MOwNiT – Układy równań liniowych - metody bezpośrednie

Przygotował: Szymon Budziak

Problem 1:

Elementy macierzy A o wymiarze $n \times n$ są określone wzorem:

$$\begin{cases} a_{1j} = 1 \\ a_{ij} = \frac{1}{i+j-1} & dla \ i \neq 1 \end{cases}$$
 $i, j = 1, \dots, n$

Przyjmij wektor x jako dowolną n-elementową permutację ze zbioru {1, -1} i oblicz wektor **b**. Następnie metodą eliminacji Gaussa rozwiąż układ równań liniowych **Ax=b** (przyjmując jako niewiadomą wektor **x**). Przyjmij różną precyzję dla znanych wartości macierzy **A** i wektora **b**. Sprawdź, jak błędy zaokrągleń zaburzają rozwiązanie dla różnych rozmiarów układu (porównaj – zgodnie z wybraną normą – wektory **x** obliczony z **x** zadany). Przeprowadź eksperymenty dla różnych rozmiarów układu.

Rozmiary układu, które zostały przetestowane w tym zadaniu to: 3, 4, ..., 80. Przyjęta prezycja to float32, float64, float 128 z biblioteki numpy.

Wyniki do zadania pierwszego

n	float type	norm	Α	В	х
3	<class 'numpy.float32'=""></class>	0e+0	[[1. 1. 1.] [0.5 0.33333334 0.25] [0.33333334 0.25 0.2]]	[1. 0.41666666 0.28333335]	[11. 1.]
3	<class 'numpy.float64'=""></class>	0e ⁺⁰⁰	[[1. 1. 1.] [0.5 0.33333333 0.25] [0.33333333 0.25 0.2]]	[1. 0.41666667 0.283333333]	[11. 1.]
3	<class 'numpy.float128'=""></class>	0e ⁺⁰	[[1. 1. 1.] [0.5 0.33333333 0.25] [0.33333333 0.25 0.2]]	[1. 0.41666667 0.283333333]	[11. 1.]
4	<class 'numpy.float32'=""></class>	6.646e ⁻¹⁵	[[1. 1. 1. 1.] [0.5 0.33333334 0.25 0.2] [0.33333334 0.25 0.2 0.16666667] [0.25 0.2 0.166666667 0.14285715]]	[0. 0.21666665 0.11666667 0.07380952]	[11. 1. -1.]
4	<class 'numpy.float64'=""></class>	3.018e ⁻¹³	[[1. 1. 1. 1.] [0.5 0.33333333 0.25 0.2] [0.33333333 0.25 0.2 0.16666667] [0.25 0.2 0.16666667 0.14285714]]	[0. 0.21666667 0.11666667 0.07380952]	[11. 1. -1.]
4	<class 'numpy.float128'=""></class>	4.187e ⁻¹⁷	[[1. 1. 1. 1.] [0.5 0.33333333 0.25 0.2] [0.33333333 0.25 0.2 0.16666667] [0.25 0.2 0.166666667 0.14285714]]	[0. 0.21666667 0.11666667 0.07380952]	[11. 1. -1.]
5	<class 'numpy.float32'=""></class>	7.282e ⁻¹³	[[1. 1. 1. 1. 1.] [0.5 0.33333334 0.25 0.2 0.16666667] [0.333333334 0.25 0.2 0.16666667 0.14285715] [0.25 0.2 0.166666667 0.14285715 0.125] [0.2 0.16666667 0.14285715 0.125 0.11111111]]	[1. 0.38333333 0.25952382 0.19880952 0.16230159]	[11. 11. 1.]
5	<class 'numpy.float64'=""></class>	9.229e ⁻¹²	[[1. 1. 1. 1. 1.] [0.5 0.33333333 0.25 0.2 0.16666667] [0.333333333 0.25 0.2 0.16666667 0.14285714] [0.25 0.2 0.166666667 0.14285714 0.125] [0.2 0.16666667 0.14285714 0.125 0.11111111]]	[1. 0.38333333 0.25952381 0.19880952 0.16230159]	[11. 11. 1.]
5	<class 'numpy.float128'=""></class>	8.671e ⁻¹⁶	[[1. 1. 1. 1. 1.] [0.5 0.33333333 0.25 0.2 0.16666667] [0.33333333 0.25 0.2 0.16666667 0.14285714] [0.25 0.2 0.166666667 0.14285714 0.125] [0.2 0.16666667 0.14285714 0.125 0.1111111]	[1. 0.38333333 0.25952381 0.19880952 0.16230159]	[11. 11. 1.]

