Machine Learning

SEU — 2022 Spring

Assignment 1

61520522 Teddy van Jerry   *Instructor: TeX - LaTeX Stack Exchange*

Problem 1: Basic Vector Operations

(1) $\|\mathbf{a}\|_2 = \sqrt{1^2+2^2+3^2} = \sqrt{14}, \quad \|\mathbf{b}\|_2 = \sqrt{(-8)^2+1^2+2^2} = \sqrt{69}.$

(2) $\|\mathbf{a}-\mathbf{b}\|_2 = \sqrt{9^2+1^2+1^2} = \sqrt{83}.$

(3) $\mathbf{a}$ and $\mathbf{b}$ are orthogonal.

Proof. The inner product of vectors $\mathbf{a}$ and $\mathbf{b}$ is

$$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T\mathbf{b} = 1 \times (-8) + 2 \times 1 + 3 \times 2 = 0, \tag{1.1}$$

therefore $\mathbf{a}$ and $\mathbf{b}$ are orthogonal. □

Problem 2: Basic Matrix Operations

According to the consensus, the matrix notation should be the bold upper-case letter like $\mathbf{A}$ or $\boldsymbol{A}$, not A.

(1)

$$\begin{aligned}[\mathbf{A}, \mathbf{I}_3] &= \left[\begin{array}{rrr:rrr} 1 & -3 & 3 & 1 & 0 & 0 \\ 3 & -5 & 3 & 0 & 1 & 0 \\ 6 & -6 & 4 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rrr:rrr} 1 & -3 & 3 & 1 & 0 & 0 \\ 0 & 4 & -6 & -3 & 1 & 0 \\ 0 & 12 & -14 & -6 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rrr:rrr} 1 & -3 & 3 & 1 & 0 & 0 \\ 0 & 4 & -6 & -3 & 1 & 0 \\ 0 & 0 & 4 & 3 & -3 & 1 \end{array}\right] \\ &\sim \left[\begin{array}{rrr:rrr} 1 & -3 & 0 & -\frac{5}{4} & \frac{9}{4} & \frac{3}{4} \\ 0 & 4 & 0 & \frac{3}{2} & -\frac{7}{2} & -\frac{3}{2} \\ 0 & 0 & 1 & \frac{3}{4} & -\frac{3}{4} & \frac{1}{4} \end{array}\right] \sim \left[\begin{array}{rrr:rrr} 1 & 0 & 0 & -\frac{1}{8} & -\frac{3}{8} & \frac{3}{8} \\ 0 & 1 & 0 & \frac{3}{8} & -\frac{7}{8} & \frac{3}{8} \\ 0 & 0 & 1 & \frac{3}{4} & -\frac{3}{4} & \frac{1}{4} \end{array}\right],\end{aligned} \tag{2.1}$$

where $\mathbf{I}_3$ is the 3×3 identity matrix. Therefore we have

$$\mathbf{A}^{-1} = \begin{bmatrix} -\frac{1}{8} & -\frac{3}{8} & \frac{3}{8} \\ \frac{3}{8} & -\frac{7}{8} & \frac{3}{8} \\ \frac{3}{4} & -\frac{3}{4} & \frac{1}{4} \end{bmatrix}. \tag{2.2}$$

The determinant of matrix $\mathbf{A}$ can be calculated as

$$\det(\mathbf{A}) = 1 \times \begin{vmatrix} -5 & 3 \\ -6 & 4 \end{vmatrix} - (-3) \times \begin{vmatrix} 3 & 3 \\ 6 & 4 \end{vmatrix} + 3 \times \begin{vmatrix} 3 & -5 \\ 6 & -6 \end{vmatrix} = 1 \times (-2) + 3 \times (-6) + 3 \times 12 = 16, \tag{2.3}$$

where $|\cdot|$ denotes the determinant.

(2) The rank of matrix $\mathbf{A}$ is 3 because as is shown in Eq. (2.2) the matrix $\mathbf{A}$ is invertible.

(3) The trace of matrix $\mathbf{A}$ is

$$\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{3} a_{ii} = 1 + (-5) + 4 = 0. \tag{2.4}$$

$$\mathbf{A} + \mathbf{A}^T = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 3 & 6 \\ -3 & -5 & -6 \\ 3 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 9 \\ 0 & -10 & -3 \\ 9 & -3 & 8 \end{bmatrix}. \tag{2.5}$$

(4)

$$\mathbf{A} + \mathbf{A}^T = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 3 & 6 \\ -3 & -5 & -6 \\ 3 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 9 \\ 0 & -10 & -3 \\ 9 & -3 & 8 \end{bmatrix}. \tag{2.6}$$

(5) $\mathbf{A}$ is not an orthogonal matrix.

