

Pracownia z analizy numerycznej

Sprawozdanie do zadania P.2.3.

Prowadzący: mgr. Filip Chudy

Wrocław, 8 stycznia 2023, 21:37

Spis treści

1. Wstęp	1
1.1. Metodologia	1
2. Rozwiązywanie równań liniowych	2
2.1. Eliminacja Gaussa	2
2.2. Rozkład QR	2
3. Algorytm eliminacji Gaussa	3
3.1. Wyniki	3
3.2. Złożoność obliczeniowa	3
4. Transformacja Householdera	3
4.1. Podstawy teoretyczne	3
4.2. Wyniki	5
4.3. Złożoność obliczeniowa	5

1. Wstęp

Fajne podpierdalanko:

algorytmy

jak użyć do rozwiązywania równań

1.1. Metodologia

W poniższej pracy zostaną porównane dwa sposoby doprowadzania macierzy kwadratowej A do postaci górnotrójkątnej: metoda eliminacji Gaussa oraz rozkład QR z transformacją Householdera. Oba te algorytmy zostaną wykorzystane do rozwiązywania układu równań

$$Ax = b$$

dla odwracalnej macierzy A oraz dowolnego wektora b . Błąd dla każdej z metod będzie obliczany jako

$$e = -\log(\|Ax - b\|).$$

Dla metody Householdera, która produkuje nam macierz ortogonalną Q oraz górnotrójkątną R , porównane również zostaną macierze

$$A - QR$$

$$Q^T A - R$$

$$Q^T Q - I$$

poprzez wyznaczenie normy wektora Bx , gdzie B to jedna z powyższych macierzy, z x jest wektorem składającym się wyłącznie z 1. Macierze będące wynikami powyższych przekształceń powinny być macierzami zerowymi, co wynika z przekształcenia $A = QR$ oraz ortogonalności macierzy Q (wtedy $Q^{-1} = Q^T$).

Dodatkowo, sprawdzona zostanie złożoność obliczeniowa każdego z algorytmów.

2. Rozwiązywanie równań liniowych

Mając dany układ równań liniowych:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases}$$

możemy go opisać w postaci macierzy. Macierz główna tego układu równań to macierz zawierająca wszystkie współczynniki przy zmiennych $[X]$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Jeśli do macierzy głównej dołączymy wektor zawierający wszystkie wyrazy wolne $[B]$, to dostaniemy macierz rozszerzoną tego układu:

$$A|B = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{array} \right].$$

Zapisanie układu równań w postaci macierzowej ma wiele zalet. Jesteśmy w stanie w szybki sposób sprawdzić, czy równanie ma jednoznaczne rozwiązanie przez sprawdzenie czy wyznacznik macierzy głównej nie jest zerowy, gdyż jeśli $AX = B$, to $A^{-1}AX = X = A^{-1}B$. Musi więc istnieć macierz odwrotna. Sprawia to również, że zapis układu jest bardziej czytelny oraz pozwala ułatwić operowanie na takim układzie równań za pomocą komputera.

2.1. Eliminacja Gaussa

Metoda eliminacji Gaussa jest algorytmem stosowanym do rozwiązywania układu równań. Polega ona na doprowadzeniu macierzy do postaci schodkowej, tzn. zawierającej niezerowe wartości tylko na głównej przekątnej. W algorytmie dozwolone są tylko operacje na wierszach i kolumnach, czyli dodawanie lub odejmowanie od wiersza (kolumny) wielokrotności innego wiersza (kolumny) oraz zamienianie kolejności dwóch wierszy (kolumn).

Alternatywnie, na kursie algebry liniowej poznaliśmy metodę na odwracanie macierzy za pomocą eliminacji Gaussa. Wtedy z lewej strony wpisujemy oryginalną macierz, z prawej macierz identyczności i dokonując operacji wierszowych na całości staramy się doprowadzić lewą macierz do macierzy identyczności. Wtedy to co, powstanie z prawej strony będzie szukaną macierzą odwrotną. W poniższej pracy nie skorzystamy z tej wariacji na temat metody eliminacji Gaussa.

2.2. Rozkład QR

Każdą macierz A $m \times n$ o wyrazach rzeczywistych taka, że $\text{rank}(A) = n$, można zapisać jako $A = QR$, gdzie R jest macierzą górnotrójkątną, a Q ma kolumny ortogonalne. Ponieważ my będziemy

rozważać macierze A będące reprezentacją jednoznacznych układów równań, to interesują nas tylko $A \in GL_n(\mathbb{R})$.

