A 50-Minute Tour of Linear Algebra

(A Supplement to ECE 403/403 Course Notes, Summer 2019)

1. Vectors

· ***Column and row vectors***

Column vectors assume the form

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_d \end{bmatrix}$$

Transpose of a column vector yields a row vector, for example,

$$\boldsymbol{u}^T = \begin{bmatrix} u_1 & u_2 & \cdots & u_d \end{bmatrix}$$

· ***Norm of vectors***

The L_p norm of a vector $\boldsymbol{u}$ is defined by

$$\|\boldsymbol{u}\|_p = \left(|u_1|^p + |u_2|^p + \cdots + |u_d|^p\right)^{1/p}$$

where $p \ge 1$ is an integer.

<u>Properties of L_p norm</u>

- $\|\boldsymbol{u}\|_p = 0$ if and only if $\boldsymbol{u} = \boldsymbol{0}$.
- $\|\alpha\boldsymbol{u}\|_p = |\alpha| \cdot \|\boldsymbol{u}\|_p$ for any scalar α.
- $\|\boldsymbol{u}+\boldsymbol{v}\|_p \le \|\boldsymbol{u}\|_p + \|\boldsymbol{v}\|_p$ (triangle inequality).

<u>Popular L_p norms</u>

L_2 norm: $\|\boldsymbol{u}\|_2 = \left(|u_1|^2 + |u_2|^2 + \cdots + |u_d|^2\right)^{1/2}$

L_1 norm: $\|\boldsymbol{u}\|_1 = |u_1| + |u_2| + \cdots + |u_d|$

L_∞ norm: $\|\boldsymbol{u}\|_\infty = \max\{|u_1|,\ |u_2|,\ \ldots, |u_d|\}$

- The unit vector along the direction of vector $\boldsymbol{u}$ is given by $\tilde{\boldsymbol{u}} = \dfrac{\boldsymbol{u}}{\|\boldsymbol{u}\|_p}$.
- Vector norms are useful for computing the "distance" between two vectors, say $\boldsymbol{u}$ and $\boldsymbol{v}$ (when they are regarded as "points"). For example,
  - ◊ the L_2 (Euclidean) distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is given by

$$\|\boldsymbol{u}-\boldsymbol{v}\|_2 = \left\| \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_d \end{bmatrix} - \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} u_1 - v_1 \\ u_2 - v_2 \\ \vdots \\ u_d - v_d \end{bmatrix} \right\|_2 = \left((u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_d - v_d)^2\right)^{1/2}$$

◊ the L_1 (Manhattan) distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is given by

$$\|\boldsymbol{u}-\boldsymbol{v}\|_1 = \left\| \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_d \end{bmatrix} - \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{bmatrix} \right\|_1 = \left\| \begin{bmatrix} u_1 - v_1 \\ u_2 - v_2 \\ \vdots \\ u_d - v_d \end{bmatrix} \right\|_1 = |u_1 - v_1| + |u_2 - v_2| + \cdots + |u_d - v_d|$$

· ***Inner product of two vectors***

Let

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_d \end{bmatrix} \text{ and } \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{bmatrix},$$

their *inner product* is defined by

$$\boldsymbol{u}^T \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \cdots + u_d v_d$$

Note that

$$\boldsymbol{u}^T \boldsymbol{v} = \boldsymbol{v}^T \boldsymbol{u} \text{ and } \boldsymbol{u}^T \boldsymbol{u} = \|\boldsymbol{u}\|_2^2 \tag{1}$$

Inner product has a geometrical interpretation:

$$\boldsymbol{u}^T \boldsymbol{v} = \|\boldsymbol{u}\|_2 \cdot \|\boldsymbol{v}\|_2 \cdot \cos\theta \tag{2}$$

where θ is the angle between vectors $\boldsymbol{u}$ and $\boldsymbol{v}$. It follows that

$$\cos\theta = \frac{\boldsymbol{u}^T \boldsymbol{v}}{\|\boldsymbol{u}\|_2 \cdot \|\boldsymbol{v}\|_2} \tag{3}$$

<u>Example</u>: Let $\boldsymbol{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, then $\cos\theta = \frac{\boldsymbol{u}^T \boldsymbol{v}}{\|\boldsymbol{u}\|_2 \cdot \|\boldsymbol{v}\|_2} = \frac{8}{\sqrt{10} \cdot \sqrt{8}} = \frac{2}{\sqrt{5}} = \sqrt{0.8} \approx 0.8944$, hence the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ is $\theta \approx \cos^{-1}(0.8944) = 26.57^\circ$.

