Cursul 11

Valori şi vectori proprii pentru matrice simetrice. Forme pătratice. Descompunerea SVD a unei matrice

11.1 Valori şi vectori proprii pentru matrice simetrice. Descompunerea ortogonală a unei matrice simetrice

Reamintim că o matrice pătratică $M \in \mathbb{R}^{n\times n}$ cu proprietatea că $M^T = M$ se numeşte *matrice simetrică*.

În machine learning matricele simetrice studiate sunt cel mai adesea matrice obţinute dintr-o matrice de date $A \in \mathbb{R}^{m\times n}$ care stochează pe coloane datele pentru n entităţi. Fiecare entitate are m caracteristici, numite atribute: $A = [X_1|X_2|\dots|X_n]$.

De exemplu, în diagnosticarea inteligentă sau în studiul eficacităţii unor medicamente în tratamentul unei boli, entităţile sunt n persoane. Pentru fiecare individ se înregistrează valorile pentru un set de m analize medicale (atribute ale indivizilor). Dacă după constituirea matricei de date se calculează versorul fiecărei coloane şi se notează cu B matricea $B = [X_1^0|X_2^0|\dots|X_n^0]$, atunci matricea simetrică $M = B^T B$ are ca elemente $M_{ij} =< X_i^0, X_j^0 >$. Cu alte cuvinte, un element M_{ij} indică similaritatea dintre indivizii i şi j. Pe de alta parte, elementele N_{ij} ale matricei $N = BB^T$ indică similaritatea dintre atribute. Informaţia importantă codificată de matricea de date A se extrage din valorile proprii şi vectorii proprii corespunzători ai matricei simetrice M, respectiv N.

Să studiem particularităţile matricelor simetrice comparativ cu matricele pătratice generale, nesimetrice.

În continuare vom presupune că $\mathbb{R}^n$ este înzestrat cu produsul scalar standard $< \cdot, \cdot >$ şi că elementele sale sunt vectori coloană. De asemenea, vom interpreta produsul Av, dintre o matrice pătratică $A \in \mathbb{R}^{n\times n}$ şi un vector coloană $v \in \mathbb{R}^n$, ca fiind vectorul w ce reprezintă efectul unui operator liniar $L : \mathbb{R}^n \to \mathbb{R}^n$, a cărui matrice în baza canonică este A, asupra vectorului v. Deci, în loc de $L(v)$ scriem Av.

Propoziţia 11.1.1 Dacă $A \in \mathbb{R}^{n\times n}$ este o matrice pătratică, atunci are loc următoarea relaţie:

$$\boxed{< Av, w >=< v, A^T w >, \ \forall\, v, w \in \mathbb{R}^n. \qquad (11.1)}$$

Demonstraţie: Exprimând produsul scalar standard astfel $< x, y >= x^T y$, membrul stâng al egalităţii ce dorim să o demonstrăm devine

$$< Av, w >= (Av)^T w = v^T A^T w,$$

iar cel drept este

$$< v, A^T w >= v^T (A^T w) = v^T A^T w,$$

deci $< Av, w >=< v, A^T w >$. □

Observaţia 11.1.1 Dacă A este o matrice simetrică, atunci, din $A^T = A$ şi relaţia (11.1), rezultă

$$< Av, w >=< v, Aw >, \ \forall \ v, w \in \mathbb{R}^n. \tag{11.2}$$

Să enunţăm particularităţile valorilor şi vectorilor proprii pentru matrice simetrice:

Propoziţia 11.1.2 **Polinomul caracteristic al unei matrice simetrice $A \in \mathbb{R}^{n\times n}$ are toate cele n rădăcini reale, adică o matrice simetrică are n valori proprii. În plus, dimensiunea fiecărui subspaţiu propriu S_λ coincide cu ordinul de multiplicitate m_λ al valorii proprii corespunzătoare, adică $dim(S_\lambda) = m_\lambda$ pentru orice valoare proprie λ.**

Propoziţia 11.1.3 **Pentru orice matrice simetrică $A \in \mathbb{R}^{n\times n}$, la valori proprii distincte corespund vectori proprii ortogonali.**

Demonstraţie: Fie $\lambda_1 \neq \lambda_2$ două valori proprii distincte ale matricei simetrice A şi $v_1 \in S_{\lambda_1}$, $v_2 \in S_{\lambda_2}$ doi vectori proprii corespunzători, adică $Av_1 = \lambda_1 v_1$, $Av_2 = \lambda_2 v_2$. Din proprietatea (11.2) a matricelor simetrice, avem că

$$< Av_1, v_2 >=< v_1, Av_2 >,$$

ceea ce este echivalent cu

$$< \lambda_1 v_1, v_2 >=< v_1, \lambda_2 v_2 >$$

sau

$$\lambda_1 < v_1, v_2 >= \lambda_2 < v_1, v_2 > \quad \Leftrightarrow \quad (\lambda_1 - \lambda_2) < v_1, v_2 >= 0.$$

Cum $\lambda_1 \neq \lambda_2$, rezultă că $\lambda_1 - \lambda_2 \neq 0$, deci $< v_1, v_2 >= 0$, adică $v_1 \perp v_2$, căci v_1, v_2 sunt vectori proprii, deci **nenuli**.

□

Fie $A \in \mathbb{R}^{n\times n}$ o matrice simetrică.

• Polinomul caracteristic al lui A, $P_n(\lambda) = \det(A - \lambda I_n)$, având n valori proprii (simple sau multiple), admite descompunerea:

$$P_n(\lambda) = (-1)^n (\lambda - \lambda_1)^{k_1} (\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_s)^{k_s}, \quad k_1 + k_2 + \cdots + k_s = n,$$

unde λ_i este rădăcină de ordin $m_{\lambda_i} = k_i$, $i = \overline{1, s}$.

• Pentru fiecare valoare proprie λ se determină subspaţiul propriu corespunzător,

$$S_\lambda = \{v \in \mathbb{R}^n \mid Av = \lambda v \Leftrightarrow (A - \lambda I_n)v = 0\}.$$

• Deoarece dimensiunea subspaţiului propriu S_{λ_i} este egală cu ordinul de multipliciate, k_i, al lui λ_i, determinăm o bază arbitrară $\mathcal{B}_i$ în S_{λ_i} (care conţine k_i vectori) şi apoi o ortonormăm folosind procedeul Gram–Schmidt şi obţinem baza ortonormată $\mathcal{B}'_i$, $i = \overline{1, s}$.

