2 - 4x_3 + x_4 &= 2 \\ 4x_1 + 2x_2 - x_3 + 2x_4 &= 1 \\ 6x_1 + x_2 + 2x_3 - x_4 &= 1 \\ 12x_1 + 6x_2 - 3x_3 + 2x_4 &= 4 \end{aligned}$$

w postaci kanonicznej.

Układ ten przyjmuje w zapisie macierzowym postać

$$\begin{bmatrix} 2 & 3 & -4 & 1 \\ 4 & 2 & -1 & 2 \\ 6 & 1 & 2 & -1 \\ 12 & 6 & -3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 4 \end{bmatrix}.$$

Czasem zapisuje się go również w postaci

$$\left[\begin{array}{cccc|c} 2 & 3 & -4 & 1 & 2 \\ 4 & 2 & -1 & 2 & 1 \\ 6 & 1 & 2 & -1 & 1 \\ 12 & 6 & -3 & 2 & 4 \end{array} \right].$$

W tym przypadku należy jednak pamiętać o tym, że poszczególnym kolumnom przypisane są odpowiednie zmienne x_1, x_2, x_3, x_4 i że kolejność przypisania może ulec zmianie w trakcie przekształceń układu, co będzie miało miejsce w naszym przykładzie.

Po przekształceniach elementarnych określonych w metodzie eliminacji Gaussa (polegających w tym przypadku na pomnożeniu pierwszego równania przez -2 i dodaniu do drugiego, pomnożeniu pierwszego równania przez -3 i dodaniu do trzeciego, pomnożeniu pierwszego równania przez -6 i dodaniu do czwartego, dalej pomnożeniu drugiego równania przez -2 i dodaniu do trzeciego, następnie pomnożeniu trzeciego równania przez -1 i dodaniu do czwartego i w końcu zamianie kolejności zapisu zmiennych x_3 i x_4) otrzymamy równoważny układ równań w zapisie macierzowym ze wszystkimi elementami pod główną przekątną równymi zeru

$$\begin{bmatrix} 2 & 3 & 1 & -4 \\ 0 & -4 & 0 & 7 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_4 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 1 \\ 0 \end{bmatrix}.$$

Z zapisu tego widać, że $r(A) = r([A, b]) = 3$. Dokonując dalszych przekształceń elementarnych określonych w metodzie eliminacji Gaussa (polegających w tym przypadku na usunięciu równania czwartego – liniowo za-

leżnego, na pomnożeniu trzeciego równania przez $1/4$ i dodaniu do pierwszego, następnie pomnożeniu drugiego równania przez $3/4$ i dodaniu do pierwszego i w końcu na podzieleniu pierwszego równania przez 2, oraz drugiego i trzeciego przez -4) otrzymamy postać kanoniczną

$$\begin{bmatrix} 1 & 0 & 0 & \frac{5}{8} \\ 0 & 1 & 0 & -\frac{7}{4} \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_4 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{3}{4} \\ -\frac{1}{4} \end{bmatrix}$$

lub

$$\begin{bmatrix} I & \tilde{A} \end{bmatrix} \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \tilde{b},$$

gdzie

$$x_B = (x_1, x_2, x_4)^\top, x_N = x_3, \tilde{A} = \begin{bmatrix} \frac{5}{8} \\ -\frac{7}{4} \\ 0 \end{bmatrix}, \tilde{b} = \begin{bmatrix} 0 \\ \frac{3}{4} \\ -\frac{1}{4} \end{bmatrix}.$$

Rozwiązanie układu równań można przedstawić wówczas następująco:

$$x_B = \tilde{b} - \tilde{A}x_N.$$

Powyższy opis ilustruje ogólny schemat: Zakładając, jak wyżej, że pierwsze m kolumn macierzy A jest liniowo niezależnych, można układ równań $Ax = b$ przekształcić metodą eliminacji Gaussa do tzw. *postaci trapezowej*: $A^*x = b^*$, gdzie:

$$[A^*, b^*] = \begin{bmatrix} a_{11}^* & a_{12}^* & \dots & a_{1r}^* & a_{1,r+1}^* & \dots & a_{1n}^* & b_1^* \\ 0 & a_{21}^* & \dots & a_{2r}^* & a_{2,r+1}^* & \dots & a_{2n}^* & b_2^* \\ \dots & & & & & & & \dots \\ 0 & 0 & \dots & a_{rr}^* & a_{r,r+1}^* & \dots & a_{rn}^* & b_r^* \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & b_{r+1}^* \\ \dots & & & & & & & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & b_n^* \end{bmatrix}$$

oraz $a_{11}^* \neq 0, \dots, a_{rr}^* \neq 0$. Elementy b_{r+1}^*, \dots, b_n^* decydują o istnieniu rozwiązania układu równań. Układ posiada rozwiązanie wtedy i tylko wtedy, gdy wszystkie one są równe zeru. W przypadku istnienia rozwiązań ich ilość zależy od r . Jeśli $r = n$, to układ posiada dokładnie jedno rozwiązanie. Jeśli $r < n$, to układ posiada nieskończenie wiele rozwiązań zależnych od $n - r$ parametrów. Jeśli przynajmniej jeden z elementów b_{r+1}^*, \dots, b_n^* jest różny od zera, to układ nie posiada rozwiązania.

Układ równań $A^*x = b^*$ można teraz (w przypadku istnienia rozwiązania) rozwiązać „od dołu do góry”. Po przekształceniach elementarnych opisanych w metodzie eliminacji Gaussa, otrzymamy równoważny mu układ równań

$$\begin{bmatrix} I & \tilde{A} \end{bmatrix} \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \tilde{b}$$

lub inaczej $x_B = \tilde{b} - \tilde{A}x_N$.

W dalszej części tego rozdziału zakładamy, że $r(A) = m$. Jak zauważyliśmy wyżej, założenie to nie ogranicza ogólności rozważań.

Rozwiązanie układu równań $Ax = b$ można również opisać inaczej. Niech w zapisie blokowym $A = \begin{bmatrix} A_B & A_N \end{bmatrix}$, przy czym A_B jest macierzą nieosobliwą ($\det A_B \neq 0$). Nietrudno zauważyć, że założenie to również nie ogranicza ogólności rozważań (w razie potrzeby można ewentualnie przenumerać zmienne x_1, \dots, x_n). Dalej, niech $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$ (na przykład $x_B = (x_1, \dots, x_m)^\top$, $x_N = (x_{m+1}, \dots, x_n)^\top$). Współrzędne wektora x_B nazywamy *zmiennymi bazowymi*, zaś współrzędne wektora x_N – *zmiennymi niebazowymi*. Omawiany układ równań można zapisać w postaci

$$\begin{bmatrix} A_B & A_N \end{bmatrix} \begin{bmatrix} x_B \\ x_N \end{bmatrix} = b$$

lub inaczej w postaci

$$A_B x_B + A_N x_N = b.$$

Po lewostronnym pomnożeniu obu stron ostatniego równania przez A_B^{-1} otrzymamy

$$x_B = A_B^{-1}b - A_B^{-1}A_Nx_N. \quad (2.3)$$

Definicja 2.2.3. Nieosobliwa podmacierz A_B stopnia m macierzy A nazywa się *macierzą bazową*. Jest ona utworzona z m liniowo niezależnych kolumn macierzy A .

Definicja 2.2.4. Niech $A = [A_B \ A_N]$, gdzie A_B jest macierzą bazową. Wektor $\bar{x} = (\bar{x}_B^\top, 0)^\top$, gdzie $\bar{x}_B = A_B^{-1}b$, nazywa się *rozwiązaniem bazowym* układu równań $Ax = b$. Współrzędne wektora x_B danego równością (2.3) nazywają się *zmiennymi bazowymi*. Pozostałe współrzędne wektora $x = (x_B^\top, x_N^\top)^\top$ będącego rozwiązaniem układu równań nazywają się *zmiennymi niebazowymi*.

Uwaga 2.2.5. Każde nieujemne rozwiązanie x układu równań $Ax = b$ nazywa się *rozwiązaniem dopuszczalnym*. Zbiór rozwiązań dopuszczalnych ZPL oznaczamy symbolem X . Rozwiązanie bazowe $(x_B, 0)$ nazywa się *dopuszczalnym rozwiązaniem bazowym*, jeśli $x_B \geq 0$.

Uwaga 2.2.6. Maksymalna liczba rozwiązań bazowych układu równań $Ax = b$ wynosi $\binom{n}{m}$. Może być ich jednak mniej, jeśli pewne podmacierze typu $m \times m$ macierzy A będą osobliwe. Ponadto niektóre z rozwiązań bazowych mogą nie być dopuszczalne.

Definicja 2.2.7. Rozwiązanie bazowe $(x_B^\top, 0)^\top$ nazywa się *zdegenerowanym*, jeśli przynajmniej jedna ze zmiennych bazowych jest równa zeru. W przeciwnym wypadku rozwiązanie bazowe nazywa się *niezdegenerowanym*.

Definicja 2.2.8. Dopuszczalne rozwiązanie x^* ZPL nazywa się *rozwiązaniem optymalnym*, jeśli nie istnieje rozwiązanie dopuszczalne ZPL o większej (albo mniejszej w przypadku minimalizacji) wartości funkcji celu niż $c^\top x^*$. Zbiór rozwiązań optymalnych ZPL oznaczamy symbolem X^* . Jeśli dopuszczalne rozwiązanie bazowe jest rozwiązaniem optymalnym, to nazywa się ono *optymalnym rozwiązaniem bazowym*.

W przypadku, gdy ZPL jest dane w postaci standardowej, jego funkcję celu możemy zapisać następująco:

$$z = c^\top x = c_B^\top x_B + c_N^\top x_N$$

Po wstawieniu w miejsce x_B wyrażenia określonego równością (2.3) otrzymamy

$$z = c_B^\top A_B^{-1} b + (c_N^\top - c_B^\top A_B^{-1} A_N) x_N \quad (2.4)$$

W zagadnieniach ekonomicznych wektor $c_B^\top A_B^{-1}$ nazywa się wektorem *cen ukrytych*, wektor $c_B^\top A_B^{-1} A_N$ – wektorem *kosztów równoważnych*, wektor $c_N^\top - c_B^\top A_B^{-1} A_N$ – wektorem *kosztów zredukowanych*, zaś przeciwny do niego wektor $c_B^\top A_B^{-1} A_N - c_N^\top$ – wektorem *ujemnych kosztów zredukowanych*. Podamy teraz interpretację tych pojęć na przykładzie zagadnienia produkcyjnego (przykład 2.1.4). Przypuśćmy, że $x = (x_B^\top, x_N^\top)^\top$ jest pewnym rozwiązaniem dopuszczalnym (wektorem udziałów surowcowych) i, że $x' = (x_B'^\top, x_N'^\top)^\top$ jest takim rozwiązaniem dopuszczalnym, że $x'_N = x_N + e_j$ i $x'_B = A_B^{-1} b - A_B^{-1} A_N x'_N$. Inaczej mówiąc, j -tą zmienną niebazową (udział j -tego surowca) zwiększono o jednostkę, a zmienne bazowe zmieniono tak, aby x' było nadal rozwiązaniem dopuszczalnym. Niech z i z' będą wartościami funkcji celu odpowiadającymi tym rozwiązaniom. Wówczas z równości (2.4) otrzymamy po prostych przekształceniach

$$z' = z - (c_B^\top A_B^{-1} A_j - c_j),$$

gdzie A_j oznacza j -tą kolumnę macierzy A . Widzimy więc, że zwiększenie udziału j -tego surowca o jednostkę (przy zachowaniu profilu produkcji) powoduje redukcję wartości funkcji celu (kosztów surowcowych) o wielkość $c_B^\top A_B^{-1} A_j - c_j$. Zauważmy, że wielkość ta jest j -tą współrzędną wektora ujemnych kosztów zredukowanych. Jeśli przy takiej zmianie rozwiązania x na x' (zwiększenie udziału j -tego surowca o jednostkę) funkcja celu (koszty surowcowe) miałyby pozostać bez zmiany, to oczywiście

$$c_j = c_B^\top A_B^{-1} A_j,$$

czyli cena jednostki j -tego surowca musiałaby być równa j -tej współrzędnej wektora kosztów równoważnych. Przy okazji zauważmy, że jeżeli x jest (bazowym) rozwiązaniem optymalnym zadania minimalizacji, to wszystkie koszty zredukowane dla zmiennych niebazowych (nie używanych surowców) są nieujemne albo inaczej ich koszty równoważne są równe co najwyżej jednostkowym cenom surowcowym. Dla rozwiązania optymalnego zadania maksymalizacji wszystkie ujemne koszty zredukowane dla zmiennych niebazowych są nieujemne.

Interpretacja ekonomiczna wektora cen ukrytych będzie podana później.

Przykład 2.2.9. Określimy wszystkie bazowe rozwiązania dopuszczalne układu równań

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}.$$

Oczywiście $r(A) = 2$.

a) Wybierając $B = \{1, 2\}$ otrzymamy $x_B = (2, 1)^\top$ i dopuszczalne rozwiązanie bazowe $x = (2, 1, 0)^\top$,

b) wybierając z kolei $B = \{2, 3\}$ otrzymamy $x_B = (\frac{5}{3}, \frac{2}{3})^\top$ i dopuszczalne rozwiązanie bazowe $x = (0, \frac{5}{3}, \frac{2}{3})^\top$.

c) natomiast wybierając $B = \{1, 3\}$ otrzymamy $x_B = (5, -1)^\top$ i niedopuszczalne rozwiązanie bazowe $x = (5, 0, -1)^\top$.

Przykład 2.2.10. Wyznamy wszystkie bazowe rozwiązania dopuszczalne ZPL:

$$\begin{array}{ll} \text{maksymalizować} & c_1x_1 + c_2x_2 \\ \text{przy ograniczeniach} & \begin{array}{ll} -x_1 + 2x_2 & \leq 8 \\ x_1 + x_2 & \leq 10 \\ 2x_1 + x_2 & \leq 16 \\ x_1 & \leq 6 \\ x_1, x_2 & \geq 0. \end{array} \end{array}$$

Postać standardowa tego zadania w zapisie macierzowym wygląda następująco:

$$\begin{array}{ll} \text{maksymalizować} & c_1x_1 + c_2x_2 \\ \text{przy ograniczeniach} & \begin{bmatrix} -1 & 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 8 \\ 10 \\ 16 \\ 6 \end{bmatrix} \end{array}$$

Oczywiście $r(A) = r([A, b]) = 4$. Niektóre z rozwiązań bazowych przedstawiają się następująco:

a) dla $B = \{2, 4, 5, 6\}$, $x_B = (4, 6, 12, 6)^\top$ i $x = (0, 4, 0, 6, 12, 6)^\top$ jest dopuszczalnym rozwiązaniem bazowym,

b) dla $B = \{1, 2, 5, 6\}$, $x_B = (4, 6, 2, 2)^\top$ i $x = (4, 6, 0, 0, 2, 2)^\top$ jest dopuszczalnym rozwiązaniem bazowym,

c) dla $B = \{1, 2, 3, 4\}$, $x_B = (6, 4, 6, 0)^\top$ i $x = (6, 4, 6, 0, 0, 0)^\top$ jest rozwiązaniem zdegenerowanym,

d) dla $B = \{1, 2, 3, 5\}$, $x_B = (6, 4, 6, 0)^\top$ i $x = (6, 4, 6, 0, 0, 0)^\top$ jest rozwiązaniem zdegenerowanym,

e) dla $B = \{1, 2, 3, 6\}$, $x_B = (6, 4, 6, 0)^\top$ i $x = (6, 4, 6, 0, 0, 0)^\top$ jest rozwiązaniem zdegenerowanym,

f) dla $B = \{1, 3, 4, 5\}$, $x_B = (6, 14, 4, 4)^\top$ i $x = (6, 0, 14, 4, 4, 0)^\top$ jest dopuszczalnym rozwiązaniem bazowym,

g) dla $B = \{3, 4, 5, 6\}$, $x_B = (8, 10, 16, 6)^\top$ i $x = (0, 0, 8, 10, 16, 6)^\top$ jest dopuszczalnym rozwiązaniem bazowym,

h) dla $B = \{1, 2, 4, 6\}$, $x_B = (\frac{24}{5}, \frac{32}{5}, -\frac{6}{5}, \frac{6}{5})^\top$ i $x = (\frac{24}{5}, \frac{32}{5}, 0, -\frac{6}{5}, 0, \frac{6}{5})^\top$ jest niedopuszczalnym rozwiązaniem bazowym.

Poza tym istnieje jeszcze 6 niedopuszczalnych rozwiązań bazowych.

Zwróćmy uwagę na to, że trzem różnym wyborom zbioru B odpowiada jedno (potrójne) rozwiązanie zdegenerowane. Czytelnikowi zalecamy narysowanie zbioru rozwiązań dopuszczalnych omawianego zadania i porównanie wierzchołków tego zbioru z wyznaczonymi bazowymi rozwiązaniami dopuszczalnymi.

Ćwiczenie 2.2.11. Przedstawić:

a) problem ogrodnika (przykład 2.1.1);

b) zagadnienie diety (przykład 2.1.3);

jako ZPL w postaci klasycznej i standardowej i rozwiązać te zadania graficznie.

Ad a) Postać klasyczna tego problemu wygląda następująco:

$$\begin{array}{ll} \text{maksymalizować} & x_1 + 2x_2 \\ \text{przy ograniczeniach} & \begin{array}{ll} x_1 + x_2 & \leq 100 \\ 6x_1 + 9x_2 & \leq 720 \\ x_2 & \leq 60 \\ x_1, x_2 & \geq 0. \end{array} \end{array}$$

Stąd otrzymujemy postać standardową (która jest jednocześnie postacią kanoniczną):

$$\begin{array}{ll} \text{maksymalizować} & x_1 + 2x_2 \\ \text{przy ograniczeniach} & \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 6 & 9 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 100 \\ 720 \\ 60 \end{bmatrix} \\ & x_1, x_2, u_1, u_2, u_3 \geq 0. \end{array}$$

Ad b) Postać klasyczna tego problemu wygląda następująco:

$$\begin{array}{ll} \text{minimalizować} & 5x_1 + 7x_2 \\ \text{przy ograniczeniach} & \begin{array}{ll} 2x_1 + x_2 & \geq 6 \\ 2x_1 + 4x_2 & \geq 12 \\ 7x_2 & \geq 4 \\ x_1, x_2 & \geq 0 \end{array} \end{array}$$

albo

$$\begin{array}{rcl}
 \text{maksymalizować} & -5x_1 - 7x_2 & \\
 \text{przy ograniczeniach} & -2x_1 - x_2 \leq -6 & \\
 & -2x_1 - 4x_2 \leq -12 & \\
 & -7x_2 \leq -4 & \\
 & x_1, x_2 \geq 0. &
 \end{array}$$

Pozostałą część zadania pozostawiamy do uzupełnienia czytelnikowi.

2.3. Podstawy teoretyczne

*Naukę buduje się z faktów, tak jak dom buduje się z kamieni.
Jednak zbiór faktów nie jest jeszcze nauką, tak jak stos kamieni
nie jest jeszcze domem.*

[H. Poincaré]

W niniejszym ustępie zestawiamy pojęcia i twierdzenia dotyczące zbiorów i funkcji wypukłych, które będą nam potrzebne w dalszej części. Niektóre z twierdzeń przedstawiamy bez dowodów. Dowody te można znaleźć w podręcznikach do analizy wypukłej, na przykład w podręcznikach J. B. Hiriarta-Urruty'ego i C. Lemaréchala [10, 11].

2.3.1. Zbiory wypukłe

Definicja 2.3.1. Zbiór $K \subset \mathbb{R}^n$ nazywa się *zbiorem wypukłym*, jeśli dla dowolnych $x, y \in K$ i dla każdego $\lambda \in [0, 1]$

$$(1 - \lambda)x + \lambda y \in K.$$

Otoczką wypukłą zbioru $S \subset \mathbb{R}^n$ nazywamy najmniejszy zbiór wypukły zawierający S . Oznaczamy ją symbolem $\text{conv } S$. *Kombinacją wypukłą* elementów $x_1, \dots, x_m \in \mathbb{R}^n$ nazywamy element $x \in \mathbb{R}^n$ postaci

$$x = \sum_{i=1}^m \lambda_i x_i,$$

gdzie $\lambda_i \geq 0$, $i = 1, \dots, m$, $\sum_{i=1}^m \lambda_i = 1$, $m \in \mathbb{N}$.

Czytelnikowi pozostawiamy dowód następującego lematu.

Lemat 2.3.2. *Podzbiór $C \subset \mathbb{R}^n$ jest wypukły wtedy i tylko wtedy, gdy do C należy dowolna kombinacja wypukła jego elementów.*

Przykład 2.3.3. Podamy teraz ważne dla zastosowań przykłady zbiorów wypukłych:

- a) Przekrój dowolnej rodziny zbiorów wypukłych jest zbiorem wypukłym.
- b) Dowolna podprzestrzeń afiniczna, w szczególności dowolna *hiperpłaszczyzna* $\{x \in \mathbb{R}^n : a^\top x = b\}$ jest zbiorem wypukłym.
- c) Dowolna *półprzestrzeń* $\{x \in \mathbb{R}^n : a^\top x \leq b\}$ jest zbiorem wypukłym.
- d) Ponieważ przekrój dowolnej liczby zbiorów wypukłych jest zbiorem wypukłym, więc w szczególności zbiory postaci

$$\bigcap_{i \in I} \{x \in \mathbb{R}^n : a_i^\top x = b_i\}$$

są wypukłe. Oczywiście algebraicznie są one zbiorami rozwiązań układu równań liniowych $a_i^\top x = b_i$, $i \in I$.

- e) Z tych samych powodów wypukłe są zbiory postaci

$$\bigcap_{i \in I} \{x \in \mathbb{R}^n : a_i^\top x \leq b_i\}.$$

W przypadku, gdy I jest zbiorem skończonym, zbiory takie nazywamy *zbiorami wielościennymi*. Szczególnym przypadkiem zbioru wielościennego jest *sympleks standardowy*, zdefiniowany następująco

$$\Delta_m = \{w = (\lambda_1, \dots, \lambda_m)^\top \in \mathbb{R}^m : \lambda_i \geq 0, i = 1, \dots, m, \sum_{i=1}^m \lambda_i = 1\}.$$

Innym przykładem zbioru wielościennego jest zbiór rozwiązań dopuszczalnych X zadania programowania liniowego. Zbiorem wielościennym jest również zbiór rozwiązań optymalnych X^* tego zadania ponieważ

$$X^* = X \cap \{x \in \mathbb{R}^n : c^\top x = c^\top x^*\},$$

gdzie x^* jest ustalonym elementem zbioru X^* . Tak więc zbiory X i X^* są wypukłe.

- f) Dowolna kula $B(y, r) = \{x \in \mathbb{R}^n : \|x - y\| \leq r\}$ jest zbiorem wypukłym.

Lemat 2.3.4. *Otoczka wypukła zbioru $S \subset \mathbb{R}^n$ jest postaci*

$$\text{conv } S = \left\{ \sum_{i=1}^m \lambda_i x_i : x_i \in S, w = (\lambda_1, \dots, \lambda_m)^\top \in \Delta_m, m \in \mathbb{N} \right\}. \quad (2.5)$$

Dowód

„ \supset ” Na mocy definicji 2.3.1 $\text{conv } S$ jest zbiorem wypukłym. Skoro $\text{conv } S \supset S$, więc z lematu 2.3.2 wynika, że $\text{conv } S$ zawiera również wszystkie kombinacje wypukłe elementów zbioru S .

„ \subset ” Nietrudno pokazać, że zbiór po prawej stronie równości (2.5) jest wypukły. Zawiera on S , więc zgodnie z definicją 2.3.1 zawiera on również $\text{conv } S$. \square

Liczbę elementów wchodzących do kombinacji wypukłych, o których mowa w równości (2.5) można ograniczyć. Zachodzi bowiem następujące twierdzenie.

Twierdzenie 2.3.5 (Carathéodory). *Otoczka wypukła zbioru $S \subset \mathbb{R}^n$ składa się ze wszystkich kombinacji wypukłych co najwyżej $n + 1$ elementów zbioru S :*

$$\text{conv } S = \left\{ \sum_{i=1}^m \lambda_i x_i : x_i \in S, w = (\lambda_1, \dots, \lambda_m)^\top \in \Delta_m, m \leq n + 1 \right\}.$$

Dowód. Niech $x \in \text{conv } S$. Zgodnie z równością (2.5) $x = \sum_{i=1}^m \lambda_i x_i$ dla $m \in \mathbb{N}$, $x_1, \dots, x_m \in S$ i $\lambda \in \Delta_m$. Spośród wszystkich przedstawień wektora x w przedstawionej postaci możemy wybrać to, dla którego liczba m jest najmniejsza. Pozostaje pokazać, że $m \leq n + 1$. Przypuśćmy, że jest przeciwnie. Niech $x'_i = (x_i, 1) \in \mathbb{R}^n \times \mathbb{R}$. Ponieważ $m > n + 1$, więc wektory x'_i , $i = 1, \dots, m$, są liniowo zależne. Oznacza to, że istnieją takie liczby $\alpha_1, \dots, \alpha_m$ nie wszystkie równe zeru, że $\sum_{i=1}^m \alpha_i x'_i = 0$, czyli $\sum_{i=1}^m \alpha_i x_i = 0$ i $\sum_{i=1}^m \alpha_i = 0$. Tak więc wśród α_i , $i = 1, \dots, m$, istnieją liczby dodatnie. Niech

$$\varepsilon_0 = \frac{\lambda_{i_0}}{\alpha_{i_0}} = \min \left\{ \frac{\lambda_i}{\alpha_i} : \alpha_i > 0, i = 1, \dots, m \right\} \quad (2.6)$$

i niech

$$\bar{\lambda}_i = \lambda_i - \varepsilon_0 \alpha_i, \quad i = 1, \dots, m.$$

Z równości (2.6) wynika, że $\bar{\lambda}_i \geq 0$ dla $i = 1, \dots, m$ i że $\bar{\lambda}_{i_0} = 0$. Zatem

$$\sum_{i=1}^m \bar{\lambda}_i x_i = \sum_{i=1}^m \lambda_i x_i - \varepsilon_0 \sum_{i=1}^m \alpha_i x_i = x$$

i

$$\sum_{i=1}^m \bar{\lambda}_i = \sum_{i=1}^m \lambda_i - \varepsilon_0 \sum_{i=1}^m \alpha_i = \sum_{i=1}^m \lambda_i = 1.$$

Ponieważ $\bar{\lambda}_{i_0} = 0$, więc widzimy, że x można przedstawić jako kombinację wypukłą $m - 1$ elementów. \square

2.3.2. Punkty ekstremalne

Definicja 2.3.6. Punkt x należący do zbioru wypukłego $K \subset \mathbb{R}^n$ nazywa się *punktem ekstremalnym* tego zbioru, jeśli nie jest środkiem odcinka łączącego dwa różne punkty zbioru K :

$$x = \frac{1}{2}(x' + x'') \text{ i } x', x'' \in K \Rightarrow x' = x''.$$

Zbiór punktów ekstremalnych zbioru K oznaczamy symbolem $\text{ext } K$. W przypadku, gdy K jest zbiorem wielościnnym, jego punkty ekstremalne nazywają się *wierzchołkami*.

Przykład 2.3.7. Podamy zbiory punktów ekstremalnych dla wybranych zbiorów wypukłych.

- a) Dla $K = [a, b] \subset \mathbb{R}$, przy czym $a \leq b$, $\text{ext } K = \{a, b\}$,
- b) dla $K = \mathbb{R}_+^n = \{x \in \mathbb{R}^n : x \geq 0\}$, $\text{ext } K = \{0\}$,
- c) dla $K = \Delta_m$, $\text{ext } K = \{e_i : i = 1, \dots, m\}$,
- d) dla $K = \{x \in \mathbb{R}^n : Ax = b\}$, $\text{ext } K = K$ jeśli układ równań $Ax = b$ posiada dokładnie jedno rozwiązanie i $\text{ext } K = \emptyset$ w przeciwnym wypadku.

Twierdzenie 2.3.8. Niech $X = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ będzie zbiorem rozwiązań dopuszczalnych zadania programowania liniowego w postaci standardowej. Wówczas $x \in X$ jest wierzchołkiem zbioru X wtedy i tylko wtedy, gdy x jest bazowym rozwiązaniem dopuszczalnym.

Dowód

(\Rightarrow) Niech x będzie wierzchołkiem zbioru X . Zapiszmy (po ewentualnym przenumowaniu zmiennych) $x = (x_1, \dots, x_k, 0, \dots, 0)^\top$, gdzie $x_i > 0$ dla $i = 1, \dots, k$. Oznaczmy $K = \{1, \dots, k\}$ i $x_K = (x_1, \dots, x_k)^\top$. Pokażemy, że macierz $A_K = [A_1, \dots, A_k]$ ma liniowo niezależne kolumny ($r(A_K) = k$). Przypuśćmy, że jest przeciwnie, tzn. $A_K a_K = 0$ dla pewnego niezerowego wektora $a_K = (\alpha_1, \dots, \alpha_k)^\top$. Niech $a = (\alpha_1, \dots, \alpha_k, 0, \dots, 0)^\top \in \mathbb{R}^n$. Oczywiście wektor a jest również niezerowy. Niech $\varepsilon > 0$ będzie liczbą na tyle małą, że $x' = x + \varepsilon a \geq 0$ i $x'' = x - \varepsilon a \geq 0$. Taka liczba istnieje, ponieważ $x_i > 0$ dla $i = 1, \dots, k$. Zauważmy, że x' i x'' są rozwiązaniami dopuszczalnymi, gdyż

$$A(x \pm \varepsilon a) = Ax \pm \varepsilon Aa = b \pm \varepsilon A_K a_K = b.$$

Ale $x = \frac{1}{2}(x' + x'')$ i $x' \neq x''$, gdyż $\varepsilon \neq 0$. Zatem x nie jest wierzchołkiem zbioru X . Uzyskana sprzeczność dowodzi, że $r(A_K) = k$, a więc

$k \leq m$, gdyż $r(A) = m$. Dołączmy kolumny A_{k+1}, \dots, A_m (po ewentualnym ich przenumowaniu) do macierzy A_K tak, aby powstała macierz $A_B = [A_1, \dots, A_m]$ była bazowa (nieosobliwa). Niech $A_N = [A_{m+1}, \dots, A_n]$. Oznaczmy $x_B = (x_1, \dots, x_k, \underbrace{0, \dots, 0}_{m-k})^\top$. Mamy więc $x = (x_B^\top, 0)^\top$ i widać, że x jest dopuszczalnym rozwiązaniem bazowym, gdyż

$$b = Ax = [A_B, A_N]x = A_B x_B$$

czyli $x_B = A_B^{-1}b$ oraz $x \geq 0$. Ponadto, jeśli $k < m$, to rozwiązanie to jest zdegenerowane.

(\Leftarrow) Niech A_B będzie macierzą bazową i niech $x = (x_B^\top, 0)^\top$, gdzie $x_B = A_B^{-1}b$ będzie bazowym rozwiązaniem dopuszczalnym. Przypuśćmy, że $x = \frac{1}{2}(x' + x'')$ dla pewnych rozwiązań dopuszczalnych $x' = (x_B'^\top, x_N'^\top)^\top$ i $x'' = (x_B''^\top, x_N''^\top)^\top$. Mamy więc

$$x_B = \frac{1}{2}(x_B' + x_B'')$$

oraz

$$0 = \frac{1}{2}(x_N' + x_N'').$$

W istocie $x_N' = x_N'' = 0$, bowiem jeśli suma nieujemnych wektorów jest wektorem zerowym, to wektory te są wektorami zerowymi. Ponadto

$$b = Ax' = [A_B, A_N]x' = A_B x_B',$$

czyli $x_B' = A_B^{-1}b$. Podobnie, $x_B'' = A_B^{-1}b$. Tak więc $x_B' = x_B''$. Zatem $x' = x''$. Pokazaliśmy więc, że x jest wierzchołkiem zbioru X . \square

Uwaga 2.3.9. Powyższe twierdzenie jest prawdziwe również dla postaci klasycznej zadania programowania liniowego. W celu pokazania tej wersji twierdzenia należy najpierw sprowadzić zbiór ograniczeń

$$X = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$$

do postaci standardowej

$$\bar{X} = \{w \in \mathbb{R}^{n+m} : \bar{A}w = b, w \geq 0\},$$

gdzie $\bar{A} = [A, I]$, $w = (x^\top, u^\top)^\top$, dla której to postaci twierdzenie zostało wyżej pokazane. Następnie należy zauważyć, że jeśli punkt $w = (x^\top, u^\top)^\top$ jest wierzchołkiem zbioru \bar{X} , to x jest wierzchołkiem zbioru X oraz jeśli x jest wierzchołkiem zbioru X i wektor $u \in \mathbb{R}^m$ jest taki, że $Ax + u = b$, to punkt $w = (x^\top, u^\top)^\top$ jest wierzchołkiem zbioru \bar{X} (dowody tych faktów pozostawiamy czytelnikowi jako ćwiczenie).

2.3.3. Projekcja metryczna na zbiór wypukły

Definicja 2.3.10. Niech $C \subset \mathbb{R}^n$ i niech $x \in \mathbb{R}^n$. Punkt $y = P_C(x) \in C$ nazywamy *projekcją metryczną* punktu x na zbiór C jeśli

$$\forall z \in C \quad \|x - y\| \leq \|x - z\|.$$

Wprowadzając projekcję metryczną można zdefiniować dla dowolnej normy, jednak będą nas interesować własności tej projekcji dla normy indukowanej przez iloczyn skalarny. W dalszym ciągu tego ustępu zakładamy więc, że $\|x\| = \sqrt{\langle x, x \rangle}$ dla pewnego iloczynu skalarnego $\langle \cdot, \cdot \rangle$.

Twierdzenie 2.3.11. Niech $C \subset \mathbb{R}^n$ będzie zbiorem niepustym, domkniętym i wypukłym. Wówczas dla dowolnego $x \in \mathbb{R}^n$ istnieje dokładnie jedna jego projekcja metryczna na C .

Dowód. Twierdzenie pokażemy najpierw dla $x = 0$. Niech $d = \inf\{\|y\| : y \in C\}$ i niech ciąg $(y_k) \subset C$ będzie wybrany tak, by $\|y_k\| \rightarrow d$. Dowód rozbijemy na trzy części.

a) Pokażemy, że (y_k) jest ciągiem Cauchy'ego. Niech $\varepsilon > 0$ i niech $k_0 \in \mathbb{N}$ będzie takie, że $\|y_k\|^2 \leq d^2 + \varepsilon/4$ dla $k \geq k_0$. Niech $k, l \geq k_0$. Oczywiście $\frac{1}{2}y_k + \frac{1}{2}y_l \in C$ ponieważ C jest wypukły. Stąd $\frac{1}{2}\|y_k + y_l\| \geq d$. Korzystając z tożsamości równoległoboku otrzymujemy w konsekwencji:

$$\|y_k - y_l\|^2 = 2\|y_k\|^2 + 2\|y_l\|^2 - \|y_k + y_l\|^2$$

$$\leq 2(d^2 + \varepsilon/4) + 2(d^2 + \varepsilon/4) - 4d^2 = \varepsilon,$$

tzn. (y_k) jest ciągiem Cauchy'ego.

b) Z a) wynika, że y_k zbiega do pewnego $y \in \mathbb{R}^n$, gdyż \mathbb{R}^n jest przestrzenią zupełną. Ponadto $y \in C$, ponieważ C jest domknięty. Stąd i z ciągłości normy wynika, że $\|y\| = d$. Oznacza to, że $y = P_C(0)$.

c) Pokażemy teraz, że projekcja metryczna określona jest jednoznacznie. Niech $y' \in C$ i niech $\|y'\| = d$. Z wypukłości C otrzymujemy $\frac{1}{2}y + \frac{1}{2}y' \in C$. Ponadto

$$d \leq \left\| \frac{1}{2}y + \frac{1}{2}y' \right\| \leq \frac{1}{2}\|y\| + \frac{1}{2}\|y'\| = d,$$

a więc $\|y + y'\| = 2d$. Korzystając powtórnie z tożsamości równoległoboku mamy:

$$\|y - y'\|^2 = 2\|y\|^2 + 2\|y'\|^2 - \|y + y'\|^2 = 2d^2 + 2d^2 - 4d^2 = 0,$$

czyli $y = y'$.

Ponieważ $P_C(x) = x + P_{C-x}(0)$, więc twierdzenie jest prawdziwe dla dowolnego $x \in \mathbb{R}^n$. \square

Uwaga 2.3.12. Istnienie projekcji metrycznej na zbiór $C \subset \mathbb{R}^n$ można pokazać prościej korzystając z ciągłości normy i z twierdzenia Weierstrassa. Natomiast przeprowadzony powyżej dowód wskazuje, że twierdzenie 2.3.11 jest prawdziwe również dla dowolnej przestrzeni Hilberta.

Twierdzenie 2.3.13. Niech $x \in \mathbb{R}^n$, $C \subset \mathbb{R}^n$ będzie zbiorem niepustym, domkniętym i wypukłym i niech $y \in C$. Wówczas następujące warunki są równoważne

- (i) $y = P_C(x)$,
- (ii) $\langle x - y, z - y \rangle \leq 0$ dla dowolnego $z \in C$.

Dowód

(i) \Rightarrow (ii). Niech $y = P_C(x)$ i niech $z \in C$. Ponadto, niech

$$z_\lambda = y + \lambda(z - y)$$

dla $\lambda \in (0, 1)$. Oczywiście $z_\lambda \in C$, ponieważ C jest wypukły. Z (i) i z własności iloczynu skalarnego mamy więc

$$\begin{aligned}\|x - y\|^2 &\leq \|x - z_\lambda\|^2 = \|x - y - \lambda(z - y)\|^2 \\ &= \|x - y\|^2 - 2\lambda\langle x - y, z - y \rangle + \lambda^2\|z - y\|^2.\end{aligned}$$

Skoro $\lambda > 0$, więc

$$\langle x - y, z - y \rangle \leq \frac{\lambda}{2}\|z - y\|^2,$$

a ponieważ λ jest dowolną liczbą z przedziału $(0, 1)$, więc musi zachodzić (ii).

