Course: Comtemporary Algorithms T.II/2019-20

Lecture 12: Linear Algebra

12 February 2020

Lecturer: Dr. Kanat T. *Scribe: Pitipat C. & Nuttapat K.*

1 Notations

Vector

$$\begin{aligned}\vec{v} &= (v_1, v2, \dots, v_d) \in \mathbb{R}^d \\ &= [v_1, v_2, \dots, v_d]^T \\ &= \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{bmatrix}\end{aligned}$$

Matrix

$$\begin{aligned}A &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{11} & a_{12} & \dots & a_{1d} \\ \vdots & & \ddots & \\ a_{n1} & \dots & \dots & a_{nd} \end{bmatrix} \in \mathbb{R}^{nxd} \\ &= [\vec{a_1}, \vec{a_2}, \dots, \vec{a_d}] && \text{(MathLab Syntax)} \\ &= \begin{bmatrix} -\vec{a_1}- \\ -\vec{a_2}- \\ -\vdots- \\ -\vec{a_d}- \end{bmatrix}\end{aligned}$$

Linear Equations

$$\underbrace{\begin{matrix} ax_1 & +bx_2 & +cx_3 = d \\ ex_1 & +fx_2 & +gx_3 = h \\ & \| & \\ \end{matrix}}$$

$$\underbrace{\begin{bmatrix} a & b & c \\ e & f & g \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} d \\ h \end{bmatrix}}_{Ax=b}$$

2 Vector-Vector Products

Let

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}, \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{bmatrix}, \vec{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_d \end{bmatrix} \qquad \vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^d$$

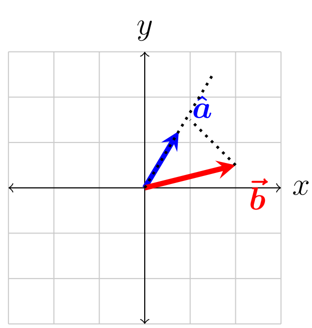
$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \qquad \vec{v} \in \mathbb{R}^n$$

2.1 Inner Product (Dot Product)

$$\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y} = \langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^{d} x_i y_i$$

Dot product is **linear**.

$$\langle \alpha\vec{x}, \vec{y} \rangle = \alpha \langle \vec{x}, \vec{y} \rangle \tag{1}$$
$$\langle \vec{x}, \vec{y} + \vec{z} \rangle = \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle \tag{2}$$



$$\langle \vec{a}, \vec{b} \rangle = \text{length of the projection of } \vec{b} \text{ onto } \vec{a}$$
$$\langle \vec{a}, \vec{b} \rangle = \|\vec{a}\| \left\|\vec{b}\right\| cos\theta$$

2.2 Outer Product

$$\vec{v} \cdot \vec{y}^T = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \dots & y_d \end{bmatrix} = \begin{bmatrix} d \\ h \end{bmatrix} = \begin{bmatrix} v_1 y_1 & v_1 y_2 & \dots & v_1 y_d \\ v_2 y_1 & v_2 y_2 & \dots & v_2 y_d \\ \vdots & & \ddots & \\ v_n y_1 & \dots & \dots & v_n y_d \end{bmatrix}$$

3 Norms

3.1 Euclidean norm (aka. "length")

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \qquad \vec{x} \in \mathbb{R}^d$$

$$\|\vec{x}\|_2 = \sqrt{\sum_i^d x_i^2} = \sqrt{\langle \vec{x}, \vec{x} \rangle}$$

$$\|\vec{x}\|_p = (\sum_i^d x_i^p)^{\frac{1}{p}}$$

$$\|\vec{x}\|_\infty = max_i^d |x_i|$$

3.2 Frobenius Norm

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{11} & a_{12} & \dots & a_{1d} \\ \vdots & & \ddots & \\ a_{n1} & \dots & \dots & a_{nd} \end{bmatrix} \in \mathbb{R}^{nxd}$$

$$\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^d (a_{ij})^2}$$

$$= \sqrt{\sum_{i=1}^n a_i^T a_i}$$

3.3 Spectral Norm

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{11} & a_{12} & \dots & a_{1d} \\ \vdots & & \ddots & \\ a_{n1} & \dots & \dots & a_{nd} \end{bmatrix} \in \mathbb{R}^{nxd}$$

$$
\begin{aligned}
\|A\| &= \|A\|_2 \\
&= max_{\vec{x}\in\mathbb{R}^m, \vec{x}\neq\vec{0}} \frac{\|A\vec{x}\|}{\|\vec{x}\|} \\
&= max_{\hat{x}\in\mathbb{R}^m, \|\hat{x}\|=1} \|A\hat{x}\|_2
\end{aligned}
$$

Simply put, find a unit vector $\hat{x}$ that maximize the length regardless of direction.