• Dacă $\mathcal{B}'_1 = \underbrace{(u_1, u_2, \dots, u_{k_1})}_{k_1 \text{ vectori}}$, $\mathcal{B}'_2 = \underbrace{(u_{k_1+1}, \dots, u_{k_1+k_2})}_{k_2 \text{ vectori}}$, $\mathcal{B}'_s = \underbrace{(u_{n-k_s+1}, \dots, u_n)}_{k_s \text{ vectori}}$ sunt baze ortonormate în subspaţiile proprii $S_{\lambda_1}, S_{\lambda_2}, \dots, S_{\lambda_s}$, concatenând aceste baze ortonormate formate din vectori proprii ai matricei A, obţinem o bază ortonormată în $\mathbb{R}^n$,

$$\mathcal{B}' = (u_1, u_2, \dots, u_{k_1}, u_{k_1+1}, \dots, u_{k_1+k_2}, \dots, u_n),$$

deoarece vectorii din bazele $\mathcal{B}'_i$ sunt ortonormaţi şi pentru că la valori proprii distincte corespund vectori proprii ortogonali, adică orice vector din $\mathcal{B}'_i$ este ortogonal pe orice vector din $\mathcal{B}'_j$, $i \neq j$.

• Notăm cu $T_{\mathcal{B}\mathcal{B}'}$ matricea de trecere de la baza canonică $\mathcal{B} = B_c$ din $\mathbb{R}^n$ la baza ortonormată $\mathcal{B}'$ formată din vectori proprii ai matricei A. Aceasta este o matrice ortogonală, fiind matricea de trecere dintre două baze ortonormate, deci $T^{-1}_{\mathcal{B}\mathcal{B}'} = T^T_{\mathcal{B}\mathcal{B}'}$ (reamintim că inversa unei matrice ortogonale este chiar transpusa sa).

• Notând cu D matricea diagonală a valorilor proprii,

$$D = \text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{k_1 \text{ ori}}, \underbrace{\lambda_2, \dots, \lambda_2}_{k_2 \text{ ori}}, \dots, \underbrace{\lambda_s, \dots, \lambda_s}_{k_s \text{ ori}}),$$

rezultă că matricea simetrică A este similară cu matricea diagonală D şi, în plus, matricea T din relaţia de similaritate este matricea ortogonală $T_{\mathcal{B}\mathcal{B}'}$, adică are loc relaţia:

$$A = T_{\mathcal{B}\mathcal{B}'} D T^{-1}_{\mathcal{B}\mathcal{B}'} = T_{\mathcal{B}\mathcal{B}'} D T^T_{\mathcal{B}\mathcal{B}'}.$$

Astfel, suntem conduşi la unul din cele mai importante rezultate aplicative din algebra liniară şi anume:

Propoziţia 11.1.4 Dacă $A \in \mathbb{R}^{n \times n}$ este o matrice simetrică ce are valorile proprii λ_i cu ordinele de multipliciate k_i, $i = \overline{1, s}$, $k_1 + k_2 + \cdots + k_s = n$, atunci există o bază ortonormată $\mathcal{B}'$ în $\mathbb{R}^n$ formată din vectori proprii ai lui A şi notând cu Q matricea de trecere de la baza canonică la baza $\mathcal{B}'$, matricea A este similară cu matricea

$$D = \text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{k_1 \text{ ori}}, \dots, \underbrace{\lambda_s, \dots, \lambda_s}_{k_s \text{ ori}}), \tag{11.3}$$

iar relaţia de similaritate este

$$A = QDQ^T. \tag{11.4}$$

Definiţia 11.1.1 Descompunerea unei matrice simetrice A în forma $A = QDQ^T$, unde D este matricea diagonală a valorilor proprii ale lui A şi $Q = T_{\mathcal{B}\mathcal{B}'}$ este o matrice ortogonală, se numeşte **descompunere ortogonală.**

Să ilustrăm această descompunere printr-un exemplu:

Exemplul 1. Să se determine valorile proprii şi subspaţiile proprii pentru matricea simetrică

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}.$$

Să se determine apoi câte o bază ortonormată în fiecare subspaţiu propriu al lui A şi o bază ortonormată în $\mathbb{R}^3$ formată din vectori proprii ai lui A. Să se scrie descompunerea ortogonală a matricei A.

- Valorile proprii ale lui A sunt $\lambda_{1,2} = 1, \lambda_3 = 7$;
- Subspaţiile proprii corespunzătoare:

$$S_{\lambda=1} = \{v = (x, y, z)^T \ : \ Av = 1v \Leftrightarrow (A - 1I_3)v = 0\}.$$

Avem

$$(A - 1I_3)v = 0 \iff \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Rezolvând sistemul de mai sus, obţinem

$$S_{\lambda=1} = \{v = (-\alpha - \beta, \alpha, \beta)^T = \alpha(-1, 1, 0)^T + \beta(-1, 0, 1)^T \ : \ \alpha, \beta \in \mathbb{R}\}.$$

Bază în acest subspaţiu este $\mathcal{B}_1 = (v_1 = (-1, 1, 0)^T, v_2 = (-1, 0, 1)^T)$. Evident că baza $\mathcal{B}_1$ nu este ortonormată. Aplicând procedeul Gram–Schmidt, obţinem baza

$$\mathcal{B}_1' = \left(q_1 = \frac{1}{\sqrt{2}}(-1, 1, 0)^T, \, q_2 = \frac{1}{\sqrt{6}}(-1, -1, 2)^T \right).$$

Pentru a determina subspaţiul propriu $S_{\lambda=7}$ determinăm soluţiile sistemului

$$(A - 7I_3)v = 0 \iff \begin{bmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Alegem drept determinant principal determinantul constituit din elementele de intersecţie ale liniilor 1,2 cu coloanele 1,2. Astfel, $z = \alpha$ este necunoscută secundară şi obţinem

$$S_{\lambda=7} = \{v = \alpha(1, 1, 1)^T \ : \ \alpha \in \mathbb{R}\}.$$

O baza ortonormată în $S_{\lambda=7}$ este

$$\mathcal{B}_2' = \left(q_3 = \frac{1}{\sqrt{3}}(1, 1, 1)^T \right),$$

iar $\mathcal{B}' = \mathcal{B}_1' \cup \mathcal{B}_2' = (q_1, q_2, q_3)$ este o bază ortonormată în $\mathbb{R}^3$ formată din vectori proprii ai matricei A. Notând

$$Q = [q_1|q_2|q_3] = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix},$$

obţinem descompunerea ortogonală $A = QDQ^T$, unde $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}$.