· ***Orthogonal Projection of a vector onto another vector***

The projection of a vector $\boldsymbol{v}$ onto another vector $\boldsymbol{u}$ is defined as a vector who has the same direction as vector $\boldsymbol{u}$ and hence assumes the form $\alpha\tilde{\boldsymbol{u}}$ where $\tilde{\boldsymbol{u}}$ is the unit vector along $\boldsymbol{u}$ and α is a scalar given by

$$\alpha = \|\boldsymbol{v}\|_2 \cos\theta = \|\boldsymbol{v}\|_2 \cdot \frac{\boldsymbol{u}^T \boldsymbol{v}}{\|\boldsymbol{u}\|_2 \cdot \|\boldsymbol{v}\|_2} = \frac{\boldsymbol{u}^T \boldsymbol{v}}{\|\boldsymbol{u}\|_2} = \tilde{\boldsymbol{u}}^T \boldsymbol{v}$$

Hence the orthogonal projection of vector $\boldsymbol{v}$ onto $\boldsymbol{u}$ is given by

$$\text{orthogonal projection of vector } \boldsymbol{v} \text{ onto vector } \boldsymbol{u} = (\boldsymbol{v}^T \tilde{\boldsymbol{u}})\tilde{\boldsymbol{u}} \tag{4}$$

2. Matrices

A matrix $\boldsymbol{A} \in R^{d \times m}$ has d rows and m columns and hence, if necessary, can be expressed explicitly in terms its columns as

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m \end{bmatrix} \text{ where each } \boldsymbol{a}_i \in R^{d \times 1}.$$

・The transpose of a matrix $\boldsymbol{A} \in R^{d \times m}$, denoted by $\boldsymbol{A}^T$, is a matrix of size $m \times d$, whose ith row is the transpose of $\boldsymbol{A}$'s ith column. That is,

$$\boldsymbol{A}^T = \begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \\ \vdots \\ \boldsymbol{a}_m^T \end{bmatrix}$$

・***Multiplying two matrices***

Given two matrices $\boldsymbol{A} \in R^{d \times m}$ and $\boldsymbol{B} \in R^{m \times k}$ with consistent dimensions (so that $\boldsymbol{A} \cdot \boldsymbol{B}$ can be performed). Multiplication $\boldsymbol{A} \cdot \boldsymbol{B}$ can be carried out in many ways.

For example, if we denote

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{r}_1^T \\ \boldsymbol{r}_2^T \\ \vdots \\ \boldsymbol{r}_d^T \end{bmatrix} \text{ and } \boldsymbol{B} = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_k \end{bmatrix}$$

then

$$\boldsymbol{A} \cdot \boldsymbol{B} = \begin{bmatrix} \boldsymbol{r}_1^T \boldsymbol{b}_1 & \boldsymbol{r}_1^T \boldsymbol{b}_2 & \cdots & \boldsymbol{r}_1^T \boldsymbol{b}_k \\ \boldsymbol{r}_2^T \boldsymbol{b}_1 & \boldsymbol{r}_2^T \boldsymbol{b}_2 & \cdots & \boldsymbol{r}_2^T \boldsymbol{b}_k \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{r}_d^T \boldsymbol{b}_1 & \boldsymbol{r}_d^T \boldsymbol{b}_2 & \cdots & \boldsymbol{r}_d^T \boldsymbol{b}_k \end{bmatrix}_{d \times k}$$

In addition, we can also write

$$\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{A} \cdot \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} \cdot \boldsymbol{b}_1 & \boldsymbol{A} \cdot \boldsymbol{b}_2 & \cdots & \boldsymbol{A} \cdot \boldsymbol{b}_k \end{bmatrix}$$

or

$$\boldsymbol{A} \cdot \boldsymbol{B} = \begin{bmatrix} \boldsymbol{r}_1^T \\ \boldsymbol{r}_2^T \\ \vdots \\ \boldsymbol{r}_d^T \end{bmatrix} \cdot \boldsymbol{B} = \begin{bmatrix} \boldsymbol{r}_1^T \boldsymbol{B} \\ \boldsymbol{r}_2^T \boldsymbol{B} \\ \vdots \\ \boldsymbol{r}_d^T \boldsymbol{B} \end{bmatrix}$$

<u>Example</u>: Let $\boldsymbol{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$. Then both

$$\boldsymbol{A}\boldsymbol{A}^T = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix} \text{ and } \boldsymbol{A}^T \boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

· ***Symmetric matrices***

· A matrix $\boldsymbol{P}$ is said to be symmetric if $\boldsymbol{P}^T = \boldsymbol{P}$.

· Matrices $\boldsymbol{A}\boldsymbol{A}^T$ and $\boldsymbol{A}^T\boldsymbol{A}$ (for any $\boldsymbol{A}$) are always symmetric (see the example above).