ь					
6	<class 'numpy.float32'=""></class>	5.440e ⁻¹¹	[[1. 1. 1. 1. 1. 1.] [0.5 0.33333334 0.25 0.2 0.16666667	[0. 0.24047618 0.13452382 0.08769841 0.06230159 0.04678932]	[11. 11. 11.]
6	<class 'numpy.float64'=""></class>	3.637e ⁻¹⁰	[[1. 1. 1. 1. 1. 1.] [0.5 0.333333333 0.25 0.2 0.16666667	[0. 0.24047619 0.13452381 0.08769841 0.06230159 0.04678932]	[11. 11. 11.]
6	<class 'numpy.float128'=""></class>	3.012e ⁻¹⁶	[[1. 1. 1. 1. 1. 1.] [0.5 0.333333333 0.25 0.2 0.16666667	[0. 0.24047619 0.13452381 0.08769841 0.06230159 0.04678932]	[11. 11. 11.]

Tabela 1: 12 pierwszych wyników z zadania pierwszego

Wnioski

Możemy zauważyć, że błędy zwiększają się wraz z rosnącym rozmiarem układu. Również można zaobserwować, że błąd jest różny dla różnych precyzji. Dla typów float32 oraz float64 błędy te są znacznie większe niż błędy dla typu float128.

Problem 2:

Powtórz eksperyment dla macierzy zadanej wzorem:

$$\begin{cases} a_{ij} = \frac{2i}{j} & dla \quad j \ge i \\ a_{ij} = a_{ji} & dla \quad j < i \end{cases}$$

$$i, j = 1, \dots, n$$

Porównaj wyniki z tym, co otrzymano w przypadku układu z punktu 1). Spróbuj uzasadnić, skąd biorą się różnice w wynikach. Sprawdź uwarunkowanie obu układów.

Rozmiary układu, które zostały przetestowane w tym zadaniu to: 3, 4, ..., 80. Przyjęta prezycja to float32, float64, float 128 z biblioteki numpy.

Wyniki do zadania drugiego

n	float type	norm	Α	В	x
3	<class 'numpy.float32'=""></class>	0e ⁺⁰	[[2. 1. 0.6666667] [1. 2. 1.3333334] [0.6666667 1.3333334 2.]]	[1.66666669 0.33333337 1.33333333]	[11. 1.]
3	<class 'numpy.float64'=""></class>	3.140e ⁻¹⁶	[[2. 1. 0.66666667] [1. 2. 1.33333333] [0.66666667 1.33333333 2.]]	[1.66666667 0.33333333 1.33333333]	[11. 1.]
3	<class 'numpy.float128'=""></class>	1.212e ⁻¹⁹	[[2. 1. 0.66666667] [1. 2. 1.33333333] [0.666666667 1.333333333 2.]]	[1.66666667 0.33333333 1.33333333]	[11. 1.]
4	<pre>class 'numpy.float32'></pre>		[1. 2. 1.3333334 1.] [0.6666667 1.3333334 2. 1.5]	[1.16666669 -0.66666663 -0.16666669 -1.]	[11. 11.]
4	<class 'numpy.float64'=""></class>	2.482e ⁻¹⁶	[[2. 1. 0.66666667 0.5] [1. 2. 1.33333333 1.] [0.66666667 1.33333333 2. 1.5] [0.5 1. 1.5 2.]]	[1.16666667 -0.66666667 -0.16666667 -1.]	[11. 11.]

