

Metody numeryczne

Janusz Miller

Katedra Informatyki Stosowanej
AGH University of Science and Technology

2.11.2021

Część V

Osobliwe układy równań liniowych Aproksymacja

- 1 Macierze prostokątne w algebrze liniowej
 - Rozwiązywanie układów równań z macierzą prostokątną
 - Rozkład macierzy wg wartości szczególnych

Spłot dotychczasowych wątków

Krótkie podsumowanie



- ① Algebra liniowa – macierze nieosobliwe
 - ① rozwiązywanie nieosobliwych układów równań,
 - ② wartości i wektory własne.
- ② Interpolacja.

Krótko o każdym punkcie:

Krótkie podsumowanie

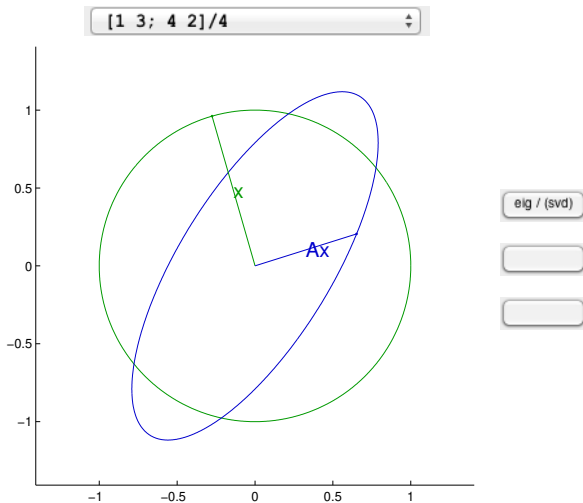
O macierzach kwadratowych, nieosobliwych



- $A \in \mathbb{R}^{n,n}$, $\text{rank}(A) = n$,
- Pomijamy macierze o elementach zespolonych,
- Dla każdej macierzy $A \in \mathbb{R}^{n,n}$ istnieje macierz ortogonalna $U \in \mathbb{R}^{n,n}$ taka, że $U^T A U = R$, gdzie $R \in \mathbb{R}^{n,n}$ jest macierzą trójkątną (z kwadratowymi blokami 1x1 lub 2x2 na przekątnej głównej).
- Macierz A jest operatorem przekształcenia przestrzeni $\mathbb{R}^n$ w $\mathbb{R}^n$.

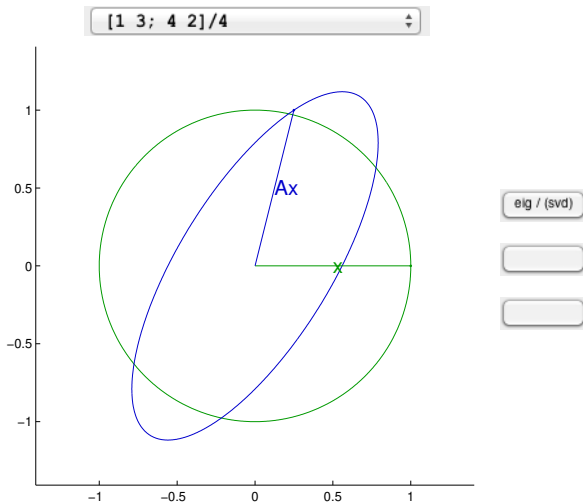
Krótkie podsumowanie

O macierzach kwadratowych, nieosobliwych c.d.



Krótkie podsumowanie

O macierzach kwadratowych, nieosobliwych c.d.



Krótkie podsumowanie

O rozwiązywaniu nieosobliwych układów równań



- Istnieje dokładnie jedno rozwiązanie.
- Algorytmy metod bezpośrednich zawierają etapy
 - 1 (bardziej kosztowny obliczeniowo) rozkładu macierzy A iloczyn macierzy trójkątnych (LU) albo ortogonalnej i trójkątnej (QR),
 - 2 (mało kosztowny) rozwiązania układu równań z macierzą trójkątną (np. metodą wstecznego postawiania).
- Algorytmy iteracyjne pozostawiamy bez kontynuacji.

