PHƯƠNG PHÁP TOÁN
TRONG MACHINE LEARNING

MỞ ĐẦU

Trong Machine Learning, thường gặp những bài toán dẫn đến việc khảo sát, ước lượng mô hình: $\mathrm{f}: \mathrm{X} \rightarrow \mathrm{Y}$, trong đó f là một ánh xạ từ tập đầu vào X sang Y là tập đầu ra. Chẳng hạn, cần ước lượng giá Y của một căn nhà có diện tích X_1, số phòng ngủ X_2, khoảng cách tới trung tâm thành phố X_3, tức là cần ước lượng giá căn nhà (đầu ra) là hàm $\mathrm{Y} = \mathrm{f}(\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_3)$. Hoặc khi cần dự đoán giới tính Y hay tuổi Z dựa theo những đặc điểm X của khuôn mặt, thì phần dự đoán giới tính Y theo đặc điểm khuôn mặt X là hồi quy logistic (định tính), phần dự đoán tuổi Z theo X là hồi quy định lượng.... Để tiếp cận những vấn đề đặt ra trong Machine Learning, cần có những công cụ không thể thiếu của toán học: Đại số tuyến tính, giải tích, Xác suất thống kê, tối ưu hóa. Để đáp ứng yêu cầu đó, tài liệu này trình bày các nội dung sau đây:

***Chương** 1. **Ma trận và định thức:** (1)* Ma trận và các dạng ma trận đặc biệt; (2)Các phép toán về ma trận, ma trận khả nghịch. (3)Định thức của ma trận vuông; (4)Hạng của ma trận và thuật toán tìm hạng ma trận; (5) Thuật toán tìm ma trận nghịch đảo.

Chương 2. Không gian véc tơ và ánh xạ tuyến tính: *(1) Khái niệm về không gian véc tơ, hệ độc lập tuyến tính.* (2)Cơ sở của không gian véc tơ và tọa độ của véc tơ trong cơ sở sắp thứ tự; (3) Không gian định chuẩn và các chuẩn thường dùng trong $\mathbb{R}^{\mathrm{n}}$và $\mathbb{R}^{\mathrm{m} \times \mathrm{n}}$; (4) Ánh xạ tuyến tính và ma trận của ánh xạ tuyến tính trong các không gian hữa hạn chiều; (5) Ma trận của phép biến đổi tuyến tính trong không gian hữu hạn chiều và thuật toán tìm ma trận chuyển cơ sở; (6) Trị riêng và véc tơ riêng của phép biến đổi tuyến tính và của ma trận.
(7) Ma trận trực giao, ma trận đồng dạng, ma trận xác định dương, bán xác định dương, chéo hóa ma trận

Chương 3. Giải tích ma trận**:** (1) Đạo hàm của hàm có giá trị là số; (2) Đạo hàm của hàm với biến ma trận; (3) Đạo hàm của hàm có đầu ra là véc tơ; (4) Các tính chất quan trọng của đạo hàm: Quy tắc tích và quy tắc chuỗi; (5) Đạo hàm của các hàm cơ bản thường gặp, tính gần đúng đạo hàm

Chương 4. Những nội dung cần thiết về Xác suất: (1) Xác suất và các công thức xác suất quan trọng; (2) Biến ngẫu nhiên và phân phối xác suất; (2) Các mô hình phân phối xác suất quan trọng; (3) Các bất đẳng thức quan trọng: BĐT Markov, BĐT Chebyshev, BĐT Chernoff, BĐT Jensen; (4) Véc tơ ngẫu nhiên và các đặc trưng: véc tơ kỳ vọng, ma trận tương quan, hàm hồi quy, véc tơ ngẫu nhiên hai chiều.

Chương 5. Một số vấn đề về tối ưu hóa: (1) Tập lồi; (2) Hàm lồi và các tính chất; (3) Kiểm tra tính lồi; (4) Bài toán tối ưu lồi

Chương 6. Các đặc trưng mẫu và các phương pháp ước lượng: (1) Mẫu và các đặc trưng mẫu quan trọng; (2) Ước lượng khoảng tin cậy cho các tham số tổng thể; (3) Các phương pháp ước lượng tham ẩn của mô hình: Phương pháp bình phương bé nhất (OLS), ước lượng hợp lý cực đại (MLE), ước lượng hậu nghiệm cực đại (MAP)

Chương 1. MA TRẬN – ĐỊNH THỨC

1.1. MA TRẬN

1.1.1. Các k.n:

- Một bảng các số (thực hoặc phức) được sắp xếp thành m hàng, n cột và đặt trong dấu (.) hoặc dấu [.] được gọi là một ma trận cấp m x n (thực hoặc phức) và m x n số này được gọi là các phần tử của ma trận
- Dùng các chữ A, B, C,... để ký hiệu một ma trận và ký hiệu: $\mathrm{A} = \left(a_{ij}\right)_{m\times n}$ để chỉ A là một ma trận cấp m x n có pần tử ở hàng i và cột j là a_{ij}.
- Kí hiệu $\mathrm{R}^{\mathrm{m}\times\mathrm{n}}$: Tập tất cả các ma trận thực cấp $m\times n$; $\mathbb{C}^{\mathrm{m}\times\mathrm{n}}$: Tập tất cả các ma trận phức cấp $m\times n$;
- $\forall\, \mathrm{A} = \left(a_{ij}\right)_{m\times n} \in \mathrm{R}^{\mathrm{m}\times\mathrm{n}}$: ***chuyển vị*** của A là ma trận $\mathrm{A}^{\mathrm{T}} = \left(\mathrm{b}_{\mathrm{ji}}\right)_{\mathrm{n}\times\mathrm{m}}$ cấp n×m trong đó: $\mathrm{b}_{\mathrm{ji}} = a_{ij}$, $\forall \mathrm{i} = \overline{1,\mathrm{m}}, \forall \mathrm{j} = \overline{1,\mathrm{n}}$. Vậy A^{T} nhận được từ A bằng cách chuyển hàng của A thành cột của A^{T} (hay cột của A thành hàng của A^{T}), nhưng vẫn giữ nguyên thứ tự giữa chúng (phép chuyển vị).
- $\forall\, \mathrm{A} = \left(a_{ij}\right)_{m\times n} \in \mathbb{C}^{\mathrm{m}\times\mathrm{n}}$: ***chuyển vị liên hợp*** của A là ma trận $\mathrm{A}^{\mathrm{H}} = \left(\mathrm{b}_{\mathrm{ji}}\right)_{\mathrm{n}\times\mathrm{m}}$ cấp n×m, trong đó: $\mathrm{b}_{\mathrm{ji}} = \overline{\mathrm{a}_{\mathrm{ij}}}$, $\forall \mathrm{i} = \overline{1,\mathrm{m}}, \forall \mathrm{j} = \overline{1,\mathrm{n}}$.

vd 1: $$\mathrm{A} = \begin{bmatrix} -1 & 0 \\ 3 & -6 \\ 4 & 9 \end{bmatrix};\ \mathrm{B} = \begin{bmatrix} -1+2\mathrm{i} & 0 \\ 3 & 6-3\mathrm{i} \\ 4+\mathrm{i} & 9-4\mathrm{i} \end{bmatrix}$$

$$\Rightarrow \mathrm{A}^{\mathrm{T}} = \begin{bmatrix} -1 & 3 & 4 \\ 0 & -6 & 9 \end{bmatrix}; B^{\mathrm{H}} = \begin{bmatrix} -1-2i & 3 & 4-i \\ 0 & 6+3i & 9+4i \end{bmatrix}$$

- Với $\mathrm{A} = \left(a_{ij}\right)_{m\times n}$, m.trận: $-\mathrm{A} = \left(-a_{ij}\right)_{m\times n}$: Ma trận đối của A
- M.trận không, cấp $m\times n$: $\mathcal{O}_{m\times n}$ là m.trận gồm m.n p.tử đều bằng 0.

- Ma trận cấp $1\times n$ (chỉ có 1 hàng) gọi là ma trận hàng
- Ma trận cấp $m\times 1$ (chỉ có 1 cột) gọi là ma trận cột
- Hàng (cột) của ma trận gồm toàn số 0 gọi là hàng không (cột không)

Chú ý: (1)Cần phân biệt 2 kí hiệu : R^{n}(hoặc $\mathbb{C}^{\mathrm{n}}$) là không gian véc tơ n chiều, mỗi véc tơ gồm n thành phần là n số, được viết dưới dạng cột (một ma trận cột). $\mathrm{R}^{\mathrm{m}\times\mathrm{n}}$ (hoặc $\mathbb{C}^{\mathrm{m}\times\mathrm{n}}$) là không gian các ma trận cấp $m\times n$, mỗi phần tử là một ma trận cấp $m\times n$. Như vậy $\mathrm{R}^{\mathrm{n}}=\mathrm{R}^{\mathrm{n}\times 1}$ ($\mathbb{C}^{\mathrm{n}}=\mathbb{C}^{\mathrm{n}\times 1}$). (2) Nếu A là ma trận thực thì $\mathrm{A}^{\mathrm{H}}=\mathrm{A}^{\mathrm{T}}$

1.1.2. Các dạng ma trận đặc biệt

a- M.trận bậc thang là m.trận thỏa 2 đ/k:

a. Hàng không (gồm toàn số 0), nếu có, phải nằm dưới mọi hàng khác không.

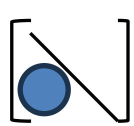
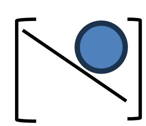
b. Với 2 hàng khác không bất kì kề nhau: p.tử khác 0 đầu tiên của hàng trên phải nằm về bên trái cột chứa p.tử khác 0 đầu tiên của hàng dưới .

$\boldsymbol{VD\ 2}$: $\mathrm{A}=\begin{bmatrix}1&2&3&0\\0&3&0&0\\1&0&0&0\\0&0&0&0\end{bmatrix}$; $\mathrm{B}=\begin{bmatrix}1&0&0&0\\0&3&0&0\\0&0&0&0\\0&1&0&0\end{bmatrix}$; $\mathrm{C}=\begin{bmatrix}2&0&1\\0&-1&9\\0&0&4\\0&0&0\end{bmatrix}$; $\mathrm{D}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&5\end{bmatrix}$

b. Ma trận vuông: Một ma trận $\mathrm{A}=\left(a_{ij}\right)_{\mathrm{n}\times n}$ cấp $n\times n$ được gọi là một ma trận vuông cấp n. Khi đó: $a_{11},a_{22},\dots,a_{\mathrm{nn}}$ được gọi là các phần tử đường chéo (đường chéo chính)

b1. Ma trận vuông A được gọi là ma trận đối xứng nếu $\mathrm{A}^{\mathrm{T}}=\mathrm{A}$, và được gọi là ma trận ***Hermitian*** nếu $\mathrm{A}^{\mathrm{H}}=\mathrm{A}$

b2. Ma trận vuông A được gọi là ma trận ***tam giác trên (tam giác dưới)***, nếu mọi phần tử nằm phía dưới (phía trên) đường chéo đều bằng 0. Ma trận tam giác trên, ma trận tam giác dưới có tên chung là ma trận tam giác. Hình ảnh:

b3. Ma trận đường chéo là ma trận vừa là tam giác trên, vừa là tam giác dưới.
b4. Ma trận đơn vị cấp n, kí hiệu $\mathrm{I_n}$ (có tài liệu kí hiệu $\mathrm{E_n}$) là ma trận đường chéo cấp n mà mọi phần tử đường chéo đều bằng 1.
Chú ý: Ma trận bậc thang nói chung và ma trận tam giác nói riêng được ứng dụng nhiều trong lý thuyết giải hệ phương trình tuyến tính.

1.1.3. Các phép toán đối với các ma trận

a. Phép cộng và phép trừ các ma trận cùng cấp: Cho $\mathrm{A} = \left(a_{ij}\right)_{m\times \mathrm{n}}; B = \left(b_{ij}\right)_{\mathrm{m}\times n}$

$$\mathrm{A} + B = \left(\mathrm{c}_{ij}\right)_{m\times\mathrm{n}} \left(\mathrm{c}_{ij} = \mathrm{a}_{ij} + \mathrm{b}_{ij}, \forall i, j\right);\ \mathrm{A} - \mathrm{B} = \mathrm{A} + (-\mathrm{B})$$

b. Phép nhân ma trận với một số: Cho ma trận $\mathrm{A} = \left(a_{ij}\right)_{m\times\mathrm{n}}$ và số λ**:**

$$\lambda.\mathrm{A} = \left(\lambda.a_{ij}\right)_{m\times\mathrm{n}}$$

VD 3: Tìm ma trận: $3.\begin{bmatrix} 1 & 0 & -2 \\ 2 & 3 & -1 \end{bmatrix} - 2.\begin{bmatrix} 3 & -4 & 2 \\ -5 & 1 & 5 \end{bmatrix}$

c. Phép nhân hai m.trận: Với: $\mathrm{A} = \left(a_{ij}\right)_{m\times p}; B = \left(b_{ij}\right)_{p\times n}$

$$A.\mathrm{B} = \left(c_{ij}\right)_{m\times n}; trong\ \text{đó}: c_{ij} = \sum_{k=1}^{p} a_{ik}\,.\,b_{kj};\ \forall i, j \qquad (1.1)$$

(A) **(B)**

$$\text{hàng i} \rightarrow \begin{bmatrix} \dots & \cdots & \cdots \end{bmatrix} . \begin{bmatrix} & \vdots & \\ & \vdots & \\ & \vdots & \end{bmatrix}$$

(cột j)

Nhận xét: Để tinh p.tử hàng i, cột j của A.B là c_{ij}, ta lấy hàng i của A nhân tương ứng với cột j của B rồi cộng lại.