Proof. Assume $\mathbf{A}$ is an orthogonal matrix, therefore

$$\mathbf{A}\mathbf{A}^T = \mathbf{I}_3, \tag{2.7}$$

Take the determinant at both side, it can be derived that

$$|\det(\mathbf{A})| = \sqrt{|\mathbf{A}||\mathbf{A}^T|} = |\det(\mathbf{I}_3)| = 1, \tag{2.8}$$

which contradicts with Eq. (2.3). Therefore, the assumption is false. □

(6) Let $f(\lambda)$ be the characteristic function of matrix $\mathbf{A}$ and

$$f(\lambda) = \begin{vmatrix} \lambda - 1 & 3 & -3 \\ -3 & \lambda + 5 & -3 \\ -6 & 6 & \lambda - 4 \end{vmatrix} = (\lambda - 4)(\lambda + 2)^2, \tag{2.9}$$

therefore the eigenvalues are $\lambda_1 = 4, \lambda_2 = \lambda_3 = -2$. Let the corresponding eigenvectors be $\boldsymbol{\alpha}_i$, $i = 1, 2, 3$.

$$(\mathbf{A} - \lambda_i \mathbf{I}_3)\boldsymbol{\alpha}_i = \mathbf{0}, \quad i = 1, 2, 3, \tag{2.10}$$

and the corresponding eigenvectors are

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} 1 & 1 & 2 \end{bmatrix}^T, \quad \boldsymbol{\alpha}_{2,3} = \begin{bmatrix} 1 & 1 + c_{2,3} & c_{2,3} \end{bmatrix}^T, \tag{2.11}$$

where $c_{2,3} \in \mathbb{R}$. Without loss of generality, we take $c_2 = 0$ and $c_3 = -1$, and we have $\boldsymbol{\alpha}_2 = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^T$ and $\boldsymbol{\alpha}_2 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T$.

(7) Use the result from Eq. (2.9), the matrix $\mathbf{A}$ can be diagonalized as

$$\boldsymbol{\Lambda} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}. \tag{2.12}$$

(8) The $\ell_{2,1}$ norm of $\mathbf{A}$ is

$$\|\mathbf{A}\|_{2,1} = \sum_{i=1}^{3} \sqrt{\sum_{j=1}^{3} a_{ij}^2} = \sqrt{46} + \sqrt{70} + \sqrt{34} \approx 20.98, \tag{2.13}$$

and the Frobenius norm of $\mathbf{A}$ is

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i,j=1,2,3} a_{ij}^2} = \sqrt{150} = 5\sqrt{6} \approx 12.247. \tag{2.14}$$

(9) The nuclear norm of $\mathbf{A}$ is

$$\|\mathbf{A}\|_* = \operatorname{tr}(\sqrt{\mathbf{A}\mathbf{A}^*}) = \sum_{i=1}^{3} \sigma_i(\mathbf{A}) \approx 14.728, \tag{2.15}$$

and the spectral norm of $\mathbf{A}$ is

$$\|\mathbf{A}\|_2 = \max \sigma_i(\mathbf{A}) \approx 12.065. \tag{2.16}$$

MATLAB Code for Check

```
A = [1, -3, 3; 3, -5, 3; 6, -6, 4]; % define the matrix A
inv(A) % calculate and print the inverse of A
det(A) % the determinant of A
rank(A) % the rank of A
trace(A) % the trace of A
A + A.' % the sum of A and the transpose of A
sum(sum(A * A.' ~= eye(3))) % check if A is orthogonal
[X, D] = eig(A) % the eigenvectors and the corresponding eigenvalues of A
sum(sqrt(sum(A .^ 2))) % l-2,1 norm of A
norm(A, 'fro') % Frobenius norm of A
sum(svd(A)) % nuclear norm of A
max(svd(A)) % spectral norm of A
```

Problem 3: Linear Equations

(1) It is evident to solve the linear equation

$$\begin{cases} x_1 = -1, \\ x_2 = 0, \\ x_3 = 1. \end{cases} \tag{3.1}$$

(2) Let

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \tag{3.2}$$

and we have $\mathbf{A}\mathbf{x} = \mathbf{b}$ as

$$\begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}. \tag{3.3}$$

(3) Since there is a unique solution shown in Eq. (3.1), we know

$$\operatorname{rank}(\mathbf{A}) = 3. \tag{3.4}$$

(4)