Krótkie podsumowanie

O wartościach i wektorach własnych



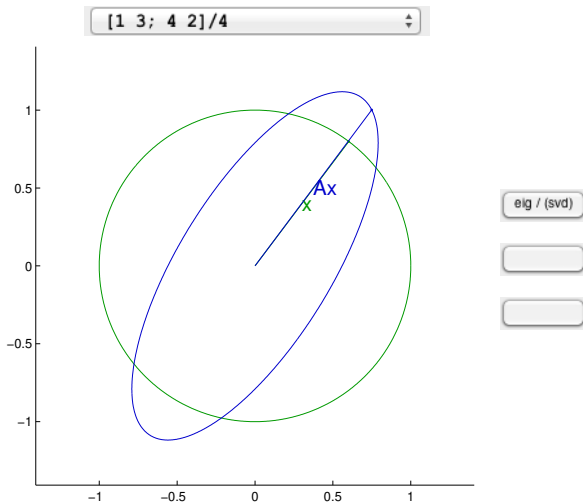
AGH

www.agh.edu.pl

- $Av = \lambda v$ - interpretacja.
- Wartości własne macierzy $A \in \mathbb{R}^{n,n}$ są rzeczywiste lub zespolone. Są wartościami na przekątnej macierzy R rozkładu $A = URU^T$ (w przypadku bloku 2x2 są pierwiastkami równania charakterystycznego tego bloku).
- Wektory własne nie zawsze istnieją i nie zawsze są jednoznaczne.
- Rzeczywista i **symetryczna** macierz A ma tylko rzeczywiste wartości własne.
 $A = U\Lambda U^T$, $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.
- Ważne np. dla wskaźnika uwarunkowania macierzy.

Krótkie podsumowanie

O wartościach i wektorach własnych c.d.



Krótkie podsumowanie

O interpolacji



Funkcja (liniowo zależna od jej współczynników)
interpolująca zadane węzły jest **jednoznacznie określona**
przez **nieosobliwy układ równań liniowych**.

Rzeczywista macierz prostokątna



$$A \in \mathbb{R}^{m,n}$$

jest operatorem przekształcenia przestrzeni $\mathbb{R}^n$
w przestrzeń $\mathbb{R}^m$

$$y = Ax, \quad x \in \mathbb{R}^n, \quad y \in \mathbb{R}^m.$$

Układy równań



Rozwiązania układów równań w sensie najmniejszych kwadratów

- ① $m < n$ - podokreślony układ równań.
Rozwiązanie: o najmniejszej normie $|| \cdot ||_2$.
- ② $m > n$ - nadokreślony układ równań.
Rozwiązanie: minimalizujące normę $|| \cdot ||_2$ residuum.

W zadaniach aproksymacji zachodzi przypadek układów nadokreślonych.

Algorytm rozwiązywania – w dalszej części wykładu „Aproksymacja”.

Równania nadokreślone c.d.

Interpolacja a aproksymacja



AGH

www.agh.edu.pl

Dla interpolacji / aproksymacji funkcjami liniowo zależnymi od ich współczynników –

- Układ równań: nieosobliwy / nadokreślony.
- Dla $i \neq j$:
Interpolacja: $x_i \neq x_j$, Aproksymacja: dopuszczalne $x_i = x_j$.

Inne niż liniowe zadanie najmniejszych kwadratów
możliwości interpolacji / aproksymacji:

- wybór funkcji nieliniowej względem jej współczynników,
- wybór innej normy.

Poniżej: Krótko o innych normach.

Wartości szczególne



Inne określenia: dekompozycja wg wartości osobliwych, singularnych (ang. *singular value decomposition* – SVD).

Twierdzenie. Każdą macierz $A \in \mathbb{R}^{m,n}$ można przedstawić w postaci iloczynu

$$A = U\Sigma V^T, \quad (1)$$

gdzie

- $U \in \mathbb{R}^{m,m}$ i $V \in \mathbb{R}^{n,n}$ są macierzami – ortogonalnymi¹,
- Kolumny U - lewe wektory szczególne, kolumny V - prawe wektory szczególne.
- $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r, 0, \dots, 0)$, $r = \text{rank}(A)$.
- σ_i – i-ta wartość szczególna macierzy A .

¹kolumny tworzą zbiór wektorów ortonormalnych

Wartości szczególne c.d.



