Notatki z Metod numerycznych

Jacek Olczyk

October 2018

Część I
Wykład

1 Rozwiązywanie układów równań liniowych

Znane metody

- $Ax = b, A \in \mathbb{R}^{n \times n}$
- Algorytm rokładem LU (elim. Gaussa) z wybraniem el. gł $O(\frac{2}{3}N^3)$.
- 1. złota myśl numeryka: Co zrobić jeśli zadanie jest za trudne? Zmienić zadanie
- Zamiast rozwiązywać układ równań, przybliżamy go
- Czy da się szybciej niż Gauss, który jest $O(n^3)$? To jak nisko da się zejść to problem otwarty, ale istnieją algorytmy lepsze niż sześcian.

2 Przybliżone rozwiązywanie układów równań

- Niech $A = M - Z$, wtedy $Ax = Mx - Zx = b$, zatem $Mx = Zx + b$
- TODO Metoda iteracji prostej Banacha $Mx_{n+1} = Zx_n + b$
- Jeśli wybierzemy M tak, by układ z macierzą M można było tanio rozwiązać, wtedy iteracja też będzie tania
- Chcemy, żeby M było dobrym przybliżeniem A, ale nie aż tak łatwo że
- Metoda Jacobiego
$$a_{kk}x_k^{n+1} = b - \sum_{j \neq k} a_{kj}x_j^n$$
- inny pomysł, metoda Gaussa-Seidela: $a_{kk}x_k^{(n+1)} = b_k - \sum_{j<k} a_{kj}x_j^{(n+1)} - \sum_{j>k} a_{kj}x_j^{(n)}$

- Uwaga: fakt życiowy. Gdy n jest bardzo duże, wówczas w A jest zazwyczaj bardzo dużo zer, o ile układ pochodzi z REAL LIFE^{TM}.
- To oznacza, że ilość elementów różnych od 0 jest rzędu $O(n)$. Mówimy wtedy że macierz jest rzadka.
- Wniosek: Jeśli A ma $O(n)$ niezerowych elementów, to mnożenie Ax kosztuje też $O(n)$. Ponadto, rozwiązanie układu z macierzą dolnotrójkątną też jest $O(n)$

3 Normy macierzowe i wektorowe

Normy wektorowe

$$||x||_p := (\sum_{i=1}^{N} |x_i|^p)^{\frac{1}{p}}$$

$$||x||_\infty := \max_i |x_i|$$

Norma macierzowa

$$||A||_p := \max_{x \neq 0} \frac{||Ax||_p}{||x||_p} = \max_{||x||_p} ||Ax||_p$$

Własności normy macierzowej

1. $$||Ax|| \leqslant ||A|| \cdots ||x|| \forall_{x \in \mathbb{R}^n}$$
2. $$||Ax||$$

   - nie dało się przeczytać tablicy
3. tu też coś było :(

4 Warunek wystarczający zbieżności klasycznej metody iteracyjnej ($A = M - Z$)

$$Mx_{k+1} = b + Zx_k \tag{$*$}$$

Niech x^* będzie dokładnym rozwiązaniem $Ax^* = b$

$$x_{k+1} = M^{-1}(b - Zx_k)$$

$$x_{k+1} - x^* = M^{-1}(b - Zx_k) - x^*$$
$$= M^{-1}(Ax^* - Zx_k) - x^*$$

$$= M^{-1}(Ax^* - (M - A)x_k) - x^*$$
$$= M^{-1}Ax^* + (I - M^{-1}A)x_k) - x^*$$
$$= -(I - M^{-1}A)x^* + (I - M^{-1}A)x_k) - x^*$$
$$= (I - M^{-1}A)(x^k - xu^*)$$

Czyli B pomnożont błąd k-ty.
Czyli $x_{n+1} - x^* = B(x_k - x^*) = B^2(x_k - x^*) \ldots = B^{k+1}(x_0 - x^*)$

Wniosek: Jeśli $||B|| < 1$, to $(*)$ zbieżna do $x*$ dla dow. $x_0 \in \mathbb{R}^N$

Twierdzenie: Metoda $(*)$ jest zbieżna do x^* z dowolnego x_0 wtw gdy $\rho(B) < 1$ gdzie $\rho(B) = \max\{|\lambda| : \lambda$ jest wartością własną $B\}$ - promień spektralny macierzy B Dowód pominięty

