Mathematics in DL: **Jan 16 2018**

Lecture 1: Linear Algebra

Lecturer: Juhee Lee *Scribes: Juhee Lee*

1.1 Numbers and Arrays

먼저 앞으로 사용하게될 용어를 정리합니다

Notations

- scalar(스칼라)는 “하나의 숫자”를 의미합니다. ex) $3, -5, \ldots$

  자연수($\mathbb{N}$), 정수($\mathbb{I}$), 유리수($\mathbb{Q}$), 실수($\mathbb{R}$), 복소수($\mathbb{C}$)등으로 분류합니다.

  만일, ’3은 실수이다.’란 문장은 ’3은 실수에 속한다.’고 표현하고 수학적 표현은 $3 \in \mathbb{R}$ 입니다.

- array(배열)란 n개의 원소의 나열 즉, $\mathbf{a} = (a_1, \cdots, a_n)$ 이고, 이를 “n차원 tuple” 이라 합니다.

- vector는 n개의 성분의 나열로 1-D 배열입니다. 즉, $\mathbf{a} = (a_1, \cdots, a_n)$ 이고, 각 성분 a_i 의 속성 혹은 데이터 타입이 같아야 합니다.

  Note. 벡터는 기하학적 의미로 사용(by directed line segment), 벡터의 데이터 성분으로는 벡터, 행렬, 함수 모두 가능합니다.

- matrix는 2-D 배열이다. 즉,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & a_{ij} & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{m \times n}$$

- tensor는 3-D 이상을 표현할 수 있는 배열이다. 즉, 각성분이 행렬로 구성된 배열이다.

앞에서 살펴본 용어에 대해 자세히 살펴보도록 하겠습니다.

vectors

Definition 1.1 *A vector $\boldsymbol{a}$ is an array of numbers*

- vector의 성분은 행렬, 함수 등 같은 형식의 성분이 들어오면 됩니다.
- row vector $\mathbf{a} = (a_1, \cdots, a_n)$
- column vector $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$

지금부터 벡터$\mathbf{a}$를 comlumn vector라고 놓습니다

- "transpose" (column vector)^T = row vector

수들이 존재하면 그들간의 연산 즉 operation을 정의합니다.

Operations of vectors

- "+ / -" 연산의 정의에서 제일 중요한것은 연산이 가능한지?? 를 찾아보는 것인데, 먼저 dimension check 후 연산 가능합니다.

  $\mathbf{a} + \mathbf{b} = (a_1 + b_1, \cdots, a_n + b_n)$

- "곱하기?" 연산의 정의에 따라 여러가지가 가능합니다.

  1. "scalar multiplication"

     $c \in \mathbb{R}, c\mathbf{a} = c\,(a_1, \cdots, a_n) = (ca_1, \cdots, ca_n)$
  2. "Hadamard product (elementwise multiplication)"

     $\mathbf{a} \odot \mathbf{b} = (a_1 + b_1, \cdots, a_n + b_n)$
  3. "dot product (Euclidean inner product)" (not just inner product)

     파이썬에선 명령어 inner product

     $\mathbf{a} \cdot \mathbf{b} = (a_1 + b_1 + \cdots + a_n + b_n) = \sum_{i=1}^{n} a_i b_i$ 기호 정리
  4. "cross product"

     $$\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1) = \det \begin{pmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} = (||a||||b||sin\theta\ n))$$

Matrices

Definition 1.2 *A matrix* $\mathbf{A}$ *is an 2D array of numbers*

$\mathbf{A}$ 는 $m \times n$ *matrix*

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & a_{ij} & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix} = (a_{ij})_{m \times n}$$

Operations of matrices

- "+" 연산은 같은 사이즈의 행렬의 연산만 가능합니다.

  $\mathbf{A} + \mathbf{B} = (a_{ij}) + (b_{ij})_{m \times n}$

- "transpose" $\mathbf{A} = (a_{ij})_{m \times n} \Rightarrow \mathbf{A}^T = (a_{ji})_{n \times m}$

  "벡터에서의 transpose도 행렬표현으로 가능"

- "곱하기?" vector에서의 곱하기연산 다 가능합니다.