11.2 Forme pătratice

În inteligenţa artificială deseori într-o etapă a unui algoritm trebuie determinate punctele în care o funcţie $f : D \subset \mathbb{R}^n \to \mathbb{R}$ de clasă C^2 ia valoarea minimă sau maximă. Problema aflării punctelor de minim sau maxim se numeşte problemă de optimizare şi se notează astfel:

$$\text{argmin} f(x), x \in D, \quad \text{argmax} f(x), x \in D$$

şi se citeşte: să se determine argumentul $x \in D$ care minimizează funcţia f sau să se determine argumentul $x \in D$ care maximizează funcţia f.

Exemple de probleme de optimizare: minimizarea erorii în clasificare sau determinarea drumului de lungime minimă pe care trebuie să îl parcurgă un agent inteligent pentru a deservi n noduri/puncte de lucru etc.

Decizia dacă un punct x_0 din D este punct de minim sau maxim pentru o funcţie f se ia analizând o formă pătratică asociată funcţiei f şi anume $\varphi(v) = d^2_{x_0} f(v, v)$, unde $d^2_{x_0} f$ este diferenţiala de ordinul 2 a funcţiei f în punctul x_0.

Considerăm spaţiul vectorial $\mathbb{R}^n$ înzestrat cu produsul scalar standard $< v, w >= v^T w$ şi $\mathcal{B}$ o bază ortonormată în $\mathbb{R}^n$ (de obicei, baza canonică; vom face convenţia că dacă nu este precizat în context cine este baza $\mathcal{B}$, aceasta va fi considerată baza canonică).

Definiţia 11.2.1 Fie $A = (a_{ij})_{i,j=\overline{1,n}}$ o matrice simetrică. Aplicaţia $q : \mathbb{R}^n \to \mathbb{R}$, definită prin

$$q(v_{\mathcal{B}}) =< v_{\mathcal{B}}, Av_{\mathcal{B}} >= v_{\mathcal{B}}^T A v_{\mathcal{B}}, \tag{11.5}$$

se numeşte *formă pătratică*.

Dacă un vector arbitrar $v \in \mathbb{R}^n$ are relativ la baza $\mathcal{B}$ coordonatele $v_{\mathcal{B}} = (x_1, x_2, \ldots, x_n)^T$, atunci expresia analitică a formei pătratice în această bază este:

$$q(v_{\mathcal{B}}) = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \tag{11.6}$$

Efectuând produsele, obţinem

$$q(v_{\mathcal{B}}) = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + \cdots + 2a_{n-1\,n}x_{n-1}x_n.$$

Expresia precedentă ilustrează de ce funcţia q se numeşte formă pătratică: expresia ei este o sumă de termeni de gradul 2 în $x_1, x_2, \ldots, x_n$, adică ceea ce se numeşte **polinom omogen de gradul 2**.

Exemplul 2. Formă pătratică q determinată de matricea simetrică $A = \begin{bmatrix} -2 & 1 \\ 1 & 3 \end{bmatrix}$ este

$$q(x_1, x_2)^T \stackrel{\text{not.}}{=} q(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -2x_1^2 + 2x_1x_2 + 3x_2^2.$$

În cele ce urmează în loc de $q(x_1, x_2, \ldots, x_n)^T$ vom scrie simplu $q(x_1, x_2, \ldots, x_n)$.

Dacă cunoaştem expresia analitică a unei forme pătratice $q : \mathbb{R}^n \to \mathbb{R}$, adică un polinom omogen de gradul 2 în $x_1, x_2, \ldots, x_n$, matricea simetrică ce o defineşte se determină astfel:

• coeficienţii termenilor $x_1^2, x_2^2, \ldots, x_n^2$ reprezintă elementele $a_{11}, a_{22}, \ldots, a_{nn}$ ale matricei simetrice A, adică elementele de pe diagonala principală a lui A;

• coeficienţii produselor x_ix_j împărţiti la 2 sunt elementele a_{ij} şi a_{ji} ale matricei A, pentru $i, j = \overline{1, n}$, $i \neq j$.

Exemplul 3. Se dă forma pătratică $q : \mathbb{R}^3 \to \mathbb{R}$, definită prin

$$q(x_1, x_2, x_3) = 2x_1^2 + 3x_1x_2 + 6x_1x_3 - 5x_2^2 - 8x_2x_3 + x_3^2.$$

Matricea simetrică asociată este

$$A = \begin{bmatrix} 2 & \frac{3}{2} & 3 \\ \frac{3}{2} & -5 & -4 \\ 3 & -4 & 1 \end{bmatrix}.$$

O formă pătratică ia valori reale care pot fi pozitive, negative sau zero. Se observă că $q(0) = 0$, deci o formă pătratică aplică vectorul nul $0 = (0, 0, \ldots, 0)$ în numărul real 0.

• Forma pătratică $q : \mathbb{R}^n \to \mathbb{R}$ care ia valori pozitive, $q(v) > 0$, oricare ar fi vectorul $v \in \mathbb{R}^n \setminus \{0\}$, se numeşte **formă pătratică pozitiv definită**.

• Dacă $q(v) \geq 0$, pentru orice $v \in \mathbb{R}^n$, atunci q se numeşte **formă semipozitiv definită** (mai precis, în acest caz q poate lua valoarea zero şi pentru vectori nenuli).

• Dacă $q(v) < 0$, oricare ar fi $v \in \mathbb{R}^n \setminus \{0\}$, atunci q se numeşte **formă negativ definită**, iar dacă $q(v) \leq 0$, $\forall\ v \in \mathbb{R}^n$, forma q se numeşte **seminegativ definită**.

