6 Dolne oszacowanie liczby mnożeń

6.1 Wyznacznik Vandermonde'a

Twierdzenie 6.1 *Dla $n > 1$ zachodzi następujący wzór:*

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ & & & \ddots & \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix} = \prod_{i<j}(x_j - x_i).$$

Dowód. Dla i równego kolejno $n, \dots, 2$ odejmujemy od i-tej kolumny rozważanej macierzy pomnożoną przez x_n kolumnę o numerze $i-1$. Po wykonaniu tej operacji w ostatnim wierszu wyznacznika pojawi się wiersz $(1, 0, \dots, 0)$. Zmienimy najwyżej znak wyznacznika, jeżeli usuniemy z otrzymanej w ten sposób macierzy pierwszą kolumnę i ostatni wiersz. Teraz z j-tego wiersza przekształcanej macierzy możemy wyłączyć $-(x_j - x_n)$. W ten sposób obliczenie wyznacznika Vandermonde'a sprowadzamy do obliczenia wyznacznika o analogicznej postaci i o wymiarze o 1 mniejszym. Dalej wystarczy powołać się na zasadę indukcji. □

6.2 Pierścienie wielomianów

Niech R będzie ciałem liczb rzeczywistych (albo dowolnym ciałem nieskończonym, w tym ciałem liczb wymiernych). Wielomianem nad ciałem R nazywamy funkcję f definiowaną wzorem postaci

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0,$$

gdzie $a_n, a_{n-1}, \dots, a_1, a_0 \in R$. Taka funkcja f przekształca R w R. Wielomiany możemy dodawać i mnożyć przyjmując zwykłe definicje tych działań. Łatwo sprawdzić, że zbiór wielomianów ze zwykłym dodawaniem i mnożeniem jest pierścieniem. Będziemy go oznaczać symbolem $R[x]$. Jest to pierścień z przemiennym mnożeniem.

Przedstawioną konstrukcję można iterować. Przyjmujemy, że

$$R[x_1, \dots, x_n, x_{n+1}] = R[x_1, \dots, x_n][x_{n+1}].$$

Nietrudno zauważyć, że $R[x_1, \dots, x_n]$ jest zbiorem wielomianów o współczynnikach należących do R ze zmiennymi $x_1, \dots, x_n$. Zbiór ten jest też pierścieniem z przemiennym mnożeniem.

Jednomianem w pierścieniu $R[x_1, \dots, x_n]$ nazywamy wyrażenie postaci

$$x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$$

(w tym wzorze ważna jest kolejność, w jakiej zostały wymienione zmienne). Wiadomo, że każdy wielomian ze zmiennymi $x_1, \dots, x_n$ jest sumą takich jednomianów pomnożonych przez współczynniki z ciała R ($x_1^0 x_2^0 \dots x_n^0$ jest szczególnym jednomianem, który we wzorach jest pomijany lub zastępowany przez 1). Ustalmy (w dowolny sposób) kolejność, w jakiej wymieniamy jednomiany. Kanonicznym przedstawieniem wielomianu nazywamy sumę jednomianów pomnożonych przez niezerowe współczynniki z R, w której jednomiany są podane w ustalonej kolejności. Jest oczywiste, że każdy wielomian ze zmiennymi $x_1, \dots, x_n$ ma kanoniczne przedstawienie. Najważniejszą własność wielomianów wyraża następujące twierdzenie:

Twierdzenie 6.2 *Jeżeli R jest (dowolnym) ciałem nieskończonym, to każdy wielomian z $R[x_1, \ldots, x_n]$ ma jednoznacznie wyznaczone kanoniczne przedstawienie.*

Dowód. Twierdzenie to dowodzimy przez indukcję ze względu na n. Dla wielomianów z jedną zmienną wynika to ze wzoru z rozdziału 6.1: bierzemy przedstawienie kononiczne wielomianu i dostatecznie duży zbiór elementów R. Dla wybranych elementów rozważamy równania, których niewiadomymi są współczynniki wielomianu, i które stwierdzają, że wartość przedstawienia kanonicznego jest równa wartości wielomianu. W ten sposób otrzymujemy układ równań liniowych, a jego macierzą główną jest wyznacznik Vandermonde'a. Niezerowość wyznacznika implikuje jednoznaczność rozwiązania rozważanego układu, a więc jednoznaczność współczynników kanonicznego przedstawienia wielomianu. Dla wielomianów z wieloma zmiennymi stosujemy analogiczne rozumowanie. □