(ii) \Rightarrow (i). Z własności iloczynu skalarnego oraz z (ii) mamy dla dowolnego $z \in C$

$$\|z - x\|^2 = \|z - y\|^2 + \|y - x\|^2 + 2\langle z - y, y - x \rangle \geq \|y - x\|^2,$$

co na mocy definicji projekcji metrycznej daje (i). \square

2.3.4. Twierdzenia o oddzielaniu i ich konsekwencje

Twierdzenie 2.3.14 (o ostrym oddzielaniu). *Niech $C \subset \mathbb{R}^n$ będzie zbiorem niepustym, domkniętym i wypukłym i niech $x \notin C$. Wówczas istnieje taki wektor $s \in \mathbb{R}^n$, że*

$$\langle s, x \rangle > \sup\{\langle s, y \rangle : y \in C\}.$$

Dowód. Niech $s = x - P_C(x)$. Wówczas z twierdzenia 2.3.13 wynika prosto, że dla każdego $y \in C$

$$\langle x - y, s \rangle \geq \|s\|^2,$$

czyli

$$\langle s, x \rangle \geq \langle s, y \rangle + \|s\|^2.$$

Zauważmy, że $s \neq 0$, bo C jest domknięty i $x \notin C$. Zatem

$$\langle s, x \rangle \geq \sup\{\langle s, y \rangle : y \in C\} + \|s\|^2 > \sup\{\langle s, y \rangle : y \in C\}.$$

\square

Wniosek 2.3.15. Niech $A, B \subset \mathbb{R}^n$ będą zbiorami wypukłymi i domkniętymi, przy czym jeden z nich jest zbiorem zwartym. Jeśli $A \cap B = \emptyset$, to istnieje taki wektor $s \in \mathbb{R}^n$, że

$$\inf\{\langle s, u \rangle : u \in A\} > \sup\{\langle s, v \rangle : v \in B\}.$$

Dowód. Niech $C = B - A = \{z \in \mathbb{R}^n : z = v - u, v \in B, u \in A\}$. Nietrudno zauważyć, że C jest zbiorem wypukłym. Ponadto C jest zbiorem domkniętym, ponieważ A i B są domknięte, a jeden z nich zwarty. Istotnie, niech $(z_n) \subset C$ i niech $z_n \rightarrow z$. Wówczas $z_n = v_n - u_n$, gdzie $u_n \in A$ i $v_n \in B$. Przypuśćmy, że zbiór A jest zwarty. Wówczas istnieje podciąg zbieżny (u_{n_k}) ciągu (u_n) . Niech $u = \lim_{k \rightarrow \infty} u_{n_k}$. Wówczas

$$v_{n_k} = z_{n_k} + u_{n_k} \rightarrow v = z + u.$$

Oczywiście $u \in A$ i $v \in B$, ponieważ A i B są zbiorami domkniętymi. Mamy więc $z = v - u \in B - A = C$. Oznacza to, że zbiór C jest domknięty. Ponieważ, $A \cap B = \emptyset$, więc $0 \notin C$. Na mocy twierdzenia o ostrym oddzieleniu istnieje taki wektor $s \in \mathbb{R}^n$, że

$$\begin{aligned} 0 &= \langle s, 0 \rangle > \sup\{\langle s, y \rangle : y \in C\} \\ &= \sup\{\langle s, v - u \rangle : u \in A, v \in B\} \\ &= \sup\{\langle s, v \rangle : v \in B\} - \inf\{\langle s, u \rangle : u \in A\}. \end{aligned}$$

□

Twierdzenie 2.3.16 (o słabym oddzielaniu). Niech $C \subset \mathbb{R}^n$ będzie zbiorem niepustym i wypukłym i niech $x \notin C$. Wówczas istnieje $s \in \mathbb{R}^n$ takie, że

$$\langle s, x \rangle \geq \sup\{\langle s, y \rangle : y \in C\}.$$

Dowód. Jeśli $x \notin \text{cl } C$, to teza wynika z twierdzenia o ostrym oddzielaniu. Niech więc $x \in \text{cl } C \setminus C$. Ponieważ ten ostatni zbiór jest zawarty w brzegu zbioru C , więc istnieje ciąg $x_k \rightarrow x$, $x_k \notin \text{cl } C$. Z twierdzenia

o ostrym oddzielaniu wynika, że istnieje ciąg $s_k \neq 0$ taki, że dla każdego $y \in C$

$$\langle s_k, x_k \rangle > \langle s_k, y \rangle$$

lub inaczej

$$\left\langle \frac{s_k}{\|s_k\|}, x_k \right\rangle > \left\langle \frac{s_k}{\|s_k\|}, y \right\rangle.$$

Ciąg $s_k/\|s_k\|$ jest ograniczony, posiada on więc podciąg zbieżny do pewnego wektora unormowanego, powiedzmy do s . Przechodząc do granicy i korzystając z ciągłości iloczynu skalarnego otrzymamy

$$\langle s, x \rangle \geq \langle s, y \rangle$$

dla każdego $y \in C$, co jest równoważne tezie twierdzenia. \square

Wniosek 2.3.17. *Niech $A, B \subset \mathbb{R}^n$ będą zbiorami wypukłymi i domkniętymi. Jeśli $A \cap B = \emptyset$, to istnieje taki wektor $s \in \mathbb{R}^n$, że*

$$\inf\{\langle s, u \rangle : u \in A\} \geq \sup\{\langle s, v \rangle : v \in B\}.$$

Dowód powyższego wniosku przeprowadza się podobnie do dowodu wniosku 2.3.15, przy czym należy w nim skorzystać z twierdzenia o słabym oddzielaniu.

Definicja 2.3.18. Ograniczony zbiór wielościenny nazywa się *wielościannem*.

Twierdzenie 2.3.19 (Minkowski). *Wielościann K jest otoczką wypukłą zbioru swoich punktów ekstremalnych:*

$$K = \text{conv ext } K.$$

Dowód indukcyjny twierdzenia Minkowskiego względem wymiaru przestrzeni oparty na twierdzeniu o oddzielaniu można znaleźć w podręcznikach do analizy wypukłej, np. w podręczniku J. B. Hiriarta-Urruty'ego i C. Lemaréchala [10, twierdzenie 2.3.4]. Podobne twierdzenie można sformułować dla zwartych i wypukłych podzbiorów dowolnej lokalnie wypukłej przestrzeni liniowo-topologicznej. Nosi ono nazwę twierdzenia Kreina–Milmana.

Uwaga 2.3.20. Prawdziwe jest również twierdzenie w pewnym sensie odwrotne do twierdzenia Minkowskiego: otoczka wypukła zbioru skończonego jest wielościanem.

Twierdzenie 2.3.21. *Jeśli zbiór rozwiązań dopuszczalnych X zadania programowania liniowego (w postaci klasycznej lub standardowej) jest niepusty, to posiada on przynajmniej jeden wierzchołek.*

Dowód. Pokażemy najpierw, że funkcja $f(x) = e^\top x$ osiąga minimum na zbiorze X , gdzie $e = (1, \dots, 1)^\top$. Niech $z \in X$. Wówczas oczywiście

$$\inf\{e^\top x : x \in X\} = \inf\{e^\top x : e^\top x \leq e^\top z, x \in X\}.$$

Ponieważ zbiór $\bar{X} = \{x \in \mathbb{R}^n : e^\top x \leq e^\top z, x \in X\}$ jest zwarty jako podzbiór domknięty zbioru zwartego $\{x \in \mathbb{R}_+^n : e^\top x \leq e^\top z\}$ i funkcja f jest ciągła, więc osiąga ona minimum na \bar{X} , równe powiedzmy α . Z postaci zbioru \bar{X} wynika, że f osiąga również minimum równe α na zbiorze X . Zbiór $X_\alpha = \{x \in X : e^\top x = \alpha\}$ jest wielościanem, więc na mocy twierdzenia Minkowskiego posiada punkt ekstremalny \bar{x} . Pokażemy, że jest on również punktem ekstremalnym zbioru X . Niech $\bar{x} = \frac{1}{2}x' + \frac{1}{2}x''$ dla $x', x'' \in X$. Mamy

$$\alpha = e^\top \bar{x} = \frac{1}{2}e^\top x' + \frac{1}{2}e^\top x'' \geq \alpha,$$

więc $e^\top x' = e^\top x'' = \alpha$. Zatem $x', x'' \in X_\alpha$. W konsekwencji $x' = x''$, bo \bar{x} jest punktem ekstremalnym zbioru X_α . \square

Uwaga 2.3.22. Liczba wierzchołków zbioru rozwiązań dopuszczalnych X zadania programowania liniowego wynosi co najwyżej:

- a) $\binom{n}{m}$ dla postaci standardowej,
- b) $\binom{n+m}{m}$ dla postaci klasycznej.

Fakt ten wynika bezpośrednio z twierdzenia 2.3.8 i z uwagi 2.2.6.

2.3.5. Funkcje wypukłe

Definicja 2.3.23. Funkcja $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ nazywa się *wypukłą*, jeśli

$$\forall_{x,y \in \mathbb{R}^n} \quad \forall_{\lambda \in [0,1]} \quad f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y).$$

Funkcja $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ nazywa się *wklęsłą*, jeśli $-f$ jest funkcją wypukłą lub inaczej

$$\forall_{x,y \in \mathbb{R}^n} \quad \forall_{\lambda \in [0,1]} \quad f((1-\lambda)x + \lambda y) \geq (1-\lambda)f(x) + \lambda f(y).$$

Lemat 2.3.24. *Funkcja wypukła $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ jest lokalnie ograniczona z góry, tzn. dla dowolnego $x \in \mathbb{R}^n$ istnieje $r > 0$ takie, że f jest ograniczona z góry na kuli $B(x, r)$.*

Dowód. Niech $x_0 = x - \frac{1}{2n}e$, $x_i = x_0 + e_i$, $i = 1, \dots, n$, gdzie $e = (1, \dots, 1)^\top$ i niech

$$\Delta = \text{conv}\{x_i : i = 0, 1, \dots, n\}.$$

Czytelnikowi pozostawiamy dowód faktu, że Δ ma niepuste wnętrze (wystarczy pokazać na przykład, że $B(x, \frac{1}{2n}) \subset \Delta$). Niech c będzie maksymalną wartością funkcji f na zbiorze wierzchołków x_0, x_1, \dots, x_n zbioru Δ . Wówczas dla $r > 0$ takiego, że $B(x, r) \subset \Delta$ i dla dowolnego $y \in B(x, r)$ mamy $y = \sum_{i=0}^n \lambda_i x_i$, dla pewnego $w = (\lambda_1, \dots, \lambda_n)^\top \in \Delta_n$, i

$$f(y) = f\left(\sum_{i=0}^n \lambda_i x_i\right) \leq \sum_{i=0}^n \lambda_i f(x_i) \leq c \sum_{i=0}^n \lambda_i = c.$$

□

Twierdzenie 2.3.25. *Funkcja wypukła $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ jest ciągła.*

Dowód. Niech $x \in \mathbb{R}^n$ i niech $h : \mathbb{R}^n \longrightarrow \mathbb{R}$, $h(z) = f(x+z) - f(x)$. Oczywiście h jest wypukła i $h(0) = 0$. Na mocy lematu 2.3.24 istnieją więc $r > 0$ i $c \in \mathbb{R}$ takie, że $h(z) \leq c$ dla $z \in B(0, r)$. Niech $z \in B(0, r)$, $z \neq 0$

i niech $y = \frac{r}{\|z\|}z$. Wówczas oczywiście $\|y\| = r$ i $z = \lambda y$ dla $\lambda = \frac{\|z\|}{r}$. Mamy więc

$$h(z) = h(\lambda y) = h(\lambda y + (1 - \lambda)0) \leq \lambda h(y) + (1 - \lambda)h(0) \leq \frac{\|z\|}{r}c.$$

Z drugiej strony

$$0 = h(0) = h\left(\frac{1}{2}z + \frac{1}{2}(-z)\right) \leq \frac{1}{2}h(z) + \frac{1}{2}h(-z),$$

czyli

$$h(z) \geq -h(-z) \geq -\frac{\|z\|}{r}c,$$

bo $-z \in B(0, r)$. Stąd $|h(z)| \leq \frac{\|z\|}{r}c$ i dla $y = x + z$ mamy

$$|f(y) - f(x)| = |h(z)| \leq \frac{c}{r}\|z\| = \frac{c}{r}\|y - x\|,$$

a więc f jest ciągła w x . □

Twierdzenie 2.3.26. *Jeśli $f : \mathbb{R}^n \rightarrow \mathbb{R}$ jest funkcją wklęsłą i K jest wielościannem, to istnieje punkt ekstremalny zbioru K , w którym f osiąga minimum na tym zbiorze.*

Dowód. Ponieważ funkcja wklęsła f jest ciągła zaś wielościann jest zbiorem zwartym, więc f osiąga minimum na zbiorze K na mocy twierdzenia Weierstrassa. Niech więc $x^* \in K$ będzie takie, że $f(x) \geq f(x^*) = f^*$ dla każdego $x \in K$. Niech $w_1, \dots, w_p \in K$ będą punktami ekstremalnymi zbioru K takimi, że $x^* = \sum_{i=1}^p \lambda_i w_i$, dla pewnego wektora $u = (\lambda_1, \dots, \lambda_p)^\top \in \Delta_p$. Punkty takie istnieją na mocy twierdzenia Minkowskiego. Przypuśćmy, że $f(w_i) > f^*$ dla $i = 1, \dots, p$. Wówczas na mocy wklęsłości funkcji f mamy

$$f^* = f(x^*) = f\left(\sum_{i=1}^p \lambda_i w_i\right) \geq \sum_{i=1}^p \lambda_i f(w_i) > \sum_{i=1}^p \lambda_i f^* = f^*.$$

Otrzymana sprzeczność dowodzi prawdziwości twierdzenia. □

Wniosek 2.3.27. Funkcja liniowa określona na wielościanie osiąga swoje kresy w wierzchołkach tego wielościanu.

Uwaga 2.3.28. Twierdzenie 2.3.26 pozostanie prawdziwe, jeśli założyć, że K jest zbiorem zwartym. Dowód jest podobny do dowodu twierdzenia 2.3.26, przy czym zamiast z twierdzenia Minkowskiego należy skorzystać z twierdzenia Kreina–Milmana.

2.3.6. Stożki

Definicja 2.3.29. Zbiór $C \subset \mathbb{R}^n$ nazywa się *stożkiem*, jeśli

- 1^o $x, y \in C \Rightarrow x + y \in C$,
- 2^o $x \in C, \alpha \geq 0 \Rightarrow \alpha x \in C$.

Uwaga 2.3.30. Stożek jest zbiorem wypukłym

Definicja 2.3.31. Niech $C \subset \mathbb{R}^n$ będzie stożkiem. Podzbiór

$$C^* = \{y \in \mathbb{R}^n : \forall x \in C \quad x^\top y \leq 0\}$$

nazywa się *stożkiem sprzężonym* z C .

Przykład 2.3.32. Następujące zbiory są stożkami:

- a) dowolna podprzestrzeń liniowa, w szczególności $\{x \in \mathbb{R}^n : Ax = 0\}$;
- b) $\{x \in \mathbb{R}^n : Ax = 0, x \geq 0\}$;
- c) $\{x \in \mathbb{R}^n : Ax \leq 0\}$;
- d) stożek sprzężony C^* z danym stożkiem C ;
- e) $\text{cone } S = \{x \in \mathbb{R}^n : x = \sum_{i=1}^p \alpha_i s_i, \alpha_i \geq 0, s_i \in S, i = 1, \dots, p, p \in \mathbb{N}\}$
(zbiór ten nazywa się *stożkiem generowanym przez podzbiór $S \subset \mathbb{R}^n$*);
w szczególności, jeśli za S przyjąć kolumny macierzy A , to otrzymany stożek ma postać

$$C = \{y \in \mathbb{R}^n : y = Au, u \geq 0\};$$

f) $\mathbb{R}_+K = \{\alpha x : \alpha \geq 0, x \in K\}$ dla wypukłego podzbioru $K \subset \mathbb{R}^n$.
 Ponadto $\mathbb{R}_+K = \text{cone } K$.

Lemat 2.3.33. Niech $C \subset \mathbb{R}^n$ będzie stożkiem domkniętym i niech $A \subset \mathbb{R}^n$ będzie zbiorem zwartym i wypukłym. Jeśli $A \cap C = \emptyset$, to istnieje taki wektor $w \in \mathbb{R}^n$, że $w^\top v \leq 0$ dla każdego $v \in C$ i $w^\top u > 0$ dla każdego $u \in A$.

Dowód. W przypadku $C = \{0\}$ wystarczy wziąć $w = P_A 0$. Niech więc $C \neq \{0\}$. Z wniosku 2.3.15 wynika, że istnieje wektor w taki, że $\alpha = \inf_{u \in A} w^\top u > \sup_{v \in C} w^\top v = \gamma$. Pozostaje więc pokazać, że $\gamma = 0$. Przypuśćmy, że $\gamma > 0$. Niech wektor $z \in C$ będzie taki, że $w^\top z > 0$. Wówczas $\beta z \in C$ dla każdego $\beta \geq 0$, gdyż C jest stożkiem. Otrzymamy wówczas

$$+\infty = \lim_{\beta \rightarrow +\infty} w^\top \beta z \leq \sup_{v \in C} w^\top v = \gamma,$$

czyli $\gamma = +\infty$. Równość ta stoi w sprzeczności z nierównością $\alpha > \gamma$. Przypuśćmy więc, że $\gamma < 0$. Otrzymamy wówczas

$$0 > \gamma = \sup_{v \in C} w^\top v \geq \lim_{\beta \rightarrow 0^+} w^\top \beta z = 0$$

dla pewnego $z \in C$. Otrzymana sprzeczność dowodzi, że $\gamma = 0$. □

Przykład 2.3.34. a) Stożkiem sprzężonym ze stożkiem

$$K = \{x \in \mathbb{R}^n : Ax \leq 0\}$$

jest stożek

$$K' = \{y \in \mathbb{R}^n : y = A^\top w, w \geq 0\}.$$

Istotnie, niech $y \in K'$, czyli $y = A^\top w$ dla pewnego $w \geq 0$. Wówczas dla dowolnego $x \in K$ mamy

$$y^\top x = (A^\top w)^\top x = w^\top Ax \leq 0.$$

W rezultacie $K' \subset K^*$. Pokażemy teraz, że $K^* \subset K'$. Przypuśćmy, że jest odwrotnie, tzn., że istnieje taki wektor $y \in K^*$, że $y \notin K'$. Ponieważ K' jest

stożkiem domkniętym i wypukłym, więc na mocy twierdzenia o oddzielaniu istnieje $u \in \mathbb{R}^n$ taki, że

$$u^\top z \leq 0 \text{ dla każdego } z \in K' \text{ i } u^\top y > 0$$

(patrz lemat 2.3.33). Stąd dla każdego $w \geq 0$ mamy

$$(Au)^\top w = u^\top A^\top w \leq 0$$

To z kolei oznacza, że $Au \leq 0$ (gdyby było inaczej, to $(Au)^\top w > 0$ dla odpowiednio dobranego $w \geq 0$) lub, co na jedno wychodzi $u \in K$. Mamy więc $u^\top y \leq 0$, bo $y \in K^*$. Uzyskana sprzeczność dowodzi, że $K^* \subset K'$.

b) Stożkiem sprzężonym ze stożkiem

$$K' = \{y \in \mathbb{R}^n : y = A^\top w, w \geq 0\}$$

jest stożek

$$K = \{x \in \mathbb{R}^n : Ax \leq 0\}.$$

Istotnie, inkluzję $K \subset (K')^*$ dowodzi się tak samo, jak w przykładzie a). Niech teraz $x \in (K')^*$. Wówczas dla dowolnego $w \geq 0$

$$(Ax)^\top w = x^\top (A^\top w) \leq 0.$$

Stąd już prosto wynika, że $Ax \leq 0$, a więc $x \in K$.

c) Stożkiem sprzężonym ze stożkiem

$$\{x \in \mathbb{R}^n : Ax \leq 0, x \geq 0\}$$

jest stożek

$$\{y \in \mathbb{R}^n : y \leq A^\top w, w \geq 0\}.$$

d) Stożkiem sprzężonym ze stożkiem

$$\{x \in \mathbb{R}^n : Ax = 0, x \geq 0\}$$

jest stożek

$$\{y \in \mathbb{R}^n : y \leq A^\top w, w \in \mathbb{R}^m\}.$$

1.

Dowody faktów c) i d) pozostawiamy czytelnikowi.

2.3.7. Lemat Farkasa

Konsekwencją własności przedstawionej w przykładzie 2.3.34.a), a właściwie jej równoważną postacią jest następujące, bardzo ważne twierdzenie, z którego będziemy korzystać w dalszej części.

Twierdzenie 2.3.35 (lemat Farkasa). *Niech $c \in \mathbb{R}^n$ i niech A będzie macierzą typu $m \times n$. Wówczas zbiór*

$$S = \{x \in \mathbb{R}^n : c^\top x < 0, Ax \leq 0\}$$

jest pusty wtedy i tylko wtedy, gdy istnieje wektor $y \in \mathbb{R}_+^m$ taki, że

$$c + A^\top y = 0.$$

Dowód. Zauważmy, że $S = \emptyset$ wtedy i tylko wtedy gdy

$$\{x \in \mathbb{R}^n : Ax \leq 0\} \subset \{x \in \mathbb{R}^n : c^\top x \geq 0\}.$$

Oznaczając $K = \{x \in \mathbb{R}^n : Ax \leq 0\}$ inkluzję tę możemy natomiast zapisać następująco:

$$\forall_{x \in K} \quad (-c)^\top x \leq 0,$$

co oznacza, że $-c \in K^*$. Zgodnie z własnością przedstawioną w przykładzie 2.3.34.a) fakt ten jest równoważny temu, że $-c = A^\top y$ dla pewnego wektora $y \geq 0$. \square

Uwaga 2.3.36. Lemat Farkasa można znaleźć w literaturze w kilku równoważnych postaciach. Podamy jeszcze jedną z nich, będącą wnioskiem z twierdzenia 2.3.35, pozostawiając jej dowód czytelnikowi.

Lemat 2.3.37 (Farkas). *Niech $c \in \mathbb{R}^n$ i niech A będzie macierzą typu $m \times n$. Wówczas dokładnie jedno z następujących stwierdzeń jest prawdziwe: albo*

(a) *układ równań $A^\top y = c$ posiada nieujemne rozwiązanie $y \in \mathbb{R}^m$*

albo

(b) *układ nierówności $Ax \geq 0, c^\top x < 0$ posiada rozwiązanie $x \in \mathbb{R}^n$.*

2.3.8. Dalsze konsekwencje twierdzeń o oddzielaniu

Niech $X = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\} \neq \emptyset$ będzie zbiorem rozwiązań dopuszczalnych ZPL w postaci standardowej. Oznaczmy

$$W = \text{ext } X = \{w_1, \dots, w_p\},$$

$$K = \text{conv } W$$

oraz

$$C = \{x \in \mathbb{R}^n : Ax = 0, x \geq 0\}.$$

Twierdzenie 2.3.38. *Zachodzi równość*

$$X = K + C,$$

tzn. zbiór rozwiązań dopuszczalnych X jest sumą kompleksową wielościanu K (będącego otoczką wypukłą wierzchołków zbioru X) i stożka C .

Dowód

„ \supset ” Niech $y \in K$ i $z \in C$. Mamy wówczas $Ay = b$, $Az = 0$ oraz $y, z \geq 0$. Zatem dla $x = y + z$ zachodzą równości

$$Ax = A(y + z) = Ay + Az = b$$

oraz nierówność $x \geq 0$. Oznacza to, że $x \in X$.

„ \subset ” Niech $x \in X$. Pokażemy, że $x = y + z$ przy pewnych $y \in K$ i $z \in C$. Jeśli $x \in K$, to wystarczy przyjąć $z = 0$. Niech więc $x \notin K$. Pokażemy najpierw, że $(x - K) \cap C \neq \emptyset$. Przypuśćmy, że jest przeciwnie. Ponieważ zbiór $x - K$ jest zwarty i wypukły, więc na mocy lematu 2.3.33, istnieje $w \in \mathbb{R}^n$ taki, że

$$\forall_{s \in x - K} \forall_{z \in C} \quad w^\top s > 0 \text{ i } w^\top z \leq 0. \quad (2.7)$$

Stąd otrzymujemy w szczególności, że

$$w^\top x > w^\top y$$

dla każdego $y \in K$. Ponieważ funkcja liniowa określona na wielościanie osiąga maksimum w punktach ekstremalnych tego wielościanu (wniosek 2.3.27), więc

$$w^\top x > w^\top y^* \quad (2.8)$$

dla pewnego wierzchołka y^* wielościanu K . Na mocy twierdzenia 2.3.8, y^* jest rozwiązaniem bazowym układu $Ax = b$, a więc istnieje macierz bazowa A_B taka, że $y^* = \begin{bmatrix} A_B^{-1}b \\ 0 \end{bmatrix}$. Wynika stąd, że

$$w^\top y^* = w_B^\top A_B^{-1}b, \quad (2.9)$$

gdzie $w = \begin{bmatrix} w_B \\ w_N \end{bmatrix}$. Postępując podobnie jak przy wyprowadzaniu równości (2.4) oraz korzystając z nierówności (2.8) i równości (2.9) otrzymamy

$$w^\top x = w_B^\top A_B^{-1}b + (w_N^\top - w_B^\top A_B^{-1}A_N)x_N > w_B^\top A_B^{-1}b,$$

gdzie $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$. Zatem

$$(w_N^\top - w_B^\top A_B^{-1}A_N)x_N > 0. \quad (2.10)$$

Zauważmy, że $x_N \geq 0$, gdyż $x \in X$. Z drugiej strony, dla $z = \begin{bmatrix} z_B \\ z_N \end{bmatrix} \in C$, mamy $z_B = -A_B^{-1}A_N z_N$. Zgodnie z (2.7) otrzymujemy więc, że

$$0 \geq w^\top z = w_B^\top z_B + w_N^\top z_N = (w_N^\top - w_B^\top A_B^{-1}A_N)z_N.$$

Ponieważ $z \in C$ jest dowolny, więc

$$(w_N^\top - w_B^\top A_B^{-1}A_N)z_N \leq 0$$

dla dowolnego $z_N \geq 0$. Uzyskaliśmy sprzeczność z nierównością (2.10), zatem $(x - K) \cap C \neq \emptyset$. Oznacza to, że istnieje $y \in K$ takie, że $x - y = z \in C$. Otrzymaliśmy więc, że $x = y + z$ dla pewnych $y \in K$ i $z \in C$. \square

Uwaga 2.3.39. Podobne twierdzenie zachodzi dla zbioru rozwiązań dopuszczalnych ZPL w postaci klasycznej. W tym przypadku zbiór C ma postać

$$C = \{x \in \mathbb{R}^n : Ax \leq 0, x \geq 0\}.$$

W dowodzie tej wersji należy skorzystać z uwagi 2.3.9.

Definicja 2.3.40. Niesprzeczne zadanie programowania liniowego nazywa się *ograniczonym*, jeśli funkcja celu $c^\top x$ jest ograniczona na zbiorze rozwiązań dopuszczalnych X z góry – w przypadku maksymalizacji, bądź z dołu – w przypadku minimalizacji.

Twierdzenie 2.3.41. *Funkcja liniowa $z = c^\top x$ jest ograniczona z góry na zbiorze*

$$X = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$$

wtedy i tylko wtedy, gdy istnieje wektor $u \in \mathbb{R}^m$ taki, że $c \leq A^\top u$. Co więcej, jeśli zachodzi którykolwiek z tych warunków, to funkcja ta osiąga maksimum na X . Maksimum to jest osiągnięte w jednym z wierzchołków zbioru X .

Dowód. Niech K i C będą określone tak, jak w twierdzeniu 2.3.38. Niech $W = \text{ext } X$. Oczywiście zbiór W jest niepusty na mocy wniosku 2.3.21. Zachodzi następująca równość

$$\sup_{x \in X} c^\top x = \max_{y \in K} c^\top y + \sup_{z \in C} c^\top z.$$

Niech $c \leq A^\top u$, gdzie $u \in \mathbb{R}^m$. Wtedy $c^\top z \leq u^\top Az = 0$ dla $z \in C$. W konsekwencji,

$$\sup_{x \in X} c^\top x = \max_{y \in K} c^\top y = \max_{w \in W} c^\top w.$$

Założmy teraz, że $\sup_{x \in X} c^\top x < +\infty$. Wówczas musi być $c^\top z \leq 0$ dla $z \in C$. W przeciwnym bowiem przypadku $c^\top(tz) \rightarrow +\infty$ dla $t \rightarrow +\infty$, przy czym $tz \in C$, bo C jest stożkiem. Otrzymaliśmy więc, że

$$c \in C^* = \{z \in \mathbb{R}^n : z \leq A^\top u, u \in \mathbb{R}^m\}$$

(por. przykład 2.3.34.d). □

Uwaga 2.3.42. Podobne twierdzenie zachodzi dla postaci klasycznej ZPL. W tym przypadku warunkiem koniecznym i wystarczającym ograniczoności funkcji $z = c^\top x$ na zbiorze

$$X = \{x \in \mathbb{R}^n : Ax \leq b\}$$

jest zajęcie równości $c = A^\top u$, dla pewnego $u \geq 0$. Dowód tego faktu pozostawiamy czytelnikowi.

Uwaga 2.3.43. Z ostatniego twierdzenia wynika, że przy poszukiwaniu rozwiązania optymalnego zadania programowania liniowego (o ile rozwiązanie takie istnieje) wystarczy ograniczyć się do rozwiązań bazowych (wierzchołków zbioru rozwiązań dopuszczalnych X). Co więcej, okazuje się, że nie musimy badać wszystkich rozwiązań bazowych. W następnym rozdziale poznamy metodę rozwiązywania ZPL polegającą na „wędrowce” z danego wierzchołka zbioru X wzdłuż odpowiedniej krawędzi do sąsiedniego wierzchołka tak, aby funkcja celu rosła. Procedurę tę powtarza się tak długo, aż „dojdziemy” do wierzchołka, w którym funkcja celu osiąga swoje maksimum.

ROZDZIAŁ 3

Metoda sympleksowa

*Matematyka nie może wypełnić życia, ale jej nieznanomość już
niejednemu je wypełniła.*

[H. Steinhaus]

3.1. Tablica sympleksowa

Rozważmy ZPL w postaci klasycznej

$$\begin{array}{ll} \text{maksymalizować} & z = c^\top x \\ \text{przy ograniczeniach} & Ax \leq b \\ & x \geq 0. \end{array} \quad (3.1)$$

Po wprowadzeniu zmiennych uzupełniających otrzymamy równoważny mu problem w postaci kanonicznej:

$$\begin{array}{ll} \text{maksymalizować} & z = c^\top x \\ \text{przy ograniczeniach} & Ax + Iu = b \\ & x, u \geq 0. \end{array} \quad (3.2)$$

Z kolei ten ostatni problem jest równoważny następującemu: Wśród wszystkich rozwiązań układu $m + 1$ równań z $n + m + 1$ niewiadomymi x_1, \dots, x_n ,

$u_1, \dots, u_m, z :$

$$\begin{array}{rclcl} z & +0^\top u & -c^\top x & = & 0 \\ 0z & +Iu & +Ax & = & b, \end{array} \quad (3.3)$$

przy czym $x_1, \dots, x_n, u_1, \dots, u_m \geq 0$, znaleźć rozwiązanie dla którego zmienna z przyjmuje największą wartość.

Ostatni układ równań można przedstawić w postaci tzw. *tablicy sympleksowej*:

a) w postaci *długiej*

$$\begin{array}{c|ccccccc|c} z & u_1 & \dots & u_m & x_1 & \dots & x_n & \\ \hline 1 & 0 & \dots & 0 & -c_1 & \dots & -c_n & = 0 \\ \hline 0 & 1 & \dots & 0 & a_{11} & \dots & a_{1n} & = b_1 \\ \hline \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \hline 0 & 0 & & 1 & a_{m1} & & a_{mn} & = b_m \\ \hline \end{array} \quad (3.4)$$

lub

b) w postaci *krótkiej*

$$\begin{array}{c|cccc|c} 1 & -x_1 & \dots & -x_n & \\ \hline 0 & -c_1 & \dots & -c_n & = z \\ \hline b_1 & a_{11} & \dots & a_{1n} & = u_1 \\ \hline \dots & \dots & \dots & \dots & \dots \\ \hline b_m & a_{m1} & \dots & a_{mn} & = u_m \\ \hline \end{array} . \quad (3.5)$$

Ostatnia tablica odpowiada równoważnej postaci układu (3.3):

$$\begin{array}{rcl} c^\top x & = & z \\ b - Ax & = & u. \end{array}$$

Przykład 3.1.1. Następujące ZPL

$$\begin{array}{rcl}
 \text{maksymalizować} & x_1 + 2x_2 & \\
 \text{przy ograniczeniach} & x_1 + x_2 \leq 100 & \\
 & 6x_1 + 9x_2 \leq 720 & \\
 & x_2 \leq 60 & \\
 & x_1, x_2 \geq 0 &
 \end{array}$$

przedstawimy w postaci tablicy sympleksowej. Zadanie to ma następującą postać kanoniczną:

$$\begin{array}{rcl}
 \text{maksymalizować} & z = & x_1 + 2x_2 \\
 \text{przy ograniczeniach} & x_1 + x_2 + u_1 & = 100 \\
 & 6x_1 + 9x_2 + u_2 & = 720 \\
 & x_2 + u_3 & = 60 \\
 & x_1, x_2, u_1, u_2, u_3 & \geq 0.
 \end{array}$$

Odpowiadające jej tablice sympleksowe wyglądają następująco:

a) postać długa

| z | u_1 | u_2 | u_3 | x_1 | x_2 | |
|-----|-------|-------|-------|-------|-------|-------|
| 1 | 0 | 0 | 0 | -1 | -2 | = 0 |
| 0 | 1 | 0 | 0 | 1 | 1 | = 100 |
| 0 | 0 | 1 | 0 | 6 | 9 | = 720 |
| 0 | 0 | 0 | 1 | 0 | 1 | = 60 |

b) postać krótka

| 1 | $-x_1$ | $-x_2$ | |
|-----|--------|--------|---------|
| 0 | -1 | -2 | = z |
| 100 | 1 | 1 | = u_1 |
| 720 | 6 | 9 | = u_2 |
| 60 | 0 | 1 | = u_3 |

Rozważmy teraz ZPL w postaci standardowej (przypominamy o założeniu $r(A) = m < n$)

$$\begin{array}{ll} \text{maksymalizować} & z = c^\top x \\ \text{przy ograniczeniach} & Ax = b \\ & x \geq 0. \end{array}$$

Zadanie to możemy sprowadzić do postaci kanonicznej wyznaczając z układu równań $Ax = b$ zmienne bazowe x_B . Można to zrobić stosując na przykład metodę eliminacji Gaussa. Po przekształceniach otrzymamy (w poniższym zapisie przyjmujemy dla uproszczenia, że $x_B = (x_1, \dots, x_m)^\top$ oraz $x_N = (x_{m+1}, \dots, x_n)^\top$):

$$\begin{array}{ll} \text{maksymalizować} & z = \tilde{c}_N^\top x_N + d \\ \text{przy ograniczeniach} & Ix_B + \tilde{A}x_N = \tilde{b} \\ & x_B, x_N \geq 0, \end{array}$$

przy czym

$$A = \begin{bmatrix} A_B & A_N \end{bmatrix},$$

$$c = \begin{bmatrix} c_B \\ c_N \end{bmatrix},$$

$$\tilde{A} = A_B^{-1} A_N,$$

$$\tilde{c}_N^\top = c_N^\top - c_B^\top A_B^{-1} A_N,$$

$$\tilde{b} = A_B^{-1} b$$

i

$$d = c_B^\top \tilde{b}.$$

Ostatni problem jest równoważny następującemu: wśród wszystkich rozwiązań układu $m + 1$ równań z $n + 1$ niewiadomymi x_1, \dots, x_n, z :

$$\begin{array}{llll} z & +0x_B & -\tilde{c}_N^\top x_N & = d \\ 0z & +Ix_B & +\tilde{A}x_N & = \tilde{b}, \end{array} \quad (3.6)$$

przy czym $x_B, x_N \geq 0$, znaleźć rozwiązanie, dla którego zmienna z przyjmuje największą wartość. Odpowiadające temu układowi równań tablice sympleksowe mają:

a) postać długą

| z | x_1 | ... | x_m | x_{m+1} | ... | x_n | |
|-----|-------|-----|-------|---------------------|-----|------------------|-----------------|
| 1 | 0 | ... | 0 | $-\tilde{c}_{m+1}$ | ... | $-\tilde{c}_n$ | $= d$ |
| 0 | 1 | ... | 0 | $\tilde{a}_{1,m+1}$ | ... | \tilde{a}_{1n} | $= \tilde{b}_1$ |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 0 | 0 | | 1 | $\tilde{a}_{m,m+1}$ | | \tilde{a}_{mn} | $= \tilde{b}_m$ |

b) postać krótką

| 1 | $-x_{m+1}$ | ... | $-x_n$ | |
|---------------|---------------------|-----|------------------|---------|
| d | $-\tilde{c}_{m+1}$ | ... | $-\tilde{c}_n$ | $= z$ |
| \tilde{b}_1 | $\tilde{a}_{1,m+1}$ | ... | \tilde{a}_{1n} | $= x_1$ |
| ... | ... | ... | ... | ... |
| \tilde{b}_m | $\tilde{a}_{m,m+1}$ | ... | \tilde{a}_{mn} | $= x_m$ |

Uwaga 3.1.2. Stała d występująca w ostatnich dwóch tablicach nie ma wpływu na rozwiązanie ZPL. Jeśli wziąć $d = 0$, to zbiór rozwiązań optymalnych nie zmieni się, natomiast optymalna wartość funkcji celu zmniejszy się o d . Widzimy więc, że otrzymana wyżej postać (3.6) jest w istocie postacią kanoniczną ZPL.

Przykład 3.1.3. Przedstawimy w postaci tablicy sympleksowej następujące ZPL

$$\begin{array}{ll}
 \text{maksymalizować} & z = x_1 + 2x_2 + 3x_3 \\
 \text{przy ograniczeniach} & x_1 + 2x_2 + x_3 = 4 \\
 & 2x_1 + x_2 + 5x_3 = 5 \\
 & x_1, x_2, x_3 \geq 0.
 \end{array}$$

Po zastosowaniu do ograniczeń metody eliminacji Gaussa otrzymamy następującą postać ZPL:

$$\begin{array}{rcll} \text{maksymalizować} & z = & x_1 & +2x_2 & 3x_3 \\ \text{przy ograniczeniach} & & x_1 & & +3x_3 & = 2 \\ & & & x_2 & -x_3 & = 1 \\ & & x_1, & x_2, & x_3 & \geq 0. \end{array}$$

Jeśli teraz wstawimy $x_1 = -3x_3 + 2$ i $x_2 = x_3 + 1$ do funkcji celu, otrzymamy następującą postać kanoniczną:

$$\begin{array}{rcll} \text{maksymalizować} & z = & & 2x_3 & + & 4 \\ \text{przy ograniczeniach} & & x_1 & & +3x_3 & = 2 \\ & & & x_2 & -x_3 & = 1 \\ & & x_1, & x_2, & x_3 & \geq 0 \end{array}$$

$$(x_B = (x_1, x_2)^\top, x_N = x_3, \tilde{A} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \tilde{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \tilde{c}_N = 2, d = 4).$$

Odpowiadające jej tablice sympleksowe mają postaci:

a) długą

$$\begin{array}{c|c|c|c|c|c} \hline \hline z & x_1 & x_2 & x_3 & & \\ \hline 1 & 0 & 0 & -2 & & = 4 \\ \hline 0 & 1 & 0 & 3 & & = 2 \\ \hline 0 & 0 & 1 & -1 & & = 1 \\ \hline \hline \end{array} ,$$

b) krótką

$$\begin{array}{c|c|c|c} \hline \hline 1 & -x_3 & & \\ \hline 4 & -2 & & = z \\ \hline 2 & 3 & & = x_1 \\ \hline 1 & -1 & & = x_2 \\ \hline \hline \end{array} .$$

W dalszej części będziemy rozważać ZPL w postaci kanonicznej i związaną z nią postać krótką tablicy sympleksowej (3.5) (zgodnie z uwagą

3.1.2, ZPL w postaci standardowej daje się sprowadzić do postaci kanonicznej i związanej z nią krótkiej formy tablicy sympleksowej (3.5)). Jak wcześniej powiedzieliśmy, postać tę można traktować jako układ $m + 1$ równań z $n + m + 1$ niewiadomymi. Tablica (3.5) jest tablicą początkową w metodzie sympleksowej. Przedstawia ona układ równań, w którym zmienne u_1, u_2, \dots, u_m (bazowe) i z przedstawione są w zależności od pozostałych zmiennych x_1, x_2, \dots, x_n (niebazowych). Ten układ równań można przedstawić także w sposób równoważny wyznaczając z niego inne m zmiennych bazowych (oznaczymy je s_1, \dots, s_m) i z w zależności od pozostałych n zmiennych niebazowych (oznaczymy je r_1, \dots, r_n). To równoważne przedstawienie zapiszemy za pomocą następującej tablicy:

$$\begin{array}{c|cccc|c}
 & 1 & -r_1 & \dots & -r_n & \\
 \hline
 & \alpha_{00} & \alpha_{01} & \dots & \alpha_{0n} & = z \\
 & \alpha_{10} & \alpha_{11} & \dots & \alpha_{1n} & = s_1 \\
 & \dots & \dots & \dots & \dots & \dots \\
 & \alpha_{m0} & \alpha_{m1} & \dots & \alpha_{mn} & = s_m
 \end{array} \quad (3.7)$$

Tablica ta nazywa się *tablicą sympleksową*. W sumie istnieje co najwyżej $\binom{n+m}{m}$ możliwości przedstawień m zmiennych bazowych (i dodatkowo z) w zależności od pozostałych n zmiennych (co najwyżej, ponieważ nie każdemu wyborowi m zmiennych x_B odpowiada nieosobliwa podmacierz A_B). Każda tablica sympleksowa przedstawia rozwiązanie bazowe układu równań (3.3) (zmienne niebazowe równają się zeru). To rozwiązanie bazowe można odczytać w zerowej kolumnie tablicy sympleksowej. Jeśli rozwiązanie to jest dopuszczalne ($\alpha_{i0} \geq 0$, $i = 1, \dots, m$), to odpowiada mu wierzchołek zbioru X rozwiązań dopuszczalnych. (Uwaga: jednemu wierzchołkowi może odpowiadać więcej tablic sympleksowych – jest to jednak możliwe wyłącznie w przypadku rozwiązań zdegenerowanych; przypadek ten omówimy dokładniej w jednym z kolejnych punktów.). Wśród wszystkich wierzchołków (bazowych rozwiązań dopuszczalnych) znajduje się rozwiązanie optymalne, o ile zbiór X^* wszystkich rozwiązań optymalnych jest niepusty (patrz: twierdzenie 2.3.41 i uwaga 2.3.43). W ogólnym przypadku nie ma jednak

potrzeby przeszukiwania wszystkich wierzchołków (bazowych rozwiązań dopuszczalnych) w celu znalezienia rozwiązania optymalnego. Metoda sympleksowa (zamiennie będziemy używać nazwy *algorytm sympleksowy*) podaje sposób na systematyczne przeszukiwanie wierzchołków.