Zauważmy, że jeśli A ma niezerowy wyznacznik, to A nie może mieć liniowo zależnych kolumn. W takim razie, wektory a_1, \dots, a_n odpowiadające kolumnom A są bazą przestrzeni \mathbb{R}^n jako maksymalny możliwy układ wektorów liniowo niezależnych. Możemy na ich podstawie stworzyć bazę ortonormalną u_1, \dots, u_n przez proces Grama-Schmidta. Wtedy dla $k = 1, \dots, n$

$$u_k = a_k - \sum_{i=1}^{k-1} \frac{\langle u_i, a_k \rangle}{\langle u_i, u_i \rangle} u_i.$$

Co więcej, dla dowolnego a_k z oryginalnej bazy możemy go zapisać za pomocą kombinacji liniowej wektorów z bazy ortonormalnej:

$$a_k = \sum_{i=1}^n c_i u_i = \sum_{i=1}^n c_i \left[a_k - \sum_{j=1}^{i-1} \frac{\langle u_j, a_k \rangle}{\langle u_j, u_j \rangle} u_j \right]$$

a ponieważ a_1, \dots, a_n były wektorami lnz, to dla $i > k$ $c_i = 0$. Niech r_k to będzie wektor zawierający współczynniki c_i dla wektora a_k :

$$r_k = \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_k \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

Czyli mamy, że

$$a_k = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} r_k$$

i dalej

$$A = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \begin{bmatrix} r_1 & r_2 & \dots & r_n \end{bmatrix}.$$

Zauważamy, że $R = \begin{bmatrix} r_1 & r_2 & \dots & r_n \end{bmatrix}$ to macierz górnotrójkątna, a Q to macierz ortogonalna.

Niech teraz A to macierz główna rozważanego układu równań, Q, R to macierze z jej rozkładu, X niech będzie wektorem wartości szukanych, a B niech będzie wektorem wyrazów wolnych. Wtedy

$$AX = B$$

$$(QR)X = B$$

i ponieważ dla macierzy ortonormalnych mamy $Q^{-1} = Q^T$, to w prosty sposób możemy zamienić powyższy układ na

$$RX = Q^T B.$$


3. Algorytm eliminacji Gaussa

3.1. Wyniki

3.2. Złożoność obliczeniowa

4. Transformacja Householdera

4.1. Podstawy teoretyczne

Transformacja Householdera [ : Householder transformation] to liniowe przekształcenie poprzez odbicie punktu wokół płaszczyzny, lub hiperpłaszczyzny, która zawiera początek układu współrzędnych.

Płaszczyzna wokół której obracamy jest zdefiniowana przez jednostkowy wektor u do niej normalny, a więc odbicie względem niej to x pomniejszony o dwa rzuty na u :

$$x' = x - 2u\langle x, u \rangle = x - 2u(u^*x)$$

co dla przestrzeni rzeczywistej wynosi

$$x' = x - 2u(u^T x).$$

Macierz tego odbicia to

$$P = I - 2uu^*$$

i jest ona Hermitowska:

$$P^* = (I - 2uu^*)^* = I^* - (2uu^*)^* = I - 2(uu^*)^* = I - 2(u^*)^*u^* = I - 2uu^* = P$$

oraz unitarna (czyli $P^*P = PP^* = I$):

$$\begin{aligned} P^*P &= P^2 = (I - 2uu^*)^2 = I - 4uu^* + 4(uu^*)^2 = \\ &= I - 4uu^* + 4u(u^*u)u^* = I - 4uu^* + 4u\langle u, u \rangle u^* = \\ &= I - 4uu^* + 4uu^* = I \end{aligned}$$

a więc w przypadku rzeczywistym dostajemy macierz symetryczną i ortogonalną, czyli taką jakiej szukamy.

Niech teraz A będzie macierzą $m \times m$, której formę QR chcemy znaleźć, a a_1, \dots, a_m będą wektorami odpowiadającymi jej kolumnom. Dalej, niech e_1, \dots, e_m będą wektorami ze standardowej bazy przestrzeni \mathbb{R}^m i ustalmy

$$\begin{aligned} v &= a_1 - \|a_1\|e_1 \\ u &= \frac{v}{\|v\|}. \end{aligned}$$

Wektor u jest jednostkowym wektorem pewnej płaszczyzny przechodzącej przez początek układu współrzędnych, możemy więc dla niego znaleźć macierz Householdera

$$P'_1 = I - 2uu^*.$$

Zauważmy, że

$$P'_1 a_1 = \begin{pmatrix} \|a_1\| \\ 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}$$

czyli zaczynamy tworzyć macierz górnotrójkątną. Proces transformacji Householdera możemy powtórzyć dla macierzy $P_1 A$ bez pierwszej kolumny i wiersza, co da nam macierz P'_2 która dla $P'_2 a_2$ daje wektor niezerowy tylko na pierwszej współrzędnej. Jednak P'_2 jest $(m-1) \times (m-1)$, więc musimy ją rozciągnąć, chociażby dodając identyczność w lewym górnym rogu. Rozciągając tę procedurę na przypadek ogólny, mamy

$$P_k = \begin{pmatrix} I_{k-1} & * \\ 0 & P'_k \end{pmatrix}$$

gdzie I_{k-1} to identyczność ale na \mathbb{R}^{k-1} .

Szukana przez nas macierz górnotrójkątna ma zatem postać

$$R = P_m \dots P_1 A$$

natomiast szukana macierz ortogonalna to

$$Q = (P_m \dots P_1)^{-1} = P_1^{-1} \dots P_m^{-1} = P_1 \dots P_m$$

z faktu, że każda z macierzy P_k jest unitarna i hermitowska.

Dla polepszenia wyników algorytmu definicja wektora v musi brać pod uwagę znak lewego górnego rogu macierzy (minora) którą będziemy poddawać transformacji Householdera. Chcemy zawsze mnożyć normę a_1 przez znak przeciwny do wspomnianego elementu macierzy.

4.2. Wyniki

4.3. Złożoność obliczeniowa