Chú ý

a. Phép nhân 2 ma trận có kể đến thứ tự các nhân tử và được thực hiện khi và chỉ khi số cột của nhân tử trước bằng số hàng của nhân tử sau.

b. Có thể tồn tại A.B mà không tồn tại B.A và nếu tồn tại cả hai tích thì nói chung: $A.\mathrm{B} \neq \mathrm{B}.\mathrm{A}$.

c. Với ma trận vuông A thì tích của k nhân tử A được viết: $\mathrm{A}.\mathrm{A} \ldots \ldots \mathrm{A} = \mathrm{A}^{\mathrm{k}}$

VD 4. Cho $\mathrm{A} = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 1 & 0 \end{bmatrix}, \mathrm{B} = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 2 \\ -2 & 0 & 6 \end{bmatrix}, \mathrm{C} = \begin{bmatrix} 4 & 0 \\ -1 & 1 \\ 0 & 2 \end{bmatrix}$. Xét các tích A.B, B.A, A.C, C.A

1.1.4. Ma trận nghịch đảo: Ma trận vuông A cấp n được gọi là khả nghịch nếu tồn tại ma trận vuông B cấp n sao cho: $\mathrm{A}.\mathrm{B} = \mathrm{B}.\mathrm{A} = \mathrm{I}_{\mathrm{n}}$. Khi đó B được gọi là ma trận nghịch đảo của A và được ký hiệu: $\mathrm{B} = \mathrm{A}^{-1}$

VD 5. Cho $A = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}$**,** Xét tính khả nghịch và tìm ma trận A^{-1} (nếu có).

Theo đ/n, có thể chỉ ra A khả nghịch và tìm được $\mathrm{A}^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}$**.**

Phương pháp dùng đ/n để giải bài toán này chỉ mang tính minh họa.

1,1.5. Các tính chất của các phép toán về ma trận

T/c 1: Phép cộng có t/c g.hoán, kết hợp, A + (-A) = 0; A + 0 = A

T/c 2: Với hai m.trận A, B, hai số thực a, b thì:

a.(A + B) = a.A + a.B; (a + b).A = a.A + b.A

T/c 3: $(a.A)^T = a.A^T; (A+B)^T = A^T + B^T;$

$(A.B)^T = B^T.A^T; (A^T)^T = A$

T/c 4: $0.A = \mathcal{O}; A.\mathcal{O} = \mathcal{O};\ \mathcal{O}.A = \mathcal{O}; 1.A = A; (-1).A = -A;$

$A_{m\times n}.I_n = A_{m\times n}; I_m.A_{m\times n} = A_{m\times n}$

T/c 5: $(A+B).C = A.C + B.C; C(A+B) = C.A + C.B$

T/c 6: Nếu A, B là các m.trận vuông cùng cấp, khả nghịch thì A.B cũng khả nghịch và: $(A.B)^{-1} = B^{-1}.A^{-1}$

1.2. Định thức của m.trận vuông: Đ.thức của m.trận vuông A là số, kí hiệu: det(A) hoặc: $|A|$

1.2.1. Đ.thức của m.trận vuông cấp 1, cấp 2, cấp 3:

a. Với A = (a) (ma trận vuông cấp 1) thì det (A) = a

Cho: $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$**:** $\quad \mathbf{det}(A) = a_{11}a_{22} - a_{21}a_{12}$

b. Với $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$;

$$\mathbf{det}(A) = a_{11}.\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$
$$= a_{11}.\,a_{22}.\,a_{33} + a_{12}.\,a_{23}a_{31} + a_{13}a_{21}a_{32}$$
$$-a_{11}a_{23}a_{32} - a_{12}.\,a_{21}a_{33} - a_{13}a_{31}a_{22}$$

1.2.2. Định thức của ma trận vuông cấp n (n > 2)**.** Cho: $\mathrm{A} = \left(a_{ij}\right)_{n\times n}$: Từ A bỏ đi (n – r) hàng và (n – r) cột, được một m.trận vuông cấp r, gọi *là m.trận vuông con cấp r* của A. Kí hiệu M_{ij} là m.trận cấp (n – 1), thu được từ A bằng cách bỏ đi hàng i và cột j. Khi đó gọi: $A_{ij} = (-1)^{i+j}.\,det(M_{ij})$ *là phần phụ đại số của p.tử* a_{ij} *trong m.trận* A. Ta đ.nghĩa:

$$\mathbf{det}(A) = a_{11}.\,A_{11} + a_{12}.\,A_{12} + \ldots + a_{1n}.\,A_{1n} \qquad (1.2)$$

Các t.chất của đ.thức: Giả sử A là m.trận vuông cấp n

T/c 1: $\mathbf{det}(A) = a_{i1}.A_{i1} + a_{i2}.A_{i2} + \ldots + a_{in}.A_{in}, \forall i = \overline{1,n}$ (1.2a)

$= a_{1j}.A_{1j} + a_{2j}.A_{2j} + \ldots + a_{nj}.A_{nj}, \forall j = \overline{1,n}$ (1.2b)

và: $a_{i1}.A_{k1} + \ldots + a_{in}.A_{kn} = a_{1j}.A_{1k} + a_{2j}.A_{2k} + \ldots + a_{nj}.A_{nk} = 0, \forall \mathrm{i} \neq k \neq j$

T/c 2: *Nếu A có 1 hàng 0 (cột 0) thì*: $\mathbf{det}(A) = \mathbf{0}$

T/c 3: $\mathbf{det}\left(\boldsymbol{A}^{\boldsymbol{T}}\right) = \mathbf{det}(\boldsymbol{A})$

T/c 4: *Nếu đổi chỗ 2 hàng (2 cột) của m.trận thì định thức chỉ đổi dấu. Đặc biệt nếu A có 2 hàng (2 cột) giống nhau thì* $\boldsymbol{det}(\boldsymbol{A}) = \mathbf{0}$

T/c 5: *Nếu nhân 1 hàng (cột) của A với số r thì định thức được nhân lên với r.*

Đặc biệt: $$\boldsymbol{det}(\boldsymbol{\alpha}.\boldsymbol{A}) = \boldsymbol{\alpha}^{\boldsymbol{n}}.\mathbf{det}\,\boldsymbol{A} \qquad (1.3)$$

T/c 6: *Nếu A có 2 hàng (2 cột) tương ứng tỉ lệ thì:* $\boldsymbol{det}(\boldsymbol{A}) = \mathbf{0}$

T/c 7: *Nếu A có 1 hàng (1 cột) mà mỗi phần tử là tổng của 2 số hạng thì:* $\boldsymbol{det}(\boldsymbol{A}) = \boldsymbol{det}(\boldsymbol{A_1}) + \boldsymbol{det}(\boldsymbol{A_2})$, *($\boldsymbol{A_1}$ và $\boldsymbol{A_2}$ là các m.trận thu được từ A bằng cách giữ nguyên các hàng (cột) khác, còn hàng (cột) nói trên chỉ giữ lại số hạng thứ nhất và thứ 2 tương ứng).* Minh họa:

$$\begin{vmatrix} a_{11} & b_{12}+c_{12} & a_{13} \\ a_{21} & b_{22}+c_{22} & a_{23} \\ a_{31} & b_{32}+c_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & b_{12} & a_{13} \\ a_{21} & b_{22} & a_{23} \\ a_{31} & b_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & c_{12} & a_{13} \\ a_{21} & c_{22} & a_{23} \\ a_{31} & c_{32} & a_{33} \end{vmatrix}$$

T/c 8: *Nếu A có 1 hàng (cột) là tổ hợp t.tính của các hàng (cột) khác thì:* $detA = 0$

Vậy khi nhân 1 hàng (cột) với số t bất kì rồi cộng vào 1 hàng (cột) khác thì det không đổi.

T/c 9: *Nếu A là m.t tam giác thì*: $\mathbf{det}(A) = a_{11}.a_{22}.....a_{nn}$

Sơ đồ tinh định thức của ma trận vuông A:

- (1) Nếu A là mt tam giác thì $\mathbf{det}(A) = a_{11}.a_{22}.....a_{nn}$
- (2) Nếu A có dấu hiệu của t/ c 2, 6, 8 thì detA = 0
- (3) Nếu không phải (1), (2) thì dùng các phép biến đổi không làm thay đổi định thức để đưa mt A về ma trận tam giác A*
- Suy ra detA = detA*

T/c 10: $$\mathbf{det}(A.B) = \boldsymbol{det}(\boldsymbol{A}).\boldsymbol{det}(\boldsymbol{B}). \qquad (1.4)$$

Do đó: $$\det(A^m) = \{\det(A)\}^m$$

T/c 11: *Đ.kiện cần và đủ để m.trận vuông A khả nghịch là* $\boldsymbol{det}(A) \neq \mathbf{0}$. *Khi đó:*

$$A^{-1} = \frac{1}{det(A)}.\begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \dots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

với: A_{ij} *là phần phụ đại số của phần tử* a_{ij} *trong m.trận A.*

VD: K/sát tinh khả nghịch và tìm ma trận nghịch đảo (nếu có) của

$$\mathrm{A} = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 2 & 1 \\ 2 & 3 & 0 \end{bmatrix}$$

Giải: Có $det(A) = 2 \neq 0$, nên A khả nghịch, A^{-1} tìm theo công thức:

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}$$

Có $A_{11} = -3, A_{12} = 2;\ A_{13} = -1$**,** $A_{21} = -3;\ A_{22} = 2,\ A_{23} = 1, A_{31} = 4,$ $A_{32} = -2; A_{33} = 0$. Vậy $A^{-1} = \frac{1}{2}\begin{bmatrix} -3 & -3 & 4 \\ 2 & 2 & -2 \\ -1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -3/2 & -3/2 & 2 \\ 1 & 1 & -1 \\ -1/2 & 1/2 & 0 \end{bmatrix}$

1.3. Hạng của m.trận

1.3.1. Đ/n: Cho m.trận $\mathrm{A} = \left(a_{ij}\right)_{m\times n}$. Hạng của A là số, kí hiệu: rank(A), là cấp cao nhất của các m.trận vuông con có đ.thức khác 0 của A.

1.3.2. NX: $\boldsymbol{r}\mathrm{ank}(\boldsymbol{A}) \le \boldsymbol{min}\{\boldsymbol{m},\boldsymbol{n}\}$**.**

Vd 1: Tìm hạng của: $\mathrm{A} = \begin{pmatrix} 1 & 2 & 3 & 3 \\ 5 & 6 & 7 & 8 \\ 4 & 4 & 4 & 5 \end{pmatrix}$

1.3.3. Tìm hạng của m.trận

a. Đ.lí: (1) Các phép biến đổi sau không làm thay đổi hạng của m.trận:

a. Các phép biến đổi sơ cấp: đổi chỗ 2 hàng (cột), nhân 1 hàng (cột) với số $\alpha \neq 0$, nhân 1 hàng (1 cột) với thừa số λ bất kì rồi cộng vào 1 hàng (1 cột) khác).

b. Phép chuyển vị.

(2) M.trận bậc thang có hạng bằng số hàng khác không của nó.

b. Tìm hạng của m.trận: Dùng các phép b.đổi trên, đưa A về dạng bậc thang A^*, khi đó: $r(A) = r(A^*) =$ số hàng khác 0 của A^*.

Vd 2: Tìm hạng của m.trận: $\begin{pmatrix} 1 & 2 & 3 & 1 & 2 \\ 2 & 3 & 1 & 2 & 3 \\ 3 & 1 & 2 & 3 & 1 \\ 3 & 5 & 4 & 3 & 5 \end{pmatrix}$

Giải: Có sơ đồ biến đổi: $\mathrm{A} = \begin{pmatrix} 1 & 2 & 3 & 1 & 2 \\ 2 & 3 & 1 & 2 & 3 \\ 3 & 1 & 2 & 3 & 1 \\ 3 & 5 & 4 & 3 & 5 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 1 & 2 \\ 0 & -1 & -5 & 0 & -1 \\ 0 & -4 & -2 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

$\to \begin{pmatrix} 1 & 2 & 3 & 1 & 2 \\ 0 & -1 & -5 & 0 & -1 \\ 0 & 0 & 18 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \mathrm{A}^*$ (bậc thang). Vậy $\mathrm{r(A) = r(A^*) = 3}$

c. Các tính chất

(1) $\mathrm{rank(A.B)} \leq \min\{\mathrm{rank(A), rank(B)}\}$

(2) $\mathrm{rank(A + B)} \leq \mathrm{rank(A)} + \mathrm{rank(B)}$

(3) $\forall \mathrm{A} \in \mathbb{R}^{\mathrm{m+n}}, \forall \mathrm{B} \in \mathbb{R}^{\mathrm{n+k}}$, thì: $\mathrm{rank(A)} + \mathrm{rank(B)} - \mathrm{n} \leq \mathrm{rank(A.B)}$

(4) Nếu A là m.trận vuông cấp n thì:

$$\text{A khả nghịch} \Leftrightarrow \det(\mathrm{A}) \neq 0 \Leftrightarrow \mathrm{rank(A)} = \mathrm{n}$$

VD. $\mathrm{A} = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 2 & 1 \\ 2 & 3 & 0 \end{bmatrix}$, có $\det(\mathrm{A}) = 2 \neq 0$, suy ra $\mathrm{rank(A)} = 3$

1.4. Thuật toán Gauss – Jordan khảo sát và tìm ma trận nghịch đảo

Với m.trận vuông: $\boldsymbol{A} = \left(\boldsymbol{a_{ij}}\right)_{\boldsymbol{n\times n}}$, ta thực hiện các bước:

B1. Dùng các phép b.đổi sơ cấp trên các hàng, đưa m.trận ghép $(A|I)$ về dạng: $(A^*|B^*)$, trong đó A^* là m.trận tam giác trên, khi đó:

- Nếu A^* có 1 p.tử đường chéo là 0, thì k.luận A không khả nghịch. Kết thúc.
- Nếu mọi p.tử đường chéo của A^* đều $\neq$ 0, thì k.luận A khả nghịch và chuyển sang B2.

