

# Dodatek A

Piotr Jacek Suchomski

## Przestrzeń liniowa. Podprzestrzeń. Przekształcenie liniowe

Niech  $K$  będzie ciałem liczbowym, a  $X$  niepustym zbiorem. Zakłada się, iż w zbiorze  $X$  określone są dwa działania:

- dodawanie elementów zbioru  $X$ :  $X \times X \ni (a, b) \mapsto a + b \in X$ ,
- mnożenie elementów zbioru  $X$  przez elementy ciała  $K$ :  $K \times X \ni (\alpha, a) \mapsto \alpha a \in X$ .

Elementy ciała  $K$  nazywane są *skalarami*, zaś elementy zbioru  $X$  - *wektorami*. Zbiór  $X$  jest *przestrzenią liniową (wektorową)* nad ciałem  $K$ , jeżeli spełnione są następujące warunki:

- $a + b = b + a \quad \forall (a, b) \in X \times X$ ,
- $(a + b) + c = a + (b + c) \quad \forall (a, b, c) \in X \times X \times X$ ,
- równanie  $a + x = b$  ma dokładnie jedno rozwiązanie w zbiorze  $X$  oznaczane jako  $x = b - a \quad \forall (a, b) \in X \times X$ ,
- $\alpha(a + b) = \alpha a + \alpha b \quad \forall (\alpha, a, b) \in K \times X \times X$ ,
- $(\alpha + \beta)a = \alpha a + \beta a \quad \forall (\alpha, \beta, a) \in K \times K \times X$ ,
- $\alpha(\beta a) = (\alpha\beta)a \quad \forall (\alpha, \beta, a) \in K \times K \times X$ ,
- $1a = a \quad \forall a \in X$ , gdzie  $1 \in K$  jest jednością ciała  $K$ .

Element  $0_X \in X$  taki, że  $\forall a \in X$  zachodzi  $a + 0_X = a$  nazywany jest *zerem* przestrzeni  $X$ . Z kolei, elementem *odwrotnym* do  $a \in X$  jest taki element  $b \in X$ , że  $a + b = 0_X$ . Element odwrotny do  $a$  oznaczany jest przez  $-a$ .  $\forall a \in X$  zachodzi  $0a = 0_X$ , gdzie  $0 \in K$  oznacza zero ciała  $K$ . Z powyższej definicji przestrzeni liniowej wynika, iż zbiór  $X$  z działaniem dodawania wektorów tworzy strukturę algebraiczną *grupy przemiennej*.

Niech  $X_i \subset X$ ,  $i \in \{1, \dots, k\}$ , oznacza rodzinę podzbiorów przestrzeni liniowej  $X$ . Przez *sumę algebraiczną* tych podzbiorów rozumie się zbiór

$$\sum_{i=1}^k X_i = \left\{ a \in X : a = \sum_{i=1}^k a_i, a_i \in X_i, i \in \{1, \dots, k\} \right\}.$$

Niech  $X$  będzie przestrzenią liniową. Podzbiór  $X_0 \subset X$  nazywany jest *podprzestrzenią liniową* przestrzeni  $X$ , jeżeli spełnia aksjomaty przestrzeni liniowej z działaniem grupowym i mnożeniem przez skalar obcięty do  $X_0$ . *Rozmaitością liniową (hiperpłaszczyzną)* w  $X$  jest każdy podzbiór o postaci:  $\{a \in X : a = \{a_0\} + X_0\}$ , gdzie  $a_0 \in X$ , zaś  $X_0 \subset X$  jest podprzestrzenią liniową w  $X$ . *Prostą* przechodzącą przez punkt  $a_0 \in X$  w kierunku  $a_1 \in X$  jest następujący podzbiór przestrzeni  $X$ :  $\{a \in X : a = a_0 + \alpha a_1, \alpha \in K\}$ . Każda prosta przechodząca przez punkt  $0_X \in X$  jest podprzestrzenią liniową przestrzeni  $X$ .

Niech  $X$  i  $Y$  będą przestrzeniami liniowymi nad tym samym ciałem  $K$ , a zbiór  $D(A) \subset X$  podprzestrzenią liniową w  $X$ . Przekształcenie (odwzorowanie)  $A : X \supset D(A) \rightarrow Y$  nazywane jest *liniowym*, gdy spełnione są następujące warunki:

- $A(\alpha a) = \alpha A(a) \quad \forall (\alpha, a) \in K \times D(A) \subset K \times X$ ,
- $A(a + b) = A(a) + A(b) \quad \forall (a, b) \in D(A) \times D(A) \subset X \times X$  .

Wartość przekształcenia liniowego  $A$  dla  $a \in D(A)$  oznacza się jako  $Aa$ . Zbiór  $D(A) \subset X$  nazywany jest *dziedziną* przekształcenia  $A$ . Przez  $\mathcal{L}(X, Y)$  oznacza się zbiór wszystkich przekształceń liniowych określonych na dziedzinach zawartych w  $X$  oraz o wartościach w  $Y$ . Należy zauważyć, iż w ogólności  $\mathcal{L}(X, Y)$  może nie być przestrzenią liniową ze względu na odmiennie zdefiniowane dziedziny przekształceń tego zbioru. Gdy rozważa się przekształcenia liniowe określone na całej przestrzeni  $X$  o wartościach w  $Y$ , zbiór  $\mathcal{L}(X, Y)$  oznaczany teraz jako  $\mathcal{L}_0(X, Y)$ , jest przestrzenią liniową. Zachodzi przy tym  $\mathcal{L}_0(X, Y) \subset \mathcal{L}(X, Y)$ . Zbiory  $\mathcal{L}(X, X)$  oraz  $\mathcal{L}_0(X, X)$  oznaczają się odpowiednio  $\mathcal{L}(X)$  oraz  $\mathcal{L}_0(X)$ .

Dalsze rozważania dotyczyć będą przestrzeni liniowych nad ciałem liczb rzeczywistych  $\mathbb{R}$  oraz zespolonych  $\mathbb{C}$ .

## Wektory i macierze

Niech  $\mathbb{R}^{m \times n}$  oznacza przestrzeń liniową *macierzy* rzeczywistych o wymiarach  $m \times n$

$$A \in \mathbb{R}^{m \times n} \Leftrightarrow A = [a_{ij}]_{i \in \{1, \dots, m\}, j \in \{1, \dots, n\}} = [a_{ij}]_{m, n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

Niekiedy elementy macierzy  $A \in \mathbb{R}^{m \times n}$  indeksujemy jako  $a_{i,j}$ , gdzie  $i \in \{1, \dots, m\}$  oraz  $j \in \{1, \dots, n\}$ .

Oprócz  *dodawania macierzy* ( $\mathbb{R}^{m \times n} \times \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ )

$$C = A + B \quad : \quad c_{ij} = a_{ij} + b_{ij}, \quad i \in \{1, \dots, m\}, j \in \{1, \dots, n\}$$

oraz  *mnożenia macierzy przez skalar* ( $\mathbb{R} \times \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ )

$$B = \alpha A \quad : \quad b_{ij} = \alpha a_{ij}, \quad i \in \{1, \dots, m\}, j \in \{1, \dots, n\}$$

do podstawowych działań zdefiniowanych dla macierzy należą:

-  *mnożenie macierzy* ( $\mathbb{R}^{m \times n} \times \mathbb{R}^{n \times p} \rightarrow \mathbb{R}^{m \times p}$ )

$$C = AB \quad : \quad c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}, \quad i \in \{1, \dots, m\}, j \in \{1, \dots, p\}$$

-  *transpozycja macierzy* ( $\mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{n \times m}$ )

$$B = A^T \quad : \quad b_{ij} = a_{ji}, \quad i \in \{1, \dots, m\}, j \in \{1, \dots, n\}.$$

Przez  $\mathbb{R}^m$  rozumie się przestrzeń liniową *wektorów kolumnowych* ( $\mathbb{R}^m \equiv \mathbb{R}^{m \times 1}$ ). Współrzędne wektora  $x \in \mathbb{R}^m$  oznacza się przy pomocy pojedynczej notacji indeksowej  $x = [x_1 \ \cdots \ x_m]^T$ .  $\forall (x, y) \in \mathbb{R}^m \times \mathbb{R}^n$  określony jest *iloczyn zewnętrzny*  $xy^T \in \mathbb{R}^{m \times n}$

$$xy^T = \begin{bmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & \ddots & \vdots \\ x_m y_1 & \cdots & x_m y_n \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

$\forall (x, y) \in \mathbb{R}^m \times \mathbb{R}^m$  określony jest *iloczyn wewnętrzny*  $x^T y \in \mathbb{R}$  tych wektorów

$$x^T y = \sum_{i=1}^m x_i y_i.$$

Macierz  $A \in \mathbb{R}^{n \times n}$  nazywana jest rzeczywistą macierzą *kwadratową*. Macierz kwadratowa  $I_n \in \mathbb{R}^{n \times n}$  o postaci

$$I_n = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

jest macierzą *jednostkową* w  $\mathbb{R}^{n \times n}$ . Jeżeli macierze kwadratowe  $A, B \in \mathbb{R}^{n \times n}$  spełniają równanie  $AB = I_n$ , wówczas o macierzy  $B$  mówi się, że jest macierzą *odwrotną* macierzy  $A$  (co zapisujemy jako  $B = A^{-1}$ ), zaś macierz  $A$  nazywa się macierzą *nieosobliwą*. Analogicznie mamy  $A = B^{-1}$ . Dla danej nieosobliwej macierzy  $A \in \mathbb{R}^{n \times n}$  macierz odwrotna  $A^{-1} \in \mathbb{R}^{n \times n}$  jest jednoznacznie wyznaczoną macierzą nieosobliwą. Dla nieosobliwej macierzy  $A \in \mathbb{R}^{n \times n}$  zachodzi  $(A^{-1})^{-1} = A$ , a ponadto  $(A^{-1})^T = (A^T)^{-1}$ , co pozwala na przyjęcie oznaczenia  $A^{-T} \equiv (A^{-1})^T$ .

