1 Algorytm Strassena mnożenia macierzy

1.1 Pierścienie

Pierścieniem nazywamy zbiór z dwoma działaniami: dodawaniem $+$ i mnożeniem $\cdot$ spełniającymi niektóre z praw przysługujących dodawaniu i mnożeniu np. liczb rzeczywistych. Tak więc oba działania w pierścieniu są łączne i mnożenie jest rozdzielne względem dodawania. Dodawanie jest przemienne. W pierścieniu są elementy neutralne dla obu działań, a więc zero i jedność spełniające dla wszystkich elementów x pierścienia równości $x+0=x$ oraz $x \cdot 1 = 1 \cdot x = x$. Zakładamy, że $0 \neq 1$. Dla dowolnego dowolnego elementu x z pierścienia istnieje element przeciwny, a więc taki element y, że $x+y=0$. Element ten jest oznaczany $-x$.

Przykładem pierścienia jest zbiór liczb całkowitych ze zwykłymi działaniami. W tym przypadku mnożenie jest także działaniem przemiennym. Taki pierścień nazywamy przemiennym. Pierścieniami są także zbiory liczb wymiernych i rzeczywistych ze zwykłymi działaniami. W tych pierścieniach każdy element różny od zera ma element odwrotny. W definicji pierścienia nie żądamy jednak ani przemienności mnożenia, ani istnienia elementów odwrotnych.

1.2 Mnożenie macierzy

Ustalmy pierścień Z. Będziemy rozważać macierze o wyrazach należących do Z. Przypuśćmy, że mamy dwie takie macierze

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix} \text{ oraz } B = \begin{bmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,p} \\ b_{2,1} & b_{2,2} & \dots & b_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \dots & b_{n,p} \end{bmatrix}.$$

Zauważmy, że liczba kolumn macierzy A jest równa liczbie wierszy B. Mając takie macierze możemy je pomnożyć. Ich iloczynem $A \cdot B$ jest macierz

$$C = \begin{bmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,p} \\ c_{2,1} & c_{2,2} & \dots & c_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m,1} & c_{m,2} & \dots & c_{m,p} \end{bmatrix}$$

o wyrazach $c_{i,j}$ danych wzorem

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} \cdot b_{k,j}$$

($c_{i,j}$ jest więc iloczynem skalarnym i-tego wiersza macierzy A i j-tej kolumny macierzy B, w podanym wzorze istotną rolę odgrywa także kolejność mnożonych przez siebie elementów). Zauważmy, że iloczyn $A \cdot B$ ma tyle wierszy, co macierz A i tyle kolumn co B.

Zauważmy, że jeżeli macierz B jest kwadratowa (tzn. $p=n$), wyrazy znajdujące się na przekątnej tej macierzy są równe 1 (tzn. $b_{i,i}=1$ dla $i=1,\dots,n$), a pozostałe wyrazy tej macierzy są równe 0, to iloczyn $A \cdot B$ jest równy A. Macierz B o podanych własnościach nazywamy jednostkową i oznaczamy symbolem 1. Podobnie, jeżeli A jest macierzą jednostkową o m wierszach i $n=m$ kolumnach, to $A \cdot B = B$.

Przypuśćmy, że mamy dane trzy macierze: A i B (takie jak wyżej) oraz

$$C = \begin{bmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,r} \\ c_{2,1} & c_{2,2} & \dots & c_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p,1} & c_{p,2} & \dots & c_{p,r} \end{bmatrix}.$$

Dla takich trzech macierzy możemy obliczyć iloczyny $(A \cdot B) \cdot C$ i $A \cdot (B \cdot C)$. W obu wypadkach otrzymujemy macierze o m wierszach i r kolumnach. Wyraz na przecięciu i-tego wiersza i j-tej kolumny macierzy $(A \cdot B) \cdot C$ jest dany wzorem

$$\sum_{l=1}^{p} (\sum_{k=1}^{n} a_{i,k} \cdot b_{k,l}) \cdot c_{l,j}.$$

Ten sam wyraz macierzy $A \cdot (B \cdot C)$ jest dany wzorem

$$\sum_{k=1}^{n} a_{i,k} \cdot (\sum_{l=1}^{p} b_{k,l} \cdot c_{l,j}).$$

Odpowiednie wyrazy macierzy $(A \cdot B) \cdot C$ i $A \cdot (B \cdot C)$ są więc równe. Oznacza to, że mnożenie macierzy jest łączne.