• Dacă pentru anumiţi vectori q ia valori pozitive, iar pe alţii valori negative, atunci q se numeşte **formă pătratică nedefinită**.

Analizând expresia analitică a formei pătratice din Exemplul 3 este greu să ne pronunţăm dacă ea este pozitiv definită, negativ definită sau nedefinită. Este însă foarte simplu să indicăm tipul formei pătratice dacă ea conţine doar termeni în x_i^2, $i = \overline{1, n}$.

Exemplul 4. Forma pătratică $q(x_1, x_2, x_3) = -3x_1^2 - x_2^2 - 4x_3^2$ este negativ definită, forma pătratică $q(x_1, x_2, x_3) = 2x_1^2 - x_2^2 + 6x_3^2$ este nedefinită, deoarece $q(1, 0, 1) = 8 > 0$ şi $q(0, 1, 0) = -1 < 0$, iar forma $q(x_1, x_2, x_3) = x_1^2 + x_2^2$ este semipozitiv definită (ea nu este pozitiv definită, căci ia valoarea 0 şi pentru vectori nenuli, de exemplu $q(0, 0, 1) = 0$).

Observăm că putem deduce rapid tipul unei forme pătratice dacă ea este definită de o matrice diagonală, care evident este simetrică:

$$\begin{aligned} q(x_1, x_2, \ldots, x_n) &= [x_1\, x_2\, \ldots\, x_n] \begin{bmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= d_1x_1^2 + d_2x_2^2 + \cdots + d_nx_n^2, \quad d_i \in \mathbb{R}. \end{aligned}$$

O formă pătratică a cărei matrice de definiţie este diagonală se zice că este în *forma canonică*.

Să exploatăm faptul că orice matrice simetrică A este similară cu o matrice diagonală, adică există $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$, ce sunt valorile proprii ale lui A, şi o matrice inversabilă $T = T_{\mathcal{B}\mathcal{B}'}$, ce este matricea de trecere de la baza canonică $\mathcal{B}$ la baza ortonormată $\mathcal{B}'$, formată din vectori proprii ai lui A, astfel încât $A = TDT^{-1} = T_{\mathcal{B}\mathcal{B}'}DT_{\mathcal{B}\mathcal{B}'}^T$, unde $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$.

Relaţia de mai sus este echivalentă cu

$$D = T_{\mathcal{B}\mathcal{B}'}^T A T_{\mathcal{B}\mathcal{B}'}. \tag{11.7}$$

Pe de altă parte, relaţia dintre coordonatele unui vector $v \in \mathbb{R}^n$ relativ la cele două baze este $v_{\mathcal{B}} = T_{\mathcal{B}\mathcal{B}'}v_{\mathcal{B}'}$. Să deducem expresia analitică a formei pătratice q definită de matricea A relativ la baza $\mathcal{B}'$, formată din vectori proprii ai lui A. Pentru aceasta notăm cu $X_1, X_2, \ldots, X_n$, coordonatele vectorului arbitrar v relativ la baza $\mathcal{B}'$, adică $v_{\mathcal{B}'} = (X_1, X_2, \ldots, X_n)^T$. Avem

$$\begin{aligned} q(v_{\mathcal{B}}) &= < v_{\mathcal{B}}, Av_{\mathcal{B}} > = < T_{\mathcal{B}\mathcal{B}'}v_{\mathcal{B}'}, AT_{\mathcal{B}\mathcal{B}'}v_{\mathcal{B}'} > \overset{(11.1)}{=} < v_{\mathcal{B}'}, T_{\mathcal{B}\mathcal{B}'}^T AT_{\mathcal{B}\mathcal{B}'}v_{\mathcal{B}'} > \overset{(11.7)}{=} < v_{\mathcal{B}'}, Dv_{\mathcal{B}'} > \\ &= \begin{bmatrix} X_1 & X_2 & \ldots & X_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \\ &= \lambda_1X_1^2 + \lambda_2X_2^2 + \cdots + \lambda_nX_n^2. \end{aligned}$$

Am arătat astfel că o formă pătratică definită de o matrice simetrică A asociază unui vector exprimat în baza $\mathcal{B}$ aceeaşi valoare ca valoarea asociată de forma pătratică definită de matricea diagonală D a valorilor proprii ale lui A aceluiaşi vector, dar exprimat în baza $\mathcal{B}'$.

În concluzie, pentru a decide tipul formei pătratice care relativ la baza ortonormată iniţială $\mathcal{B}$ are matricea simetrică A, se determină valorile proprii ale lui A.

- **Dacă toate valorile proprii sunt pozitive ($\lambda_i > 0, \forall\ i = \overline{1,n}$), atunci forma pătratică este pozitiv definită;**
- **Dacă toate valorile proprii sunt mai mari sau egale cu 0, atunci forma pătratică este semipozitiv definită;**
- **Dacă $\lambda_i < 0, \forall\ i = \overline{1,n}$, atunci forma este negativ definită, respectiv seminegativ definită dacă $\lambda_i \leq 0, \forall\ i = \overline{1,n}$;**
- **Dacă cel puţin o valoare proprie este pozitivă şi alta negativă, atunci forma pătratică este nedefinită.**

Exemplul 5. Se dă forma pătratică $q : \mathbb{R}^2 \to \mathbb{R}$, $q(x_1, x_2) = 4x_1^2 - 4x_1x_2 + x_2^2$. Să se determine matricea formei pătratice, valorile ei proprii şi să se precizeze tipul formei.

Matricea formei pătratice este

$$A = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}.$$

Valorile proprii ale lui A sunt $\lambda_1 = 0, \lambda_2 = 5$, deci forma pătratică este semipozitiv definită.