### Liniowa niezależność. Baza

Niech  $\{a_i \in \mathbb{R}^m\}_{i=1}^n$  będzie zbiorem wektorów. Wektor  $a \in \mathbb{R}^m$ , dla którego zachodzi

$$a = \sum_{i=1}^n \alpha_i a_i, \quad \alpha_i \in \mathbb{R}, \quad i \in \{1, \dots, n\}$$

jest *kombinacją liniową* wektorów  $\{a_i \in \mathbb{R}^m\}_{i=1}^n$ . Niech  $P \subset \mathbb{R}^m$  będzie podzbiorem przestrzeni  $\mathbb{R}^m$ . Zbiorowi temu można przyporządkować następujący podzbiór  $\text{span } P \subset \mathbb{R}^m$

$$\text{span } P = \left\{ a \in \mathbb{R}^m : a = \sum_{i=1}^k \alpha_i a_i, \quad k \in \mathbb{N}, \quad \alpha_i \in \mathbb{R}, \quad a_i \in P, \quad i \in \{1, \dots, k\} \right\}$$

nazywany *powłoką liniową* zbioru  $P$ . Powłoka  $\text{span } P$  jest najmniejszą (ze względu na inkluzję) podprzestrzenią liniową przestrzeni  $\mathbb{R}^m$  zawierającą zbiór  $P$ . O powłoce  $\text{span } P$  mówi się także jako o podprzestrzeni liniowej *rozpiętej* na zbiorze  $P$ .

Niech  $P_1$  i  $P_2$  będą podzbiorem przestrzeni liniowej  $\mathbb{R}^m$ . Zbiory te są *wzajemnie liniowo niezależne*, gdy  $P_1 \neq \{0_m\}$ ,  $P_2 \neq \{0_m\}$  oraz  $\text{span } P_1 \cap \text{span } P_2 = \{0_m\}$ . A zatem, żadnego niezerowego wektora ze zbioru  $P_1$  nie można przedstawić w postaci kombinacji liniowej wektorów ze zbioru  $P_2$  i na odwrót. Zbiór  $P \subset \mathbb{R}^m$  jest *liniowo niezależny*, jeżeli  $0_m \notin P$  oraz  $\forall a \in P$  przekrój  $\{a\} \cap \text{span } \{P \setminus \{a\}\}$  jest zbiorem pustym:  $\{a\} \cap \text{span } \{P \setminus \{a\}\} = \emptyset$ .

Z powyższego wynika, że wektory  $\{a_i \in \mathbb{R}^m\}_{i=1}^n$  są *liniowo niezależne* (tworzą układ wektorów liniowo niezależnych), jeżeli

$$\left( \sum_{j=i}^n \alpha_j a_j = 0_m, \quad \alpha_i \in \mathbb{R}, \quad i \in \{1, \dots, n\} \right) \Rightarrow (\alpha_i = 0, \quad i \in \{1, \dots, n\}) .$$

W przeciwnym razie wektory  $\{a_i \in \mathbb{R}^m\}_{i=1}^n$  nazywane są wektorami *liniowo zależnymi* – co oznacza, że istnieją takie skalary  $\alpha_i \in \mathbb{R}$ ,  $i \in \{1, \dots, n\}$ , nie wszystkie równe zero, dla których powyższa kombinacja liniowa jest wektorem zerowym  $0_m \in \mathbb{R}^m$ . Zbiór wszystkich kombinacji liniowych wektorów  $\{a_i \in \mathbb{R}^m\}_{i=1}^n$  tworzy powłokę liniową  $\text{span } \{a_i\}_{i=1}^n$

$$\mathbb{R}^m \supset \text{span } \{a_i\}_{i=1}^n = \left\{ a \in \mathbb{R}^m : a = \sum_{i=1}^n \alpha_i a_i, \quad \alpha_i \in \mathbb{R}, \quad i \in \{1, \dots, n\} \right\} .$$

Niech  $b \in \text{span } \{a_j\}_{j=1}^n$ . W przypadku, w którym wektory  $\{a_i\}_{i=1}^n$  są liniowo niezależne, wektor  $b$  jest jednoznacznie określoną kombinacją liniową tych wektorów.

Niech  $\{S_i \subset \mathbb{R}^m\}_{i=1}^k$  będzie rodziną podprzestrzeni liniowych przestrzeni  $\mathbb{R}^m$ . Sumą tych podprzestrzeni jest zbiór

$$S_1 + \dots + S_k = \sum_{i=1}^k S_i = \left\{ s \in \mathbb{R}^m : s = \sum_{i=1}^k a_i, \quad a_i \in S_i, \quad i \in \{1, \dots, k\} \right\}$$

zaś ich *przekrojem* – zbiór

$$S_1 \cap \dots \cap S_k = \bigcap_{i=1}^k S_i = \{s \in \mathbb{R}^m : s \in S_i, \quad i \in \{1, \dots, k\}\} .$$

Ponieważ  $0_m \in S_i$ ,  $i \in \{1, \dots, k\}$ , zatem przekrój dowolnych podprzestrzeni liniowych przestrzeni  $\mathbb{R}^m$  jest niepusty. Suma oraz przekrój podprzestrzeni  $\{S_i \subset \mathbb{R}^m\}_{i=1}^k$  są podprzestrzeniami liniowymi przestrzeni  $\mathbb{R}^m$ .

Podprzestrzeń  $S \subset \mathbb{R}^m$  jest *sumą prostą* podprzestrzeni  $\{S_i \subset \mathbb{R}^m\}_{i=1}^k$ , co zapisuje się jako

$$S = S_1 \oplus \dots \oplus S_k = \bigoplus_{i=1}^k S_i$$

jeżeli  $\forall s \in S$  posiada jednoznaczną reprezentację

$$s = \sum_{i=1}^k a_i, \quad a_i \in S_i, \quad i \in \{1, \dots, k\} .$$

Podprzestrzeń  $S \subset \mathbb{R}^m$  jest sumą prostą podprzestrzeni  $\{S_i \subset \mathbb{R}^m\}_{i=1}^k$  wtedy i tylko wtedy, gdy:

- $S = S_1 + \dots + S_k$ ,
- $(S_1 + \dots + S_i) \cap S_{i+1} = 0_m, i \in \{1, \dots, k-1\}$  (co oznacza, że podprzestrzenie  $S_i, i \in \{1, \dots, k\}$ , są wzajemnie liniowo niezależne).

Zbiór wektorów  $\{a_i \subset \mathbb{R}^m\}_{i=1}^k$  nazywa się *bazą* podprzestrzeni  $S \subset \mathbb{R}^m$ , jeżeli

- $a_i \in S, i \in \{1, \dots, k\}$ ,
- $a_i, i \in \{1, \dots, k\}$ , są liniowo niezależne,
- $S = \text{span} \{a_i\}_{i=1}^k$ .

Wszystkie bazy podprzestrzeni  $S \subset \mathbb{R}^m$  mają taką samą liczbę elementów – określaną jako *wymiar* tej podprzestrzeni ( $\dim S$ ). Gdy  $S = \{0_m\}$ , przyjmuje się  $\dim S = 0$ .

Niech  $S \subset \mathbb{R}^m$  będzie  $k$ -wymiarową podprzestrzenią liniową. Wtedy:

- bazę  $S$  tworzy każdy zbiór  $k$  liniowo niezależnych wektorów należących do  $S$ ,
- każdy zbiór  $k+1$  wektorów należących do  $S$  jest zbiorem wektorów liniowo zależnych,
- każdy zbiór wektorów liniowo niezależnych należących do  $S$  da się uzupełnić do bazy tej podprzestrzeni,
- jeżeli wektory  $a_i \in S, i \in \{1, \dots, k\}$ , tworzą bazę podprzestrzeni  $S$  to  $\forall a \in S$  posiada jednoznaczną reprezentację w tej bazie

$$a = \sum_{i=1}^k \alpha_i a_i, \alpha_i \in \mathbb{R}, i \in \{1, \dots, k\}$$

(skalary  $\alpha_i, i \in \{1, \dots, k\}$ , nazywane są *współzrędnymi* wektora  $a$  w bazie  $\{a_i\}_{i=1}^k$ ).