Jeżeli rozkład (1) zastosujemy dla macierzy symetrycznej $A^T A$, to otrzymamy

$$A^T A = V \Sigma^T U^T U \Sigma V^T = V \Sigma^T \Sigma V^T = V \Sigma^2 V^T.$$

$\Sigma^2 = \text{diag}(\sigma_1^2, \dots, \sigma_r^2, 0, \dots, 0)$, zwyczajowo – uporządkowane w porządku malejącym.

Analogicznie

$$A A^T = U \Sigma^2 U^T.$$

Zatem σ_i^2 jest wartością własną macierzy $A^T A$ i $A A^T$.

Wektory szczególne a wektory własne



- Macierz prostokątna nie ma wektorów własnych (Ax ma inny rozmiar niż x).
- Odpowiednikami wektorów własnych macierzy kwadratowych są dla macierzy prostokątnych wektory szczególne (singularne).
- Para wektorów (v_i, u_i) jest parą wektorów szczególnych (v_i – prawym, w przestrzeni $\mathbb{R}^n$, a u_i – lewym, w przestrzeni $\mathbb{R}^m$) jeżeli
 - 1 $Av_i = \sigma_i u_i$ oraz
 - 2 istnieje wektor v_j ortogonalny do v_i taki, że wektor $u_j = \frac{1}{\sigma_j} Av_j$ jest ortogonalny do u_i .Liczby σ_i i σ_j są wartościami szczególnymi.

Wektory szczególne a wektory własne



Powyższy warunek dla wszystkich wektorów szczególnych można zapisać w postaci macierzowej

$$AV = U\Sigma \quad (2)$$

równoważnej (po prawostronnym pomnożeniu (1) przez V) równości (1).

Inaczej mówiąc:

Zapiszmy (2) w postaci r równań (r jest rzędem macierzy A)

$$Av_1 = \sigma_1 u_1$$

$$\dots$$

$$Av_r = \sigma_r u_r.$$

Jeżeli wektory v_i i v_j są prawymi wektorami szczególnymi w przestrzeni „wejściowej”, to ich obrazy, odpowiednio $\sigma_i u_i$ i $\sigma_j u_j$ w przestrzeni „wyjściowej” są ortogonalne (i są lewymi wektorami szczególnymi).

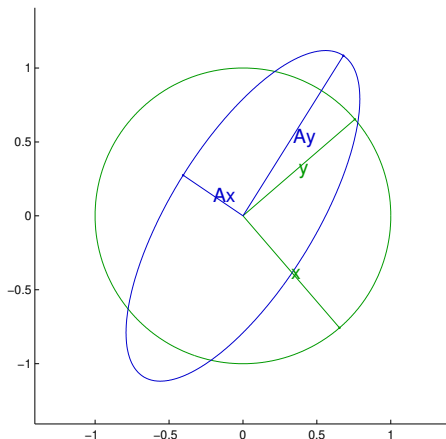
Wektory szczególne a wektory własne

Interpretacja graficzna



Wektory szczególne: **prawe** i **lewe**

[1 3; 4 2]/4



svd / (eig)

Wartości i wektory szczególne

Interpretacja macierzy U, Σ, V



Przekształcenie przestrzeni $\mathbb{R}^n$ operatorem A w przestrzeń $\mathbb{R}^m$ jest złożeniem

- 1 obrotu V w $\mathbb{R}^n$,
- 2 skalowania Σ ,
- 3 obrotu U w $\mathbb{R}^m$.

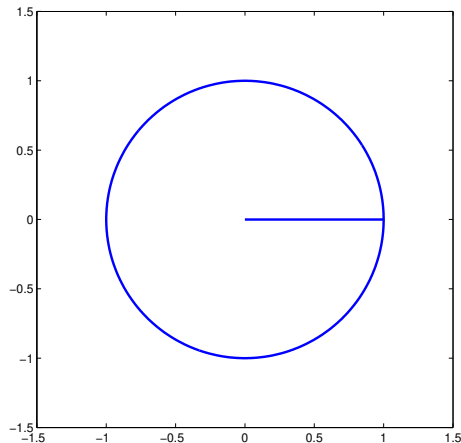
Można porównać liczbę „stopni swobody” macierzy A z liczbą parametrów ww. przekształceń, np. dla $n=2$ lub 3 .