1.3 Dodawanie macierzy i pierścień $M_n(Z)$

Przypuśćmy, że mamy dwie macierze

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix} \text{ oraz } B = \begin{bmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,n} \\ b_{2,1} & b_{2,2} & \dots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \dots & b_{m,n} \end{bmatrix}$$

o tych samych wymiarach. Takie macierze możemy dodać. Sumą $A + B$ macierzy A i B jest macierz C dana wzorem

$$C = \begin{bmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \dots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \dots & a_{2,n} + b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & \dots & a_{m,n} + b_{m,n} \end{bmatrix}.$$

Nietrudno zauważyć, że dodawanie macierzy jest łączne i przemienne. Jeżeli do macierzy A dodamy macierz o takich samych wymiarach wypełnioną zerami (oznaczamy ją symbolem 0), to w wyniku otrzymamy macierz A. Łatwo wskazać macierz przeciwną do danej. Można też dowieść prawo rozdzielności mnożenia macierzy względem dodawania.

Dla dowolnego pierścienia Z i liczby naturalnej n definiujemy zbiór $M_n(Z)$. Jest to zbiór macierzy kwadratowych o wymiarach $n \times n$ i wyrazach z pierścienia Z. Zbiór ten jest zamknięty ze względu na mnożenie i dodawanie macierzy. Udowodnione lub omówione własności tych działań świadczą o tym, że $M_n(Z)$ z dodawaniem i mnożeniem macierzy jest pierścieniem.

1.4 Pewna własność iloczynu macierzy

Przypuśćmy, że mamy dany wektor $\vec{a} = (a_1, a_2, \ldots, a_n) \in R^n$ i rosnący ciąg liczb naturalnych $1 = n_1 < \ldots < n_k < n_{k+1} = n+1$ (n_{k+1} jest pomocniczym oznaczeniem liczby $n+1$, a nie wyrazem rozważanego ciągu). Przyjmijmy, że $\vec{a}_i = (a_{n_i}, a_{n_i+1}, \ldots, a_{n_{i+1}-1})$ dla $i = 1, \ldots, k$. Nietrudno zauważyć, że wektor (ciąg) $\vec{a}$ jest konkatenacją wektorów (ciągów) $\vec{a}_1, \ldots, \vec{a}_k$. Jest oczywiste, że dla danego ciągu liczb $n_1, \ldots, n_k$ o podanych własnościach wektor $\vec{a}$ możemy utożsamiać z wektorem $(\vec{a}_1, \ldots, \vec{a}_k)$ o wyrazach wektorowych. Wektory o wyrazach wektorowych też możemy mnożyć skalarnie: odpowiadające sobie współrzędne (a więc wektory) mnożymy skalarnie i dodajemy otrzymane liczby.

Jeżeli mamy dane liczby $n_1, \ldots, n_k$ i dwa wektory $\vec{a}, \vec{b} \in R^n$, to iloczyn skalarny $\vec{a} \cdot \vec{b}$ możemy obliczać na dwa sposoby: po pierwsze, zgodnie z definicją

$$\vec{a} \cdot \vec{b} = \sum_{j=1}^{n} a_j \cdot b_j,$$

a po drugie, możemy dane wektory przekształcić w wektory o wyrazach wektorowych i dopiero wtedy pomnożyć je skalarnie

$$(\vec{a}_1, \ldots, \vec{a}_k) \cdot (\vec{b}_1, \ldots, \vec{b}_k) = \sum_{i=1}^{k} \vec{a}_i \cdot \vec{b}_i = \sum_{i=1}^{k} \sum_{j=n_i}^{n_{i+1}-1} a_j \cdot b_j.$$