Niech  $S_1, S_2 \subset \mathbb{R}^m$  oznaczają podprzestrzenie liniowe przestrzeni  $\mathbb{R}^m$ . Wymiar sumy  $S_1 + S_2$  tych podprzestrzeni określony jest wzorem

$$\dim(S_1 + S_2) = \dim S_1 + \dim S_2 - \dim(S_1 \cap S_2).$$

W przypadku sumy prostej  $S_1 \oplus S_2$  otrzymuje się

$$\dim(S_1 \oplus S_2) = \dim S_1 + \dim S_2.$$

Wzory powyższe można bez trudu uogólnić na większą liczbę składników; przykładowo dla sumy prostej podprzestrzeni zachodzi

$$\dim \bigoplus_{i=1}^k S_i = \sum_{i=1}^k \dim S_i.$$

### Macierz przekształcenia liniowego

Macierz  $A \in \mathbb{R}^{m \times n}$  traktować można jako macierz pewnego przekształcenia liniowego  $\mathbb{R}^n \rightarrow \mathbb{R}^m$ :  $x \mapsto Ax$ . Zbiór wszystkich takich macierzy jest bowiem izomorficzny ze zbiorem  $\mathcal{L}_0(\mathbb{R}^n, \mathbb{R}^m)$  wszystkich przekształceń liniowych przestrzeni  $\mathbb{R}^n$  w przestrzeń  $\mathbb{R}^m$ . Wektor  $y = Ax \in \mathbb{R}^m$  jest *obrazem* wektora  $x \in \mathbb{R}^n$ , zaś dla danego wektora  $y \in \mathbb{R}^m$  każdy wektor  $x \in \mathbb{R}^n$ , dla którego zachodzi  $y = Ax$  jest *przeciwoobrazem* wektora  $y$ . Zauważmy, że gdy  $y \notin A(\mathbb{R}^n)$  taki przeciwoobraz jest zbiorem pustym. Dla  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  wyróżnia się dwie podprzestrzenie tego przekształcenia, często określane także jako podprzestrzenie związane z macierzą  $A \in \mathbb{R}^{m \times n}$ . Zbiór  $\text{Im } A \subset \mathbb{R}^m$  zdefiniowany jako

$$\text{Im } A = \{y \in \mathbb{R}^m : \exists x \in \mathbb{R}^n \vdash y = Ax\}$$

nazywany jest *obrazem* (*przestrzenią wartości*, *przestrzenią zasięgową*) przekształcenia  $A$  (macierzy  $A$ ). Zbiór  $\text{Ker } A \subset \mathbb{R}^n$ , określane jako *przestrzeń zerowa* (*jądro*) odwzorowania  $A$  (macierzy  $A$ ), zdefiniowany jest następująco

$$\text{Ker } A = \{x \in \mathbb{R}^n : Ax = 0_m\}.$$

Dla macierzy  $A = [a_1 \ \dots \ a_n] \in \mathbb{R}^{m \times n}$ ,  $a_i \in \mathbb{R}^m$ ,  $i \in \{1, \dots, n\}$ , obraz  $\text{Im } A$  jest podprzestrzenią liniową przestrzeni  $\mathbb{R}^m$  rozpiętą na kolumnach tej macierzy:  $\text{Im } A = \text{span}\{a_i\}_{i=1}^n$ . O  $\text{Im } A$  mówi się jako o *kolumnowej podprzestrzeni* danej macierzy  $A$ . Przestrzeń zerowa  $\text{Ker } A$  macierzy  $A \in \mathbb{R}^{m \times n}$  jest podprzestrzenią liniową przestrzeni  $\mathbb{R}^n$ .

*Rząd przekształcenia liniowego*  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  (*rząd macierzy*  $A$ ) definiowany jest jako

$$\text{rank } A = \dim \text{Im } A.$$

Ponieważ zachodzi  $\text{rank } A = \text{rank } A^T$ , zatem:

- rząd macierzy  $A$  jest równy maksymalnej liczbie liniowo niezależnych kolumn oraz maksymalnej liczbie liniowo niezależnych wierszy tej macierzy,
- $\text{rank } A \leq \min\{m, n\}$  .

Macierz  $A \in \mathbb{R}^{m \times n}$  jest macierzą *regularną* (macierzą o *pełnym rzędzie*), gdy  $\text{rank } A = \min\{m, n\}$ . Dla dowolnego przekształcenia liniowego  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  (dowolnej macierzy  $A \in \mathbb{R}^{m \times n}$ ) zachodzi

$$\dim \text{Ker } A + \dim \text{Im } A = \dim \text{Ker } A + \text{rank } A = \dim \mathbb{R}^n = n .$$

Wynika stąd, iż dla macierzy kwadratowej  $A \in \mathbb{R}^{n \times n}$  poniższe warunki są równoważne:

- macierz  $A$  jest nieosobliwa,
- $\text{Ker } A = \{0_n\}$ ,
- macierz  $A$  ma pełny rząd ( $\text{rank } A = n$ ) .

W oparciu o podane wyżej definicje można pokazać, że:

- $\text{rank } a = \begin{cases} 0 & \text{gdy } a = 0_n \\ 1 & \text{gdy } a \neq 0_n \end{cases} \quad \forall a \in \mathbb{R}^n,$
- $\text{rank } (ab^T) = \begin{cases} 0 & \text{gdy } (a = 0_n) \vee (b = 0_m) \\ 1 & \text{gdy } (a \neq 0_n) \wedge (b \neq 0_m) \end{cases} \quad \forall (a, b) \in \mathbb{R}^n \times \mathbb{R}^m,$
- $\text{rank } (A + B) \leq \text{rank } A + \text{rank } B \quad \forall (A, B) \in \mathbb{R}^{m \times n} \times \mathbb{R}^{m \times n},$
- $\text{rank } (AB) \leq \min\{\text{rank } A, \text{rank } B\} \quad \forall (A, B) \in \mathbb{R}^{m \times n} \times \mathbb{R}^{n \times p}$  .

Dla dowolnych  $A \in \mathbb{R}^{m \times n}$  oraz  $B \in \mathbb{R}^{n \times p}$  zachodzi

$$\begin{aligned} \text{rank } AB &= \text{rank } B - \dim \text{Im } B \cap \text{Ker } A \\ &= \text{rank } A - \dim \text{Im } A^T \cap \text{Ker } B^T . \end{aligned}$$

Podprzestrzeń liniowa  $S \subset \mathbb{R}^n$  jest podprzestrzenią *niezmienniczą* (*inwariantną*) przekształcenia  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , gdy  $\forall a \in S$  zachodzi  $Aa \in S$ . Oznacza to, że obraz podprzestrzeni  $S$  względem przekształcenia  $A$  zawiera się w  $S$ :  $AS \subset S$ . Podprzestrzeń taką określa się także jako  $A$ -inwariantną.

## Ortogonalność

Zbiór wektorów  $\{x_i \in \mathbb{R}^m\}_{i=1}^k$  jest *ortogonalny*, gdy  $x_i^T x_j = 0$ ,  $i \neq j$ ,  $i, j \in \{1, \dots, k\}$ . Jeżeli ponadto  $x_i^T x_i = 1$ ,  $i \in \{1, \dots, k\}$ , zbiór ten jest określany jako *ortonormalny* (dla zbioru takiego zachodzi zatem  $x_i^T x_j = \delta_{ij}$ , gdzie  $\delta_{ij}$ ,  $i, j \in \{1, \dots, k\}$ , oznacza symbol Kroneckera). Fakt  $x_i^T x_j = 0$  zapisujemy jako  $x_i \perp x_j$ . O podprzestrzeniach  $S_i \subset \mathbb{R}^m$ ,  $i \in \{1, \dots, k\}$ , mówi się, że stanowią rodzinę podprzestrzeni *wzajemnie ortogonalnych*, gdy  $\forall (x, y) \in S_i \times S_j \subset \mathbb{R}^m \times \mathbb{R}^m$  zachodzi  $x^T y = 0$ ,  $i \neq j$ ,  $i, j \in \{1, \dots, k\}$ . Niech  $S \subset \mathbb{R}^m$  oznacza podprzestrzeń liniową przestrzeni  $\mathbb{R}^m$ . *Ortogonalnym uzupełnieniem* tej podprzestrzeni jest zbiór

$$\mathbb{R}^m \supset S^\perp = \{x \in \mathbb{R}^m : x^T y = 0 \quad \forall y \in S\} .$$

Zachodzi:

- $S^\perp$  jest podprzestrzenią liniową przestrzeni  $\mathbb{R}^m$ ,
- $\dim S + \dim S^\perp = m$ ,
- $\forall A \in \mathbb{R}^{m \times n} \quad \text{Im } A^\perp = \text{Ker } A^T$  oraz  $\text{Im } (A^T)^\perp = \text{Ker } A$  .

