Sprawozdanie - Laboratorium 2
Dekompozycja LU, odwracanie macierzy

Marta Dychała

17 marca 2021

1 Wstęp teoretyczny

Rozkład LU macierzy to jedna z metod rozwiązywania układów równań liniowych. Polega ona na zapisaniu macierzy współczynników $\mathbf{A}$ w postaci iloczynu dwóch macierzy trójkątnych: dolnej $\mathbf{L}$ oraz górnej $\mathbf{U}$, przy czym elementy na diagonali jednej z tych macierzy (u nas jest to macierz $\mathbf{L}$) są równe 1:

$$\mathbf{A} = \mathbf{L} \cdot \mathbf{U}$$

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}}_{\mathbf{A}} = \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix}}_{\mathbf{L}} \cdot \underbrace{\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}}_{\mathbf{U}} \tag{1}$$

Dany jest układ równań liniowych w postaci macierzowej:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \tag{2}$$

Po zastosowaniu metody LU układ równań (2) przyjmuje następującą postać:

$$\mathbf{L} \cdot \mathbf{U} \cdot \mathbf{x} = \mathbf{L} \cdot (\mathbf{U} \cdot \mathbf{x}) = \mathbf{L} \cdot \mathbf{y} = \mathbf{b},$$

przez co rozwiązanie układu równań (2) sprowadza się do rozwiązania dwóch układów równań z macierzami trójkątnymi:

$$\mathbf{U} \cdot \mathbf{x} = \mathbf{y},$$
$$\mathbf{L} \cdot \mathbf{y} = \mathbf{b},$$

które są układami trywialnymi, dla których łatwo jest wyznaczyć elementy zbiór rozwiązań:

$$y_i = \frac{1}{l_{ii}}[b_i - \sum_{j=1}^{i-1} l_{ij} y_j], i = 2, 3, .., n,$$

$$x_n = \frac{y_n}{u_{nn}},$$

$$x_i = \frac{1}{u_{ii}}[y_i - \sum_{j=i+1}^{n} u_{ij} x_j], i = n-1, n-2, ..., 1.$$

Omawiana właściwość rozkładu LU jest bardzo przydatna przy operacji odwracania macierzy. Zgodnie ze wzorem:

$$A \cdot A^{-1} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{n1} & a'_{n2} & \cdots & a'_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 1 \end{bmatrix} = \mathbf{I}$$

jako macierz wynikową **b** można przyjąć kolejną kolumnę macierzy jednostkowej; wtedy otrzymane wyniki są kolejnymi kolumnami macierzy odwrotnej. Biorąc pod uwagę obliczenia numeryczne jest to całkiem wydajna metoda odwracania macierzy, ponieważ operuje ona w miejscu, przez co nie jest konieczne rezerwowanie dodatkowej pamięci na inną macierz. Wiedząc, że na przekątnej jednej z macierzy powstałej z rozkładu LU znajdują się same jedynki, macierze **L** i **U** można połączyć w jedną macierz składającą się z części macierzy L pod diagonalą oraz części U powyżej i (w naszym przypadku) na diagonali.

2 Zadanie do wykonania

2.1 Opis problemu

W trakcie laboratoriów zadaniem było znalezienie rozkładu LU poniższych macierzy. Jak można zauważyć, macierz B różni się od macierzy A tylko elementem w pierwszej kolumnie pierwszego wiersza:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1.1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

Aby znaleźć rozkład LU powyższych macierzy, należało użyć procedury ***ludcmp***, która wykonuje rozkład LU w miejscu, tj. nadpisuje przekazaną macierz do procedury jej rozkładem LU. Elementy tak utworzonej macierzy znajdujące się poniżej jej diagonali, to fragment macierzy **L**, natomiast elementy na przekątnej i powyżej niej, to górna część macierzy **U**. Ponadto do procedury ***ludcmp*** przekazano wektor permutacji typu integer oraz zmienną permutacji typu float, które są istotne przy obliczaniu macierzy odwrotnej.

Kolejnym krokiem było obliczenie macierzy odwrotnej za pomocą uzyskanego wcześniej rozkładu LU. W tym celu użyto procedury ***lubksb***, która jako argumenty przyjmuje wcześniej uzyskany rozkład LU, wektor i zmienną permutacji. W celu znalezienia macierzy odwrotnej, procedura ta musi być użyta na wektorach wyrazów wolnych, które to są kolejnymi kolumnami macierzy jednostkowej 3×3:

$$b_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, b_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, b_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Procedura ***lubksb*** użyta na wektorze wyrazów wolnych nadpisuje go kolejną kolumną otrzymanej macierzy odwrotnej. Fakt ten można wykorzystać, używając procedury ***lubksb*** na macierzy jednostkowej, wywołując ją dla każdego jej wiersza. Wtedy otrzymaną macierz wynikową z macierzy jednostkowej wystarczy transponować, co można zrobić bez tworzenia dodatkowej macierzy.

