"A mathematician, like a painter or a poet, is a maker of patterns. If his patterns are more permanent than theirs, it is because they are made with ideas."

Godfrey Harold Hardy

6

Descompunerea valorilor singulare

Compresia datelor

Transmiterea eficienta si stocarea unui volum mare de date au devenit o problema majora a lumii tehnologiei. In cele ce urmeaza vom prezenta o metoda de compresie a datelor in asa fel incat acestea sa poata fi transmise mai rapid si stocate pe mai putin spatiu. Reamintim ca rata de compresie a datelor se calculeaza cu formula:

$$\text{rata compresie} = \frac{\text{marime fisier necomprimat}}{\text{marime fisier comprimat}}$$

Astfel un fisier de 10 MB care este comprimat in unul de 2MB va avea o rata de compresie 5 : 1.

Spre exemplu, o fotografie alb-negru poate fi scanata si apoi stocata ca o matrice A atasand fiecarui pixel o valoare numerica in functie de nivelul de gri al acestuia.. Daca folosim 256 nivele diferite de gri (0= alb, 255 = negru), atunci elementele matricei ar fi numere intregi cuprinse intre 0 si 255. Imaginea poate fi recuperata din matricea A afisand pixelii colorati in functie de nivelele lor de gri. Daca matricea este de tip $m \times n$ atunci am putea sa stocam toate cele $m \cdot n$ date individual. O alternativa consta in gasirea unei descompuneri a matricei pentru care sa fie nevoie de mai putine date. Spre exemplu, daca am avea o descompunere de tipul:

$$A = \sigma_1 \boldsymbol{u_1} \cdot \boldsymbol{v_1^T} + \sigma_2 \boldsymbol{u_2} \cdot \boldsymbol{v_2^T} + \ldots + \sigma_r \boldsymbol{u_r} \cdot \boldsymbol{v_r^T}$$

unde $\mathbf{u_i} \in \mathcal{M}_{1\times m}(\mathbb{R})$, $\mathbf{v_i} \in \mathcal{M}_{1\times n}(\mathbb{R})$ si $\sigma_i > 0$, atunci nu trebuie sa stocam decat rm date pentru cei r de u_i, rn date pentru cei r de v_i si r date pentru cei r de σ_i. In concluzie $r+rm+rn = r(m+n+1)$ date. Ce nu e suficient de evident la aceasta descompunere este ca fiecare termen este de fapt o matrice $m\times n$. Tinand cont de modul in care am modelat matematic stocarea fotografiei alb-negru o astfel de matrice adauga cateva nuante de gri fiecarui pixel al imaginii urmand ca toate impreuna sa realizeze reconstructia totala a acesteia. Sa presupunem acum ca unii dintre acesti σ_i sunt foarte mici atunci eliminand termenii corespunzatori lor vom obtine ceea ce numim o aproximatie de rang k a lui A:

$$A = \sigma_1 \boldsymbol{u_1} \cdot \boldsymbol{v_1^T} + \sigma_2 \boldsymbol{u_2} \cdot \boldsymbol{v_2^T} + \ldots + \sigma_r \boldsymbol{u_k} \cdot \boldsymbol{v_k^T}, \quad k < r$$

In acest fel am putea stoca doar aceste $k(m + n + 1)$ date si vom obtine o aproximare a imaginii intiale, suficient de buna pentru scopul nostru. Spre exemplu o aproximare de rang 100 a unei imagini cu 1000×1000 pixeli are nevoie doar de $100(1000 + 1000 + 1) = 200.100$ date, astfel rezulta o compresie cu o rata de compresie de aproape $5 : 1$.

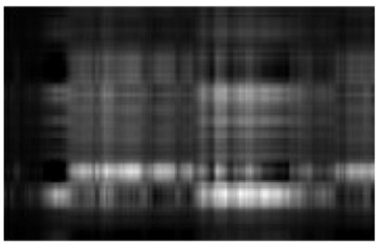
Mai jos aveti cateva aproximari ale versiunii alb-negru a imaginii de la inceputul capitolului:

Reconstructie folosind matricea intreaga

Reconstructie folosind o aproximare de rang 100

Reconstructie folosind o aproximare de rang 10

Reconstructie folosind o aproximare de rang 3



Mai jos aveti un cod Matlab care genereaza astfel de aproximari ale imaginii tigrului:

```
% Read the picture of the tiger, and convert to black and white.
tiger = rgb2gray(imread('tiger.jpg'));

% Downsample, just to avoid dealing with high-res images.
tiger = im2double(imresize(tiger, 0.5));

% Compute SVD of this tiger
[U, S, V] = svd(tiger);

% Plot the magnitude of the singular values (log scale)
sigmas = diag(S);
figure; plot(log10(sigmas)); title('Singular Values (Log10 Scale)');
figure; plot(cumsum(sigmas) / sum(sigmas)); title('Cumulative Percent of Total
Sigmas');

% Show full-rank tiger
figure; subplot(4, 2, 1), imshow(tiger), title('Full-Rank Tiger');

% Compute low-rank approximations of the tiger, and show them
ranks = [200, 100, 50, 30, 20, 10, 3];
for i = 1:length(ranks)
    % Keep largest singular values, and nullify others.
    approx_sigmas = sigmas; approx_sigmas(ranks(i):end) = 0;

    % Form the singular value matrix, padded as necessary
    ns = length(sigmas);
    approx_S = S; approx_S(1:ns, 1:ns) = diag(approx_sigmas);

    % Compute low-rank approximation by multiplying out component matrices.
    approx_tiger = U * approx_S * V';

    % Plot approximation
    subplot(4, 2, i + 1), imshow(approx_tiger), title(sprintf('Rank %d Tiger',
ranks(i)));
end
```

Amprente digitale


Original Reconstructie

In 1924 FBI a inceput sa colecteze amprente digitale si acum are mai mult de 100 de milioane de astfel de amprente in fisierele sale. Pentru a reduce costul de depozitare, FBI a inceput sa lucreze cu Laboratorul National din Los Alamos, precum si cu alte grupuri, pentru a gasi metode de comprimare ale fisierelor, care contineau amprentele in forma electronica. In figura alaturata avem amprenta originala si reconstructia ei, dintr-un fisier comprimat la o rata 26 : 1, prin metoda descompunerii valorilor singulare.

Sinteza teorie

- orice matrice $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ poate fi descompusa sub forma

$$A = \begin{pmatrix} u_1 \mid & u_2 \mid \ldots \mid u_r \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_r \end{pmatrix} \begin{pmatrix} v_1^t \\ v_2^t \\ \vdots \\ v_r^t \end{pmatrix}$$

unde u_i si v_i sunt vectori coloana iar $\sigma_i > 0$ sunt numere reale pozitive numite valori singulare ale matricei.

- descompunerea de mai sus este echivalenta cu

$$A = \sigma_1 \boldsymbol{u_1} \cdot \boldsymbol{v_1^t} + \sigma_2 \boldsymbol{u_2} \cdot \boldsymbol{v_2^t} + \ldots + \sigma_r \boldsymbol{u_r} \cdot \boldsymbol{v_r^t}$$

si poarta numele de decompunerea valorilor singulare

Remarca

Numarul r al valorilor singulare este de fapt rangul matricei A. Mai mult, daca asezam valorile singulare intr-o ordine descrescatoare

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$$

atunci aceste valori ne dau de fapt distantele de la matricea A la multimile matricelor de rang mai mic

$$\sigma_{k+1} = \min_{rang(B)=k} dist(A, B), \quad k < r = rang(A)$$

Aceasta relatie se interpreteaza in felul urmator: distanta de la A la cea mai "apropiata" matrice de rang k este σ_{k+1}. Distanta dintre doua matrice se calculeaza cu formula

$$dist(A, B) = \max_{\|u\|=1} ||(A - B)u||$$

unde u este un vector coloana si norma lui $\|u\|$ este cea obisnuita. De remarcat ca spatiul matricelor este un spatiu metric, pe care putem defini distante extrem de utile in probleme de aproximare.

Algoritmul de obtinere a descompunerii SVD

- valorile singulare $\sigma_1, \sigma_2, \ldots, \sigma_r$ se afla cu formulele

$$\sigma_1 = \sqrt{\lambda_1}, \ \sigma_2 = \sqrt{\lambda_2}, \ldots, \sigma_r = \sqrt{\lambda_r}$$

unde $\lambda_1, \lambda_2, \ldots, \lambda_r$ sunt valorile proprii nenule ale matricei $A^t A$.