Wektory  $\{a_i \in \mathbb{R}^m\}_{i=1}^k$  tworzą ortonormalną bazę podprzestrzeni liniowej  $S \subset \mathbb{R}^m$ , gdy są ortonormalne (a zatem liniowo niezależne!) oraz  $S = \text{span } \{a_i\}_{i=1}^k$ . W przypadku, gdy  $\dim S = k < m$ , bazę tę można zawsze rozszerzyć do bazy ortonormalnej  $\{a_1, \dots, a_k, a_{k+1}, \dots, a_m\}$  w  $\mathbb{R}^m$ . Zachodzi wówczas  $S^\perp = \text{span } \{a_i\}_{i=k+1}^m$ . Niech zatem wektory  $a_i \in \mathbb{R}^m$ ,  $i \in \{1, \dots, m\}$ , tworzą ortonormalną bazę przestrzeni  $\mathbb{R}^m$ .  $\forall a \in \mathbb{R}^m$  w bazie tej określona jest jednoznaczna reprezentacja

$$a = \sum_{i=1}^m (a_i^T a) a_i .$$

Dla dowolnej (niekoniecznie ortonormalnej) bazy  $\{a_i \in \mathbb{R}^m\}_{i=1}^m$  istnieje i jest jednoznacznie określona *baza sprzężona*  $\{b_i \in \mathbb{R}^m\}_{i=1}^m$ , taka, że:

- zbiór  $\{b_i \in \mathbb{R}^m\}_{i=1}^m$  jest bazą w przestrzeni  $\mathbb{R}^m$ ,
- $b_i^T a_j = \delta_{ij}$ ,  $i, j \in \{1, \dots, m\}$  .

O zbiorach  $\{a_i \in \mathbb{R}^m\}_{i=1}^m$  oraz  $\{b_i \in \mathbb{R}^m\}_{i=1}^m$  mówi się, że stanowią *biortogonalny* układ wektorów w  $\mathbb{R}^m$ . W takim przypadku bazę sprzężoną  $\{b_i \in \mathbb{R}^m\}_{i=1}^m$  tworzą kolumny macierzy  $[b_1 \ \dots \ b_m] = [a_1 \ \dots \ a_m]^{-T}$ . Należy podkreślić, iż w przypadku  $k$ -wymiarowej podprzestrzeni  $S \subset \mathbb{R}^m$ ,

$k < m$  baza sprzężona danej bazy  $\{a_i \in S\}_{i=1}^k$  tej podprzestrzeni może nie istnieć lub też może być określona niejednoznacznie.

Macierz  $A = [a_1 \ \cdots \ a_n] \in \mathbb{R}^{m \times n}$ ,  $m \geq n$ , jest *kolumnowo ortogonalna*, jeżeli jej kolumny  $a_i$ ,  $i \in \{1, \dots, n\}$ , są niezerowe oraz tworzą układ ortogonalny. Macierz  $A \in \mathbb{R}^{m \times n}$ ,  $m \geq n$ , jest zatem kolumnowo ortogonalna wtedy i tylko wtedy, gdy  $A^T A = I_n$ . Macierz  $A \in \mathbb{R}^{m \times n}$ ,  $m \leq n$ , jest *wierszowo ortogonalna*, gdy  $A^T$  jest kolumnowo ortogonalna. Macierz  $A \in \mathbb{R}^{m \times n}$ ,  $m \leq n$ , jest zatem wierszowo ortogonalna wtedy i tylko wtedy, gdy  $AA^T = I_m$ .

W przypadku macierzy kwadratowych  $A \in \mathbb{R}^{n \times n}$  wyróżnia się:

- macierze *symetryczne*:  $A^T = A$ ,
- macierze *skośnie symetryczne*:  $A^T = -A$ ,
- macierze *dodatnio określone*:  $x^T A x > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0_n\}$ ,
- macierze *dodatnio półokreślone (nieujemnie określone)*:  $x^T A x \geq 0 \quad \forall x \in \mathbb{R}^n$ ,
- macierze *nieokreślone*:  $\exists(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$  takie, że  $(x^T A x)(y^T A y) < 0$ ,
- macierze *ortogonalne*:  $A^T A = I_n$ ,
- macierze *nilpotentne*:  $\exists k \in \mathbb{N}$  takie, że  $A^k = 0_{n \times n}$ ,
- macierze *idempotentne*:  $A^2 = A$ .

Dla dowolnej macierzy ortogonalnej  $A \in \mathbb{R}^{n \times n}$  zachodzi:

- $A$  jest kolumnowo oraz wierszowo ortogonalna,
- $A$  jest nieosobliwa ( $A^{-1} = A^T$ ), przy czym macierz odwrotna  $A^{-1}$  jest także macierzą ortogonalną,
- $|\det A| = 1$ ,
- przekształcenie  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  jest przekształceniem  $\mathbb{R}^n$  na  $\mathbb{R}^n$ , a jego niezmiennikiem jest iloczyn wewnętrzny wektorów ( $\forall(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$  zachodzi bowiem  $(Ax)^T(Ay) = x^T y$ ).

Niech  $A, B \in \mathbb{R}^{n \times n}$ . Jeżeli istnieje taka nieosobliwa macierz  $X \in \mathbb{R}^{n \times n}$ , że  $B = X^T A X$ , wówczas o macierzach  $A$  i  $B$  mówi się, że są *kongruentne*. Niezmiennikami relacji kongruencji są cechy symetrii, skośnej symetrii oraz określoności macierzy. Macierze  $A$  i  $B$ , dla których obowiązuje relacja  $B = X^{-1} A X$  są macierzami *podobnymi*. Niezmiennikami relacji podobieństwa macierzy są widma tych macierzy, ich wyznaczniki oraz ślady.

## Rzuty

Niech  $\mathbb{R}^n = M \oplus W$ , gdzie  $M, W \subset \mathbb{R}^n$  oznaczają pewne podprzestrzenie liniowe przestrzeni  $\mathbb{R}^n$ .  $\forall x \in \mathbb{R}^n$  istnieje wówczas jednoznaczne przedstawienie  $x = x_M + x_W$  takie, że  $x_M \in M$  oraz  $x_W \in W$ . Przekształcenie liniowe  $P : \mathbb{R}^n \rightarrow M$ , zdefiniowane jako

$$x \mapsto Px = x_M$$

nazywane jest *rzutem* przestrzeni  $\mathbb{R}^n$  na podprzestrzeń  $M$  *równoległym* do podprzestrzeni  $W$ . Podprzestrzeń  $M$  określana jest jako przestrzeń *rzutowania*, a podprzestrzeń  $W$  jako *kierunek rzutowania* przekształcenia  $P$ . Z powyższej definicji wynika, że:

- $\text{Im } P = M$  oraz  $\text{Ker } P = W$ ,
- gdy  $W = \mathbb{R}^n$  rzut  $P$  jest przekształceniem zerowym ( $P = 0_{n \times n}$ ),
- gdy  $W = \{0_n\}$  rzut  $P$  jest przekształceniem identycznościowym ( $P = I_n$ ).

Macierz  $P \in \mathbb{R}^{n \times n}$  przekształcenia rzutowego  $P : \mathbb{R}^n \rightarrow \mathbb{R}^n$  jest macierzą idempotentną ( $P^2 = P$ ). Prawdziwe jest także twierdzenie głoszące, że każde przekształcenie liniowe  $P : \mathbb{R}^n \rightarrow \mathbb{R}^n$  o idempotentnej macierzy  $P$  jest przekształceniem rzutowym. Przestrzenią rzutowania tego przekształcenia jest obraz  $\text{Im } P$ , a kierunkiem rzutowania jądro  $\text{Ker } P$ , przy czym w omawianym przypadku zachodzi  $\text{Im } P = \{a \in \mathbb{R}^n : Pa = a\}$  oraz  $\text{Ker } P = \{a \in \mathbb{R}^n : a = b - Pb, b \in \mathbb{R}^n\}$ . Jeżeli  $P \in \mathbb{R}^{n \times n}$  jest macierzą przekształcenia rzutowego wówczas  $I_n - P$  jest macierzą przekształcenia rzutowego przestrzeni  $\mathbb{R}^n$  na podprzestrzeń  $\text{Ker } P$  równoległego do podprzestrzeni  $\text{Im } P$ . W przypadku, gdy  $W = M^\perp$  rzut przestrzeni  $\mathbb{R}^n$  na podprzestrzeń  $M$  równoległy do podprzestrzeni  $M^\perp$  nazywany jest *rzutem ortogonalnym*  $\mathbb{R}^n$  na  $M$ .

