9/1/2021

last time:
→ discussed matrix-vector and matrix-matrix multiplication.
→ trace of a matrix.

Lemma: $\underline{\underline{A}}, \underline{\underline{B}} \in \mathbb{R}^{n \times n}$, Then

$$tr(\underline{\underline{A}}\,\underline{\underline{B}}) = tr(\underline{\underline{B}}\,\underline{\underline{A}}).$$

Proof let $\underline{\underline{C}} = \underline{\underline{A}}\,\underline{\underline{B}} = (c_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$

The diagonal element c_{kk} is the k^{th} row of $\underline{\underline{A}}$, $\hat{\underline{a}}_k^T$, times the k^{th} column of $\underline{\underline{B}}$, $\underline{b}_k$

$$c_{kk} = \hat{\underline{a}}_k^T \underline{b}_k = \sum_{i=1}^{n} a_{ki} b_{ik}.$$

So, $tr(\underline{\underline{A}}\,\underline{\underline{B}}) = tr(\underline{\underline{C}}) = \sum_{k=1}^{n} \sum_{i=1}^{n} a_{ki} b_{ik}$

If $\underline{\underline{D}} = \underline{\underline{B}}\,\underline{\underline{A}}$, then

$$d_{kk} = \hat{b}_k^T a_k = \sum_{i=1}^{n} b_{ki} a_{ik}$$

and, $tr(\underline{\underline{B}}\,\underline{\underline{A}}) = tr(\underline{\underline{D}}) = \sum_{k=1}^{n} \sum_{i=1}^{n} b_{ki} a_{ik}$

$$\Rightarrow tr(\underline{\underline{B}}\,\underline{\underline{A}}) = tr(\underline{\underline{A}}\,\underline{\underline{B}}).$$

Definition: (transpose)

let $\underline{\underline{A}} \in \mathbb{R}^{m\times n}$. The transpose of $\underline{\underline{A}}$, $\left(\underline{\underline{A}} = (a_{ij})\right)$.

denoted $\underline{\underline{A}}^T$, is defined as

$$\underline{\underline{A}}^T = (a_{ji}). \in \mathbb{R}^{n\times m}.$$

Example: $\underline{\underline{A}} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \\ -5 & \pi \end{pmatrix}.$

then $\underline{\underline{A}}^T$ is defined as $\underline{\underline{A}}^T = \begin{pmatrix} 2 & 4 & -5 \\ 1 & 2 & \pi \end{pmatrix}.$

Fact: $(\underline{\underline{A}}\,\underline{\underline{B}})^T = \underline{\underline{B}}^T \underline{\underline{A}}^T.$

Remark: If $\underline{v} \in \mathbb{R}^n$ is a (column) vector of length n, we can view $\underline{v}$ as an $n\times 1$ matrix. As we have been using before, $\underline{v}^T$ is the (row) vector corresponding to $\underline{v}$ that we can view as a $1\times n$ matrix.

<u>Remark</u>: The space of matrices of size m by n (m rows and n columns) itself forms a vector space. This space is what we call

$$\mathbb{R}^{m \times n}.$$

Since we can define norms of vectors, we can define the norm of a matrix. There are many types of matrix norms ... one "Natural" one is the Frobenius norm:

<u>Definition</u>: let $\underline{\underline{A}} \in \mathbb{R}^{m \times n}$. The <u>Frobenius</u> <u>norm</u> of $\underline{A}$ is

$$\|\underline{\underline{A}}\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}.$$

This is "sort of like" the Euclidean norm of a vector generalized to a matrix.

<u>Fact</u>: $\|\underline{\underline{A}}\|_F = \left(tr(\underline{\underline{A}}\,\underline{\underline{A}}^T) \right)^{1/2}.$

Definition: let $\underline{u}, \underline{v} \in \mathbb{R}^n$.