- vectorii $v_1, v_2, \ldots, v_r$ sunt vectori proprii ortonormati corespunzatori valorilor proprii $\lambda_1, \lambda_2, \ldots, \lambda_r$

$\Longrightarrow$ daca avem o valoare proprie multipla se aplica procedeul Gram-Schmidt pentru a obtine vectori proprii ortonormati corespunzatori acesteia

$\Longrightarrow$ pentru valorile proprii simple (ordine de multiplicitate 1) se imparte la norma sa un vector propriu corespunzator

• vectorii u se obtin cu formulele

$$u_i = \frac{1}{\sigma_i} A v_i, \qquad i = \overline{1,r}$$

Remarca

De retinut ca matricea $A^t A$ este simetrica si orice matrice simetrica este dia-gonalizabila iar matricea care o diagonalizeaza poate fi aleasa cu coloane reprezentate de catre vectori ortonormati. Mai mult orice valoare proprie a lui $A^t A$ va fi pozitiva. Asadar algoritmul de mai sus va functiona intot-deauna si nu se pune problema ca dimensiunile subspatiilor proprii sa nu coincida cu multiplicitatile algebrice corespunzatoare.

Forme patratice

• daca $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ este un vector linie si $A \in \mathcal{M}_n(\mathbb{R})$ este o matrice simetrica, numim expresia

$$q_A(\bar{\boldsymbol{x}}) = \bar{\boldsymbol{x}} \cdot A \cdot \bar{\mathbf{x}}^t$$

forma patratica asociata lui A, adica

$$q_A(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} a_{ii}^2 x_i^2 + \sum_{i,j=1}^{n} a_{ij} x_i x_j, \qquad \bar{\boldsymbol{x}} = (x_1, x_2, \ldots, x_n)$$

• o forma patratica q_A este pozitiv definita $\Longleftrightarrow q_A(\bar{\boldsymbol{x}}) > 0, \quad \forall\, \bar{\boldsymbol{x}} \in \mathbb{R}^n$

$\Longleftrightarrow$ A este pozitiv definita $\Longleftrightarrow$ toti minorii principali sunt strict pozitivi

• o forma patratica q_A este negativ definita $\Longleftrightarrow q_A(\bar{\boldsymbol{x}}) < 0, \quad \forall\, \bar{\boldsymbol{x}} \in \mathbb{R}^n$

$\Longleftrightarrow$ A este negativ definita $\Longleftrightarrow$ minorii principali alterneaza ca semn si primul este strict negativ

• daca P este matricea care diagonalizeaza pe A (adica $A = PDP^{-1}$) atunci prin transformarea $\bar{\boldsymbol{x}}^t = P\bar{\boldsymbol{y}}^t$ forma patratica devine

$$q_D(\bar{\boldsymbol{y}}) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2$$

• o forma patratica este pozitiv definita $\Longleftrightarrow$ toate valorile proprii sunt strict pozitive

• o forma patratica este negativ definita $\Longleftrightarrow$ toate valorile proprii sunt strict negative

Probleme de optimizare

Teorema de extrem conditionat

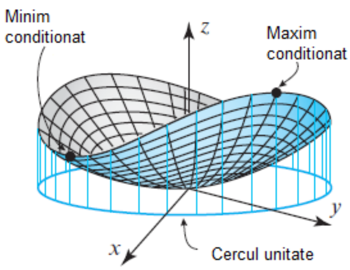
Fie $A \in \mathcal{M}_n(\mathbb{R})$ o matrice simetrica a carui valori proprii sunt:

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$$

Atunci forma patratica $q_A(\bar{\boldsymbol{x}})$ atinge o valoare maxima si una minima pe multimea vectorilor $\bar{\boldsymbol{x}}$ pentru care $\|\bar{\boldsymbol{x}}\| = 1$.
Valoarea maxima este λ_n si este atinsa intr-un vector propriu corespunzator lui λ_n.
Valoarea minima este λ_1 si este atinsa intr-un vector propriu corespunzator lui λ_1.

Remarca

Pentru a vizualiza problema, in cazul $\bar{\boldsymbol{x}} = (x, y) \in \mathbb{R}^2$, forma patratica q_A poate fi reprezentata ca fiind o suprafata $z = \bar{\boldsymbol{x}} \cdot A \cdot \bar{\boldsymbol{x}}^t$ intr-un sistem de axe Oxyz. Acum $\|\bar{\boldsymbol{x}}\| = 1$ devine cercul unitate $x^2 + y^2 = 1$.