Niech  $\{v_i \in \mathbb{R}^n\}_{i=1}^k$  będzie dowolną ortonormalną bazą podprzestrzeni  $M$  o wymiarze  $\dim M = k$ . Macierz  $P = VV^T$ , gdzie  $V = [v_1 \ \cdots \ v_k] \in \mathbb{R}^{n \times k}$  jest ową jednoznacznie wyznaczoną macierzą rzutu ortogonalnego.  $\forall x \in \mathbb{R}^n$  istnieje jednoznaczne przedstawienie  $x = x_M + x_{M^\perp}$  takie, że  $x_M \in M$  oraz  $x_{M^\perp} \in M^\perp$ . Ponieważ  $x_M$  można wyrazić jako

$$x_M = \sum_{i=1}^k (v_i^T x_M) v_i$$

zaś dla  $x_{M^\perp}$  zachodzi  $V^T x_{M^\perp} = 0_k$ , zatem

$$\begin{aligned} VV^T x &= VV^T x_M = \sum_{i=1}^k (v_i^T x_M) V(V^T v_i) = \sum_{i=1}^k (v_i^T x_M) V e_i \\ &= \sum_{i=1}^k (v_i^T x_M) v_i = x_M \end{aligned}$$

gdzie przez  $e_i \in \mathbb{R}^k$  oznaczono  $i$ -ty wersor przestrzeni  $\mathbb{R}^k$ ,  $i \in \{1, \dots, k\}$ . Macierz  $P$  rzutu ortogonalnego jest przeto symetryczną macierzą idempotentną ( $P^2 = P$ ,  $P^T = P$ ).

Dla dowolnej podprzestrzeni  $M \subset \mathbb{R}^n$  prawdziwe jest także następujące twierdzenie: jeżeli macierz  $P \in \mathbb{R}^{n \times n}$  spełnia warunki:  $\text{Im } P = M$ ,  $P^2 = P$  oraz  $P^T = P$ , wówczas  $P$  jest macierzą rzutu ortogonalnego przestrzeni  $\mathbb{R}^n$  na podprzestrzeń  $M$ . Z idempotentności  $P$  wynika bowiem, że  $P$  jest macierzą rzutu  $\mathbb{R}^n$  na  $M$  równoległego do  $\text{Ker } P$ . Ponieważ  $\text{Ker } P = \text{Im } P^{T^\perp} = \text{Im } P^\perp$ , zatem kierunek rzutowania macierzy  $P$  jest ortogonalny do przestrzeni rzutowania  $M$ .

### Zagadnienie własne

Zagadnienie własne, związane z macierzą  $A \in \mathbb{R}^{n \times n}$  (liniowym operatorem  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  ( $\mathbb{C}^n \rightarrow \mathbb{C}^n$ )), polega na wyznaczeniu zbioru *wektorów własnych*  $\{x_i \in \mathbb{R}^n\}_{i=1}^n$ , czyli zbioru kierunków  $A$ -niezmienniczych oraz zbioru *wartości własnych (widma)*  $\{\lambda_i \in \mathbb{R} (\mathbb{C})\}_{i=1}^n$ , dla których zachodzi

$$Ax_i = \lambda_i x_i, \quad x_i \neq 0_n, \quad i \in \{1, \dots, n\}.$$

Aby jednorodne liniowe równanie  $(A - \lambda_i I_n)x_i = 0_n$  posiadało niezerowe rozwiązanie  $x_i \in \text{Ker}(A - \lambda_i I_n)$ ,  $x_i \neq 0_n$ , musi obowiązywać nierówność

$$\dim \text{Ker}(A - \lambda_i I_n) > 0, \quad i \in \{1, \dots, n\}.$$

Wynika stąd, że wartości własne  $\{\lambda_i\}_{i=1}^n$  wyznaczymy, rozważając warunek

$$\text{rank}(A - \lambda_i I_n) < n, \quad i \in \{1, \dots, n\}.$$

Wartości własne macierzy  $A$  muszą przeto spełniać następujące algebraiczne równanie stopnia  $n$

$$\det(A - \lambda_i I_n) = 0, \quad i \in \{1, \dots, n\}.$$

Twierdzenie *Cayley'a-Hamiltona* głosi, że każda macierz  $A \in \mathbb{R}^{n \times n}$  spełnia swoje *równanie charakterystyczne*:  $\varphi_A(A) = 0_{n \times n}$ , gdzie  $\varphi_A(\lambda) = \det(\lambda I_n - A)$  jest *charakterystycznym wielomianem* tej macierzy.

## Rozkład macierzy według wartości szczególnych

Niech  $A \in \mathbb{R}^{m \times n}$ . Dowolną taką macierz można przedstawić w postaci iloczynu

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

gdzie:  $U \in \mathbb{R}^{m \times m}$  oraz  $V \in \mathbb{R}^{n \times n}$  są macierzami ortogonalnymi ( $U^T = U^{-1}$ ,  $V^T = V^{-1}$ ), zaś  $\Sigma \in \mathbb{R}^{m \times n}$  jest macierzą blokowo diagonalną

$$\Sigma = \begin{bmatrix} \Sigma_r & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix}$$

o diagonalnej podmacierzy  $\Sigma_r \in \mathbb{R}^{r \times r}$

$$\Sigma_r = \text{diag} \{ \sigma_i \}_{i=1}^r = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & \sigma_r \end{bmatrix},$$

gdzie  $r = \text{rank } A$ , zaś nieujemne liczby

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

to wartości szczególne macierzy  $A$  (*osobliwe, singularne*, stąd *svd - singular value decomposition*). Wygodnie jest przyjąć, że  $\sigma_{r+1} = \cdots = \sigma_{\min\{m,n\}} = 0$ . Dana macierz  $A \in \mathbb{R}^{m \times n}$  posiada przeto  $r$  dodatnich wartości szczególnych oraz  $\min\{m,n\} - r$  zerowych wartości szczególnych. Kolumny  $u_i \in \mathbb{R}^m$ ,  $i \in \{1, \dots, m\}$  ortonormalnej macierzy  $U = [u_1 \ \cdots \ u_m]$  są *lewymi wektorami szczególnymi* macierzy  $A$ , zaś kolumny  $v_i \in \mathbb{R}^n$ ,  $i \in \{1, \dots, n\}$  ortonormalnej macierzy  $V = [v_1 \ \cdots \ v_n]$  nazywane są *prawymi wektorami szczególnymi* macierzy  $A$ . Obowiązują następujące użyteczne równości:

- $U^T A V = \Sigma$ ,
- $u_i^T u_j = \delta_{ij}$ ,  $i, j \in \{1, \dots, m\}$ ,
- $v_i^T v_j = \delta_{ij}$ ,  $i, j \in \{1, \dots, n\}$ .

Macierzy  $A \in \mathbb{R}^{m \times n}$  można w jednoznaczny sposób przyporządkować macierz *pseudoodwrotną*  $A^+ \in \mathbb{R}^{n \times m}$ . Pseudoodwrotność rozumiana jest w sensie Moore'a-Penrose'a, co oznacza, że macierz  $A^+$  musi spełniać następujące relacje, które traktować można jako definicję omawianego typu uogólnionej odwrotności:

- $A^+A = (A^+A)^T$ ,
- $AA^+ = (AA^+)^T$ ,
- $AA^+A = A$ ,
- $A^+AA^+ = A^+$  .

Operacja pseudoodwracania macierzy charakteryzuje się następującymi własnościami:

- $A^{++} = (A^+)^+ = A$ ,
- $(A^T)^+ = (A^+)^T$ ,
- dla nieosobliwej macierzy  $A \in \mathbb{R}^{n \times n}$  mamy  $A^+ = A^{-1}$ ,
- jeżeli dany jest rozkład *svd* macierzy  $A \in \mathbb{R}^{m \times n}$ , wówczas rząd macierzy  $A$  można wyznaczyć w oparciu o indeks najmniejszej niezerowej wartości szczególnej tej macierzy

$$\text{rank } A = \arg \min_{1 \leq i \leq \min\{m,n\}} \sigma_i |_{\sigma_i > 0} .$$

Macierz pseudoodwrotną  $A^+$  wyrazić można, korzystając z rozkładu macierzy  $A \in \mathbb{R}^{m \times n}$  według jej wartości szczególnych. Niech zatem będzie dany taki rozkład  $A = U\Sigma V^T$ , gdzie:  $U \in \mathbb{R}^{m \times m}$ ,  $\Sigma \in \mathbb{R}^{m \times n}$ , zaś  $V \in \mathbb{R}^{n \times n}$ . Macierz  $A^+$  przyjmuje postać

$$A^+ = V\Sigma^+U^T,$$

w której

$$\Sigma^+ = \begin{bmatrix} \Sigma_r^{-1} & 0_{r \times (m-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (m-r)} \end{bmatrix}, \quad \Sigma^+ \in \mathbb{R}^{n \times m},$$

zaś

$$\Sigma_r^{-1} = \text{diag} \{ \sigma_i^{-1} \}_{i=1}^r = \begin{bmatrix} \sigma_1^{-1} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & \sigma_r^{-1} \end{bmatrix} .$$