B2. Dùng b.đổi trên, đưa $(A^*|B^*)$ về dạng: $(I|B)$, suy ra: $A^{-1} = B$

<u>*Vd 1*</u>: K/s tính khả nghịch và tìm m.trận nghịch đảo (nếu có) của:

$$A = \begin{bmatrix} -1 & 0 & 1 & -1 \\ 2 & 1 & -3 & 4 \\ 1 & 2 & -2 & 8 \\ -2 & 0 & 2 & 1 \end{bmatrix}; B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 3 & 5 \\ 3 & 3 & 4 & 5 \\ 1 & 1 & 1 & 2 \end{bmatrix};$$

<u>*Vd 2*</u>(Bt): Với m.trận A, B trong vd 1, hãy tìm m.trận X sao cho: $(X - A).A^T = B$

<u>*Vd 4*</u> (Bt): Cho A là m.trận đường chéo. Tìm A^k

<u>*Vd 6*</u> (Bt): Tìm đ/k cho m để m.trận $A = \begin{bmatrix} m+1 & 1 & 3 \\ 2 & m+2 & 0 \\ 2m & 1 & 3 \end{bmatrix}$ khả nghịch

Chương 2. Không gian véc tơ và ánh xạ tuyến tính

2.1. Không gian véc tơ (KGVT)

2.1.1. Các k/n về không gian véc tơ

a. Đ/n KGVT. Cho tập X $\neq \emptyset$, trên đó x.định 2 phép toán: phép cộng hai p.tử của X (kí hiệu: (+)) và phép nhân p.tử của X với một số (thực hoặc phức) (kí hiệu: (.)), thỏa các đ/k sau:

(1) $\forall x, y \in X$ thì: $x + y \in X$ và: $y + x = x + y$.

(2) $\forall x, y, z \in X$ thì: $(x + y) + z = x + (y + z)$(do đó viết: $x + y + z$)

(3) Tồn tại p.tử ∈ X, kí hiệu: **0** (gọi là p.tử không), sao cho: $x + \mathbf{0} = x, \forall x \in X$.

(4) $\forall x \in X, \exists$ p.tử $\in X$, kí hiệu: $-x$, gọi là p.tử đối của x, sao cho: $x + (-x) = \mathbf{0}$.

(5) $\forall x \in X$, với mọi số λ, thì: $\lambda . x \in X$; $1. x = x$,

(6) $\alpha.(\beta.x) = (\alpha\beta).x, \forall x \in X, \forall$số α, β.

(7) $\alpha.(x + y) = \alpha.x + \alpha.y, \forall x, y \in X, \forall$số α.

(8) $(\alpha + \beta).x = \alpha.x + \beta.x, \forall x \in X, \forall$số α, β.

Khi đó X cùng với 2 phép toán này được gọi là một k.gian véc tơ (hay k.gian t.tính). Mỗi p.tử của X gọi là một véc tơ (hay một điểm)

- Nếu phép nhân chỉ thực hiện với số thực thì X gọi là k.gian véc tơ thực. Nếu phép nhân thực hiện với số phức thì X gọi là k.gian véc tơ phức. Trong tài liệu này ta chỉ xét k.gian véc tơ thực.

b. Các ví dụ:

<u>*Vd 1*</u>: Tập: $\mathbb{R}^n = \{x = (x_1, x_2, \dots, x_n): x_1, x_2, \dots, x_n \in \mathbb{R}\}$, với phép cộng (+) và phép nhân với số thực:

$\forall x, y \in \mathbb{R}^n: x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n); \alpha \in \mathbb{R}$:

$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n);\ \alpha . x = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$

là một không gian véc tơ (véc tơ **0** =(0, 0,..., 0))

<u>*Vd 2*</u>: Với phép cộng 2 m.trận và phép nhân m.trận với số thực (phức), $\mathrm{R}^{\mathrm{m}\times\mathrm{n}}$ (hoặc $\mathbb{C}^{\mathrm{m}\times\mathrm{n}}$) trở thành k.gian véc tơ (véc tơ **0** = $\mathcal{O}_{m\times n}$) .

<u>*Vd 3*</u>: $C_{[a,\, b]}$ là tập tất cả các hàm số liên tục trên [a, b], với phép cộng hai hàm số, phép nhân hàm số với số thực, trở thành một k.gian véc tơ (véc tơ **0** là hàm đồng nhất = 0 trên [a, b].

<u>*Vd 4*</u>: $P_n[x]$ là tập tất cả các đa thức của biến x, có bậc không quá n, với phép cộng các h.số, phép nhân h.số với số thực, là k.gian véc tơ.

Trong chương này , chúng ta đặc biệt khảo sát một số nội dung quan trọng liên quan đến các không gian véc tơ: $\mathbb{R}^n$, $\mathrm{R}^{\mathrm{m}\times\mathrm{n}}$ (hoặc $\mathbb{C}^{\mathrm{m}\times\mathrm{n}}$).

***2.1.2.* Hệ véc tơ độc lập tuyến tính, phụ thuộc tuyến tính**

Trong không gian véc tơ X trên trường số K (thực hoặc phức), cho họ véc tơ $\mathrm{S} = \subset \mathrm{X}, \mathrm{S} \neq \emptyset$ và m số $\mathrm{x}_1, x_2, \dots, x_{\mathrm{m}} \in \mathrm{K}$.

- Véc tơ: $\mathrm{u} = \mathrm{x}_1 . \mathrm{u}_1 + x_2 . \mathrm{u}_2 , \dots + x_{\mathrm{m}} . \mathrm{u}_{\mathrm{m}}$ (m = 1,2,...) được gọi là ***một tổ hợp tuyến tính*** (hữu hạn) của các véc tơ của S.

- *Bao tuyến tính của hệ véc tơ* S, kí hiệu là Spam (S), là tập tất cả các tổ hợp tuyến tính các véc tơ của S:

Spam (S) = $\{\mathrm{u} \in \mathrm{X}: \text{ u là một tổ hợp tuyến tính các véc tơ của S}\}$

NX: Spam (S) là một k.gian con của X và là không gian con nhỏ nhất chứa S. Bởi vậy Spam (S) còn gọi là không gian con sinh bởi S. Nếu Spam (S) = X thí nói S là một hệ sinh của X. Như vậy S là một hệ sinh của Spam (S).

- S được gọi một *hệ véc tơ độc lập tuyến tính*, nếu mọi véc tơ trong hệ đều không thể là tổ hợp tuyến tính của các véc tơ còn lại của hệ, tức là:

$$\mathrm{x}_1.\mathrm{u}_1 + x_2.\mathrm{u}_2, \ldots + x_\mathrm{m}.\mathrm{u}_\mathrm{m} = \mathcal{O} \Leftrightarrow \mathrm{x}_1 = x_2 = \cdots = x_\mathrm{m} = 0$$

- S được gọi một *hệ véc tơ phụ thuộc tuyến tính*, nếu S không độc lập tuyến tính, tức là có một véc tơ trong hệ là tổ hợp tuyến tính của các véc tơ còn lại của hệ, hay là: tồn tại các số $x_1, x_2, \ldots, x_\mathrm{m}$ không đồng thời bằng 0 sao cho:

$$\mathrm{x}_1.\mathrm{u}_1 + x_2.\mathrm{u}_2, \ldots + x_\mathrm{m}.\mathrm{u}_\mathrm{m} = \mathcal{O}$$

- S được gọi một *hệ véc tơ độc lập tuyến tính cực đại trong X, nếu S độc lập tuyến tính và* $\forall \mathrm{x} \in \mathrm{X}, \text{hệ } \mathrm{S}' = \{\mathrm{S}, \mathrm{x}\}$ là hệ phụ thuộc tuyến tính***.***

NX: **(1)** Nếu hệ S có chứa véc tơ không $\mathcal{O}$ thì S phụ thuộc tuyến tính;

(2) Với $\mathrm{B} \subset S$ thì: Nếu S độc lập tuyến tính thì B cũng độc lập tuyến tính, nếu B phụ thuộc tuyến tính thì S cũng phụ thuộc tuyến tính.

(3) Giả sử S và B là các hệ véc tơ độc lập tuyến tính cực đại trong X. Khi đó nếu S gồm m véc tơ thì B cũng gồm m véc tơ, nếu S có vô số véc tơ thì V cũng có vô số véc tơ.

2.1.3. Cơ sở và số chiều của không gian véc tơ

- Hệ véc tơ S được gọi là một cơ sở của kg véc tơ X nếu: **(1)** Spam (S) = X, **(2)** S là hệ độc lập tuyến tính.

Vậy S được gọi là một cơ sở của kg véc tơ X khi và chỉ khi mỗi véc tơ $\mathrm{x} \in \mathrm{X}$ đều có biểu diễn duy nhất dưới dạng một tổ hợp tuyến tính các véc tơ của S:

$$\mathrm{x} = \mathrm{x}_1.\mathrm{u}_1 + x_2.\mathrm{u}_2\,, \ldots + x_{\mathrm{m}}.\mathrm{u}_{\mathrm{m}}\ (\mathrm{m} \in \mathbb{N}^*; \mathrm{u}_1, \ldots, \mathrm{u}_{\mathrm{m}} \in \mathrm{S}; x_1, \ldots, x_{\mathrm{m}} \in \mathrm{K})$$

Khi đó x_{k} được gọi là tọa độ của véc tơ x tương ứng với véc tơ u_{k} trong cơ sở S.

NX: Nếu S và B đều là các cơ sở của X thì đều là các hệ độc lập tuyến tính cực đại trong X, hơn nữa chúng có cùng số lượng các véc tơ.

- Với S là một cơ sở của X, Kí hiệu: $\dim(\mathrm{X}) = \begin{cases} \mathrm{n}, \text{nếu S gồm n véc tơ} \\ +\infty, \text{nếu S có vô số véc tơ} \end{cases}$

Ta gọi $\dim(\mathrm{X})$ là số chiều của không gian véc tơ X. Nếu $\dim(\mathrm{X}) = \mathrm{n}$, ta nói X là kg hữu hạn chiều (n chiều). Nếu $\dim(\mathrm{X}) = +\infty$ ta nói X là kg vô hạn chiều.

- ***Hạng của một hệ véc tơ*** B là số véc tơ độc lập tuyến tính cực đại trong B.

VD. **(1)** $\Phi = \{e_1, e_2, \ldots, e_n\}, (e_j = \big(0, \ldots, 0, \underset{j}{1}, 0, \ldots, 0\big), j = 1,2, \ldots, n)$ là một cơ sở trong $\mathbb{R}^n$***, gọi là cơ sở chính tắc của*** $\mathbb{R}^n$***.*** Vì thế $\mathbb{R}^n$ là kg n chiều

(2) $\Phi = \{\varphi_0, \varphi_1, \ldots, \varphi_n\}$ $(\varphi_0(x) = 1, \varphi_1(x) = x, \ldots, \varphi_n(x) = x^n, \forall x \in \mathbb{R})$ là một cơ sở ***(gọi là cơ sở chính tắc)*** trong $P_n[x]$**.** Vậy $P_n[x]$ là kg n + 1 chiều.

(3) $\wp[x]$ (tập tấ cả các đa thức) và $C_{[a,\, b]}$ là các kg vô hạn chiều

(4) Trong kg véc tơ $\mathrm{R}^{\mathrm{m}\times\mathrm{n}}$, hệ véc tơ

$$\Phi = \left\{\varphi_{\mathrm{ij}} = (\mathrm{a_{rs}})_{\mathrm{m}\times\mathrm{n}}\text{: } \mathrm{a_{rs}} = \begin{cases} 1, \text{khi } \mathrm{r = i, s = j} \\ 0, \text{khi } \mathrm{r \neq i}, \text{hoặc } \mathrm{s \neq j} \end{cases}, \mathrm{i} = \overline{1, \mathrm{m}}\,, \mathrm{j} = \overline{1, \mathrm{n}}\right\},$$

là một cơ sở, gọi là cơ sở chính tắc của $\mathrm{R}^{\mathrm{m}\times\mathrm{n}}$.

2.2. Không gian định chuẩn

Mục này trình bày cách đo "độ lớn", "khoảng cách" giữa các điểm hoặc các tập dữ liệu

2.2.1. Các k/n về không gian định chuẩn- Cho kgvt X trên trường số K,

ĐN1. Mỗi ánh xạ: $\|.\| : \mathrm{X} \to [0, +\infty)$ được gọi là một chuẩn trên kgvt X, nếu thỏa các điều kiện:

(a) $\|x\| = 0 \Leftrightarrow \mathrm{x} = \mathcal{O}$ (véc tơ không);

(b) $\|\alpha.x\| = |\alpha|.\mathrm{d(x)}, \forall \mathrm{x} \in \mathrm{X}, \forall \alpha \in \mathrm{K}$;

(c) $\|x + y\| \le \|x\| + \|y\|, \forall \mathrm{x, y} \in \mathrm{X}$

Khi đó $(\mathrm{X}, \|.\|)$ gọi là một không gian định chuẩn, gọi $\|x\|$ là chuẩn (hay độ dài) của véc tơ x và $\mathrm{d(x, y)} = \|x - y\|$ là khoảng cách giữa hai phần tử x và y.

Véc tơ x có độ dài = 1 gọi là véc tơ đơn vị.

- Với kg. định chuẩn $(\mathrm{X}, \|.\|)$, $\mathrm{a} \in \mathrm{X}, \mathrm{r} > 0$:

$\mathrm{B(a, r)} = \{\mathrm{u} \in \mathrm{X}: \|\mathrm{u - a}\| < \mathrm{r}\}$: Hình cầu mở tâm a, bán kính r

$\overline{\mathrm{B}}\mathrm{(a, r)} = \{\mathrm{u} \in \mathrm{X}: \|\mathrm{u - a}\| \le \mathrm{r}\}$: Hình cầu đóng tâm a, bán kính r

$\mathrm{S(a, r)} = \{\mathrm{u} \in \mathrm{X}: \|\mathrm{u - a}\| = \mathrm{r}\}$: Mặt cầu tâm a, bán kính r

ĐN 2. Mỗi ánh xạ: $\langle .\,,.\rangle: X\times X \to \mathbb{R}$ được gọi là một tích vô hướng trên X, nếu thỏa mãn các đ/k:

(a) $\langle x, y\rangle = \langle y, x\rangle, \forall x, y \in X$

(b) $\langle x, x\rangle \geq 0, \forall x \in X$, và$\langle x, x\rangle = 0 \Leftrightarrow x = \mathcal{O}$

(c) $|\langle x, y\rangle|^2 \leq \langle x, x\rangle + \langle y, y\rangle, \forall x, y \in X$.