$$\begin{aligned}
[\mathbf{A}, \mathbf{I}_3] &= \left[\begin{array}{rrr:rrr} 2 & 2 & 3 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ -1 & 2 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rrr:rrr} 2 & 2 & 3 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{array}\right] \sim \left[\begin{array}{rrr:rrr} 1 & 1 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{array}\right] \\
&\sim \left[\begin{array}{rrr:rrr} 1 & 1 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 0 & -2 & -\frac{3}{2} & -\frac{1}{2} & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{array}\right] \sim \left[\begin{array}{rrr:rrr} 1 & 1 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 0 & -1 & -\frac{3}{4} & -\frac{1}{4} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} & \frac{3}{2} & 1 \end{array}\right] \sim \left[\begin{array}{rrr:rrr} 1 & 1 & 0 & 2 & -9 & -6 \\ 0 & -1 & 0 & -1 & 5 & 3 \\ 0 & 0 & 1 & -1 & 6 & 4 \end{array}\right] \\
&\sim \left[\begin{array}{rrr:rrr} 1 & 0 & 0 & 1 & -4 & -3 \\ 0 & 1 & 0 & 1 & -5 & -3 \\ 0 & 0 & 1 & -1 & 6 & 4 \end{array}\right],
\end{aligned} \tag{3.5}$$

therefore the inverse of $\mathbf{A}$ is

$$\mathbf{A}^{-1} = \begin{bmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{bmatrix}. \tag{3.6}$$

The determinant of $\mathbf{A}$ can be calculated as

$$\det(\mathbf{A}) = 2 \times \begin{vmatrix} -1 & 0 \\ 2 & 1 \end{vmatrix} - 2 \times \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} + 3 \times \begin{vmatrix} 1 & -1 \\ -1 & 2 \end{vmatrix} = 2 \times (-1) - 2 \times 1 + 3 \times 1 = -1. \tag{3.7}$$

(5) As is shown in Eq. (3.4), $\mathbf{A}$ is invertible and with the result in Eq. (3.6)

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \begin{bmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \tag{3.8}$$

and it is exactly the same result with Eq. (3.1).

(6) The inner product

$$\langle \mathbf{x}, \mathbf{b} \rangle = \mathbf{x}^T\mathbf{b} = 1 \times 1 + 0 \times (-1) + 1 \times 2 = 1, \tag{3.9}$$

and the outer product is

$$\mathbf{x} \otimes \mathbf{b} = \mathbf{x}\mathbf{b}^T = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 1 & -2 \\ 0 & 0 & 0 \\ 1 & -1 & 2 \end{bmatrix}. \tag{3.10}$$

(7) $\|\mathbf{b}\|_1 = |1| + |-1| + |2| = 4, \quad \|\mathbf{b}\|_2 = \sqrt{1^2 + (-1)^2 + 2^2} = \sqrt{6}, \quad \|\mathbf{b}\|_\infty = \max\{|1|, |-1|, |2|\} = 2.$

(8) Let $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}^T$, we have

$$\mathbf{y}^T\mathbf{A}\mathbf{y} = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = 2y_1^2 - y_2^2 + y_3^2 + 3y_1y_2 + 2y_2y_3 + 2y_1y_3, \tag{3.11}$$

and

$$\nabla_{\mathbf{y}}\, \mathbf{y}^T\mathbf{A}\mathbf{y} = \begin{bmatrix} \frac{\partial}{\partial y_1}\mathbf{y}^T\mathbf{A}\mathbf{y} \\ \frac{\partial}{\partial y_2}\mathbf{y}^T\mathbf{A}\mathbf{y} \\ \frac{\partial}{\partial y_3}\mathbf{y}^T\mathbf{A}\mathbf{y} \end{bmatrix} = \begin{bmatrix} 4y_1 + 3y_2 + 2y_3 \\ 3y_1 - 2y_2 + 2y_3 \\ 2y_1 + 2y_2 + 2y_3 \end{bmatrix}. \tag{3.12}$$

(9) The equation $\mathbf{A}_1\mathbf{x} = \mathbf{b}_1$ can be represented as

$$\begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 2 \end{bmatrix}. \tag{3.13}$$

(10) $\operatorname{rank}(\mathbf{A}_1) = 3$.

Proof. On one hand, $\operatorname{rank}(\mathbf{A}_1) \geq \operatorname{rank}(\mathbf{A}) = 3$ which is shown in Eq. (3.4). On the other hand, $\operatorname{rank}(\mathbf{A}_1) \leq \min\{3, 4\} = 3$. Therefore, $\operatorname{rank}(\mathbf{A}_1) = 3$. We can also find the first three equations are linearly independent while the last equation is actually the same with the third equation which makes it meaningless. □

(11) Yes.

Proof. Since $\operatorname{rank}(\mathbf{A}_1) = \|\mathbf{x}\|_0$, i.e. rank of $\mathbf{A}_1$ is equal to the dimension of $\mathbf{x}$, the formula can be solved with a unique solution the same as Eq. (3.1). □