1. “scalar multiplication”
   $c \in \mathbb{R}, c\mathbf{A} = (ca_{ij})_{m\times n}$
2. “Hadamard product (elementwise multiplication)”
   $\mathbf{A} \odot \mathbf{B} = (a_{ij}b_{ij})_{m\times n}$
3. “product” (matrix multiplication)
   $\mathbf{AB} = (a_{ij})_{m\times p}(b_{ij})_{p\times n} = (c_{ij})_{m\times n}$ where $c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj} = \mathbf{a}_i^T \cdot \mathbf{b}_j$
   즉 c_{ij} 는 (i-th row of A)·(j-th column of B)
   $\begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \mathbf{b}_2 \cdots \mathbf{b}_n \end{pmatrix}$
4. “tensor product” (Kronecker product)

$$\mathbf{A}_{m\times n} \otimes \mathbf{B}_{p\times q} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & a_{13}\mathbf{B} & \ldots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & a_{23}\mathbf{B} & \ldots & a_{2n}\mathbf{B} \\ \vdots & \vdots & a_{ij}\mathbf{B} & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & a_{m3}\mathbf{B} & \ldots & a_{mn}\mathbf{B} \end{pmatrix}$$

   Quiz. $\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \otimes \begin{pmatrix} 3 & 0 \\ 4 & -5 \end{pmatrix}$
   solution.

$$\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \otimes \begin{pmatrix} 3 & 0 \\ 4 & -5 \end{pmatrix} \quad = \quad \begin{pmatrix} 1\begin{pmatrix} 3 & 0 \\ 4 & -5 \end{pmatrix} & 2\begin{pmatrix} 3 & 0 \\ 4 & -5 \end{pmatrix} \\ 2\begin{pmatrix} 3 & 0 \\ 4 & -5 \end{pmatrix} & -1\begin{pmatrix} 3 & 0 \\ 4 & -5 \end{pmatrix} \end{pmatrix}$$

Quiz. $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 \mathbf{a}_2 \end{pmatrix}$ $\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 \mathbf{b}_2 \end{pmatrix}$. 이때, $\mathbf{a}_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, $\mathbf{a}_2 = \begin{pmatrix} 4 \\ 5 \\ 3 \end{pmatrix}$, $\mathbf{b}_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$, $\mathbf{b}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

1. $3\mathbf{A} + 4 =$을 계산하여라.
2. $3\mathbf{A}^T + \mathbf{b}_1$ 을 계산하여라.
3. 다음 중 matrix multiplication 계산이 가능한 것은?

1) $\begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \mathbf{b}_2 \end{pmatrix}$ 2) $\begin{pmatrix} \mathbf{b}_1 \mathbf{b}_2 \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \mathbf{a}_2 \end{pmatrix}$ 3) $\begin{pmatrix} \mathbf{a}_1 \mathbf{a}_2 \end{pmatrix} \begin{pmatrix} \mathbf{b}_1^T \\ \mathbf{b}_2^T \end{pmatrix}$

알아두면 좋은 Tip!

Note.

- $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
- $\mathbf{A}(\mathbf{BC}) = \mathbf{A}(\mathbf{BC})$: 세개의 수의 동시 연산은 수학에선 없다. 한번에 두개의 개산을 묶는 것이 기본 룰

  Quiz. $\mathbf{AB} \overset{?}{=} \mathbf{BA}$
- $\mathbf{a}^T\mathbf{b} = \mathbf{b}^T\mathbf{a}$

- $(\mathbf{A}^T)^T = \mathbf{A}$
- $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
- $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$

  Quiz. $(c\mathbf{A})^T = ?$

1.2 Identity matrix and Inverse matrix

Definition 1.3 $\mathbf{I}_n$ *is the identity matrix of* n

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Definition 1.4 *The matrix inverse of* $\mathbf{A}$ *is denoted as* $\mathbf{A}^{-1}$ *and is defined as the matrix such as*

$$\mathbf{AB} = \mathbf{I}_n = \mathbf{BA}, \quad \mathbf{B} = \mathbf{A}^{-1}$$

*역행렬은 존재하면 유일!

Inverse matrix 활용

- linear system $\mathbf{Ax} = \mathbf{b}$의 해를 구하는 방법!
  - $\mathbf{A}^{-1}$가 존재하면, 해는 $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.
  - $\mathbf{A}^{-1}$를 구할 수 없는 rectangular matrix에서, $\mathbf{Ax} = \mathbf{b}$는 pseudo inverse의 개념을 이용해서 방정식을 근을 구합니다.

Determinant of a square($n \times n$) matrix

Definition 1.5 *The determinant of* $\mathbf{A} = (a_{ij})_{n\times n}$ *is defined as*

$$\det(\mathbf{A}) = |A| = \sum_{j=1}^{n} \boldsymbol{a}_{1j}\mathbf{C}_{1j} = \boldsymbol{a}_{11}\mathbf{C}_{11} + \boldsymbol{a}_{22}\mathbf{C}_{22} + \cdots + \boldsymbol{a}_{nn}\mathbf{C}_{nn}$$

where the minor $\mathbf{M}_{ij}$ *of the element* $\boldsymbol{a}_{ij}$ *is the determinant of the matrix obtained by deleting the ith row and jth column of* $\mathbf{A}$. *The cofactor* $\mathbf{C}_{ij}$ *is given by* $\mathbf{C}_{ij} = (-1)^{i+j}\mathbf{M}_{ij}$

* determinant 의미는 parallel pipe의 부피

- $\mathbf{A}^{-1}$가 존재하려면, $\det(\mathbf{A}) \neq 0$.

