Machine Learning exam

I. Lý thuyết.

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

1) Vector $\vec{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ và $\vec{v_2} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ độc lập tuyến tính, vì với mỗi số thực $c_1, c_2 \in \mathbb{R}$ bất kỳ:

$$c_1 v_1 + c_2 v_2 = A\vec{c} = 0$$

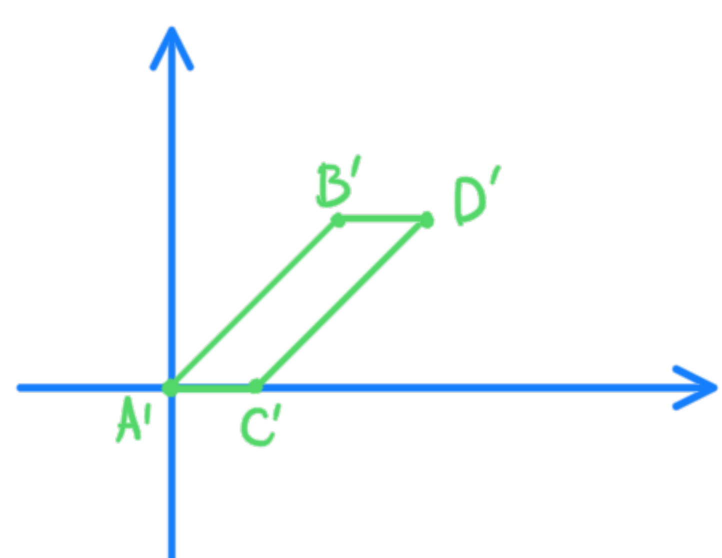
có nghiệm $\vec{c} \neq \vec{0}$ khi và chỉ khi A là ma trận singular.
Mà $\det(A) = 2$, do đó A là ma trận khả nghịch.

↳ Chỉ có $\vec{c} = \vec{0}$ thỏa mãn $A\vec{c} = 0$. Như vậy $\vec{v_1}$ và $\vec{v_2}$ độc lập tuyến tính. □

2) Biến đổi mỗi điểm với ma trận A:

$A\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\quad$ $A\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$A\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ $\quad$ $A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$



↳ Linear transformation với ma trận A cho hình mới A'B'D'C' với diện tích là 2.

3) $A = PDP^{-1}$, trong đó $\begin{cases} P: \text{ ma trận eigenvectors của } A. \\ D: \text{ ma trận eigenvalues của } A. \end{cases}$

☆ Tìm eigenvalues λ_i của A:

$$A\vec{x} = \lambda\vec{x} \Leftrightarrow (A-\lambda I)\vec{x} = 0$$

Ma trận A có eigenvector $\vec{x} \neq \vec{0}$ khi và chỉ khi $A - \lambda I$ là ma trận singular.

$\Leftrightarrow \det(A - \lambda I) = 0$

$\Leftrightarrow \begin{vmatrix} 1-\lambda & 2 \\ 0 & 2-\lambda \end{vmatrix} = 0$

$\Leftrightarrow (1-\lambda)(2-\lambda) = 0$

$\Leftrightarrow \begin{cases} \lambda_1 = 1 \\ \lambda_2 = 2 \end{cases}$. Từ đó: $D = \begin{bmatrix} 1 & \\ & 2 \end{bmatrix}$.

☆ Tìm eigenvectors của A:

+) $\lambda_1 = 1$: $\begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \vec{x}_1 \end{bmatrix} = 0 \Leftrightarrow \vec{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

+) $\lambda_2 = 2$: $\begin{bmatrix} -1 & 2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \vec{x}_2 \end{bmatrix} = 0 \Leftrightarrow \vec{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Từ đó: $P = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. Suy ra $P^{-1} = \frac{1}{\det P} C^T = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$

(C là cofactor matrix của P: $C = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$)

☆ Đáp án:

$$\boxed{\underset{A}{\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}} = \underset{P}{\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}} \underset{D}{\begin{bmatrix} 1 & \\ & 2 \end{bmatrix}} \underset{P^{-1}}{\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}}}$$