Ghi chú: Giả sử X là không gian véc tơ có tích vô hướng $\langle .\,,.\rangle$

(1) Tích vô hướng $\langle .\,,.\rangle$ cảm sinh một chuẩn $\|.\|$ trên X: $\|x\| = \langle x, x\rangle^{1/2}$

và một khoảng cách: $d(x, y) = \|x - y\| = \langle x - y, x - y\rangle^{1/2}$

(2) Nếu $\langle x, y\rangle = 0$, ta nói x và y trực giao nhau, kí hiệu: $x \perp y$.

- Hệ véc tơ S được gọi là ***hệ trực giao*** nếu: $\forall x, y \in X, x \neq y$ thì $x \perp y$.
- Một cơ sở S gồm các véc tơ trực giao nhau gọi là một ***cơ sở trực giao***
- Hệ véc tơ S được gọi là ***hệ trực chuẩn*** nếu S là hệ trực giao, đồng thời mọi véc tơ trong hệ đều là véc tơ đơn vị.
- Một cơ sở gồm các véc tơ trực chuẩn gọi là một ***cơ sở trực chuẩn***

2.1.2. Một số chuẩn thường dùng các không gian $\mathbb{R}^n$ và $R^{m\times n}$

Trên một không gian véc tơ có thể đưa ra nhiều chuẩn khác nhau, tức là nhiều cách đo độ dài véc tơ khác nhau.

a. Chuẩn Euclide trong $\mathbb{R}^{\mathbf{n}}$: Đây là chuẩn thông dụng nhất Trên $\mathbb{R}^n$**,** xây dựng như sau: $\forall x, y \in \mathbb{R}^n: x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n)$

Tích vô hướng: $\langle x, y\rangle = \sum_{i=1}^{n} x_i . y_i$ cảm sinh chuẩn sau gọi là chuẩn Euclide*:*

$$\|x\| = \langle x, x\rangle^{1/2} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \quad (2.1)$$

và khoảng cách Euclide:

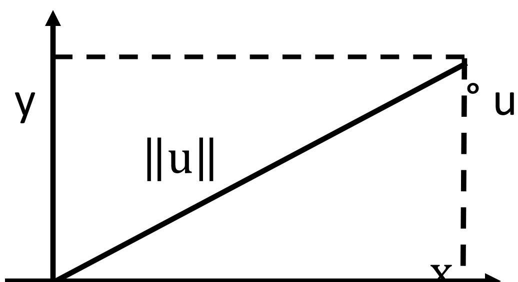
$$d(x,y) = \|x-y\| = \langle x-y, x-y\rangle^{1/2} = \left(\sum_{i=1}^{n}(x_i - y_i)^2\right)^{1/2} \quad (2.2)$$

$\mathbb{R}^n$ với tích vô hướng nói trên được gọi là không gian Euclide n chiều

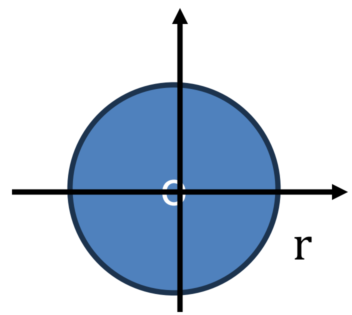
b. Chuẩn Max trong $\mathbb{R}^n$: $\forall x \in \mathbb{R}^n: x = (x_1, x_2, \dots, x_n)$

$$\|x\|_{max} = \max\{|x_i|, i = \overline{1,n}\} \quad (2.3)$$

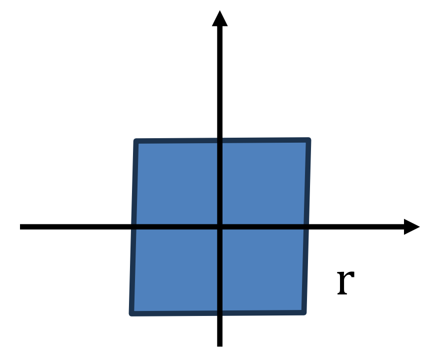
Minh họa trong $\mathbb{R}^2$:



$\|u\|_{max} = \max\{|x|, |y|\} = |x|$



Hình cầu Euclide



Hình cầu Max

c. Chuẩn $\|.\|_p$ (l_p-norm) trong $\mathbb{R}^n$: $\forall x \in \mathbb{R}^n: x = (x_1, x_2, \dots, x_n)$

$$\|x\|_p = \left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p} \ (p > 1) \quad (2.4)$$

NX: (1). $\|x\|_2$ ***chính là chuẩn Euclide***

(2). $\|x\|_\infty = \lim_{p\to+\infty}\|x\|_p = |x_s| = \max\{|x_i|:\ i = \overline{1,n}\} = \|x\|_{max}$

d. Không gian $\mathrm{R}^{m\times n}$ *với chuẩn Frobenius, kí hiệu* $\|.\|_F$*:*

$$\forall A = (a_{ij})_{m\times n} \in \mathrm{R}^{m\times n}:\ \ \|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2} \qquad (2.5)$$

VD: Với $A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & -4 & -1 \end{bmatrix}, \|A\|_F = \sqrt{1+4+9+16+1} = \sqrt{31}$

2.3. Một số bài toán cơ bản trong $\mathbb{R}^n$

2.3.1. Tọa độ véc tơ theo cơ sở: Ta đã biết

$$\Phi = \{e_1, e_2, \dots, e_n\}, (e_j = \left(0, \dots, 0, \underset{j}{1}, 0, \dots, 0\right), j = 1,2, \dots, n)$$

là một cơ sở của $\mathbb{R}^n$ và vì vậy $\mathbb{R}^n$ là kgvt n chiều và mọi cơ sở của $\mathbb{R}^n$ đều gồm n véc tơ.

G.sử $S = \{u_1, u_2, \dots, u_n\}$ là một cơ sở có sắp thứ tự của $\mathbb{R}^n$, khi đó mỗi $x \in \mathbb{R}^n$ có biểu diễn duy nhất: $x = t_1.u_1 + t_2.u_2, \dots + t_n.u_n$

và $t_1, t_2, \dots, t_n$ là ***các tọa độ của véc tơ x trong cơ sở sắp thứ tự S***. Kí hiệu:

$(x)_S = (t_1, t_2, \dots, t_n)$ (gọi là ***dòng hay hàng tọa độ của x trong S***)

$[x]_S = {(x)_S}^T$ (chuyển vị của $(x)_S$) (gọi là ***cột tọa độ của x trong S***)

Trên $\mathbb{R}^n$ với tích vô hướng Euclide, cơ sở Φ là một cơ sở trực chuẩn, còn gọi là cơ sở chính tắc của $\mathbb{R}^n$.

$\forall x \in \mathbb{R}^n: x = (x_1, x_2, \dots, x_n) \Leftrightarrow x = x_1.e_1 + x_2.e_2 + \cdots + x_n.e_n$

Như vậy $(x_1, x_2, \dots, x_n)$ chính là dòng tọa độ của x trong cơ sở chính tắc Φ

2.3.2. Tìm hạng của một hệ véc tơ trong $\mathbb{R}^n$

$\forall u \in \mathbb{R}^n$, để đơn giản ta kí hiệu dòng và cột tọa độ trong cơ sở chính tắc của u là: (u) và [u] tương ứng.

Cho hệ véc tơ $B = \{u_1, u_2, \dots, u_m\} \subset \mathbb{R}^n$, hạng của hệ véc tơ B là số kí hiệu: rank(B) = số véc tơ độc lập tuyến tính cực đại trong B. Dễ thấy nếu m > n thi B phụ thuộc tuyến tính, vì thế:

$0 \le \text{rank}(B) \le n\,;\ \text{rank}(B) = 0 \Leftrightarrow B\ \text{chỉ chứa véc tơ không}\,.$

Gọi: $([u_1]\ \ [u_2] \dots [u_m])$ là ma trận liên kết cột của hệ B, $\big((u_1)\ \ (u_2) \dots (u_m)\big)$ là m.trận liên kết hàng của hệ B. Một trong 2 m.trận này ta kí hiệu là: B^*.

Đ.lý. Hạng của một ma trận A bằng hạng của hệ véc tơ cột (hoặc hàng) của A

Từ đó suy ra: rank (B) = rank (B^*), và như vậy bài toán tìm hạng của hệ véc tơ B chuyển về bài toán tìm hạng của m.trận liên kết (hàng hoặc cột) của B.

VD1. Cho B = {u = (1,0,2,1), v = (2,2,1,2), w =(0,2,0,1), z = (-1,0,1,0)}, tìm rank(B)

Ta tìm sơ đồ biến đổi m.trận l.kết hàng của B:

$$B^* = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & 2 & 1 & 2 \\ 0 & 2 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{bmatrix} \xrightarrow[\text{h4+h1}\to\text{h4}]{\text{h2}-2\text{h1}\to\text{h2}} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 2 & -3 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 3 & 1 \end{bmatrix} \xrightarrow[\text{h4}-(\text{h3}-\text{h2})\to\text{h4}]{\text{h3}-\text{h2}\to\text{h3}} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 2 & -3 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = B^{**}$$

(B^{**} là ma trận bậc thang)

Vậy rank(B) = rank(B^*) = rank (B^{**}) = 3

Chú ý: (1)Hệ véc tơ B là một cơ sở trong $\mathbb{R}^n$ $\Leftrightarrow$ B gồm đúng n véc tơ đ.lập t.tính

(2) Trong các không gian $\mathbb{R}^n, M_{n\times m}(\mathbb{R}), P_n[x]$, khi xét ma trận liên kết của một hệ véc tơ, người ta thường xét trong cơ sở chính tắc và khi nói tới ma trận liên kết của một hệ véc tơ mà không giải thích gì thêm, thì hiểu là xét trong cơ sở chính tắc.

2.3.3. Đổi cơ sở trong $\mathbb{R}^{\mathrm{n}}$

a. Đặt vấn đề. Đổi cơ sở là một vấn đề thường gặp khi ta muốn tìm tọa độ của một véc tơ trong cơ sở mới. Trong k.g.v.t n chiều, cho 2 cơ sở được sắp thứ tự: $S = \{s_1, s_2, \dots, s_n\}$, B $= \{u_1, u_2, \dots, u_n\}$. V/đề đặt ra là: khi biết tọa độ của véc tơ u trong cơ sở S là $[u]_S$, cần tìm tọa độ $[u]_B$ của u trong cơ sở B.

Biểu diễn các véc tơ của S qua B, ta có:

$$s_j = a_{1j}.u_1 + a_{2j}.u_2 + \cdots + a_{nj}.u_n, j = 1,2, \dots, n \qquad (2.6)$$

Ma trận $\mathbf{A} = \left(\boldsymbol{a_{ij}}\right)_{\mathbf{n}\times\boldsymbol{n}} = ([\boldsymbol{s_1}]_{\boldsymbol{B}}, [\boldsymbol{s_2}]_{\boldsymbol{B}}, \dots, [\boldsymbol{s_n}]_{\boldsymbol{B}})$ (cột thứ j của A là cột tọa độ $[\mathrm{s_j}]_{\mathrm{B}}$ của véc tơ $\mathrm{s_j}$ trong cơ sở B), gọi là **ma trận chuyển từ B sang S. Còn kí hiệu : $\mathbf{A} = \boldsymbol{P}_{(\mathbf{B}\to\mathbf{S})}$.**

A là ma trận khả nghịch (vì r(A) = rank S = n). Khi đó với mọi véc tơ u, ta có:

$$u = \sum_{j=1}^{n} x_j s_j = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}x_j\right).u_i \qquad (2.7)$$

Từ đó cho thấy: $\quad [\boldsymbol{u}]_{\boldsymbol{B}} = \boldsymbol{A}.[\boldsymbol{u}]_{\boldsymbol{S}}$ (2.8)

Từ (2.8) ta có: $\quad [\boldsymbol{u}]_{\boldsymbol{S}} = \boldsymbol{A}^{-\mathbf{1}}.[\boldsymbol{u}]_{\boldsymbol{B}}$ (2.9)

Điều này cho thấy, nếu $\mathbf{A} = \boldsymbol{P}_{(\mathbf{B}\to\mathbf{S})}$**, thì** $\boldsymbol{A}^{-\mathbf{1}} = \mathbf{P}_{(\mathbf{S}\to\mathbf{B})}$ (2.10)

b. Thuật toán tìm ma trận chuyển cơ sở trong $\mathbb{R}^n$:

$$(\boldsymbol{B}^*|\boldsymbol{S}^*) \xrightarrow{\textit{các phép biến đổi sơ cấp trên các hàng}} (\boldsymbol{I}|\boldsymbol{A}).$$ **Suy ra:** $\boldsymbol{P}_{(\mathbf{B}\to\mathbf{S})} = \boldsymbol{A}$.

VD2. Trong $\mathbb{R}^3$ cho: B = {v_1= (1,0,1), v_2= (1, 1, 0), v_3= (0, 1, 1)};

S= {s_1 = (1, 2, 3), s_2 = (0, 1, 0), s_3 = (0, 0, 1)}

(1) C/m S và B là các cơ sở của $\mathbb{R}^3$; (2) Tìm ma trận chuyển từ B sang S; (3) Tìm tọa độ của véc tơ: u = 2. s_1 – 3s_2 + 5.s_3 trong cơ sở B.

