

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 rozkładem 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 LIFETM.
- 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

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

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

Norma macierzowa

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

Własności normy macierzowej

1.

$$\|Ax\| \leq \|A\| \cdot \|x\| \quad \forall x \in \mathbb{R}^n$$

2.

$$\|Ax\|$$

- nie dało się przeczytać tablicy

3. tu też coś było :(

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

$$Mx_{k+1} = b + Zx_k \quad (*)$$

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

$$\begin{aligned} x_{k+1} &= M^{-1}(b - Zx_k) \\ x_{k+1} - x^* &= M^{-1}(b - Zx_k) - x^* \\ &= M^{-1}(Ax^* - Zx_k) - x^* \end{aligned}$$

$$\begin{aligned}
&= M^{-1}(Ax^* - (M - A)x_k) - x^* \\
&= M^{-1}Ax^* + (I - M^{-1}A)x_k - x^* \\
&= -(I - M^{-1}A)x^* + (I - M^{-1}A)x_k \\
&= (I - M^{-1}A)(x_k - x^*)
\end{aligned}$$

Czyli B pomnożony błąd k -ty.

Czyli $x_{n+1} - x^* = B(x_k - x^*) = B^2(x_{k-1} - x^*) \dots = B^{k+1}(x_0 - x^*)$

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

Twierdzenie: Metoda $(*)$ jest zbieżna do x^* z dowolnego x_0 wtw gdy $\rho(B) < 1$ gdzie $\rho(B) = \max\{|\lambda| : \lambda \text{ jest wartością własną } B\}$ - 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 $x_n \in \mathbb{R}^n$)

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 diagonalu i ułamekami na reszcie, pierwszy wiersz to $0, a_{12}/a_{11}, a_{13}/a_{11} \dots$

Żeby uzyskać B odejmujemy I .

$$\|B\|_\infty = \max_i w_i$$

$w_i = \sum_j |b_{i,j}| = \sum_{j \neq i} |a_{ij}/a_{ii}| = \frac{1}{|a_{ii}|} \sum_{j \neq i} |a_{ij}| < 1$ 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

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

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:

$$\|x_k - X^*\|_C \leq \|y - x^*\|_C \forall y \in x_0 + K_k$$

lub

$$\|r_k\| \leq \|b - Ay\|_C \forall y \in x_0 + K_k$$

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 $x^T A x > 0$ 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. $\|x\|_A := \sqrt{x^T A x}$ określa normę wektorową (norma energetyczna indukowana przez A)

Iterację metody gradientów sprzężonych definiujemy następująco:

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

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

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 $\|x_k - x^*\|_A \leq 2\left(\frac{\sqrt{\alpha}-1}{\sqrt{\alpha}+1}\right)^k \|x_0 - x^*\|_A$ gdzie $\alpha = \lambda_{\max}(A)/\lambda_{\min}(A)$.

6 Zagadnienia własne

Dla $A \in R^{N \times N}$ znaleźć parę własną (λ, x) , że $Ax = \lambda x$ oraz $x \neq 0$. λ pierwiastkiem wielomianu charakterystycznego: $\det(A - \lambda I) = 0$ Gdy $A = A^T$ to wartości i wektory własne rzeczywiste, istnieje Q ortogonalna $A = Q * L * Q^T$ (L to tylko lambdy na przekątnej)

3 podstawowe klasy zadań obliczeniowych dla zagadnień własnych:

1. ekstremalne wartości własne (największa, najmniejsza, etc) i odp. wektory (PageRank)

2. wartości własne bliskie zadanej wartości (wieżowce w Japonii)

3. pełne zadanie własne

Wyznaczanie wektora odpowiadającego dominującej wartości własnej (zakładamy że istnieje dokładnie jedna wartość własna że jej moduł ostro większy od innych modułów)

$$||Ax|| = ||\lambda * x|| = |\lambda| * ||x|| = |\lambda|$$

(bo $||x|| = 1$)

Metoda potęgowa x_0 startowy o normie 1

$$x_{n+1} = A * x_n$$

$$x_{n+1} := x_{n+1} / ||x_{n+1}||$$

skąd nazwa:

$$x_{n+1} = A * x_n = A * Ax_{n-1} = A^2 * x_{n-1} = \dots = A^{n+1} * x_0$$

nie robić tego w ten sposób, bo A jest duże (ale rzadkie) i będzie coraz mniej rzadkie! Lepiej iteracyjnie, bo tanio mnożyć przez rzadką macierz

Twierdzenie o zbieżności tej metody: Załóżmy, że A diagonalizowalna - istnieje Y nieosobliwe że YAY^{-1} tworzy macierz diagonalną

$A * y_i = \lambda * y_i$ gdzie y_i to kolumna Y

$$x_0 = \sum_1^n \alpha_i * y_i$$

$$x_n = A^n * x_0 = A^{n-1} * (A * x_0) =$$

$$= A^{n-1} * \sum_1^n \alpha_i * y_i =$$

$$= A^{n-1} * \sum_1^n \alpha_i * \lambda_i * y_i =$$

$$= \sum_1^n \alpha_i * \lambda_i^n * y_i =$$

$$= \lambda_1^n * \sum_1^n \alpha_i * (\lambda_i / \lambda_1)^n * y_i$$

Jeżeli λ_1 dominujące, to $\lambda_i / \lambda_1^n \rightarrow 0$ ($\lambda_1 \neq 0$)

Odzyskanie wartości własnej na podstawie przybliżenia (znaleźć takie przybliżenie λ że norma przybliżenia $A * x - \lambda * x$ minimalna) - jest to zadanie najmniejszych kwadratów iloraz Rayleigh
Transformacje spektrum: 1. Jeżeli λ ww A to $\lambda - \mu$ ww $A - \mu * I$ 2. Jeżeli λ ww A nieosobliwego to $1/(\lambda)$ ww A^{-1}