4) Ta có $A^4 = PDP^{-1}PDP^{-1}PDP^{-1}PDP^{-1}$

+) Với mỗi $P^{-1}P = I$, ta rút gọn A^4 thành:

$A^4 = PD^4P^{-1}$

(từ đó có công thức $A^k = PD^kP^{-1}$)

$$A^4 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & \\ & 16 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

$$A^4 = \begin{bmatrix} 1 & 32 \\ 0 & 16 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

$$A^4 = \begin{bmatrix} 1 & 30 \\ \dots \end{bmatrix}$$

$\lfloor 0 \quad 16 \rfloor$

5) Sử dụng tính chất của determinant:

(a) $\det(AB) = \det A \det B$

(b) $\det A^{-1} = \frac{1}{\det A}$.

(c) Determinant của ma trận chéo là tích của mỗi entry trên đường chéo đó.

+) Ta có ma trận A khả nghịch $\Rightarrow \det A \neq 0$ và do đó $\lambda_i \neq 0$. Như vậy A có thể chéo hóa:

$$A = PDP^{-1}$$

Từ tính chất (a)(b)(c) ta có:

$$\det A = \det P \det D \det P^{-1}$$

$$\Rightarrow \det A = \det P \left(\prod_{i=1}^{n} \lambda_i\right) \frac{1}{\det P}$$

$$\Rightarrow \boxed{\det A = \prod_{i=1}^{n} \lambda_i} \quad \square$$

II. Thực hành

3) Nhận xét:

+) g(x) là hàm số lẻ: $g(-x) = -g(x)$.

+) Domain của g(x): $(-\infty, \infty)$. Hàm g(x) liên tục và có đạo hàm với mọi x.

+) $$\begin{cases} \lim\limits_{x \to +\infty} g(x) = 1 \\ \lim\limits_{x \to -\infty} g(x) = 0 \\ g(0) = \frac{1}{2} \end{cases}$$

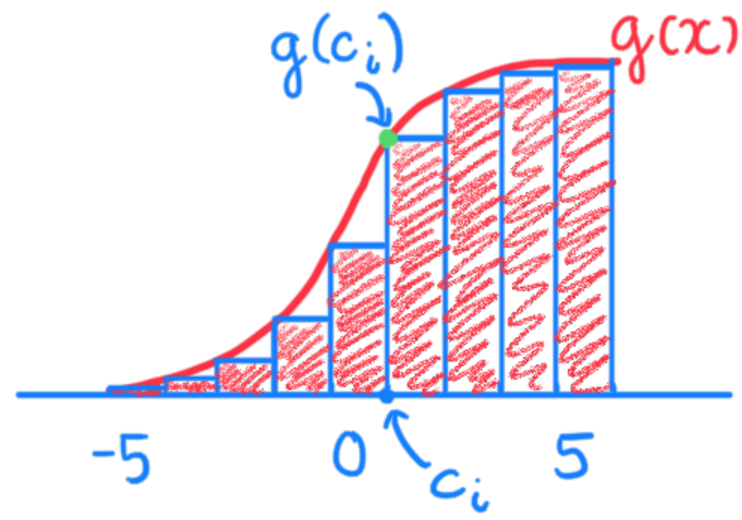
+) g(x) là hàm không tuyến tính, thường sử dụng trong binary classification với mục đích phân loại dữ liệu dựa trên xác suất rơi vào 1 trong 2 class.
(có thể dùng cho multiclass, tuy nhiên không tốt bằng softmax).

5) Sử dụng phương pháp Left Riemann sum:

+) Chia domain (a, b) thành n phần bằng nhau.

+) Vẽ các hình chữ nhật với đáy $\frac{b-a}{n}$ và cạnh bên trái là giá trị của $g(c_i)$ tại điểm c_i. Cạnh bên phải bằng cạnh bên trái.

+) Tổng diện tích của các hình chữ nhật là tích phân của $g(x)$ từ a đến b:



$$\int_a^b g(x)\,dx = \sum_{i=1}^{n} \left(\frac{b-a}{n}\right) g(c_i)$$