Geometric, problema aflarii minimului si maximului, relativ la conditionarea $\|\bar{\boldsymbol{x}}\| = 1$, se reduce la aflarea punctului aflat la cota z cea mai mare si respectiv cota z cea mai mica pe intersectia dintre suprafata si cilindrul drept format de cercul unitate

Fie $f : \mathbb{R}^n \to \mathbb{R}$ o functie pentru care derivatele partiale de ordin doi exista si sunt continue. Fie $\bar{\boldsymbol{a}} = (a_1, a_2, \ldots, a_n)$ un punct critic al lui f. Atunci f are un punct de minim local in $\bar{\boldsymbol{a}}$ daca matricea Hessiana:

$$H(\bar{\boldsymbol{a}}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 x_1}(\bar{\boldsymbol{a}}) & \frac{\partial^2 f}{\partial x_1 x_2}(\bar{\boldsymbol{a}}) & \ldots & \frac{\partial^2 f}{\partial x_1 x_n}(\bar{\boldsymbol{a}}) \\ \frac{\partial^2 f}{\partial x_2 x_1}(\bar{\boldsymbol{a}}) & \frac{\partial^2 f}{\partial x_2 x_2}(\bar{\boldsymbol{a}}) & \ldots & \frac{\partial^2 f}{\partial x_2 x_n}(\bar{\boldsymbol{a}}) \\ \vdots & \vdots & \ldots & \vdots \\ \frac{\partial^2 f}{\partial x_n x_1}(\bar{\boldsymbol{a}}) & \frac{\partial^2 f}{\partial x_n x_2}(\bar{\boldsymbol{a}}) & \ldots & \frac{\partial^2 f}{\partial x_n x_n}(\bar{\boldsymbol{a}}) \end{pmatrix}$$

este pozitiv definita. Analog, f are un punct de maxim local in $\bar{\boldsymbol{a}}$ daca matricea Hessiana este negativ definita. Daca matricea $H(\bar{\boldsymbol{a}})$ are si valori proprii pozitive si negative atunci avem un punct sa.

Probleme rezolvate

Problema 1

Aflati valorile singulare ale matricei

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$$

si apoi determinati descompunerea valorilor singulare.

Solutie: Intai aflam valorile singulare care prin definitie sunt valorile proprii ale matricei A^tA. Avem:

$$A^tA = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Polinomul caracteristic este:

$$p(\lambda) = \det(A^tA - \lambda I) = \lambda^2 - 4\lambda + 3$$

iar valorile proprii vor fi $\lambda_1 = 3$ si $\lambda_2 = 1$. Deci valorile singulare ale matricei A sunt:

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{3}, \quad \sigma_2 = \sqrt{\lambda_2} = 1$$

In continuare aflam subspatiile proprii corespunzatoare lui λ_1 si λ_2 :

$$S_{\lambda_1} = \{v : (A^tA - \lambda_1 I)v = \bar{0}\} = \left\{ v = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\}$$

Asadar $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ este un vector dintr-o baza a acestuia dar avem nevoie de un vector de norma 1 si prin urmare $v_1 = \frac{v}{\|v\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ este vectorul cautat Analog gasim:

$$S_{\lambda_2} = \{v : (A^tA - \lambda_2 I)v = \bar{0}\} = \left\{ v = \begin{pmatrix} \alpha \\ -\alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\}$$

si prin urmare $v_2 = \frac{v}{\|v\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ este vectorul cautat.

Pentru a afla u_1 si u_2 aplicam formula $u_i = \frac{1}{\sigma_i} A v_i$ si gasim:

$$u_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \end{pmatrix}$$

respectiv:

$$u_2 = 1 \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$$

In final descompunerea cautata este:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{6}}{3} & 0 \\ \frac{\sqrt{6}}{6} & \frac{-\sqrt{2}}{2} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Problema 2

Aflati minimul si maximul formei patratice

$$q(x, y) = 5x^2 + 5y^2 + 4xy$$

atunci cand $x^2 + y^2 = 1$.