6.3 Programy liniowe

Ustalmy nieskończone ciało R. Przypuśćmy, że mamy dane dwa zbiory zmiennych: zbiór zmiennych wejściowych i zbiór zmiennych (pozostałych, a więc jeżeli x jest zmienną, to nie jest zmienną wejściową). Programem liniowym nazywamy ciąg instrukcji podstawiania postaci $x := y + z$, $x := y - z$ oraz $x := y * z$ takich, że x jest zmienną, a y i z są zmiennymi wejściowymi, zmiennymi lub elementami R, spełniający następujące warunki:

1) dla dowolnej zmiennej x w tym ciągu występuje najwyżej jedna instrukcja postaci $x := \ldots$,

2) jeżeli w pewnym miejscu tego ciągu występuje instrukcja postaci $x := y + z$ i y jest zmienną; to w tym ciągu na miejscu wcześniejszym znajduje się instrukcja postaci $y := \ldots$; to samo dotyczy z, i analogiczna własność jest prawdziwa dla pozostałych typów instrukcji.

Przypuśćmy, że rozważamy program liniowy i $x_1, \ldots, x_n$ są wszystkimi zmiennymi wejściowymi występującymi w tym programie. Jeżeli $x := y + z$ jest instrukcją, to x, y i z będziemy nazywać elementami tej instrukcji. Analogicznie definiujemy elementy pozostałych typów instrukcji. Zdefiniujemy wartość wszystkich elementów wszystkich instrukcji występująch w danym programie. Wartości elementów będą należeć do pierścienia $R[x_1, \ldots, x_n]$. Jeżeli elementem jest zmienna wejściowa lub liczba należąca do R, to jego wartością jest on sam (czyli wielomian, który jest jednomianem zawierającym jedną zmienną, lub wielomian stałym). Jeżeli elementem jest zmienna x (występująca w programie), to po raz pierwszy w rozważanym programie występuje w instrukcji, która może mieć np. postać $x := y + z$. Wtedy wartością x jest suma wartości elementów y i z. Jeżeli pierwsza instrukcja podstawiająca pod x ma inną postać, to pojęcie wartości definiujemy analogicznie. Warunki wymagane od programu liniowego gwarantują, że podana definicja jest poprawna.

6.4 Przestrzeń $R^m[x_1, \ldots, x_n]$

Przestrzeń $R^m[x_1, \ldots, x_n]$ definiujemy jako zbiór m-elementowych ciągów o wyrazach należących do $R[x_1, \ldots, x_n]$. Nietrudno zauważyć, że zbiór ten z naturalną definicją dodawania elementów tego zbioru jest przestrzeń liniowa nad ciałem R.

W przestrzeni $R^m[x_1, \ldots, x_n]$ będziemy rozważać specyficzne pojęcie liniowej niezależności wektorów. Będziemy mówić, że wektory $\vec{v}_1, \ldots, \vec{v}_k$ są liniowo niezależne modulo R^m, jeżeli dla dowolnych elementów $a_1, \ldots, a_k \in R$ warunek $\sum_{i=1}^{k} a_i \vec{v}_i \in R^m$ implikuje, że $a_1 = \ldots = a_k = 0$.

6.5 Główny rezultat

Twierdzenie 6.3 *Niech $x_1, \ldots, x_n$ i $y_1, \ldots, y_m$ będą rozłącznymi zbiorami zmiennych wejściowych. Przypuśćmy, że mamy macierz M o r liniowo niezależnych modulo R^m wierszach należących do $R^m[x_1, \ldots, x_n]$. Wtedy dowolny program liniowy obliczający iloczyn $M\vec{y}$ macierzy M i kolumny $\vec{y}$ o wyrazach $y_1, \ldots, y_m$ zawiera przynajmniej r instrukcji polecających wykonanie mnożenia.*

Dowód. Przypuśćmy, że mamy liniowy program L obliczającym $M\vec{y}$, i że znajduje się w nim $s < r$ instrukcji mnożenia. Załóżmy, że lewymi stronami tych instrukcji są zmienne $z_1, \ldots, z_s$. Usuwamy z programu L instrukcje mnożenia, a zmienne $z_1, \ldots, z_s$ uważamy za nowe zmienne wejściowe. Otrzymujemy pewien program liniowy L' bez mnożeń.

Wartości zmiennych z programu L' są wielomianami z pierścienia

$$R[x_1, \ldots, x_n, y_1, \ldots, y_m, z_1, \ldots, z_s].$$

Ponieważ w L' nie występują instrukcje mnożenia, więc są to wielomiany stopnia pierwszego o współczynnikach naturalnych (kombinacje liniowe o współczynnikach naturalnych zmiennych).

Niech $z'_1, \ldots, z'_s$ będą wartościami zmiennych $z_1, \ldots, z_s$ w programie L. Jest oczywiste, że $z'_i \in R[x_1, \ldots, x_n, y_1, \ldots, y_m]$.