Wartości i wektory szczególne

Interpretacja macierzy U, Σ, V



$A = [1 \ 3; 4 \ 2]/4$. Figura w przestrzeni wejściowej $\mathbb{R}^n$.

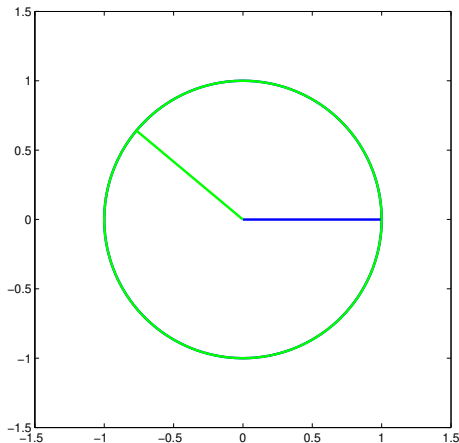


Wartości i wektory szczególne

Interpretacja macierzy U, Σ, V



$A = [1 \ 3; 4 \ 2]/4$. Figura po obrocie macierzą V^T .

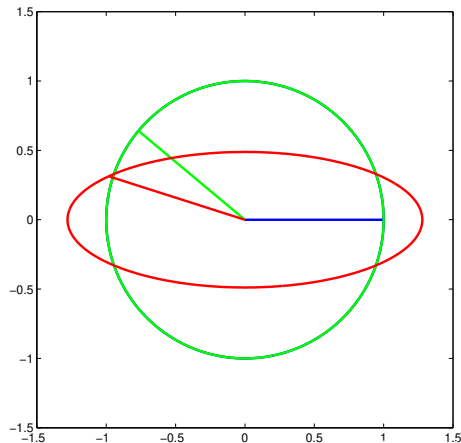


Wartości i wektory szczególne

Interpretacja macierzy U, Σ, V



$A = [1 \ 3; 4 \ 2]/4$. Figura po skalowaniu
 $\sigma_1 = 1.2792$, $\sigma_2 = 0.4886$.

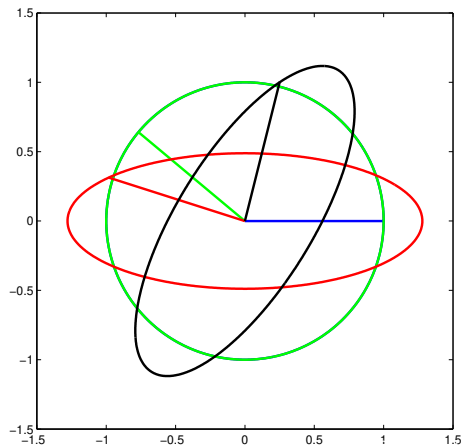


Wartości i wektory szczególne

Interpretacja macierzy U, Σ, V



$A = [1 \ 3; 4 \ 2]/4$. Figura po obrocie macierzą U w przestrzeni wyjściowej $\mathbb{R}^m$.



Wartości szczególne c.d.

Normy macierzy



AGH

Normy macierzy wyrażone za pomocą wartości szczególnych:

- $\|A\|_2 = \max_i \sigma_i(A),$
- nuklearna $\|A\|_N = \sum_{i=1}^{\min\{m,n\}} \sigma_i(A),$
- Frobeniusa

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} = \sqrt{\sum_{i=1}^{\min\{m,n\}} \sigma_i^2(A)}.$$

Principal Component Analysis - PCA



AGH

- „Kompresja” danych – zamiast olbrzymich macierzy - kilka współczynników istotnych dla opisu badanej/szukanej zależności.
- Pominięcie wpływów nieistotnych, szumów, błędów numerycznych, pomiarowych, przypadkowych.
- Zastosowanie w sieciach neuronowych, głębokim uczeniu maszynowym.
- Trzeba pamiętać, że przekształcenie $x \rightarrow y: y = Ax$ jest przekształceniem liniowym, a nie afinicznym (przekształcenie zera w zero).
- Porównanie z zadaniem najmniejszych kwadratów.