În analiza matematică, unei funcţii $f : D \subset \mathbb{R}^n \to \mathbb{R}$, de clasă C^2 pe D, i se asociază matricea simetrică Hess$(f)(x_0)$ a derivatelor parţiale de ordinul 2 într-un punct critic x_0, numită Hessiana funcţiei f în x_0. Elementele a_{ij} ale acestei matrice sunt

$$a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0), \text{ pentru } i, j = \overline{1,n}.$$

Dacă x_0 este un punct critic al funcţiei f, adică

$$\frac{\partial f}{\partial x_i}(x_0) = 0, \forall\ i = \overline{1,n},$$

atunci tipul formei pătratice având ca matrice Hessiana funcţiei f în punctul x_0 indică dacă punctul x_0 este punct de maxim, punct de minim sau punct şa (este punct critic, dar nu este punct de extrem).

În Fig.11.1 este ilustrat graficul unei forme pătratice

$$q : \mathbb{R}^2 \to \mathbb{R},\ q(x_1, x_2) = \lambda_1 x_1^2 + \lambda_2 x_2^2$$

pentru cazul q – pozitiv definită ($\lambda_1, \lambda_2 > 0$), negativ definită ($\lambda_1, \lambda_2 < 0$), respectiv nedefinită ($\lambda_1 > 0, \lambda_2 < 0$).

Observăm că în primul caz $(0, 0)$ este punct de minim pentru că $q(x_1, x_2) > 0$ pentru orice $v = (x_1, x_2)^T \neq (0, 0)^T$, în al doilea caz este punct de maxim şi în al treilea este punct şa. O funcţie $f : D \subset \mathbb{R}^2 \to \mathbb{R}$, de clasă C^r pe D, $r \geq 2$, este aproximată în vecinătatea unui punct critic x_0 de o astfel de formă pătratică şi, deci, tipul extremal al punctului critic depinde de tipul punctului $(0, 0)$ pentru forma pătratică asociată.

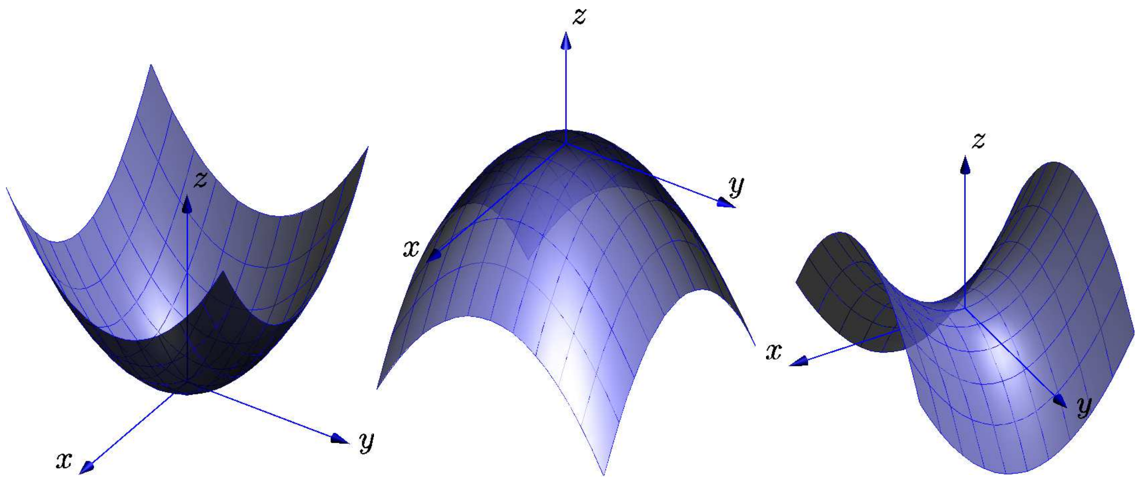

Fig.11.1: Graficele a trei forme pătratice aduse la forma canonică $q(x_1, x_2) = \lambda_1 x_1^2 + \lambda_2 x_2^2$: în stânga forma este pozitiv definită, în centru forma este negativ definită şi în dreapta forma este nedefinită.

11.3 Descompunerea singulară a unei matrice

Necesitatea de a minimiza volumul de informaţie ce trebuie să fie stocată sau transmisă printr-un canal de comunicaţie a condus la dezvoltarea unor metode de reducere a dimensiunii acestora (comprimarea datelor). Una din metodele de comprimare a datelor oferite de algebra liniară se bazează pe descompunerea singulară a unei matrice $A \in \mathbb{R}^{m\times n}$.

11.3.1 Noţiuni şi rezultate preliminare

Definiţia 11.3.1 O matrice simetrică $A \in \mathbb{R}^{n\times n}$ pentru care forma pătratică asociată q este pozitiv definită, adică

$$q(v) =< v, Av > > 0, \forall v \in \mathbb{R}^n \setminus \{0\}, \tag{11.8}$$

se numeşte *matrice pozitiv definită*, iar dacă este verificată relaţia

$$q(v) =< v, Av > \geq 0, \forall\ v \in \mathbb{R}^n, \tag{11.9}$$

se numeşte *matrice simetrică semipozitiv definită*.

Propoziţia 11.3.1 *O matrice simetrică este semipozitiv definită dacă şi numai dacă are toate valorile proprii mai mari sau egale cu zero.*

Demonstraţie: Fie $A \in \mathbb{R}^{n\times n}$ o matrice simetrică semipozitiv definită. Fiind simetrică, A are n valori proprii (distincte sau nu). Fie $\lambda \in \mathbb{R}$ o valoare proprie şi v un vector propriu corespunzător, adică $Av = \lambda v$. Au loc următoarele echivalenţe:

$$< v, Av > \geq 0 \quad \Leftrightarrow \quad < v, \lambda v > \geq 0 \quad \Leftrightarrow \quad \lambda < v, v > \geq 0.$$

v, fiind vector propriu, este nenul, deci $< v, v > > 0$. Astfel, $\lambda < v, v > \geq 0$ dacă şi numai dacă $\lambda \geq 0$. $\square$

Considerăm o matrice arbitrară cu m linii şi n coloane, $A \in \mathbb{R}^{m\times n}$. Matricele asociate A^TA, AA^T sunt matrice pătratice de tip $n\times n$, respectiv $m\times m$ şi simetrice, deoarece coincid cu transpusele lor. De exemplu, $(A^TA)^T = A^T(A^T)^T = A^TA$.