Recall that we can think of the vectors $\underline{u}$ and $\underline{v}$ as $n \times 1$ matrices, with $n = \#$ of rows
$1 = \#$ of columns.

The outer product of $\underline{u}$ and $\underline{v}$ is denoted $\underline{u}\,\underline{v}^T$ and is defined as the $n \times n$ matrix:

$$\underline{u}\,\underline{v}^T = \begin{pmatrix} | & | & & | \\ v_1\underline{u} & v_2\underline{u} & \cdots & v_n\underline{u} \\ | & | & & | \end{pmatrix},$$

i.e. all column of $\underline{u}\underline{v}^T$ are scalar multiples of $\underline{u}$.

Ex: Calculate the Frobenius norm of $\underline{\underline{A}} = \begin{pmatrix} 1 & 2 \\ -3 & 4 \end{pmatrix}$.

$$\|\underline{\underline{A}}\|_F = \left(1^2 + 2^2 + (-3)^2 + (4)^2\right)^{1/2}$$

$$= \left(1 + 4 + 9 + 16\right)^{1/2}$$

$$= \sqrt{30}.$$

An aside: What really is a matrix?

Answer: It is a representation of a linear map between vector spaces with respect to the "standard basis."

Definition: let $\underline{e}_i \in \mathbb{R}^m$ be a vector with zeros in every component, except 1 in the i^{th} component.

Ex: $\underline{e}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $\underline{e}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}$, $\underline{e}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}$.

$\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ is a collection of vectors that is called the standard basis of $\mathbb{R}^3$.

Why? Because any vector $\underline{a} = \begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix} \in \mathbb{R}^3$ can be represented as a sum of $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$

$$\underline{a} = \begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix} = a_1\underline{e}_1 + a_2\underline{e}_2 + a_3\underline{e}_3.$$

Definition: A function $T: \mathbb{R}^m \to \mathbb{R}^n$ is called a linear transformation if it satisfies, $\forall\, \underline{u}, \underline{v} \in \mathbb{R}^m$ and $\forall\, \gamma, \beta \in \mathbb{R}$: $T(\gamma\underline{u} + \beta\underline{v}) = \gamma T(\underline{u}) + \beta T(\underline{v})$.

Since a linear transformation can be "split" across sums of vectors, if we know

$$T(\underline{e}_i) = ? \quad \forall\, i = 1, \dots, m$$

then we know everything about the linear transformation.

Consider the case $m = n = 3$.

Assume we know:

$$T(\underline{e}_1) = T\left(\begin{pmatrix}1\\0\\0\end{pmatrix}\right) = \underline{a}_1$$

$$T(\underline{e}_2) = T\left(\begin{pmatrix}0\\1\\0\end{pmatrix}\right) = \underline{a}_2$$

$$T(\underline{e}_3) = T\left(\begin{pmatrix}0\\0\\1\end{pmatrix}\right) = \underline{a}_3$$

Then for any vector $\underline{u} = \begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix} \in \mathbb{R}^3$,

$$T(\underline{u}) = T\left(\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}\right) = T\left(u_1\underline{e}_1 + u_2\underline{e}_2 + u_3\underline{e}_3\right)$$

$$\left(\begin{array}{l}\text{using the fact}\\\text{that } T \text{ is}\\\text{a linear}\\\text{transformation}\end{array}\right) = u_1 T(\underline{e}_1) + u_2 T(\underline{e}_2) + u_3 T(\underline{e}_3)$$

$$= u_1 \underline{a}_1 + u_2 \underline{a}_2 + u_3 \underline{a}_3$$

$$= \begin{pmatrix} | & | & | \\ \underline{a}_1 & \underline{a}_2 & \underline{a}_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}.$$

$$= \underline{\underline{A}}\, \underline{u}.$$

In summary, $\underline{\underline{A}}$ is a representation of a linear transformation T, where the columns of $\underline{\underline{A}}$ are

T applied to the standard basis vectors.