3.2. Opis metody sympleksowej

W metodzie sympleksowej przeprowadzane są następujące czynności:

- a) sprowadzenie zadania do tablicy sympleksowej (3.5) – odpowiada ona początkowi układu współrzędnych,
- b) znalezienie dopuszczalnego rozwiązania bazowego – wierzchołka zbioru X (jest to tzw. *I faza* metody sympleksowej),
- c) wędrowanie od wierzchołka do wierzchołka zbioru X po krawędziach, wzdłuż których funkcja celu rośnie (a przynajmniej nie maleje), tak długo, aż osiągnięty zostanie wierzchołek, w którym funkcja celu osiąga maksimum na X (jest to tzw. *II faza* metody sympleksowej).

Czynności te zostaną opisane w następnych punktach.

Uwaga 3.2.1. Przedstawimy teraz, jakie podstawowe informacje można odczytać z tablicy sympleksowej.

- a) Rozwiązanie bazowe można odczytać w zerowej kolumnie tablicy sympleksowej (3.7) (wektor $(\alpha_{10}, \dots, \alpha_{m0})^\top$). Pierwszy element α_{00} tej kolumny zawiera wartość funkcji celu, która odpowiada temu rozwiązaniu bazowemu. Rozwiązanie to jest więc dopuszczalne dokładnie wtedy, gdy wektor $(\alpha_{10}, \dots, \alpha_{m0})^\top$ jest nieujemny.
- b) Dopuszczalne, niezdegenerowane rozwiązanie bazowe przedstawione tablicą sympleksową (3.7) jest optymalne wtedy i tylko wtedy, gdy wszystkie elementy zerowego wiersza, poza zerowym elementem tej tablicy są nieujemne, czyli $\alpha_{01}, \dots, \alpha_{0n} \geq 0$. W tym przypadku bowiem

zwiększenie jakiegokolwiek zmiennej niebazowej nie spowoduje zwiększenia wartości funkcji celu. Dla takiej tablicy optymalna wartość funkcji celu znajduje się w zerowym wierszu i zerowej kolumnie tablicy (element α_{00}).

- c) Może się zdarzyć, że tablica sympleksowa (3.7) przedstawia rozwiązanie optymalne, natomiast niektóre elementy α_{0j} , $j = 1, \dots, n$, znajdujące się w zerowym wierszu tej tablicy są ujemne. Może to jednak mieć miejsce wyłącznie w przypadku degeneracji. Sytuacja ta będzie omówiona później.

3.2.1. Wymiana zmiennej bazowej (piwotyzacja)

Założmy, że w układzie równań

$$\begin{array}{ccccccc} \alpha_{00} & -\alpha_{01}r_1 & -\dots & -\alpha_{0j}r_j & -\dots & -\alpha_{0n}r_n & = z \\ \alpha_{10} & -\alpha_{11}r_1 & -\dots & -\alpha_{1j}r_j & -\dots & -\alpha_{1n}r_n & = s_1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_{i0} & -\alpha_{i1}r_1 & -\dots & -\alpha_{ij}r_j & -\dots & -\alpha_{in}r_n & = s_i \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_{m0} & -\alpha_{m1}r_1 & -\dots & -\alpha_{mj}r_j & -\dots & -\alpha_{mn}r_n & = s_m \end{array}$$

z niewiadomymi s_1, \dots, s_m (zmiennne bazowe), r_1, \dots, r_n (zmiennne niebazowe) i z , współczynnik $\alpha_{ij} \neq 0$. Wówczas możemy wykonać *wymianę zmiennej bazowej* s_i ze zmienną niebazową r_j . W ten sposób uzyskamy nowy układ zmiennych bazowych

$$s_1, \dots, s_{i-1}, r_j, s_{i+1}, \dots, s_m,$$

który możemy wyrazić za pomocą pozostałych niewiadomych

$$r_1, \dots, r_{j-1}, s_i, r_{j+1}, \dots, r_n$$

(nowych zmiennych niebazowych). W tym przypadku otrzymamy

$$\frac{\alpha_{i0}}{\alpha_{ij}} - \frac{\alpha_{i1}}{\alpha_{ij}}r_1 - \dots - \frac{\alpha_{i,j-1}}{\alpha_{ij}}r_{j-1} - \frac{1}{\alpha_{ij}}s_i - \frac{\alpha_{i,j+1}}{\alpha_{ij}}r_{j+1} - \dots - \frac{\alpha_{in}}{\alpha_{ij}}r_n = r_j$$

i po wstawieniu do pozostałych równań otrzymamy

$$\begin{aligned} & (\alpha_{k0} - \frac{\alpha_{i0}\alpha_{kj}}{\alpha_{ij}}) - (\alpha_{k1} - \frac{\alpha_{i1}\alpha_{kj}}{\alpha_{ij}})r_1 - \dots - (\alpha_{k,j-1} - \frac{\alpha_{i,j-1}\alpha_{kj}}{\alpha_{ij}})r_{j-1} \\ & - \frac{\alpha_{kj}}{\alpha_{ij}}s_i - (\alpha_{k,j+1} - \frac{\alpha_{i,j+1}\alpha_{kj}}{\alpha_{ij}})r_{j+1} - \dots - (\alpha_{kn} - \frac{\alpha_{in}\alpha_{kj}}{\alpha_{ij}})r_n = s_k, \end{aligned}$$

$k = 0, 1, \dots, m$, $k \neq i$ (oznaczono tu zmienną z jako s_0).

Przy tej wymianie bazy zamieniono tylko dwie zmienne s_i i r_j , tzn. stara i nowa baza różnią się dokładnie na jednej pozycji. Te dwie bazy odpowiadają więc dwóm wierzchołkom zbioru X , które połączone są wspólną krawędzią.

Po wymianie bazy otrzymaliśmy więc nowy układ równań równoważny staremu. Odpowiadająca mu tablica sympleksowa ma następującą krótką postać:

| 1 | $-r_1$ | ... | $-r_{j-1}$ | $-s_i$ | $-r_{j+1}$ | ... | $-r_n$ | |
|------------------------|------------------------|-----|--------------------------|------------------------|--------------------------|-----|------------------------|-------------|
| $\bar{\alpha}_{00}$ | $\bar{\alpha}_{01}$ | ... | $\bar{\alpha}_{0,j-1}$ | $\bar{\alpha}_{0j}$ | $\bar{\alpha}_{0,j+1}$ | ... | $\bar{\alpha}_{0n}$ | $= z$ |
| $\bar{\alpha}_{10}$ | $\bar{\alpha}_{11}$ | ... | $\bar{\alpha}_{1,j-1}$ | $\bar{\alpha}_{1j}$ | $\bar{\alpha}_{1,j+1}$ | ... | $\bar{\alpha}_{1n}$ | $= s_1$ |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| $\bar{\alpha}_{i-1,0}$ | $\bar{\alpha}_{i-1,1}$ | ... | $\bar{\alpha}_{i-1,j-1}$ | $\bar{\alpha}_{i-1,j}$ | $\bar{\alpha}_{i-1,j+1}$ | ... | $\bar{\alpha}_{i-1,n}$ | $= s_{i-1}$ |
| $\bar{\alpha}_{i0}$ | $\bar{\alpha}_{i1}$ | ... | $\bar{\alpha}_{i,j-1}$ | $\bar{\alpha}_{ij}$ | $\bar{\alpha}_{i,j+1}$ | ... | $\bar{\alpha}_{in}$ | $= r_j$ |
| $\bar{\alpha}_{i+1,0}$ | $\bar{\alpha}_{i+1,1}$ | ... | $\bar{\alpha}_{i+1,j-1}$ | $\bar{\alpha}_{i+1,j}$ | $\bar{\alpha}_{i+1,j+1}$ | ... | $\bar{\alpha}_{i+1,n}$ | $= s_{i+1}$ |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| $\bar{\alpha}_{m0}$ | $\bar{\alpha}_{m1}$ | ... | $\bar{\alpha}_{m,j-1}$ | $\bar{\alpha}_{mj}$ | $\bar{\alpha}_{m,j+1}$ | ... | $\bar{\alpha}_{mn}$ | $= s_m$ |

gdzie

$$\bar{\alpha}_{ij} = \frac{1}{\alpha_{ij}},$$

$$\bar{\alpha}_{il} = \frac{\alpha_{il}}{\alpha_{ij}}, \text{ dla } l \neq j,$$

$$\bar{\alpha}_{kj} = -\frac{\alpha_{kj}}{\alpha_{ij}}, \text{ dla } k \neq i,$$

$$\bar{\alpha}_{kl} = \alpha_{kl} - \frac{\alpha_{il}\alpha_{kj}}{\alpha_{ij}} \text{ dla } l \neq j \text{ i } k \neq i.$$

Opisane czynności wymiany zmiennej bazowej będziemy nazywać *piwotyzacją* (od ang. *pivot* – oś) element α_{ij} nazywa się *elementem głównym* (lub *osią*), i -ty wiersz nazywa się *wierszem głównym*, zaś j -ta kolumna – *kolumną główną*.

Opisane reguły wymiany zmiennych można schematycznie przedstawić następująco

$$\begin{bmatrix} p & q \\ r & s \end{bmatrix} \longrightarrow \begin{bmatrix} \frac{1}{p} & \frac{q}{p} \\ -\frac{r}{p} & s - \frac{qr}{p} \end{bmatrix}, \quad (3.8)$$

gdzie p jest elementem głównym, q – dowolnym elementem w wierszu głównym, r – dowolnym elementem w kolumnie głównej, zaś s – dowolnym pozostałym elementem.

3.2.2. Krawędzie zbioru rozwiązań dopuszczalnych

Na podstawie tablicy sympleksowej można również wyznaczyć krawędzie, które łączą wierzchołek odpowiadający tej tablicy z sąsiednimi wierzchołkami. Załóżmy, że tablica sympleksowa

| | | | | | | | |
|---------------|---------------|--------|---------------|--------|---------------|---------|--|
| | 1 | $-r_1$ | ... | $-r_j$ | ... | $-r_n$ | |
| α_{00} | α_{01} | ... | α_{0j} | ... | α_{0n} | $= z$ | |
| α_{10} | α_{11} | ... | α_{1j} | ... | α_{1n} | $= s_1$ | |
| ... | ... | ... | | ... | ... | ... | |
| α_{i0} | α_{i1} | | α_{ij} | ... | α_{in} | $= s_i$ | |
| ... | ... | ... | ... | ... | ... | ... | |
| α_{m0} | α_{m1} | ... | α_{mj} | | α_{mn} | $= s_m$ | |

przedstawia niezdegenerowane dopuszczalne rozwiązanie bazowe ($\alpha_{i0} > 0$, $i = 1, \dots, m$). Jeśli nadamy zmiennej niebazowej r_j wartość $t \geq 0$, przy czym pozostałe zmienne niebazowe r_k (dla $k \neq j$) przyjmujemy równe zeru, to otrzymamy

$$s_i(t) = -\alpha_{ij}t + \alpha_{i0}$$

$i = 1, \dots, m$. Równania te są parametrycznymi równaniami prostej zawierającej odpowiednią krawędź wychodzącą z wierzchołka, który przedstawia powyższa tablica. Idąc wzdłuż tej krawędzi pozostaniemy w zbiorze rozwiązań dopuszczalnych, dopóki $s_i \geq 0$, $i = 1, \dots, m$. Nietrudno zauważyć, że

$$s_i(t) \geq 0 \text{ dla wszystkich } i \Leftrightarrow t \leq \min_{i: \alpha_{ij} > 0} \frac{\alpha_{i0}}{\alpha_{ij}}. \quad (3.9)$$

Wartościom zmiennej $t \in [0, \min_{i: \alpha_{ij} > 0} \frac{\alpha_{i0}}{\alpha_{ij}}]$ odpowiadają więc punkty na odpowiedniej krawędzi zbioru rozwiązań dopuszczalnych wychodzącej z wierzchołka, który przedstawia powyższa tablica.

Ćwiczenie 3.2.2. Określić wszystkie krawędzie wychodzące z wierzchołka odpowiadającego tablicy sympleksowej podanej w przykładzie 3.1.1.

3.2.3. Piwotyzacja przy znanym dopuszczalnym rozwiązaniu bazowym (II faza metody sympleksowej)

W punkcie tym zakładamy, że wszystkie współrzędne wektora $(\alpha_{10}, \dots, \alpha_{m0})^\top$ w tablicy sympleksowej (3.7) są **nieujemne**. Wówczas odpowiadające tej tablicy rozwiązanie bazowe jest dopuszczalne. Jeśli założenie to jest spełnione dla startowej tablicy sympleksowej (3.5), czyli innymi słowy $b \geq 0$, to odpowiadający jej wierzchołek 0 (początek układu współrzędnych) jest rozwiązaniem dopuszczalnym ZPL (3.1). Wówczas tablica ta jest tablicą startową w algorytmie sympleksowym. W każdej iteracji algorytmu przeprowadza się wymianę bazy w taki sposób, aby wartość funkcji celu odpowiadająca kolejnej tablicy sympleksowej była nie mniejsza niż wartość odpowiadająca tablicy poprzedniej. Zakładamy ponadto, że nie pojawią się rozwiązania zdegenerowane (rozpatrzymy je później).

Podane poniżej reguły piwotyzacji odnoszą się do postaci klasycznej ZPL i do postaci krótkiej tablicy sympleksowej (3.7).

Algorytm 3.2.3 (sympleksowy, II faza)

Krok 1. (*wyбір kolumny głównej*). Wybrać dowolną kolumnę j_0 , dla której $\alpha_{0j_0} < 0$ (najczęściej wybiera się j_0 takie, że

$$\alpha_{0j_0} = \min\{\alpha_{0j} : \alpha_{0j} < 0\}.$$

Jeśli brak takiego j_0 , to tablica sympleksowa przedstawia rozwiązanie optymalne (patrz uwaga 3.2.1.b).

Krok 2. (*wyбір wiersza głównego*). Wśród wszystkich i , dla których $\alpha_{ij_0} > 0$, wybrać i_0 dla którego

$$\frac{\alpha_{i_0 0}}{\alpha_{i_0 j_0}} = \min_{i: \alpha_{ij_0} > 0} \frac{\alpha_{i 0}}{\alpha_{i j_0}}.$$

Jeśli brak takiego i ($\alpha_{ij_0} \leq 0$ dla wszystkich i), to funkcja celu jest nieograniczona z góry na zbiorze X rozwiązań dopuszczalnych.

Krok 3. (*piwotyżacja*). Wymienić zmienną bazową s_{i_0} ze zmienną niebazową r_{j_0} i przetransformować tablicę sympleksową zgodnie z regułą (3.8).

Uwaga 3.2.4. a) Wybór wiersza głównego podany w kroku 2 nie wyprowadza ze zbioru rozwiązań dopuszczalnych. Fakt ten wynika z równoważności (3.9).

b) Jeśli dla tablicy sympleksowej (3.7) przedstawiającej bazowe, niezdegenerowane rozwiązanie dopuszczalne dla wszystkich i zachodzą nierówności $\alpha_{ij_0} \leq 0$, to ZPL przedstawione za pomocą tej tablicy jest nieograniczone. Dla dowodu tego faktu wystarczy zauważyć, że w tym przypadku zwiększanie zmiennej niebazowej r_j do $+\infty$ nie wyprowadza ze zbioru rozwiązań dopuszczalnych, natomiast wartość funkcji celu rośnie nieograniczenie.

c) Jeśli nie wystąpi degeneracja, to po wykonaniu skończenie wielu iteracji II fazy algorytmu sympleksowego albo wyznaczymy rozwiązanie optymalne, albo stwierdzimy, że rozwiązań takich brak. Fakt ten wynika z obserwacji, że wierzchołków zbioru rozwiązań dopuszczalnych (tablic sympleksowych) jest skończenie wiele, a w przypadku braku degeneracji każda iteracja prowadzi do zwiększenia wartości funkcji celu.

Przykład 3.2.5. Za pomocą algorytmu sympleksowego wyznaczmy rozwiązanie optymalne ZPL podanego w przykładzie 3.1.1.

Początkowa tablica sympleksowa ma postać (zaznaczono element główny):

| 1 | $-x_1$ | $-x_2$ | |
|-----|--------|--------|---------|
| 0 | -1 | -2 | $= z$ |
| 100 | 1 | 1 | $= u_1$ |
| 720 | 6 | 9 | $= u_2$ |
| 60 | 0 | 1 | $= u_3$ |

Dwie kolejne piwotyzacje dają następujące tablice sympleksowe:

| 1 | $-x_1$ | $-u_3$ | |
|-----|--------|--------|---------|
| 120 | -1 | 2 | $= z$ |
| 40 | 1 | -1 | $= u_1$ |
| 180 | 6 | -9 | $= u_2$ |
| 60 | 0 | 1 | $= x_2$ |

i

| 1 | $-u_2$ | $-u_3$ | |
|-----|----------------|----------------|---------|
| 150 | $\frac{1}{6}$ | $\frac{1}{2}$ | $= z$ |
| 10 | $-\frac{1}{6}$ | $\frac{1}{2}$ | $= u_1$ |
| 30 | $\frac{1}{6}$ | $-\frac{3}{2}$ | $= x_1$ |
| 60 | 0 | 1 | $= x_2$ |

Ostatnia tablica przedstawia rozwiązanie optymalne: $x^* = (30, 60)$. Optymalna wartość funkcji celu wynosi $z = 150$. Na tym samym przykładzie przedstawimy kolejne wymiany bazy przeprowadzane na długiej postaci tablicy sympleksowej. Przekształcenia te są w istocie identyczne z metodą eliminacji Gaussa, w której element główny wyznacza się zgodnie z regułami podanymi w algorytmie sympleksowym. Otrzymujemy kolejne tablice sympleksowe z zaznaczonymi w nich elementami głównymi:

| z | u_1 | u_2 | u_3 | x_1 | x_2 | |
|-----|-------|-------|-------|-------|-------|---------|
| 1 | 0 | 0 | 0 | -1 | -2 | $= 0$ |
| 0 | 1 | 0 | 0 | 1 | 1 | $= 100$ |
| 0 | 0 | 1 | 0 | 6 | 9 | $= 720$ |
| 0 | 0 | 0 | 1 | 0 | 1 | $= 60$ |

| z | u_1 | u_2 | u_3 | x_1 | x_2 | |
|-----|-------|----------------|----------------|---|-------|-------|
| 1 | 0 | 0 | 2 | -1 | 0 | = 120 |
| 0 | 1 | 0 | -1 | 1 | 0 | = 40 |
| 0 | 0 | 1 | -9 | 6 | 0 | = 180 |
| 0 | 0 | 0 | 1 | 0 | 1 | = 60 |
| | | | | | | |
| z | u_1 | u_2 | u_3 | x_1 | x_2 | |
| 1 | 0 | $\frac{1}{6}$ | $\frac{1}{2}$ | 0 | 0 | = 150 |
| 0 | 1 | $-\frac{1}{6}$ | $\frac{1}{2}$ | 0 | 0 | = 10 |
| 0 | 0 | $\frac{1}{6}$ | $-\frac{3}{2}$ | 1 | 0 | = 30 |
| 0 | 0 | 0 | 1 | 0 | 1 | = 60 |

Jak widać, otrzymaliśmy to samo rozwiązanie optymalne co poprzednio: $x^* = (30, 60)$.

Uwaga 3.2.6. W części przedstawionych w tym rozdziale przykładów podajemy kolejne tablice sympleksowe od początkowej do tablicy przedstawiającej rozwiązanie optymalne. Powinno to ułatwić czytelnikowi samodzielne przeprowadzanie wymiany zmiennej bazowej i porównanie wyników. Do wymiany zmiennej bazowej (piwotyzacji) można użyć również komputera i zastosować odpowiednie pakiety matematyczne. Niektóre z nich można znaleźć w Internecie, na przykład na stronie:

<http://www.princeton.edu/~rvdb/LPbook/>

Stosowanie tych pakietów polega najczęściej na wskazaniu przez użytkownika elementu głównego tablicy sympleksowej, natomiast odpowiedni program wyznacza kolejną tablicę.

3.2.4. Przypadki szczególne

Rozpatrujemy zadanie programowania liniowego w postaci klasycznej i zakładamy, że punktem startowym w algorytmie sympleksowym jest dopusz-

czalne rozwiązanie bazowe ($b \geq 0$) oraz, że w trakcie jego realizacji nie występuje degeneracja (rozpatrzmy ją później).

Funkcja celu nieograniczona (zbiór rozwiązań optymalnych pusty)

Przypadek ten wystąpi dokładnie wówczas, gdy dla pewnej tablicy sympleksowej odpowiadającej rozwiązaniu dopuszczalnemu ($\alpha_{i0} \geq 0$ dla $i = 1, \dots, m$) istnieje kolumna j_0 , dla której $\alpha_{0j_0} < 0$ i dla której wszystkie pozostałe elementy tej kolumny są niedodatnie ($\alpha_{kj_0} \leq 0$, dla $k = 1, \dots, m$):

| | | | | | | | |
|--|---------------|---------------|---------|------------|---------|---------------|---------|
| | 1 | $-r_1$ | \dots | $-r_{j_0}$ | \dots | $-r_n$ | |
| | α_{00} | α_{01} | \dots | < 0 | \dots | α_{0n} | $= z$ |
| | ≥ 0 | α_{11} | \dots | ≤ 0 | \dots | α_{1n} | $= s_1$ |
| | \dots | \dots | \dots | | \dots | \dots | \dots |
| | ≥ 0 | α_{i1} | | ≤ 0 | \dots | α_{in} | $= s_i$ |
| | \dots | \dots | \dots | \dots | \dots | \dots | \dots |
| | ≥ 0 | α_{m1} | \dots | ≤ 0 | | α_{mn} | $= s_m$ |
| | | | | | | | |

Przykład 3.2.7. Rozważmy następujące ZPL:

$$\begin{array}{ll}
 \text{maksymalizować} & x_1 + x_2 \\
 \text{przy ograniczeniach} & x_1 - 2x_2 \leq 1 \\
 & -x_1 + x_2 \leq 1 \\
 & -2x_1 + 4x_2 \leq 2 \\
 & x_1, x_2 \geq 0.
 \end{array}$$

Początkowa tablica sympleksowa ma postać

| | | | | |
|--|---|--------|--------|---------|
| | 1 | $-x_1$ | $-x_2$ | |
| | 0 | -1 | -1 | $= z$ |
| | 1 | 1 | -2 | $= u_1$ |
| | 1 | -1 | 1 | $= u_2$ |
| | 2 | -2 | 4 | $= u_3$ |

Po piwotyzacji otrzymamy

| | 1 | $-u_1$ | $-x_2$ | |
|--|---|--------|--------|---------|
| | 1 | 1 | -3 | $= z$ |
| | 1 | 1 | -2 | $= x_1$ |
| | 2 | 1 | -1 | $= u_2$ |
| | 4 | 2 | 0 | $= u_3$ |

Teraz jeśli $u_1 = 0$ i $x_2 = t \geq 0$, to wszystkie zmienne bazowe pozostaną nieujemne i $z = 1 + 3t \rightarrow +\infty$, gdy $t \rightarrow +\infty$. Funkcja celu jest więc nieograniczona na zbiorze X rozwiązań dopuszczalnych, w konsekwencji problem nie posiada rozwiązania optymalnego.

Wiele rozwiązań optymalnych

Ten przypadek wystąpi wtedy, gdy istnieje tablica sympleksowa, której odpowiada optymalne rozwiązanie bazowe ($\alpha_{i0} \geq 0$, dla $i = 1, \dots, m$ oraz $\alpha_{0j} \geq 0$ dla $j = 1, \dots, n$) i istnieje para (i_0, j_0) , $i_0 \in \{1, \dots, m\}$, $j_0 \in \{1, \dots, n\}$, dla której $\alpha_{0j_0} = 0$, $\alpha_{i_0 0} > 0$ i $\alpha_{i_0 j_0} > 0$:

| | 1 | $-r_1$ | ... | $-r_{j_0}$ | ... | $-r_n$ | |
|--|---------------|------------------|-----|-----------------|-----|------------------|-------------|
| | α_{00} | ≥ 0 | ... | $= 0$ | ... | ≥ 0 | $= z$ |
| | > 0 | α_{11} | ... | α_{1j_0} | ... | α_{1n} | $= s_1$ |
| | ... | ... | ... | ... | ... | ... | ... |
| | > 0 | $\alpha_{i_0 1}$ | ... | > 0 | ... | $\alpha_{i_0 n}$ | $= s_{i_0}$ |
| | ... | ... | ... | ... | ... | ... | ... |
| | > 0 | α_{m1} | ... | α_{mj_0} | ... | α_{mn} | $= s_m$ |

Uwaga: dla rozwiązania zdegenerowanego, przypadek ten może mieć miejsce również pod innymi warunkami.

Przykład 3.2.8. Rozważmy następujące ZPL:

$$\begin{array}{rcl}
 \text{maksymalizować} & 4x_1 + 2x_2 & \\
 \text{przy ograniczeniach} & x_1 - 3x_2 \leq 4 & \\
 & x_2 \leq 7 & \\
 & 2x_1 + x_2 \leq 10 & \\
 & x_1, x_2 \geq 0. &
 \end{array}$$

Początkowa tablica sympleksowa ma postać

| 1 | $-x_1$ | $-x_2$ | |
|----|--------|--------|---------|
| 0 | -4 | -2 | $= z$ |
| 4 | 1 | -3 | $= u_1$ |
| 7 | 0 | 1 | $= u_2$ |
| 10 | 2 | 1 | $= u_3$ |

natomiast kolejne tablice mają postać

| 1 | $-u_1$ | $-x_2$ | |
|----|--------|--------|---------|
| 16 | 4 | -14 | $= z$ |
| 4 | 1 | -3 | $= x_1$ |
| 7 | 0 | 1 | $= u_2$ |
| 2 | -2 | 7 | $= u_3$ |

i

| 1 | $-u_1$ | $-u_3$ | |
|----------------|----------------|----------------|---------|
| 20 | 0 | 2 | $= z$ |
| $\frac{34}{7}$ | $\frac{1}{7}$ | $\frac{3}{7}$ | $= x_1$ |
| $\frac{47}{7}$ | $\frac{2}{7}$ | $-\frac{1}{7}$ | $= u_2$ |
| $\frac{2}{7}$ | $-\frac{2}{7}$ | $\frac{1}{7}$ | $= x_2$ |

Ostatnia tablica odpowiada rozwiązaniu optymalnemu $x^* = (\frac{34}{7}, \frac{2}{7})^\top$. Po dokonaniu wymiany zmiennych u_1 i u_2 otrzymamy nową tablicę

| 1 | $-u_2$ | $-u_3$ | |
|----------------|----------------|----------|---------|
| 20 | 0 | 2 | $= z$ |
| $\frac{3}{2}$ | $-\frac{1}{2}$ | \times | $= x_1$ |
| $\frac{47}{2}$ | $\frac{7}{2}$ | \times | $= u_1$ |
| 7 | 1 | \times | $= x_2$ |

(3.10)

która odpowiada nowemu bazowemu rozwiązaniu optymalnemu $x^{**} = (\frac{3}{2}, 7)^\top$. Ponieważ w tym przypadku kolejna dopuszczalna piwotyżacja doprowadzi do rozwiązania x^* , więc problem posiada tylko dwa dopuszczalne bazowe rozwiązania optymalne: x^* i x^{**} . Odpowiadają one dwóm wierzchołkom w zbiorze rozwiązań optymalnych X^* . Zgodnie z twierdzeniami 2.3.19 i 2.3.26 mamy:

$$X^* = \{x : x = \lambda x^* + (1 - \lambda)x^{**}, \lambda \in [0, 1]\} = \\ \{x : x = (\frac{34}{7}\lambda + \frac{3}{2}(1 - \lambda), \frac{2}{7}\lambda + 7(1 - \lambda), \lambda \in [0, 1]\}.$$

Przykład 3.2.9. Wyznamy wszystkie rozwiązania optymalne zadania:

$$\begin{array}{ll} \text{maksymalizować} & -\frac{9}{8}x_1 + \frac{15}{16}x_2 + \frac{9}{16}x_3 \\ \text{przy ograniczeniach} & \begin{array}{l} 3x_1 + 2x_2 + 6x_3 \leq \frac{27}{7} \\ -3x_1 + \frac{5}{2}x_2 + \frac{3}{2}x_3 \leq 4 \\ -\frac{3}{4}x_1 - x_2 + \frac{8}{7}x_3 \leq \frac{5}{9} \\ x_1, x_2, x_3 \geq 0. \end{array} \end{array}$$

Początkowa tablica sympleksowa ma postać

| 1 | $-x_1$ | $-x_2$ | $-x_3$ | |
|----------------|----------------|------------------|-----------------|---------|
| 0 | $\frac{9}{8}$ | $-\frac{15}{16}$ | $-\frac{9}{16}$ | $= z$ |
| $\frac{27}{7}$ | 3 | 2 | 6 | $= u_1$ |
| 4 | -3 | $\frac{5}{2}$ | $\frac{3}{2}$ | $= u_2$ |
| $\frac{5}{9}$ | $-\frac{3}{4}$ | -1 | $\frac{8}{7}$ | $= u_3$ |

Po piwotyżacji otrzymamy

| 1 | $-x_1$ | $-u_2$ | $-x_3$ | |
|-----------------|------------------|----------------|-----------------|---------|
| $\frac{1}{2}$ | 0 | $\frac{3}{8}$ | 0 | $= z$ |
| $\frac{23}{35}$ | $\frac{27}{5}$ | $-\frac{4}{5}$ | $\frac{24}{5}$ | $= u_1$ |
| $\frac{8}{5}$ | $-\frac{6}{5}$ | $\frac{2}{5}$ | $\frac{3}{5}$ | $= x_2$ |
| $\frac{97}{45}$ | $-\frac{39}{20}$ | $\frac{2}{5}$ | $\frac{61}{35}$ | $= u_3$ |

Ostatnia tablica odpowiada rozwiązaniu optymalnemu $x' = (0, \frac{8}{5}, 0)^\top$. Dwa następne bazowe rozwiązania optymalne x'' bądź x''' otrzymamy po wymianie zmiennych niebazowych x_1 bądź x_3 ze zmienną bazową u_1 :

| | 1 | $-u_1$ | $-u_2$ | $-x_3$ | |
|--|------------------|-----------------|---------------|-----------------|---------|
| | $\frac{1}{2}$ | 0 | $\frac{3}{8}$ | 0 | $= z$ |
| | $\frac{23}{189}$ | $\frac{5}{27}$ | \times | $\frac{8}{45}$ | $= x_1$ |
| | $\frac{110}{63}$ | $\frac{2}{9}$ | \times | $\frac{5}{3}$ | $= x_2$ |
| | $\frac{67}{28}$ | $\frac{13}{36}$ | \times | $\frac{73}{21}$ | $= u_3$ |

albo

| | 1 | $-x_1$ | $-u_2$ | $-u_1$ | |
|--|------------------|---------------|---------------|----------------|---------|
| | $\frac{1}{2}$ | 0 | $\frac{3}{8}$ | 0 | $= z$ |
| | $\frac{23}{168}$ | $\frac{9}{8}$ | \times | $\frac{5}{24}$ | $= x_3$ |
| | $\frac{85}{56}$ | — | \times | — | $= x_2$ |
| | + | — | \times | — | $= u_3$ |

Mamy: $x'' = (\frac{23}{189}, \frac{110}{63}, 0)^\top$ oraz $x''' = (0, \frac{85}{56}, \frac{23}{168})^\top$. Dalsze piwotyzacje nie przyniosą nowych rozwiązań optymalnych. Tak więc

$$X^* = \{x \in \mathbb{R}^3 : x = \lambda_1 x' + \lambda_2 x'' + \lambda_3 x''', \lambda_1, \lambda_2, \lambda_3 \geq 0, \lambda_1 + \lambda_2 + \lambda_3 = 1\}.$$

Zbiór rozwiązań optymalnych nieograniczony

Przypadek ten może wystąpić dokładnie wtedy, gdy istnieje tablica sympleksowa odpowiadająca rozwiązaniu optymalnemu ($\alpha_{i0} \geq 0$ dla $i = 1, \dots, m$ oraz $\alpha_{0j} \geq 0$ dla $j = 1, \dots, n$), dla której istnieje j_0 takie, że $\alpha_{0j_0} = 0$ i dla

wszystkich i zachodzi $\alpha_{i0} = 0$ (degeneracja) bądź $\alpha_{ij_0} \leq 0$:

| | 1 | $-r_1$ | ... | $-r_{j_0}$ | ... | $-r_n$ | |
|---------------|---------------|--------|-----------------|------------|---------------|--------|---------|
| α_{00} | ≥ 0 | ... | $= 0$ | ... | ≥ 0 | | $= z$ |
| ≥ 0 | α_{11} | ... | ≤ 0 | ... | α_{1n} | | $= s_1$ |
| ... | ... | ... | | ... | ... | | ... |
| $= 0$ | α_{i1} | | α_{ij_0} | ... | α_{in} | | $= s_i$ |
| ... | ... | ... | ... | ... | ... | | ... |
| ≥ 0 | α_{m1} | ... | ≤ 0 | | α_{mn} | | $= s_m$ |

Przykład 3.2.10. Rozważmy następujące ZPL:

$$\begin{aligned}
 &\text{maksymalizować} && -2x_1 + 4x_2 \\
 &\text{przy ograniczeniach} && -2x_1 + x_2 \leq 1 \\
 & && -x_1 + 2x_2 \leq 4 \\
 & && x_1, x_2 \geq 0.
 \end{aligned}$$

Początkowa tablica ma postać

| | 1 | $-x_1$ | $-x_2$ | |
|---|----|---|--------|---------|
| 0 | 2 | -4 | | $= z$ |
| 1 | -2 | 1 | | $= u_1$ |
| 4 | -1 | 2 | | $= u_2$ |

Dwie następne tablice mają postać

| | 1 | $-x_1$ | $-u_1$ | |
|---|---|--------|--------|---------|
| 4 | -6 | 4 | | $= z$ |
| 1 | -2 | 1 | | $= x_2$ |
| 2 | 3 | -2 | | $= u_2$ |

i

| | 1 | $-u_2$ | $-u_1$ | |
|---------------|---------------|----------------|--------|---------|
| 8 | 2 | 0 | | $= z$ |
| $\frac{7}{3}$ | $\frac{2}{3}$ | $-\frac{1}{3}$ | | $= x_2$ |
| $\frac{2}{3}$ | $\frac{1}{3}$ | $-\frac{2}{3}$ | | $= x_1$ |

Ostatnia tablica odpowiada bazowemu rozwiązaniu optymalnemu $x^* = (\frac{2}{3}, \frac{7}{3})^\top$. Dla ostatniej tablicy żadna piwotyzacja nie jest dopuszczalna,

ponieważ wszystkie elementy w jedynej dopuszczalnej kolumnie głównej są ujemne. Biorąc teraz $u_2 = 0$ i $u_1 = t \geq 0$, otrzymamy $z = 8$, $x_1 = \frac{2}{3} + \frac{2}{3}t \geq 0$, $x_2 = \frac{7}{3} + \frac{1}{3}t \geq 0$. Tak więc zbiór rozwiązań optymalnych X^* jest półprostą:

$$X^* = \{x \in \mathbb{R}^2 : x = x^* + \left(\frac{2}{3}, \frac{1}{3}\right)^\top t, t \geq 0\}.$$

Przykład 3.2.11. Rozważmy następujące ZPL:

$$\begin{array}{ll} \text{maksymalizować} & -\frac{25}{6}x_1 + \frac{5}{2}x_2 + \frac{4}{7}x_3 \\ \text{przy ograniczeniach} & \begin{array}{ll} -x_1 + \frac{4}{9}x_2 - 2x_3 & \leq \frac{4}{9} \\ -\frac{2}{3}x_1 + \frac{2}{5}x_2 + \frac{2}{5}x_3 & \leq \frac{2}{5} \\ -2x_1 - \frac{3}{2}x_2 + \frac{1}{2}x_3 & \leq 6 \\ x_1, x_2, x_3 & \geq 0. \end{array} \end{array}$$

Początkowa tablica sympleksowa ma postać

| | | $-x_1$ | $-x_2$ | $-x_3$ | |
|--|---------------|----------------|----------------|----------------|---------|
| | 0 | $\frac{25}{6}$ | $-\frac{5}{2}$ | $-\frac{4}{7}$ | $= z$ |
| | $\frac{4}{9}$ | -1 | $\frac{4}{9}$ | -2 | $= u_1$ |
| | $\frac{2}{5}$ | $-\frac{2}{3}$ | $\frac{2}{5}$ | $\frac{2}{5}$ | $= u_2$ |
| | 6 | -2 | $-\frac{3}{2}$ | $\frac{1}{2}$ | $= u_3$ |

Teraz możliwy jest wybór jednego z trzech elementów głównych: $\alpha_{12} = \frac{4}{9}$ albo $\alpha_{22} = \frac{2}{5}$ albo $\alpha_{23} = \frac{2}{5}$. Wybór drugiego z nich daje kolejną tablicę sympleksową:

| | 1 | $-x_1$ | $-u_2$ | $-x_3$ | |
|--|----------------|-----------------|----------------|-----------------|---------|
| | $\frac{5}{2}$ | 0 | $\frac{25}{4}$ | $\frac{27}{14}$ | $= z$ |
| | 0 | $-\frac{7}{27}$ | \times | \times | $= u_1$ |
| | 1 | $-\frac{5}{3}$ | \times | \times | $= x_2$ |
| | $\frac{15}{2}$ | $-\frac{12}{5}$ | \times | \times | $= u_3$ |

Tablica ta przedstawia zdegenerowane bazowe rozwiązanie optymalne $x' = (0, 1, 0)^\top$. Jeśli teraz przyjąć $u_2 = x_3 = 0$ i $x_1 = t \geq 0$, otrzymamy $z = \frac{5}{2}$, $x_1 = t$, $x_2 = 1 + \frac{5}{3}t \geq 0$, $x_3 = 0$. Zbiór rozwiązań optymalnych X^* zawiera więc półprostą opisaną powyższymi równaniami parametrycznymi.

Wybór w początkowej tablicy sympleksowej elementu głównego $\alpha_{12} = \frac{4}{9}$ prowadzi do kolejnych tablic

| 1 | $-x_1$ | $-u_1$ | $-x_3$ | | 1 | $-x_1$ | $-u_1$ | $-u_2$ | |
|----------------|------------------|-----------------|-------------------|---------|----------------|----------|----------|----------|---------|
| $\frac{5}{2}$ | $\frac{135}{24}$ | $\frac{45}{8}$ | $-\frac{331}{28}$ | $= z$ | $\frac{5}{2}$ | + | + | + | $= z$ |
| 1 | $-\frac{4}{9}$ | $\frac{9}{4}$ | $\frac{9}{2}$ | $= x_2$ | 1 | \times | \times | \times | $= x_2$ |
| 0 | $\frac{7}{30}$ | $-\frac{9}{10}$ | $\frac{11}{5}$ | $= u_2$ | 0 | \times | \times | \times | $= x_3$ |
| $\frac{15}{2}$ | \times | $\frac{27}{8}$ | $-\frac{25}{4}$ | $= u_3$ | $\frac{15}{2}$ | \times | \times | \times | $= u_3$ |

Ostatnia przedstawia otrzymane już uprzednio rozwiązanie $x' = (0, 1, 0)^\top$. Z tej tablicy nie można jednak odczytać, że zbiór X^* jest nieograniczony. Przyczyną jest pojawiająca się degeneracja. Zauważmy ponadto, że dla tej tablicy dowolna piwotyzacja jest niedopuszczalna.