Propoziţia 11.3.2 *Dacă $A \in \mathbb{R}^{m\times n}$ este o matrice de tip $m\times n$, atunci matricele simetrice asociate A^TA, AA^T sunt semipozitiv definite.*

Demonstraţie: Deoarece produsul scalar al unui vector cu el însuşi este mai mare sau egal cu zero, avem că pentru orice vector $v \in \mathbb{R}^n$: $< Av, Av >\geq 0$. Dar

$$< Av, Av > \stackrel{(11.1)}{=} < v, A^TAv >,$$

ceea ce implică $< v, A^TAv > \geq 0, \forall\, v \in \mathbb{R}^n$, adică matricea A^TA este semipozitiv definită. Analog, $< A^Tv, A^Tv > \geq 0$ şi din

$$< A^Tv, A^Tv > \stackrel{(11.1)}{=} < v, AA^Tv >, \forall\, v \in \mathbb{R}^m$$

rezultă că şi matricea AA^T este semipozitiv definită. □

Propoziţia 11.3.3 *Rangul matricei A^TA coincide cu rangul matricei $A \in \mathbb{R}^{m\times n}$.*

11.4 Calculul descompunerii singulare a unei matrice

Propoziţia 11.4.1 (**Descompunerea singulară**) Pentru orice matrice $A \in \mathbb{R}^{m\times n}$ de rang r există două matrice ortogonale $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ şi numerele reale pozitive

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

astfel încât A se descompune în produsul $A = U\Sigma V^T$, adică

$$\underbrace{A}_{m\times n} = \underbrace{U}_{m\times m} \underbrace{\begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \sigma_r & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix}}_{m\times n} \underbrace{V^T}_{n\times n}. \tag{11.10}$$

Definiţia 11.4.1 Descompunerea de mai sus se numeşte **descompunerea singulară** a matricei A (**singular value decomposition, SVD**). Valorile pozitive $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ din matricea Σ se numesc **valorile singulare** ale matricei A, vectorii u_i, **vectori singulari stângi**, iar vectorii v_i, **vectori singulari drepţi**, $i = \overline{1, r}$.

Pentru a interpreta descompunerea SVD şi pentru a prezenta aplicaţii ale ei, definim câteva noţiuni şi rezultate de calcul matriceal:

• Produsul exterior a doi vectori $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$ este o matrice de tip $m \times n$ obţinută înmulţind vectorul coloană u cu vectorul linie v^T, adică uv^T.

Dacă $u = (x_1, x_2, \ldots, x_m)^T$ şi $v = (y_1, y_2, \ldots, y_n)^T$, atunci produsul lor exterior este

$$uv^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \ldots & y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 & x_1y_2 & \ldots & x_1y_n \\ x_2y_1 & x_2y_2 & \ldots & x_2y_n \\ \vdots & \vdots & \ldots & \vdots \\ x_my_1 & x_my_2 & \ldots & x_my_n \end{bmatrix}.$$

Prin urmare, matricea produs exterior, $B = uv^T$, are elementele $b_{ij} = x_iy_j$, $i = \overline{1,m}$, $j = \overline{1,n}$.

Dacă vectorii u, v sunt nenuli, atunci coloanele matricei uv^T sunt proporţionale, deci rangul matricei B este 1.

Propoziţia 11.4.2 Produsul $U\Sigma V^T$ din descompunerea SVD a unei matrice $A \in \mathbb{R}^{m\times n}$ de rang r este egal cu

$$U\Sigma V^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_1^T + \cdots + \sigma_r u_1 v_r^T.$$

Prin urmare, ideea de bază a descompunerii SVD a unei matrice $A \in \mathbb{R}^{m\times n}$ de rang r este că matricea A se poate descompune ca o combinaţie liniară cu coeficienţi pozitivi şi descrescători a r matrice de rangul 1, $M_j = u_j v_j^T$, $j = \overline{1,r}$:

$$A = U\Sigma V^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T.$$

Aproximarea de rang k a matricei A: Dacă în descompunerea SVD a unei matrice ultimele valori singulare sunt mici (apropiate de zero), atunci, renunţând la termenii ce le conţin, obţinem o aproximare a matricei A:

$$A_k = \sigma_1(u_1 v_1^T) + \cdots + \sigma_k(u_k v_k^T).$$

Matricea aproximare A_k se factorizează astfel:

$$A_k = \underbrace{[u_1|u_2|\ldots|u_k]}_{m\times k} \underbrace{\begin{bmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & \sigma_k \end{bmatrix}}_{k\times k} \underbrace{\begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_k^T \end{bmatrix}}_{k\times n} = U_k \Sigma_k V_k^T$$

şi are rangul k.

Ori de câte ori se aproximează un element al unei mulţimi înzestrate cu o distanţă (o astfel de mulţime se numeşte *spaţiu metric*) cu alt element al aceleaşi mulţimi, ne interesează "cât de bună este acea aproximare", evaluând distanţa dintre element şi aproximantul său.

În mulţimea matricelor din $\mathbb{R}^{m\times n}$ se definesc diferite norme şi atunci $dist(A, B) = \|A-B\|$. O normă este cea definită de produsul scalar a două matrice:

$$< A, B >= \text{trace}(A^T B)$$

şi anume

$$\|A\| = \sqrt{< A, A >} = \sqrt{\text{trace}(A^T A)}.$$

Această normă se numeşte **norma Frobenius** a unei matrice şi pentru a o distinge de alte norme se notează $\|A\|_F$.

Teorema 11.4.1 (Teorema Eckart) Fie matricea $A \in \mathbb{R}^{m\times n}$ şi $A_k \in \mathbb{R}^{m\times n}$ aproximarea sa de rang k. Dintre toate matricele $B \in \mathbb{R}^{m\times n}$ de rang k, distanţa de la A la B este minimă pentru $B = A_k$, adică

$$\min_{B\,|\,\text{rang}(B)=k} \|A - B\|_F = \|A - A_k\|_F = \sqrt{\sigma_{k+1}^2 + \cdots + \sigma_r^2}.$$

Cu alte cuvinte, A_k este cea mai bună aproximare de rang k a matricei A. Această proprietate se exploatează în numeroase domenii din Computer Science, printre care: comprimarea datelor, în general, şi a imaginilor, în particular, information retrieval, machine learning.

