

Wstęp do metod numerycznych

Eliminacja Gaussa

P. F. Góra

<http://th-www.if.uj.edu.pl/zfs/gora/>

2017

Co można zrobić z układem równań

... tak, aby jego rozwiązania się nie zmieniły?

Rozważam układ równań (przykład 3×3 dla oszczędności miejsca):

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \quad (1)$$

1. Równania można zapisać w innej kolejności:

$$\begin{cases} a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \quad (2)$$

Odpowiada to **permutacji wierszy macierzy układu równań, z jednoczesną permutacją kolumny wyrazów wolnych.**

2. Równania można dodać stronami, po pomnożeniu przez dowolną stałą różną od zera:

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ (z \cdot a_{11} + a_{31})x_1 + (z \cdot a_{12} + a_{32})x_2 + (z \cdot a_{13} + a_{33})x_3 = z \cdot b_1 + b_3 \end{array} \right. \quad (3)$$

Odpowiada to **zastąpieniu jednego wiersza macierzy układu równań przez dowolną kombinację liniową tego wiersza z innymi, z jednoczesną analogiczną operacją na kolumnie wyrazów wolnych.**

3. We wszystkich równaniach można przestawić kolejność, w jakiej pojawiają się zmienne:

$$\begin{cases} a_{11}x_1 + a_{13}x_3 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{23}x_3 + a_{22}x_2 = b_2 \\ a_{31}x_1 + a_{33}x_3 + a_{32}x_2 = b_3 \end{cases} \quad (4)$$

Odpowiada to **permutacji *kolumn* macierzy układu równań, z jednoczesną permutacją kolumny niewiadomych.**

Eliminacja Gaussa

Rozpatrzmy jeszcze raz układ równań

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \quad (5)$$

Podzielmy pierwsze równanie stronami przez a_{11}

$$\begin{cases} x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3 = \frac{b_1}{a_{11}} \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \quad (6)$$

Teraz mnożymy pierwsze z równań (6) przez a_{21} i odejmijmy stronami od

drugiego, a następnie mnożymy pierwsze z równań (6) przez a_{31} i odejmijmy stronami od **trzeciego**. Otrzymujemy

$$\left\{ \begin{array}{l} x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3 = \frac{b_1}{a_{11}} \\ \left(a_{22} - \frac{a_{21}a_{12}}{a_{11}}\right)x_2 + \left(a_{23} - \frac{a_{21}a_{13}}{a_{11}}\right)x_3 = b_2 - \frac{a_{21}}{a_{11}}b_1 \\ \left(a_{32} - \frac{a_{31}a_{12}}{a_{11}}\right)x_2 + \left(a_{33} - \frac{a_{31}a_{13}}{a_{11}}\right)x_3 = b_3 - \frac{a_{31}}{a_{11}}b_1 \end{array} \right. \quad (7a)$$

Przepiszmy to w postaci (tylko zmiana oznaczeń!)

$$\left\{ \begin{array}{l} x_1 + a'_{12}x_2 + a'_{13}x_3 = b'_1 \\ a'_{22}x_2 + a'_{23}x_3 = b'_2 \\ a'_{32}x_2 + a'_{33}x_3 = b'_3 \end{array} \right. \quad (7b)$$

W układzie równań (7b) pierwsza zmienna, x_1 , występuje wyłącznie w pierwszym równaniu. Tego równania już nie przekształcamy, natomiast z pozo-

stałymi równaniami postępujemy analogicznie: dzielimy drugie stronami przez a'_{22} i odpowiednio mnożąc, odejmujemy od trzeciego. Otrzymujemy

$$\begin{cases} x_1 + a'_{12}x_2 + a'_{13}x_3 = b'_1 \\ x_2 + a''_{23}x_3 = b''_2 \\ a''_{33}x_3 = b''_3 \end{cases} \quad (8)$$

Teraz pierwsza zmienna występuje wyłącznie w pierwszym równaniu, druga — w pierwszym i w drugim. Gdyby równań było więcej, moglibyśmy to postępowanie kontynuować.

Ostatecznie otrzymalibyśmy równanie postaci

$$\left\{ \begin{array}{l} x_1 + \bullet x_2 + \bullet x_3 + \dots + \bullet x_N = \tilde{b}_1 \\ \quad x_2 + \bullet x_3 + \dots + \bullet x_N = \tilde{b}_2 \\ \quad \quad x_3 + \dots + \bullet x_N = \tilde{b}_3 \\ \quad \quad \quad \dots \quad \quad \dots \\ \quad \quad \quad \quad \quad x_N = \tilde{b}_N \end{array} \right. \quad (9)$$

gdzie symbole \bullet oznaczają *jakieś* współczynniki, dające się wyliczyć z pierwotnych współczynników równania, \tilde{b}_i są przekształconymi w toku całej procedury wyrazami wolnymi.

Równanie w postaci (9) nazywamy układem równań *z macierzą trójkątną górną*. Algorytm prowadzący od (5) do (9) nazywamy *eliminacją Gaussa*.

Dygresja: Złożoność obliczeniowa

Niech N oznacza liczbę danych wejściowych pewnego algorytmu. Niech $\mathcal{M}(N)$ oznacza liczbę operacji, jaką algorytm ten wykonuje dla N danych. Mówimy, że **algorytm ma złożoność obliczeniową $O(\mathcal{P}(N))$** jeżeli

$$\exists N_0 \in \mathbb{N}, A_1, A_2 > 0 \forall N > N_0: A_1 \cdot \mathcal{P}(N) \leq \mathcal{M}(N) \leq A_2 \cdot \mathcal{P}(N) \quad (10)$$

Złożoność obliczeniowa eliminacji Gaussa

Aby usunąć zmienną x_1 z jednego wiersza, należy wykonać $O(N)$ operacji. Ponieważ zmienną x_1 musimy usunąć z $N-1$ wierszy, musimy łącznie wykonać $O(N^2)$ operacji. Ponieważ musimy to samo zrobić ze zmiennymi x_2, x_3, \dots , ostatecznie musimy wykonać $O(N^3)$ operacji.