W końcu wybór w początkowej tablicy sympleksowej elementu głównego $\alpha_{23} = \frac{2}{5}$ prowadzi do tablicy

| 1 | $-x_1$ | $-x_2$ | $-u_2$ | |
|----------------|-----------------|------------------|----------------|---------|
| $\frac{4}{7}$ | $\frac{45}{14}$ | $-\frac{27}{14}$ | $\frac{10}{7}$ | $= z$ |
| $\frac{22}{9}$ | $-\frac{13}{3}$ | $\frac{22}{9}$ | 5 | $= u_1$ |
| 1 | $-\frac{5}{2}$ | 1 | $\frac{5}{2}$ | $= x_3$ |
| $\frac{11}{2}$ | $-\frac{7}{6}$ | -2 | $-\frac{5}{4}$ | $= u_3$ |

Mamy teraz dwie dopuszczalne możliwości wyboru elementu głównego: $\alpha_{12} = \frac{22}{9}$ bądź $\alpha_{22} = 1$. Wybór dowolnego z nich prowadzi do wyznaczonego już uprzednio rozwiązania x' . Tak więc

$$X^* = \{x \in \mathbb{R}^3 : x = x' + (1, \frac{5}{3}, 0)^\top t, t \geq 0\}.$$

3.2.5. Piwotyzacja przy nieznanym dopuszczalnym rozwiązaniu bazowym (I faza metody sympleksowej)

Jeśli wektor b w początkowej tablicy sympleksowej (3.5) ma przynajmniej jedną ujemną współrzędną, to tablica ta przedstawia niedopuszczalne rozwiązanie bazowe. Opiszemy teraz, jak można w tym przypadku otrzymać dopuszczalne rozwiązanie bazowe za pomocą algorytmu sympleksowego.

Uzyskanie bazowego rozwiązania dopuszczalnego

Zacznijmy od następującego przykładu.

Przykład 3.2.12. Rozważmy następujące ZPL:

$$\begin{array}{ll} \text{maksymalizować} & x_1 + x_2 \\ \text{przy ograniczeniach} & -\frac{1}{3}x_1 + x_2 \leq 3 \\ & 4x_1 + 3x_2 \leq 24 \\ & -2x_1 - 3x_2 \leq -6 \\ & x_1, x_2 \geq 0. \end{array}$$

Początkowa tablica ma postać

| 1 | $-x_1$ | $-x_2$ | |
|----|----------------|--------|---------|
| 0 | -1 | -1 | $= z$ |
| 3 | $-\frac{1}{3}$ | 1 | $= u_1$ |
| 24 | 4 | 3 | $= u_2$ |
| -6 | -2 | -3 | $= u_3$ |

i przedstawia niedopuszczalne rozwiązanie bazowe $(x, u) = (0, 0, 3, 24, -6)$. Naszym pierwszym celem jest teraz uzyskanie nieujemnej wartości zmiennej uzupełniającej u_3 . Cel ten osiągniemy, jeśli wybierzemy u_3 jako tzw. *potencjalną funkcję celu*, którą będziemy maksymalizować przy pozostałych spełnionych ograniczeniach. Odpowiednie piwotyzacje będą prowadzone

tak długo, aż $u_3 \geq 0$. W naszym przypadku otrzymamy już po pierwszej piwotyzacji tablicę

| 1 | $-x_1$ | $-u_1$ | |
|----|---|--------|---------|
| 3 | $-\frac{4}{3}$ | 1 | $= z$ |
| 3 | $-\frac{1}{3}$ | 1 | $= x_2$ |
| 15 | 5 | -3 | $= u_2$ |
| 3 | -3 | 3 | $= u_3$ |

która przedstawia dopuszczalne rozwiązanie bazowe. Można więc ją traktować jako startową tablicę sympleksową w II fazie algorytmu sympleksowego i postępować dalej według opisanych już reguł piwotyzacji. Otrzymamy wówczas następującą tablicę:

| 1 | $-u_2$ | $-u_1$ | |
|----|----------------|----------------|---------|
| 7 | $\frac{4}{15}$ | $\frac{1}{5}$ | $= z$ |
| 4 | $\frac{1}{15}$ | $\frac{4}{5}$ | $= x_2$ |
| 3 | $\frac{1}{15}$ | $-\frac{3}{5}$ | $= x_1$ |
| 12 | $\frac{3}{5}$ | $\frac{6}{5}$ | $= u_3$ |

która odpowiada rozwiązaniu optymalnemu $x^* = (3, 4)^\top$.

Przedstawimy teraz I fazę algorytmu sympleksowego, która podaje sposób wyznaczenia dopuszczalnego rozwiązania bazowego bądź podaje regułę pozwalającą stwierdzić, że brak jest takiego rozwiązania. Poniższy algorytm dotyczy postaci klasycznej ZPL i związanej z nią krótkiej postaci tablicy sympleksowej (3.7).

Algorytm 3.2.13 (sympleksowy, I faza)

Krok 1. (*wyбір pomocniczej funkcji celu*). Wybrać największe $k \geq 1$, dla którego $\alpha_{k0} < 0$:

$$k = \max\{i \geq 1 : \alpha_{i0} < 0\}.$$

Jeśli brak takiego k ($\alpha_{i0} \geq 0$ dla $i = 1, \dots, m$), to tablica odpowiada dopuszczalnemu rozwiązaniu bazowemu (w tym przypadku przechodzi się do II fazy algorytmu sympleksowego).

Krok 2. (*wyбір kolumny głównej*). Wybrać j_0 , dla którego $\alpha_{kj_0} < 0$ (najczęściej wybiera się j_0 takie, że

$$\alpha_{kj_0} = \min\{\alpha_{kj} : \alpha_{kj} < 0\}.$$

Jeśli brak jest takiego j_0 ($\alpha_{kj} \geq 0$ dla $j = 1, \dots, n$), to brak jest rozwiązania dopuszczalnego (problem jest sprzeczny).

Krok 3. (*wyбір wiersza głównego*). Wśród wszystkich $i \geq k$, dla których $\frac{\alpha_{i0}}{\alpha_{ij_0}} > 0$, wybrać i_0 , dla którego iloraz ten osiąga minimum:

$$\frac{\alpha_{i_0 0}}{\alpha_{i_0 j_0}} = \min\left\{\frac{\alpha_{i0}}{\alpha_{ij_0}} : i \geq k \text{ i } \frac{\alpha_{i0}}{\alpha_{ij_0}} > 0\right\}. \quad (3.11)$$

(Zauważmy, że takie i_0 istnieje, gdyż $\alpha_{k0} < 0$ i $\alpha_{kj_0} < 0$.)

Krok 4. (*piwotyzacja*). Wymienić zmienną bazową s_{i_0} ze zmienną niebazową r_{j_0} i przetransformować tablicę sympleksową zgodnie z regułą piwotyzacji (3.8). Przejść do kroku 1.

Uwaga 3.2.14. a) Dla $i > k$ elementy α_{i0} pozostaną po piwotyzacji nieujemne:

- $\bar{\alpha}_{i_0 0} = \frac{\alpha_{i_0 0}}{\alpha_{i_0 j_0}} > 0$ zgodnie z równością (3.11).
- Dla $i \neq i_0$ mamy $\bar{\alpha}_{i0} = \alpha_{i0} - \alpha_{ij_0} \frac{\alpha_{i_0 0}}{\alpha_{i_0 j_0}}$. Jeśli $\alpha_{ij_0} \leq 0$, to $\bar{\alpha}_{i0} \geq \alpha_{i0} \geq 0$, gdyż $\frac{\alpha_{i_0 0}}{\alpha_{i_0 j_0}} > 0$ zgodnie z równością (3.11). Jeśli natomiast $\alpha_{ij_0} > 0$, to dla $i \geq k$ mamy $\bar{\alpha}_{i0} = \alpha_{ij_0} \left(\frac{\alpha_{i0}}{\alpha_{ij_0}} - \frac{\alpha_{i_0 0}}{\alpha_{i_0 j_0}} \right) \geq 0$ zgodnie z równością (3.11).

b) Element α_{k0} wzrośnie po piwotyzacji, gdyż

$$\bar{\alpha}_{k0} = \alpha_{k0} - \alpha_{kj_0} \frac{\alpha_{i_0 0}}{\alpha_{i_0 j_0}} > \alpha_{k0}$$

i $\alpha_{kj_0} < 0$ i $\frac{\alpha_{i_0 0}}{\alpha_{i_0 j_0}} > 0$ zgodnie z równością (3.11).

c) Załóżmy, że ZPL posiada rozwiązanie dopuszczalne. (W przeciwnym wypadku sprzeczność ograniczeń zostanie stwierdzona po skończeniu wielu iteracjach I fazy algorytmu sympleksowego. W przypadku niewystąpienia degeneracji kolejne tablice przedstawiają bowiem różne rozwiązania bazowe, których jest skończenie wiele). Wówczas z b) wynika, że po skończeniu wielu iteracjach $\bar{\alpha}_{k0}$ będzie dodatnie (w wyniku każdej piwotyzacji α_{k0} rośnie). Ponadto z a) wynika, że wówczas zbiór wierszy z ujemnymi elementami α_{i0} zmniejszy się. Zatem po skończeniu wielu iteracjach I faza algorytmu sympleksowego doprowadzi do bazowego rozwiązania dopuszczalnego.

d) Zanim wybierze się k w kroku 1., można pozamieniać wiersze tablicy sympleksowej tak, by $\alpha_{i0} < 0$ dla $i = 1, \dots, k$ i $\alpha_{i0} \geq 0$ dla $i = k + 1, \dots, m$. Prowadzi to często do szybszego wyznaczenia rozwiązania dopuszczalnego. Należy jednak wówczas pamiętać o tym, aby po takiej operacji kolejną ewentualną zamianę wierszy przeprowadzić dopiero wtedy, gdy w wyniku piwotyzacji α_{k0} osiągnie nieujemną wartość.

Uwaga 3.2.15. Rozważmy teraz ograniczenia zadania programowania liniowego w postaci kanonicznej

$$\begin{aligned} Ax + Iu &= b \\ x, u &\geq 0 \end{aligned}$$

i pomnóżmy przez -1 te równania, dla których prawa strona jest ujemna ($b_i < 0$). Otrzymany układ równań zapiszmy w postaci

$$A'z = b',$$

przy czym $z = (x^\top, u^\top)^\top$. Bazowe rozwiązanie dopuszczalne wyjściowego zadania można również wyznaczyć wprowadzając tzw. *zmiennne sztuczne* $y \in \mathbb{R}^m$ i następnie rozwiązując problem pomocniczy

$$\begin{aligned} &\text{maksymalizować} && -e^\top y \\ &\text{przy ograniczeniach} && A'z + y = b' \\ &&& z, y \geq 0. \end{aligned} \tag{3.12}$$

Można pokazać, że wyjściowe ZPL posiada rozwiązanie dopuszczalne $\bar{z} = (\bar{x}^\top, \bar{u}^\top)^\top$ wtedy i tylko wtedy, gdy $(\bar{z}, 0)$ jest rozwiązaniem optymalnym problemu (3.12). Zauważmy przy tym, że dla ostatniego problemu istnieje rozwiązanie dopuszczalne (punkt $(0, b') \geq 0$ spełnia podane w nim ograniczenia) i posiada rozwiązanie optymalne (funkcja celu jest ograniczona przez 0).

Zbiór rozwiązań dopuszczalnych pusty

Przykład 3.2.16. Rozważmy następujące ZPL:

$$\begin{array}{ll} \text{maksymalizować} & x_1 - x_2 \\ \text{przy ograniczeniach} & 2x_1 + x_2 \leq 2 \\ & x_1 + 2x_2 \leq 2 \\ & x_1 + x_2 \geq 2 \\ & x_1, x_2 \geq 0. \end{array}$$

Początkowa tablica sympleksowa ma postać

| 1 | $-x_1$ | $-x_2$ | |
|----|--------|--------|---------|
| 0 | -1 | 1 | $= z$ |
| 2 | 2 | 1 | $= u_1$ |
| 2 | 1 | 2 | $= u_2$ |
| -2 | -1 | -1 | $= u_3$ |

Ponieważ odpowiada ona niedopuszczalnemu rozwiązaniu bazowemu, należy zastosować I fazę algorytmu sympleksowego. Jako funkcję pomocniczą wybierzemy u_3 . Aby skrócić liczbę dokonywanych pivotyzacji, przesuniemy trzeci wiersz w miejsce pierwszego. Otrzymamy kolejne tablice symplek-

sowe

| 1 | $-x_1$ | $-x_2$ | |
|----|---|--------|---------|
| 0 | -1 | 1 | $= z$ |
| -2 | -1 | -1 | $= u_3$ |
| 2 | 2 | 1 | $= u_1$ |
| 2 | 1 | 2 | $= u_2$ |

,

| 1 | $-u_1$ | $-x_2$ | |
|----|----------------|--|---------|
| 1 | $\frac{1}{2}$ | $\frac{3}{2}$ | $= z$ |
| -1 | $\frac{1}{2}$ | $-\frac{1}{2}$ | $= u_3$ |
| 1 | $\frac{1}{2}$ | $\frac{1}{2}$ | $= x_1$ |
| 1 | $-\frac{1}{2}$ | $\frac{3}{2}$ | $= u_2$ |

| 1 | $-u_1$ | $-u_2$ | |
|----------------|---------------|---------------|---------|
| \times | \times | \times | $= z$ |
| $-\frac{2}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ | $= u_3$ |
| \times | \times | \times | $= x_1$ |
| \times | \times | \times | $= x_2$ |

Ostatnia tablica wskazuje na to, że zbiór rozwiązań dopuszczalnych X jest pusty (ograniczenia $u_3 = -\frac{1}{3}u_1 - \frac{1}{3}u_2 - \frac{2}{3}$, $u_1, u_2, u_3 \geq 0$ są sprzeczne).

Przykład 3.2.17. Rozważmy następujące ZPL:

$$\begin{aligned}
 &\text{maksymalizować} && \frac{1}{2}x_1 - x_2 - x_3 \\
 &\text{przy ograniczeniach} && -\frac{1}{2}x_1 + 2x_2 + x_3 \leq 2 \\
 & && \frac{1}{2}x_1 - 2x_2 + x_3 \leq -3 \\
 & && x_2 - x_3 \leq 2 \\
 & && x_1, x_2 \geq 0.
 \end{aligned}$$

Dwie pierwsze tablice sympleksowe mają postać (u_2 wybieramy jako pomocniczą funkcję celu):

| 1 | $-x_1$ | $-x_2$ | $-x_3$ | |
|----|----------------|---|--------|---------|
| 0 | $-\frac{1}{2}$ | 1 | 1 | $= z$ |
| 2 | $-\frac{1}{2}$ | 2 | 1 | $= u_1$ |
| -3 | $\frac{1}{2}$ | -2 | 1 | $= u_2$ |
| 2 | 0 | 1 | -1 | $= u_3$ |

,

| 1 | $-x_1$ | $-u_1$ | $-x_3$ | |
|----------|----------|----------|----------|---------|
| \times | \times | \times | \times | $= z$ |
| \times | \times | \times | \times | $= x_2$ |
| -1 | 0 | 1 | 2 | $= u_2$ |
| \times | \times | \times | \times | $= u_3$ |

Ostatnia tablica wskazuje na to, że brak jest rozwiązań dopuszczalnych.

3.3. Degeneracja

Matematyka uczy, że należy liczyć się z zerami.

Algorytm sympleksowy umożliwia wyznaczenie w skończenie wielu krokach bazowego rozwiązania optymalnego (II faza) lub bazowego rozwiązania dopuszczalnego (I faza) przy założeniu, że nie wystąpią po drodze rozwiązania zdegenerowane. Założenie to jest wykorzystywane w istotny sposób w celu pokazania, że funkcja celu (pomocnicza funkcja celu) rośnie w wyniku każdej piwotyzacji, z czego z kolei wynika zbieżność algorytmu w skończenie wielu krokach. Bez tego założenia może powstać cykl kolejnych piwotyzacji, dających rozwiązania zdegenerowane, które nie są optymalne. W ten sposób powstanie nieskończony ciąg rozwiązań bazowych nie przybliżających się nawet do rozwiązania optymalnego (II faza) bądź dopuszczalnego (I faza).

Przykład 3.3.1. Rozważmy następujące ZPL:

$$\begin{array}{ll}
 \text{maksymalizować} & \frac{3}{4}x_1 - 20x_2 + \frac{1}{2}x_3 - 6x_4 \\
 \text{przy ograniczeniach} & \begin{array}{l}
 \frac{1}{4}x_1 - 8x_2 - x_3 + 9x_4 \leq 0 \\
 \frac{1}{2}x_1 - 12x_2 - \frac{1}{2}x_3 + 3x_4 \leq 0 \\
 x_3 \leq 1 \\
 x_1, x_2, x_3, x_4 \geq 0.
 \end{array}
 \end{array}$$

Początkowa tablica sympleksowa ma postać

| | 1 | $-x_1$ | $-x_2$ | $-x_3$ | $-x_4$ | |
|--|---|----------------|--------|----------------|--------|---------|
| | 0 | $-\frac{3}{4}$ | 20 | $-\frac{1}{2}$ | 6 | $= z$ |
| | 0 | $\frac{1}{4}$ | -8 | -1 | 9 | $= u_1$ |
| | 0 | $\frac{1}{2}$ | -12 | $-\frac{1}{2}$ | 3 | $= u_2$ |
| | 1 | 0 | 0 | 1 | 0 | $= u_3$ |

i odpowiada rozwiązaniu zdegenerowanemu $(x, u) = (0, 0, 0, 0, 0, 0, 1)$. Wybierając za każdym razem przy kolejnych piwotyzacjach kolumnę główną j_0 , dla której $-c_{j_0} = \min\{-c_j : j = 1, \dots, n\}$, i pierwszy – spośród możliwych do wyboru – wiersz główny, otrzymamy kolejne tablice sympleksowe:

| 1 | $-u_1$ | $-x_2$ | $-x_3$ | $-x_4$ | |
|---|--------|--------|----------------|--------|---------|
| 0 | 3 | -4 | $-\frac{7}{2}$ | 33 | $= z$ |
| 0 | 4 | -32 | -4 | 36 | $= x_1$ |
| 0 | -2 | 4 | $\frac{3}{2}$ | -15 | $= u_2$ |
| 1 | 0 | 0 | 1 | 0 | $= u_3$ |

| 1 | $-u_1$ | $-u_2$ | $-x_3$ | $-x_4$ | |
|---|----------------|---------------|---------------|-----------------|---------|
| 0 | 1 | 1 | -2 | 18 | $= z$ |
| 0 | -12 | 8 | 8 | -84 | $= x_1$ |
| 0 | $-\frac{1}{2}$ | $\frac{1}{4}$ | $\frac{3}{8}$ | $-\frac{15}{4}$ | $= x_2$ |
| 1 | 0 | 0 | 1 | 0 | $= u_3$ |

| 1 | $-u_1$ | $-u_2$ | $-x_1$ | $-x_4$ | |
|---|----------------|----------------|-----------------|-----------------|---------|
| 0 | -2 | 3 | $\frac{1}{4}$ | -3 | $= z$ |
| 0 | $-\frac{3}{2}$ | 1 | $\frac{1}{8}$ | $-\frac{21}{2}$ | $= x_3$ |
| 0 | $\frac{1}{16}$ | $-\frac{1}{8}$ | $-\frac{3}{64}$ | $\frac{3}{16}$ | $= x_2$ |
| 1 | $\frac{3}{2}$ | -1 | $-\frac{1}{8}$ | $\frac{21}{2}$ | $= u_3$ |

| 1 | $-u_1$ | $-u_2$ | $-x_1$ | $-x_2$ | |
|---|---------------|----------------|----------------|----------------|---------|
| 0 | -1 | 1 | $-\frac{1}{2}$ | 16 | $= z$ |
| 0 | 2 | -6 | $-\frac{5}{2}$ | 56 | $= x_3$ |
| 0 | $\frac{1}{3}$ | $-\frac{2}{3}$ | $-\frac{1}{4}$ | $\frac{16}{3}$ | $= x_4$ |
| 1 | -2 | 6 | $\frac{5}{2}$ | -56 | $= u_3$ |

| 1 | $-x_3$ | $-u_2$ | $-x_1$ | $-x_2$ | |
|---|----------------|-----------------------|----------------|--------|---------|
| 0 | $\frac{1}{2}$ | -2 | $-\frac{7}{4}$ | 44 | $= z$ |
| 0 | $\frac{1}{2}$ | -3 | $-\frac{5}{4}$ | 28 | $= u_1$ |
| 0 | $-\frac{1}{6}$ | $\boxed{\frac{1}{3}}$ | $\frac{1}{6}$ | -4 | $= x_4$ |
| 1 | 1 | 0 | 0 | 0 | $= u_3$ |

| 1 | $-x_3$ | $-x_4$ | $-x_1$ | $-x_2$ | |
|---|----------------|--------|----------------|--------|---------|
| 0 | $-\frac{1}{2}$ | 6 | $-\frac{3}{4}$ | 20 | $= z$ |
| 0 | -1 | 9 | $\frac{1}{4}$ | -8 | $= u_1$ |
| 0 | $-\frac{1}{2}$ | 3 | $\frac{1}{2}$ | -12 | $= u_2$ |
| 1 | 1 | 0 | 0 | 0 | $= u_3$ |

Widzimy, że ostatnia tablica jest równoważna pierwszej oraz, że kolejne piwotyzacje dokonywane zgodnie z podaną wyżej regułą nie dadzą żadnego nowego bazowego rozwiązania dopuszczalnego. Jeśli jednak w początkowej tablicy wybierzemy trzecią kolumnę jako kolumnę główną, otrzymamy następujące tablice:

| 1 | $-x_1$ | $-x_2$ | $-u_3$ | $-x_4$ | |
|---------------|-----------------------|--------|---------------|--------|---------|
| $\frac{1}{2}$ | $-\frac{3}{4}$ | 20 | $\frac{1}{2}$ | 6 | $= z$ |
| 1 | $\frac{1}{4}$ | -8 | 1 | 9 | $= u_1$ |
| $\frac{1}{2}$ | $\boxed{\frac{1}{2}}$ | -12 | $\frac{1}{2}$ | 3 | $= u_2$ |
| 1 | 0 | 0 | 1 | 0 | $= x_3$ |

| 1 | $-u_2$ | $-x_2$ | $-u_3$ | $-x_4$ | |
|---------------|---------------|--------|---------------|----------------|---------|
| $\frac{5}{4}$ | $\frac{3}{2}$ | 2 | $\frac{5}{4}$ | $\frac{21}{2}$ | $= z$ |
| $\frac{3}{4}$ | × | × | × | × | $= u_1$ |
| 1 | × | × | × | × | $= x_1$ |
| 1 | × | × | × | × | $= x_3$ |

Ostatnia tablica odpowiada rozwiązaniu optymalnemu $x = (1, 0, 1, 0)$ z optymalną wartością funkcji celu $z = \frac{5}{4}$.

Degeneracji można uniknąć zakłócając prawą stronę ograniczeń, czyli rozwiązując problem

$$\begin{array}{ll} \text{maksymalizować} & c^\top x \\ \text{przy ograniczeniach} & Ax \leq b + \varepsilon \\ & x \geq 0, \end{array}$$

przy czym $\varepsilon = (\varepsilon_1, \dots, \varepsilon_m)^\top$ i współrzędne $\varepsilon_i > 0$, $i = 1, \dots, m$, wybrane zostały dostatecznie małe. Można pokazać, że wyjściowy problem (3.1) posiada rozwiązanie wtedy i tylko wtedy, gdy problem z zakłóconą prawą stroną posiada rozwiązanie dla każdego $\varepsilon > 0$. Jeśli wziąć za $\varepsilon_i > 0$, $i = 1, \dots, m$, wystarczająco małe liczby losowe, to prawdopodobieństwo, że nastąpi degeneracja jest równe zeru.

Uwaga 3.3.2. Dla praktycznej realizacji algorytmu sympleksowego w przypadku wystąpienia degeneracji wystarczy, aby, wyłącznie w celu wyznaczenia wiersza głównego, dodać dostatecznie małe $\varepsilon > 0$ do prawej strony tych równań, dla których $b_i = 0$.

Przykład 3.3.3. Rozpatrzmy jeszcze raz problem podany w przykładzie 3.3.1. Jeśli zastosujemy powyższą uwagę, otrzymamy następujący ciąg tablic sympleksowych:

| | 1 | $-x_1$ | $-x_2$ | $-x_3$ | $-x_4$ | |
|---------------|--|--------|----------------|--------|--------|---------|
| 0 | $-\frac{3}{4}$ | 20 | $-\frac{1}{2}$ | 6 | | $= z$ |
| ε | $\frac{1}{4}$ | -8 | -1 | 9 | | $= u_1$ |
| ε | $\frac{1}{2}$ | -12 | $-\frac{1}{2}$ | 3 | | $= u_2$ |
| 1 | 0 | 0 | 1 | 0 | | $= u_3$ |

| | 1 | $-u_2$ | $-x_2$ | $-x_3$ | $-x_4$ | |
|---------------|---------------|----------------|---|----------------|----------------|---------|
| | 0 | $\frac{3}{2}$ | 2 | $-\frac{5}{4}$ | $\frac{21}{2}$ | $= z$ |
| ε | | $-\frac{1}{2}$ | -2 | $-\frac{3}{4}$ | $\frac{15}{2}$ | $= u_1$ |
| ε | 2 | | -24 | -1 | 6 | $= x_1$ |
| 1 | 0 | 0 | 1 | 0 | | $= u_3$ |
| | | | | | | |
| | 1 | $-u_2$ | $-x_2$ | $-u_3$ | $-x_4$ | |
| | $\frac{5}{4}$ | $\frac{3}{2}$ | 2 | $\frac{5}{4}$ | $\frac{21}{2}$ | $= z$ |
| $\frac{3}{4}$ | \times | \times | \times | \times | \times | $= u_1$ |
| 1 | \times | \times | \times | \times | \times | $= x_1$ |
| 1 | \times | \times | \times | \times | \times | $= x_3$ |
| | | | | | | |

Widzimy, że ostatnia tablica odpowiada rozwiązaniu optymalnemu $x^* = (1, 0, 1, 0)^\top$.

3.4. Zrewidowana metoda sympleksowa

Zwróćmy uwagę na to, że w metodzie sympleksowej przekształcenia tablicy sympleksowej polegają na aktualizacji macierzy niebazowej A_N (pivotyzacje). Jeśli jednak popatrzymy na postać zmiennych bazowych (2.3) i odpowiadającą im wartość funkcji celu (2.4), to widzimy, że do ich wyznaczenia wystarczy aktualizacja odwrotności macierzy bazowej A_B odpowiadająca aktualizacji bazy. W przypadku, gdy liczba ograniczeń jest dużo mniejsza od liczby zmiennych, macierz A_B jest dużo mniejsza niż macierz A_N . Wówczas wygodniej jest aktualizować macierz A_B^{-1} . W rezultacie otrzymujemy tak zwaną *zrewidowaną metodę sympleksową*, którą teraz opiszemy dokładniej.

Rozważmy ZPL w postaci standardowej

$$\begin{array}{ll}
 \text{maksymalizować} & z = c^\top x \\
 \text{przy ograniczeniach} & Ax = b \\
 & x \geq 0,
 \end{array}$$

dla której macierz A typu $m \times n$ ma pełen rząd wierszowy ($r(A) = m$). Oznaczmy $A = \begin{bmatrix} A_B & A_N \end{bmatrix}$, $x^\top = \begin{bmatrix} x_B^\top & x_N^\top \end{bmatrix}$ i $c^\top = \begin{bmatrix} c_B^\top & c_N^\top \end{bmatrix}$. Długa postać tablicy sympleksowej przedstawia się wówczas następująco

| z | x_B^\top | x_N^\top | |
|-----|------------|---------------------------------------|-------------------------|
| 1 | 0 | $-(c_N^\top - c_B^\top A_B^{-1} A_N)$ | $= c_B^\top A_B^{-1} b$ |
| 0 | I | $A_B^{-1} A_N$ | $= A_B^{-1} b$ |

Przy każdej wymianie zmiennej bazowej aktualizowane są więc wszystkie elementy tej tablicy. Zauważmy jednak, że wystarczy aktualizować A_B^{-1} , $c_B^\top A_B^{-1}$, $A_B^{-1} b$ i $c_B^\top A_B^{-1} b$. Na ich podstawie można w razie potrzeby wyznaczyć wszystkie elementy tej tablicy. W szczególności, jeśli liczba ograniczeń m jest dużo mniejsza niż liczba zmiennych n , to przy takiej aktualizacji zaoszczędzone zostaną operacje mnożenia przez macierz A_N . Jeśli więc przy wymianie zmiennej bazowej ograniczamy się do aktualizacji wielkości A_B^{-1} , $c_B^\top A_B^{-1}$, $A_B^{-1} b$ i $c_B^\top A_B^{-1} b$, to realizowana jest wówczas tzw. *zrewidowana metoda sympleksowa*. Należy jednak podkreślić, że jest to metoda równoważna omawianej do tej pory metodzie sympleksowej (w obu metodach generowane są te same kolejne rozwiązania bazowe).

Przedstawimy teraz, jak można aktualizować wielkości A_B^{-1} , $c_B^\top A_B^{-1}$, $A_B^{-1} b$ i $c_B^\top A_B^{-1} b$. Rozważmy w tym celu macierz \bar{A} typu $(m+2) \times (m+2)$ przedstawioną w postaci blokowej

$$\bar{A} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & c_B^\top \\ b & 0 & A_B \end{bmatrix}.$$

Zauważmy, że jest ona nieosobliwa i, że

$$\bar{A}^{-1} = \begin{bmatrix} -1 & 0 & 0 \\ c_B^\top A_B^{-1} b & -1 & c_B^\top A_B^{-1} \\ A_B^{-1} b & 0 & A_B^{-1} \end{bmatrix}.$$

Potrzebne elementy tablicy sympleksowej uzyskamy wówczas mnożąc macierz \bar{A}^{-1} przez macierz \hat{A} typu $(m+2) \times (n-m+1)$ przedstawioną w postaci blokowej

$$\hat{A} = \begin{bmatrix} 1 & 0 \\ 0 & c_N^\top \\ 0 & A_N \end{bmatrix}.$$

Mamy bowiem

$$\bar{A}^{-1} \hat{A} = \begin{bmatrix} -1 & 0 \\ c_B^\top A_B^{-1} b & c_N^\top - c_B^\top A_B^{-1} A_N \\ A_B^{-1} b & A_B^{-1} A_N \end{bmatrix}.$$

Reguły wyboru elementu głównego są takie same, jak w zwykłej metodzie sympleksowej. Jako kolumnę główną można wybrać kolumnę j -tą, gdzie j jest takie, że j -ta współrzędna wektora $-(c_N^\top - c_B^\top A_B^{-1} A_N)$ jest ujemna. Kolumną główną jest więc j -ta kolumna macierzy

$$\begin{bmatrix} c_N^\top - c_B^\top A_B^{-1} A_N \\ A_B^{-1} A_N \end{bmatrix}.$$

Zapiszmy ją podobnie jak w ustępie 3.2.2 w postaci $(\alpha_{0j}, \alpha_{1j}, \dots, \alpha_{mj})^\top$. Również podobnie jak w zwykłej metodzie sympleksowej, jako numer wier-

sza głównego można wybrać i takie, że $\frac{\alpha_{i0}}{\alpha_{ij}} = \min\{\frac{\alpha_{k0}}{\alpha_{kj}} : \alpha_{kj} > 0\}$. Wymianie elementu głównego odpowiada teraz aktualizacja macierzy \bar{A}^{-1} . Załóżmy, że wystarczy w tym celu zaktualizować macierz A_B^{-1} . Niech B^+ będzie zaktualizowaną bazą, tzn. $B^+ = B \cup \{j\} \setminus \{i\}$. Wówczas macierz A_{B^+} powstaje z macierzy A_B przez dodanie tzw. *poprawki rzędu pierwszego*

$$A_{B^+} = A_B + (a_j - a_i)e_i^\top,$$

gdzie a_i oznacza kolumnę macierzy A odpowiadającą zmiennej usuwanej z bazy, zaś a_j – kolumnę macierzy A odpowiadającą zmiennej wprowadzanej do bazy. Macierz $A_{B^+}^{-1}$ można wyznaczyć korzystając z tzw. *formuły Shermana–Morrisona*

$$A_{B^+}^{-1} = A_B^{-1} - \frac{A_B^{-1}(a_j - a_i)e_i^\top A_B^{-1}}{e_i^\top A_B^{-1}a_j}$$

(porównaj podręcznik A. Kiełbasińskiego i H. Schwetlicka [15]). Zwróćmy uwagę na fakt, że $e_i^\top A_B^{-1}a_j \neq 0$, gdyż wielkość ta jest elementem głównym. Pozostawiamy czytelnikowi sprawdzenie, że powyższy wzór opisuje rzeczywiście macierz odwrotną do A_{B^+} . Istnieją również inne możliwości aktualizacji macierzy \bar{A}^{-1} . Zainteresowanych odsyłamy do książki P. Kalla [13].

ROZDZIAŁ 4

Dualizm w programowaniu liniowym

Matematyka jest sztuką nadawania różnym rzeczom tych samych nazw.

[H. Poincaré]

4.1. Definicja i przykłady

Zanim zdefiniujemy zadanie dualne do ZPL i omówimy jego związki z wyjściowym zadaniem wróćmy do przykładu 2.1.1. Problem tam przedstawiony ma następującą postać:

$$\begin{array}{ll} \text{maksymalizować} & x_1 + 2x_2 \\ \text{przy ograniczeniach} & \begin{array}{ll} x_1 + x_2 & \leq 100 \\ 6x_1 + 9x_2 & \leq 720 \\ x_2 & \leq 60 \\ x_1, x_2 & \geq 0 \end{array} \end{array}$$

Jest jasne, że jeśli punkt $x = (x_1, x_2)^\top$ jest rozwiązaniem dopuszczalnym tego zadania, to optymalna wartość funkcji celu f^* jest równa co

najmniej wielkości $c^\top x = x_1 + 2x_2$. Na przykład, dla punktu dopuszczalnego $x = (60, 40)^\top$ otrzymujemy $c^\top x = 140 \leq f^*$. Nie wiemy jednak, na ile dobre jest to oszacowanie. Moglibyśmy je ocenić, gdybyśmy również umieli dobrze oszacować f^* z góry. W celu wyznaczenia takiego oszacowania pomnożmy pierwsze ograniczenie przez $1/2$, drugie – przez $1/12$ i dodajmy je do trzeciego ograniczenia. Otrzymamy wówczas nierówność $x_1 + 9/4x_2 \leq 170$, a ponieważ $x_1 + 2x_2 \leq x_1 + 9/4x_2$, więc wnioskujemy stąd, że wartość funkcji celu w tym zadaniu nie przekracza 170. Czynność taką możemy przeprowadzić również dla innego układu współczynników, przez które mnożymy odpowiednie ograniczenia. Na przykład, jeśli pomnożymy pierwsze ograniczenie przez $1/2$, drugie – przez $1/12$, trzecie zaś przez $3/4$ i dodamy otrzymane nierówności do siebie, to otrzymamy nierówność $x_1 + 2x_2 \leq 155$, która musi być również spełniona, jeśli tylko spełnione są ograniczenia wyjściowego zadania. Widzimy więc, że wielkość 155 jest lepszym niż poprzednie ograniczeniem z góry optymalnej wartości funkcji celu. Z przeprowadzonych rozważań wynika, że optymalna wartość funkcji celu mieści się w przedziale $[140, 155]$. Zastanówmy się, czy można ten przedział jeszcze bardziej zawężyć? Aby odpowiedzieć na to pytanie, pomnożmy pierwsze ograniczenie przez y_1 , drugie przez y_2 zaś trzecie przez y_3 , przy czym $y_1, y_2, y_3 \geq 0$, i dodajmy otrzymane nierówności do siebie. Otrzymamy wówczas nierówność

$$(y_1 + 6y_2)x_1 + (y_1 + 9y_2 + y_3)x_2 \leq 100y_1 + 720y_2 + 60y_3.$$

Zauważmy, że jeśli tylko spełnione będą nierówności $y_1 + 6y_2 \geq 1$ oraz $y_1 + 9y_2 + y_3 \geq 2$, to w konsekwencji spełniona będzie również nierówność

$$x_1 + 2x_2 \leq 100y_1 + 720y_2 + 60y_3.$$

Najlepsze ograniczenie z dołu dla prawej strony ostatniej nierówności otrzymamy dla wektora $y = (y_1, y_2, y_3)^\top$ będącego rozwiązaniem następującego ZPL:

$$\begin{array}{ll} \text{minimalizować} & 100y_1 + 720y_2 + 60y_3 \\ \text{przy ograniczeniach} & \begin{array}{ll} y_1 + 6y_2 & \geq 1 \\ y_1 + 9y_2 + y_3 & \geq 2 \\ y_1, y_2, y_3 & \geq 0. \end{array} \end{array} \quad (4.1)$$

Ograniczenie to będzie również najlepszym z możliwych do osiągnięcia w podany sposób górnym ograniczeniem optymalnej wartości funkcji celu f^* . W istocie optymalna wartość funkcji celu dla zadania (4.1) wynosi 150 (jest ona osiągnięta dla $y = (0, 1/6, 1/2)^\top$) i jest równa optymalnej wartości funkcji celu dla wyjściowego zadania. W dalszej części tego rozdziału pokażemy, że własność ta przysługuje dowolnemu ograniczonemu ZPL.

Przejdźmy wobec tego do sytuacji ogólniejszej. Dane jest zadanie programowania liniowego w postaci klasycznej

$$\begin{array}{ll} \text{maksymalizować} & c^\top x \\ \text{przy ograniczeniach} & \begin{array}{l} Ax \leq b \\ x \geq 0, \end{array} \end{array} \quad (4.2)$$

gdzie $c, x \in \mathbb{R}^n$, A jest macierzą typu $m \times n$ i $b \in \mathbb{R}^m$.

Definicja 4.1.1. Zadanie programowania liniowego

$$\begin{array}{ll} \text{minimalizować} & b^\top y \\ \text{przy ograniczeniach} & \begin{array}{l} A^\top y \geq c \\ y \geq 0, \end{array} \end{array} \quad (4.3)$$

gdzie $y \in \mathbb{R}^m$, nazywa się *zadaniem dualnym* do zadania (4.2). Zadanie (4.2) nazywa się wówczas *zadaniem pierwotnym*.

Uwaga 4.1.2. Zadanie (4.2) jest zadaniem dualnym do zadania (4.3). W celu pokazania tego faktu należy zapisać zadanie (4.3) w postaci klasycznej (jako zadanie maksymalizacji), wyznaczyć zadanie do niego dualne i zauważyć, że jest nim (4.2). Szczegóły pozostawiamy czytelnikowi. Oba zadania są więc wzajemnie dualne.

Rozważmy teraz zadanie programowania liniowego w postaci standardowej:

$$\begin{array}{ll} \text{maksymalizować} & c^\top x \\ \text{przy ograniczeniach} & \begin{array}{l} Ax = b \\ x \geq 0. \end{array} \end{array} \quad (4.4)$$

Można je przedstawić w równoważnej mu postaci klasycznej:

$$\begin{array}{ll} \text{maksymalizować} & c^\top x \\ \text{przy ograniczeniach} & \begin{bmatrix} A \\ -A \end{bmatrix} x \leq \begin{bmatrix} b \\ -b \end{bmatrix} \\ & x \geq 0. \end{array}$$

Zgodnie z definicją 4.1.1 zadaniem dualnym do tego ostatniego jest zadanie

$$\begin{array}{ll} \text{minimalizować} & b^\top(v - w) \\ \text{przy ograniczeniach} & A^\top(v - w) \geq c \\ & v, w \geq 0, \end{array}$$

gdzie $v, w \in \mathbb{R}^m$. Po wprowadzeniu nowej zmiennej $y = v - w$ ostatnie zadanie przyjmie postać

$$\begin{array}{ll} \text{minimalizować} & b^\top y \\ \text{przy ograniczeniach} & A^\top y \geq c. \end{array} \quad (4.5)$$

Zauważmy, że na y nie nakłada się ograniczenia nieujemności. Tak więc zadanie (4.5) jest dualne do zadania (4.4). Można również łatwo pokazać, że zadanie (4.4) jest dualne do (4.5). Dowód tego faktu pozostawiamy czytelnikowi. Oba zadania (4.4) i (4.5) są więc wzajemnie dualne.