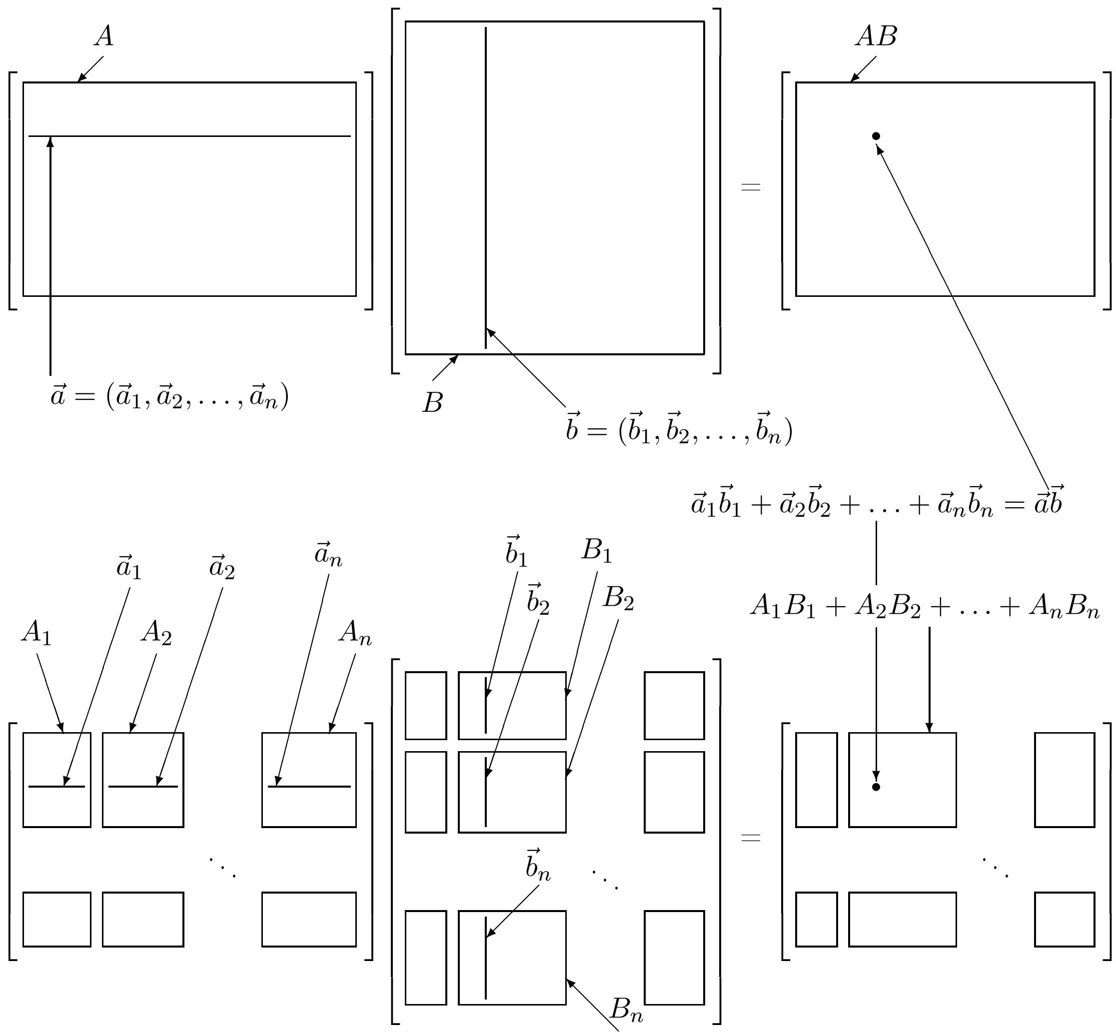
Nietrudno zauważyć, że w obu przypadkach sumujemy te same zbiory składników, a więc otrzymujemy te same liczby. W ten sposób uzasadniliśmy wzór

$$\vec{a} \cdot \vec{b} = (\vec{a}_1, \ldots, \vec{a}_k) \cdot (\vec{b}_1, \ldots, \vec{b}_k).$$

