2 Dolne oszacowanie liczby mnożeń

2.1 Wyznacznik Vandermonde'a

Twierdzenie 2.1 *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. □

2.2 Pierścienie wielomianów

Niech R będzie dowolnym ciałem nieskończonym, np. ciałem liczb wymiernych. Wielomianem nad ciałem R nazywamy funkcję f definiowaną wzorem postaci

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

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 ze zmiennymi $x_1, \dots, x_n$, o współczynnikach należących do R. Zbiór ten też jest 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 2.2 *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 Vandermonde'a z rozdziału 2.1: 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. □

2.3 Programy liniowe

2.3.1 Pojęcie programu liniowego

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 stałymi należącymi do 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.

2.3.2 Wartość zmiennej w programie liniowym

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 bedziemy 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 stała należąca do R, to jego wartością jest on sam (czyli wielomian, który jest jednomianem zawierającym jedną zmienną, lub wielomian stały). Jeżeli elementem jest zmienna x (występująca w programie), to po raz pierwszy w rozważanym programie występuje w instrukcji, która ma 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.

2.3.3 Programy obliczające funkcje

Przypuśćmy, że mamy program liniowy P ze zmiennymi wejściowymi $x_1, \ldots, x_n$ i funkcje (wielomiany) $f_1, \ldots, f_m \in R[x_1, \ldots, x_n]$. O programie P mówimy, że oblicza funkcje $f_1, \ldots, f_m$, jeżeli w tym programie są zmienne $z_1, \ldots, z_m$, których wartościami są odpowiednio funkcje $f_1, \ldots, f_m$.

2.3.4 Złożoność zadania polegającego na obliczeniu funkcji

Przypuśćmy, że mamy dane funkcje $f_1, \ldots, f_m \in R[x_1, \ldots, x_n]$. Złożonością zadania polegającego na obliczeniu funkcji $f_1, \ldots, f_m$ nazywamy liczbę $L^{tot}(f_1, \ldots, f_m)$ będącą długością najkrótszego programu obliczającego $f_1, \ldots, f_m$.

W razie potrzeby, zdefiniowaną złożoność będziemy nazywać całkowitą. Oprócz złożoności całkowitej będziemy rozważać złożoność multiplikatywną.

Instrukcję mnożenia $x := y * z$ będziemy nazywać istotną, jeżeli zarówno y jak i z są zmiennymi (roboczymi lub wejściowymi), czyli nie są stałymi. Złożonością multiplikatywną zadania polegającego na obliczeniu funkcji $f_1, \ldots, f_m$ nazywamy najmniejszą spośród liczb, które są liczbami istotnych instrukcji mnożenia (instrukcji mnożenia zmiennych) w programach obliczających funkcje $f_1, \ldots, f_m$. Złożoność multiplikatywną zadania obliczenia tych funkcji będziemy oznaczać symbolem $L(f_1, \ldots, f_m)$.

Złożonością (całkowitą lub multiplikatywną) problemu mnożenia macierzy o wymiarach $n \times n$ będziemy nazywać złożoność zadania polegającego na obliczeniu wszystkich funkcji $\sum_{k=1}^{n} x_{i,k} y_{k,j}$ dla $i, j = 1, \ldots n$. Symbole $x_{i,k}$ i $y_{k,j}$ są zmiennymi wejściowymi reprezentującymi odpowiednie wyrazy danych macierzy.

2.4 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 $a_1 = \ldots = a_k = 0$.

2.5 Dolne oszacowanie złożoności multiplikatywnej

Twierdzenie 2.3 *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 problem obliczenia iloczynu $M\vec{y}$ macierzy M i kolumny $\vec{y}$ o wyrazach $y_1, \ldots, y_m$ ma złożoność multiplikatywną $\geq r$ (dowolny program liniowy obliczający iloczyn $M\vec{y}$ macierzy M i kolumny $\vec{y}$ o wyrazach $y_1, \ldots, y_m$ zawiera przynajmniej r istotnych instrukcji mnożenia).*

Dowód. Przypuśćmy, że mamy liniowy program P obliczającym $M\vec{y}$, i że znajduje się w nim $s < r$ istotnych instrukcji mnożenia. Załóżmy, że lewymi stronami tych instrukcji są zmienne $z_1, \ldots, z_s$. Usuwamy z programu P instrukcje mnożenia, a zmienne $z_1, \ldots, z_s$ uważamy za nowe zmienne wejściowe. Otrzymujemy pewien program liniowy P' bez mnożeń. Symbolami w_x^P i $w_x^{P'}$ będziemy oznaczać wartość zmiennej x odpowiednio w programie P i P'.