Solutie: Forma patratica poate fi exprimata in forma matriciala ca:

$$q(x, y) = 5x^2 + 5y^2 + 4xy = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Se arata usor ca valorile proprii sunt $\lambda_1 = 7$ si $\lambda_2 = 3$. Un vector propriu corespunzator lui λ_1 este $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ si unul corespunzator lui λ_2 este $v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Pentru a putea aplica teorema de extrem conditionat trebuie sa normalizam acesti vectori:

$$u_1 = \frac{v_1}{\|v_1\|} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \qquad u_2 = \frac{v_2}{\|v_2\|} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

In concluzie:
Maximul conditionat are valoarea 7 si se obtine in punctul $(x, y) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$
Minimul conditionat are valoarea 3 se se obtine in punctul $(x, y) = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$

Probleme propuse

A. Consolidare cunostinte

Problema A.1. *Aratati ca daca $A \in M_{m\times n}(\mathbb{R})$ este o matrice oarecare, atunci A^tA are valori proprii reale si pozitive.*

Problema A.2. *Pentru orice transformare liniara $T : \mathbb{R}^n \to \mathbb{R}^p$ putem schimba sistemul de coordonate in $\mathbb{R}^n$ si $\mathbb{R}^p$ astfel incat relativ la noile sisteme de coordonate T sa fie o scalare neuniforma, cu factori posibil nuli.*

Indiciu: teorema generala de descompunere a valorilor singulare

B. Tehnica de calcul

Problema B.1. *Aflati descompunerea valorilor singulare pentru matricele:*

$$A = \begin{pmatrix} 1 & 0 \\ -2 & -2 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & 1 & 1 \\ \sqrt{2} & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Problema B.2. *Sa se determine o baza ortonormata in care transformarea liniara $T : \mathbb{R}^3 \to \mathbb{R}^3$ de matrice $A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{-1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}$ are forma diagonala.*

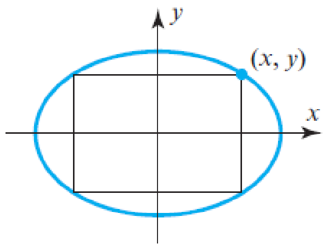
Problema B.3. *Aflati punctele critice ale functiei:*

$$f(x, y) = \frac{x^3}{3} + xy^2 - 8xy + 3$$

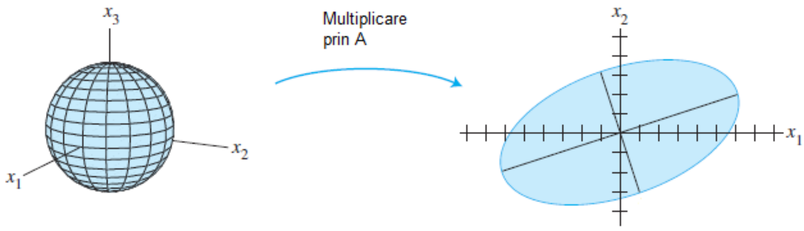
si folositi valorile proprii ale Hessianei pentru a decide care dintre ele sunt puncte de minim local, maxim local sau puncte sa.

C. Probleme cu caracter practic-aplicativ

Problema C.1. *Un dreptunghi este inscris in elipsa $4x^2 + 9y^2 = 36$ ca in figura. Folositi teoria formelor patratice pentru a afla valorile pozitive ale lui x si y pentru care dreptunghiul inscris are aria maxima.*



Problema C.2. *Fie matricea* $A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$, *atunci aplicatia liniara indusa* $\bar{\boldsymbol{x}} \to A\bar{\boldsymbol{x}}$ *transforma sfera unitate* $\{\bar{\boldsymbol{x}} \in \mathbb{R}^3 : \|\bar{\boldsymbol{x}}\| = 1\}$ *din* $\mathbb{R}^3$ *intr-o elipsa din* $\mathbb{R}^2$*, la fel ca in figura. Aflati un vector unitate* $\bar{\boldsymbol{x}}$ *pentru care lungimea* $\|A\bar{\boldsymbol{x}}\|$ *este maxima si calculati aceasta lungime.*



Bibliografie

[1] H. Anton, C. Rorres. *Elementary Linear Algebra. Applications Version* Ed. Wiley, 2014.

[2] O. Bundau, A. Juratoni *Exercitii si probleme de algebra liniara*, Ed. Politehnica, 2012.

[3] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000.

[4] D. Lay. *Linear Algebra and its Applications*, Ed. Addison-Wesley, 2012.