Przypuśćmy, że mamy macierze A i B, które możemy pomnożyć. Przyjmijmy, że A ma m wierszy i n kolumn. Macierz B ma więc n wierszy. Przyjmijmy, że B ma p kolumn. Rozważmy trzy rosnące ciągi liczb $1 = n_1 < \ldots < n_t \le n$, $1 = m_1 < \ldots < m_u \le m$ oraz $1 = p_1 < \ldots < p_v \le p$. Ciąg $n_1, \ldots, n_t$ wyznacza sposób podziału wierszy (sposób grupowania kolumn) macierzy A. Podobnie ciąg $m_1, \ldots, m_u$ wyznacza sposób podziału kolumn macierzy A. Mając te dwa ciągi możemy podzielić macierz A na $t \cdot u$ podmacierzy. Macierz A możemy utożsamiać z otrzymaną w ten sposób macierzą o wyrazach macierzowych (na rys. 1 macierz ta jest przedstawiona pod macierzą A). Podobnie, ciągi $p_1, \ldots, p_v$ i $n_1, \ldots, n_t$ pozwalają podzielić macierz B i B możemy utożsamiać z otrzymaną w ten sposób macierzą o wyrazach macierzowych. Zauważmy, że wiersze macierzy A i kolumny macierzy B dzielimy w ten sam sposób. W takim przypadku otrzymane macierze o wyrazach macierzowych możemy pomnożyć przez siebie, a więc pierwszy z czynników ma wiersze długości takiej, jak kolumny drugiego, a także potrzebne iloczyny macierzy otrzymanych z podziału A i B są określone. Zauważmy także, że AB ma m wierszy i p kolumn, a więc możemy podzielić ją na mniejsze macierze w sposób wyznaczony przez ciągi $m_1, \ldots, m_u$ i $p_1, \ldots, p_v$.

Mnożenie macierzy ma następującą własność: jeżeli weźmiemy dwie macierze A i B, pomnożymy je, a następnie iloczyn AB podzielimy na podmacierze tak, jak to zostało opisane, to otrzymamy dokładnie iloczyn macierzy otrzymanych przez opisany podział macierzy A i B. Uzasadnienie tego faktu jest przedstawione na rysunku 1. Własność ta jest konsekwencją podanej własności iloczynu skalarnego.

Niech $A \in M_{n \cdot m}(Z)$. Podzielmy w naturalny sposób macierz A na n^2 macierzy o wymiarach $m \times m$. Otrzymaną w ten sposób macierz należącą do pierścienia $M_n(M_m(Z))$ oznaczmy symbolem $i_{n,m}(A)$. W ten sposób zdefiniowaliśmy funkcję $i_{n,m} : M_{n \cdot m}(Z) \to M_n(M_m(Z))$.



Rysunek 1: Dwa sposoby obliczania iloczynu macierzy

Twierdzenie 1.1 *Funkcja $i_{n,m}$ jest izomorfizmem pierściemi $M_{n \cdot m}(Z)$ oraz $M_n(M_m(Z))$, a więc jest wzajemnie jednoznaczna i typu "na", oraz zachowuje dodawanie i mnożenie, czyli dla dowolnych $A, B \in M_{n \cdot m}(Z)$ zachodzą następujące równości*

$$i_{n,m}(A + B) = i_{n,m}(A) + i_{n,m}(B),$$
$$i_{n,m}(A \cdot B) = i_{n,m}(A) \cdot i_{n,m}(B).$$

1.5 Mnożenie macierzy o wymiarach 2×2 według Strassena

Przypuśćmy, że iloczynem macierzy

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \text{ oraz } B = \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix}$$

jest macierz

$$C = \begin{bmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{bmatrix}.$$

Wyrazy macierzy C możemy znaleźć wykonując operacje wymienione w następującej tabelce:

$$\begin{aligned}
m_1 &= (a_{1,2} - a_{2,2})(b_{2,1} + b_{2,2}), \\
m_2 &= (a_{1,1} + a_{2,2})(b_{1,1} + b_{2,2}), \\
m_3 &= (a_{1,1} - a_{2,1})(b_{1,1} + b_{1,2}), \\
m_4 &= (a_{1,1} + a_{1,2})b_{2,2}, \\
m_5 &= a_{1,1}(b_{1,2} - b_{2,2}), \\
m_6 &= a_{2,2}(b_{2,1} - b_{1,1}), \\
m_7 &= (a_{2,1} + a_{2,2})b_{1,1}, \\
c_{1,1} &= m_1 + m_2 - m_4 + m_6, \\
c_{1,2} &= m_4 + m_5, \\
c_{2,1} &= m_6 + m_7, \\
c_{2,2} &= m_2 - m_3 + m_5 - m_7.
\end{aligned}$$

Nietrudno zauważyć, że podane wzory pozwalają mnożyć macierze o wymiarach 2×2, o wyrazach należących do dowolnego, także nieprzemiennego pierścienia. Przyjmijmy, że $\mathcal{S}_2[D_Z, M_Z](A, B)$ oznacza algorytm obliczający iloczyn macierzy A i B (o wymiarach 2×2) z elementami z pierścienia Z, wykonujący podane obliczenia w ten sposób, że operacja dodawania $+$ polega na wykonaniu procedury D_Z, a operacja mnożenia – procedury M_Z.