Giải: 1/ Có $\det(S^*) = 1 \neq 0, \det(B^*) = 2 \neq 0$, nên S, B là các cơ sở của $\mathbb{R}^3$. Có sơ đồ biến đổi:

$$(B^*|S^*) = \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 3 & 0 & 1 \end{array}\right) \xrightarrow[\text{h3}-\text{h1}+\text{h2}\to\text{h3}]{\text{h2}-\text{h3}-\text{h1}\to\text{h2}} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & -1 \\ 0 & 0 & 2 & 4 & 1 & 1 \end{array}\right)$$

$$\xrightarrow{2h1-h2\to h1} \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 2 & -1 & 1 \\ 0 & 2 & 0 & 0 & 1 & -1 \\ 0 & 0 & 2 & 4 & 1 & 1 \end{array}\right) \xrightarrow{\substack{\frac{1}{2}h1\to h1 \\ \frac{1}{2}h2\to h2;\ \frac{1}{2}h3\to h3}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -0{,}5 & 0{,}5 \\ 0 & 1 & 0 & 0 & 0{,}5 & -0{,}5 \\ 0 & 0 & 1 & 2 & 0{,}5 & 0{,}5 \end{array}\right)$$

$$\Rightarrow P_{(B\to S)} = A = \begin{pmatrix} 1 & -0{,}5 & 0{,}5 \\ 0 & 0{,}5 & -0{,}5 \\ 2 & 0{,}5 & 0{,}5 \end{pmatrix}.$$

2/ Ta có: $[u]_B = A.[u]_S = \begin{pmatrix} 1 & -0,5 & 0,5 \\ 0 & 0,5 & -0,5 \\ 2 & 0,5 & 0,5 \end{pmatrix}.\begin{pmatrix} 2 \\ -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 6 \\ -4 \\ 1 \end{pmatrix}$

VD3 (BT). Trong $\mathbb{R}^4$ cho:

B = {v_1= (1,1,1,1), v_2= (0, 1, 1,1), v_3= (1, 0,1,1), v_4 = (1,1,0,1)};

S = {s_1 = (1, 1, 0, 0), s_2 = (0, 1, 1, 0), s_3 = (0, 0,1, 1), s_4 = (1, 0, 0, 0)}

(1) C/m S và B là các cơ sở của $\mathbb{R}^4$; (2)Tìm ma trận chuyển từ B sang S.

(3) Tìm tọa độ của véc tơ: u trong cơ sở B, biết $(u)_S = (1, -2, 3, -4)$.

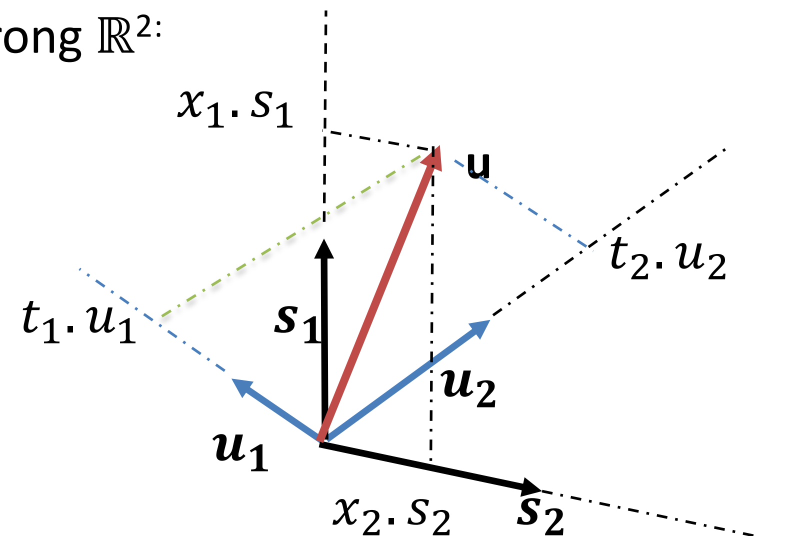
c. Minh họa hình học.

Để dễ hình dung, xét 2 cơ sở

$S = \{s_1, s_2\}$, B $= \{u_1, u_2\}$ trong $\mathbb{R}^2$:

$u = x_1.s_1, +x_2.s_2$

$= t_1.u_1, +t_2.u_2$



2.3.4. Quá trình trực giao hóa của Gram-Smidt

Cho V là kgvt có tích vô hướng, $B = \{v_1, v_2, \dots, v_n\}$ là họ véc tơ độc lập tuyến tính của V. Từ B, cần xây dựng một họ véc tơ trực chuẩn $S = \{u_1, u_2, \dots, u_n\}$ sao cho span(S) = span(B). Quá trình gồm các bước:

B1: Chọn: $\mathrm{u}_1 = \frac{1}{\|\mathrm{v}_1\|}.\mathrm{v}_1$, ta có: $\|\mathrm{u}_1\| = 1$ và: $\mathrm{span}\{\mathrm{u}_1\} = \mathrm{span}\{\mathrm{v}_1\}$

B2: Tìm u_2 để họ $\{\mathrm{u}_1, \mathrm{u}_2\}$ trực chuẩn, bằng cách đặt: $\mathrm{u}_2' = \mathrm{v}_2 + \mathrm{tu}_1$, và chọn t từ: $\langle \mathrm{u}_2', \mathrm{u}_1\rangle = 0 \Leftrightarrow \mathrm{t} = -\langle \mathrm{v}_2, \mathrm{u}_1\rangle$. Đặt: $\mathrm{u}_2 = \frac{1}{\|\mathrm{u}_2'\|}.\mathrm{u}'_2$, thì $\{\mathrm{u}_1, \mathrm{u}_2\}$ là họ trực chuẩn và ta có: $\mathrm{span}(\{\mathrm{u}_1, \mathrm{u}_2\}) = \mathrm{span}(\{\mathrm{v}_1, \mathrm{v}_2\})$.

B3: *G*iả sử đã xây dựng được họ: $S_{k-1} = \{u_1, u_2, \dots, u_{k-1}\}$ là họ trực chuẩn sao cho: $span(S_i) = span(\{v_1, v_2, \dots, v_i\}), \forall i = \overline{1, k-1}$. Khi đó ta tiếp tục xây dựng: $u_k' = v_k + t_1 u_1 + t_2 u_2 + \cdots + t_{k-1} u_{k-1}$,

Với t_j tìm được từ đ/k: $\langle u_k', u_j\rangle = 0, hay\ t_j = -\langle v_k, u_j\rangle, \forall j = 1, \overline{k-1}$.

Từ đó chọn: $u_k = \frac{1}{\|u_k'\|}.u_k'$, và nhận được: $S_k = \{u_1, u_2, \dots, u_k\}$ là họ trực chuẩn mà: $span(S_k) = span(\{v_1, v_2, \dots, v_k\})$.

Quá trình xây dựng họ véc tơ trực chuẩn $S = \{u_1, u, \dots, u_n\}$ nói trên được gọi là quá trình trực chuẩn hóa Gram- Smidt.

Như vậy bằng quy nạp, ta đã chứng minh được kết quả sau:

Đ.lí 4: Với $B=\{v_1,v_2,\ldots,v_n\}$ là họ véc tơ độc lập tuyến tính trong không gian véc tơ thực V, ta tìm được họ véc tơ trực chuẩn $S=\{u_1,u_2,\ldots,u_n\}$ sao cho: $span(\{u_1,u_2,\ldots,u_k\})=span(\{v_1,v_2,\ldots,v_k\}), với\ mọi\ k=1,2,\ldots,n$.

VD7: Trong $\mathbb{R}^3$,bằng quá trình Gram-Smidt hãy trực chuẩn hóa hệ véc tơ:

$B=\{v_1=(1,1,1),v_2=(1,1,0),v_3=(0,1,1)\}$

Giải: Dễ thấy B là họ véc tơ độc lập tuyến tính trong $\mathbb{R}^3$.

Đặt $u_1=\frac{1}{\|v_1\|}.v_1=\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)$; $u_2'=v_2+tu_1,$ *với* $t=-\langle v_2,u_1\rangle=-\frac{2}{\sqrt{3}}\Rightarrow u_2'=\left(\frac{1}{3},\frac{1}{3},-\frac{2}{3}\right)$. *Chọn:* $u_2=\frac{1}{\|u_2'\|}.u_2'=\left(\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}},-\frac{2}{\sqrt{6}}\right)$.

Đặt: $u_3'=v_3+au_1+bu_2, với: a=-\langle v_3,u_1\rangle=-\dfrac{2}{\sqrt{3}}; b=-\langle v_3,u_2\rangle=\dfrac{1}{\sqrt{6}}$

$\Rightarrow u_3'=\left(-\frac{1}{2},\frac{1}{2},0\right)$. Chọn $u_3=\frac{1}{\|u_3'\|}.u_3'=\left(-\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},0\right)$. Hệ cần tím là:

$$S=\left\{u_1=\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right),u_2=\left(\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}},-\frac{2}{\sqrt{6}}\right),u_3=\left(-\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},0\right)\right\}.$$

VD8(BT): Trong $\mathbb{R}^3$,bằng quá trình Gram-Smidt hãy trực chuẩn hóa hệ véc tơ:

$B=\{v_1=(1,-1,1),v_2=(-1,1,1),v_3=(1,1,-1)\}$.

2.4. Ánh xạ tuyến tính

2.4.1. Các khái niệm về ánh xạ tuyến tính

ĐN1. Cho X, Y là các kgvt trên cùng một trường số K (thực hoặc phức)
Mỗi ánh xạ $\mathrm{f}: \mathrm{X} \rightarrow \mathrm{Y}$ được gọi là một ánh xạ tuyến tính (hay một đồng cấu tuyến tính) từ X sang Y, nếu có các tính chất:

a/ Tính cộng tính: $\mathrm{f}(\mathrm{u}+\mathrm{v})=\mathrm{f}(\mathrm{u})+\mathrm{f}(\mathrm{v}), \forall \mathrm{u}, \mathrm{v} \in \mathrm{X}$

b/ Tính thuần nhất: $\mathrm{f}(\alpha . \mathrm{u})=\alpha . \mathrm{f}(\mathrm{u}), \forall \alpha \in \mathrm{K}, \forall \mathrm{u} \in \mathrm{X}$,

- Với ánh xạ tuyến tính $\mathrm{f}: \mathrm{X} \rightarrow \mathrm{Y}$, ký hiệu:

$\mathrm{Im}(\mathrm{f})=\{\mathrm{f}(\mathrm{u}): \mathrm{u} \in \mathrm{X}\}$: gọi là ảnh của f

$\mathrm{Ker}(\mathrm{f})=\{\mathrm{u} \in \mathrm{X}: \mathrm{f}(\mathrm{u})=\mathcal{O}_{\mathrm{Y}}$ (véc tơ không của Y$\}$: Gọi là nhân của f.

NX. $\mathrm{f}: \mathrm{X} \rightarrow \mathrm{Y}$ là ánh xạ tuyến tính khi và chỉ khi:

$$\mathrm{f}(\alpha . \mathrm{u}+\beta . \mathrm{v})=\alpha . \mathrm{f}(\mathrm{u})+\beta . \mathrm{f}(\mathrm{v}), \forall \mathrm{u}, \mathrm{v} \in \mathrm{X}, \forall \alpha, \beta \in \mathrm{K}$$

Định lý 1: Giả sử f: X →Y là ánh xạ tuyến tính. Khi đó:

1/ $f(\mathcal{O}_X) = \mathcal{O}_Y$.

2/ Imf là không gian con của X, Kerf là không gian con của Y.

3/ f là toàn ánh $\Leftrightarrow \mathrm{Im}(\mathrm{f})=\mathrm{Y}$; f là đơn ánh $\Leftrightarrow \mathrm{Ker}(\mathrm{f})=\{\mathcal{O}_{\mathrm{X}}\}$

4/ Nếu S là hệ véc tơ độc lập tuyến tính trong X thì f(S) là hệ véc tơ độc lập tuyến tính trong Imf

5/ Nếu X là không gian véc tơ n chiều thì: dim Imf + dim Kerf = n.

Ta gọi dim(Imf) là hạng của ánh xạ tuyến tính f, dim Kerf là số khuyết của f.

VD1. f : $R^{m\times n} \to R^{n\times m}, f(A) = f(A^T)$ (phép chuyển vị) là một ánh xạ truyến tính từ $R^{m\times n}$ sang $R^{n\times m}$. f là một song ánh, vì $\mathrm{Im(f)} = R^{n\times m}$ (toàn ánh) và $\mathrm{Ker(f)} = \{\mathrm{A} \in \mathrm{R^{m\times n}}\text{: } \mathrm{A^T} = \mathcal{O}_{\mathrm{n\times m}}\} = \mathcal{O}_{\mathrm{m\times n}}$ (đơn ánh)

Suy ra: hạng của f = dim(Im(f)) = dim($R^{n\times m}$) = m.n, và f có số khuyết = 0.

VD2 (BT). Cho f : $\mathrm{R^2} \to \mathrm{R^3}\text{: } \mathrm{f(x,y) = (x, y, x+y)}$

a. C/m f là ánh xạ tuyến tínhb. Tìm Im(f) và Ker(f). Tìm hạng và số khuyết của f.

ĐN2. Mỗi ánh xạ tuyến tính f : X → X từ kgvt X vào chính nó gọi là một phép biến đổi tuyến tính trong X.

VD3. f : $\mathrm{R^2} \to \mathrm{R^2}\text{: } \mathrm{f(x,y) = (y,x)}, \forall \mathrm{(x,y)} \in \mathrm{R^2}$ (phép đối xứng qua đường phân giác thứ nhất) là một phép biến đổi tuyến tính trong $\mathrm{R^2}$ và dễ thấy là một song ánh.