Przykład 4.1.3. Powróćmy do zagadnienia analizy działalności gospodarczej (przykład 2.1.2), które daje się przedstawić jako zadanie programowania liniowego w postaci klasycznej:

$$\begin{array}{ll} \text{maksymalizować} & c_1x_1 + c_2x_2 + \dots + c_nx_n \\ \text{przy ograniczeniach} & \begin{array}{ll} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n & \leq b_2 \\ \dots & \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n & \leq b_m \\ x_1, \dots, x_n & \geq 0, \end{array} \end{array}$$

gdzie x jest profilem produkcji, b wektorem zasobów surowcowych, c wektorem zysków jednostkowych, zaś element a_{ij} macierzy A wskazuje na ilość

jednostek i -tego surowca potrzebną do produkcji jednostki j -tego towaru. Zadaniem dualnym do powyższego jest zadanie

$$\begin{array}{ll}
 \text{minimalizować} & b_1 y_1 + b_2 y_2 + \dots + b_m y_m \\
 \text{przy ograniczeniach} & a_{11} y_1 + a_{21} y_2 + \dots + a_{m1} y_m \geq c_1 \\
 & a_{12} y_1 + a_{22} y_2 + \dots + a_{m2} y_m \geq c_2 \\
 & \dots \\
 & a_{1n} y_1 + a_{2n} y_2 + \dots + a_{mn} y_n \geq c_n \\
 & y_1, y_2, \dots, y_m \geq 0.
 \end{array}$$

Ostatnie zadanie ma następującą interpretację ekonomiczną. Niech y_i będzie ceną jednostki i -tego surowca, $i = 1, \dots, m$. Wówczas j -te ograniczenie zadania dualnego mówi, że udział kosztów przy produkcji jednostki j -tego towaru wynosi co najmniej c_j , $i = 1, \dots, m$, $j = 1, \dots, m$. Natomiast zadanie polega na znalezieniu takiego wektora cen surowcowych, dla którego całkowita wartość użytych do produkcji surowców jest minimalna.

4.2. Twierdzenia o dualności

Twierdzenie 4.2.1 (słabe twierdzenie o dualności). *Jeśli $x \in \mathbb{R}^n$ jest rozwiązaniem dopuszczalnym zadania pierwotnego (4.2) i $y \in \mathbb{R}^m$ jest rozwiązaniem dopuszczalnym zadania dualnego (4.3), to zachodzi nierówność*

$$c^\top x \leq b^\top y.$$

Dowód. Niech x, y będą rozwiązaniami dopuszczalnymi odpowiednio dla zadania pierwotnego i dualnego, tzn. $x, y \geq 0$, $Ax \leq b$ i $A^\top y \geq c$. Wówczas zachodzą następujące nierówności i równości:

$$c^\top x \leq (A^\top y)^\top x = y^\top Ax \leq y^\top b = b^\top y.$$

□

Wniosek 4.2.2. *Jeśli obydwa zadania (4.2) i (4.3) są dopuszczalne, tzn.*

$$X = \{x : Ax \leq b, x \geq 0\} \neq \emptyset$$

i

$$Y = \{y : A^T y \geq c, y \geq 0\} \neq \emptyset,$$

to posiadają one rozwiązania optymalne, odpowiednio x^* i y^* .

Wniosek 4.2.3. *Jeśli x i y są rozwiązaniami dopuszczalnymi odpowiednio dla zadania pierwotnego (4.2) i dualnego (4.3) i zachodzi równość $c^T x = b^T y$, to x jest rozwiązaniem optymalnym zadania pierwotnego i y jest rozwiązaniem optymalnym zadania dualnego.*

Ćwiczenie 4.2.4. Dane jest zadanie

$$\begin{array}{ll} \text{maksymalizować} & 2x_1 + x_2 \\ \text{przy ograniczeniach} & \begin{array}{ll} x_1 + x_2 & \leq 6 \\ -x_1 + x_2 & \leq 5 \\ x_1 - 2x_2 & \leq 2 \\ x_1, x_2 & \geq 0 \end{array} \end{array}$$

i zadanie do niego dualne

$$\begin{array}{ll} \text{minimalizować} & 6y_1 + 5y_2 + 2y_3 \\ \text{przy ograniczeniach} & \begin{array}{ll} y_1 - y_2 + y_3 & \geq 2 \\ y_1 + y_2 - 2y_3 & \geq 1 \\ y_1, y_2, y_3 & \geq 0. \end{array} \end{array}$$

Sprawdzić, czy para $(x^{*\top}; y^{*\top})^\top = (\frac{14}{3}, \frac{4}{3}, \frac{5}{3}, 0, \frac{1}{3})^\top$ jest rozwiązaniem zadania pierwotnego i dualnego.

Wniosek 4.2.5 *Jeśli zadanie pierwotne (4.2) jest nieograniczone, to zadanie dualne (4.3) nie posiada rozwiązania dopuszczalnego. Podobnie, jeśli zadanie dualne (4.3) jest nieograniczone, to zadanie pierwotne (4.2) nie posiada rozwiązania dopuszczalnego.*

Twierdzenie 4.2.6 (mocne twierdzenie o dualności). *Jeśli zadanie pierwotne (4.4) posiada rozwiązanie optymalne x^* , to zadanie dualne (4.5) posiada rozwiązanie optymalne y^* i zachodzi równość*

$$c^T x^* = b^T y^*.$$

Dowód. Niech x^* będzie rozwiązaniem optymalnym zadania (4.4). Wówczas istnieje bazowe rozwiązanie optymalne \bar{x} (funkcja liniowa określona na zbiorze wypukłym osiąga swoje maksimum w punktach ekstremalnych). Niech $A = \begin{bmatrix} A_B & A_N \end{bmatrix}$ i $\bar{x} = (\bar{x}_B^\top, 0)^\top$, $\bar{x}_B = A_B^{-1}b$, gdzie A_B jest macierzą bazową. Niech $\bar{y} = A_B^{-\top}c_B$, gdzie $c = (c_B^\top, c_N^\top)^\top$ i $A_B^{-\top}$ oznacza macierz $(A_B^\top)^{-1}$. Wówczas \bar{y} jest rozwiązaniem dopuszczalnym problemu dualnego (4.5), bo

$$A^\top \bar{y} = \begin{bmatrix} A_B^\top \\ A_N^\top \end{bmatrix} A_B^{-\top} c_B = \begin{bmatrix} c_B^\top \\ (A_B^{-1} A_N)^\top c_B \end{bmatrix} \geq \begin{bmatrix} c_B^\top \\ c_N^\top \end{bmatrix} = c.$$

Nierówność wynika z następującego rozumowania: Niech f^* będzie optymalną wartością funkcji celu zadania (4.4). Mamy więc dla $x \in X$

$$\begin{aligned} f^* &= c^\top \bar{x} = (c_B^\top, c_N^\top) \begin{bmatrix} \bar{x}_B \\ 0 \end{bmatrix} \\ &\geq c^\top x = (c_B^\top, c_N^\top) \begin{bmatrix} x_B \\ x_N \end{bmatrix} \\ &= (c_B^\top, c_N^\top) \begin{bmatrix} A_B^{-1}b - A_B^{-1}A_N x_N \\ x_N \end{bmatrix} \\ &= c_B^\top A_B^{-1}b - c_B^\top A_B^{-1}A_N x_N + c_N^\top x_N \\ &= f^* - c_B^\top A_B^{-1}A_N x_N + c_N^\top x_N, \end{aligned}$$

czyli

$$c_B^\top A_B^{-1}A_N x_N \geq c_N^\top x_N$$

dla dowolnego x_N . Biorąc więc $x_N = e_i$, $i = m+1, \dots, n$, otrzymamy

$$c_B^\top A_B^{-1}A_N \geq c_N^\top.$$

Ponadto

$$c^\top \bar{x} = c_B^\top \bar{x}_B = c_B^\top A_B^{-1}b = (A_B^{-\top} c_B)^\top b = \bar{y}^\top b = b^\top \bar{y}$$

i na mocy wniosku 4.2.3, punkt \bar{y} jest rozwiązaniem optymalnym zadania (4.5). \square

Uwaga 4.2.7. Mocne twierdzenie o dualności prawdziwe jest również dla postaci klasycznej (4.2) i (4.3).

Wniosek 4.2.8. *Jeśli zadania pierwotne i dualne są dopuszczalne ($X \neq \emptyset$ i $Y \neq \emptyset$), to posiadają one rozwiązania optymalne.*

Wniosek 4.2.9. *Niech x i y będą rozwiązaniami dopuszczalnymi odpowiednio dla zadania pierwotnego (4.2) i dualnego (4.3). Wówczas x i y są rozwiązaniami optymalnymi tych zadań wtedy i tylko wtedy, gdy*

$$c^\top x = b^\top y.$$

Dowód. Wniosek wynika z mocnego twierdzenia o dualności i z wniosku 4.2.3. \square

Uwaga 4.2.10. a) Zadaniu pierwotnemu i dualnemu odpowiada w istocie ta sama tablica sympleksowa. Zadania te można bowiem przedstawić jako układy równań:

(pierwotne ZPL)

$$\begin{aligned} z &= c^\top x & i \\ b - Ax &= u \\ x, u &\geq 0 \end{aligned}$$

(dualne ZPL)

$$\begin{aligned} z &= b^\top y \\ -c + A^\top y &= v \\ y, v &\geq 0. \end{aligned}$$

Wspólna początkowa tablica sympleksowa obu zadań ma postać:

| | | | | | | | | |
|-----------|-------|--------------------------------------|----------|--------|----------|--------|---------|-------------|
| | | zmienne niebazowe pierwotnego ZPL | | | | | | |
| | | 1 | $-x_1$ | ... | $-x_n$ | | | |
| | | 1 | 0 | $-c_1$ | ... | $-c_n$ | $= z$ | |
| zmienne | y_1 | b_1 | a_{11} | ... | a_{1n} | | $= u_1$ | zmienne |
| niebazowe | ... | ... | ... | ... | ... | | ... | bazowe |
| dualnego | y_m | b_m | a_{m1} | ... | a_{mn} | | $= u_m$ | pierwotnego |
| ZPL | | | | | | | | ZPL |
| | | | | | | | | |
| | | z | v_1 | ... | v_n | | | |
| | | zmienne bazowe dualnego ZPL | | | | | | |

Tablicy tej odpowiada rozwiązanie bazowe zadania pierwotnego $(x, u) = (0, b)$ i rozwiązanie bazowe zadania dualnego $(y, v) = (0, -c)$ (oczywiście w postaci kanonicznej dla obu zadań).

b) Przy piwotyzacji, w której a_{ij} jest elementem głównym, wymienione zostają zmienne u_i z x_j (zadanie pierwotne) i y_i z v_j (zadanie dualne).

c) W każdej tablicy zmienna dualna y_i znajduje się naprzeciw zmiennej uzupełniającej zadania pierwotnego u_i , $i = 1, \dots, m$, oraz zmienna pierwotna x_j znajduje się naprzeciw zmiennej uzupełniającej zadania dualnego v_j , $j = 1, \dots, n$.

d) Rozwiązania bazowe zadania pierwotnego (x, u) i zadania dualnego (y, v) odpowiadające tej samej tablicy sympleksowej posiadają następującą własność: $y_i u_i = 0$, $i = 1, \dots, m$, i $x_j v_j = 0$, $j = 1, \dots, n$, (jest to tzw. *własność komplementarności*).

e) Jeśli zerowy wiersz tablicy sympleksowej (wektor $-c$ w przypadku, gdy jest to początkowa tablica) jest nieujemny, to tablica ta przedstawia dopuszczalne rozwiązanie zadania dualnego. Wówczas o rozwiązaniu zadania pierwotnego przedstawionego przez tę tablicę mówimy, że jest *dualnie dopuszczalne*. Jeśli ponadto przynajmniej jeden z elementów tego wiersza jest równy zero ($c_j = 0$ dla pewnego j , w przypadku gdy jest to początkowa tablica), to rozwiązanie zadania dualnego przedstawionego przez tę tablicę

jest zdegenerowane. Wówczas o rozwiązaniu zadania pierwotnego przedstawionego przez tę tablicę mówimy, że jest *dualnie zdegenerowane*.

f) Jeśli zerowa kolumna tablicy sympleksowej (wektor b w przypadku, gdy jest to początkowa tablica) jest nieujemna, to tablica ta przedstawia rozwiązanie dopuszczalne zadania pierwotnego. Mówi się również, że rozwiązanie to jest *pierwotnie dopuszczalne*. Jeśli ponadto przynajmniej jeden z elementów tej kolumny jest równy zero ($b_i = 0$ dla pewnego i , w przypadku gdy jest to początkowa tablica), to rozwiązanie zadania pierwotnego przedstawionego przez tę tablicę jest zdegenerowane. Mówi się również, że rozwiązanie to jest *pierwotnie zdegenerowane*.

g) Jeśli zerowy wiersz i zerowa kolumna tablicy sympleksowej są nieujemne ($-c \geq 0, b \geq 0$ w przypadku, gdy jest to początkowa tablica), to tablica ta przedstawia rozwiązanie optymalne zadania pierwotnego i rozwiązanie optymalne zadania dualnego.

h) Współrzędne y_i^* , $i = 1, \dots, m$, bazowego rozwiązania optymalnego y^* zadania dualnego nazywają się również *cenami ukrytymi* (z dowodu mocnego twierdzenia o dualności wynika, że ta definicja jest równoważna definicji cen ukrytych podanej w ustępie 2.2). Pokazują one, jak szybko wzrasta funkcja celu zadania pierwotnego w zależności od odpowiednich współrzędnych b_i , $i = 1, \dots, m$, wektora b : jeśli b_i zwiększy się o jednostkę, to funkcja celu wzrośnie o y_i^* jednostek (o ile baza odpowiadająca bazowemu rozwiązaniu optymalnemu nie zmieni się po tej operacji). Istotnie, niech y^* będzie bazowym rozwiązaniem optymalnym zadania dualnego. Dalej, niech

$$b' = b + e_i = (b_1, \dots, b_{i-1}, b_i + 1, b_{i+1}, \dots, b_m)$$

i niech y' będzie bazowym rozwiązaniem optymalnym zadania dualnego do zakłóconego zadania:

$$\begin{array}{ll} \text{maksymalizować} & c^\top x \\ \text{przy ograniczeniach} & Ax \leq b' \\ & x \geq 0. \end{array}$$

Wówczas $y^* = y'$ (obu rozwiązaniom odpowiada ta sama baza). Niech z i z' będą optymalnymi wartościami funkcji celu odpowiednio dla zadania

wyjściowego i zakłóconego. Otrzymamy następujące równości

$$z' = b'^\top y' = b'^\top y^* = b^\top y^* + y_i^* = z + y_i^*.$$

Widzimy więc, że y_i^* wyraża szybkość przyrostu optymalnej wartości funkcji celu w zależności od zmiany i -tej współrzędnej wektora b . Własność ta ma ważne znaczenie ekonomiczne.

Przykład 4.2.11. Dane jest zagadnienie analizy działalności gospodarczej, któremu odpowiada następujące zadanie programowania liniowego:

$$\begin{array}{ll} \text{maksymalizować} & 2x_1 + 4x_2 + 3x_3 \\ \text{przy ograniczeniach} & \begin{array}{ll} x_1 + 4x_2 + 3x_3 & \leq 240 \\ 2x_1 + x_2 + 5x_3 & \leq 300 \\ x_1 + x_2 + x_3 & \leq 200 \\ x_1, x_2, x_3 & \geq 0. \end{array} \end{array}$$

Wyznamy rozwiązania optymalne zadań pierwotnego i dualnego i podamy interpretację ekonomiczną otrzymanych wyników.

Zadanie dualne ma postać

$$\begin{array}{ll} \text{minimalizować} & 240y_1 + 300y_2 + 200y_3 \\ \text{przy ograniczeniach} & \begin{array}{ll} y_1 + 2y_2 + y_3 & \geq 2 \\ 4y_1 + y_2 + y_3 & \geq 4 \\ 3y_1 + 5y_2 + y_3 & \geq 3 \\ y_1, y_2, y_3 & \geq 0. \end{array} \end{array}$$

Początkowa tablica sympleksowa ma dla obu problemów postać

| | 1 | $-x_1$ | $-x_2$ | $-x_3$ | |
|-------|-------|--------|--------|--------|---------|
| 1 | 0 | -2 | -4 | -3 | $= z$ |
| y_1 | 240 | 1 | 4 | 3 | $= u_1$ |
| y_2 | 300 | 2 | 1 | 5 | $= u_2$ |
| y_3 | 200 | 1 | 1 | 1 | $= u_3$ |
| | Π | Π | Π | Π | |
| | z | v_1 | v_2 | v_3 | |

Po dwóch piwotyzacjach otrzymamy tablicę

| | 1 | $-u_2$ | $-u_1$ | $-x_3$ | |
|-------|------------------|---------------|---------------|----------------|---------|
| 1 | $377\frac{1}{7}$ | $\frac{4}{7}$ | $\frac{6}{7}$ | $\frac{17}{7}$ | $= z$ |
| v_2 | $25\frac{5}{7}$ | \times | \times | \times | $= x_2$ |
| v_1 | $137\frac{1}{7}$ | \times | \times | \times | $= x_1$ |
| y_3 | $37\frac{1}{7}$ | \times | \times | \times | $= u_3$ |
| | \parallel | \parallel | \parallel | \parallel | |
| | z | y_2 | y_1 | v_3 | |

Tablica ta przedstawia rozwiązanie optymalne zadania pierwotnego $x^* = (137\frac{1}{7}, 25\frac{5}{7}, 0)$ i rozwiązanie optymalne zadania dualnego $y^* = (\frac{6}{7}, \frac{4}{7}, 0)$. Optymalna wartość funkcji celu dla obu zadań wynosi $z^* = 377\frac{1}{7}$. Współrzędne wektora y^* są cenami ukrytymi czynników produkcyjnych (surowców) R_1, R_2 i R_3 . Jeśli dostępna ilość pierwszego surowca zwiększy się o jednostkę, to zysk producenta wzrośnie o $\frac{6}{7}$ jednostek. Podobna interpretacja ekonomiczna dotyczy drugiej współrzędnej wektora y^* : producentowi opłaca się kupić pewną dodatkową ilość drugiego surowca za cenę co najwyżej $\frac{4}{7}$ jednostek pieniężnych za jego jednostkę (lub innymi słowy: jeśli cena ta wyniesie co najwyżej $\frac{4}{7}$, to zysk producenta nie zmniejszy się). Zwiększenie dostępnej ilości trzeciego surowca nie wpłynie na zmianę zysku. Tak więc jeśli cena trzeciego surowca na rynku jest dodatnia, to zakup tego surowca spowoduje zmniejszenie zysku. Przyczyną jest niepełne wykorzystanie trzeciego surowca przy optymalnym profilu produkcji: pozostało go $37\frac{1}{7}$ jednostek, bo tyle wynosi wartość zmiennej uzupełniającej u_3^* . Zwróćmy przy okazji uwagę na własność komplementarności: $y_3^* u_3^* = 0$).

Przykład 4.2.12. Powróćmy teraz do problemu podanego w przykładzie 2.1.1 i rozwiązanego w przykładzie 3.2.5. Tablica sympleksowa odpowiada-

jąca rozwiązaniu optymalnemu ma postać

| | 1 | $-u_2$ | $-u_3$ | |
|-------|-----|----------------|----------------|---------|
| 1 | 150 | $\frac{1}{6}$ | $\frac{1}{2}$ | $= z$ |
| y_1 | 10 | $-\frac{1}{6}$ | $\frac{1}{2}$ | $= u_1$ |
| v_1 | 30 | $\frac{1}{6}$ | $-\frac{3}{2}$ | $= x_1$ |
| v_2 | 60 | 0 | 1 | $= x_2$ |
| | z | y_2 | y_3 | |

Optymalne rozwiązanie wynosi dla zadania pierwotnego $x^* = (30, 60)$, zaś dla zadania dualnego $y^* = (0, \frac{1}{6}, \frac{1}{2})$. Wielkości te mają następującą interpretację ekonomiczną. Ogrodnik uzyska największy zysk 150 zł, jeśli przeznaczy 30 m² swojego ogrodu na róże i 60 m² na goździki. Zakup dodatkowej powierzchni pod uprawę zysk ten zmniejszy ($y_1^* = 0$; ma to związek z dodatnią wartością zmiennej uzupełniającej $u_1^* = 10$ – tyle metrów kwadratowych powierzchni ogrodu jest niewykorzystanych przy optymalnym rozwiązaniu). Ogrodnikowi opłaci się zainwestować więcej niż 720 zł, o ile zdobycie dodatkowej złotówki (kredyt) kosztować go będzie co najwyżej $\frac{1}{6}$ zł. Opłaci mu się również przeznaczyć więcej niż 60 m² na goździki, o ile koszt zagospodarowania 1 m² dodatkowej powierzchni wyniesie co najwyżej $\frac{1}{2}$ zł. Zwróćmy uwagę na własność komplementarności: $y_i^* u_i^* = 0, i = 1, 2, 3$.

Twierdzenie 4.2.13 (o komplementarności). *Niech x i y będą rozwiązaniami dopuszczalnymi odpowiednio dla zadania pierwotnego (4.2) i dualnego (4.3). Wówczas x i y są rozwiązaniami optymalnymi wtedy i tylko wtedy, gdy*

$$y^\top (Ax - b) = 0 \quad (4.6)$$

i

$$x^\top (c - A^\top y) = 0. \quad (4.7)$$

Dowód. Niech x i y będą rozwiązaniami dopuszczalnymi odpowiednio dla zadania pierwotnego (4.2) i dualnego (4.3). Wówczas, podobnie jak w dowodzie słabego twierdzenia o dualności, zachodzą następujące

nierówności i równości:

$$c^\top x \leq (A^\top y)^\top x = y^\top Ax \leq y^\top b = b^\top y. \quad (4.8)$$

Jeśli x i y są rozwiązaniami optymalnymi, to $c^\top x = b^\top y$ na mocy mocnego twierdzenia o dualności. W konsekwencji

$$c^\top x = (A^\top y)^\top x = y^\top Ax = y^\top b = b^\top y,$$

czyli $(c - A^\top y)^\top x = 0$ i $y^\top (Ax - b) = 0$. Załóżmy teraz, że zachodzą równości (4.6) i (4.7). Wówczas z (4.8) wynika, że $c^\top x = b^\top y$. W konsekwencji x i y są rozwiązaniami optymalnymi na mocy słabego twierdzenia o dualności. \square

Uwaga 4.2.14. Niech a^i oznacza i -ty wiersz oraz a_j - j -tą kolumnę macierzy A . Jeśli x i y są rozwiązaniami dopuszczalnymi odpowiednio dla zadania pierwotnego (4.2) i dualnego (4.3), to równości (4.6) i (4.7) można zapisać w postaci

$$y_i(a^i x - b_i) = 0 \quad (\text{lub równoważnie} \quad y_i u_i = 0), \quad (4.9)$$

$i = 1, \dots, m$, i

$$x_j(a_j^\top y - c_j) = 0 \quad (\text{lub równoważnie} \quad x_j v_j = 0), \quad (4.10)$$

$j = 1, \dots, n$. Wynika to z faktu, że jeśli suma liczb nieujemnych jest równa zeru, to liczby te są równe zeru.

Wniosek 4.2.15. Niech x i y będą rozwiązaniami dopuszczalnymi odpowiednio dla zadania pierwotnego (4.2) i dualnego (4.3). Wówczas x i y są rozwiązaniami optymalnymi wtedy i tylko wtedy, gdy

$$y_i > 0 \Rightarrow a^i x = b_i \quad (\text{lub równoważnie} \quad a^i x < b_i \Rightarrow y_i = 0),$$

$i = 1, \dots, m$, i

$$x_j > 0 \Rightarrow a_j^\top y = c_j \quad (\text{lub równoważnie} \quad a_j^\top y > c_j \Rightarrow x_j = 0),$$

$j = 1, \dots, n$.

4.3. Dualny algorytm sympleksowy

Zadanie dualne (4.3) można równoważnie przedstawić jako następujące zadanie maksymalizacji:

$$\begin{array}{ll} \text{maksymalizować} & -b^\top y \\ \text{przy ograniczeniach} & -A^\top y \leq -c \\ & y \geq 0. \end{array}$$

Ostatniemu zadaniu odpowiada poniższa tablica sympleksowa

| | | | | |
|---------|-----------|---------|-----------|---------|
| 1 | $-y_1$ | \dots | $-y_m$ | |
| 0 | b_1 | \dots | b_m | $= z$ |
| $-c_1$ | $-a_{11}$ | \dots | $-a_{m1}$ | $= v_1$ |
| \dots | \dots | \dots | \dots | \dots |
| $-c_n$ | $-a_{1n}$ | \dots | $-a_{mn}$ | $= v_n$ |

Przypuśćmy, że tablica ta przedstawia dopuszczalne rozwiązanie bazowe (czyli $c \leq 0$). Jeśli teraz zastosujemy do tej tablicy reguły pivotyzacji opisane w punkcie 3.2.3, to jako kolumnę główną wybierzemy i_0 , dla którego $b_{i_0} < 0$ oraz jako wiersz główny - j_0 , dla którego

$$\frac{-c_{j_0}}{-a_{i_0 j_0}} = \min \left\{ \frac{-c_j}{-a_{i_0 j}} : -a_{i_0 j} > 0 \right\}.$$

Następnie wymieniamy zmienną bazową v_{j_0} ze zmienną niebazową y_{i_0} i przekształcamy odpowiednio tablicę sympleksową. Zastosowaliśmy tu tzw. *pierwotny algorytm sympleksowy* do zadania dualnego. To samo otrzymamy stosując tzw. *dualny algorytm sympleksowy* do zadania pierwotnego. Poniżej przedstawimy ten algorytm przyjmując, że zadaniu pierwotnemu

odpowiada tablica:

| | 1 | $-r_1$ | \dots | $-r_n$ | |
|--|---------------|---------------|---------|---------------|---------|
| | α_{00} | α_{01} | \dots | α_{0n} | $= z$ |
| | α_{10} | α_{11} | \dots | α_{1n} | $= s_1$ |
| | \dots | \dots | \dots | \dots | \dots |
| | α_{m0} | α_{m1} | \dots | α_{mn} | $= s_m$ |
| | | | | | |

oraz że rozwiązanie przedstawione za pomocą tej tablicy jest dualnie dopuszczalne ($\alpha_{0j} \geq 0$ dla $j = 1, \dots, n$).

Algorytm 4.3.1 (dualny sympleksowy)

Krok 1. (*wyбір wiersza głównego*). Wybrać dowolny wiersz $i_0 \geq 1$, dla którego $\alpha_{i_0 0} < 0$ (najczęściej wybiera się i_0 takie, że $\alpha_{i_0 0} = \min\{\alpha_{i_0 0} : \alpha_{i_0 0} < 0\}$). Jeśli brak jest takiego i_0 , to tablica przedstawia rozwiązanie dopuszczalne. (Jeśli dodatkowo $\alpha_{0j} \geq 0$ dla $j = 1, \dots, n$, to rozwiązanie to jest optymalne.)

Krok 2. (*wyбір kolumny głównej*). Wśród wszystkich kolumn $j \geq 1$, takich, że $\alpha_{i_0 j} < 0$, wybrać kolumnę j_0 , dla której

$$\frac{\alpha_{0j_0}}{\alpha_{i_0 j_0}} = \max\left\{\frac{\alpha_{0j}}{\alpha_{i_0 j}} : \alpha_{i_0 j} < 0\right\}.$$

Jeśli brak jest takiej kolumny ($\alpha_{i_0 j} \geq 0$ dla wszystkich $j \geq 1$), to problem pierwotny nie posiada rozwiązania dopuszczalnego.

Krok 3. (*piwotyzacja*). Wymienić zmienną bazową s_{i_0} ze zmienną niebazową r_{j_0} i przetransformować tablicę sympleksową zgodnie z regułą (3.8).

Uwaga 4.3.2. Dualny algorytm sympleksowy dla zadania pierwotnego i pierwotny algorytm sympleksowy dla zadania dualnego są sobie równoważne.

Uwaga 4.3.3. Dualny algorytm sympleksowy nadaje się w szczególności do wyznaczenia rozwiązania dopuszczalnego zadania pierwotnego, dla którego początkowa tablica sympleksowa przedstawia rozwiązanie bazowe, które jest dualnie dopuszczalne. Będzie to wykorzystane przy programowaniu liniowym całkowitoliczbowym.

Przykład 4.3.4. Rozważmy teraz problem diety podany w przykładzie 2.1.3:

$$\begin{array}{ll} \text{minimalizować} & 5x_1 + 7x_2 \\ \text{przy ograniczeniach} & \begin{array}{ll} 2x_1 + x_2 & \geq 6 \\ 2x_1 + 4x_2 & \geq 12 \\ 7x_2 & \geq 4 \\ x_1, x_2 & \geq 0. \end{array} \end{array}$$

Zadanie to można sprowadzić do postaci klasycznej:

$$\begin{array}{ll} \text{maksymalizować} & -5x_1 - 7x_2 \\ \text{przy ograniczeniach} & \begin{array}{ll} -2x_1 - x_2 & \leq -6 \\ -2x_1 - 4x_2 & \leq -12 \\ -7x_2 & \leq -4 \\ x_1, x_2 & \geq 0. \end{array} \end{array}$$

Początkowa tablica sympleksowa ma postać

| | | | |
|-----|--------|--------|---------|
| 1 | $-x_1$ | $-x_2$ | |
| 0 | 5 | 7 | $= z$ |
| -6 | -2 | -1 | $= u_1$ |
| -12 | -2 | -4 | $= u_2$ |
| -4 | 0 | -7 | $= u_3$ |

Ponieważ odpowiada ona rozwiązaniu bazowemu, które jest niedopuszczalne, zadanie możemy rozwiązać przy pomocy dwufazowego algorytmu sym-

pleksowego. Otrzymamy kolejno następujące tablice:

| | | |
|---|---|--|
| $\begin{array}{c ccc c} 1 & -x_1 & -x_2 & & \\ \hline 0 & 5 & 7 & & = z \\ \hline -6 & -2 & -1 & & = u_1 \\ -12 & -2 & -4 & & = u_2 \\ -4 & 0 & -7 & & = u_3 \\ \hline \end{array}$ | , | $\begin{array}{c ccc c} 1 & -x_1 & -u_3 & & \\ \hline -4 & 5 & 1 & & = z \\ \hline -\frac{38}{7} & -2 & -\frac{1}{7} & & = u_1 \\ -\frac{68}{7} & -2 & -\frac{4}{7} & & = u_2 \\ \frac{4}{7} & 0 & -\frac{1}{7} & & = x_2 \\ \hline \end{array}$ |
| $\begin{array}{c ccc c} 1 & -u_2 & -u_3 & & \\ \hline -\frac{198}{7} & \frac{5}{2} & -\frac{3}{7} & & = z \\ \hline \frac{30}{7} & -1 & \frac{3}{7} & & = u_1 \\ \frac{34}{7} & -\frac{1}{2} & \frac{2}{7} & & = x_1 \\ \frac{4}{7} & 0 & -\frac{1}{7} & & = x_2 \\ \hline \end{array}$ | , | $\begin{array}{c ccc c} 1 & -u_2 & -u_1 & & \\ \hline -24 & \frac{3}{2} & 1 & & = z \\ \hline 10 & -\frac{7}{3} & \frac{7}{3} & & = u_3 \\ 2 & \frac{1}{6} & -\frac{2}{3} & & = x_1 \\ 2 & -\frac{1}{3} & \frac{1}{3} & & = x_2 \\ \hline \end{array}$ |

Otrzymaliśmy rozwiązanie optymalne $x^* = (2, 2)$. Istnieje również możliwość wyznaczenia rozwiązania optymalnego stosując dualny algorytm sympleksowy. Otrzymamy wówczas kolejno następujące tablice:

| | | |
|--|---|--|
| $\begin{array}{c ccc c} 1 & -x_1 & -x_2 & & \\ \hline 0 & 5 & 7 & & = z \\ \hline -6 & -2 & -1 & & = u_1 \\ -12 & -2 & -4 & & = u_2 \\ -4 & 0 & -7 & & = u_3 \\ \hline \end{array}$ | , | $\begin{array}{c ccc c} 1 & -x_1 & -u_2 & & \\ \hline 1 & -21 & \frac{3}{2} & \frac{7}{4} & = z \\ \hline -3 & -\frac{3}{2} & -\frac{1}{4} & & = u_1 \\ 3 & \frac{1}{2} & -\frac{1}{4} & & = x_2 \\ 17 & \frac{7}{2} & -\frac{4}{7} & & = u_3 \\ \hline \end{array}$ |
| $\begin{array}{c ccc c} 1 & -u_1 & -u_2 & & \\ \hline -24 & 1 & \frac{3}{2} & & = z \\ \hline 2 & -\frac{2}{3} & \frac{1}{6} & & = x_1 \\ 2 & \frac{1}{3} & -\frac{1}{3} & & = x_2 \\ 10 & \frac{7}{3} & -\frac{7}{3} & & = u_3 \\ \hline \end{array}$ | | |

Otrzymaliśmy to samo rozwiązanie optymalne co poprzednio: $x^* = (2, 2)$.

ROZDZIAŁ 5

O złożoności obliczeniowej ZPL

By dojść do źródła, trzeba płynąć pod prąd.

[S. J. Lec]

5.1. Złożoność obliczeniowa algorytmu sympleksowego

W rozdziale 3. stwierdziliśmy, że użycie metody sympleksowej pozwala na wyznaczenie rozwiązania optymalnego zadania programowania liniowego w skończenie wielu iteracjach. Liczba tych iteracji nie przekracza liczby wszystkich wierzchołków wielościanu będącego zbiorem rozwiązań dopuszczalnych ZPL, która – jak stwierdziliśmy – jest dla postaci klasycznej tego zadania równa co najwyżej $\binom{m+n}{m}$. Już tak prosty wielościan, jak kostka $[0, 1]^n$ ma 2^n wierzchołków, mimo że można go zadać za pomocą $2n$ ograniczeń. Liczba ta już dla niewielkich n jest ogromna i gdyby algorytm sympleksowy musiał wykonać rzeczywiście aż tyle iteracji, byłby mało przydatny z praktycznego punktu widzenia. Z drugiej strony, doświadczenie wskazuje, że najczęściej algorytm sympleksowy zatrzymuje się po

wykonaniu co najwyżej $2m \div 3m$ iteracji. Chciałoby się jednak wiedzieć, czy w każdym przypadku mamy gwarancję, że algorytm sympleksowy nie prowadzi do „zbyt wielu” iteracji potrzebnych do wyznaczenia rozwiązania optymalnego. Problem polega na tym, jak określić pojęcie „zbyt wiele”. Poza tym nie chodzi tu wyłącznie o liczbę iteracji, ale również, a może nawet przede wszystkim o liczbę operacji arytmetycznych, które należy wykonać w celu zakończenia działania algorytmu. Problemami szybkości zbieżności algorytmów zajmuje się teoria złożoności obliczeniowej, przy czym nie chodzi tu o szybkość w sensie czasu (bo ta zależy również od urządzeń, na których wykonywane są obliczenia), a o szybkość w sensie liczby wspomnianych operacji arytmetycznych, które trzeba wykonać, aby algorytm zakończył swoje działanie. Podstawowym pojęciem tej teorii jest tak zwana długość danych wejściowych, czyli liczba bitów potrzebna do zapisu danych określających problem, który ma być rozwiązany. W przypadku zadania programowania liniowego danymi są macierz $A = [a_{ij}]_{m \times n}$ oraz wektory $b = (b_1, \dots, b_m)$ i $c = (c_1, \dots, c_n)$. W tym przypadku długość danych wejściowych wyraża się wzorem

$$L = \sum_{i=0}^m \sum_{j=0}^n \log_2(|a_{ij}| + 1) + \log_2(mn) + 1,$$

gdzie $a_{i0} = b_i$, $i = 1, \dots, m$, $a_{0j} = c_j$, $j = 1, \dots, n$, $a_{00} = 1$. Do analizy złożoności obliczeniowej algorytmu sympleksowego wystarczy nam jednak wyłącznie liczba iteracji (piwotyzacji), ponieważ każda iteracja składa się z takiej samej liczby operacji arytmetycznych. Ponadto jeśli założymy, że wszystkie dane wejściowe są mniej więcej jednakowego rzędu, to można pokazać, że długość danych wejściowych można oszacować z góry przez liczbę w przybliżeniu proporcjonalną do m i do n .

Wśród algorytmów służących rozwiązaniu zadanego problemu wyróżnia się algorytmy o *wielomianowej złożoności obliczeniowej* (to znaczy takie, dla których liczbę wykonywanych operacji można oszacować z góry wielomianem zależnym od długości danych wejściowych L) i o *wykładniczej złożoności obliczeniowej* (to znaczy takie, dla których liczba wykonywanych operacji rośnie wykładniczo wraz ze wzrostem długości danych wejściowych).

Złożoność obliczeniowa algorytmu sympleksowego była przedmiotem badań od początku jego powstania, czyli od połowy XX wieku. Jednak dopiero w roku 1972 Klee i Minty pokazali, że złożoność ta jest wykładnicza podając przykład problemu, dla którego liczba iteracji algorytmu sympleksowego niezbędnych do otrzymania rozwiązania rośnie wykładniczo wraz z wymiarem zadania. Przykład Klee i Minty'ego ma postać:

$$\begin{array}{ll} \text{maksymalizować} & \sum_{j=1}^n 10^{n-j} x_j \\ \text{przy ograniczeniach} & 2 \sum_{j=1}^{i-1} 10^{i-j} x_j + x_i \leq 100^{i-1}, \quad i = 1, \dots, n \\ & x \geq 0. \end{array}$$

Zbiór rozwiązań dopuszczalnych jest tutaj zdeformowaną kostką n -wymiarową. Kolejne rozwiązania wyznaczone przy pomocy metody sympleksowej są różnymi wierzchołkami tej kostki. Przykład jest skonstruowany tak, by wszystkie wierzchołki kostki zostały „odwiedzone” nim wyznaczone zostanie rozwiązanie optymalne. Stąd już nietrudno wywnioskować, że algorytm sympleksowy ma wykładniczą złożoność obliczeniową. Szczegóły można znaleźć w książce R. J. Vanderbeia [21]. Czytelnikowi polecamy prześledzenie kolejnych tablic dla $n = 3$. Zadanie ma wówczas postać

$$\begin{array}{ll} \text{maksymalizować} & 100x_1 + 10x_2 + x_3 \\ \text{przy ograniczeniach} & x_1 \leq 1 \\ & 20x_1 + x_2 \leq 100 \\ & 200x_1 + 20x_2 + x_3 \leq 10000 \\ & x_1, x_2, x_3 \geq 0. \end{array}$$

W związku z wykładniczą złożonością algorytmu sympleksowego rodzi się pytanie, czy dla ZPL istnieje algorytm o złożoności wielomianowej. Pozytywną odpowiedź na to pytanie podał gruziński matematyk L. Chaczjan w roku 1979. Opracował on tak zwaną *metodę elipsoidalną*, którą zastosował do rozwiązywania układu nierówności liniowych, ale można ją zastosować również do ZPL. Zadanie programowania liniowego

$$\begin{array}{ll} \text{maksymalizować} & c^\top x \\ \text{przy ograniczeniach} & Ax \leq b \\ & x \geq 0, \end{array}$$

można bowiem na podstawie twierdzeń o dualności przedstawić w postaci zagadnienia wyznaczenia rozwiązania układu nierówności liniowych

$$\begin{aligned} Ax &\leq b \\ A^\top y &\geq c \\ c^\top x &\geq b^\top y \\ x &\geq 0 \\ y &\geq 0. \end{aligned}$$

Metoda Chacjana polega na konstrukcji ciągu elipsoid o coraz to mniejszej objętości, których środki są zbieżne do rozwiązania zadania. Szczegóły można znaleźć w książce S. Walukiewicza [22]. Mimo, że metoda ta ma złożoność wielomianową, ma ona dzisiaj wyłącznie znaczenie teoretyczne. W praktyce zachowuje się ona bowiem gorzej – poza pewnymi sztucznie dobranymi przykładami – niż metoda sympleksowa. Poza tym w ciągu ostatnich kilkunastu lat skonstruowano dla zadań programowania liniowego inne metody

o wielomianowej złożoności obliczeniowej, które w praktyce skutecznie konkurują z metodą sympleksową. Pierwszą z takich metod była metoda punktów wewnętrznych podana w roku 1984 przez 28-letniego N. Karmarkara [14].

5.2. Metoda Karmarkara punktów wewnętrznych

Metoda Karmarkara punktów wewnętrznych służy do rozwiązania zadania programowania liniowego przedstawionych w tzw. *postaci normalnej*:

$$\begin{aligned} &\text{minimalizować} && c^\top x \\ &\text{przy ograniczeniach} && Ax = 0 \\ & && x \in \Delta_n, \end{aligned} \tag{5.1}$$

gdzie

$$\Delta_n = \{x \in \mathbb{R}^n : e^\top x = 1, x \geq 0\},$$

$e = (1, \dots, 1)^\top$. Ponadto zakłada się, że

- (i) A jest macierzą typu $m \times n$ pełnego rzędu wierszowego, tzn. $r(A) = m$,
- (ii) $Ae = 0$, innymi słowy tzw. *barycentrum sympleksu* (lub inaczej *środek ciężkości sympleksu*) $\frac{1}{n}e \in \Delta_n$ jest punktem dopuszczalnym,
- (iii) optymalna wartość funkcji celu $z^* = 0$.

