Lecture #1. 9-4-18

- Review syllabus
- Prereqs: - Linear algebra
  - Calculus
  - Some idea about statistics: mean, variance, etc.
- Where are we going?

Review of linear algebra: let $A \in \mathbb{R}^{m \times n}$ mean

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & & & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & & & a_{mn} \end{pmatrix}$$

i.e. A is a matrix of real numbers with m rows and n columns.

- If $A \in \mathbb{R}^{n \times n}$, i.e. $m=n$ above, A is square.

Recall that A acts on a vector $\vec{x} \in \mathbb{R}^{n \times 1}$ on the left as:

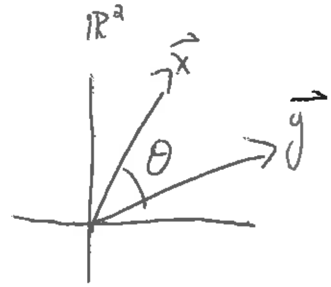
$$A\vec{x} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & & & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & & & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \in \mathbb{R}^{m \times 1}$$

If A is itself an element of $\mathbb{R}^{n \times 1}$, i.e. $\underset{\vec{y}}{\overset{}{A}}$ is a vector, we get:

$\vec{y}\cdot\vec{x} = \sum_{i=1}^{n} x_i y_i \in \mathbb{R}^{1\times 1}$. This is sometimes called the *dot product* of $\vec{x}$ and $\vec{y}$.

The dot product $\vec{x}\cdot\vec{y} = \sum_{i=1}^{n} x_i y_i$ is an example of an *inner product space*, and leads to a notion of *angle* between $\vec{x}, \vec{y}$:



$$\angle(\vec{x},\vec{y}) = \arccos\left(\frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\cdot\|\vec{y}\|}\right),$$

where $\|\vec{x}\| = \sqrt{\vec{x}\cdot\vec{x}}$

$$\|\vec{x}\| = \sqrt{\vec{x}\cdot\vec{x}} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

is the (ℓ^2)-norm of $\vec{x}$. It is a notion of length. It is common to consider vectors $\vec{x}$ with $\|\vec{x}\|=1$, i.e. *normalized* $\vec{x}$, in data science.

- Suppose $A = A^T$, where $(A^T)_{ij} = A_{ji}$, i.e. A is equal to its transpose. Then we say A is *symmetric*.
- Symmetric matrices have a powerful analytic tool available: the eigendecomposition/spectral decomposition.

- Any symmetric matrix $A \in \mathbb{R}^{n \times n}$ may be written as

$$A = U^T \Lambda U,$$

where 1) U is orthogonal, i.e. $U = \begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{pmatrix}$ where 1) $\vec{u}_i \in \mathbb{R}^{1 \times n}$ (*)
2) $\vec{u}_i \cdot \vec{u}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

2) $\Lambda = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$ is a diagonal matrix

Note that since the rows of U satisfy (*), $U^T U = U U^T = I$

$$= \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

Why is this valuable? It can be shown that the rows of U form a <u>basis</u> for $\mathbb{R}^n$, i.e. $\forall \vec{x} \in \mathbb{R}^{n \times 1}$, $\exists\, \alpha_1, \dots, \alpha_n \in \mathbb{R}$ s.t. "such that"

$$\vec{x} = \sum_{i=1}^{n} \alpha_i \vec{u}_i^{\,T}$$

Then, in the "U basis,"

$$
\begin{aligned}
& A\vec{x} \\
&= U^T \Lambda U \vec{x} \\
&= U^T \Lambda U \left(\sum_{i=1}^{n} \alpha_i \vec{u}_i^{\,T} \right) \\
&= U^T \Lambda \begin{pmatrix} \vec{u}_1 \\ \vdots \\ \vec{u}_n \end{pmatrix} \left(\sum_{i=1}^{n} \alpha_i \vec{u}_i^{\,T} \right)
\end{aligned}
$$

$$U^T \Lambda \begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots \\ \alpha_n\end{pmatrix}$$

$$= U^T \begin{pmatrix}\lambda_1\alpha_1\\ \lambda_2\alpha_2\\ \vdots\\ \lambda_n\alpha_n\end{pmatrix}$$

$$= \sum_{i=1}^n \lambda_i \alpha_i \vec{u_i}^T$$

- In summary, $\vec{x} = \sum_{i=1}^n \alpha_i \vec{u_i}^T \xmapsto{A} \sum_{i=1}^n \alpha_i \lambda_i \vec{u_i}^T$, so that A simply _scales_ $\alpha_i \mapsto \lambda_i \alpha_i$. This is much simpler than what a typical matrix-vector multiplication does.

- Phrased differently, the rows of U are _eigenvectors_ of A satisfying

$$A\vec{u_i}^T = \lambda_i \vec{u_i}^T$$

Remark: One could equivalently write $A = U \Lambda U^T$, where the columns of U are the eigenvectors.