Rozważmy ortonormalne bazy podprzestrzeni  $\text{Im } A \subset \mathbb{R}^m$  oraz  $\text{Ker } A \subset \mathbb{R}^n$  związanych z macierzą  $A \in \mathbb{R}^{m \times n}$ . Wymiary tych podprzestrzeni wynoszą  $\dim \text{Im } A = r$  oraz  $\dim \text{Ker } A = n - r$ , gdzie  $r = \text{rank } A$ . Zakłada się nietrywialny przypadek  $r \geq 1$ . Bazę przestrzeni kolumnowej  $\text{Im } A$  macierzy  $A$  można utworzyć z każdego obrazu  $r$  liniowo niezależnych kolumn tej macierzy,

otrzymanego w dowolnym nieosobliwym przekształceniu liniowym. Bazę tę można następnie poddać ortonormalizacji, uzyskując poszukiwaną bazę ortonormalną. Przykładową bazę ortonormalną można także wyznaczyć, korzystając z rozkładu *svd* macierzy  $A$ :  $A = U\Sigma V^T$ . W tym celu dokonuje się następującej partycji macierzy czynnikowych  $U \in \mathbb{R}^{m \times m}$  oraz  $V \in \mathbb{R}^{n \times n}$ :

$$\begin{aligned} U &= [U_r \ \bar{U}_r] \in \mathbb{R}^{m \times m}, \quad U_r \in \mathbb{R}^{m \times r}, \quad \bar{U}_r \in \mathbb{R}^{m \times (m-r)} \\ V &= [V_r \ \bar{V}_r] \in \mathbb{R}^{n \times n}, \quad V_r \in \mathbb{R}^{n \times r}, \quad \bar{V}_r \in \mathbb{R}^{n \times (n-r)}. \end{aligned}$$

Macierz  $A$  posiada zatem następującą reprezentację

$$A = U\Sigma V^T = U_r \Sigma_r V_r^T.$$

Jak łatwo zauważyć

$$\text{rank } U_r = \text{rank } \Sigma_r = \text{rank } V_r = r$$

co oznacza, że kolumny macierzy  $U_r$ , czyli odpowiednie lewe szczególne wektory rozważanej macierzy  $A$ , stanowią poszukiwaną ortonormalną bazę podprzestrzeni  $\text{Im } A$

$$\text{Im } A = \text{span } \{u_i\}_{i=1}^r.$$

Rozważmy z kolei przestrzeń zerową  $\text{Ker } A$  danej macierzy  $A \in \mathbb{R}^{m \times n}$ . Przyjmujemy, że rozpatrujemy nietrywialny przypadek  $r < n$ . Ponieważ przestrzeń  $\text{Ker } A$  jest ortogonalnym uzupełnieniem przestrzeni kolumnowej macierzy  $A^T \in \mathbb{R}^{n \times m}$ , co zapisujemy jako  $\text{Ker } A = \text{Im } A^{T\perp}$ , a zatem na podstawie równości:

$$\begin{aligned} A^T &= V\Sigma^T U^T = V_r \Sigma_r^T U_r^T \\ \text{Im } A^T &= \text{span } \{v_i\}_{i=1}^r \end{aligned}$$

oraz własności macierzy  $V$  wnioskujemy, że poszukiwaną ortonormalną bazą podprzestrzeni  $\text{Ker } A$  jest zbiór kolumn macierzy  $\bar{V}_r$  (czyli odpowiednie prawe wektory szczególne rozważanej macierzy  $A$ )

$$\text{Ker } A = \text{span } \{v_i\}_{i=r+1}^n, \quad \dim \text{Ker } A = \text{rank } \bar{V}_r = n - r.$$

Mamy ponadto

$$\text{Ker } A^T = \text{span } \{u_i\}_{i=r+1}^m$$

a także:

$$\begin{aligned} \text{span } \{u_i\}_{i=1}^r \oplus \text{span } \{u_i\}_{i=r+1}^m &= \mathbb{R}^m \\ \text{span } \{v_i\}_{i=1}^r \oplus \text{span } \{v_i\}_{i=r+1}^n &= \mathbb{R}^n \end{aligned}$$

oraz:

$$\begin{aligned}\text{span} \{u_i\}_{i=1}^r &= (\text{span} \{u_i\}_{i=r+1}^m)^\perp \\ \text{span} \{v_i\}_{i=1}^r &= (\text{span} \{v_i\}_{i=r+1}^n)^\perp.\end{aligned}$$

Poszukując macierzy ortogonalnych projekcji na podprzestrzenie związane z macierzą  $A \in \mathbb{R}^{m \times n}$ , przyjmijmy następujące oznaczenia:

$P_{\text{Im } A}$  – macierz ortogonalnej projekcji  $\mathbb{R}^m$  na  $\text{Im } A$ ,

$P_{\text{Ker } A}$  – macierz ortogonalnej projekcji  $\mathbb{R}^n$  na  $\text{Ker } A$ ,

$P_{\text{Im } A^\perp}$  – macierz ortogonalnej projekcji  $\mathbb{R}^m$  na  $\text{Im } A^\perp$ ,

$P_{\text{Ker } A^\perp}$  – macierz ortogonalnej projekcji  $\mathbb{R}^n$  na  $\text{Ker } A^\perp$ .

Dodatkowo dla macierzy transponowanej:

$P_{\text{Im } A^T}$  – macierz ortogonalnej projekcji  $\mathbb{R}^n$  na  $\text{Im } A^T$ ,

$P_{\text{Ker } A^T}$  – macierz ortogonalnej projekcji  $\mathbb{R}^m$  na  $\text{Ker } A^T$ .

Uwzględniając powyższe ustalenia, otrzymujemy następujące reprezentacje macierzy wymienionych projekcji:

$$\begin{aligned}P_{\text{Im } A} &= U_r U_r^T, & P_{\text{Im } A} &\in \mathbb{R}^{m \times m}, \\ P_{\text{Ker } A} &= \bar{V}_r \bar{V}_r^T, & P_{\text{Ker } A} &\in \mathbb{R}^{n \times n}, \\ P_{\text{Im } A^\perp} = P_{\text{Ker } A^T} &= \bar{U}_r \bar{U}_r^T, & P_{\text{Im } A^\perp}, P_{\text{Ker } A^T} &\in \mathbb{R}^{m \times m}, \\ P_{\text{Ker } A^\perp} = P_{\text{Im } A^T} &= V_r V_r^T, & P_{\text{Ker } A^\perp}, P_{\text{Im } A^T} &\in \mathbb{R}^{n \times n}.\end{aligned}$$

Ponadto:

$$\begin{aligned}\text{rank } P_{\text{Im } A} &= r, \\ \text{rank } P_{\text{Ker } A} &= n - r, \\ \text{rank } P_{\text{Im } A^\perp} = \text{rank } P_{\text{Ker } A^T} &= m - r, \\ \text{rank } P_{\text{Ker } A^\perp} = \text{rank } P_{\text{Im } A^T} &= n - r.\end{aligned}$$

Po wykonaniu elementarnych przekształceń uzyskujemy użyteczne formuły:

$$\begin{aligned}AA^+ &= U_r U_r^T, \\ A^+A &= V_r V_r^T,\end{aligned}$$

z których wynikają następujące reprezentacje macierzy projekcji, wyrażone przy użyciu macierzy pseudoodwrotnej:

$$\begin{aligned} P_{\text{Im } A} &= AA^+, \\ P_{\text{Ker } A^\perp} &= A^+A. \end{aligned}$$

Zauważmy wreszcie, że na podstawie macierzy pseudoodwrotnej łatwo jest także określić macierze projekcji  $P_{\text{Im } A^\perp}$  oraz  $P_{\text{Ker } A}$ :

$$\begin{aligned} P_{\text{Im } A^\perp} &= I_m - AA^+, \\ P_{\text{Ker } A} &= I_n - A^+A. \end{aligned}$$

### Norma

Niech  $X$  oznacza przestrzeń liniową nad ciałem  $K$ . Funkcję  $\|\cdot\|$  określoną na  $X$ ,  $\|\cdot\| : X \rightarrow \mathbb{R}$ , nazywamy *normą*, jeżeli:

$$\begin{aligned} \|x\| &\geq 0 \quad \forall x \in X, \\ \|\alpha x\| &= |\alpha| \cdot \|x\| \quad \forall x \in X, \forall \alpha \in K, \\ \|x + y\| &\leq \|x\| + \|y\| \quad \forall x, y \in X, \\ \|x\| = 0 &\Leftrightarrow x = 0_X. \end{aligned}$$

