Linear algebra for computational statistics II

Jong-June Jeon

September, 2022

Department of Statistics, University of Seoul

Things to know

- basic operation of matrix
- spanning space, null space
- projection and geometry
- linear map and matrix

Linear space

Linear space and matrix

Step 2

벡터는 숫자의 순서열로서 단순히 어떤 숫자들의 모임이지만, 내적이 정의된 공간에서의 원소로 이해할 수 있다. 벡터 공간의 원소로써 벡터를 다루면 앞서 정의한 벡터의 연산 과정을 시각화하여 더 깊은 이해를 얻을 수 있다. 행렬은 벡터공간에서 정의된 선형변환(함수)이며 행렬의 연산이 이 선형변환에 정의된 함수의 연산이라는 사실을 알 수 있다. 여기서는 벡터 공간의 정의, 행렬이 곧 선형변환임을 이해한다.

Vector space

- operation rule (addition, scalar multiplication, ...)
- completeness of elements : (identity and inverse)

Vector space: example

- $\mathbb{R}^n$ is vector space?
- The set of $\mathbb{R}^{p\times p}$ is vector space?

Before answering the above question, check the operation rules and elements of indentity in your vector space.

Let A be $m \times n$ matrix and x be n matrix (n dimensional comlumn vector).

- Write an example of A and x and compute Ax. Where does the result lie on?
- Choose an other x' and compute Ax'.
- Choose two constant a and b and compute $A(ax)$ and $A(bx')$ and $A(ax) + A(bx')$.
- Compute $A(ax + bx')$.

- Write an example of A and x and compute Ax. Where does the result lie on?
  A moves $x \in \mathbb{R}^n$ on $Ax \in \mathbb{R}^m$.
- Choose an other x' and compute Ax'.
  A also moves $x' \in \mathbb{R}^n$ on $Ax' \in \mathbb{R}^m$.
- Choose two constant a and b and compute $A(ax)$ and $A(bx')$ and $A(ax) + A(bx')$.
- Compute $A(ax + bx')$.

Note that $A(ax) + A(bx') = A(ax + bx')$, which implies that A moves elements in $\mathbb{R}^n$ to $\mathbb{R}^n$ with satisfying an special property.

Definition: Linear map

Let V and W be vector spaces and let $\mathcal{L}$ be map from V to W.

- $\mathcal{L}(x + y) = \mathcal{L}(x) + \mathcal{L}(y)$ for all $x, y \in V$
- $\mathcal{L}(cx) = c\mathcal{L}(x)$ for a scalar c.

Matrix and linear map

Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces, and consider a linear map $\mathcal{L}$ from $\mathcal{V}$ to $\mathcal{W}$. In particular, let $\mathcal{V} = \mathbb{R}^p$ and $\mathcal{W} = \mathbb{R}^n$, then $\mathcal{L}(\mathbf{0}) = \mathbf{0}$, and

$$\mathcal{L}(ax + bx') = a\mathcal{L}(x) + b\mathcal{L}(x')$$

for all $x, x' \in \mathbb{R}^p$ and all $a, b \in \mathbb{R}$.

Thus, $n \times p$ matrix can be regarded as a linear map. Moreover, we can consider one-to-one correspondence between linear map and matrix.

Matrix and linear map

- Matrix addition: let A and B be $n \times p$ matrix, and denote the corresponding linear map by $\mathcal{L}_A$ and $\mathcal{L}_B$. $A + B$ is also $n \times p$ matrix and $\mathcal{L}_{A+B}$ be the correspondent linear map to $A + B$. Then, $\mathcal{L}_{A+B} = \mathcal{L}_A + \mathcal{L}_B$.

$$(A + B)x = Ax + Bx$$

Matrix and linear map

- Matrix multiplication: let A and B be $n \times k$ and $k \times p$ matrix, and denote the corresponding linear map by $\mathcal{L}_A$ and $\mathcal{L}_B$. AB is $n \times p$ matrix and $\mathcal{L}_{AB}$ be the correspondent linear map to AB. Then, $\mathcal{L}_{AB} = \mathcal{L}_A \circ \mathcal{L}_B$ (Composition of functions: 합성함수)

$$x \mapsto Ax \mapsto B(Ax)$$

$$W \in \mathbb{R}^{n \times p} \text{ if and only if } W : \mathbb{R}^p \mapsto \mathbb{R}^n \text{ is linear.}$$

- When $n < p$ W is called a (linear) encoder (압축).
- When $n > p$ W is called a (linear) decoder (해제).

Let $W = [W_1, \cdots, W_p] \in \mathbb{R}^{n \times p}$ and $a = (a_1, \cdots, a_p)^\top \in \mathbb{R}^p$.

$$W(a) = a_1 W_1 + \cdots + a_p W_p \in \mathbb{R}^n$$

$W(a)$ is the image of W or the range of $\mathcal{L}_W$. Note that $W(a)$ is a linear combination of column vectors of W. Suppose that we gather all elements of $W(a)$ when $n > p$. This recovers $\mathbb{R}^n$? Or when $n \leq p$ this always recovers $\mathbb{R}^n$?.