5.2.1. Transformacja ZPL do postaci normalnej

W ustępie tym pokażemy, że dowolne zadanie programowania liniowego można sprowadzić do postaci normalnej. Sprowadzimy bowiem do niej ZPL w postaci klasycznej

$$\begin{array}{ll} \text{maksymalizować} & c^\top x \\ \text{przy ograniczeniach} & \begin{array}{ll} Ax & \leq b \\ x & \geq 0. \end{array} \end{array} \quad (5.2)$$

Korzystając z twierdzeń o dualności zadanie to można – jak już zauważyliśmy wcześniej – sprowadzić do wyznaczenia rozwiązania układu nierówności liniowych: $Ax \leq b$, $A^\top y \geq c$, $x \geq 0$, $y \geq 0$, $c^\top x \geq b^\top y$. Zapiszmy ten układ w postaci

$$Gu \leq d, \quad (5.3)$$

gdzie

$$G = \begin{bmatrix} A & 0 \\ 0 & -A^\top \\ -c^\top & b^\top \\ -I & 0 \\ 0 & -I \end{bmatrix}, \quad u = \begin{bmatrix} x \\ y \end{bmatrix}, \quad d = \begin{bmatrix} b \\ -c \\ 0 \end{bmatrix}. \quad (5.4)$$

Rozważmy następujące ZPL

$$\begin{array}{ll} \text{maksymalizować} & z \\ \text{względem} & (u, z) \in \mathbb{R}^{n+m} \times \mathbb{R} \\ \text{przy ograniczeniach} & Gu + ze \leq d. \end{array} \quad (5.5)$$

Lemat 5.2.1. *Układ nierówności (5.3) posiada rozwiązanie u^* wtedy i tylko wtedy gdy $(u^*, 0)$ jest rozwiązaniem optymalnym zadania (5.5).*

Dowód. Jeśli $(u^*, 0)$ jest rozwiązaniem optymalnym zadania (5.5), to oczywiście u^* jest rozwiązaniem zadania (5.3). Niech teraz u^* będzie rozwiązaniem zadania (5.3). Dla dowolnego rozwiązania dopuszczalnego (u, z) zadania (5.5) mamy $z \leq 0$. Przypuśćmy bowiem, że istnieje para (u, z) taka, że $Gu + ez \leq d$ i $z > 0$. Z nierówności (5.4) otrzymamy

$$\begin{aligned} Ax &< Ax + ze \leq b \\ -A^\top y &< -A^\top y + ze \leq -c \\ -c^\top x + b^\top y &< -c^\top x + b^\top y + ze \leq 0 \\ -x &< -x + ze \leq 0 \\ -y &< -y + ze \leq 0 \end{aligned}$$

i w konsekwencji $Ax < b$, $A^\top y > c$, $c^\top x > b^\top y$, $x > 0$, $y > 0$, co jest sprzeczne ze słabym twierdzeniem o dualności. W rezultacie, para $(u^*, 0)$ jest rozwiązaniem optymalnym tego zadania. \square

Uwaga 5.2.2. W dalszej części rozpatrywać będziemy zadanie (5.5), o którym zakładamy, że 0 jest jego wartością optymalną. Zgodnie z lematem 5.2.1 i twierdzeniami o dualności oznacza to, że zadanie (5.2) posiada rozwiązanie optymalne.

Zadanie dualne do zadania (5.5) ma postać

$$\begin{aligned} &\text{minimalizować} && d^\top v \\ &\text{przy ograniczeniach} && G^\top v = 0 \\ & && e^\top v = 1 \\ & && v \geq 0, \end{aligned} \tag{5.6}$$

w którym optymalna wartość funkcji celu wynosi 0, podobnie jak dla zadania (5.5). Bez szkody dla ogólności rozważań możemy założyć, że macierz G^\top jest pełnego rzędu wierszowego. W razie potrzeby możemy bowiem z układu $G^\top v = 0$ usunąć równania będące liniową kombinacją pozostałych. Wprowadźmy teraz sztuczną zmienną s i przekształćmy powyższe zadanie

do postaci

$$\begin{array}{ll} \text{minimalizować} & d^\top v + Ms \\ \text{przy ograniczeniach} & \begin{array}{l} G^\top v - G^\top es = 0 \\ e^\top v + s = 1 \\ v, s \geq 0, \end{array} \end{array} \quad (5.7)$$

gdzie $M > 0$. Jest oczywiste, że barycentrum sympleksu standardowego jest rozwiązaniem dopuszczalnym tego zadania. Pokażemy, że dla odpowiednio dużego M zadania (5.7) i (5.6) są sobie równoważne.

Lemat 5.2.3. *Jeśli M jest odpowiednio duże, to v^* jest rozwiązaniem optymalnym zadania (5.6) wtedy i tylko wtedy, gdy $(v^*, 0)$ jest rozwiązaniem zadania (5.7). Ponadto to drugie zadanie ma również wartość optymalną równą zero.*

Dowód. W celu pokazania równoważności obu zadań pokażemy równoważność zadań do nich dualnych. Zadaniem dualnym do (5.6) jest zadanie (5.5), zaś zadaniem dualnym do (5.7) jest zadanie

$$\begin{array}{ll} \text{maksymalizować} & z \\ \text{względem} & (u, z) \in \mathbb{R}^{n+m} \times \mathbb{R} \\ \text{przy ograniczeniach} & \begin{array}{l} Gu + ze \leq d \\ e^\top Gu + z \leq M. \end{array} \end{array} \quad (5.8)$$

Z mocnego twierdzenia o dualności wynika, że zadania (5.5) i (5.8) są dopuszczalne, gdyż zadania do nich dualne są ograniczone. Niech więc (u', z') i (u'', z'') będą rozwiązaniami dopuszczalnymi odpowiednio dla zadania (5.5) i (5.8) i niech $\zeta = \min\{z', z''\}$. Wobec tego wprowadzenie do obu zadań dodatkowego ograniczenia $z \geq \zeta$ przekształci je w zadania im odpowiednio równoważne. Dla $M \geq e^\top d + (1 - p)\zeta$, gdzie p jest liczbą nierówności w układzie (5.3), mamy więc dla rozwiązania dopuszczalnego (u, z) pierwszego zadania

$$e^\top Gu + pz \leq e^\top d.$$

Zatem

$$\begin{aligned} e^\top Gu + z &= e^\top Gu + pz + (1-p)z \\ &\leq e^\top d + (1-p)\zeta \\ &\leq M, \end{aligned}$$

czyli (u, z) jest również rozwiązaniem dopuszczalnym drugiego zadania. Widzimy więc, że dla odpowiednio dużego M drugie ograniczenie w zadaniu (5.8) jest spełnione, o ile tylko spełnione jest pierwsze ograniczenie. Zatem usunięcie tego drugiego ograniczenia z zadania (5.8) nie zmieni zbioru rozwiązań dopuszczalnych. W tej sytuacji jasne jest, że zadania (5.5) i (5.8) są identyczne, zatem zadania do nich dualne są sobie równoważne. Ponadto z twierdzenia o komplementarności wynika, że dla rozwiązania optymalnego (v^*, s^*) zadania (5.7) zajdzie równość $s^* = 0$, gdyż w drugim ograniczeniu zadania (5.8) nierówność jest zawsze ostra, o ile tylko $M > e^\top d + (1-p)\zeta$. Stąd już prosto wynika teza lematu. \square

Widzimy więc, że zadanie (5.7) ma postać (5.1), przy czym spełnione są warunki (i)-(iii).

Definicja 5.2.4. Rozwiązanie dopuszczalne x zadania (5.1) nazywa się ściśle dopuszczalne, jeśli wszystkie jego współrzędne są dodatnie, czyli $x \gg 0$.

5.2.2. Idea metody Karmarkara

W dalszej części pokażemy, że obierając punkt $x_1 = \frac{1}{n}e$ jako punkt startowy i postępując zgodnie z metodą Karmarkara wygenerujemy ciąg (x_k) takich rozwiązań ściśle dopuszczalnych zadania (5.1), że

$$c^\top x_{k+1} \leq \exp\left(-\frac{k}{5n}\right) c^\top x_1,$$

czego konsekwencją jest wielomianowa złożoność metody. W każdej iteracji metody wykonywane jest tzw. przekształcenie rzutowe sympleksu standardowego Δ_n zachowujące jego wierzchołki i „ściągające” kolejne (ściśle

dopuszczalne) przybliżenie rozwiązania do barycentrum sympleksu. Przekształcenie to opiszemy dokładniej w dalszej części. Zaletą takiego podejścia jest uniknięcie operacji kombinatorycznych na elementach brzegu sympleksu. Jak już wiemy, takie operacje wykonywane w algorytmie sympleksowym (piwotyzacje) prowadzą do wykładniczej złożoności tego algorytmu. W metodzie punktów wewnętrznych wykonywany jest krok w kierunku przeciwnym do gradientu przetransformowanej zgodnie z przekształceniem rzutowym funkcji celu, rzutowanego na przetransformowane ograniczenia równościowe. Następnie wykonanie odwrotnego przekształcenia rzutowego prowadzi do wyznaczenia kolejnego przybliżenia rozwiązania. Przypuśćmy, że funkcja celu zadania (5.1) nie jest stale równa zero na zbiorze rozwiązań dopuszczalnych $D = \{x \in \mathbb{R}^n : Ax = 0, e^\top x = 1, x \geq 0\}$. Wówczas miarą szybkości zbliżania się do rozwiązania optymalnego jest tzw. *funkcja potencjału*

$$f(x) = n \ln c^\top x - \sum_{j=1}^n \ln \xi_j = \sum_{j=1}^n \ln \frac{c^\top x}{\xi_j},$$

gdzie $x = (\xi_1, \dots, \xi_n)^\top$.

Pokażemy, że funkcja potencjału jest dobrze określona na zbiorze rozwiązań ściśle dopuszczalnych zadania (5.1), tzn., że

$$c^\top x > 0 \text{ dla } x \in \text{ri } D = \{x \in \mathbb{R}^n : Ax = 0, e^\top x = 1, x \gg 0\}.$$

Jest jasne, że $c^\top x \geq 0$ dla $x \in \text{ri } D$, gdyż optymalna wartość funkcji celu zadania (5.1) wynosi zero. Przypuśćmy, że $c^\top x = 0$ dla pewnego $x \in \text{ri } D$. Niech $S = \{s_1, \dots, s_m\}$ będzie zbiorem punktów ekstremalnych zbioru D i niech $S_1 = \{s \in S : c^\top s = 0\}$ oraz $S_2 = \{s \in S : c^\top s > 0\}$. Oczywiście $S = S_1 \cup S_2$ oraz $S_2 \neq \emptyset$, gdyż założyliśmy, że funkcja celu nie jest stale równa zero na zbiorze D . Z twierdzenia Minkowskiego wynika, że $x = \sum_{i=1}^m \lambda_i s_i$ dla pewnych stałych $\lambda_1, \dots, \lambda_m$ takich, że $\sum_{i=1}^m \lambda_i = 1$ i $\lambda_i \geq 0, i = 1, \dots, m$. Wobec tego

$$0 = c^\top x = \sum_{i=1}^m \lambda_i c^\top s_i = \sum_{s_i \in S_2} \lambda_i c^\top s_i > 0.$$

Otrzymana sprzeczność dowodzi, że funkcja celu zadania (5.1) jest dodatnia na zbiorze $\text{ri } D$, w konsekwencji funkcja potencjału jest dobrze określona na zbiorze rozwiązań ściśle dopuszczalnych tego zadania.

5.2.3. Przekształcenie rzutowe

Niech $\bar{x} = (\bar{\xi}_1, \dots, \bar{\xi}_n)^\top \in \text{ri } \Delta_n = \{x \in \mathbb{R}^n : e^\top x = 1, x \gg 0\}$. Oznaczmy

$$\bar{X} = \text{diag } \bar{x} = \begin{bmatrix} \bar{\xi}_1 & 0 & \dots & 0 \\ 0 & \bar{\xi}_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \bar{\xi}_n \end{bmatrix}.$$

Oczywiście macierz \bar{X} jest nieosobliwa. Określmy przekształcenie $T : \Delta_n \rightarrow \Delta_n$ następująco

$$y = Tx = \frac{1}{e^\top \bar{X}^{-1}x} \bar{X}^{-1}x.$$

Odwzorowanie to będziemy nazywać *przekształceniem rzutowym*. Nietrudno pokazać, że $T\bar{x} = \frac{1}{n}e$, $Te_j = e_j$, $j = 1, \dots, n$, oraz, że T jest odwzorowaniem wzajemnie jednoznacznym i

$$x = T^{-1}y = \frac{1}{e^\top \bar{X}y} \bar{X}y.$$

Zauważmy ponadto, że T jest złożeniem dwóch przekształceń, $T = T_2 \circ T_1$. Przekształcenie T_1 odwzorowuje Δ_n w $\mathbb{R}_+^n \setminus \{0\}$ zgodnie z równością

$$u = T_1x = \frac{1}{n} \bar{X}^{-1}x.$$

Natomiast przekształcenie T_2 odwzorowuje $\mathbb{R}_+^n \setminus \{0\}$ w Δ_n zgodnie z równością

$$y = T_2u = \frac{1}{e^\top u} u.$$

Odwzorowanie T_1 nazywa się *skalowaniem* i jest przekształceniem liniowym. Obrazem punktu \bar{x} w tym odwzorowaniu jest barycentrum sympleksu, tzn.

$T_1\bar{x} = \frac{1}{n}e$. Ponadto odwzorowanie to przeprowadza wierzchołki e_i sympleksu standardowego Δ_n w punkty $(n\bar{\xi}_i)^{-1}e_i$, $i = 1, \dots, n$, będące wierzchołkami „przeskalowanego” sympleksu

$$T_1(\Delta) = \{x \in \mathbb{R}^n : e^\top \bar{X}^{-1}x = n, x \geq 0\}.$$

Odwzorowanie T_2 nazywa się *rzutem środkowym*. Zauważmy, że $T_2u \in \Delta_n$ dla dowolnego $u \geq 0, u \neq 0$, $T_2(\frac{1}{n}e) = \frac{1}{n}e$ i $T_2(\alpha e_i) = e_i$ dla dowolnego $\alpha > 0$. Zauważmy ponadto, że w wyniku działania odwzorowania T_2 punkt $u \in \mathbb{R}_+^n \setminus \{0\}$ zostaje „ściągnięty” wzdłuż promienia przechodzącego przez punkt u i przez początek układu do punktu na sympleksie standardowym Δ_n . Stąd nazwa tego odwzorowania – rzut środkowy. Odwzorowanie to nie jest jednak liniowe. Ma ono natomiast ważną własność, którą mają również przekształcenia liniowe.

Lemat 5.2.5. *Rzut środkowy T_2 przekształca odcinek o końcach w zbiorze $\mathbb{R}_+^n \setminus \{0\}$ w odcinek o końcach w sympleksie standardowym Δ_n .*

Dowód. Przypomnijmy, że odcinkiem o końcach $a, b \in \mathbb{R}^n$ jest zbiór

$$[a, b] = \{x \in \mathbb{R}^n : x = (1 - \lambda)a + \lambda b, \lambda \in [0, 1]\}.$$

Niech $a, b \in \mathbb{R}_+^n \setminus \{0\}$. Niech $u = (1 - \lambda)a + \lambda b$ dla pewnego $\lambda \in [0, 1]$ i niech $y = T_2u$. Pokażemy, że $T_2u \in [T_2a, T_2b]$. Mamy $e^\top u > 0$ i

$$\begin{aligned} y &= \frac{1}{e^\top u}[(1 - \lambda)a + \lambda b] \\ &= \frac{(1 - \lambda)e^\top a}{e^\top u}u \cdot \frac{a}{e^\top a} + \frac{\lambda e^\top b}{e^\top u} \cdot \frac{b}{e^\top b} \\ &= \mu T_2a + \nu T_2b \end{aligned}$$

dla $\mu = (e^\top u)^{-1}(1 - \lambda)e^\top a$, $\nu = (e^\top u)^{-1}\lambda e^\top b$. Zauważmy, że $\mu, \nu \in [0, 1]$, gdyż $a, b \in \mathbb{R}_+^n$ oraz, że $\mu + \nu = 1$. Zatem $y \in [T_2a, T_2b]$. \square

Wniosek 5.2.6. *Przekształcenie rzutowe T przekształca wielościan o wierzchołkach w zbiorze $\mathbb{R}_+^n \setminus \{0\}$ w wielościan o wierzchołkach w sympleksie standardowym Δ_n .*

W wyniku transformacji projekcyjnej wyjściowe zadanie (5.1) przybiera równoważną postać

$$\text{minimalizować } (e^\top \bar{X}y)^{-1} c^\top \bar{X}y$$

$$\begin{aligned} \text{przy ograniczeniach} \quad & A\bar{X}y = 0 \\ & y \in \Delta_n. \end{aligned}$$

Ponieważ optymalna wartość funkcji celu $z^* = 0$, więc definiując $\bar{A} = A\bar{X}$ i $\bar{c} = \bar{X}c$ i pomijając czynnik $(e^\top \bar{X}y)^{-1}$ w funkcji celu otrzymujemy zadanie równoważne:

$$\begin{aligned} \text{minimalizować } & \bar{c}^\top y \\ \text{przy ograniczeniach } & \bar{A}y = 0 \\ & y \in \Delta_n, \end{aligned} \tag{5.9}$$

którego postać jest identyczna z (5.1).

5.2.4. Relaksacja ograniczeń

Trudność rozwiązania zadania (5.9) wynika z tego, że ograniczenia $y \in \Delta_n$ nie są gładkie. W metodzie punktów wewnętrznych zastępuje się je ograniczeniami gładkimi

$$y \in B'_n\left(\frac{1}{n}e, \rho\right) = \{y \in \mathbb{R}^n : e^\top y = 1, \|y - \frac{1}{n}e\| \leq \rho\},$$

gdzie $\rho > 0$ jest dobrane tak, aby $B'_n(\frac{1}{n}e, \rho) \subset \Delta_n$. Zobaczymy, że w tym przypadku rozwiązanie zadania przyjmuje jawną postać. Czytelnikowi zostawimy jako ćwiczenie sprawdzenie, że

$$B'_n\left(\frac{1}{n}e, r\right) \subset \Delta_n \tag{5.10a}$$

dla $r = \frac{1}{\sqrt{n(n-1)}}$ i, że ten promień jest optymalny (tzn. dla $r > \frac{1}{\sqrt{n(n-1)}}$ inkluzja ta już nie zachodzi) oraz, że

$$\Delta_n \subset B'_n\left(\frac{1}{n}e, R\right) \tag{5.11}$$

dla $R = \sqrt{\frac{n-1}{n}}$.

Zajmiemy się teraz zadaniem

$$\begin{aligned} & \text{minimalizować} && \bar{c}^\top y \\ & \text{przy ograniczeniach} && \begin{aligned} \bar{A}y &= 0 \\ e^\top y &= 1 \\ \frac{1}{2}\|y - \frac{1}{n}e\|^2 &\leq \frac{1}{2}\rho^2, \end{aligned} \end{aligned} \quad (5.12)$$

gdzie $\rho > 0$, które – zgodnie z tym co powiedziano wyżej – powstało z zadania (5.9) przez zastąpienie ograniczeń niegładkich $y \in \Delta_n$ ograniczeniami gładkimi $y \in B'_n(\frac{1}{n}e, \rho)$. Ostatnie ograniczenie w powyższym zadaniu jest oczywiście równoważne nierówności $\|y - \frac{1}{n}e\| \leq \rho$. Podaną w tym zadaniu postać tego ograniczenia przyjęto ze względu na łatwiejsze wyznaczenie punktu Kuhna–Tuckera.

Lemat 5.2.7. *Rozwiązaniem zadania (5.12) jest*

$$y^+(\rho) = -\rho \frac{1}{\|d\|} d + \frac{1}{n}e, \quad (5.13)$$

gdzie

$$d = [I - \bar{A}^\top (\bar{A} \bar{A}^\top)^{-1} \bar{A} + \frac{1}{n}ee^\top] \bar{c}.$$

Dowód. Funkcja celu w zadaniu (5.12) jest liniowa, zaś zbiór rozwiązań dopuszczalnych jest zwarty i wypukły. Z twierdzenia 2.3.26 wynika więc, że rozwiązanie tego zadania istnieje i jest osiągnięte w punkcie ekstremalnym tego zbioru, tzn. dla y takiego, że $\|y - \frac{1}{n}e\| = \rho$. W celu wyznaczenia tego rozwiązania znajdziemy punkt stacjonarny funkcji Lagrange’a $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}$,

$$\mathcal{L}(y, \lambda, \mu, \nu) = \bar{c}^\top y + \lambda^\top \bar{A}y + \mu(e^\top y - 1) + \nu(\frac{1}{2}\|y - \frac{1}{n}e\|^2 - \frac{1}{2}\rho^2).$$

Mamy

$$\begin{aligned} \nabla_y \mathcal{L}(y, \lambda, \mu, \nu) &= \bar{c} + \bar{A}^\top \lambda + \mu e + \nu(y - \frac{1}{n}e) &= 0 \\ \nabla_\lambda \mathcal{L}(y, \lambda, \mu, \nu) &= \bar{A}y &= 0 \\ \nabla_\mu \mathcal{L}(y, \lambda, \mu, \nu) &= e^\top y - 1 &= 0 \\ \nabla_\nu \mathcal{L}(y, \lambda, \mu, \nu) &= \frac{1}{2}\|y - \frac{1}{n}e\|^2 - \frac{1}{2}\rho^2 &= 0 \end{aligned} \quad (5.14)$$

Z pierwszego równania otrzymujemy

$$y = -\frac{1}{\nu}(\bar{c} - \bar{A}^\top \lambda - \mu e + \frac{1}{n}e) \quad (5.15)$$

Po uwzględnieniu tej równości oraz założenia $\bar{A}e = 0$, drugie równanie układu (5.14) przybierze postać

$$\bar{A} \bar{A}^\top \lambda = -\bar{A} \bar{c}, \quad (5.16)$$

zaś równanie trzecie – postać

$$\mu = \frac{1}{n}e^\top \bar{c}.$$

Macierz $\bar{A} \bar{A}^\top$ jest nieosobliwa, gdyż macierz A ma – zgodnie z założeniem – pełny rząd wierszowy oraz macierz \bar{X} jest nieosobliwa. W konsekwencji macierz \bar{A} ma pełny rząd wierszowy. Wobec tego rozwiązaniem równania (5.16) jest

$$\lambda = -(\bar{A} \bar{A}^\top)^{-1} \bar{A} \bar{c}.$$

Wstawiając wyznaczone λ i μ do równości (5.15) otrzymamy

$$y - \frac{1}{n}e = -\frac{1}{\nu}[I - \bar{A}^\top(\bar{A} \bar{A}^\top)^{-1}\bar{A} + \frac{1}{n}ee^\top]\bar{c}.$$

Po uwzględnieniu ostatniego równania układu (5.14) otrzymamy dwa punkty stacjonarne funkcji Lagrange’a: $y = \frac{1}{n}e \pm \rho \frac{1}{\|d\|}d$. Ponieważ mnożnik Lagrange’a ν odpowiadający ostatniemu ograniczeniu w zadaniu (5.12) jest nieujemny, otrzymujemy, że jedynym punktem Kuhna–Tuckera jest

$$y = \frac{1}{n}e - \rho \frac{1}{\|d\|}d,$$

gdzie

$$d = [I - \bar{A}^\top(\bar{A} \bar{A}^\top)^{-1}\bar{A} + \frac{1}{n}ee^\top]\bar{c}.$$

Ponieważ rozwiązanie zadania (5.12) istnieje, jest nim więc otrzymany punkt y . \square

Uwaga 5.2.8. Postać wektora d podaną w lemacie 5.2.7 można zapisać prościej:

$$d = [I - B^\top (BB^\top)^{-1} B] \bar{c}, \quad (5.17)$$

gdzie

$$B = \begin{bmatrix} \bar{A} \\ e^\top \end{bmatrix}, \quad (5.18)$$

gdź

$$(BB^\top)^{-1} = \begin{bmatrix} (\bar{A} \bar{A}^\top)^{-1} & 0 \\ 0 & \frac{1}{n} \end{bmatrix}.$$

W tej sytuacji otrzymane rozwiązanie ma elegancką interpretację geometryczną przedstawioną w następującym lemacie.

Lemat 5.2.9. *Wektor d dany równością (5.17) jest rzutem ortogonalnym wektora \bar{c} na podprzestrzeń zerową macierzy B ,*

$$\ker B = \{y \in \mathbb{R}^n : By = 0\}.$$

W konsekwencji $\bar{c}^\top d = \|d\|^2$.

Dowód. W celu pokazania lematu wystarczy sprawdzić, że $d \in \ker B$ i że $(\bar{c} - d) \perp \ker B$. Sprawdzenie tych faktów pozostawiamy czytelnikowi. \square

Przedstawimy teraz oszacowanie optymalnej wartości funkcji celu zadania (5.12) przydatne w późniejszej analizie.

Lemat 5.2.10. *Dla $y^+(\rho)$ danego równością (5.13) zachodzi nierówność*

$$\bar{c}^\top y^+(\rho) \leq (1 - \rho) \frac{\bar{c}^\top e}{n}. \quad (5.19)$$

Dowód. Z lematu 5.2.9. otrzymujemy

$$\bar{c}^\top y^+(\rho) = \frac{1}{n} \bar{c}^\top e - \rho \frac{1}{\|d\|} \bar{c}^\top d = \frac{1}{n} \bar{c}^\top e - \rho \|d\|. \quad (5.20)$$

Z drugiej strony dla $\rho = R = \sqrt{\frac{n-1}{n}}$ mamy

$$\bar{c}^\top y^+(R) \leq 0, \quad (5.21)$$

gdyż optymalna wartość funkcji celu zadania (5.1) wynosi zero i zachodzi inkluzja $\Delta_n \subset B'_n(\frac{1}{n}e, R)$. Korzystając z równości (5.20) i z nierówności (5.21) możemy teraz oszacować z dołu $\|d\|$. Mamy

$$\|d\| \geq \frac{1}{nR} \bar{c}^\top e \geq \frac{1}{n} \bar{c}^\top e,$$

gdyż $R \leq 1$. Podstawiając to oszacowanie do równości (5.20) otrzymamy tęzę lematu. \square

5.2.5. Opis metody Karmarkara punktów wewnętrznych

Przypuśćmy, że w k -tej iteracji wyznaczyliśmy przybliżenie $\bar{x} = x_k$ rozwiązania zadania (5.1). Kolejna, $(k+1)$ -sza iteracja metody Karmarkara polega na przekształceniu zadania przy pomocy odwzorowania rzutowego przeprowadzającego \bar{x} w barycentrum $\bar{y} = \frac{1}{n}e$ sympleksu Δ_n , następnie wyznaczeniu rozwiązania $y^+ = \frac{1}{n}e - \rho \frac{1}{\|d\|}d$ przekształconego zadania (5.12) dla parametru $\rho > 0$ gwarantującego pozostanie w zbiorze rozwiązań ściśle dopuszczalnych, a następnie wykonaniu odwrotnej transformacji projekcyjnej x^+ punktu y^+ . Otrzymujemy w ten sposób kolejne przybliżenie $x_{k+1} = x^+$. Możemy to zapisać w postaci następującego schematu

$$\begin{array}{ccc} \bar{x} & & x^+ \\ T \downarrow & & \uparrow T^{-1} \\ \bar{y} & \longrightarrow & y^+ \end{array}$$

Pokażemy w dalszej części tego ustępu, że w wyniku takich operacji funkcja potencjału zmniejsza się o stałą wielkość. Własność ta będzie mieć z kolei istotne znaczenie dla dowodu wielomianowej złożoności metody. Przedstawimy najpierw powyższy opis w postaci ścisłej procedury.

Algorytm 5.2.11 (punktów wewnętrznych)

Wejście: Macierz A typu $m \times n$ i wektor $c \in \mathbb{R}^n$ spełniające warunki (i)-(iii), parametr tolerancji optymalności l .

Wyjście: Punkt $\bar{x} \in \mathbb{R}^n$ dopuszczalny dla zadania (5.1) taki, że

$$c^\top \bar{x} \leq 2^{-l} \frac{1}{n} c^\top e.$$

Krok 0. (*inicjalizacja*)

- (a) Położyć $\bar{x} = \frac{1}{n}e$.
- (b) Położyć $k = 1$.

Krok 1. (*kryterium zatrzymania*). Jeśli $c^\top \bar{x} \leq 2^{-l} \frac{1}{n} c^\top e$, to algorytm zatrzymuje się.

Krok 2. (*relaksacja przetransformowanego zadania*)

- (a) Położyć $\bar{X} = \text{diag } \bar{x}$.
- (b) Położyć $\bar{c} = \bar{X}c$ i $\bar{A} = A\bar{X}$.
- (c) Wyznaczyć d ze wzoru (5.17).
- (d) Położyć $y^+ = \frac{1}{n}e - \frac{\alpha}{n} \frac{1}{\|d\|} d$ dla $\alpha \in (0, 1)$.

Krok 3. (*odwrotna transformacja projekcyjna*)

- (a) Położyć $x^+ = (e^\top \bar{X}y)^{-1} \bar{X}y^+$.
- (b) Położyć $\bar{x} = x^+$.
- (c) Zwiększyć k o 1 i przejść do kroku 1.

5.2.6. Zbieżność i złożoność obliczeniowa metody punktów wewnętrznych

Pokażemy, że w wyniku pojedynczej iteracji powyższego algorytmu funkcja potencjału zmniejsza się o stałą wartość. W konsekwencji funkcja celu zadania (5.1) zbiega geometrycznie do 0. W tym celu podamy najpierw pewne własności funkcji potencjału

$$f(x) = f(x, c) = n \ln c^\top x - \sum_{j=1}^n \ln \xi_j.$$

Oznaczmy

$$\bar{f}(y) = f(y, \bar{c}) = f(y, \bar{X}c)$$

Lemat 5.2.12. *Niech $\bar{x} \in \text{ri } \Delta_n$. Wówczas dla $y = Tx = (e^\top \bar{X}^{-1}x)^{-1} \bar{X}^{-1}x$ zachodzi równość*

$$\bar{f}(y) = f(x) + \ln(\det \bar{X}).$$

Dowód. Mamy

$$\begin{aligned} \bar{f}(y) &= n \ln c^\top \bar{X} \frac{\bar{X}^{-1}x}{e^\top \bar{X}^{-1}x} - \sum_{j=1}^n \ln \frac{\xi_j / \bar{\xi}_j}{e^\top \bar{X}^{-1}x} \\ &= n \ln c^\top x - n \ln e^\top \bar{X}^{-1}x - \sum_{j=1}^n (-\ln \bar{\xi}_j + \ln \xi_j - \ln e^\top \bar{X}^{-1}x) \\ &= n \ln c^\top x - \sum_{j=1}^n \ln \xi_j + \sum_{j=1}^n \ln \bar{\xi}_j = f(x) + \ln(\det \bar{X}). \end{aligned}$$

□

Lemat 5.2.13. *Dla y^+ wyznaczonego w kroku 2(d) algorytmu zachodzi nierówność*

$$n \ln \bar{c}^\top y^+ \leq n \ln \left(\frac{1}{n} \bar{c}^\top e \right) - \alpha.$$

Dowód. Korzystając z nierówności (5.19) dla $\rho = \frac{\alpha}{n}$ i z oczywistej nierówności $\ln(1+x) \leq x$ prawdziwej dla $x > -1$ otrzymujemy

$$\begin{aligned} \ln \bar{c}^\top y^+ &\leq \ln\left(1 - \frac{\alpha}{n}\right) \frac{\bar{c}^\top e}{n} \\ &= \ln\left(1 - \frac{\alpha}{n}\right) + \ln \frac{\bar{c}^\top e}{n} \\ &\leq -\frac{\alpha}{n} + \ln \frac{\bar{c}^\top e}{n}. \end{aligned}$$

Mnożąc powyższe nierówności przez n otrzymamy tezę lematu. \square

Lemat 5.2.14. Dla $y = (\eta_1, \dots, \eta_n)^\top \in B'_n(\frac{1}{n}e, \frac{\alpha}{n})$ zachodzi nierówność

$$-\sum_{j=1}^n \ln \eta_j \leq -\sum_{j=1}^n \ln \frac{1}{n} + \frac{\alpha^2}{2(1-\alpha)^2}.$$

Dowód. Rozwijając we wzór Taylora funkcję \ln w otoczeniu punktu 1 otrzymamy

$$\ln(1 + (n\eta_j - 1)) = n\eta_j - 1 - \frac{1}{2}(n\eta_j - 1)^2 \frac{1}{(1 + \Theta(n\eta_j - 1))^2}$$

dla pewnego $\Theta \in (0, 1)$. Skoro $y \in B'_n(\frac{1}{n}e, \frac{\alpha}{n})$, więc dla każdej współrzędnej wektora y zajdzie oczywista nierówność $n\eta_j \geq 1 - \alpha$, $j = 1, \dots, n$. Wobec tego mamy

$$\ln(n\eta_j) \geq n\eta_j - 1 - \frac{1}{2(1-\alpha)^2}(n\eta_j - 1)^2.$$

Ponieważ $\sum_{j=1}^n n\eta_j = n$ i

$$\sum_{j=1}^n (n\eta_j - 1)^2 = \|ny - e\|^2 = n^2 \|y - \frac{e}{n}\|^2 \leq \alpha^2,$$

więc

$$\sum_{j=1}^n \ln(n\eta_j) \geq \frac{\alpha^2}{2(1-\alpha)^2},$$

skąd otrzymujemy bezpośrednio tezę lematu. \square

Twierdzenie 5.2.15. *Jeśli w algorytmie 5.2.11 przyjąć $\alpha = \frac{1}{3}$, to w każdej iteracji tego algorytmu funkcja potencjału zmniejsza się co najmniej o $\frac{1}{5}$.*

Dowód. Z lematów 5.2.13 i 5.2.14 otrzymujemy

$$\begin{aligned}\bar{f}(y^+) &= n \ln \bar{c}^\top y^+ - \sum_{j=1}^n \ln \eta_j \\ &\leq n \ln\left(\frac{1}{n} \bar{c}^\top e\right) - \alpha + \frac{\alpha^2}{2(1-\alpha)^2} - \sum_{j=1}^n \ln \frac{1}{n} \\ &= \bar{f}\left(\frac{e}{n}\right) - \alpha + \frac{\alpha^2}{2(1-\alpha)^2}.\end{aligned}$$

Biorąc $\alpha = \frac{1}{3}$ dostajemy więc $\alpha - \frac{\alpha^2}{2(1-\alpha)^2} = \frac{5}{24} > \frac{1}{5}$ i

$$\bar{f}(y^+) \leq \bar{f}\left(\frac{e}{n}\right) - \frac{1}{5}. \quad (5.22)$$

Z kolei stosując lemat 5.2.12 dla $y = \frac{e}{n} = T(\bar{x})$ i dla $y = y^+ = T(x^+)$ otrzymujemy równość

$$\bar{f}\left(\frac{e}{n}\right) = f(\bar{x}) + \ln(\det \bar{X})$$

oraz

$$\bar{f}(y^+) = f(x^+) + \ln(\det \bar{X}).$$

Zatem na mocy nierówności (5.22) mamy

$$\begin{aligned}f(x^+) &= \bar{f}(y^+) - \ln(\det \bar{X}) \\ &\leq \bar{f}\left(\frac{e}{n}\right) - \frac{1}{5} - \ln(\det \bar{X}) \\ &= f(\bar{x}) - \frac{1}{5}.\end{aligned}$$

□

Oznaczmy symbolem x_k punkt \bar{x} otrzymany w k -tej iteracji algorytmu 5.2.11. Jeśli w algorytmie tym przyjąć $\alpha = \frac{1}{3}$, to z twierdzenia 5.2.15 wynika oczywiście, że

$$f(x_{k+1}) \leq f(x_1) - \frac{k}{5}. \quad (5.23)$$

Zapiszemy tę nierówność w języku funkcji celu zadania (5.1). Zauważmy przede wszystkim, że funkcja $h : \Delta_n \rightarrow R$ określona następująco:

$$h(x) = h(\xi_1, \dots, \xi_n) = \xi_1 \cdot \dots \cdot \xi_n$$

osiąga swoje maksimum w barycentrum sympleksu Δ_n . Wobec tego oznaczając $x_i = (\xi_{1i}, \dots, \xi_{ni})^\top$ otrzymamy

$$\sum_{j=1}^n \ln \xi_{j,k+1} \leq \sum_{j=1}^n \ln \frac{1}{n} = \sum_{j=1}^n \ln \xi_{j1}.$$

W konsekwencji z nierówności (5.23) wynika, że

$$\begin{aligned} n \ln c^\top x_{k+1} &\leq n \ln c^\top x_1 - \frac{k}{5} + \sum_{j=1}^n \ln \xi_{j,k+1} - \sum_{j=1}^n \ln \xi_{j1} \\ &\leq n \ln c^\top x_1 - \frac{k}{5}. \end{aligned}$$

Stąd już prosto otrzymujemy nierówność

$$c^\top x_{k+1} \leq \exp\left(-\frac{k}{5n}\right) c^\top x_1. \quad (5.24)$$

Ponieważ optymalna wartość funkcji celu tego zadania wynosi 0 i zbiór rozwiązań dopuszczalnych tego zadania jest zwarty, więc konsekwencją tej nierówności jest zbieżność pewnego podciągu (x_{n_k}) do rozwiązania optymalnego. Ważniejszą konsekwencją tej nierówności jest jednak zbieżność w czasie wielomianowym. Przypuśćmy, że wszystkie dane wejściowe są całkowite. Nie zmniejsza to w praktyce ogólności rozważań, ponieważ wszystkie dane w maszynie cyfrowej mają skończoną reprezentację. Wobec tego w wyniku

pomnożenia wszystkich danych wejściowych przez pewną stałą całkowitą będą one całkowite, a zadanie minimalizacji otrzyma postać równoważną. Przypuśćmy ponadto, że stałą l w kroku 1 algorytmu 5.2.11 przyjęto równą długości danych wejściowych L (lub równą liczbie do niej proporcjonalnej) zaś stałą α w kroku 2(d) równą $\frac{1}{3}$. Wówczas z nierówności (5.24) wynika, że dla $k > 5nL$ zachodzą nierówności

$$c^\top x_{k+1} < \exp\left(-\frac{k}{5n}\right) c^\top x_1 < \exp(-L) c^\top x_1.$$

Oznacza to, że po wykonaniu co najwyżej $O(nL)$ iteracji algorytm zatrzyma się dając w wyniku rozwiązanie ε -optymalne dla $\varepsilon = \exp(-L) c^\top \frac{e}{n}$. W każdej iteracji tego algorytmu nakład obliczeniowy jest zdominowany przez odwrócenie macierzy BB^\top , które – jak można pokazać – wymaga $O(n^3)$ operacji arytmetycznych. Stąd już nietrudno otrzymać wielomianową zbieżność algorytmu Karmarkara.

Twierdzenie 5.2.16. *Algorytm Karmarkara punktów wewnętrznych wymaga co najwyżej $O(nL)$ iteracji i $O(n^4L)$ operacji arytmetycznych w celu uzyskania rozwiązania optymalnego.*

Uwaga 5.2.17. W praktyce algorytm Karmarkara daje rozwiązanie z dużą dokładnością już po 20-30 iteracjach niezależnie od rozmiaru zadania. Pamiętać jednak należy o tym, że w każdej iteracji należy odwrócić macierz BB^\top , która dla dużych zadań może mieć duży rozmiar. Niemniej jednak, metoda Karmarkara skutecznie konkuruje w praktyce z metodą sympleksową. Ponadto metoda ta dała impuls do powstania w ostatnich kilkunastu latach wielu nowych metod, na przykład metod skalowania afinicznego, jednorodnej metody samodualnej czy też metod podążania za ścieżką. Ich przedstawienie wykracza jednak poza ramy tego podręcznika. Zainteresowani nimi czytelnicy mogą znaleźć dokładniejsze informacje na ich temat w książce R. J. Vanderbeia [21].