Twierdzenie: Jeśli macierz A jest ściśle diagonalnie dominująca, tzn zachodzi $|a_n| > \sum_{j \neq i} |a_{ij}|$ dla $i = 1..N$ to metoda Jacobiego jest zbieżna (dla dowolnych $x_n \in R^n$)

Dowód. Zbadajmy macierz iteracji.

$$||B||_\infty = ||I - M^{-1}A||_\infty$$

M^-1 dla macierzy diagonalnej to podnoszenie wszystkich elementów do -1.
$M^{-1}A = I +$ macierz z zerami na diagonali i ułamkami na reszcie, pierwszy wiersz to $0, a_{12}/a_{11}, a_{13}/a_{11} \ldots$
Żeby uzyskać B odejmujemy I.
$||B||_\infty = \max_i w_i$
$w_i = \sum_j |b_{i,j}| = \sum_{j \neq i} |a_{ij}/a_{ji}| = \frac{1}{|a_{ii}|} \sum_{j \neq i} |a_{ij}| < 1$ zatem norma B jest mniejsza od 1 więc normy są zbieżne. $\square$

5 Metody iteracyjne oparte na normalizacji w przestrzeni Kryłowa

k-ta przestrzeń Kryłowa

$$K_k = r_0, Ar_0, ..., A^{k-1}r_0$$

gdzie $r_k := b - Ax_k$ - reszta na k-tej iteracji

Metoda iteracyjna

- $x_k + 1 \in K_k$ przesunięta o x_0
- $x_k + 1$ normalizuje pewną miarę błędu na $x_0 + K_k$
- Na przykład:
  $$||x_k - X^*||_C \leqslant ||y - x^*||_C \forall_{y \in x_0 + K_k}$$
  lub
  $$||r_k|| \leqslant ||b - Ay||_C \forall_{y \in x_0 + K_k}$$
  gdzie $C = C^T > 0$

5.1 Metoda gradientów sprzężonych (CG - Conjugate Gradient) dla macierzy $A = A^T > 0$

5.1.1 Fakty o macierzach symetrycznych i dodatnio określonych

Niech $A = A^T > 0$ (symetryczna i dodatnio określona, a co za tym idzie $x^T Ax > 0$ dla $x \neq 0$). Wtedy:

1. Wartości własne są rzeczywiste a wektory własne są ortogonalne (czyli $A = Q \Lambda Q^T$, gdzie Q jest ortogonalna, a Λ jest diagonalna)
2. $||x||_A := \sqrt{x^T Ax}$ określa normę wektorową (norma energetyczna indukowana przez A)

Iterację metody gradientów sprzężonych definiujemy następująco:

$$x_{k+1} \in x_0 + K_k$$

$$||x_{k+1} - x^*||_A \leqslant ||y - x^*||_A \forall_{y \in x_0 + K_k}$$

Ale przecież potrzebujemy mieć rozwiązanie żeby to zrobić!

Fakt. Można stąd wyprowadzić algorytm iteracyjny, który na podstawie kilku poprzednio wyznaczonych wektorów wyznaczy x_{k+1} kosztem jednego mnożenia przez macierz A i $O(N)$

Twierdzenie. Po k iteracjach metody CG błąd $||x_k - x^*||_A \leqslant 2(\frac{\sqrt{\alpha}-1}{\sqrt{\alpha}+1})^k ||x_0 - x^*||_A$ gdzie $\alpha = \lambda_{max}(A)/\lambda_{min}(A)$.

Część II
Ćwiczenia

6 Układy nadokreślone - kontynuacja

6.1 Zadanie 1.

Macierz Hessenberga - to macierz trójkątna górna, z tym że niezerowe elementy mogą być jeden element pod diagonalą.

$$\begin{Bmatrix} x & \dots & x & x \\ x & x & \dots & x \\ & x & & \end{Bmatrix}$$

Jak najmniejszym kosztem znaleźć rozkład QR tej macierzy? Metodą Householdera? Nie ma jak wykorzystać zer na dole. Obrotami Givensa.