Spanned column space

- Spanned column space of W is the range of W or $\mathcal{L}_W$.

$$\mathcal{C}(W) = \{\sum_{j=1}^{p} a_j W_j \in \mathbb{R}^n : a_j \in \mathbb{R}, 1 \leq j \leq p\}$$

- It is clear that $\mathcal{C}(W) \subset \mathbb{R}^n$.
- When $n > p$, how much rich $\mathcal{C}(W)$ is? (the dimension of $\mathcal{C}(W)$)

linear independence

Let V be vector space. A linear independence or linear relation among vectors $w_1, ..., w_n \in V$ is $a_1 w_1 + \cdots + a_n w_n = \mathbf{0}$ implies that all a_ks are zero.

dimension of vector space V

Let V be vector space and $v_1, ..., v_k \in V$. The dimension of V is the maximum number of k where $v_1, ..., v_k$ are linearly independent.

dimension of vector space $\mathcal{C}(W)$

The dimension of $\mathcal{C}(W)$ is the maximum number of k where $W_1, ..., W_k$ are linearly independent. The dimension of $\mathcal{C}(W)$ is called of the (column) rank of W $rank(W)$. It is known that

$$rank(W) = rank(W^\top)$$

Basis of V

If $w_1, ..., w_n \in V$ are linearly independent, and $\mathcal{C}([w_1, ..., w_n]) = V$, then $w_1, ..., w_n$ is called a basis of V. Here n is the dimension of V denoted by $\dim(V)$.

- Null space of $W \in \mathbb{R}^{n \times p}$:

$$\mathcal{N}(W) = \{a \in \mathbb{R}^p : Wa = 0\}$$

Dimensionality Theorem

$$dim(\mathcal{C}(W)) + dim(\mathcal{N}(W)) = p$$

When $dim(\mathcal{C}(W)) = p$, W is called full-column rank.

Basis of $\mathbb{R}^n$

If $w_1, ..., w_n \in \mathbb{R}^n$ are linearly independent, then the set $\{w_1, ..., w_n\}$ is called a basis of $\mathbb{R}^n$. Note that basis is not unique.

Basis of $\mathbb{R}^n$

Recall that $W \in \mathbb{R}^{n \times p}$ is a linear map

$$\mathcal{L} : a \in \mathbb{R}^p \mapsto Wa \in \mathbb{R}^n$$

We have seen that $\mathcal{C}(W)$ is the range of the $\mathcal{L}$ and the richness of the space is measured by $dim(\mathcal{C}(W))$, the column rank of W.

Matrix and linear map*

- Let $\{v_1, \cdots, v_p\}$ be ordered basis of $\mathbb{R}^p$ and $w_1, \cdots, w_p \in \mathbb{R}^n$. Then, there exists $\mathcal{L}$ such that it is the unique linear map from $\mathbb{R}^p$ to $\mathbb{R}^n$ and $\mathcal{L}(v_j) = w_j$. (The image of the basis in $\mathbb{R}^p$ uniquely determines the corresponding linear map.)
- (Matrix representation) Let $\{v_1, \cdots, v_p\}$ and $\mathbb{R}^p$ and $\{w_1, \cdots, w_n\} \in \mathbb{R}^n$ be basis of $\mathbb{R}^p$ and $\mathbb{R}^n$. A linear map $\mathcal{L}$ is completely characterized by p elements, r_j. Moreover r_j's are uniquely represented by $\{w_1, \cdots, w_n\}$. That is, $\mathcal{L}(v_j) = r_j = \sum_{i=1}^n a_{ij} w_i$ for $j = 1, \cdots, p$. That is, the matrix (a_{ij}) is the representation of the linear map $\mathcal{L}$ with the basis $\{v_1, \cdots, v_p\}$ and $\{w_1, \cdots, w_n\}$.

Useful linear map

- Identity linear map: identity matrix
- Elementary operations:
    - Let $e_j \in \mathbb{R}^p$ is a unit column vector where the jth element is 1 and $\pi = (\pi_1, \cdots, \pi_n)$ is a permutation of $(1, \cdots, p)$, where $\pi_j \in \{1, \cdots, p\}$ for $j = 1, \cdots, n$. Then, $E_\pi = (e_{\pi_1}, \cdots, e_{\pi_n})' \in \mathbb{R}^n \times \mathbb{R}^p$ is a linear map that rearranges the elements according to π.

$$E_\pi = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad x = (x_1, x_2, x_3)',$$

then $E_\pi x = (x_3, x_1, x_2)'$

Useful linear map*

- Elementary operations:
    - Let $n = p$ and $E_\pi = (0, \cdots, 0, e_{\pi_k}, 0, \cdots, 0)' \in \mathbb{R}^n \times \mathbb{R}^p$. What is this operation $I + aE_\pi$ with $a \in \mathbb{R}$?
    - Suppose that $E_\pi X$ is well defined, then what is the operational meaning of the E_π?
    - Suppose that XE'_π is well defined, the what is the operational meaning of E'_π?