4	<class 'numpy.float128'=""></class>	1.212e ⁻¹⁹	[[2. 1. 0.66666667 0.5] [1. 2. 1.33333333 1.] [0.666666667 1.33333333 2. 1.5] [0.5 1. 1.5 2.]]	[1.16666667 -0.66666667 -0.16666667 -1.]	[11. 11.]
5	<class 'numpy.float32'=""></class>	2.482e ⁻¹⁶	[[2. 1. 0.6666667 0.5 0.4] [1. 2. 1.3333334 1. 0.8] [0.66666667 1.33333334 2. 1.5 1.2] [0.5 1. 1.5 2. 1.6] [0.4 0.8 1.2 1.6 2.]]	[1.56666669 0.13333338 1.03333336 0.60000002 1.20000002]	[11. 11. 1.]
5	<class 'numpy.float64'=""></class>	4.154e ⁻¹⁶	[[2. 1. 0.66666667 0.5 0.4] [1. 2. 1.33333333 1. 0.8] [0.666666667 1.333333333 2. 1.5 1.2] [0.5 1. 1.5 2. 1.6] [0.4 0.8 1.2 1.6 2.]]	[1.56666667 0.133333333 1.03333333 0.6 1.2]	[11. 11. 1.]
5	<class 'numpy.float128'=""></class>	2.484e ⁻¹⁹	[[2. 1. 0.66666667 0.5 0.4] [1. 2. 1.33333333 1. 0.8] [0.666666667 1.333333333 2. 1.5 1.2] [0.5 1. 1.5 2. 1.6] [0.4 0.8 1.2 1.6 2.]]	[1.56666667 0.133333333 1.03333333 0.6 1.2]	[11. 11. 1.]
6	<class 'numpy.float32'=""></class>	3.140e ⁻¹⁶	[[2. 1. 0.6666667 0.5 0.4 0.333333334] [1. 2. 1.3333334 1. 0.8 0.6666667] [0.6666667 1.33333334 2. 1.5 1.2 1.] [0.5 1. 1.5 2. 1.6 1.33333334] [0.4 0.8 1.2 1.6 2. 1.6666666] [0.333333334 0.6666667 1. 1.33333334 1.6666666 2.]]	[1.23333335 -0.53333333 0.03333336 -0.73333335 -0.46666661 -1.000000009]	[11. 11. 11.]
6	<class 'numpy.float64'=""></class>	9.742e ⁻¹⁶	[[2. 1. 0.66666667 0.5 0.4 0.33333333] [1. 2. 1.33333333 1. 0.8 0.66666667] [0.66666667 1.333333333 2. 1.5 1.2 1.] [0.5 1. 1.5 2. 1.6 1.333333333] [0.4 0.8 1.2 1.6 2. 1.666666667] [0.333333333 0.666666667 1. 1.333333333333333333333333333333333	[1.23333333 -0.53333333 0.03333333 -0.73333333 -0.46666667 -1.]	[11. 11. 11.]
6	<class 'numpy.float128'=""></class>	1.626e ⁻¹⁹	[[2. 1. 0.66666667 0.5 0.4 0.33333333] [1. 2. 1.33333333 1. 0.8 0.66666667] [0.66666667 1.33333333 2. 1.5 1.2 1.] [0.5 1. 1.5 2. 1.6 1.33333333] [0.4 0.8 1.2 1.6 2. 1.66666667] [0.333333333 0.666666667 1. 1.333333333	[1.23333333 -0.53333333 0.03333333 -0.73333333 -0.46666667 -1.]	[11. 11. 11.]