Gdy  $X = \mathbb{R}^n$  lub  $\mathbb{C}^n$ , wtedy dla  $x = [x_1 \ \cdots \ x_n]^T$  mamy:

$$\begin{aligned} \|x\|_p &= \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} \\ \|x\|_\infty &= \max_i |x_i|. \end{aligned}$$

Dla  $X = \mathbb{R}^{m \times n}$  lub  $\mathbb{C}^{m \times n}$  przy  $A = [a_{ij}]_{m,n}$  mamy:

$$\begin{aligned} \|A\|_1 &= \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}| \\ \|A\|_\infty &= \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}| \\ \|A\|_F &= \left( \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2} \\ \|A\|_s &= \bar{\sigma}(A) = \max_{1 \leq i \leq \min\{m,n\}} \sigma_i(A). \end{aligned}$$

Norma  $\| \cdot \|_F$  to norma *Frobeniusa* (macierzowa norma *euklidesowa*, norma *Schura*), zaś  $\| \cdot \|_s$  jest macierzową normą *spektralną*. Niech  $A \in \mathbb{R}^{m \times n}$  ( $\mathbb{C}^{m \times n}$ ). Rozważając wektorowe normy w przestrzeniach  $\mathbb{R}^m$  ( $\mathbb{C}^m$ ) oraz  $\mathbb{R}^n$  ( $\mathbb{C}^n$ ), zdefiniować można odpowiednie *indukowane* macierzowe normy (normy liniowego operatora  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  ( $\mathbb{C}^n \rightarrow \mathbb{C}^m$ ))

$$\|A\|_{ind} = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Kładąc wektorowe euklidesowe normy, mamy  $\| \cdot \|_{ind} \equiv \| \cdot \|_2 = \| \cdot \|_s$ .

### Liniove zadanie najmniejszych kwadratów

Rozważmy następujące liniove zadanie najmniejszych kwadratów

$$Ax = b,$$

gdzie  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  (przy czym zwykle zakłada się  $m \geq n$ ), w którym poszukiwane rozwiązanie  $x \in \mathbb{R}^n$  spełnia następujący warunek

$$x = \arg \min_{y \in \mathbb{R}^n} \|y\|_2.$$

Z powyższego wynika, że jako rozwiązanie  $x$  przyjmuje się to spośród rozwiązań  $y \in \mathbb{R}^n$  problemu minimalizacji euklidesowej normy residualnego wektora

$$y = \arg \min_{z \in \mathbb{R}^n} \|Az - b\|_2$$

które posiada najmniejszą euklidesową normę. Problem minimalizacji normy  $\|Az - b\|_2$  może bowiem, w ogólności, nie posiadać jednoznacznego rozwiązania.

Przystępując do rozwiązania liniowego zadania najmniejszych kwadratów, zauważmy przede wszystkim, że euklidesowa norma danego wektora jest niezmiennicza względem dowolnego ortogonalnego przekształcenia:  $\|x\|_2 = \|Tx\|_2$ ,  $\forall x \in \mathbb{R}^n$ ,  $\forall T \in \mathbb{R}^{n \times n} : T^{-1} = T^T$ . Niech  $\text{rank } A = r < n$ . Stosując rozkład *svd*  $A = U\Sigma V^T$ ,  $U \in \mathbb{R}^{m \times m}$ ,  $V \in \mathbb{R}^{n \times n}$  oraz  $\Sigma \in \mathbb{R}^{m \times n}$ , uzyskujemy następującą reprezentację euklidesowej normy wektora residualnego

$$\|Ax - b\|_2 = \|U\Sigma V^T x - b\|_2 = \|\Sigma V^T x - U^T b\|_2 = \|\Sigma \tilde{x} - \tilde{b}\|_2,$$

gdzie:

$$\begin{aligned} \tilde{x} &= V^T x, & \tilde{x} &= [\tilde{x}_1 \ \cdots \ \tilde{x}_n]^T \in \mathbb{R}^n, \\ \tilde{b} &= U^T b, & \tilde{b} &= [\tilde{b}_1 \ \cdots \ \tilde{b}_m]^T \in \mathbb{R}^m. \end{aligned}$$

Z kolei, uwzględniając postać macierzy  $\Sigma$ , otrzymujemy

$$\|\Sigma \tilde{x} - \tilde{b}\|_2 = \sqrt{\sum_{i=1}^r (\sigma_i \tilde{x}_i - \tilde{b}_i)^2 + \sum_{i=r+1}^m \tilde{b}_i^2}.$$

Powyższe wyrażenie, traktowane jako funkcja niewiadomych  $\{\tilde{x}_i\}_{i=1}^r$ , przyjmuje wartość minimalną przy

$$\tilde{x}_i = \frac{\tilde{b}_i}{\sigma_i}, \quad i \in \{1, \dots, r\}.$$

Zauważmy, iż wartość ta nie zależy od  $\{\tilde{x}_i\}_{i=r+1}^n$

$$\|\Sigma \tilde{x} - \tilde{b}\|_2 \Big|_{\tilde{x}_i = \tilde{b}_i / \sigma_i, i \in \{1, \dots, r\}} = \sqrt{\sum_{i=r+1}^m \tilde{b}_i^2}.$$

Aby zatem norma  $\|x\|_2 = \|V\tilde{x}\|_2 = \|\tilde{x}\|_2$  miała minimalną wartość, współrzędne  $\{\tilde{x}_i\}_{i=r+1}^n$  należy przyrównać do zera

$$\tilde{x}_i = 0, \quad i \in \{r+1, \dots, n\}.$$

Powyższe rozumowanie można przedstawić w następującej sformalizowanej postaci, korzystając z pojęcia macierzy pseudoodwrotnej

$$\tilde{x} = \Sigma^+ \tilde{b} = \Sigma^+ U^T b \quad \Rightarrow \quad x = V \tilde{x} = V \Sigma^+ U^T b = A^+ b.$$

W przypadku, gdy macierz  $A \in \mathbb{R}^{m \times n}$ ,  $m \geq n$ , posiada pełny kolumnowy rząd,  $\text{rank } A = n$ , zadanie minimalizacji euklidesowej normy wektora residualnego  $Ax - b$  posiada jednoznaczne rozwiązanie, wynikające z następującego układu *równań normalnych*

$$A^T A x = A^T b.$$

W rozważanym przypadku

$$x = (A^T A)^{-1} A^T b$$

co stanowi szczególną postać poprzednio określonego ogólnego przepisu  $x = A^+ b$  rozwiązania liniowego zadania najmniejszych kwadratów, przy czym teraz

$$A^+ = (A^T A)^{-1} A^T.$$

Gdy macierz  $A \in \mathbb{R}^{m \times n}$ ,  $m \geq n$ , jest macierzą o niepełnym kolumnowym rzędzie,  $\text{rank } A = r < n$ , zadanie minimalizacji euklidesowej normy wektora residualnego  $Ax - b$  nie posiada jednoznacznego rozwiązania.

**Zadania**

**Zadanie 1.** Pokaż, że  $\forall A \in \mathbb{R}^{n \times n}$  zarówno  $\text{Im } A$  jak i  $\text{Ker } A$  są podprzestrzeniami  $A$ -inwariantnymi.

**Zadanie 2.** Niech  $S \subset \mathbb{R}^m$  będzie podprzestrzenią liniową przestrzeni  $\mathbb{R}^m$ . Pokaż, że:  $S^\perp$  jest podprzestrzenią liniową przestrzeni  $\mathbb{R}^m$  oraz  $\dim S^\perp = m - \dim S$ .

**Zadanie 3.** Pokaż, że dowolne niezerowe ortogonalne wektory są liniowo niezależne.

**Zadanie 4.** Udowodnij, że relacje kongruencji oraz podobieństwa są relacjami równoważności w  $\mathbb{R}^{n \times n}$ .

**Zadanie 5.** Pokaż, że:

- macierz dodatnio określona jest nieosobliwa, przy czym odwrotność macierzy dodatnio określonej jest macierzą dodatnio określoną,
- odwrotności nieosobliwych macierzy symetrycznych, nieosobliwych jednostkowych macierzy trójkątnych oraz nieosobliwych macierzy trójkątnych są odpowiednio nieosobliwymi macierzami symetrycznymi, jednostkowymi macierzami trójkątnymi oraz macierzami trójkątnymi,
- macierz  $A^T A$ , gdzie  $A \in \mathbb{R}^{m \times n}$ , jest macierzą dodatnio półokreśloną; macierz ta jest dodatnio określona wtedy i tylko wtedy, gdy  $\text{rank } A = n$  ( $A$  jest macierzą o pełnym kolumnowym rzędzie).