6.1.1 Obroty Givensa - przypomnienie

G_{ij} - macierz Givensa $b = G_{ij}a, b_j = 0$ $\cos\phi = \frac{a_i}{\sqrt{a_i^2+a_j^2}}$ $\sin\phi = \frac{a_j}{\sqrt{a_i^2+a_j^2}}$

6.1.2 Zamiana macierzy Hessenberga w górnotrójkątną obrotami Givensa

$$(G_{ij}a)_j = -\frac{a_i a_j}{\sqrt{a_i^2+a_j^2}} + \frac{a_i a_j}{\sqrt{a_i^2+a_j^2}} = 0$$

$$G_{n-1\ n} \dots G_{i\ i+1} \dots G_{12}A = R$$

6.1.3 Koszt

Robimy $n-1$ iteracji. Dla $G_{i\ i+1}$ trzeba wykonać jeden pierwiastek, $w_i = cw_i + sw_{i+1}$ oraz $w_{i+1} = -sw_i + cw_{i+1}$, łącznie $4(n-1)$ mnożeń. Łącznie $4\sum_{i=1}^{n-1} n-i = 4\sum_{i=1}^{n-1} i = \frac{4n(n-1)}{2} \sim 2n^2$

6.2 Zadanie 2.

Dane są punkty $(-1,-1),(0,2),(1,0),(2,1)$. Znajdź prostą $y = ax+b$ najlepiej przybliżającą te punkty (w sensie LZNK). (x_i, y_i) - punkty $y(x_i) - y_i$ - co chcemy zminimalizować Policzmy normę: $\min_{a,b} \sum_{i=1}^{4}(y(x_i)-y_i)^2$ Niewiadome to a oraz b, więc niech $z = \begin{Bmatrix} a \\ b \end{Bmatrix}$ oraz $d = \{y_i\}_{i=1,2,3,4}$

$$A = \begin{Bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{Bmatrix}$$

chcemy min $||Az-d||_2$ Uwaga: w ten sposób odległość między punktami a prostą liczymy w pionie, a nie najbliższą (to dobrze, tak działą LZNK). Uwaga 2: LZNK nie działa dla równania $y=a+e^{bx}$, ale dla $y=a+be^x$ już tak!

7 Normy

dodatniość, liniowość z modułem, nierówność trójkąta wektorowe p-te:

$$||x||_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$$

$$||x||_\infty = \max_i |x_i|$$

Normy macierzowe $A \in \mathbb{R}^{n\times n}$ normy indukowane $||A|| = \sup_{x\neq 0} \frac{||Ax||}{||x||} = \sup_{||x||=1} ||Ax||$ p-te normy macierzowe $||A||_p = \sup_{||x||_p=1} ||Ax||_p, p = 1, 2, \ldots, \infty$ wszystkie poza $1, 2, \infty$ zwykle się pomija

7.1 Własności norm indukowanych macierzy

1. $||Ax|| \leqslant ||A||||x||$ - z definicji mamy $||A|| \geqslant \frac{||Ax||}{||x||}$

2. $||AB|| \leqslant ||A||||B||, A, B \in \mathbb{R}^{n\times n}$ - bo

$$||ABx|| \leqslant ||A||||Bx|| \leqslant ||A||||B||||x||$$

oraz

$$||AB|| = \sup_{x\neq 0} \frac{||ABx||}{||x||} \leqslant ||A||||B||$$

Fakt. W przestrzenuiach skończonego wymiaru wszystkie normy spełniają równanie: $\exists_{c_1,c_2>0}\forall_x c_1||x||_1 \leqslant ||x||2 \leqslant c_2||x||_1$, gdzie normy są dowolne (niekoniecznie 1-a i 2-a)

Zależności między normami $x \in \mathbb{R}^n$ $c_1||x||_1 \leqslant ||x||2 \leqslant c_2||x||_1$ Jak c_1 się ma do c_2? $||x||_1^2 = (\sum_i |x_i|)^2 \geqslant ||x||_2^2$ $||x||_1 \leqslant> ||x||_2$ $||x||_1 \geqslant ||x||_\infty$ $n||x||_\infty \geqslant ||x||_1$ $||x||_\infty \leqslant ||x||_2$ $\sqrt{n}||x||_\infty \geqslant ||x||_2$ $||x||_1 \leqslant n||x||_\infty \leqslant n||x||_2$ zatem $? = n$

Nierówności $\frac{1}{n}||A||_2 \leqslant \frac{1}{\sqrt{n}}||A||_\infty \leqslant ||A||_2 \leqslant \sqrt{n}||A||_1 \leqslant n||A||_2$