ROZDZIAŁ 6

Zadanie transportowe

*Żadna inna nauka nie umacnia tak wiary w siłę ludzkiego ducha,
jak matematyka.*

[H. Steinhaus]

6.1. Podstawowe pojęcia

6.1.1. Zbilansowane zadanie transportowe

Dla określonego towaru dana jest sieć m magazynów (lub dostawców) A_1, \dots, A_m i n sklepów (lub odbiorców) B_1, \dots, B_n . Zapas magazynu (lub podaż dostawcy) A_i wynosi a_i jednostek towaru, $i = 1, \dots, m$. Sklep (lub odbiorca) B_j potrzebuje b_j jednostek towaru, $j = 1, \dots, n$. Zakładamy, że

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j \quad (6.1)$$

(łączna podaż jest równa łącznemu popytowi). Koszty transportu jednostki towaru z magazynu A_i do sklepu B_j wynoszą c_{ij} jednostek pieniężnych, $i = 1, \dots, m$, $j = 1, \dots, n$. Należy określić plan transportowy o minimalnych kosztach zaspokajający zapotrzebowania wszystkich sklepów (czyli, przy podanym założeniu, wyczerpujący łączne zapasy magazynów).

Oznaczając przez x_{ij} ilość towaru transportowanego z magazynu A_i do sklepu B_j , $i = 1, \dots, m$, $j = 1, \dots, n$, powyższe zagadnienie można zapisać w następującej formie:

$$\begin{array}{ll} \text{minimalizować} & \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \\ \text{przy ograniczeniach} & \begin{array}{ll} \sum_{j=1}^n x_{ij} = a_i, & i = 1, \dots, m \\ \sum_{i=1}^m x_{ij} = b_j, & j = 1, \dots, n \\ x_{ij} \geq 0, & i = 1, \dots, m, j = 1, \dots, n. \end{array} \end{array} \quad (6.2)$$

Zadanie (6.2) nazywa się *zbilansowanym zadaniem transportowym*. Macierz

$$C = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mn} \end{bmatrix}$$

nazywa się *macierzą jednostkowych kosztów transportowych* lub krótko *macierzą kosztów*, zaś macierz

$$X = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \dots & \dots & \dots \\ x_{m1} & \dots & x_{mn} \end{bmatrix}$$

– *macierzą zmiennych decyzyjnych* lub *planem transportowym*.

6.1.2. Niezbilansowane zadanie transportowe

Z praktycznego punktu widzenia założenie

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$$

wydać się za mocne. Jeśli nie jest ono spełnione, to po odpowiedniej modyfikacji ograniczeń otrzymamy *niezbilansowane zadanie transportowe*. Jeśli zachodzi nierówność

$$\sum_{i=1}^m a_i > \sum_{j=1}^n b_j$$

(podaż jest większa od popytu), to niezbilansowane zadanie transportowe przybiera postać

$$\begin{array}{ll} \text{minimalizować} & \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \\ \text{przy ograniczeniach} & \sum_{j=1}^n x_{ij} \leq a_i, \quad i = 1, \dots, m \\ & \sum_{i=1}^m x_{ij} = b_j, \quad j = 1, \dots, n \\ & x_{ij} \geq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n. \end{array}$$

Nowe ograniczenia oznaczają, że nie wszystkie zapasy magazynów zostaną wyczerpane (nie cała podaż dostawców znajdzie nabywcę). W tym przypadku zadanie to można sprowadzić do zbilansowanego zadania transportowego wprowadzając fikcyjny sklep (fikcyjnego odbiorcę) B_{n+1} z zapotrzebowaniem (popytem)

$$b_{n+1} = \sum_{i=1}^m a_i - \sum_{j=1}^n b_j$$

i zerowymi kosztami transportowymi z dowolnego magazynu (od dowolnego dostawcy):

$$c_{i,n+1} = 0, \quad i = 1, \dots, m.$$

Jeśli natomiast zachodzi nierówność

$$\sum_{i=1}^m a_i < \sum_{j=1}^n b_j,$$

(podaż jest mniejsza od popytu), to niezbilansowane zadanie transportowe przybiera postać

$$\begin{array}{ll} \text{minimalizować} & \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \\ \text{przy ograniczeniach} & \sum_{j=1}^n x_{ij} = a_i, \quad i = 1, \dots, m \\ & \sum_{i=1}^m x_{ij} \leq b_j, \quad j = 1, \dots, n \\ & x_{ij} \geq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n. \end{array}$$

Nowe ograniczenia oznaczają, że zapotrzebowanie (popyt) nie wszystkich sklepów (odbiorców) zostanie zaspokojone. W tym przypadku zadanie to

można sprowadzić do zbilansowanego zadania transportowego wprowadzając fikcyjny magazyn (fikcyjnego dostawcę) A_{m+1} z zapasem (podażą)

$$a_{m+1} = \sum_{j=1}^n b_j - \sum_{i=1}^m a_i$$

i zerowymi kosztami transportowymi do dowolnego sklepu (odbiorcy):

$$c_{m+1,j} = 0, \quad j = 1, \dots, m.$$

W dalszej części będziemy rozpatrywać wyłącznie zbilansowane zadanie transportowe, gdyż – jak zauważyliśmy – nie prowadzi to do ograniczenia ogólności rozważań.

6.1.3. Macierz transportowa

Zbilansowane zadanie transportowe ma następującą postać macierzową

$$\begin{array}{ll} \text{minimalizować} & c^\top x \\ \text{przy ograniczeniach} & Ax = \begin{bmatrix} a \\ b \end{bmatrix} \\ & x \geq 0, \end{array}$$

gdzie $c = (c_{11}, \dots, c_{1n}, \dots, c_{m1}, \dots, c_{mn})^\top$, $x = (x_{11}, \dots, x_{1n}, \dots, x_{m1}, \dots, x_{mn})^\top$, $a = (a_1, \dots, a_m)^\top$, $b = (b_1, \dots, b_n)^\top$ zaś macierz A ma postać

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & \dots & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 1 & 1 & \dots & 1 \\ 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 1 & \dots & 0 & \dots & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 1 & \dots & 0 & 0 & 0 & 1 \end{bmatrix}$$

lub w zapisie blokowym

$$A = \begin{bmatrix} e^\top & 0 & \dots & 0 \\ 0 & e^\top & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e^\top \\ I_n & I_n & \dots & I_n \end{bmatrix},$$

gdzie $e = (1, \dots, 1)^\top \in \mathbb{R}^n$, zaś I_n jest macierzą jednostkową typu $n \times n$. Macierz A typu $(m+n) \times (mn)$ nazywa się *macierzą transportową*.

Oznaczmy przez A^i i -ty wiersz macierzy A , $i = 1, \dots, m+n$, i przez A_{ij} - jej $((i-1)n+j)$ -tą kolumnę, $i = 1, \dots, m$, $j = 1, \dots, n$ (kolumna ta odpowiada zmiennej x_{ij}).

Twierdzenie 6.1.1. *Macierz transportowa A ma rząd $r(A) = m+n-1$.*

Dowód. Zachodzi oczywista równość

$$\sum_{i=1}^m A^i = \sum_{i=m+1}^{m+n} A^i.$$

Wiersze macierzy A są więc liniowo zależne, w konsekwencji $r(A) < m+n$. Ponadto macierz ta zawiera $m+n-1$ liniowo niezależnych kolumn $A_{1n}, A_{2n}, \dots, A_{mn}, A_{11}, A_{12}, \dots, A_{1,n-1}$, ponieważ w macierzy utworzonej z tych kolumn:

$$[A_{1n}, A_{2n}, \dots, A_{mn}, A_{11}, A_{12}, \dots, A_{1,n-1}] =$$

$$\begin{bmatrix} 1 & 0 & \dots & 0 & 1 & 1 & \dots & 1 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 \\ 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \end{bmatrix}$$

pierwsze $m + n - 1$ wierszy jest liniowo niezależnych (macierz utworzona z ostatniej macierzy przez skreślenie ostatniego wiersza ma postać

$$A_B = \begin{bmatrix} I_m & R \\ 0 & I_{n-1} \end{bmatrix},$$

w konsekwencji jest ona nieosobliwa). □

Zadanie transportowe można teoretycznie rozwiązać przy pomocy algorytmu sympleksowego. Byłoby to jednak zbyt pracochłonne – zauważmy, że już dla $m = n = 10$ długa forma tablicy sympleksowej ma 20 wierszy i 100 kolumn. Dzięki szczególnej postaci macierzy transportowej A zadanie transportowe można rozwiązać znacznie prościej, niemniej jednak główne etapy wyznaczania rozwiązania są takie same, jak w metodzie sympleksowej: wyznaczenie bazowego rozwiązania dopuszczalnego, a następnie wyznaczenie bazowego rozwiązania optymalnego. Ponadto, jak się później przekonamy, w drugim etapie będziemy otrzymywać kolejno te same rozwiązania bazowe co w drugiej fazie metody sympleksowej, ale przy mniejszym nakładzie obliczeniowym.

Do opisu odpowiednich metod służących rozwiązywaniu zadania transportowego używa się tak zwanej *tablicy transportowej*, która ma następującą postać:

| | B ₁ | B ₂ | ... | B _n | |
|----------------|----------------|----------------|-----|----------------|-------|
| A ₁ | x_{11} | x_{12} | ... | x_{1n} | a_1 |
| A ₂ | x_{21} | x_{22} | ... | x_{2n} | a_2 |
| ... | ... | ... | ... | ... | ... |
| A _m | x_{m1} | x_{m2} | ... | x_{mn} | a_m |
| | b_1 | b_2 | ... | b_n | |

Tablica ta odpowiada rozwiązaniu dopuszczalnemu zbilansowanego zadania transportowego dokładnie wtedy, gdy wszystkie x_{ij} są nieujemne, ostatni element dowolnego wiersza (kolumny) jest równy sumie poprzednich elementów tego wiersza (tej kolumny) i suma wszystkich elementów ostatniego wiersza równa jest sumie wszystkich elementów ostatniej kolumny.

Opiszemy teraz, kiedy tablica transportowa przedstawia rozwiązanie bazowe. Niech $I = \{1, \dots, m\}$, $J = \{1, \dots, n\}$, oraz K niech będzie podzbiorem zbioru $I \times J$:

$$K \subset I \times J = \{(i, j) : i = 1, \dots, m, j = 1, \dots, n\}.$$

Podzbiór $\{i\} \times J$ nazywa się *wierszem zbioru $I \times J$* , $i = 1, \dots, m$, zaś podzbiór $I \times \{j\}$, – *kolumną zbioru $I \times J$* , $j = 1, \dots, n$.

Zbiorowi K przyporządkowany jest wzajemnie jednoznacznie podzbiór A_K zbioru kolumn macierzy transportowej A :

$$A_K = \{A_{ij} : (i, j) \in K\}.$$

Definicja 6.1.2. Podzbiór K nazywa się *cyklem*, jeśli w każdym wierszu i w każdej kolumnie zbioru $I \times J$ znajdują się 0 lub 2 elementy zbioru K .

Przykład 6.1.3. Elementy podzbioru $K \subset I \times J$ zostały w poniższej tablicy oznaczone symbolem \times .

| $I \setminus J$ | 1 | 2 | 3 | 4 | 5 | 6 |
|-----------------|----------|----------|---|----------|---|----------|
| 1 | \times | | | \times | | |
| 2 | | | | | | |
| 3 | | \times | | | | \times |
| 4 | \times | \times | | | | |
| 5 | | | | | | |
| 6 | | | | \times | | \times |

Podzbiór K jest cyklem.

Twierdzenie 6.1.4. Układ kolumn A_L macierzy transportowej A , gdzie $L \subset I \times J$, jest liniowo zależny wtedy i tylko wtedy, gdy podzbiór $L \subset I \times J$ zawiera pewien cykl K .

Dowód powyższego twierdzenia można znaleźć na przykład w książce W. Grabowskiego [9].

Uwaga 6.1.5. Dana jest tablica transportowa

| | B ₁ | B ₂ | ... | B _n | |
|----------------|----------------|----------------|-----|----------------|-------|
| A ₁ | x_{11} | x_{12} | ... | x_{1n} | a_1 |
| A ₂ | x_{21} | x_{22} | ... | x_{2n} | a_2 |
| ... | ... | ... | ... | ... | ... |
| A _m | x_{m1} | x_{m2} | ... | x_{mn} | a_m |
| | b_1 | b_2 | ... | b_n | |

która przedstawia rozwiązanie dopuszczalne zadania transportowego. Zgodnie z twierdzeniami 6.1.1 i 6.1.4 rozwiązanie to jest rozwiązaniem bazowym dokładnie wtedy, gdy podzbiór

$$L = \{(i, j) \in I \times J : x_{ij} \neq 0\}$$

posiada co najwyżej $m + n - 1$ elementów i nie zawiera cyklu. Jeśli ponadto zbiór L posiada dokładnie $m + n - 1$ elementów, to rozwiązanie to jest niezdegenerowane.

Przykład 6.1.6. a) Macierz transportowa

| | B ₁ | B ₂ | B ₃ | B ₄ | |
|----------------|----------------|----------------|----------------|----------------|----|
| A ₁ | 6 | 4 | | | 10 |
| A ₂ | | 1 | 7 | | 8 |
| A ₃ | | | 1 | 6 | 7 |
| | 6 | 5 | 8 | 6 | |

przedstawia niezdegenerowane dopuszczalne rozwiązanie bazowe (puste miejsca oznaczają zera).

b) Macierz transportowa

| | B ₁ | B ₂ | B ₃ | B ₄ | |
|----------------|----------------|----------------|----------------|----------------|----|
| A ₁ | 5 | | | 5 | 10 |
| A ₂ | | | 8 | | 8 |
| A ₃ | 1 | 5 | | 1 | 7 |
| | 6 | 5 | 8 | 6 | |

przedstawia rozwiązanie dopuszczalne, które jednak nie jest rozwiązaniem bazowym.

c) Macierz transportowa

| | B ₁ | B ₂ | B ₃ | B ₄ | |
|----------------|----------------|----------------|----------------|----------------|----|
| A ₁ | 5 | 5 | | | 10 |
| A ₂ | | | 8 | | 8 |
| A ₃ | 1 | | 0 | 6 | 7 |
| | 6 | 5 | 8 | 6 | |

przedstawia dopuszczalne rozwiązanie bazowe. Zmienna x_{33} jest zmienną bazową, ale ponieważ przyjmuje ona wartość zero, więc rozwiązanie przedstawione za pomocą tej tablicy jest zdegenerowane.

6.1.4. Dualne zadanie transportowe

Zadanie dualne do zbilansowanego zadania transportowego ma postać

$$\begin{array}{ll}
 \text{maksymalizować} & \sum_{i=1}^m a_i u_i + \sum_{j=1}^n b_j v_j \\
 \text{przy ograniczeniach} & u_i + v_j \leq c_{ij}, \quad i \in I, j \in J \\
 & u_i, v_j \in \mathbb{R}, \quad i \in I, j \in J.
 \end{array}$$

Uzasadnienie tego faktu pozostawiamy czytelnikowi.

Uwaga 6.1.7. Niech X będzie rozwiązaniem dopuszczalnym zadania transportowego i niech (u, v) będzie rozwiązaniem dopuszczalnym dualnego zadania transportowego. Z twierdzenia o komplementarności wynika, że X i (u, v) są rozwiązaniami optymalnymi odpowiednio zadania pierwotnego i dualnego wtedy i tylko wtedy, gdy

$$x_{ij}(c_{ij} - u_i - v_j) = 0 \text{ dla wszystkich } (i, j) \in I \times J.$$

Własność ta będzie później wykorzystana przy wyznaczaniu rozwiązania optymalnego zadania transportowego.

6.2. Wyznaczanie dopuszczalnego rozwiązania bazowego

6.2.1. Opis ogólny

Opiszemy teraz, jak można otrzymać początkowe dopuszczalne rozwiązanie bazowe dla zadania transportowego (6.2).

Najpierw rozpatrzmy przypadki trywialne:

- $m = 1$

W tym przypadku $I = \{1\}$ i wektor $x = (x_{11}, \dots, x_{1n})$ o współrzędnych

$$x_{ij} = b_j, \quad j = 1, \dots, n, \quad i \in I, \quad (6.3)$$

jest jedynym rozwiązaniem dopuszczalnym zadania transportowego. Rozwiązanie to jest więc jednocześnie rozwiązaniem optymalnym.

- $n = 1$

W tym przypadku $J = \{1\}$ i wektor $x = (x_{11}, \dots, x_{m1})$ o współrzędnych

$$x_{ij} = a_i, \quad i = 1, \dots, m, \quad j \in J, \quad (6.4)$$

jest jedynym rozwiązaniem dopuszczalnym zadania transportowego. Podobnie jak w poprzednim przypadku, rozwiązanie to jest jednocześnie rozwiązaniem optymalnym.

Opiszemy teraz ogólną postać metody wyznaczania początkowego dopuszczalnego rozwiązania bazowego dla dowolnych m i n . Niech $I = \{1, \dots, m\}$, $J = \{1, \dots, n\}$ i niech $K = \emptyset$.

Schemat iteracyjny 6.2.1**Krok 1.**

- Jeśli $\#J = 1$, to:
 - 1.1. położyć $K := K \cup I \times J$,
 - 1.2. położyć $x_{ij} := a_i$, $(i, j) \in I \times J$,
 - 1.3. algorytm zatrzymuje się: X jest dopuszczalnym rozwiązaniem bazowym.
- Jeśli $\#I = 1$, to:
 - 1.4. położyć $K := K \cup I \times J$,
 - 1.5. położyć $x_{ij} := b_j$, $(i, j) \in I \times J$,
 - 1.6. algorytm zatrzymuje się: X jest dopuszczalnym rozwiązaniem bazowym.
- Jeśli $\min\{\#I, \#J\} \geq 2$, to:
 - 1.7. wybrać parę $(p, q) \in I \times J$,
 - 1.8. położyć $K := K \cup \{(p, q)\}$,
 - 1.9. położyć $x_{pq} := \min\{a_p, b_q\}$.

Krok 2.

- Jeśli $a_p < b_q$, to:
 - 2.1. położyć $b_q := b_q - a_p$ i $a_p := 0$,
 - 2.2. położyć $I := I \setminus \{p\}$ (p -ty dostawca zostaje wyeliminowany).
- Jeśli $a_p \geq b_q$, to:
 - 2.3. położyć $a_p := a_p - b_q$ i $b_q := 0$,

2.4. położyć $J := J \setminus \{q\}$ (q -ty odbiorca zostaje wyeliminowany).

- Ze zredukowanym zadaniem transportowym przejść do kroku 1.

Twierdzenie 6.2.2. *Opisana w powyższym schemacie iteracyjnym metoda pozwala na wyznaczenie dopuszczalnego rozwiązania bazowego X w co najwyżej $m + n - 2$ iteracjach.*

Dowód można znaleźć na przykład w podręczniku [9].

Uwaga 6.2.3. Jeśli w kroku 1.7. za każdym razem wybiera się parę $(p, q) \in I \times J$, taką, że $\min\{a_p, b_q\} > 0$, to otrzymane rozwiązanie jest niezdegenerowane.

Przykład 6.2.4. Rozpatrzmy zadanie transportowe z 3 dostawcami i 4 odbiorcami, dla którego popyt, podaż oraz macierz kosztów podane są w poniższej tabeli:

| | B ₁ | B ₂ | B ₃ | B ₄ | |
|----------------|----------------|----------------|----------------|----------------|----|
| A ₁ | 7 | 2 | 4 | 7 | 10 |
| A ₂ | 9 | 5 | 3 | 3 | 8 |
| A ₃ | 7 | 7 | 6 | 4 | 7 |
| | 6 | 5 | 8 | 6 | |

Jeśli wybierzemy w kolejnych iteracjach opisanego wyżej schematu $(p, q) = (1, 3), (2, 1), (3, 2)$ w kroku 1.7, to otrzymamy następujące dopuszczalne rozwiązanie bazowe

$$X = \begin{bmatrix} 0 & 0 & 8 & 2 \\ 6 & 0 & 0 & 2 \\ 0 & 5 & 0 & 2 \end{bmatrix}.$$

Podamy teraz konkretne algorytmy pozwalające wyznaczyć dopuszczalne rozwiązanie bazowe. Są one szczególnymi przypadkami opisanego wyżej schematu iteracyjnego. Dokładniej, w każdym z podanych algorytmów, w kroku 1.7 powyższego schematu iteracyjnego para (p, q) wyznaczana jest w specjalny sposób.

6.2.2. Metoda kąta północno-zachodniego

Krok 1.7 ma w *metodzie kąta północno-zachodniego* postać:

1.7. wybrać taką parę $(p, q) \in I \times J$, że

$$p = \min\{i : i \in I\}, \quad q = \min\{j : j \in J\}.$$

Innymi słowy, wybieramy pary magazyn-sklep w kolejności zgodnej z ich numeracją.

Przykład 6.2.5. Rozpatrzmy zadanie transportowe z podażą, popytem i macierzą kosztów podanymi w przykładzie 6.2.4 Postępując zgodnie z metodą kąta północno-zachodniego otrzymujemy następujące dopuszczalne rozwiązanie bazowe:

$$X = \begin{bmatrix} 6 & 4 & & \\ & 1 & 7 & \\ & & 1 & 6 \end{bmatrix}.$$

Funkcja celu ma dla tego rozwiązania wartość $z = 106$.

Sprawdzimy teraz, czy jest to rozwiązanie optymalne. Zgodnie z uwagą 6.1.7, X jest optymalne dokładnie wtedy, gdy

$$x_{ij}(c_{ij} - u_i - v_j) = 0 \text{ dla wszystkich } (i, j) \in I \times J,$$

gdzie (u, v) jest rozwiązaniem dopuszczalnym zadania dualnego do rozpatrywanego zadania transportowego. Dla podanych współrzędnych x_{ij} rozwiązania X zachodzić więc muszą poniższe równości

$$u_1 + v_1 = 7$$

$$u_1 + v_2 = 2$$

$$u_2 + v_2 = 5$$

$$u_2 + v_3 = 3$$

$$u_3 + v_3 = 6$$

$$u_3 + v_4 = 4.$$

Wykorzystamy posiadany „stopień swobody” ($r(A) = 3 + 4 - 1 = 6$) przypisując jednej ze zmiennych dowolną wartość, na przykład $v_1 = 0$. Po prostych rachunkach otrzymamy rozwiązanie powyższego układu:

$$(u, v) = (7, 10, 13, 0, -5, -7, -9).$$

Teraz sprawdzimy, czy spełnione są pozostałe ograniczenia zadania dualnego. W tym celu musimy wyznaczyć tak zwane *ujemne koszty zredukowane*

$$\bar{c}_{ij} = c_{ij} - u_i - v_j,$$

gdzie (i, j) odpowiadają zmiennym niebazowym. Jeśli wszystkie \bar{c}_{ij} są nieujemne, to rozwiązanie jest optymalne i odwrotnie. Jednak w naszym przypadku tak nie jest, bo na przykład $\bar{c}_{31} = -6$. Tak więc otrzymane rozwiązanie nie jest optymalne.

6.2.3. Metoda elementu minimalnego macierzy kosztów

Krok 1.7 ma w metodzie *elementu minimalnego macierzy kosztów* postać:

1.7. wybrać parę $(p, q) \in I \times J$, taką, że

$$c_{pq} = \min\{c_{ij} : i \in I, j \in J\}.$$

Innymi słowy, wybieramy kolejno pary magazyn-sklep o minimalnym koszcie transportowym. Można się więc spodziewać, że otrzymane w ten sposób rozwiązanie dopuszczalne będzie lepsze niż w metodzie kąta północno-zachodniego.

Przykład 6.2.6. Rozpatrzmy zadanie transportowe z podażą, popytem i macierzą kosztów podanymi w przykładzie 6.2.4. Postępując wedle metody elementu minimalnego macierzy kosztów otrzymujemy następujące dopuszczalne rozwiązanie bazowe:

$$X = \begin{bmatrix} 5 & 5 & & \\ & & 8 & 0 \\ 1 & & & 6 \end{bmatrix}.$$

Funkcja celu ma dla tego rozwiązania wartość $z = 100$. Zmienne dualne możemy wyznaczyć podobnie jak w przykładzie 6.2.5. Otrzymamy układ równań

$$\begin{aligned} u_1 + v_1 &= 7 \\ u_1 + v_2 &= 2 \\ u_2 + v_3 &= 3 \\ u_2 + v_4 &= 3 \\ u_3 + v_1 &= 7 \\ u_3 + v_4 &= 4. \end{aligned}$$

Jeśli położymy $u_1 = 0$, to otrzymamy rozwiązanie $(u, v) = (0, -1, 0, 7, 2, 4, 4)$. Nietrudno sprawdzić, że wszystkie ujemne koszty zredukowane są w tym przypadku nieujemne, a więc X jest rozwiązaniem optymalnym.

6.2.4. Metoda aproksymacyjna Vogla

Krok 1.7 ma w metodzie *aproksymacyjnej Vogla* postać:

1.7.a) Dla każdego $i \in I$ wyznaczyć różnicę $dz_i = c_{il} - c_{ik}$ między drugim co do wielkości (licząc od najmniejszego) elementem c_{il} i najmniejszym elementem c_{ik} i -tego wiersza macierzy kosztów zredukowanego zadania transportowego ($k, l \in J$).

1.7.b) Dla każdego $j \in J$ wyznaczyć różnicę $ds_j = c_{lj} - c_{kj}$ między drugim co do wielkości (licząc od najmniejszego) elementem c_{lj} i najmniejszym elementem c_{kj} j -tej kolumny macierzy kosztów zredukowanego zadania transportowego ($k, l \in I$).

1.7.c) Wyznaczyć

$$\max\{dz_i, ds_j : i \in I, j \in J\}.$$

Jeśli maksimum to jest osiągnięte dla p -tego wiersza, to wybrać parę (p, q) , dla której

$$c_{pq} = \min\{c_{pj} : j \in J\}$$

i przejść do kroku 1.8. Jeśli natomiast maksimum to jest osiągnięte dla q -tej kolumny, to wybrać parę (p, q) , dla której

$$c_{pq} = \min\{c_{iq} : i \in I\}.$$

Przykład 6.2.7. Rozpatrzmy zadanie transportowe z podażą, popytem i macierzą kosztów podanymi w przykładzie 6.2.4. Postępując zgodnie z metodą aproksymacyjną Vogla otrzymujemy następujące dopuszczalne rozwiązanie bazowe

$$X = \begin{bmatrix} & 5 & 5 \\ & 3 & 5 \\ 6 & & 1 \end{bmatrix}.$$

Funkcja celu ma dla tego rozwiązania wartość $z = 100$. Z przykładu 6.2.6 wynika, że wartość ta jest minimalna, tak więc X jest rozwiązaniem optymalnym.

6.3. Algorytm transportowy

Mówi się często, że liczby rządzą światem. Pewne jest tylko: liczby pokazują, jak świat jest rządzony.

[J.W. Goethe]

6.3.1. Podstawowe własności

Dane jest zbilansowane zadanie transportowe (6.2) i bazowe rozwiązanie dopuszczalne X . Niech

$$x_{i_1 j_1}, x_{i_2 j_2}, \dots, x_{i_p j_p}$$

będą zmiennymi bazowymi, gdzie $p = m + n - 1$ i niech x_{ij} będzie zmienną niebazową.

Twierdzenie 6.3.1. *Zbiór*

$$\{(i, j), (i_1, j_1), (i_2, j_2), \dots, (i_p, j_p)\}$$

zawiera dokładnie jeden cykl

$$L = \{(i, j), (i, l_1), (k_1, l_1), (k_1, l_2), \dots, (k_r, l_r), (k_r, j)\}.$$

Dowód pomijamy.

Uwaga 6.3.2. Cykl L nazywa się *pętlą bazową*. Liczba elementów cyklu jest oczywiście parzysta. Zbiory

$$L^+ = \{(i, j), (k_1, l_1), \dots, (k_r, l_r)\}$$

i

$$L^- = \{(i, l_1), (k_1, l_2), \dots, (k_r, j)\}$$

nazywają się *półcyklami*. Każdy element półcyklu L^+ jest końcem pionowej krawędzi cyklu L i każdy element półcyklu L^- jest końcem poziomej krawędzi cyklu L .

Przykład 6.3.3. Rozpatrzmy zadanie transportowe z przykładu 6.2.4 i rozwiązanie bazowe

$$X = \begin{bmatrix} 6 & 4 & & \\ & 1 & 7 & \\ & & 1 & 6 \end{bmatrix}.$$

Niech $(i, j) = (3, 1)$. Oczywiście x_{31} jest zmienną niebazową. Pętla bazowa L ma postać

$$L = \{(3, 1), (3, 3), (2, 3), (2, 2), (1, 2), (1, 1)\},$$

natomiast półcykle mają postaci

$$L^+ = \{(3, 1), (2, 3), (1, 2)\}$$

i

$$L^- = \{(3, 3), (2, 2), (1, 1)\}.$$

Uwaga 6.3.4. Dla bazowego rozwiązania dopuszczalnego X ze zmiennymi bazowymi $x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_pj_p}$ zachodzą równości

$$I = \{i_1, i_2, \dots, i_p\}$$

i

$$J = \{j_1, j_2, \dots, j_p\}.$$

Równości te oznaczają po prostu, że żaden z dostawców i żaden z odbiorców nie został pominięty w planie transportowym X .

Przykład 6.3.5. Rozpatrzmy układ równań podany w przykładzie 6.2.6. Macierz współczynników tego układu ma postać

$$\begin{bmatrix} 1 & & & 1 & & & \\ & 1 & & & 1 & & \\ & & 1 & & & 1 & \\ & & & 1 & & & 1 \\ & & & & 1 & & \\ & & & & & 1 & \\ & & & & & & 1 \end{bmatrix}.$$

Widzimy, że wiersze tej macierzy są jednocześnie kolumnami macierzy transportowej

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & & & & & \\ & & & & 1 & 1 & 1 & 1 & \\ & & & & & & & 1 & 1 & 1 & 1 \\ 1 & & & & 1 & & & 1 & & & \\ & 1 & & & & 1 & & & 1 & & \\ & & 1 & & & & 1 & & & 1 & \\ & & & 1 & & & & 1 & & & 1 \end{bmatrix}$$

odpowiadającymi zmiennym bazowym $x_{11}, x_{12}, x_{22}, x_{23}, x_{33}, x_{34}$. Nie jest to przypadek – własność ta zachodzi dla dowolnego rozwiązania bazowego zbilansowanego zadania transportowego.

Twierdzenie 6.3.6. *Niech X będzie bazowym rozwiązaniem dopuszczalnym ze zmiennymi bazowymi $x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_pj_p}$. Układ równań*

$$\begin{aligned} u_{i_1} + v_{j_1} &= c_{i_1j_1} \\ u_{i_2} + v_{j_2} &= c_{i_2j_2} \\ &\dots \\ u_{i_p} + v_{j_p} &= c_{i_pj_p} \end{aligned} \tag{6.5}$$

posiada nieskończenie wiele rozwiązań zależnych od jednego parametru.

Dowód. Macierzą współczynników układu równań jest A_B^\top , gdzie A_B jest podmacierzą macierzy transportowej A , której (liniowo niezależne) kolumny odpowiadają zmiennym bazowym $x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_pj_p}$. Zatem

$$r(A_B^\top) = p = m + n - 1.$$

Łatwo zauważyć, że rząd macierzy rozszerzonej układu (6.5) jest również równy p , gdyż jest to układ p równań. Na mocy twierdzenia Kroneckera–Capellego układ ten posiada nieskończenie wiele rozwiązań zależnych od $m + n - p = 1$ parametru (liczba zmiennych w układzie jest równa $m + n$ na mocy uwagi 6.3.4). \square

Twierdzenie 6.3.7. *Ujemne koszty zredukowane $\bar{c}_{ij} = c_{ij} - u_i - v_j$ są niezależne od przyjętego rozwiązania układu równań (6.5).*

Dowód. Niech

$$L = \{(i, j), (i, l_1), (k_1, l_1), (k_1, l_2), \dots, (k_r, l_r), (k_r, j)\}$$

będzie pętlą bazową. Zachodzą równości

$$\begin{aligned} \bar{c}_{ij} &= c_{ij} - u_i - v_j \\ &= c_{ij} - (u_i + v_{l_1}) + (u_{k_1} + v_{l_1}) - (u_{k_1} + v_{l_2}) + \dots \\ &\quad + (u_{k_r} + v_{l_r}) - (u_{k_r} + v_j) \\ &= c_{ij} - c_{il_1} + c_{k_1l_1} - c_{k_1l_2} + \dots + c_{k_rl_r} - c_{k_rj}. \end{aligned}$$

\square

Uwaga 6.3.8. Zgodnie z dowodem powyższego twierdzenia, ujemne koszty zredukowane można również wyznaczyć ze wzoru

$$\bar{c}_{ij} = c_{ij} - c_{il_1} + c_{k_1l_1} - c_{k_1l_2} + \dots + c_{k_rl_r} - c_{k_rl_j}.$$

6.3.2. Ogólny opis algorytmu transportowego

Startując z dopuszczalnego rozwiązania bazowego (które można otrzymać jedną z podanych uprzednio metod) algorytm transportowy pozwala na wyznaczenie rozwiązania optymalnego po wykonaniu skończenia wielu iteracji. Podobnie jak w algorytmie sympleksowym, każda iteracja polega na wymianie bazy: jedna ze zmiennych bazowych zostaje zastąpiona jedną ze zmiennych niebazowych. Aby wybrać zmienną niebazową, która zostanie wzięta do bazy, należy wyznaczyć ujemne koszty zredukowane $\bar{c}_{ij} = c_{ij} - u_i - v_j$ dla wszystkich zmiennych niebazowych. W tym celu należy najpierw wyznaczyć zmienne dualne u_i i v_j , $i = 1, \dots, m$, $j = 1, \dots, n$, które odpowiadają aktualnej bazie. Można je wyznaczyć korzystając z twierdzenia o komplementarności (patrz uwaga 6.1.7):

$$x_{ij} \text{ jest zmienną bazową} \Rightarrow c_{ij} = u_i + v_j.$$

Do bazy można włączyć zmienną bazową, której ujemne koszty zredukowane są ujemne lub przynajmniej niedodatnie (wybór ten odpowiada wyborowi kolumny głównej w algorytmie sympleksowym). Najczęściej wybiera się zmienną niebazową, której odpowiada najmniejszy ujemny koszt zredukowany. Następnie wybiera się zmienną bazową, która przy danej wymianie bazy opuszcza tę bazę. W tym celu wyznacza się pętlę bazową dla aktualnej bazy i dla ustalonej już zmiennej niebazowej, która ma być wzięta do bazy (patrz twierdzenie 6.3.1). Wybrać należy najmniejszą zmienną bazową x_{ij} taką, że $(i, j) \in L^-$ (odpowiada to wyborowi wiersza głównego w algorytmie sympleksowym). Przy wymianie bazy, wartości zmiennych x_{ij} odpowiadających pętli bazowej zostają wyznaczone tak, aby zmienna opuszczająca bazę przyjęła wartość 0 i aby spełnione były nadal wszystkie ograniczenia (odpowiada to piwotyzacji w algorytmie sympleksowym).

6.3.3. Algorytm transportowy

Dane jest zbilansowane zadanie transportowe z m dostawcami i n odbiorcami. Niech $a \in \mathbb{R}^m$ będzie wektorem zapasów (podaży), $b \in \mathbb{R}^n$ – wektorem zapotrzebowań (popytu) i C – macierzą kosztów.

Algorytm 6.3.9 (transportowy)

Krok 1. Wyznaczyć dopuszczalne rozwiązanie bazowe X i odpowiadającą mu bazę B (zbiór indeksów zmiennych bazowych).

Krok 2. (*wyznaczenie zmiennych dualnych*). Wyznaczyć dowolne rozwiązanie układu $m + n - 1$ równań

$$c_{ij} = u_i + v_j, (i, j) \in B$$

z $m + n$ niewiadomymi (na przykład nadać jednej ze zmiennych ustaloną wartość i wyznaczyć pozostałe $m + n - 1$ zmienne z układu równań).

Krok 3. (*kryterium zatrzymania*)

(a) Wyznaczyć ujemne koszty zredukowane

$$\bar{c}_{ij} = c_{ij} - u_i - v_j$$

dla $(i, j) \notin B$.

(b) Jeśli wszystkie \bar{c}_{ij} są nieujemne, to X jest rozwiązaniem optymalnym i algorytm zatrzymuje się.

Krok 4. (*wymiana bazy*)

(a) Wybrać parę $(i_0, j_0) \notin B$, taką że

$$\bar{c}_{i_0 j_0} = \min\{\bar{c}_{ij} : (i, j) \notin B\}$$

(para (i_0, j_0) zostanie włączona do bazy).

(b) Wyznaczyć pętlę bazową

$$L = \{(i_0, j_0), (i_0, l_1), (k_1, l_1), (k_1, l_2), \dots, (k_r, l_r), (k_r, j_0)\}$$

zawartą w zbiorze $\{(i_0, j_0)\} \cup B$ i odpowiadające jej półcykle L^+ i L^- .

(c) Wybrać parę $(i_k, j_k) \in L^-$, dla której

$$x_{i_k j_k} = \Delta = \min\{x_{ij} : (i, j) \in L^-\}$$

(para (i_k, j_k) zostanie usunięta z bazy).

(d) Położyć

$$x_{i_0 j_0} = \Delta.$$

(e) Położyć

$$x_{ij} = x_{ij} - \Delta$$

dla wszystkich $(i, j) \in L^-$.

(f) Położyć

$$x_{ij} = x_{ij} + \Delta$$

dla wszystkich $(i, j) \in L^+$.

(g) Z nowym rozwiązaniem bazowym X przejść do kroku 2.

Twierdzenie 6.3.10. *Jeśli w trakcie realizacji algorytmu transportowego nie wystąpi degeneracja, to rozwiązanie optymalne otrzymuje się w skończonej liczbie iteracji.*

Dowód twierdzenia jest podobny do dowodu skończonej zbieżności algorytmu sympleksowego.

Uwaga 6.3.11. Do tej pory nie pokazano skończonej zbieżności algorytmu transportowego w przypadku wystąpienia degeneracji, jak i również nie znaleziono przykładu zadania transportowego z występującą w trakcie realizacji algorytmu transportowego degeneracją, dla którego algorytm ten nie byłby zbieżny w skończonej liczbie kroków.

Uwaga 6.3.12. Ponieważ w trakcie realizacji algorytmu transportowego jedynymi działaniami arytmetycznymi są dodawanie i odejmowanie, więc otrzymane rozwiązanie będzie całkowitoliczbowe, o ile zapasy i zapotrzebowania będą liczbami całkowitymi.

Przykład 6.3.13. Rozpatrzmy zadanie transportowe określone w przykładzie 6.2.4 i rozwiązanie bazowe tego zadania

$$X = \begin{bmatrix} 6 & 4 & & \\ & 1 & 7 & \\ & & 1 & 6 \end{bmatrix}.$$

Za pomocą algorytmu transportowego wyznaczmy rozwiązanie optymalne. Otrzymany kolejno tablice transportowe, w których naniesiono wartości zmiennych bazowych (tłusty druk), wartości zmiennych dualnych odpowiadających zmiennym bazowym i wartości ujemnych kosztów zredukowanych odpowiadających zmiennym niebazowym. W tablicach tych w lewym górnym rogu każdego okienka podano odpowiednie wartości macierzy kosztów. Każdorazowo pod tablicą podano wartości funkcji celu.

| v | 7 | 2 | 0 | -2 | |
|-----|---------------|---------------|---------------|---------------|------------|
| u | | | | | |
| 0 | 7 6 | 2 4 | 4 4 | 7 9 | 10 |
| 3 | 9 -1 | 5 1 | 3 7 | 3 2 | 8 |
| 6 | 7 -6 | 7 -1 | 6 1 | 4 6 | 7 |
| | 6 | 5 | 8 | 6 | a b |

Wartość funkcji celu $z = 106$.

| v u | 7 | 2 | 0 | 4 | |
|------------|---------------|---------------|---------------|---------------|------------|
| 0 | 7 5 | 2 5 | 4 4 | 7 3 | 10 |
| 3 | 9 -1 | 5 0 | 3 8 | 3 -4 | 8 |
| 0 | 7 1 | 7 5 | 6 6 | 4 6 | 7 |
| | 6 | 5 | 8 | 6 | a b |

Wartość funkcji celu $z = 100$.

| v u | 7 | 2 | 4 | 4 | |
|------------|---------------|---------------|---------------|---------------|------------|
| 0 | 7 5 | 2 5 | 4 0 | 7 3 | 10 |
| -1 | 9 3 | 5 4 | 3 8 | 3 0 | 8 |
| 0 | 7 1 | 7 5 | 6 2 | 4 6 | 7 |
| | 6 | 5 | 8 | 6 | a b |

Wartość funkcji celu $z = 100$.