Tabela 2: 12 pierwszych wyników z zadania drugiego

Porównanie wyników z zadania pierwszego z wynikami z zadania drugiego

n	float type	ex 1 norm	ex 2 norm	ex 1 X given	ex 2 X given	ex 1 X	ex 2 X
3	<class 'numpy.float32'=""></class>	0e ⁺⁰	0e ⁺⁰	[1-11]	[1-11]	[11. 1.]	[11. 1.]
3	<class 'numpy.float64'=""></class>	0e ⁺⁰⁰	3.140e ⁻¹⁶	[1-11]	[1-11]	[11. 1.]	[11. 1.]
3	<class 'numpy.float128'=""></class>	0e ⁺⁰	1.212e ⁻¹⁹	[1-11]	[1-11]	[11. 1.]	[11. 1.]
4	<class 'numpy.float32'=""></class>	6.646e ⁻¹⁵	0e ⁺⁰	[1-11-1]	[1-11-1]	[11. 11.]	[11. 11.]
4	<class 'numpy.float64'=""></class>	3.018e ⁻¹³	2.482e ⁻¹⁶	[1-11-1]	[1-11-1]	[11. 11.]	[11. 11.]
4	<class 'numpy.float128'=""></class>	4.187e ⁻¹⁷	1.212e ⁻¹⁹	[1-11-1]	[1-11-1]	[11. 11.]	[11. 11.]
5	<class 'numpy.float32'=""></class>	7.282e ⁻¹³	2.482e ⁻¹⁶	[1-11-11]	[1-11-11]	[11. 11. 1.]	[11. 11. 1.]
5	<class 'numpy.float64'=""></class>	9.229e ⁻¹²	4.154e ⁻¹⁶	[1-11-11]	[1-11-11]	[11. 11. 1.]	[11. 11. 1.]
5	<class 'numpy.float128'=""></class>	8.671e ⁻¹⁶	2.484e ⁻¹⁹	[1-11-11]	[1-11-11]	[11. 11. 1.]	[11. 11. 1.]
6	<class 'numpy.float32'=""></class>	5.440e ⁻¹¹	3.140e ⁻¹⁶	[1-11-11-1]	[1-11-11-1]	[11. 11. 1. -1.]	[11. 11. 1. -1.]
6	<class 'numpy.float64'=""></class>	3.637e ⁻¹⁰	9.742e ⁻¹⁶	[1-11-11-1]	[1-11-11-1]	[11. 11. 1. -1.]	[11. 11. 1. -1.]
6	<class 'numpy.float128'=""></class>	3.012e ⁻¹⁶	1.626e ⁻¹⁹	[1-11-11-1]	[1-11-11-1]	[11. 11. 1. -1.]	[11. 11. 1. -1.]
7	<class 'numpy.float32'=""></class>	3.695e ⁻⁰⁹	2.482e ⁻¹⁶	[1-11-11-11]	[1-11-11-11]	[11. 11. 1. -1. 1.]	[11. 11. 1. -1. 1.]
7	<class 'numpy.float64'=""></class>	1.360e ⁻⁰⁸	1.694e ⁻¹⁵	[1-11-11-11]	[1-11-11-11]	[11. 1. -0.99999999 0.99999999 -0.99999999 1.]	[11. 11. 1. -1. 1.]
7	<class 'numpy.float128'=""></class>	8.865e ⁻¹³	1.131e ⁻¹⁸	[1-11-11-11]	[1-11-11-11]	[11. 11. 1. -1. 1.]	[11. 11. 1. -1. 1.]

8	<class 'numpy.float32'=""></class>	6.747e ⁻⁰⁹	6.080e ⁻¹⁶	[1-11-11-11-1]	[1-11-11-11-1]	[11. 11. 1. -1. 11.]	[11. 11. 1. -1. 11.]
8	<class 'numpy.float64'=""></class>	1.203e ⁻⁰⁷	4.672e ⁻¹⁵	[1-11-11-11-1]	[1-11-11-11-1]	[11. 1.00000001 -1.00000003 1.00000007 -1.00000008 1.00000005 -1.00000001]	[11. 11. 1. -1. 11.]
8	<class 'numpy.float128'=""></class>	5.106e ⁻¹¹	1.280e ⁻¹⁸	[1-11-11-11-1]	[1-11-11-11-1]	[11. 11. 1. -1. 11.]	[11. 11. 1. -1. 11.]