  Quiz. $\det(\mathbf{A})^{-1} = ?$ 1/det(A)

  Quiz. $\det(\mathbf{A})^T = ?$ det(A)

  ex. Determinant를 구하여라.

1. $\mathbf{A} = \begin{pmatrix} 2 & -3 \\ 1 & 2 \end{pmatrix} \Rightarrow |A| = 7$
2. $\mathbf{B} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} \Rightarrow |B| = 0$
3. $\mathbf{C} = \begin{pmatrix} -1 & 3 & 2 \\ 0 & -2 & 1 \\ 1 & 0 & -2 \end{pmatrix} \Rightarrow |C| = 1/3$

   $\mathbf{C}_{11} = -1, \mathbf{C}_{12} = 5, \mathbf{C}_{13} = 4$, M도 체크.

ex. Determinant 이용해서 $\mathbf{C}^{-1}$를 구하여라.

$$\mathbf{C} = \begin{pmatrix} -1 & 3 & 2 \\ 0 & -2 & 1 \\ 1 & 0 & -2 \end{pmatrix}$$

step1. matrix of cofactors $\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \mathbf{C}_{23} \\ \mathbf{C}_{31} & \mathbf{C}_{32} & \mathbf{C}_{33} \end{pmatrix} = \begin{pmatrix} 4 & 1 & 2 \\ 6 & -2 & 3 \\ 7 & 1 & 2 \end{pmatrix}$

step2. transpose => adjoint matrix of C i.e, $adj(\mathbf{C}) = \begin{pmatrix} 4 & 6 & 7 \\ 1 & 0 & 1 \\ 2 & 3 & 2 \end{pmatrix}$

step3. $\mathbf{C}^{-1} = \frac{1}{\det(\mathbf{C})} adj(\mathbf{C}) = 1/3 \begin{pmatrix} 4 & 6 & 7 \\ 1 & 0 & 1 \\ 2 & 3 & 2 \end{pmatrix}$

- 증명은, $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$로 확인해보세요.

1.3 Norm

Norm

- vector

  $\|\mathbf{A}\|_1 = \sum_{i=1}^{n} |a_i|$

  $\|\mathbf{A}\|_\infty = \max_{1 \le i \le n} |a_i|$

  $\|\mathbf{A}\|_2 = \left(\sum_{i=1}^{n} |a_i|^2\right)^{1/2}$

- matrix $\mathbf{A}_{m \times n}$

  $\|\mathbf{A}\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|$

  $\|\mathbf{A}\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|$

  $\|\mathbf{A}\|_2 = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2\right)^{1/2} = (trace(A^* A))^{1/2}$ where $A^* = conjugatetranpose$

ex. $\mathbf{A} = \begin{pmatrix} 4 & 6 & 7 \\ 1 & 0 & 1 \\ 2 & 3 & 2 \end{pmatrix}$

$\|\mathbf{A}\|_1 = \max(4+1+2; 6+0+3; 7+1+2) = \max(7,9,10) = 10$

$\|\mathbf{A}\|_\infty = \max(4+6+7; 1+0+1; 2+3+2) = \max(17,2,7) = 17$

$\|\mathbf{A}\|_2 = \sqrt{4^2+6^2+7^2+1r+0^2+1^2+2^2+3^2+2^2}$

1.4 Matrix decomposition

decomposition

- $\mathbf{A}_{n\times n} = \mathbf{LU}$ L: lower triangular matrix, U: upper triangular matrix
- $\mathbf{A}_{n\times n} = \mathbf{QR}$ Q: orthogonal matrix, R: upper triangular matrix
- $\mathbf{A}_{n\times n} = \mathbf{PDP}^{-1}$ P: column matrix formed from eigenvector of A D: diagonal matrix with eigenvalue
- $\mathbf{A}_{m\times n} = \mathbf{U}\sum\mathbf{V}^T$ U,V: orthogonal matrix formed from eigenvector of AA^T,A^TA, $\sum$: diagonal matrix with eigenvalue$^{1/2}$
- 실제, 프로그래밍에서는 이름만 입력하면 분해가능합니다.

References

[1] http://cs229.stanford.edu/section/cs229-linalg.pdf

[2] Advanced Engineering Mathematics ch7-8, Erwin Kreyszig, Wiely

[3] Elementary Linear Algebra, 7th ed., Ron Larson

[4] Deep Learning ch2, Ian Goodfellow and Yoshua Bengio and Aaron Courville, The MIT Press

[5] https://en.wikipedia.org