**Zadanie 6.** Niech  $A \in \mathbb{R}^{n \times n}$ . Czy mogą obowiązywać następujące relacje:

- a)  $\dim \text{Im } A = 0$ ,   b)  $\dim \text{Ker } A = 0$ ,   c)  $\text{Im } A \subset \text{Ker } A$ ,  
 d)  $\text{Ker } A \subset \text{Im } A$ ,   e)  $\text{Ker } A = \text{Im } A$ ,   f)  $\text{Ker } A = \emptyset$ ,  
 g)  $\text{Im } A = \mathbb{R}^n$ ,   h)  $\text{Ker } A = \mathbb{R}^n$ ,   i)  $\text{Ker } A \perp \text{Im } A$  ?

Odpowiedź uzasadnij (wystarczy przykłady lub kontrprzykłady).

**Zadanie 7.** Czy  $A \text{Ker } A \subset \text{Ker } A \quad \forall A \in \mathbb{R}^{n \times n}$ ? Kiedy  $\text{Ker } A \subset A \text{Ker } A$ ?

**Zadanie 8.** Niech  $A \in \mathbb{R}^{m \times n}$ . Pokaż, że

$$\text{rank } A^T A = \text{rank } A A^T = \text{rank } A$$

oraz:

$$\begin{aligned} \text{Im } A^T A &= \text{Im } A^T, & \text{Im } A A^T &= \text{Im } A \\ \text{Ker } A^T A &= \text{Ker } A, & \text{Ker } A A^T &= \text{Ker } A^T. \end{aligned}$$

**Zadanie 9.** Niech  $A \in \mathbb{R}^{m \times n}$  oraz  $B \in \mathbb{R}^{n \times p}$ . Udowodnij poniższą równość

$$\dim \text{Ker } AB = \dim \text{Ker } B + \dim \text{Ker } A \cap \text{Im } B.$$

**Zadanie 10.** Niech  $A \in \mathbb{R}^{m \times n}$  oraz  $B \in \mathbb{R}^{n \times p}$ . Udowodnij poniższe implikacje:

$$\begin{aligned} \text{rank } B = n &\Rightarrow \text{rank } AB = \text{rank } A \wedge \text{Im } AB = \text{Im } A \\ \text{rank } A = n &\Rightarrow \text{rank } AB = \text{rank } B \wedge \text{Ker } AB = \text{Ker } B. \end{aligned}$$

**Zadanie 11.** Pokaż, że iloczyn macierzy ortogonalnych jest macierzą ortogonalną.

**Zadanie 12.** Pokaż, że euklidesowa norma w  $\mathbb{R}^n$  jest niezmiennicza względem ortogonalnych przekształceń. Co możesz powiedzieć o innych normach?

**Zadanie 13.** Pokaż, że macierz rzutu ortogonalnego przestrzeni  $\mathbb{R}^n$  na daną podprzestrzeń  $M$  ma postać wyznaczoną w sposób jednoznaczny.

**Zadanie 14.** Czy  $\forall A \in \mathbb{R}^{n \times n}$  obowiązuje równość  $A^+ A = A A^+$ ?

**Zadanie 15.** Czy dla dowolnych macierzy  $A$  oraz  $B$  o stosownych wymiarach zawsze zachodzi  $(AB)^+ = B^+ A^+$ ? Co powiesz o macierzach nieosobliwych?

**Zadanie 16.** Niech  $A \in \mathbb{R}^{m \times n}$ . Pokaż, że dla  $A$  o pełnym kolumnowym rzędzie zachodzi  $A^+ = (A^T A)^{-1} A^T$ , zaś dla  $A$  o pełnym rzędzie wierszowym mamy  $A^+ = A^T (A A^T)^{-1}$ .

**Zadanie 17.** Wektory  $a_i \in \mathbb{R}^m$ ,  $i \in \{1, \dots, m\}$ , tworzą pewną ortonormalną bazę w  $\mathbb{R}^m$ . Pokaż, że  $\forall a \in \mathbb{R}^m$  w bazie tej określona jest jednoznaczna reprezentacja  $a = \sum_{i=1}^m (a_i^T a) a_i$ .

**Zadanie 18.** Udowodnij, że pseudoodwrotność  $A^+$  dowolnej macierzy  $A$  charakteryzują następujące własności:  $(A^+)^+ = A$  oraz  $(A^T)^+ = (A^+)^T$ . Pokaż, że dla nieosobliwych macierzy zachodzi  $A^+ = A^{-1}$ .

**Zadanie 19.** Oblicz  $A^+$  dla:  $A = 1$ ,  $A = 0$ ,  $A = [1 \ 0]$ ,  $A = [1 \ 1]^T$ ,  $A = [0 \ 0]^T$  oraz  $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ . Wyznacz rozkład *svd* tych macierzy.

**Zadanie 20.** Pokaż, że  $AB = 0_{n \times m} \Leftrightarrow B^+A^+ = 0_{m \times n}$ .

**Zadanie 21.** Niech  $U, V \subset \mathbb{R}^n$ . Pokaż, że  $\mathbb{R}^n = U \oplus V \Leftrightarrow U \cap V = 0_n \wedge U + V = \mathbb{R}^n$ .

**Zadanie 22.** Niech  $A = U\Sigma V^T$  oznacza rozkład *svd* danej macierzy  $A$ . Sprawdź, że macierz  $V\Sigma^+U^T$  spełnia 'aksjomaty' pseudoodwrotności  $A^+$  tej macierzy w sensie Moore'a-Penrose'a.

**Zadanie 23.** Pokaż, że  $\forall A \in \mathbb{R}^{m \times n}$ :

$$\begin{aligned} A^T &= A^T A A^+ = A^+ A A^T, & A^+ &= A^T (A A^T)^+ = (A^T A)^+ A^T \\ (A^T A)^+ &= A^+ (A^T)^+, & (A A^T)^+ &= (A^T)^+ A^+ \end{aligned}$$

oraz

$$\text{Im } A^+ = \text{Im } A^T = \text{Im } A^+ A, \quad \text{Ker } A^+ = \text{Ker } A^T = \text{Ker } A A^+.$$

**Zadanie 24.** Dane są  $A \in \mathbb{R}^{m \times n}$  oraz  $b \in \mathbb{R}^m$ . Niech  $b \in \text{Im } A$ . Pokaż, że równanie  $Ax = b$  posiada w  $\mathbb{R}^n$  rozwiązanie opisane ogólną formułą  $x = x_0 + x_1$ , gdzie  $x_0 \in \mathbb{R}^n$  spełnia równość  $Ax_0 = b$ , zaś  $x_1 \in \text{Ker } A$ . Co będzie, gdy  $b \notin \text{Im } A$ ?

**Zadanie 25.** Zbadaj możliwości MATLABowych funkcji: `norm`, `rank`, `eig`, `inv`, `pinv`, `qr`, `svd`, `orth`, `null` oraz `cond`. Co w MATLABie oznacza `eps` oraz `flops`?

**Zadanie 26.** Rozważmy liniowe zadanie najmniejszych kwadratów  $Ax = b$ ,  $A \in \mathbb{R}^{m \times n}$ . Niech  $m \geq n$ . MATLAB udostępnia tu dwie możliwości:

`x=pinv(A)*b`, co odpowiada wektorowi  $x = A^+b$  o minimalnej euklidesowej normie,

$\mathbf{x}=\mathbf{A}\backslash\mathbf{b}$ , co odpowiada wektorowi  $x$  o co najwyżej  $r$  niezerowych współrzędnych.

W przypadku zadania z macierzą  $A \in \mathbb{R}^{m \times n}$ ,  $m < n$ , rozwiązanie odpowiedniego niedookreślonego liniowego zadania najmniejszych kwadratów  $Ax = b$ , zapisane w MATLABowym kodzie jako  $\mathbf{x}=\mathbf{A}\backslash\mathbf{b}$ , prowadzi do wektora  $x$ , który:

- minimalizuje euklidesową normę wektora residualnego  $Ax - b$ ,
- charakteryzuje się minimalną wartością euklidesowej normy,
- posiada co najwyżej  $r$  niezerowych współrzędnych, gdzie  $\text{rank } A = r \leq m$ .

Rozwiązanie  $x = A^+b$ , będąc wolnym od ostatniego z wymienionych wyżej ograniczeń, charakteryzuje się z reguły mniejszą wartością normy euklidesowej.

Niech  $x = [-1 \ 0 \ 2]^T$  oraz  $y = [-2 \ -1 \ 1]^T$ . Skomentuj wyniki następujących MATLABowych esperymentów:  $\mathbf{x}\backslash\mathbf{y}$ ,  $\mathbf{y}/\mathbf{x}$ ,  $(\mathbf{x}'\backslash\mathbf{y}')$ ,  $\mathbf{x}'\mathbf{y}/\mathbf{x}'\mathbf{x}$ ,  $\text{pinv}(\mathbf{x})$ ,  $\text{pinv}(\mathbf{x})*\mathbf{y}$  oraz  $(\text{pinv}(\mathbf{x}')*\mathbf{y}')$ .  $\square$