Dacă A este o matrice imagine ai cărei pixeli au diverse nivele de gri, între negru şi alb, un element al matricei A fiind codul nivelului de gri $c \in \{0, 1, 2, \dots, 255\}$ sau normalizat $c \in [0, 1]$ (depinde de tipul de imagine şi limbajul de programare care o citeşte; de exemplu, în Python imaginile în nivele de gri din mulţimea $\{0, 1, 2, \dots, 255\}$ sunt convertite la citire în imagini cu nivelul de gri în $[0, 1]$, adică dacă codul pentru gri este $c = 135$, el este convertit în $135/256$), atunci, determinând descompunerea SVD a matricei $A = \sum_{i=1}^r \sigma_i u_i v_i^T$ şi renunţând la termenii ce au coeficienţii $\sigma_{k+1}, \dots, \sigma_r$ suficient de mici în comparaţie cu $\sigma_1, \sigma_2, \dots, \sigma_k$, obţinem o aproximare A_k a imaginii A. Aceasta este o modalitate de comprimare a imaginii în scopul stocării sau transmiterii ei pe un canal de comunicaţie, adică în locul transmiterii vectorilor $u_1, u_2, \dots, u_r$, $v_1, v_2, \dots, v_r$ şi, respectiv, a valorilor singulare $\sigma_1, \sigma_2, \dots, \sigma_r$ se transmit doar vectorii $u_1, \dots, u_k$, $v_1, \dots, v_k$ şi valorile singulare $\sigma_1, \sigma_2, \dots, \sigma_k$, $k < r$.

Descompunerea trunchiată A_k a unei imagini A filtrează o parte din zgomotul conţinut în imagine fără a pierde o informaţie semnificativă din ea.

11.5 Contrucţia matricelor U, V, Σ din descompunerea SVD

• **Construcţia matricei V**:

Matricea $A^T A$, fiind o matrice de tip $n \times n$ simetrică şi semipozitiv definită, are n valori proprii $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. Construind câte o bază ortonormată în fiecare subspaţiu propriu şi reunind aceste baze obţinem baza ortonormată în $\mathbb{R}^n$ formată din vectori proprii ai matricei $A^T A$, $\mathcal{B}'_n = (v_1, v_2, \dots, v_n)$. Notăm cu $V = [v_1|v_2|\dots|v_n]$ matricea de trecere de la baza canonică la baza $\mathcal{B}'_n$, care evident este matrice ortogonală. Astfel, matricea $A^T A$ este similară cu matricea diagonală a valorilor proprii:

$$M = A^T A = V \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) V^T.$$

Dar cum două matrice similare au aceelaşi rang, rezultă că

$$rang(\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)) = rang(A^T A).$$

Cum $rang(A^T A) = rang(A) = r$, matricea diagonală diag$(\lambda_1, \lambda_2, \ldots, \lambda_n)$ are rangul r, deci doar primele r valori proprii în ordinea descrescătoare sunt nenule:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 \text{ şi } \lambda_{r+1} = \cdots = \lambda_n = 0.$$

• **Construcţia matricei U**:
Vectorii proprii ortonormaţi $v_1, v_2, \ldots, v_r$ corespund la valori proprii λ_i nenule:

$$A^T A(v_i) = \lambda_i v_i, \ i = \overline{1, n}.$$

Considerăm vectorii din $\mathbb{R}^m$, $w_i = Av_i$, $i = \overline{1, r}$. Să arătăm că sistemul $w_1, w_2, \ldots, w_r$ este un sistem ortogonal de vectori:

$$< w_i, w_j >=< Av_i, Av_j >=< v_i, (A^T A)v_j >=< v_i, \lambda_j v_j >= \lambda_j < v_i, v_j >= \lambda_j \delta_{ij}.$$

Astfel, pentru $i \neq j$, $< w_i, w_j >= \lambda_j \cdot 0 = 0$, deci vectorii w_i, $i = \overline{1, r}$, sunt ortogonali. Pentru $i = j$ avem $< w_i, w_i >= \lambda_i$, ceea ce este echivalent cu $\|w_i\|^2 = \lambda_i$ sau echivalent

$$\|w_i\| = \sqrt{\lambda_i} = \sigma_i, i = \overline{1, r}.$$

Normând sistemul de vectori $(w_1, w_2, \ldots, w_r)$, obţinem sistemul ortonormat de r vectori din $\mathbb{R}^m$, $(u_1, u_2, \ldots, u_r)$, unde

$$u_i = \frac{w_i}{\|w_i\|} = \frac{w_i}{\sigma_i} = \frac{Av_i}{\sigma_i}, i = \overline{1, r}.$$

Prin urmare, are loc următoarea relaţie între vectorii ortonormaţi $v_1, v_2, \ldots, v_r$ din $\mathbb{R}^n$ şi vectorii ortonormaţi $u_1, u_2, \ldots, u_r$ din $\mathbb{R}^m$:

$$Av_i = \sigma_i u_i, \quad i = \overline{1, r}.$$

Sistemul de vectori ortonormaţi $(u_1, u_2, \ldots, u_r)$ îl completăm la o bază ortonormată

$$(u_1, u_2, \ldots, u_r, u_{r+1}, \ldots, u_m)$$

în $\mathbb{R}^m$ şi anume:

• Determinăm o bază arbitrară în $Null(A^T)$. Deoarece A^T are rangul lui A, adică r, rezultă că dimensiunea lui este $m - r$. Baza arbitrară $(t_1, t_2, \ldots, t_{m-r})$ se ortonormează cu metoda Gram-Schmidt şi obţinem din ea baza ortonormată $(u_{r+1}, u_{r+2}, \ldots, u_m)$ în $Null(A^T)$.

• Baza ortonormată $(u_{r+1}, \ldots, u_m)$ din $Null(A^T)$ completează sistemul $(u_1, u_2, \ldots, u_r)$ la o bază ortonormată în $\mathbb{R}^m$, $(u_1, \ldots, u_r, u_{r+1}, \ldots, u_m)$.

Notăm cu U matricea de trecere de la baza canonică din $\mathbb{R}^m$ la baza ortonormată găsită anterior, $(u_1, u_2, \ldots, u_m)$. Matricea $U = [u_1|u_2|\ldots|u_m]$ este chiar matricea ortogonală din descompunerea SVD a matricei A.