Fact 3.1. *if A is symmetric,* $\|A\| = \lambda_{max}$

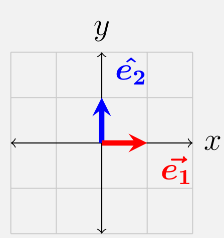
4 Linear (In)dependence

$$X = \{\vec{x_1}, \vec{x_2}, \dots, \vec{x_k}\} \subseteq \mathbb{R}^d$$

Span

$$Span(X) = \left\{ z | z = \sum_{i=1}^{n} \alpha_i \vec{x_i}, \alpha_i \in \mathbb{R} \right\}$$

Example



$$
\begin{aligned}
Span(\{\vec{e_1}, \vec{e_2}\}) &= \mathbb{R}^2 \\
(x, y) &= x \cdot \vec{e_1} + y \cdot \vec{e_2} \qquad\qquad \forall x, y \in \mathbb{R}
\end{aligned}
$$

Definition 4.1. Linearly dependent: If $\vec{z} \in Span(X)$, $\vec{z}$ is **linearly dependent** on X.

Definition 4.2. Basis: If $Span(X) = V$, X forms a **basis** for V.

Definition 4.3. Linearly Independent: X is **linearly independent** if $\forall i, x_i$ is linearly independent of $X|\{x_i\}$. In other words, x_i cannot be formed by linear combination of $X|\{x_i\}$.

5 Rank

Definition 5.1. Rank: rank(X) is the size of the largest subset of X that are linearly independent.

$$A = \begin{bmatrix} -\vec{a_1}- \\ -\vec{a_2}- \\ -\vdots- \\ -\vec{a_n}- \end{bmatrix}_{nxd}$$

$$\begin{aligned} rank(A) &= rank(\{\vec{a_1}, \ldots, \vec{a_n}\}) \\ &\leq min(n, d) \end{aligned}$$

Fact 5.2. *Row rank = Column rank*

Definition 5.3. Full Rank: matrix A has **full rank** if $rank(A) = min(n, d)$.

6 Inverse

Definition 6.1. Full Rank: matrix A is **square** if number of rows is equal to number of columns.

A square matrix A may or may not have an inverse, but if the inverse, A^{-1}, exists, it is the unique matrix satisfying

$$A^{-1}A = AA^{-1} = \underbrace{I}_{identitymatrix} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

Theorem 6.2. *Let A be a square matrix,*

$$\textit{A is invertible} \iff \textit{A has a full rank}$$

7 Orthogonality

Definition 7.1. Let $\vec{x}, \vec{y} \in \mathbb{R}^d$,

$$\langle \vec{x}, \vec{y} \rangle = 0 \equiv \vec{x} \text{ and } \vec{y} \text{ are } \textbf{orthogonal}$$

Definition 7.2. Let A be a matrix, if every column c_i of A is a unit vector (normalized with $\|c_i\| = 1$) and each column is orthogonal to other column vectors, A is **orthonormal**.

Definition 7.3. if every column c_i of A is a unit vector (normalized with $\|c_i\| = 1$) and each column is orthogonal to other column vectors, A is **orthonormal**.

Definition 7.4. A square matrix $U \in \mathbb{R}^{nxn}$ whose rows and columns are orthonormal is called an **orthogonal matrix**.

Fact 7.5. *Let $U \in \mathbb{R}^{nxn}$ be an orthogonal matrix,*

$$U^T U = I = UU^T \tag{3}$$
$$U^{-1} = U^T \tag{4}$$

Lemma 7.6. *If Q is orthonormal, then $\forall \vec{x} \in \mathbb{R}^n$, $\|Q\vec{x}\|_2 = \|\vec{x}\|_2$*

Proof.

$$\begin{aligned}
\|Q\vec{x}\|_2^2 &= (Q\vec{x})^T(Q\vec{x}) \\
\|Q\vec{x}\|_2^2 &= \vec{x}^T \underbrace{Q^T Q}_{I} \vec{x} \\
\|Q\vec{x}\|_2^2 &= \vec{x}^T\vec{x} \\
\|Q\vec{x}\|_2^2 &= \|\vec{x}\|_2^2 \\
\|Q\vec{x}\|_2 &= \|\vec{x}\|_2
\end{aligned}$$