Odwrotna metoda potęgowa na zadania typu 2:

Wartości własne $(A - \mu * I)^{-1}$ to $1/(\lambda - \mu)$ Kiedy największe? Kiedy μ blisko λ_i to wtedy $1/(\lambda_i - \mu)$ dominującą ww

RQI raileigh quotient iteration, bardzo szybko zbieżne ale niekoniecznie do najbliższego oryginałowi ww TODO

3. pełny problem - metoda QR TODO

Część II

Ćwiczenia

7 Układy nadokreślone - kontynuacja

7.1 Zadanie 1.

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

$$\begin{bmatrix} x & \dots & & & x \\ x & x & \dots & & x \\ & x & x & \dots & x \\ & & \ddots & & x \\ & & & x & x & x \\ & & & & x & x \end{bmatrix}$$

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

7.1.1 Obroty Givensa - przypomnienie

1. G_{ij} - macierz Givensa
2. $b = G_{ij}a$
3. $b_j = 0$
4. $\cos \phi = \frac{a_i}{\sqrt{a_i^2 + a_j^2}}$

$$5. \sin \phi = \frac{a_j}{\sqrt{a_i^2 + a_j^2}}$$

7.1.2 Zamiana macierzy Hessenberga w górnotrójkątną obrotami Givensa

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

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

7.1.3 Koszt

Robimy $n - 1$ iteracji.

Dla G_{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.

Wszystko razem: $4 \sum_{i=1}^{n-1} n - i = 4 \sum_{i=1}^{n-1} i = \frac{4n(n-1)}{2} \sim 2n^2$.

7.2 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).

Zadane punkty oznaczamy jako (x_i, y_i) .

Zatem to, co chcemy zminimalizować to $y(x_i) - y_i$.

Policzmy normę: $\min_{a,b} \sum_{i=1}^4 (y(x_i) - y_i)^2$

Niewiadome to a oraz b , więc niech:

$$z = \begin{bmatrix} a \\ b \end{bmatrix} \quad d = [y_i]_{i=1,2,3,4} \quad A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix}$$

Teraz wystarczy użyć LZNK aby obliczyć $\min \|Az - d\|_2$:

TODO: policzyć to

Uwaga: w ten sposób odległość między punktami a prostą liczymy w pionie, a nie najbliższą (to dobrze, tak działa LZNK).

Uwaga 2: LZNK nie działa dla równania $y = a + e^{bx}$, ale dla $y = a + be^x$ już tak!

8 Normy

Przypomnienie definicji: Norma $\|\cdot\| : V \rightarrow \mathbb{R}^+$ spełnia następujące warunki:

1. $\|u + v\| \leq \|u\| + \|v\|$
2. $\|\alpha v\| = |\alpha| \|v\|$
3. $\|v\| = 0 \implies v = 0$ - wektor zerowy

p -te normy wektorowe:

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

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

Normy macierzowe Niech $A \in \mathbb{R}^{n \times n}$.
Macierzowe normy indukowane są postaci

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|$$

p -te normy macierzowe

$$\|A\|_p = \sup_{\|x\|_p=1} \|Ax\|_p, p = 1, 2, \dots, \infty$$

wszystkie poza 1, 2, ∞ zwykle się pomija

8.1 Własności norm indukowanych macierzy

1. $\|Ax\| \leq \|A\| \|x\|$ - z definicji mamy $\|A\| \geq \frac{\|Ax\|}{\|x\|}$
2. $\|AB\| \leq \|A\| \|B\|, A, B \in \mathbb{R}^{n \times n}$ - bo

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

oraz

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

Fakt. W przestrzeniach skończonego wymiaru wszystkie normy spełniają równanie: $\exists_{c_1, c_2 > 0} \forall_x c_1 \|x\|_1 \leq \|x\|_2 \leq c_2 \|x\|_1$, gdzie normy są dowolne (niekoniecznie pierwsza i druga)

8.2 Zależności między normami

Niech $x \in \mathbb{R}^n$, a normy będą p -te.

$$\|x\|_1^2 = \left(\sum_i |x_i|\right)^2 \geq \|x\|_2^2$$

$$\|x\|_1 \leq \alpha \|x\|_2$$

Jakie α wybrać?

$$\|x\|_1 \geq \|x\|_\infty$$

$$n\|x\|_\infty \geq \|x\|_1$$

$$\|x\|_\infty \leq \|x\|_2$$

$$\sqrt{n}\|x\|_\infty \geq \|x\|_2$$

$$\|x\|_1 \leq n\|x\|_\infty \leq n\|x\|_2$$

Zatem $\alpha = n$.

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

8.3 Wzory na normy macierzowe

$$\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$$

$$\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$$

Zatem $\|A^T\|_1 = \|A\|_\infty$.

Norma druga (spektralna)

$$\|A\|_2 = \max_{\lambda \in \delta(A^T A)} \sqrt{\lambda}$$

Gdzie $\delta(M)$ jest zbiorem wartości własnych macierzy M .

Jeśli Q ortogonalna: $\|Q\|_2 = 1$, co za tym idzie $\|I\|_2 = 1$ λ - wartość własna oraz v - wektor własny spełniają $Av = \lambda v$

Norma Frobeniusa (Euklidesowa)

$$\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}$$

Nie jest normą indukowaną, bo dla wszystkich norm indukowanych zachodzi $\|I\| = \sup_{x \neq 0} \frac{\|Ix\|}{\|x\|} = 1$, a $\|I\|_F = \sqrt{n}$ - zatem nie pochodzi od drugiej normy wektorowej!