Tabela 3: 18 pierwszych wyników z porównania wyników z zadania pierwszego z zadaniem drugim

Wnioski

Dzięki tabelom możemy zauważyć, że jedną z przyczyn złych wyników dla układu równań z pierwszego problemu jest metoda wybierania elementu wiodącego. W implementacji z Problemu 1 algorytm eliminacji Gaussa jako zawsze jako element wiodący wybiera kolejne elementy przekatnej macierzy. W przypadku macierzy z Problemu 1, elementy na przekątnej redukują się na tyle, że osiągają wartość rzędu nawet 10⁻¹⁸. W przypadku macierzy z Problemu 2 jest to wielkość rzędu 10⁻², więc jest tutaj znaczna różnica. Mała wartość w przypadku Problemu 2 stanowi problem, ponieważ poszczególne wiersze w metodzie eliminacji Gaussa są dzielone przez element wiodący, to znaczy mnożone przez jego odwrotność. Jeżeli element ten ma małą wartość (< 1), to wiersze mnożone są przez duża wartość, więc błędy stają się duże w stosunku do współczynników oryginalnej macierzy. Możemy zauważyć, że wskaźniki uwarunkowania dla macierzy w problemie 1 są znacznie większe od wskaźników uwarunkowania w problemie 2. Oznacza to, że niewielki błąd znacznie wpływa na wyniki.

Problem 3:

Powtórz eksperyment dla jednej z macierzy zadanej wzorem poniżej (macierz i parametry podane w zadaniu indywidualnym). Następnie rozwiąż układ metodą przeznaczoną do rozwiązywania układów z macierzą trójdiagonalną. Porównaj wyniki otrzymane dwoma metodami (czas, dokładność obliczeń i zajętość pamięci) dla różnych rozmiarów układu. Przy porównywaniu czasów należy pominąć czas tworzenia układu. Opisz, jak w metodzie dla układów z macierzą trójdiagonalną przechowywano i wykorzystywano macierz A.

(*m*,*k* - parametry zadania):

$$\begin{cases} a_{i,i} = -m \cdot i - k \\ a_{i,i+1} = i \\ a_{i,i-1} = \frac{m}{i} & dla \quad i > 1 \\ a_{i,j} = 0 & dla \quad j < i-1 \quad oraz \quad j > i+1 \end{cases}$$
 $i, j = 1, \dots, n$

parametry zadania: k = 8, m = 3.

Rozmiary układu, które zostały przetestowane w tym zadaniu to: 3, 4, ..., 80.