**Złożoność obliczeniowa eliminacji Gaussa
wynosi $O(N^3)$.**

Backsubstitution

Rozpatrzmy układ równań w postaci (9). Ostatnie równanie jest rozwiązane ze względu na x_N . Podstawiamy to rozwiązanie do wszystkich poprzednich równań. Teraz drugie od dołu równanie ma tylko jedną nieznaną zmienną — x_{N-1} , a coś takiego umiemy rozwiązać. Podstawiamy to rozwiązanie do równania trzeciego od dołu i do poprzednich. Teraz trzecie od dołu równanie zawiera tylko jedną zmienną, x_{N-2} . Rozwiązujemy, podstawiamy do poprzednich i tak dalej...

Ponieważ wyeliminowanie jednej zmiennej wymaga $O(N)$ operacji, a musimy wyeliminować N zmiennych, cały koszt rozwiązania układu z macierzą trójkątną górną za pomocą algorytmu *backsubstitution* wynosi $O(N^2)$. Jest to *niewiele* w porównaniu z kosztem eliminacji Gaussa.

Całkowity koszt rozwiązania układu N równań liniowych za pomocą eliminacji Gaussa z następującym *backsubstitution* wynosi $O(N^3)$.

Czy coś może pójść źle?

Cały algorytm zawali się, jeżeli w którymś momencie trzeba będzie wykonać dzielenie przez zero

$$a_{11} = 0 \text{ lub } a'_{22} = 0, \text{ lub } a''_{33} = 0 \text{ itd.}$$

Przykład

Układu równań

$$\begin{cases} y + z = 1 \\ x + y + z = 2 \\ 2x - z = 0 \end{cases} \quad (11)$$

nie da się doprowadzić do postaci trójkątnej górnej za pomocą eliminacji Gaussa. Jeśli jednak przestawimy pierwszy wiersz z drugim lub z trzecim, eliminacja Gaussa powiedzie się.

Ze względów numerycznych staramy się także unikać dzielenia przez liczby bardzo małe co do wartości bezwzględnej. *Formalnie*, w arytmetyce dokładnej, jest to wykonalne, ale *w praktyce* może to doprowadzić do bardzo znacznej utraty dokładności, tak, że ostateczny wynik będzie numerycznie bezwartościowy.

Wybór elementu podstawowego

Przypuśćmy, że na pewnym etapie eliminacji Gaussa mamy układ równań

$$\left\{ \begin{array}{cccccccc} x_1 & + & \dots & + & \dots & & \dots & & \dots & = & b_1 \\ & & x_2 & + & \dots & & \dots & & \dots & = & b_2 \\ & & & & \dots & & \dots & & \dots & = & \dots \\ & & & & & a_{kk}x_k & + & a_{k,k+1}x_{k+1} & + & \dots & = & b_k \\ & & & & & a_{k+1,k}x_k & + & a_{k+1,k+1}x_{k+1} & + & \dots & = & b_{k+1} \\ & & & & & \dots & & \dots & & \dots & = & \dots \\ & & & & & a_{Nk}x_k & + & a_{N,k+1}x_{k+1} & + & \dots & = & b_N \end{array} \right. \quad (12)$$

“Powinniśmy” teraz dzielić przez a_{kk} . Zamiast tego wśród współczynników $a_{kk}, a_{k+1,k}, a_{k+2,k}, \dots, a_{Nk}$ **wyszukujemy największy co do modułu**, permutujemy wiersze tak, aby ten największy co do modułu znalazł się w pozycji diagonalnej i dzielimy przez niego. Współczynnik wypromowany do pozycji diagonalnej nazywa się **elementem podstawowym** (ang. pivot). Ten krok algorytmu nazywa się **częściowym wyborem elementu podstawowego**. Dalej postępujemy jak poprzednio.

Koszt wyszukania jednego elementu podstawowego wynosi $O(N)$. Jeżeli robimy to w każdym kroku, całkowity koszt jest rzędu $O(N^2)$, a więc jest mały w porównaniu ze złożonością obliczeniową samej eliminacji Gaussa. Wynika z tego, iż częściowego wyboru elementu podstawowego należy zawsze dokonywać, gdyż nie zwiększa to znacznie kosztu całej procedury, może natomiast zapewnić numeryczną stabilność algorytmu.

Zamiast szukać elementu podstawowego wyłącznie w jednej kolumnie, można szukać największego co do modułu współczynnika wśród wszystkich $a_{i,j}$, $k \leq i, j \leq N$. Po znalezieniu, należy tak spermutować wiersze i kolumny układu równań, aby element podstawowy znalazł się w pozycji diagonalnej. Nazywa się to *pełnym wyborem elementu podstawowego*. Zauważmy, że koszt numeryczny wynosi $O(N^3)$, a więc staje się porównywalny z kosztem całej eliminacji Gaussa, ponadto zaś permutacja kolumn wymaga późniejszego odwikłania permutacji elementów rozwiązania, co

jest kłopotliwe. Pełny wybór elementu podstawowego zapewnia większą stabilność numeryczną, niż wybór częściowy, ale w praktyce jest rzadko używany, ze wskazanych wyżej powodów.

Do skutecznego przeprowadzenia eliminacji Gaussa potrzebna jest znajomość kolumny wyrazów wolnych, gdyż wyrazy wolne także są przekształcane i permutowane w czasie eliminacji.