Pojęcie wartości zmiennej (elementu) w programie liniowym można rozszerzyć tak, aby zmiennym wejściowym można było nadawać dowolne wartości początkowe. Zgodnie z nową definicją wartością zmiennej wejściowej x nie musi być x, a może nią być ustalony element R lub pewien wielomian. Aby otrzymać nową definicję wartości zmiennej, wystarczy w podanej zmienić fragment określający wartości zmiennych wejściowych żądając, że są to ustalone wartości. Oprócz zwykłych wartości zmiennych w programie L', będziemy także rozważać wartości, które zostały zdefiniowane przy założeniu, że wartości początkowe zmiennych z_i są równe z'_i, a wartościami pozostałych zmiennych wejściowych programu L' są one same. Zgodnie z tą ostatnią definicją wartości zmiennych z programu L' są wielomianami z $R[x_1, \ldots, x_n, y_1, \ldots, y_m]$. Nietrudno też zauważyć, że te wartości są dokładnie takie, jak wartości zmiennych obliczone dla programu L. Wobec tego możemy uważać, że program L' także oblicza iloczyn $M\vec{y}$ (poszczególne elementy tego iloczynu są rozumianymi w szczególny sposób wartościami odpowiednich zmiennych programu L').

Teraz porównamy oba sposoby liczenia wartości zmiennych programu L'. Można wykazać, że wartość dowolnej zmiennej obliczona przy założeniu, że zmienne z_i mają wartości początkowe równe z'_i, jest równa zwykłej wartości tej zmiennej, w której zamiast z_i wstawiliśmy wielomian z'_i. Stąd wynika, że wartości zmiennych w programie L są liniowymi kombinacjami o współczynnikach naturalnych wielomianów-zmiennych $x_1, \ldots, x_n$ oraz $y_1, \ldots, y_m$, i wielomianów $z'_1, \ldots, z'_s$. Jeżeli wektor wielomianów

$z'_1, \ldots, z'_s$ oznaczymy symbolem $\vec{z'}$, to w wartość każdej zmiennej z programu L można zapisać w postaci $\vec{n}\vec{z'} + w$ dla pewnego ciągu liczb naturalnych $\vec{n}$ i pewnego wielomianu $w \in R[x_1, \ldots, x_n, y_1, \ldots, y_m]$ stopnia pierwszego.

W podanej postaci można zapisać wszystkie zmienne pamiętające elementy iloczynu $M\vec{y}$. Mamy więc

$$M\vec{y} = N\vec{z'} + \vec{W}$$

dla pewnej macierzy N o wyrazach naturalnych i pewnego wektora $\vec{W}$ o współrzędnych będących wielomianami stopnia 1 ze zmiennymi $x_1, \ldots, x_n$ i $y_1, \ldots, y_m$ Nietrudno ustalić, że N jest macierzą o r wierszach i s kolumnach.

Ponieważ N ma więcej wierszy niż kolumn $(s < r)$, więc wiersze N są liniowo zależnymi wektorami w przestrzeni R^s. Istnieje więc niezerowy wektor $\vec{u} \in R^r$ taki, że $\vec{u}N = \vec{0}$. Stąd otrzymujemy, że

$$\vec{u}M\vec{y} = \vec{u}N\vec{z'} + \vec{u}\vec{W} = \vec{u}\vec{W}.$$

Ta równość jest możliwa tylko wtedy, gdy $\vec{u}M \in R^m$. W przecinym razie pewna współrzędna $\vec{u}M$ jest wielomianem dodatniego stopnia i - wobec tego - $\vec{u}M\vec{y}$ jest wielomianej stopnia przynajmniej równego 2 (nic się nie uprości, ponieważ mnożymy przez wektor złożony z różnych zmiennych). Z drugiej strony $\vec{u}M\vec{y}$ jest równy $\vec{u}\vec{X}$, a więc jest wielomianem pierwszego stopnia. Z twierdzenia 6.2 wynika, że nie jest to możliwe. W ten sposób dowiedliśmy, że $\vec{u}M \in R^m$.

Fakt $\vec{u}M \in R^m$ też implikuje sprzeczność. Liniowa niezależność modulo R^m wierszy macierzy M implikuje bowiem, że $\vec{u} = 0$, a to przeczy temu, że wektor $\vec{u}$ został wybrany jako niezerowy. Sprzeczność ta świadczy o tym, że $r \leqslant s$. $\square$

Wniosek 6.4 *Obliczenie iloczynu dwóch macierzy o wymiarach $n \times n$ wymaga wykonania przynajmniej n^2 mnożeń.*