Wyniki do zadania trzeciego

n	given X	gaussian X	gaussian norm	gaussian time	thomas X	thomas norm	thomas time
3	[1-11]	[11. 1.]	2.220e ⁻¹⁶	0,0002269	[11. 1.]	2.220e ⁻¹⁶	0,00003558
4	[1-11-1]	[11. 11.]	2.220e ⁻¹⁶	0,0002043	[11. 11.]	2.220e ⁻¹⁶	0,00002810
5	[1-11-11]	[11. 11. 1.]	2.220e ⁻¹⁶	0,0007462	[11. 11. 1.]	2.220e ⁻¹⁶	0,00003398
6	[1-11-11-1]	[11. 11. 11.]	2.220e ⁻¹⁶	0,0004420	[11. 11. 11.]	2.220e ⁻¹⁶	0,00005049
7	[1-11-11-11]	[11. 11. 11. 1.]	2.482e ⁻¹⁶	0,0004010	[11. 11. 11. 1.]	2.482e ⁻¹⁶	0,00003878
8	[1-11-11-11-1]	[11. 11. 11. 1. -1.]	2.220e ⁻¹⁶	0,0003461	[11. 11. 11. 11.]	2.220e ⁻¹⁶	0,00004120
9	[1-11-11-11]	[11. 11. 11. 1. -1. 1.]	2.220e ⁻¹⁶	0,0003666	[11. 11. 11. 11. 1.]	2.220e ⁻¹⁶	0,00004317
10	[1-11-11-11-11-1]	[11. 11. 11. 1. -1. 11.]	2.220e ⁻¹⁶	0,0004321	[11. 11. 11. 11. 11.]	2.220e ⁻¹⁶	0,00004760
11	[1-11-11-11-11-1 1]	[11. 11. 11. 1. -1. 11. 1.]	2.220e ⁻¹⁶	0,0004914	[11. 11. 11. 11. 11. 1.]	2.220e ⁻¹⁶	0,00004898
12	[1-11-11-11-11-1 1-1]	[11. 11. 11. 1. -1. 11. 11.]	2.220e ⁻¹⁶	0,0005996	[11. 11. 11. 11. 11. 11.]	2.220e ⁻¹⁶	0,00005383
13	[1-11-11-11-11-1 1-11]	[11. 11. 11. 1. -1. 11. 11. 1.]	2.220e ⁻¹⁶	0,0006640	[11. 11. 11. 11. 11. 11. 11. 11. 11. 11. 11. 11. 11. 11. 11. 11. 11. 11. 11. 11	2.220e ⁻¹⁶	0,0000570
14	[1-11-11-11-11-1 1-11-1]	[11. 11. 11. 11. 11. 11. 11. 11. 11.]	2.220e ⁻¹⁶	0,0005951	[11. 11. 11. 11. 11. 11. 11.]	2.220e ⁻¹⁶	0,0000304
15	[1-11-11-11-11-1 1-11-11]	[11. 11. 11. 1. -1. 11. 11. 11. 1.]	3.140e ⁻¹⁶	0,0004147	[11. 11. 11. 11. 11. 11. 11. 1.]	3.140e ⁻¹⁶	0,00003125
16	[1-11-11-11-11-1 1-11-11-1]	[11. 11. 11. 1. -1. 11. 11. 11. 11.]	3.510e ⁻¹⁶	0,0005396	[11. 11. 11. 11. 11. 11. 11. 11.]	3.510e ⁻¹⁶	0,00003511
17	[1-11-11-11-11-1 1-11-11-11]	[11. 11. 11. 1. -1. 11. 11. 11. 11. 1.]	4.710e ⁻¹⁶	0,0005404	[11. 11. 11. 11. 11. 11. 11. 11. 1.]	4.710e ⁻¹⁶	0,00003630
18	[1-11-11-11-11-1 1-11-11-11-1]	[11. 11. 11. 1. -1. 11. 11. 11. 11. 11.]	4.299e ⁻¹⁶	0,0006642	[11. 11. 11. 11. 11. 11. 11. 11. 11.]	4.299e ⁻¹⁶	0,00003610
19	[1-11-11-11-11-1	[11. 11. 11. 1. -1. 11. 11. 11. 11. 11. 1.]	4.710e ⁻¹⁶	0,0006317	[11. 11. 11. 11. 11. 11. 11. 11. 11. 1.]	4.710e ⁻¹⁶	0,00003860
20	[1-11-11-11-11-1 1-11-11-11-11-1]	[1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1,]	4.154e ⁻¹⁶	0,0008322	[11. 11. 11. 11. 11. 11. 11. 11. 11. 11.]	4.154e ⁻¹⁶	0,00004184

Tabela 4: 18 pierwszych wyników z zadania trzeciego

Wnioski

Z Tabeli 4 możemy zaobserwować, że niezależnie której metody użyjemy to błędy (norm) są takie same. Wyniki te są przewidywalne, ponieważ metody Gaussa i Thomasa są bardzo podobne, a w niektórych miejscach identyczne. Metoda Thomasa ogranicza się tylko do działania na elementach trójdiagonali. Jeśli chodzi o czasy działania to metoda Thomasa jest znacznie szybsza od metody eliminacji Gaussa. Nie jest to jednak zaskakujące, ponieważ wykonuje ona znacznie mniej operacji, ograniczając się tylko do trójdiagonali macierzy. Metoda gaussa jest skutecznym i prostym sposobem na rozwiazywanie układów równań liniowych. W niektórych przypadkach jest ona jednak wolna z powodu błędów dokładności, które są spowodowane słabym uwarunkowaniem układów wejściowych lub sposobu wybierania elementu wejściowego. W przypadku, gdy macierz jest trójdiagonalna, to warto stosować zamiast metody eliminacji Gaussa metodę Thomasa. Jest to uproszczona wersja algorytmu Gaussa, która daje wyniki o tej samej dokładności jednak działa znacznie szybciej.

Literatura

- Wykład nr 8 dr Rycerz z przedmiotu MOwNiT
- Wikipedia na temat algorytmu Gaussa oraz algorytmu Thomasa