VD4. f : $R^{\mathrm{n}\times n} \to R^{n\times \mathrm{n}}\text{: } \mathrm{f(A) = A^*}, \forall \mathrm{A} \in R^{n\times \mathrm{n}}$ ($\mathrm{A^*}$ là m.trận đường chéo với đường chéo là đường chéo của A). Dễ dàng kiểm chứng f là một phép biến đổi tuyến tính trong $R^{n\times \mathrm{n}}$ có hạng(f) = dim(Im(f)) = n và số khuyết là: dim(Ker(f)) = n.n – n = (n – 1).n

VD5. Xét: $f\text{: } P_n[x] \to P_n[x]\text{: } f(p) = p'(phép\ lấy\ \text{đạo}\ hàm), \forall p \in P_n[x]$. Dễ thấy f là phép biến đổi tuyến tính trong $P_n[x]$ với Im(f) = $P_{n-1}[x]$, Ker(f) = $P_0[x]$ (các đa thức bậc 0), có hạng(f) = dim($P_{n-1}[x]$) = n và số khuyết = 1

VD6 (BT). Cho $A = \left(a_{ij}\right)_{m\times n}$, c/m: $f\text{: } \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{m}}\text{: } [f(x)] = A.[x]$, ($[x]$ là cột tọa độ của x trong cơ sở chính tắc) là ánh xạ tuyến tính. Hãy mô tả Ker(f)

ĐN3. Giả sử X, Y là các kgvt trên cùng trường số K. Mỗi ánh xạ tuyến tính từ X sang Y được gọi là một đơn cấu nếu f đơn ánh, và gọi là một toàn cấu nếu f toàn ánh. Nếu f là một song ánh thì được gọi là một đẳng cấu, khi đó X và Y được gọi là hai kgvt đẳng cấu với nhau.

Đ.lí 1. Hai kgvt hữu hạn chiều X, Y trên cùng một trường K là đẳng cấu với nhau khi và chỉ khi chúng có cùng số chiều (dim(X) = dim(Y)).

Nhờ vào kết quả trên mà đ/v các phép biến đổi tuyến tính trên các kgvt hữu hạn chiều, chỉ cần k/sát trên không gian $\mathbb{R}^{\mathrm{n}}$.

2.4.2. Ánh xạ tuyến tính trong không gian hữu hạn chiều

Giả sử X là không gian véc tơ n chiều có cơ sở B = $\{u_1, u_2, \ldots, u_n\}$, Y là không gian véc tơ m chiều có cơ sở S = $\{s_1, s_2, \ldots, s_m\}$, $\mathrm{f} : \mathrm{X} \to \mathrm{Y}$ là ánh xạ tuyến tính. Vì mọi véc tơ u ∈ X đều có biểu diễn duy nhất bằng tổ hợp tuyến tính của các véc tơ của B nên f hoàn toàn được xác định bởi ảnh của các véc tơ của B là $f(u_1), f(u_2), \ldots, f(u_n)$, tức là hoàn toàn được xác định bởi ma trận

$$A = ([f(u_1)]_S \quad [f(u_2)]_S \quad \ldots \quad [f(u_n)]_S)_{m\times n}$$

(cột thứ j của A là cột tọa độ $[f(u_j)]_S$ của véc tơ $f(u_j)$ trong cơ sở S).

ĐN4. Gọi $\boldsymbol{A} = ([\boldsymbol{f}(\boldsymbol{u_1})]_{\boldsymbol{S}} \quad [\boldsymbol{f}(\boldsymbol{u_2})]_{\boldsymbol{S}} \quad \ldots \quad [\boldsymbol{f}(\boldsymbol{u_n})]_{\boldsymbol{S}})_{\boldsymbol{m\times n}}$ **(2.11)**

là m.trận của ánh xạ tuyến tính f trong cặp cơ sở B – S.

Với mỗi ảnh $f(u_j)$ biểu diễn trong cơ sở S:

$$f(u_j) = \sum_{i=1}^{m} a_{\mathrm{ij}}\,.\,v_i, j = \overline{1,n} \Rightarrow A = \left(a_{\mathrm{ij}}\right)_{m\times n}$$

thì: $\forall u = \sum_{j=1}^{n} x_j.\,u_j \in \mathrm{X}$: $f(u) = f(\sum_{j=1}^{n} x_j.\,u_j) = \sum_{j=1}^{n} x_j.\,f(\,u_j)$

$$= \sum_{j=1}^{n} x_j.\sum_{i=1}^{m} a_{\mathrm{ij}}\,.\,v_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{\mathrm{ij}}\,x_j\right).\,v_i$$

Suy ra: $$[f(u)]_S = A.\,[u]_B \qquad (2.12)$$

Nhận xét:

a. Có sự tương ứng 1 – 1 các ma trận $\mathrm{R}^{\mathrm{m\times n}}$ và các ánh xạ tuyến tính từ không gian n chiều X sang không giam m chiều Y (trên cùng một trường số) khi cố định cặp cơ sở B – S: Với mỗi ánh xạ tuyến tính $\mathrm{f} : \mathrm{X} \to \mathrm{Y}$, ta có một ma trận của f và với mỗi ma trận $A \in \mathrm{R}^{\mathrm{m\times n}}$, có một ánh xạ tuyến tính $\mathrm{f} : \mathrm{X} \to \mathrm{Y}$ nhận A làm ma trận của nó trong cặp cơ sở B – S: $[f(u)]_S = A.\,[u]_B, \forall \mathrm{u} \in \mathrm{X}$

b. Hạng của ma trận A của ánh xạ tuyến tính f trong cặp cơ sở B – S bằng số chiều của Imf.

VD *8*: Cho ánh xạ f: $\mathbb{R}^3 \to \mathbb{R}^4$, f(x, y, z) = (2x, x + y, x – y , x + z)

(1) Chứng minh f là ánh xạ tuyến tính; (2) Tìm Imf, Kerf; (3)Tìm ma trận A của f trong cặp cơ sở chính tắc của $\mathbb{R}^3$ và $\mathbb{R}^4$.

Giải: (1) (S/v tự kiểm chứng)

2/ $Kerf = \{(x, y, z) \in \mathbb{R}^3$: (2x, x + y, x – y , x + z) $= (0,0,0,0)\} = \{(0,0,0)\}$

Cơ sở chính tắc của $\mathbb{R}^3$ là $\Phi_3 = \{e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)\}$. Ta có:

$f(e_1) = (2, 1, 1, 1)$, $f(e_2) = (0, 1, -1, 0)$, $f(e_3) = (0, 0, 0, 1)$

$Imf = Span\{f(e_1), f(e_2), f(e_3)\} = Span\{(2,1,1,1), (0,1,-1,0), (0,0,0,1)\}$.

Imf là kg con ba chiều của $\mathbb{R}^4$, nhận $\{f(e_1), f(e_2), f(e_3)\}$ làm một cơ sở.

3/ Cơ sở chính tắc của $\mathbb{R}^4$ là $\Phi_4 = \{u_1 = (1,0,0,0), u_2 = (0,1,0,0), u_3 = (0,0,0,1),$ $u_4 = (0,0,0,1)\}$.

Ta có: $f(e_1) = (2, 1, 1, 1) = 2.u_1 + u_2 + u_3 + u_4$; $f(e_2) = (0, 1, -1, 0) = 0.u_1 + 1.u_2 - 1.u_3 + 0.u_4$; $f(e_3) = (0, 0, 0, 1) = 0.u_1 + 0.u_2 + 0.u_3 + u_4$;

Ma trận của f trong cặp cơ sở chính tắc của $\mathbb{R}^3$ và $\mathbb{R}^4$ là: $A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$

Ma trận A có hạng bằng 3 (số chiều của Imf).

Từ định nghĩa, dễ dàng nhận được kết quả sau:

Đ.lý 2. Giả sử f, g là các ánh xạ tuyến tính từ kgvt n chiều X sang kgvt m chiều Y và có các ma trận A và G tương ứng trên cặp cơ sở B – S.

Khi đó các ánh xạ tuyến tính: $a.f\,(a \in K)$ và $f + g$ có ma trận tương ứng là $a.A$ và $A + G$ trong cặp cơ sở B – S.

Chú ý: (1) Với phép biến đổi tuyến tính $f: \mathrm{X} \longrightarrow \mathrm{X}$, và B là một cơ cở sắp thứ tự của X, m.trận của f trong cặp B – B gọi là m.trận của f trong cơ sở B.
(2) Nếu $f: \mathbb{R}^{\mathrm{n}} \longrightarrow \mathbb{R}^{\mathrm{m}}$ là ánh xạ tuyến tính, thì ma trận của f trong cặp cơ sở chính tắc của $\mathbb{R}^{\mathrm{n}}$ và của $\mathbb{R}^{\mathrm{m}}$ được gọi đơn giản là ma trận chính tắc của f

VD 9: Giả sử ánh xạ tuyến tính $f: \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ có ma trận chính tắc $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$ và ánh xạ tuyến tính $h: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ có ma trận chính tắc $H = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \end{pmatrix}$. Hãy tìm $h \circ f(2, -1)$

Giải: Ánh xạ $h \circ f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ là ánh xạ tuyến tính có ma trận chính tắc là:

$$C = H.A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \end{pmatrix} . \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$

Vậy : $\big(h \circ f(2, -1)\big)^T = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} . \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$, hay: $h \circ f(2, -1) = (3, -1)$.

2.4.3. Ma trận của phép biến đổi tuyến tính qua phép đổi cơ sở

Trong mục này ta luôn giả thiết V là không gian véc tơ n chiều.

a. Ma trận đồng dạng

ĐN5: Giả sử A và B là hai m.trận vuông cấp n. Nói ***B đồng dạng với A***, nếu tồn tại một ma trận vuông cấp n khả nghịch P sao cho:

$$B = P^{-1}.A.P \qquad (2.13)$$

Nhận xét: Nếu B đồng dạng với A thì A cũng đồng dạng với B, vì:
$B = P^{-1}.A.P \Leftrightarrow A = P.B.P^{-1} = Q^{-1}.B.Q \ (Q = P^{-1})$.
Vì vậy ta nói: A và B đồng dạng với nhau, và ký hiệu: $\boldsymbol{A \sim B}$.
Đ.lý 5: Giả sử $f: \mathrm{V} \longrightarrow \mathrm{V}$ là phép biến đổi tuyến tính trên V. Khi đó nếu A là m.trận của f trong cơ sở B và A' là m.trận của của f trong cơ sở B', thì A và A' đồng dạng với nhau và: $A' = P^{-1}.A.P$, trong đó: $P = P_{(B \to B')}$.

C/m: Theo giả thiết, ta có: $\begin{cases} A.[x]_B = [f(x)]_B \\ A'.[x]_{B'} = [f(x)]_{B'} \\ [x]_B = P.[x]_{B'} \end{cases}, \forall x \in V.$

Từ đó: $[f(x)]_{B'} = P^{-1}.[f(x)]_B = P^{-1}.A.[x]_B = P^{-1}.A.P.[x]_{B'}$
Hệ thức này cho thấy ma trận của f trong cơ sở B' là: $A' = P^{-1}.A.P$.
VD 10**:** Cho: $f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2: f(x,y) = (x+y, 4y-2x)$. Hãy tìm ma trận của f trong cơ sở $B' = \{u_1 = (1,1), u_2 = (1,2)\}$.
Giải: Với cơ sở chính tắc: $\Phi = \{e_1 = (1,0), e_2 = (0,1)\}$, ta có:
$f(e_1) = (1,-2), f(e_2) = (1,4)$, vậy f có m.trận chính tắc: $\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$. M.trận chuyển từ Φ sang B' là: $P = P_{(\Phi \to B')} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ có: $P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$. Từ đó, theo đ.lí 5, ta có ma trận của f trong cơ sở B' là: $A' = P^{-1}.A.P = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

2.4.4. Trị riêng và véc tơ riêng của phép biến đổi tuyến tính

Cho X là không gian véc tơ trên trường K, $f\colon \mathrm{X} \to \mathrm{X}$ là một phép biến đổi tuyến tính. Nếu có $x \in \mathrm{X}, \mathrm{x} \neq \mathcal{O}$ và số $\lambda \in K$, sao cho: $f(x) = \lambda . x$ thì ta nói: λ là một trị riêng của f và x là véc tơ riêng của f ứng với trị riêng λ.

VD10: Với ánh xạ tuyến tính f: $\mathbb{R}^2 \to \mathbb{R}^2$, f(x, y) = (y, x), ta có:

* $f(e_1) = f(1,0) = (0,1) = e_2; f(e_2) = f(0,1) = (1,0) = e_1$. Vậy f có ma trận chính tắc: $A = ([f(e_1)], [f(e_2)]) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

* $f(x, y) = \lambda(x, y) \Leftrightarrow$ (y, x) = λ.(x, y) ⇔ λ = 1, x = y, hoặc λ = - 1, x = - y, hoặc λ = 0, x = y = 0. Vậy f có hai trị riêng: λ = 1, λ = - 1. Các véc tơ riêng ứng với λ = 1 là: (x, x) (x ≠ 0). Các véc tơ riêng ứng với λ = -1 là: (x, - x) (x ≠ 0).

VD11: Tìm trị riêng và véc tơ riêng của: $f\colon \mathbb{R}^3 \to \mathbb{R}^3 \colon f(x, y, z) = (z, x, y)$,

VD12*:* $C^1(\mathbb{R})$ là k.gian các hàm số một biến số thực, có đạo hàm liên tục trên $\mathbb{R}$, và $C(\mathbb{R})$ là k.gian các hàm số một biến số thực, liên tục trên $\mathbb{R}$. Với phép cộng các hàm và phép nhân một hàm với một số thực thì $C^1(\mathbb{R}), C(\mathbb{R})$ là các không gian véc tơ thực ($C^1(\mathbb{R})$ là không gian con của $C(\mathbb{R})$). Xét ánh xạ: $d\colon C^1(\mathbb{R}) \to C(\mathbb{R})\colon d(f) = f'$ thì d là ánh xạ tuyến tính. Khi đó hệ thức:

$$d(f) = \lambda . f \ (f \neq 0) \Leftrightarrow f' = \lambda . f \ (f \neq 0) \Leftrightarrow f(x) = c . e^{\lambda x} (c \neq 0), \forall x \in \mathbb{R}.$$

Cho thấy mọi số thực λ đều là trị riêng của ánh xạ d (phép lấy đạo hàm) và khi đó các véc tơ riêng tương ứng là họ hàm: $f(x) = c . e^{\lambda x} (c \neq 0)$.