Po otrzymaniu macierzy odwrotnej kolejnym zadaniem było wyznaczyć wskaźniki uwarunkowania macierzy **A** oraz **B**. W tym celu dla macierzy **A** i **B** oraz macierzy do nich odwrotnych stosuje się normę:

$$||M||_{1,\infty} = max_{1 \leqslant i,j \leqslant} |m_{i,j}|, \tag{3}$$

czyli znajduje się w macierzach wartości o największym module. Wskaźnik uwarunkowania macierzy to iloczyn normy pewnej macierzy oraz macierzy do niej odwrotnej:

$$\kappa = ||M||_{1,\infty} \cdot ||M^{-1}||_{1,\infty} \tag{4}$$

Wartość wskaźnika uwarunkowania pokazuje, czy macierz jest dobrze uwarunkowana, tj. niepodatna na zaburzenia w operacjach na tej macierzy. Po obliczeniu wskaźnika uwarunkowania macierzy dla macierzy **A** i **B**, ostatnim krokiem było obliczenie iloczynu macierzy AA^{-1} oraz BB^{-1}. Ponieważ mnożymy przez siebie macierze odwrotne, to spodziewanym wynikiem jest macierz jednostkowa.

2.2 Wyniki

Dla macierzy **A** wskaźnik uwarunkowania wynosi: $\kappa(\mathbf{A}) \approx 1.8e + 21$, z czego wynika, że macierz **A** jest bardzo źle uwarunkowana. Tak wysoki wskaźnik uwarunkowania powstał dlatego, że wyznacznik macierzy A jest równy zero, ale ze względu na nieumiejętność dokładnego obliczania operacji na ułamkach dziesiętnych przez komputer (ułamki dziesiętne są okresowe w systemie binarnym), dla programu mogła się to okazać bardzo mała liczba zbliżona do zera. Dlatego też procedura ***lubksb*** obliczyła macierz odwrotną do macierzy A, nawet jeśli taka teoretycznie nie istnieje:

$$A^{-1} = \begin{bmatrix} -5e+19 & 1e+20 & -5e+19 \\ 1e+20 & -2e+20 & 1e+20 \\ -5e+19 & 1e+20 & -5e+19 \end{bmatrix}$$

Bardzo duża wartość wskaźnika uwarunkowania macierzy A rzutuje także na iloczyn AA^{-1}, przez co otrzymana macierz nie jest macierzą jednostkową:

$$AA^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ 3.51844e+13 & -7.03687e+13 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Okazuje się, że zmiana wartości jednego z elementów macierzy może drastycznie wpłynąć na wartość wskaźnika uwarunkowania, jak to się stało w przypadku macierzy **B**. Dla tej macierzy wskaźnik uwarunkowania wynosi $\kappa(\mathbf{B}) = 333.000122$, co świadczy o znacznie lepszym uwarunkowaniu macierzy **B** od **A**. Dzieje się tak też dlatego, że wyznacznik macierzy **B** jest różny od 0 i wynosi -0.3. Dlatego też macierz odwrotna do macierzy **B** wydaje się być stosunkowo dobrze dopasowana:

$$B^{-1} = \begin{bmatrix} 10 & -20 & 10 \\ -20 & 37 & -18 \\ 10 & -17.3333 & 8.33334 \end{bmatrix}$$

Powyższe czynniki sprawiają, że iloczyn BB^{-1} obliczony w programie wydaje się być prawie poprawny:

$$BB^{-1} = \begin{bmatrix} 1 & 0 & 1.90735e-06 \\ -3.8147e-06 & 1 & 3.8147e-06 \\ 0 & 0 & 1 \end{bmatrix}$$

Niektóre wartości poza diagonalą macierzy BB^{-1} są różne od zera, co początkowo wydaje się niepoprawnością obliczeń, jednak te wartości są bardzo zbliżone do zera, co wynika też z błędów numerycznych. Można zatem przyjąć, że rozwiązanie iloczynu BB^{-1} jest całkiem zadowalające.

3 Wnioski

Ponieważ metoda LU działa w miejscu, fakt ten czyni ją bardzo wydajną metodą odwracania macierzy pozwalającą zaoszczędzić na pamięci. Z drugiej strony, ponieważ w tej metodzie nie sprawdza się, czy nastąpiło dzielenie przez zero, bardzo łatwo o zdestabilizowanie wyników względem danych wejściowych, o czym świadczy chociażby wysoki wskaźnik uwarunkowania dla macierzy **A**. Mimo to problemy tego typu można wyeliminować np. poprzez obliczenie wyznacznika macierzy, która ma być odwracana i sprawdzić czy jej moduł nie jest zbyt mały, albo sprawdzić czy wskaźnik uwarunkowania macierzy nie wychodzi zbyt duży.