Etapizarea calculelor pentru determinarea descompunerii singulare a unei matrice $A \in \mathbb{R}^{m\times n}$

• se calculează produsul $M = A^T A$;

• se determină polinomul caracteristic al matricei simetrice semipozitiv definite $M = A^T A$, $P_n(\lambda) = \det(M - \lambda I_n)$, şi i se determină rădăcinile $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$;

• se determină câte o bază în fiecare subspaţiu propriu al matricei $M = A^T A$, care apoi se ortonormează folosind procedeul Gram-Schmidt; reuniunea acestor baze conduce la o bază ortonormată în $\mathbb{R}^n$ formată din vectori proprii, $(v_1, v_2, \ldots, v_n)$;

• se constituie matricea $V = [v_1|v_2|\ldots|v_n]$;

• se separă vectorii $v_1, v_2, \ldots, v_r$ ce corespund valorilor proprii nenule $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$ şi se calculează valorile singulare

$$\boxed{\sigma_i = \sqrt{\lambda_i},\ i = \overline{1, r};}$$

• se determină vectorii ortonormaţi $u_1, u_2, \ldots, u_r$ din $\mathbb{R}^m$ cu formula

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

• se determină o bază arbitrară în subspaţiul $Null(A^T)$ rezolvând sistemul liniar şi omogen $A^T x = 0$. Aplicând procedeul Gram-Schmidt, se ortonormează baza determinată, obţinând astfel $m - r$ vectori ortonormaţi $u_{r+1}, \ldots, u_m$, deoarece dimensiunea lui $Null(A^T)$ este egală cu $m - r$, adică cu numărul m de coloane ale matricei A^T minus $\mathrm{rang}(A^T) = \mathrm{rang}(A) = r$;

• se constituie matricea $U = [u_1|u_2|\ldots|u_m]$;

• se scrie descompunerea SVD a matricei A, $A = U\Sigma V^T$, ţinând seama că matricea Σ are aceleaşi dimensiuni ca şi A, adică este de tip $m \times n$.

Exemplul 6. Să se determine descompunearea singulară a matricei $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

• Calculăm

$$M = A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

• $P_2(\lambda) = det(M - \lambda I_2) = (2 - \lambda)^2 - 1$ şi $\lambda_1 = 3, \lambda_2 = 1$ (le-am ordonat descrescător).

• Determinăm câte o bază în fiecare subspaţiu propriu:

$$S_{\lambda=3} = \{v = \alpha(1, 1)^T, \alpha \in \mathbb{R}\},\ \mathcal{B}_1 = ((1, 1)^T),$$

$$S_{\lambda=1} = \{v = \beta(1, -1)^T, \beta \in \mathbb{R}\},\ \mathcal{B}_2 = ((1, -1)^T).$$

Vectorii celor două baze sunt ortogonali, deoarece corespund la valori proprii distincte ale unei matrice simetrice (A^TA). Îi normăm şi obţinem baza ortonormată în $\mathbb{R}^2$

$$\mathcal{B}' = (v_1 = (\sqrt{2}/2, \sqrt{2}/2)^T, v_2 = (\sqrt{2}/2, -\sqrt{2}/2)^T))$$

şi matricea

$$V = [v_1|v_2] = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}.$$

• Calculăm valorile singulare: $\sigma_1 = \sqrt{3}, \sigma_2 = \sqrt{1} = 1$.

• Determinăm coordonatele vectorilor $u_i = \dfrac{1}{\sigma_i} A v_i$, $i = 1, 2$, unde 2 este rangul matricei A^TA:

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

$$u_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}.$$

• Completăm sistemul ortonormat (u_1, u_2) la o bază ortonormată în $\mathbb{R}^3$. Teoretic, ar trebui să determinăm o bază în subspaţiul soluţiilor sistemului $A^Tx = 0$, adică

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Notând $x_2 = \gamma$, avem

$$\text{Null}(A^T) = \{v = \gamma(-1, 1, 1)^T \; : \; \gamma \in \mathbb{R}\},$$

deci $u_3 = (-1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})^T$.

În acest caz special am fi putut determina pe $u_3 = u_1 \times u_2$.

• Matricea U este

$$U = [u_1|u_2|u_3] = \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \end{bmatrix}.$$

• Descompunerea SVD a matricei A este

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

11.6 Probleme propuse

1. Să se determine factorizarea $A = QDQ^T$ a matricei simetrice $A = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$.

2. Matricea

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

admite valoarea proprie 1. Mai admite A şi alte valori proprii? Argumentaţi! Să se factorizeze matricea $A = QDQ^T$, unde Q este o matrice ortogonală şi D este o matrice diagonală.

3. Să se factorizeze matricea simetrică $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ în $A = QDQ^T$, unde Q este o matrice ortogonală, iar D este o matrice diagonală.

4. Determinaţi expresia analitică a formei pătratice $Q : \mathbb{R}^3 \to \mathbb{R}$, definită de matricea simetrică

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

apoi stabiliţi natura acestei forme pătratice.

5. Să se determine matricea formei pătratice $Q : \mathbb{R}^3 \to \mathbb{R}$, definită prin

$$Q(x_1, x_2, x_3) = 3x_1^2 + 8x_1x_2 - 8x_1x_3 - x_2^2 - 4x_2x_3 + x_3^2,$$

apoi stabiliţi natura acestei forme pătratice.

6. Să se aducă forma pătratică $Q(x_1, x_2) = -2\sqrt{2}x_1x_2 + x_2^2$ la forma canonică şi apoi să se stabilească natura ei, adică dacă Q este o formă pătratică pozitiv definită, negativ definită sau nedefinită.

7. Dacă matricea unei forme pătratice are valorile proprii $-1, 2, 0$, ce puteţi spune despre natura formei pătratice? Justificaţi!

8. Să determine descompunerea singulară a matricei $A = [2\ 1\ -2]$.

9. Să se determine descompunerea singulară a matricei $A = \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix}$ şi aproximaţia A_1 a acesteia.

10. Determinaţi descompunerea singulară a matricei $A = \begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix}$.