Wartości $w_x^{P'}$ zmiennych z programu P' są wielomianami z pierścienia

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

Ponieważ w P' nie występują instrukcje mnożenia, więc są to wielomiany stopnia pierwszego (kombinacje liniowe zmiennych wejściowych).

Jest też oczywiste, że $w_{z_i}^P \in R[x_1, \ldots, x_n, y_1, \ldots, y_m]$.

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. Symbolem $w_x^{P'}[a_1, \ldots, a_s]$ oznaczamy obliczoną w ten sposób wartość zmiennej x przy założeniu, że wartości początkowe zmiennych z_i są równe a_i dla $i = 1, \ldots, s$, a wartościami pozostałych zmiennych wejściowych programu P' są one same. Zgodnie z tą definicją wartości $w_x^{P'}[w_{z_1}^P, \ldots, w_{z_s}^P]$ są wielomianami z $R[x_1, \ldots, x_n, y_1, \ldots, y_m]$. Nietrudno też zauważyć, że $w_x^{P'}[w_{z_1}^P, \ldots, w_{z_s}^P] = w_x^P$ dla wszystkich zmiennych x. Wobec tego możemy uważać, że program P' także oblicza iloczyn $M\vec{y}$ (poszczególne elementy tego iloczynu są rozumianymi w szczególny sposób wartościami odpowiednich zmiennych programu P').

Teraz porównamy oba sposoby liczenia wartości zmiennych programu P'. Można wykazać, że dla dowolnej zmiennej x mamy

$$w_x^{P'}[w_{z_1}^P, \ldots, w_{z_s}^P](x_1, \ldots, x_n, y_1, \ldots, y_m) = w_x^{P'}(x_1, \ldots, x_n, y_1, \ldots, y_m, w_{z_1}^P, \ldots, w_{z_s}^P).$$

Stąd wynika, że wartości zmiennych w programie P są liniowymi kombinacjami wielomianów-zmiennych $x_1, \ldots, x_n$ oraz $y_1, \ldots, y_m$, i wielomianów $w_{z_1}^P, \ldots, w_{z_s}^P$. Jeżeli wektor wielomianów $w_{z_1}^P, \ldots, w_{z_s}^P$ oznaczymy symbolem $\vec{Z}$, to w wartość każdej zmiennej z programu P można zapisać w postaci $\vec{c}\vec{Z} + v$ dla pewnego wektora $\vec{c}$ i pewnego wielomianu $v \in R[x_1, \ldots, x_n, y_1, \ldots, y_m]$ stopnia pierwszego.

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

$$M\vec{y} = C\vec{Z} + \vec{V}$$

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

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

$$\vec{u}M\vec{y} = \vec{u}C\vec{Z} + \vec{u}\vec{V} = \vec{u}\vec{V}.$$

Ta równość jest możliwa tylko wtedy, gdy $\vec{u}M \in R^m$. W przeciwnym razie pewna współrzędna $\vec{u}M$ jest wielomianem dodatniego stopnia i - wobec tego - $\vec{u}M\vec{y}$ jest wielomianem 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 $\vec{u}\vec{V}$, a więc jest wielomianem pierwszego stopnia. Z twierdzenia 2.2 wynika, że nie jest to możliwe. W ten sposób dowiedliśmy, że $\vec{u}M \in R^m$.

Fakt $\vec{u}M \in R^m$ 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 $r \leq s$. □

Wniosek 2.4 *Obliczenie iloczynu dwóch macierzy o wymiarach $n \times n$ wymaga wykonania przynajmniej n^2 istotnych mnożeń.* □