□

Lemma 7.7. *If Q is orthonormal, then for any matrix A*

$$\|QA\| = \|A\|$$

Proof.

$$\begin{aligned}
\|Q\vec{x}\|_2 &= max_{\hat{x}\in\mathbb{R}^n, \|\hat{x}\|=1} \left\| Q \underbrace{(A\hat{x})}_{\text{also a vector}} \right\|_2 \\
&= max_{\hat{x}\in\mathbb{R}^n, \|\hat{x}\|=1} \|A\hat{x}\|_2 && \text{(by Lemma 7.6)} \\
&= \|A\| && \text{(by definition of Spectral norm)}
\end{aligned}$$

□

8 Eigen Value & Eigen Vector

Let A be $n \times n$ matrix then

$$Ax = \lambda x$$

We call x is the "Eigen Vector" of A with the corresponding "Eigen Value" λ. The intuition here is that x maintains the direction when multiply by A, hence it only changes the length.

It is possible that for eigen value, λ,

$$\begin{aligned} Ax &= \lambda x \\ Ay &= \lambda y \\ &\text{then} \\ A(\alpha x + \beta y) &= \lambda(\alpha x + \beta y) \end{aligned}$$

Definition 8.1. Let A, B be matrix, A, B are similar if there exists an invertible matrix P such that

$$A = P^{-1}BP$$

If A and B are similar, then they have the same eigen value

Proof.

$$\begin{aligned} A &= P^{-1}BP \\ PA &= BP \quad (1) \\ \text{Let } Ax &= \lambda x \\ \boldsymbol{P}Ax &= \lambda \boldsymbol{P}x \\ \boldsymbol{BP}x &= \lambda Px \quad \text{from } (2) \\ B\boldsymbol{Px} &= \lambda \boldsymbol{Px} \quad \blacksquare. \end{aligned}$$

Thus we have eigen vecgor Px with corresponding eigenvalue λ □

Definition 8.2. Matrix A is diagonalizable if A is similar to a diagonal matrix. Recall that a diagonal matrix is matrix where all x_{ij} are 0 except when $i = j$.

Theorem 8.3. *A is diagonalizable if and only if A has n linearly independent Eigen vector*

Proof (=>). Suppose there exists matrix D such that $D = P^{-1}AP$ then

$$\begin{aligned} D &= P^{-1}AP \\ PD &= AP \\ \begin{bmatrix} | & | & \cdots & | \\ P_1 & P_2 & \cdots & P_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} \delta_1 & & & 0 \\ & \delta_2 & & \\ & & \ddots & \\ 0 & & & \delta_n \end{bmatrix} &= A \begin{bmatrix} | & | & \cdots & | \\ P_1 & P_2 & \cdots & P_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix} \\ \begin{bmatrix} | & | & \cdots & | \\ \delta_1 P_1 & \delta_2 P_2 & \cdots & \delta_n P_n \\ | & | & \cdots & | \end{bmatrix} &= \begin{bmatrix} | & | & \cdots & | \\ AP_1 & AP_2 & \cdots & AP_n \\ | & | & \cdots & | \end{bmatrix} \end{aligned}$$

Since P is invertible then P has full rank and therefore P is linearly independent □

Proof (<=). Do the [=>] part in reverse order $\square$

Definition 8.4. A is orthogonally diagonalizable (OD) if there exists P which is orthogonal such that $P^{-1}AP$ is diagonal.

Claim 8.5. *If A is OD then $A = PDP^T$*

$$\begin{aligned} A &= PDP^T \\ &= \sum i = 1^n d_{ii} P_i P_i^T \end{aligned}$$

Theorem 8.6 (Spectral Theorem for Real Matrices)**.**
Let A be a real symmetric matrix, then

1. *The eigen value λ_i are reals and the eigen vectors are real*
2. *A is OD so*

$$\begin{aligned} A &= \begin{bmatrix} | & \dots & | \\ V_1 & \dots & V_n \\ | & \dots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \begin{bmatrix} - & V_1 & - \\ & \vdots & \\ - & V_n & - \end{bmatrix} \\ &= VDV^T \\ &= \sum_{i=1}^{n} \lambda_n V_i V_i^T \end{aligned}$$

Theorem 8.7 (The Fundamental Theorem of Symmetric Matrices)**.**
A real matrix A is OD $\iff$ A is symmetric.