· ***Eigenvalues and eigenvectors of a symmetric matrix***

Let $\boldsymbol{P}$ be a symmetric matrix of size $d \times d$. A scalar λ and a nonzero vector $\boldsymbol{u}$ are said to be an *eigenvalue* and associated *eigenvector* of $\boldsymbol{P}$, respectively, if

$$\boldsymbol{P}\boldsymbol{u} = \lambda \boldsymbol{u} \tag{5}$$

Note that (5) implies that the equation $(\lambda \boldsymbol{I} - \boldsymbol{P})\boldsymbol{u} = \boldsymbol{0}$ has nonzero solution, therefore the determinant of $\lambda \boldsymbol{I} - \boldsymbol{P}$ must be equal to zero if λ is an eigenvalue of $\boldsymbol{P}$. This in turn suggests that the eigenvalues of matrix $\boldsymbol{P}$ can be obtained by solving the algebraic equation

$$\det(\lambda \boldsymbol{I} - \boldsymbol{P}) = 0 \text{ for } \lambda \tag{6}$$

<u>Properties</u>

· All eigenvalues of a symmetric matrix are real-valued.

· For a symmetric $\boldsymbol{P}$ of size $d \times d$, there exist d orthonormal eigenvectors. That is, if we denote them by $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_d$, then

$$\boldsymbol{u}_i^T \boldsymbol{u}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Consequently, if we use these eigenvectors to form a matrix $\boldsymbol{U} = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_d]$, then we have

$$\boldsymbol{U}^T\boldsymbol{U} = \underbrace{\begin{bmatrix} \boldsymbol{u}_1^T \\ \boldsymbol{u}_2^T \\ \vdots \\ \boldsymbol{u}_d^T \end{bmatrix}}_{\boldsymbol{U}^T} \cdot \underbrace{[\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_d]}_{\boldsymbol{U}} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \boldsymbol{I} \tag{7}$$

A matrix $\boldsymbol{U}$ satisfying $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^T = \boldsymbol{I}$ is called an orthogonal matrix.

Moreover, if $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_d$ are orthonormal eigenvalues of matrix $\boldsymbol{P}$, then by definition we have

$$\boldsymbol{P} \cdot \underbrace{[\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_d]}_{\boldsymbol{U}} = [\boldsymbol{P}\boldsymbol{u}_1 \quad \boldsymbol{P}\boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{P}\boldsymbol{u}_d] = [\lambda_1\boldsymbol{u}_1 \quad \lambda_2\boldsymbol{u}_2 \quad \cdots \quad \lambda_d\boldsymbol{u}_d] = \underbrace{[\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_d]}_{\boldsymbol{U}} \underbrace{\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_d \end{bmatrix}}_{\Lambda}$$

That is, $\boldsymbol{P}\boldsymbol{U} = \boldsymbol{U}\Lambda$ which in conjunction with (7) implies that

$$\boldsymbol{P} = \boldsymbol{U}\Lambda\boldsymbol{U}^T \tag{8}$$

In the literature, (8) is called *eigen-decomposition* of symmetric matrix $\boldsymbol{P}$.

<u>Example</u> Let $\boldsymbol{P} = \begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix}$. To find its eigenvalues, we solve the characteristic equation

$$\det(\lambda \boldsymbol{I} - \boldsymbol{P}) = \det \begin{bmatrix} \lambda - 6 & -1 \\ -1 & \lambda - 2 \end{bmatrix} = \lambda^2 - 8\lambda + 11 = 0$$

to obtain

$$\lambda_{1,2} = 4 \pm \sqrt{5} \approx 6.2361, 1.7639$$

Eigenvectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ that are associated with λ_1 and λ_2 can be found by computing nonzero solutions of the linear equations

$$(\lambda_i \boldsymbol{I} - \boldsymbol{P})\boldsymbol{u} = \boldsymbol{0} \quad \text{for } i = 1, 2$$

・Note: There is a MATLAB function, **`eigs`**, for eigen-decomposition of a symmetric matrix $\boldsymbol{P}$:

```
[U,L] = eigs(P,q);
```

which returns q eigenvectors (as columns of **`U`**) that are associated with the q largest eigenvalues (along the diagonal of matrix **`L`** in descent order). For the 2 × 2 matrix **`P`** given above, **`eigs(P,2)`** returns with

$$L = \begin{bmatrix} 6.2361 & 0 \\ 0 & 1.7639 \end{bmatrix}$$

and **`U = [u1 u2]`** where

$$\boldsymbol{u}_1 = \begin{bmatrix} -0.9732 \\ -0.2298 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} 0.2298 \\ -0.9732 \end{bmatrix}$$

Obviously **`U`** is an orthogonal matrix.