2.5. Trị riêng và véc tơ riêng của ma trận

2.5.1. Ma trận trực giao.

ĐN1. $\mathrm{U} \in \mathrm{R}^{\mathrm{n}\times\mathrm{n}}$ được gọi là một *ma trận trực giao* nếu: $\mathbf{U}.\boldsymbol{U}^{\boldsymbol{T}} = \boldsymbol{U}^{\boldsymbol{T}}.\boldsymbol{U} = \boldsymbol{I}_{\boldsymbol{n}}$

ĐN2. $\mathrm{U} \in \mathbb{C}^{\mathrm{n}\times\mathrm{n}}$ được gọi là một ma trận Unitary nếu $\mathbf{U}.\boldsymbol{U}^{\mathrm{H}} = \boldsymbol{U}^{\mathrm{H}}.\boldsymbol{U} = \boldsymbol{I}_{\boldsymbol{n}}$

VD. Xét: $\mathrm{I_n}$, $\mathrm{A} = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha\ ) \end{bmatrix}$, $B = \frac{1}{3}\begin{bmatrix} -2 & 2 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}$; $\mathrm{U} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$

Dễ thấy $\mathrm{I_n}$, A, B là các ma trận trực giao, U là ma trận Unitary

Các tính chất của ma trận trực giao

T/c1. Nếu U là m.trận trực giao thì $\mathrm{U^T}$ cũng là m.trận trực giao và det(U) = 1

T/c2. Ma trận vuông $\mathbf{U}$ là m.trận trực giao ⇔ U khả nghịch và $\mathrm{U^{-1} = U^T}$

T/c3. M.trận trực giao U tương ứng với phép quay véc tơ trong $\mathbb{R}^{\mathrm{n}}$:

$$\forall \mathrm{x, y} \in \mathbb{R}^{\mathrm{n}}, \text{thì: } (\mathrm{Ux})^{\mathrm{T}}.(\mathrm{Uy}) = \mathrm{x^T U^T.Uy} = \mathrm{x^T}\big(\mathrm{U^T.U}\big).\mathrm{y} = \mathrm{x^T.y}$$

tức là phép biến đổi trực giao bảo toàn tích vô hướng của hai véc tơ

(ý nghĩa hình học: M.trận trực giao U xác định một ánh xạ tuyến tính:

$$\mathrm{f}: \mathrm{R^n} \to \mathrm{R^n}:\ \mathrm{f}\,(\mathrm{x}) = \mathrm{U.x}$$

mà nó bảo toàn góc giữa hai véc tơ.)

2.5.2. Trị riêng và véc tơ riêng của một ma trận vuông.

a. ĐN. Cho m.trận vuông $A \in \mathrm{R}^{\mathrm{n}\times\mathrm{n}}$ và kgvt X trên trường số K. Nếu có số $\lambda \in$ K và véc tơ $\mathrm{x} \in \mathrm{X}, \mathrm{x} \neq \mathcal{O}$ sao cho: $\mathrm{A.x} = \lambda.\mathrm{x}$, thì ta nói λ là một trị riêng của ma trận A và x là một véc tơ riêng của A tương ứng với trị riêng λ.

NX: (1) Trị riêng và véc tơ riêng của ma trận vuông A cũng chính là trị riêng và véc tơ riêng của phép biến đổi tuyến tính trong không gian n chiều V, nhận A làm ma trận trong một cơ sở B nào đó của V. **(2)**Nếu x là véc tơ riêng của A ứng với trị riêng λ, thì với mọi hằng số $c \neq 0$, véc tơ $c.x$ cũng là một véc tơ riêng của A ứng với trị riêng λ.

VD1: Với $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ và véc tơ $x = (2, 1)$, ta có: $A.x^T = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2.x^T$. Vậy $\lambda = 2$ là một trị riêng của A và $x = (2, 1)$ là một véc tơ riêng của A ứng với trị riêng $\lambda = 2$.

b. Phương trình đặc trưng. Việc tồn tại trị riêng λ và véc tơ riêng tương ứng của ma trận A là sự tồn tại nghiệm không tầm thường của hệ phương trình tuyến tính: $A.x^T = \lambda.x^T$ (2.14a)

hay: $(A - \lambda I).x^T = 0.$ (2.14b)

Điều này có nghĩa là: $\det(A - \lambda I) = 0$ (2.14c)

(2.14c)) là phương trình để tìm trị riêng λ, gọi là *phương trình đặc trưng* của ma trận vuông A. $\det(A - \lambda I)$ là một đa thức đối với λ, gọi là *đa thức đặc trưng* của ma trận A.

Chú ý:

• (2.14c) là phương trình đại số cấp n nên có n nghiệm trên trường số phức. Vậy một ma trận vuông cấp n có đúng n trị riêng (thực hoặc phức, đơn hoặc bội) khi xét trong không gian véc tơ phức. Số bội của nghiệm λ của phương trình đặc trưng (2.14c) được gọi là số bội đại số của λ.

• Với λ là một nghiệm phức thì nó là một trị riêng của ma trận A nếu kgvt là trên trường phức. x là một véc tơ riêng ứng với trị riêng λ khi và chỉ khi x là nghiệm khác không của hệ phương trình tuyến tính thuần nhất (2.14b). Như đã biết: tập nghiệm của (2.14b) là một không gian con của $\mathbb{R}^{\mathrm{n}}$, gọi là không gian con riêng ứng với trị riêng λ, ký hiệu là N_λ:

$N_\lambda = \{x \in \mathbb{R}^{\mathrm{n}}\colon\ (A - \lambda I).x^T\ = \mathcal{O}\}$

Khi đó số chiều của N_λ được gọi là số bội hình học của λ.

Đ.lý 1(*về trị riêng của ma trận đối xứng*). Nếu A là ma trận vuông cấp n, đối xứng thì A có đúng n trị riêng thực (gồm cả số bội của nghiệm) và n véc tơ riêng trực chuẩn tương ứng.

2.5.3. Cách tìm trị riêng và véc tơ riêng của ma trận vuông A:

Từ các kết quả trên, có thể tìm trị riêng và véc tơ riêng của ma trận vuông A theo các bước sau:

B1. Giải phương trình đặc trưng: $\det(A - \lambda I) = 0$, mỗi nghiệm của phương trình này là một trị riêng.

B2. Với mỗi trị riêng λ, ta giải hệ phương trình: $(A - \lambda I).X = \mathcal{O}$,

Từ đó tìm được không gian con riêng tương ứng N_λ.

Tập các véc tơ riêng của A tương ứng với trị riêng λ là $N_\lambda \setminus \{\mathcal{O}\}$

VD2: Tìm các trị riêng và các véc tơ riêng tương ứng của ma trận: $A = \begin{bmatrix} -2 & 5 \\ -`1 & 2 \end{bmatrix}$

Giải: A có phương trình đặc trưng:

$$\det(A-\lambda I)=0 \Leftrightarrow \begin{vmatrix} -2-\lambda & 5 \\ -1 & 2-\lambda \end{vmatrix}=0 \Leftrightarrow \lambda^2+1=0 \Leftrightarrow \left[\begin{matrix} \lambda=\mathrm{i} \\ \lambda=-i \end{matrix}\right. ,$$

Vậy A không có trị riêng thực (mà có 2 trị riêng phức $\lambda = i$ và $\lambda = -i$)
Vì thế A không có véc tơ riêng trên kgvt thực.
Tuy nhiên nếu xét trên kgvt phức thì hệ phương trình:

$$(A-\lambda I).X=0 \Leftrightarrow \begin{cases} (-2-\lambda)x+5y=0 \\ -x+(2-\lambda)y=0 \end{cases} \Leftrightarrow \begin{cases} (\lambda^2+1)y=0 \\ x=(2-\lambda)y \end{cases}.$$

• $Với\ \lambda = i, hệ\ có\ tập\ nghiệm$:
$N_\lambda = \{(x,y): y = t(tùy\ ý), x = (2-i)t\} = \{t.(2-i,1): t \in \mathbb{R}\}$
$= span\{u_1\}(u_1 = (2-i,1))$
Tập các véc tơ riêng tương ứng là: $\{t.(2-i,1): 0 \neq t \in \mathbb{R}\}$
• $Với\ \lambda = -i, hệ\ có\ tập\ nghiệm$: $N_\lambda = \{(x,y): y = t(tùy\ ý), x = (2+i)t\}$
$= \{t.(2+i,1): t \in \mathbb{R}\} = span\{u_1\}(u_1 = (2+i,1))$
Tập các véc tơ riêng tương ứng là: $\{t.(2+i,1): 0 \neq t \in \mathbb{R}\}$)
VD 3: Tìm các trị riêng và các véc tơ riêng tương ứng của ma trận:

$$B = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Giải: Phương trình đặc trưng:

$$\det(B-\lambda.I)=0 \Leftrightarrow \begin{vmatrix} 3-\lambda & -2 & 0 \\ -2 & 3-\lambda & 0 \\ 0 & 0 & 5-\nu \end{vmatrix} = 0 \Leftrightarrow \left[\begin{array}{l} \lambda = 1(đơn) \\ \lambda = 5(bội\ 2) \end{array}\right.$$

Hệ phương trình: $(B-\lambda I).X=\mathcal{O} \Leftrightarrow \begin{bmatrix} 3-\lambda & -2 & 0 \\ -2 & 3-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$\Leftrightarrow \begin{cases} (3-\lambda)x-2y=0 \\ -2x+(3-\lambda)y=0\,. \\ (5-\lambda)z=0 \end{cases}$ $(*)$. Từ đó:

• Với $\lambda=1$, (*) có k.gian nghiệm: $N_\lambda=\{(t,t,0): t\in\mathbb{R}\}=span\{(1,1,0)\}$

Các véc tơ riêng tương ứng là: $\{(t,t,0): 0\neq t\in\mathbb{R}\}$

• Với $\lambda=5$, (*) có k.gian nghiệm: $N_\lambda=\{(t,-t,s): t,s\in\mathbb{R}\}=$

$\{t.(1,-1,0)+s.(0,0,1), t,s\in\mathbb{R}\}=span\{(1,-1,0),(0,0,1)\}$

Các véc tơ riêng tương ứng là: $\{(t,-t,s): t,s\in\mathbb{R}, t^2+s^2>0\}$

***2.5.4.* Trị riêng và véc tơ riêng của phép biến đổi tuyến tính trong không gian hữu hạn chiều.**

Giả sử V là không gian véc tơ n chiều, $f:\ \mathrm{V}\rightarrow V$ là phép biến đổi tuyến tính và $\mathrm{B}=\ \{\mathrm{u}_1,\mathrm{u}_2,\dots,\mathrm{u}_\mathrm{n}\}$ là một cơ sở sắp thứ tự trên V. Trong cơ sở B, f có ma trận A là ma trận vuông cấp n: $A=([f(\mathrm{u}_1)]_B,[f(\mathrm{u}_2)]_B,\dots,[f(\mathrm{u}_\mathrm{n})]_B)$

Vì: $f(x) = \lambda.x \Leftrightarrow A.[x]_B = \lambda.[x]_B$
nên véc tơ x là véc tơ riêng của f, ứng với trị riêng λ, khi và chỉ khi véc tơ $[x]_B$ là véc tơ riêng của ma trận A ứng với trị riêng λ.
Do đó việc tìm giá trị riêng và véc tơ riêng của một phép biến đổi tuyến tính $f\colon \mathrm{V} \to V$ trong không gian n chiều V được chuyển về tìm giá trị riêng và véc tơ riêng của ma trận theo các bước sau:
B1. Xác định một cơ sở sắp thứ tự $\mathrm{B} = \{\mathrm{u}_1, \mathrm{u}_2, \dots, \mathrm{u}_\mathrm{n}\}$ trên V và tìm ma trận A của phép biến đổi tuyến tính f trong cơ sở B.
B2. Tìm trị riêng và véc tơ riêng của ma trận A. Lưu ý rằng: mỗi véc tơ riêng của ma trận A là một véc tơ $(\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_\mathrm{n}) \in \mathbb{R}^\mathrm{n}$.
Khi đó véc tơ: $x = \mathrm{x}_1.\mathrm{u}_1 + \mathrm{x}_2.\mathrm{u}_2 + \cdots + \mathrm{x}_\mathrm{n}.\mathrm{u}_\mathrm{n}$ $(\in V)$ là véc tơ riêng của phép biến đổi tuyến tính f.
VD4: Chứng minh phép chuyển vị ma trận vuông cấp 2 là một phép biến đổi tuyến tính và tìm tất cả các trị riêng và véc tơ riêng của phép chuyển vị này.
Giải: Phép chuyển vị ma trận vuông cấp 2 là ánh xạ $f\colon M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$: $f(D) = D^T, \forall D \in M_{2\times 2}(\mathbb{R})$ Dễ thấy f là một ánh xạ tuyến tính. Để tìm trị riêng và véc tơ riêng của f, ta thực hiện các bước:
B1. Xét cơ sở chính tắc của $M_{2\times 2}(\mathbb{R})$ là:

$$\Phi = \left\{\varphi_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \varphi_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \varphi_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \varphi_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\},$$

ta có: $f(\varphi_{11}) = \varphi_{11}, f(\varphi_{12}) = \varphi_{21}, f(\varphi_{21}) = \varphi_{12}, f(\varphi_{22}) = \varphi_{22}$, nên trong cơ sở này, f có ma trận:

$$A = ([f(\varphi_{11})]_\Phi, [f(\varphi_{12})]_\Phi, [f(\varphi_{21})]_\Phi, [f(\varphi_{22})]_\Phi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

B2. Tìm trị riêng và véc tơ riêng của ma trận A.