Ostatnia tablica transportowa przedstawia rozwiązanie optymalne. Widzimy teraz, że druga z kolei tablica przedstawia również rozwiązanie optymalne. Jednak nie zostało to stwierdzone przez algorytm transportowy. Przyczyną jest pojawiająca się pierwotna i dualna degeneracja. Podobnie jak w przypadku algorytmu sympleksowego, możemy – dzięki dualnej degeneracji – wyznaczyć inne rozwiązanie optymalne wybierając w ostatniej

tablicy parę $(i, j) = (1, 3)$, która ma być wprowadzona do bazy. Otrzymamy wówczas tablicę

| v | | 7 | 2 | 4 | 4 | |
|-----|--|---------------|---------------|---------------|---------------|------------|
| u | | | | | | |
| 0 | | 7 0 | 2 5 | 4 5 | 7 3 | 10 |
| -1 | | 9 3 | 5 4 | 3 3 | 3 5 | 8 |
| 0 | | 7 6 | 7 5 | 6 2 | 4 1 | 7 |
| | | 6 | 5 | 8 | 6 | a b |

Tablica ta przedstawia rozwiązanie optymalne dualnie zdegenerowane.

ROZDZIAŁ 7

Programowanie całkowitoliczbowe

Dobry Bóg stworzył liczby całkowite, resztę wymyślili ludzie.

[L. Kronecker]

7.1. Definicje i przykłady wprowadzające

Definicja 7.1.1. Przez zadanie programowania (liniowego) całkowitoliczbowego rozumiemy zadanie maksymalizacji lub minimalizacji (liniowej) funkcji celu przy ograniczeniach (liniowych) równościowych lub nierównościowych określonej w obszarze $\mathbb{Z}^p \times \mathbb{R}^{n-p}$, gdzie $1 \leq p \leq n$, $n, p \in \mathbb{N}$, czyli w obszarze wektorów o niektórych współrzędnych całkowitych. W postaci klasycznej zadanie to ma postać:

$$\begin{array}{ll} \text{maksymalizować} & c^\top x \\ \text{przy ograniczeniach} & Ax \leq b \\ & x \geq 0 \\ & x_i \in \mathbb{Z} \text{ dla } i = 1, \dots, p, \end{array}$$

gdzie $1 \leq p \leq n$. Jeśli $p = n$, to zadanie nazywa się *czysto całkowitoliczbowe*, w przeciwnym wypadku nazywa się ono zadaniem *mieszanym*

całkowitoliczbowym.

Definicja 7.1.2. Przez zadanie *programowania (liniowego) dyskretnego* lub *binarnego* rozumiemy zadanie maksymalizacji lub minimalizacji (liniowej) funkcji celu przy ograniczeniach (liniowych) równościowych lub nierównościowych określonej w obszarze $\{0, 1\}^n$, czyli w obszarze wektorów o współrzędnych binarnych. W postaci klasycznej zadanie to ma postać:

$$\begin{array}{llll} \text{maksymalizować} & c^\top x & & \\ \text{przy ograniczeniach} & Ax & \leq & b \\ & x & \geq & 0 \\ & x_i & \in & \{0, 1\} \text{ dla } i = 1, \dots, n. \end{array}$$

Przykład 7.1.3 (*zagadnienie rozkroju*). Z bel papieru o szerokości 210 cm należy wykroić co najmniej:

- 30 rolek o szerokości $a = 62$ cm,
- 60 rolek o szerokości $b = 55$ cm,
- 60 rolek o szerokości $c = 40$ cm.

Należy wyznaczyć plan rozkroju minimalizujący liczbę użytych bel.

Wśród wszystkich rozkrojów beli papieru można wyróżnić rozkroje efektywne, tzn. takie, dla których pozostałości po rozkroju nie da się dalej rozkroić na rolki o żądanej szerokości. Przykładowo, niech $abcc$ oznacza rozkrój beli na 4 rolki o szerokościach a, b, c, c . Nietrudno sprawdzić, że wszystkie możliwe rozkroje efektywne to: $aaa, aab, aacc, abb, abcc, accc, bbcc, bbcc, bccc, cccc$. Oznaczmy przez x_i liczbę bel krojonych według i -tego z wyżej wymienionych rozkrojów efektywnych. Wówczas zadanie można sformułować jako zadanie programowania liniowego całkowitoliczbowego

w następujący sposób:

$$\begin{array}{rcll}
 \text{minimalizować} & x_1 + \dots + x_{10} & & \\
 \text{przy ograniczeniach} & 3x_1 + 2x_2 + 2x_3 + x_4 + x_5 + x_6 & \geq & 30 \\
 & x_2 + 2x_4 + x_5 + 3x_7 + 2x_8 + x_9 & \geq & 60 \\
 & 2x_3 + 2x_5 + 3x_6 + x_7 + 2x_8 + 3x_9 + 5x_{10} & \geq & 60 \\
 & x_1, \dots, x_{10} & \geq & 0 \\
 & x_1, \dots, x_{10} & \in & \mathbb{Z}.
 \end{array}$$

Przykład 7.1.4 (*całkowitoliczbowe zadanie analizy działalności gospodarczej*). Zakłady lotnicze produkują dwa typy samolotów A i B. W celu wyprodukowania jednego samolotu typu A trzeba zainwestować 7 mln zł, zaś typu B – 6 mln zł. Z uwagi na popyt, liczba wyprodukowanych samolotów typu B nie może przekroczyć 125% liczby wyprodukowanych samolotów typu A. Zakłady mogą zainwestować maksymalnie 42 mln zł. Zysk ze sprzedaży jednego samolotu typu A wynosi 10 mln zł, zaś typu B – 9 mln zł. Ile samolotów typu A i B powinny produkować zakłady, aby zapewnić sobie największy zysk?

Niech x_1 i x_2 będzie liczbą produkowanych samolotów typu A i odpowiednio B. Zadanie to można wówczas sformułować jako zadanie programowania liniowego całkowitoliczbowego w następujący sposób:

$$\begin{array}{rcll}
 \text{maksymalizować} & 10x_1 + 9x_2 & & \\
 \text{przy ograniczeniach} & 7x_1 + 6x_2 & \leq & 42 \\
 & -5x_1 + 4x_2 & \leq & 0 \\
 & x_1, x_2 & \geq & 0 \\
 & x_1, x_2 & \in & \mathbb{Z}.
 \end{array}$$

Przykład 7.1.5 (*zagadnienie przydziału*). n pracownikom należy przydzielić n czynności. Wiadomo, że koszt wykonania j -tej czynności przez i -tego pracownika wynosi c_{ij} , $i = 1, \dots, n$, $j = 1, \dots, n$. Przydziału należy dokonać tak, aby:

a) każdy z pracowników wykonywał dokładnie jedną czynność i każda czynność była przydzielona dokładnie jednemu pracownikowi,

b) plan przydziałów miał minimalny koszt.

Niech

$$x_{ij} = \begin{cases} 1, & \text{gdy } i\text{-temu pracownikowi zostanie przydzielona } j\text{-ta czynność,} \\ 0, & \text{w przeciwnym razie.} \end{cases}$$

Zadanie można wówczas sformułować jako zadanie dyskretnego programowania liniowego w następujący sposób:

$$\begin{array}{ll} \text{minimalizować} & \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \\ \text{przy ograniczeniach} & \sum_{j=1}^n x_{ij} = 1, \quad i = 1, \dots, n \\ & \sum_{i=1}^n x_{ij} = 1, \quad j = 1, \dots, n \\ & x_{ij} \in \{0, 1\}, \quad i, j = 1, \dots, n. \end{array}$$

Zwróćmy przy okazji uwagę na fakt, że zadanie to jest szczególnym przypadkiem zadania transportowego w obszarze zmiennych binarnych. Zgodnie z uwagą 6.3.12 może ono być więc rozwiązane algorytmem transportowym. Istnieją jednak efektywniejsze metody jego rozwiązania.

Przykład 7.1.6 (*zagadnienie plecakowe*). Turysta ładuje plecak dokonując wyboru spośród czterech przedmiotów. Załóżmy, że przedmioty te ważą kolejno 3, 2, 4 i 1 kg. Każdemu z nich turysta przypisuje określoną użyteczność, którą wyraża przy pomocy liczb. Niech użyteczność ta wynosi odpowiednio 3, 4, 2 i 3. Maksymalny ciężar wziętych przedmiotów wynosi 9 kg. Które przedmioty powinien zabrać ze sobą turysta, aby osiągnąć maksymalną użyteczność, pod warunkiem, że ich całkowita waga nie przekracza 9 kg? Przyjmuje się przy tym, że łączna użyteczność przedmiotów jest sumą poszczególnych użyteczności.

Niech

$$x_j = \begin{cases} 1, & \text{jeśli turysta weźmie ze sobą } j\text{-ty przedmiot,} \\ 0, & \text{w przeciwnym razie.} \end{cases}$$

Zadanie można wówczas sformułować jako zadanie dyskretnego programowania liniowego w następujący sposób:

$$\begin{array}{ll} \text{maksymalizować} & 3x_1 + 4x_2 + 2x_3 + 3x_4 \\ \text{przy ograniczeniach} & 3x_1 + 2x_2 + 4x_3 + x_4 \leq 9 \\ & x_j \in \{0, 1\}, \quad j = 1, 2, 3, 4. \end{array}$$

Przykład 7.1.7 (*zagadnienie komiwojażera*). Handlowiec chce odwiedzić n miejscowości (klientów) startując ze swojej miejscowości i wracając tam po skończonej podróży. W jakiej kolejności powinien odwiedzać te miejscowości, aby łączna przebyta trasa (łączne koszty podróży) była(y) minimalna(e).

Niech

$$x_{ij} = \begin{cases} 1, & \text{jeśli bezpośrednio po } i\text{-tej odwiedzana będzie } j\text{-ta miejscowość,} \\ 0, & \text{w przeciwnym razie.} \end{cases}$$

Ponieważ – dla wyznaczonej trasy – bezpośrednio po każdej miejscowości handlowiec może odwiedzić tylko jedną miejscowość, więc dla dowolnego i , $i = 1, \dots, n$, musi być spełniona równość $\sum_{j=1}^n x_{ij} = 1$. Podobnie, bezpośrednio przed wizytą w dowolnej miejscowości handlowiec mógł odwiedzić dokładnie jedną miejscowość. Wobec tego dla dowolnego j , $j = 1, \dots, n$, spełniona jest równość $\sum_{i=1}^n x_{ij} = 1$. Z każdą trasą związana jest pewna permutacja p zbioru n elementowego, którą można określić następująco: jeśli $x_{ij} = 1$, to $p(i) = j$. Wyżej przedstawione równości, które muszą być spełnione dla dowolnej trasy oraz fakt, że $x_{ij} \in \{0, 1\}$ dla $i, j = 1, \dots, n$, gwarantują, że permutacja ta jest dobrze określona. Z warunków zadania wynika, że permutacja ta powinna składać się z jednego cyklu. Zatem permutacje, które rozkładają się na cykle krótsze niż n nie odpowiadają rozwiązaniu dopuszczalnemu. Po tych uwagach możemy przystąpić do przedstawienia powyższego problemu jako zadania dyskretnego programowania liniowego. Niech c_{ij} oznacza odległość (koszt podróży) od i -tej do j -tej miejscowości, $i, j = 1, \dots, n$. Opisany problem można sformułować

w następujący sposób:

$$\begin{array}{ll}
 \text{minimalizować} & \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \\
 \text{przy ograniczeniach} & \sum_{j=1}^n x_{ij} = 1, \quad i = 1, \dots, n \\
 & \sum_{i=1}^n x_{ij} = 1, \quad j = 1, \dots, n \\
 & x_{ij} \in \{0, 1\} \quad i, j = 1, \dots, n \\
 & \text{i krótkie cykle} \\
 & \text{są zabronione.}
 \end{array}$$

Krótki cykl oznacza tutaj cykl o długości krótszej niż n .

Uwaga 7.1.8. Każde dyskretne zadanie programowania (liniowego) można sprowadzić do (liniowego) zadania całkowitoliczbowego zastępując każde ograniczenie $x_i \in \{0, 1\}$ ograniczeniem $0 \leq x_i \leq 1$, $x_i \in \mathbb{Z}$, $i = 1, \dots, n$.

Uwaga 7.1.9. Każde zadanie programowania (liniowego) całkowitoliczbowego z ograniczonymi zmiennymi można sprowadzić do zadania programowania (liniowego) dyskretnego zastępując każdą ograniczoną zmienną całkowitoliczbową $0 \leq x_i \leq d_i$, $x_i \in \mathbb{Z}$ jej rozwinięciem dwójkowym

$$x_i = 2^0 y_1 + 2^1 y_2 + \dots + 2^{r_j-1} y_{r_j}, \quad y_j \in \{0, 1\}, \quad j = 1, \dots, r_j,$$

$i = 1, \dots, n$. W wielu zadaniach optymalizacji zmienne są albo z definicji ograniczone albo wartości je ograniczające można wyznaczyć. Tak więc ograniczoność zmiennych nie jest istotnym zawężeniem rozważań. Bardziej istotnym problemem jest fakt, że omawiane sprowadzenie do zadania programowania dyskretnego nie jest najczęściej efektywne.

Przykład 7.1.10 Rozważmy zadanie programowania całkowitoliczbowego omawiane w przykładzie 7.1.4:

$$\begin{array}{ll}
 \text{maksymalizować} & 10x_1 + 9x_2 \\
 \text{przy ograniczeniach} & 7x_1 + 6x_2 \leq 42 \\
 & -5x_1 + 4x_2 \leq 0 \\
 & x_1, x_2 \geq 0 \\
 & x_1, x_2 \in \mathbb{Z}.
 \end{array}$$

Rozwiązanie optymalne tego zadania bez uwzględnienia całkowitoliczowości zmiennych wynosi $x^* = (2\frac{26}{29}, 3\frac{18}{29}) \simeq (2.90, 3.62)$ z optymalną wartością funkcji celu $z^* \simeq 61.55$. Rozwiązanie to nie jest jednak całkowitoliczowe. Najbliższa od x^* para o współrzędnych całkowitych $x' = (3, 4)$ nie spełnia warunków zadania. Najbliższe od x^* rozwiązanie dopuszczalne wynosi $x'' = (3, 3)$ z wartością funkcji celu $z'' = 57$. Nie jest to jednak rozwiązanie optymalne. Jest nim bowiem $\bar{x} = (6, 0)$ z optymalną wartością funkcji celu $\bar{z} = 60$. Zauważmy przy okazji, że w omawianym przykładzie \bar{x} jest najbardziej odległym od x^* rozwiązaniem dopuszczalnym.

Powyższy przykład wskazuje na to, że wyznaczenie rozwiązania optymalnego zadania programowania całkowitoliczowego wymaga dokładniejszej analizy. W kolejnych ustępach przedstawimy niektóre metody rozwiązywania takich zadań.

7.2. Metoda Gomory'ego

Rozpocznijmy od następującego przykładu.

Przykład 7.2.1. Rozważmy następujące zadanie programowania liniowego całkowitoliczowego:

$$\begin{array}{ll}
 \text{maksymalizować} & x_1 + x_2 \\
 \text{przy ograniczeniach} & x_1 + 2x_2 \leq 32 \\
 & 18x_1 + 3x_2 \leq 224 \\
 & x_1, x_2 \geq 0 \\
 & x_1, x_2 \in
 \end{array} \quad (P)$$

Opuśćmy najpierw ograniczenie całkowitoliczowości zmiennych i rozwiążmy powstałe w ten sposób zadanie (P') za pomocą metody sympleksowej.

Otrzymamy następującą początkową tablicę sympleksową:

| 1 | $-x_1$ | $-x_2$ | |
|-----|--------|--------|---------|
| 0 | -1 | -1 | $= z$ |
| 32 | 1 | 2 | $= u_1$ |
| 224 | 18 | 3 | $= u_2$ |

Po dwóch piwotyzacjach otrzymamy tablicę

| 1 | $-u_2$ | $-u_1$ | |
|----------------|-----------------|-----------------|---------|
| $\frac{63}{3}$ | $\frac{1}{33}$ | $\frac{5}{11}$ | $= z$ |
| $\frac{32}{3}$ | $-\frac{1}{33}$ | $\frac{6}{11}$ | $= x_2$ |
| $\frac{32}{3}$ | $\frac{2}{33}$ | $-\frac{1}{11}$ | $= x_1$ |

(7.1)

Rozwiązanie optymalne zadania (P') $x^* = (10\frac{2}{3}, 10\frac{2}{3})$, $u^* = (0, 0)$ nie spełnia warunku całkowitoliczbowości zmiennych zadania (P). Zauważmy, że w zadaniu (P) zmienne uzupełniające u_1, u_2 są całkowite, gdyż wszystkie dane w tym zadaniu są całkowite.

Z równania

$$x_1 - \frac{1}{11}u_1 + \frac{2}{33}u_2 = 10\frac{2}{3} \quad (7.2)$$

wynika nierówność

$$x_1 - u_1 \leq 10\frac{2}{3},$$

gdyż $-1 = \lceil -\frac{1}{11} \rceil \leq -\frac{1}{11}$ i $0 = \lceil \frac{2}{33} \rceil \leq \frac{2}{33}$. Ponieważ x_1, u_1 są całkowite, ostatnia nierówność jest równoważna nierówności

$$x_1 - u_1 \leq 10$$

lub inaczej – układowi

$$\begin{aligned} x_1 - u_1 + u_3 &= 10 \\ u_3 &\geq 0. \end{aligned}$$

Jeśli teraz od ostatniego równania odejmiemy równanie (7.2), to otrzymamy równanie

$$-\frac{10}{11}u_1 - \frac{2}{33}u_2 + u_3 = -\frac{2}{3}.$$

Zauważmy, że rozwiązanie optymalne zadania (P') nie spełnia ostatniego równania. Równanie to musi być jednak spełnione przez dowolne rozwiązanie optymalne zadania (P). Jeśli dołączymy to równanie do układu przedstawionego za pomocą tablicy (7.1), to otrzymamy tablicę:

| 1 | $-u_2$ | $-u_1$ | |
|----------------|-----------------|------------------|---------|
| $\frac{63}{3}$ | $\frac{1}{33}$ | $\frac{5}{11}$ | $= z$ |
| $\frac{32}{3}$ | $-\frac{1}{33}$ | $\frac{6}{11}$ | $= x_2$ |
| $\frac{32}{3}$ | $\frac{2}{33}$ | $-\frac{1}{11}$ | $= x_1$ |
| $-\frac{2}{3}$ | $-\frac{2}{33}$ | $-\frac{10}{11}$ | $= u_3$ |

Związane z tą tabelą zadanie maksymalizacji (P'') jest równoważne zadaniu (P) i może być rozwiązane na przykład przy pomocy dualnego algorytmu sympleksowego. Po jednej jego iteracji otrzymamy tablicę:

| 1 | $-u_3$ | $-u_1$ | |
|----|-----------------|--------|---------|
| 21 | $\frac{1}{2}$ | 0 | $= z$ |
| 11 | $-\frac{1}{2}$ | 1 | $= x_2$ |
| 10 | 1 | -1 | $= x_1$ |
| 11 | $-\frac{33}{2}$ | 15 | $= u_2$ |

Tablica ta przedstawia rozwiązanie optymalne zadania (P''), $x'' = (10, 11)$. Ponieważ rozwiązanie to jest całkowitoliczbowe, więc jest ono również rozwiązaniem optymalnym zadania (P).

W powyższym przykładzie zastosowaliśmy tzw. *metodę cięć Gomory'ego*, którą opiszemy teraz dla ogólnego przypadku zadania programowania li-

niowego czysto całkowitoliczbowego:

$$\begin{array}{ll}
 \text{maksymalizować} & c^\top x \\
 \text{przy ograniczeniach} & Ax \leq b \\
 & x \geq 0 \\
 & x_i \in \mathbb{Z}, \quad i = 1, \dots, n,
 \end{array} \tag{7.3}$$

przy czym elementy A i b są liczbami całkowitymi. Najpierw należy rozwiązać powyższe zadanie bez uwzględniania ograniczenia całkowitoliczbowości zmiennych. Jeśli rozwiązanie optymalne x^* tego zadania jest całkowitoliczbowe, to jest ono równocześnie rozwiązaniem optymalnym wyjściowego zadania (7.3). W przeciwnym wypadku niech

$$\begin{array}{c|cccc|c}
 & 1 & -r_1 & \dots & -r_n & \\
 \hline
 & h_{00} & h_{01} & \dots & h_{0n} & = z \\
 \hline
 & h_{10} & h_{11} & \dots & h_{1n} & = s_1 \\
 & \dots & \dots & \dots & \dots & \dots \\
 & h_{i0} & h_{i1} & \dots & h_{in} & = s_i \\
 & \dots & \dots & \dots & \dots & \dots \\
 & h_{m0} & h_{m1} & \dots & h_{mn} & = s_m \\
 \hline
 & & & & &
 \end{array} \tag{7.4}$$

będzie tablicą sympleksową odpowiadającą temu rozwiązaniu, przy czym wartość h_{i0} nie jest całkowita. Równanie i -te układu przedstawionego za pomocą tej tablicy ma postać

$$s_i + \sum_{j=1}^n h_{ij} r_j = h_{i0}. \tag{7.5}$$

Ponieważ $[h_{ij}] \leq h_{ij}$ i $r_j \geq 0$, więc zachodzi nierówność

$$s_i + \sum_{j=1}^n [h_{ij}] r_j \leq h_{i0}.$$

W konsekwencji

$$s_i + \sum_{j=1}^n [h_{ij}] r_j \leq [h_{i0}],$$

ponieważ wszystkie zmienne są całkowite. Ostatnia nierówność jest równoważna układowi

$$s_i + \sum_{j=1}^n [h_{ij}] r_j + u_{m+1} = [h_{i0}] \quad (7.6)$$

$$u_{m+1} \geq 0.$$

Jeśli teraz odejmiemy równanie (7.5) od równania (7.6), otrzymamy równanie

$$\sum_{j=1}^n ([h_{ij}] - h_{ij}) r_j + u_{m+1} = [h_{i0}] - h_{i0}, \quad (7.7)$$

przy czym $u_{m+1} \geq 0$. Równanie to „odetnie” rozwiązanie x^* , ponieważ $[h_{i0}] - h_{i0} < 0$ i wszystkie zmienne niebazowe r_j są równe 0. Równanie to musi być jednak spełnione dla każdego dopuszczalnego (czyli całkowitoliczbowego) rozwiązania zadania (7.3) i dla pewnego $u_{m+1} \geq 0$. Jeśli więc równanie to dołączymy jako dodatkowe ograniczenie do wyjściowego zadania, otrzymamy zadanie równoważne. W praktyce równanie to dołącza się do tablicy sympleksowej (7.4). W efekcie otrzymamy tablicę

| | 1 | $-r_1$ | ... | $-r_n$ | |
|--|---------------------|---------------------|-----|---------------------|-------------|
| | h_{00} | h_{01} | ... | h_{0n} | $= z$ |
| | h_{10} | h_{11} | ... | h_{1n} | $= s_1$ |
| | ... | ... | ... | ... | ... |
| | h_{i0} | h_{i1} | ... | h_{in} | $= s_i$ |
| | ... | ... | ... | ... | ... |
| | h_{m0} | h_{m1} | ... | h_{mn} | $= s_m$ |
| | $[h_{i0}] - h_{i0}$ | $[h_{i1}] - h_{i1}$ | ... | $[h_{in}] - h_{in}$ | $= u_{m+1}$ |

która nie przedstawia rozwiązania optymalnego, gdyż $[h_{i0}] - h_{i0} < 0$. Tablicę tę można dalej piwotyzować, najlepiej przy pomocy dualnego algorytmu sympleksowego, z uwagi na dualną dopuszczalność. Jeśli otrzymane rozwiązanie będzie całkowitoliczbowe, to będzie ono rozwiązaniem wyjściowego zadania. W przeciwnym wypadku opisaną procedurę należy powtarzać aż do otrzymania rozwiązania całkowitoliczbowego.

7.3. Metody podziału i ograniczeń

Metody podziału i ograniczeń (ang: *branch and bound methods*) stosuje się do zadań maksymalizacji (bądź minimalizacji) funkcji f na zbiorze skończonym G :

$$\max_{x \in G} f(x)$$

Opiszemy te metody dla zadań maksymalizacji. Jest to cała klasa metod posiadających pewne wspólne cechy. Na pewnej rodzinie $\mathcal{D} \subset 2^G$ zawierającej wszystkie jednoelementowe podzbiory zbioru G definiuje funkcję φ (*ograniczenie górne*) posiadającą następujące własności:

$$1^0 \ (G_1, G_2 \in \mathcal{D} \text{ i } G_1 \subset G_2) \Rightarrow \varphi(G_1) \leq \varphi(G_2),$$

$$2^0 \ a \in G \Rightarrow \varphi(\{a\}) = f(a).$$

Funkcję φ dobiera się specjalnie do rodzaju zadania.

Uwaga 7.3.1. Z własności $1^0 - 2^0$ wynika, że funkcja φ ma również własność

$$x \in D \in \mathcal{D} \Rightarrow f(x) = \varphi(\{x\}) \leq \varphi(D).$$

Z tego powodu nazywa się ona ograniczeniem górnym.

Przykład 7.3.2. Rozważmy następujące zadanie programowania liniowego czysto całkowitoliczbowego

$$\begin{array}{ll} \text{maksymalizować} & c^\top x \\ \text{przy ograniczeniach} & Ax \leq b \\ & 0 \leq x \leq h \\ & x \in \mathbb{Z}^n, \end{array}$$

przy czym $h \in \mathbb{R}^n$. Niech G będzie zbiorem rozwiązań dopuszczalnych tego zadania:

$$G = \{x \in \mathbb{Z}^n : Ax \leq b, \ 0 \leq x \leq h\},$$

Oczywiście G jest zbiorem skończonym. Niech rodzina \mathcal{D} składa się ze wszystkich zbiorów $G(d; g) \subset G$ postaci:

$$G(d; g) = \{x \in \mathbb{Z}^n : Ax \leq b, d \leq x \leq g\},$$

gdzie $d, g \in \mathbb{Z}^n, 0 \leq d \leq g \leq h$. Zdefiniujmy

$$\varphi(G(d; g)) = \max\{c^\top x : Ax \leq b, d \leq x \leq g\}.$$

Nietrudno zauważyć, że funkcja φ ma własności $1^0 - 2^0$, gdzie $f(x) = c^\top x$.

W każdej iteracji dowolnej *metody podziału i ograniczeń* zbiór G rozkłada się na (możliwie rozłączne) podzbiory $G_1, G_2, \dots, G_{m_k} \in \mathcal{D}$:

$$G = \bigcup_{i=1}^{m_k} G_i.$$

Innymi słowy, zadanie $\max_{x \in G} f(x)$ zostaje rozłożone (lub podzielone) na zadania cząstkowe $\max_{x \in G_i} f(x)$, $i = 1, \dots, m_k$ (*branching*).

Następnie wylicza się ograniczenia górne $\varphi(G_i)$ dla zadań cząstkowych

$$\max_{x \in G_i} f(x),$$

$i = 1, 2, \dots, m_k$ (*bounding*). Na podstawie wartości tych ograniczeń górnych rozstrzyga się, które ze zbiorów G_i muszą być dalej rozłożone. Jeśli przy okazji stwierdzi się, że

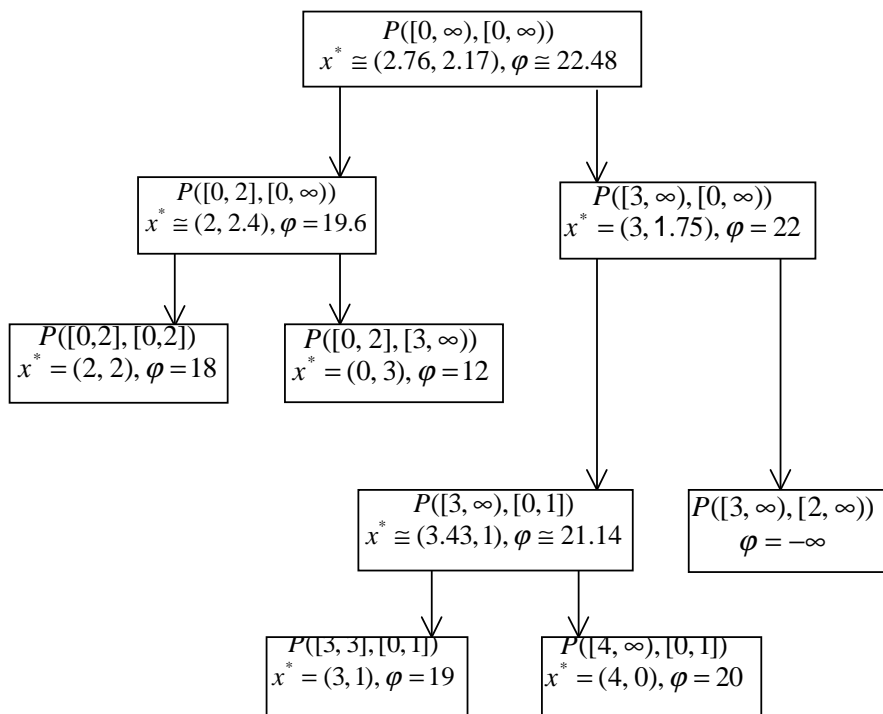
$$\varphi(G_i) = \max_{x \in G_i} f(x),$$

to wartości te mogą służyć jako *ograniczenie dolne* optymalnej wartości funkcji celu $f^* = \max_{x \in G} f(x)$.

Przykład 7.3.3. Rozważmy następujące zadania programowania liniowego całkowitoliczbowego:

$$\begin{array}{ll} \text{maksymalizować} & 5x_1 + 4x_2 \\ \text{przy ograniczeniach} & 7x_1 + 4x_2 \leq 28 \\ & 3x_1 + 10x_2 \leq 30 \\ & x_1, x_2 \geq 0 \\ & x_1, x_2 \in \mathbb{Z}. \end{array} \quad (7.8)$$

Oznaczmy przez $P([a_1, b_1], [a_2, b_2])$ zadanie (7.8) z dodatkowymi ograniczeniami $a_1 \leq x_1 \leq b_1$, $a_2 \leq x_2 \leq b_2$. Dalej, oznaczmy każdorazowo przez φ maksymalną wartość funkcji celu dla zadania cząstkowego $P([a_1, b_1], [a_2, b_2])$ z pominięciem założenia o całkowitoliczbowości zmiennych. Poniższy diagram przedstawia metodę podziału i ograniczeń dla zadania (7.8).



Rozwiązaniem optymalnym zadania (7.8) jest $x^* = (4, 0)$ i optymalna wartość funkcji celu wynosi $\varphi = 20$.

Powróćmy teraz do własności ograniczenia górnego $\varphi : \mathcal{D} \rightarrow \mathbb{R}$:

$$1^0 \ (G_1, G_2 \in \mathcal{D} \text{ i } G_1 \subset G_2) \Rightarrow \varphi(G_1) \leq \varphi(G_2),$$

$$2^0 \ a \in G \Rightarrow \varphi(\{a\}) = f(a).$$

Własności te nie opisują jednoznacznie funkcji φ . Aby otrzymać szybciej rozwiązanie, funkcja φ powinna być dopasowana do konkretnego zadania. Przy konstrukcji rodziny \mathcal{D} i funkcji φ powinny być wzięte pod uwagę następujące kryteria:

1. Wartości funkcji φ powinny być wyznaczane możliwie prosto.
2. Wartości funkcji φ powinny możliwie dokładnie szacować z góry maksymalną wartość funkcji f , tzn. wartość

$$\Delta(G_i) = \varphi(G_i) - \max_{x \in G_i} f(x)$$

powinna być możliwie mała.

Przy wyborze funkcji φ problem polega często na tym, że im lepiej funkcja ta spełnia pierwsze kryterium, tym gorzej spełnione jest kryterium drugie i odwrotnie. Przy konstrukcji funkcji φ należy więc znaleźć odpowiedni kompromis. Najczęściej funkcję $\varphi : \mathcal{D} \rightarrow \mathbb{R}$ definiuje się następująco:

$$\varphi(G_i) = \max_{x \in H_i} f(x),$$

przy czym $G_i \subset H_i$ i podzbiór H_i wybiera się tak, aby $\varphi(G_i)$ można było stosunkowo prosto wyznaczyć (mówi się o *relaksacji* wyjściowego zadania). Na przykład w zadaniu programowania liniowego całkowitoliczbowego lub dyskretnego można opuścić te ograniczenia, których usunięcie prowadzi do zadania programowania liniowego (patrz przykład 7.3.2). W tym przypadku relaksacja taka nazywa się *relaksacją LP*. Inną możliwością jest tak zwana *relaksacja Lagrange'a*. Powstaje ona w ten sposób, że usuwane ograniczenia dodaje się do funkcji celu z odpowiednimi wagami proporcjonalnymi do tak zwanych *mnożników Lagrange'a* (mnożniki te są w istocie równe zmiennym dualnym).

7.3.1. Ogólny opis metod podziału i ograniczeń

Rozważmy zadanie

$$\max_{x \in G} f(x). \quad (7.9)$$

Zakładamy, że rodzina $\mathcal{D} \subset 2^G$ i funkcja $\varphi : \mathcal{D} \rightarrow \mathbb{R}$ zostały zdefiniowane. Ponadto, niech $\varphi(\emptyset) = -\infty$.

Algorytm 7.3.4 (podziału i ograniczeń)

Krok 0. (inicjalizacja)

- (a) Położyć $\mathcal{G} = \{G\}$.
- (b) Wyznaczyć $\varphi(G)$.
- (c) Wybrać $\underline{x} \in G$ i położyć $\underline{\varphi} = f(\underline{x})$ (ewentualnie położyć $\underline{\varphi} = -\infty$).
- (d) Wybrać *kryterium optymalności* $\varepsilon \geq 0$.

Krok 1. (ograniczenie)

- (a) Wyznaczyć

$$\bar{\varphi} = \max_{G_i \in \mathcal{G}} \varphi(G_i).$$

Jeśli $\bar{\varphi} = -\infty$, to zadanie (7.9) jest sprzeczne – algorytm zatrzymuje się.

- (b) Jeśli stwierdzi się, że

$$\varphi(G_i) = f(x_i) \text{ dla pewnego } x_i \in G_i,$$

to usunąć G_i z rodziny \mathcal{G} . Jeśli dodatkowo

$$f(x_i) > \underline{\varphi},$$

to położyć

$$\underline{\varphi} = f(x_i) \text{ i } \underline{x} = x_i$$

($f(x_i)$ jest nowym ograniczeniem dolnym i x_i jest kandydatem na rozwiązanie optymalne).

- (c) Usunąć z rodziny \mathcal{G} te podzbiory G_i , dla których

$$\varphi(G_i) \leq \underline{\varphi}$$

(podzbiory te nie będą więcej rozważane, gdyż nie zawierają one rozwiązania optymalnego).

- (d) Jeśli $\mathcal{G} = \emptyset$, to \underline{x} jest rozwiązaniem optymalnym zadania (7.9) – algorytm zatrzymuje się.
- (e) Jeśli $\bar{\varphi} - \underline{\varphi} \leq \varepsilon$, to \underline{x} jest rozwiązaniem ε -optymalnym zadania (7.9) – algorytm zatrzymuje się.

Krok 2. (podział)

- (a) Wybrać zbiór $\tilde{G} \in \mathcal{G}$ i rozłożyć go na (możliwie rozłączne) podzbiory

$$\tilde{G}_i \in \mathcal{D}, i \in L : \tilde{G} = \bigcup_{i \in L} \tilde{G}_i.$$

- (b) Położyć

$$\mathcal{G} = \mathcal{G} \cup \{\tilde{G}_i : i \in L\} \setminus \{\tilde{G}\}.$$

- (c) Wyznaczyć $\varphi(G_i)$ dla wszystkich nowych $G_i \in \mathcal{G}$.
- (d) Przejść do kroku 1.

Uwaga 7.3.5. Przy zastosowaniu relaksacji LP:

a) można stwierdzić, czy

$$\varphi(G_i) = f(x_i) \text{ dla pewnego } x_i \in G_i,$$

w kroku 1b), jeśli na przykład rozwiązanie optymalne dla zadania zrelaksowanego jest całkowitoliczbowe.

b) w kroku 2a) wybiera się taki zbiór $\tilde{G} = G(d; g) \in \mathcal{G}$, dla którego wartość $\varphi(G(d; g))$ jest osiągnięta dla pewnego $x^* = (x_1^*, \dots, x_n^*)^\top$ z przynajmniej jedną niecałkowitą współrzędną, powiedzmy x_i^* . Wówczas zbiór $\tilde{G} = G(d; g)$ rozkłada się na dwa rozłączne podzbiory

$$\tilde{G}_1 = G(d; g_1, \dots, g_{i-1}, \lfloor x_i^* \rfloor, g_{i+1}, \dots, g_n)$$

i

$$\tilde{G}_2 = G(d_1, \dots, d_{i-1}, \lfloor x_i^* \rfloor + 1, d_{i+1}, \dots, d_n; g),$$

gdzie $\lfloor x \rfloor$ oznacza część całkowitą liczby x .

Ćwiczenie 7.3.6. Prześledzić algorytm 7.3.4 dla zadania programowania liniowego całkowitoliczbowego podanego w przykładzie 7.1.4.

Literatura

- [1] Alevras, D., Padberg, M. W., *Linear Optimization, Problems and Extensions*, Springer-Verlag, Berlin 2001.
- [2] Brdyś, M., Ruszczyński, A., *Metody optymalizacji w zadaniach*, WNT, Warszawa 1985.
- [3] Domschke, W., Drexl, A., *Einführung in Operations Research*, Springer-Verlag, Berlin 1995.
- [4] Findeisen, W., Szymanowski, J., Wierzbicki, A., *Teoria i metody obliczeniowe optymalizacji*, PWN, Warszawa 1980.
- [5] Fletcher, R., *Practical Methods of Optimization*, John Wiley, Chichester 1987.
- [6] Galas, Z., Nykowski, I. (red.), *Zbiór zadań z programowania matematycznego*, część I, PWN, Warszawa 1986.
- [7] Gass, S. I., *Programowanie liniowe, metody i zastosowania*, PWN, Warszawa 1963.
- [8] Gill, P., Murray, W., Wright, M. H., *Practical Optimization*, Academic Press, London 1981.
- [9] Grabowski, W., *Programowanie matematyczne*, PWE, Warszawa, 1980.

- [10] Hiriart-Urruty, J.-B., Lemaréchal, C., *Convex Analysis and Minimization Algorithms, Vol I*, Springer-Verlag, Berlin 1993.
- [11] Hiriart-Urruty, J.-B., Lemaréchal, C., *Fundamentals of Convex Analysis*, Springer-Verlag, Berlin 2001.
- [12] Kantorowicz, L., Gorstko, A., *Optymalne decyzje ekonomiczne*, PWE, Warszawa 1976.
- [13] Kall, P., *Mathematische Methoden des Operations Research*, Teubner, Stuttgart 1976.
- [14] Karmarkar, N., A new polynomial-time algorithm for linear programming, *Combinatorica*, **4** (1984) 373-395.
- [15] Kielbasiński, A., Schwetlick, H., *Numeryczna algebra liniowa*, WNT, Warszawa 1992.
- [16] Koopmans, T. C. (red.), *Activity Analysis of Production and Allocation*, Cowles Commission Monograph 13, John Willey & Sons, Inc., New York 1951.
- [17] Musielak, J., *Wstęp do analizy funkcjonalnej*, PWN, Warszawa 1989.
- [18] Nemhauser, G. L., Rinnooy Kan, A. H. G., Todd, M. J. (red.), *Optimization, Handbook in Operations Research and Management Science, Vol. 1*, North-Holland, Amsterdam (1989).
- [19] Owen, G., *Teoria gier*, PWN, Warszawa 1975.
- [20] Sysło, M., Deo, N., Kowalik, J. S., *Algorytmy optymalizacji dyskretnej*, PWN, Warszawa 1993.
- [21] Vanderbei, R. J., *Linear Programming, Foundation and Extensions*, Kluwer, Boston 1997.
- [22] Walukiewicz, S., *Programowanie dyskretne*, PWN, Warszawa, 1986. (tłumaczenie angielskie: *Integer Programming*, PWN, Warszawa 1991.)