1.6 Algorytm Strassena

Przypuśćmy, że mamy dane macierze A i B o wymiarach $2^n \times 2^n$ i o wyrazach z pierścienia Z, funkcje (programy) D_Z i M_Z obliczające sumę i iloczyn elementów Z. Niech $D_{M_m(Z)}$ będzie naturalnym programem obliczającym sumę dwóch macierzy o wymiarach $m \times m$ wykorzystującym program D_Z. (Pierścień $M_1(Z)$ utożsamiamy z Z, $D_{M_1(Z)}$ i D_Z to ten sam program). Zdefiniujemy algorytm Strassena $\mathcal{S}[D_Z, M_Z](A, B)$ obliczający iloczyn $A \cdot B$ (funkcję zwracającą ten iloczyn), wykorzystujący do dodawania i mnożenia elementów Z odpowiednio algorytmy D_Z i M_Z.

Algorytm $\mathcal{S}[D_Z, M_Z](A, B)$:

1) $m :=$ wymiar A;

2) jeżeli $m = 1$, to mnożymy wyrazy A i B stosując procedurę M_Z, zwracamy macierz, której jedynym wyrazem jest otrzymany iloczyn i przerywamy wykonywanie algorytmu;

3) $\mathcal{S} := i^{-1}_{2,m/2}(\mathcal{S}_2[D_{M_{m/2}(Z)}, \mathcal{S}[D_Z, M_Z]](i_{2,m/2}(A), i_{2,m/2}(B)))$.

Poprawność algorytmu $\mathcal{S}$ wynika z twierdzenia 1.1. Obliczenie iloczynu dowolnych macierzy można zrealizować korzystając z wzoru

$$\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} A \cdot B & 0 \\ 0 & 0 \end{bmatrix}$$

w następujący sposób:

1) powiększ A i B dopisując pewną liczbę wierszy i kolumn wypełnionych zerami tak, aby otrzymać możliwie małe macierze o wymiarach będących potęgą dwójki; przypuśćmy, że w ten sposób otrzymaliśmy odpowiednio macierze A' i B';

2) $C' := \mathcal{S}[D_Z, M_Z](A', B')$;

3) zwróć macierz otrzymaną z C' przez usunięcie tylu (zerowych) wierszy i kolumn, ile zostało dodane w kroku 1).

Niech $S'_D(n)$ i $S'_M(n)$ oznacza odpowiednio liczbę dodawań i mnożeń wykonanych przez algorytm $\mathcal{S}$ uruchomiony z danymi rozmiaru 2^n. Nietrudno zauważyć, że

$$S'_M(n+1) = 7 \cdot S'_M(n) \text{ oraz } S'_M(0) = 1.$$

Stąd $S'_M(n) = 7^n$. Funkcja $S'_D(n)$ spełnia równości

$$S'_D(n+1) = 7 \cdot S'_D(n) + 18 \cdot (2^n)^2 = 7 \cdot S'_D(n) + 18 \cdot 4^n \text{ oraz } S'_D(0) = 0.$$

Łatwo sprawdzić przez indukcję, że jest to funkcją dana wzorem

$$S'_D(n) = 6 \cdot (7^n - 4^n).$$

Niech $S_D(n)$ i $S_M(n)$ oznacza odpowiednio liczby dodawań i mnożeń wykonanych przez podany algorytm uruchomiony z macierzami rozmiaru $n \times n$. Wtedy $S_D(n) = S'_D(m)$ oraz $S_M(n) = S'_M(m)$ dla pewnej liczby spełniającej $2^m < 2 \cdot n$. Wobec tego

$$S_M(n) = S'_M(m) = 7^m = 2^{m \log_2 7} = (2^m)^{\log_2 7} < 7 \cdot n^{\log_2 7}$$

oraz

$$S_D(n) = S'_D(m) = 6 \cdot (7^m - 4^m) < 6 \cdot 7^m < 42 \cdot n^{\log_2 7}.$$