- Xét phương trình đặc trưng: $\det(A - \lambda.I) = 0$

$$\Leftrightarrow \begin{vmatrix} 1-\lambda & 0 & 0 & 0 \\ 0 & -\lambda & 1 & 0 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1-\lambda \end{vmatrix} = 0 \Leftrightarrow (\lambda - 1)^3(\lambda + 1) = 0 \Leftrightarrow \left[\begin{array}{l} \lambda = -1(\text{đơn}) \\ \lambda = 1(bội\ 3) \end{array}\right.$$

- Hệ phương trình: $(A - \lambda I).X = \mathcal{O}$

$$\Leftrightarrow \begin{bmatrix} 1-\lambda & 0 & 0 & 0 \\ 0 & -\lambda & 1 & 0 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix} . \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} (1-\lambda)x = 0 \\ -\lambda y + z = 0 \\ y - \lambda z = 0 \\ (1-\lambda)t = 0 \end{cases} \qquad (**)$$

* Ứng với $\lambda = -1$, (**) có nghiệm: $(x = 0, y = -z, z\ (tùy\ ý), t = 0)$, tức là:

$$N_{(\lambda=-1)} = \{(0, -z, z, 0): z \in \mathbb{R}\} = Span\{(0,-1,1,0)\}$$

* Ứng với $\lambda = 1$, (**) có nghiệm: $(x\ (tùy\ ý), y(tùy\ ý), z = y, t\ (tùy\ ý))$, tức là: $N_{(\lambda=1)} = \{(x, y, y, t): x, y, t \in \mathbb{R}\} = Span\{(1,0,0,0), (0,1,1,0), (0,0,0,1)\}$

Từ đó suy ra: phép chuyển vị ma trận vuông cấp 2 là ánh xạ f đang xét có các trị riêng: $\lambda = -1$ (đơn) và $\lambda = 1$ (bội 3), trong đó:

- Ứng với $\lambda = -1$, có các véc tơ riêng: $D = \begin{bmatrix} 0 & -z \\ z & 0 \end{bmatrix}, \forall z \in \mathbb{R}, z \neq 0$

- Với $\lambda = 1$, có các véc tơ riêng: $D = \begin{bmatrix} x & y \\ y & t \end{bmatrix}, \forall x, y, t \in \mathbb{R}, x^2 + y^2 + t^2 > 0.$

2.6. Chéo hóa ma trận

2.6.1. Vấn đề chéo hóa một ma trận. Đ/v một phép biến đổi tuyến tính f trong k.gian n chiều V, ma trận của f trong cơ sở B nào đó phụ thuộc vào B. Khi cơ sở B thay đổi thì ma trận của f cũng thay đổi. Vấn đề đặt ra là:

(1) Người ta mong muốn tìm được một cơ sở nào đó của V sao cho m.trận của f trong đó có dạng đơn giản là dạng chéo. Tức là có hay không một cơ sở của V sao cho m.trận của f là ma trận chéo?

(2) Đ/v V là kgvt n chiều có tích vô hướng, thì có hay không một cơ sở trực giao trong V sao cho m.trận của f trong cơ sở này là m.trận chéo?

ĐN1. Cho ma trận vuông A. Ta nói: Ma trận A chéo hóa được, nếu tồn tại một ma trận khả nghịch P sao cho $\mathrm{P}^{-1}.\mathrm{A}.\mathrm{P}$ là một ma trận đường chéo. Khi đó nói ma trận P làm chéo hóa ma trận A. Ngoài ra nếu P là một ma trận trực giao thì ta nói A chéo hóa trực giao được.

2.6.2. Giải bài toán chéo hóa ma trận

Đ.Lí 1. $\mathrm{A} \in \mathrm{R}^{\mathrm{n}\times\mathrm{n}}$ chéo hóa trực giao được ⇔ A có n véc tơ riêng trực chuẩn.
Có thể tóm tắt chứng minh định lý này như sau:

- Nếu A có n véc tơ riêng trực chuẩn: $p_1, p_2, \dots, p_n$ ứng với các trị riêng $\lambda_1, \lambda_2, \dots, \lambda_n$, thì ma trận $\mathrm{P} = \big[[p_1]\ [p_2] \dots [p_n]\big]$ là ma trận trực giao sẽ làm chéo hóa A và: $P^{-1}.A.P = D = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.
- Nếu A chéo hóa trực giao được bởi ma trận trực giao P gồm các cột $p_1, p_2, \dots, p_n$ thì đương nhiên các cột này là các véc tơ trực giao khác không, hơn nữa có thể xem chúng là các véc tơ trực chuẩn (vì nếu không thì ta chuẩn hóa các véc t ơ này). Như vậy: $P^{-1}.A.P = D$ là ma trận chéo với các phần tử chéo là $\lambda_1, \lambda_2, \dots, \lambda_n$. Khi đó $\lambda_1, \lambda_2, \dots, \lambda_n$ là các trị riêng của A và $p_1, p_2, \dots, p_n$ là các véc tơ riêng tương ứng.

Đ.Lí 2. a/ A chéo hóa trực giao được ⇔ A là ma trận đối xứng.

b/ Nếu A là m.trận đối xứng thì: **(b1)** các véc tơ riêng thuộc các k.gian riêng khác nhau sẽ trực giao nhau theo tích vô hướng Euclid trong $\mathbb{R}^{\mathrm{n}}$; **(b2)** số bội hình học của mỗi trị riêng bằng số bội đại số của nó, tức là: nếu λ là một nghiệm bội k của phương trình đặc trưng thì k.gian riêng N_λ sẽ có số chiều là k.

NX: (1) A là chéo hóa được ⇔ A đồng dạng với một m.trận đường chéo.

(2) Nếu P làm chéo hóa A: $P^{-1}.A.P = D$, thì: $\mathrm{A}^{\mathrm{k}} = \mathrm{P}.\mathrm{D}^{\mathrm{k}}.\mathrm{P}^{-1}$, $\forall \mathrm{k} \in \mathbb{N}^*$ và nếu A khả nghịch thì: $\mathrm{A}^{-1} = \mathrm{P}.\mathrm{D}^{-1}.\mathrm{P}^{-1}$

Quy trình chéo hóa trực giao một ma trận A đối xứng như sau:
B1. Tìm các trị riêng $\lambda_1, \lambda_2, \dots, \lambda_n$ (có thể trùng nhau do nghiệm bội) và các k.gian riêng tương ứng. Tìm các véc tơ cơ sở cho mỗi k.gian riêng.
B2. Áp dụng quá trình trực giao hóa Gram-Smidt cho mỗi cơ sở của các k.gian riêng để thu được cơ sở trực chuẩn cho mỗi k.gian này. Từ đó nhận được n véc tơ riêng trực chuẩn: $p_1, p_2, \dots, p_n$.
B3. Lập ma trận P gồm n cột là n véc tơ $p_1, p_2, \dots, p_n$. Ma trận P là trực giao, sẽ làm chéo hóa A: $P^{-1}.A.P = D(các\ phần\ tử\ chéo\ l$à $\lambda_1, \lambda_2, \dots, \lambda_n)$.

VD 5. Tìm m.trận trực giao P làm chéo hóa ma trận: $B = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

Giải: Trong ví dụ 3, ta đã tìm được các trị riêng $\lambda_1 = 1$(đơn) và $\lambda_2 = 5$(bội 2) và N_{λ_1} có cơ sở là một véc tơ $v_1 = (1,1,0)$, N_{λ_2} có cơ sở là 2 véc tơ $v_2 = (1,-1,0), v_3 = (0,0,1)$. Áp dụng quá trình trực giao hóa Gram-Smidt:
- Đối với cơ sở $\{v_1\}$, nhận được cơ sở trực chuẩn $\{\, u_1 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right)\}$
- Đối với cơ sở $\{v_2, v_3\}$, nhận được cơ sở trực chuẩn $\{u_2 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0\right), u_3 = (0,0,1)\}$. Ma trận trực giao làm chéo hóa B là:

$$P = \left[[\mathrm{u}_1]\ \ [\mathrm{u}_2]\ \ [\mathrm{u}_3]\right]$$

2.6.3. Ma trận xác định dương và bán xác định dương.

a. ĐN. Cho ma trận đối xứng Cho $A, B \in \mathbb{R}^{n\times n}$, gọi:

(1) A là m.trận ***xác định dương*** (*positive definite*), kí hiệu: $A \succ 0$ nếu:

$$(x).A.[x] > 0, \forall x \in \mathbb{R}^n, x \neq \mathcal{O}$$

(2) A là m.trận ***bán xác định dương*** (*positive semidefinite*),

kí hiệu: $A \succcurlyeq 0$ nếu: $(x).A.[x] \geq 0, \forall x \in \mathbb{R}^n$

(3) A là m.trận ***xác định âm*** (*negative definite*), kí hiệu: $A \prec 0$, nếu:

$$(x).A.[x] < 0, \forall x \in \mathbb{R}^n, x \neq \mathcal{O}$$

(4) A là m.trận ***bán xác định âm*** (*negative semidefinite*),

kí hiệu: $A \preccurlyeq 0$, nếu: $(x).A.[x] \leq 0, \forall x \in \mathbb{R}^n$

(5) Kí hiệu $A \succ B(\succcurlyeq, \prec, \preccurlyeq) \Leftrightarrow A - B \succ 0(A \succcurlyeq B, A \prec B, A \preccurlyeq B)$

(6) A là ma trận không xác định dấu, nếu $\exists u, v \in \mathbb{R}^n$: $(u).A.[u] > 0$, $(v).A.[v] < 0$

Chú ý: (a) $A \succ 0$ (hoặc $A \prec 0$**)** $\Rightarrow A \succcurlyeq 0$ **(hoặc** $A \preccurlyeq 0$**)**

(b) Một ma trận Hermitian $A \in \mathbb{C}^{n\times n}$ được gọi là xác định dương, nếu:

$$(x)^H.A.[x] > 0, \forall x \in \mathbb{C}^n, x \neq \mathcal{O}$$

(Tương tự cho các k/n Ma trận Hermitian bán xác định dương, xác định âm, bán xác định âm).

(c) Thực tế người ta thường quan tâm đến tính bán xác định dương

(d) Các đặc tính nói trên của ma trận được dùng trong các bài toán tối ưu, cực trị.

b. VD6. Xét các ma trận: $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Dễ thấy:

A là xác định dương, B là bán xác định dương, $-$A là xác định âm, $-$B là bán xác định âm, C không phải bán xác định dương, không phải bán xác định âm.

c. Các tính chất

T/c 1. Ma trận xác định dương có tất cả các trị riêng đều là các số thực dương

T/c 2. Nếu A là xác định dương thì A khả nghịch và det(A) > 0 và:

$(x).A.[x] = 0 \Leftrightarrow A.[x] = \mathcal{O}$

T/c 3. (*Tiêu chuẩn Sylvester*). Với ma trận đối xứng $A = (a_{ij})_{n \times n}$

a. A là xác định dương (bán xác định dương) ⇔ mọi m.trận vuông con góc trái $A_k = (a_{ij})_{k \times k}$ (k = 1, 2,..., n) của A đều có det(A_k) > 0 (det$(A_k) \geq 0$).

b. A là xác định âm ⇔ mọi ma trận vuông con góc trái $A_k = (a_{ij})_{k \times k}$ (k = 1, 2,..., n) của A có det(A_k) > 0 (det$(A_k) \geq 0$). nếu k chẵn và det(A_k) < 0 (det$(A_k) \leq 0$). nếu k lẻ.

T/c 4. $\forall B \in R^{m \times n}$ thì $A = B^H.B$ là m.trận xác định dương

T/c 5. (Khai triển Cholesky): Mọi m.trận Hermitian bán xác định dương đều biểu diễn được duy nhất dưới dạng: $A = L.L^H$, trong đó L là một m.trận tam giác dưới có các phần tử trên đường chéo đều là các số thực dương.

T/c 6. Nếu A là m.trận bán xác định dương thì: $(x).A.[x] = 0 \Leftrightarrow A.[x] = \mathcal{O}$

VD7. Cho các m.trận

$$A = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 1 & -1 \\ -1 & -1 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, C = \begin{bmatrix} -5 & -2 & 1 \\ -2 & -1 & 1 \\ 1 & 1 & -3 \end{bmatrix}, D = \begin{bmatrix} 1 & 1-i \\ i & -1 \\ 0 & 1 \end{bmatrix}$$

- Xét A, có: $\det(A_1) = 5 > 0, \det(A_2) = \begin{vmatrix} 5 & 2 \\ 2 & 1 \end{vmatrix} = 1 > 0$, $\det(A_3) = \det(A) = 1 > 0$. Vậy A là m.trận xác định dương

- Xét B: $\det(B_1) = 1 > 0, \det(B_2) = \det(B) = 0$. Vậy B bán xác định dương

- Xét C: $\det(C_1) = -5 < 0, \det(C_2) = \begin{vmatrix} -5 & -2 \\ -2 & -1 \end{vmatrix} = 1 > 0$, $\det(C_3) = \det(C) = -1 < 0$. Vậy C là m.trận xác định âm

- Xét m.trận $U = D^H . D = \begin{bmatrix} 1 & -i & 0 \\ 1+i & -1 & 1 \end{bmatrix} . \begin{bmatrix} 1 & 1-i \\ i & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$

Theo T/c 4 hoặc theo tiêu chuẩn Sylvester, ta có U là m.trận xác định dương.

2.6.4. Vết của ma trận vuông. Cho $A = \left(a_{ij}\right)_{n\times n} \in R^{n\times n}$

a. ĐN. Vết của m.trận A, kí hiệu trace(A), là tổng các phần tử trên đường chéo: $\mathrm{trace}(A) = a_{11} + a_{22} + \cdots + a_{nn}$

b. VD8. Trong VD7, có trace(A) = 5 + 1 + 3 = 9

c. Các tính chất.

(1) $\mathrm{trace}\left(A^T\right) = \mathrm{trace}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$ (Tổng các trị riêng)

$\mathrm{trace}(A.B) = \mathrm{trace}(B.A)$; $\mathrm{trace}(A+B) = \mathrm{trace}(A) + \mathrm{trace}(B)$

(2). $\forall X$ (khả nghịch) $\in R^{n\times n}$ *:* $\mathrm{trace}(X.A.X^{-1}) = \mathrm{trace}(X^{-1}.A.X) = \mathrm{trace}(A)$

(3). $\forall X \in R^{m\times n}$: $\mathrm{trace}\left(X.X^T\right) = \mathrm{trace}\left(X^T.X\right) = \|X\|_F^2$

Vết của m.trận vuông có nhiều ứng dụng trong tối ưu.