# 3.11 Eigenvalues and Eigenvectors, Spectral Representation

## 3.11.1 Eigenvalues and Eigenvectors

A vector $\varphi$ is **eigenvector** of a matrix $K$, if $K\varphi$ is parallel to $\varphi$ and $\varphi \neq 0$, i.e.,

$$K\varphi = k\varphi$$

$k$ is the eigenvalue.

If $\psi$ is eigenvector of ${}^tK$, then its components satisfy

$$\sum_i \psi_i K_{ij} = k\psi j$$

or

$$\psi K = k\psi$$

and $\psi$ is called **left eigenvector** of $K$; $\varphi$ is called **right eigenvector** of $K$.

The equation for $\varphi$ can be written as

$$(K - k\mathbf{1})\varphi = 0$$

This has a non-zero solution $\varphi$ if and only if

$$det\ (K - k\mathbf{1}) = 0$$

For an $n \times n$ matrix, the determinant is a polynomial of degree $n$ in $k$ with at most $n$ district roots. For every root 1 eigenvector. For a repeated root, there may be an many linearly indepedent eigenvectors as the multiplicity of the root, but in general there may be no more than one.

For further discussion, assume matrices of the form that $det\ (K - k\mathbf{1}) = 0$ with $n$ distinct simple roots

$k_i (i = 1, \cdots, n)$ with eigenvectors $\varphi_i$.

$\varphi_i$ linearly independent and span space $K$ $\rightarrow$ basis vectors.

Further:

$$det\ (K - k\mathbf{1}) = det\ ({}^tK - k\mathbf{1})$$

$\Rightarrow$ for each $k_i$, there is a left eigenvector $\psi_i$ and a right eigenvector $\varphi_i$ such that

$$K\varphi_i = k_i\varphi_i \quad and \quad \psi_i K = k_i\psi_i$$

For $k_i \neq k_j$, $\psi_i \perp \varphi_j$ and $\psi_j \perp \varphi_i$:

$$k_j(\psi_i, \varphi_j) = (\psi_i, K\varphi_j) = (\psi_i\ {}^tK, \varphi_j) = k_i(\psi_i, \varphi_j)$$

Since $\psi_i$ perpendicular to all vectors $\varphi_j$ with $j \neq i$ $\rightarrow$ $\psi_i$ cannot be perpendicular to $\varphi_i$, then it would be zero.

$\Rightarrow$ $(\psi_i, \varphi_i) \neq 0$ $\rightarrow$ normalized such that

$$(\psi_i, \varphi_i) = 1$$

$\rightarrow$ $(\psi_i, \varphi_j) = \delta_{ij}$

$\rightarrow$ Sets $\{\psi_i\}$ and $\{\varphi_i\}$ are called **reciprocal**.

### 3.11.2 Spectral Representation

Use set $\{\varphi_i\}$ as basis $\Rightarrow$

Each vector $f$ can be written as:

$$f = \sum_i f_i \varphi_i$$

where $f_i$ are the **components** of $f$ with respect to basis $\varphi_i$. Components are inner product of $f$ with left eigenvectors:

$$(\psi_j, f) = \sum_i f_i(\psi_j, \varphi_i) = \sum_k f_i \delta_{ij} = f_j$$

$$\rightarrow \quad \mathbf{f} = \sum_{\mathbf{i}} \varphi_{\mathbf{i}}(\psi_{\mathbf{i}}, \mathbf{f})$$

Use left eigenvectors as basis (reciprocal basis) $\rightarrow$

$$f = \sum_i (f, \varphi_i)\psi_i$$

**Notation**: $(ab)_{ij} \triangleq a_i b_j$

Then identity can be represented as

$$\mathbf{1} = \sum_i \varphi_i \psi_i$$

Apply $K$ on vector $f$:

$$Kf = \sum_i K\varphi_i(\psi_i, f) = \sum_i k_i\varphi_i(\psi_i, f)$$

With respect to basis $\{\varphi_i\}, K$ acts like a diagonal matrix. Since above equation is valid for arbitrary $f \;\Rightarrow$

$$\mathbf{K} = \sum_{\mathbf{i}} \mathbf{k_i}\boldsymbol{\varphi}_\mathbf{i}\boldsymbol{\psi}_\mathbf{i}$$

This is the **spectral representation** of matrix $K$.

The set of eigenvalues $k_i$ is called the **spectrum** of $K$.

$\rightarrow$ For any positive $n$:

$$K^n = \sum_i k_i^n \; \varphi_i\psi_i$$

and interpret $K^0 \triangleq \mathbf{1}$

For the inverse: $K^{-1} = \sum_i k_i^{-1}\varphi_i\psi_i$.

If any eigenvalue $k_i = 0$, then $K$ does not have an inverse.

## 3.12 Singular Value Decomposition (SVD)

**Theorem**: Let $A$ be an arbitrary (complex) $m \times n$ matrix

$$A \in M(m \times n, \mathbf{C})$$

1. There exists a unitary matrix $U \in M(m \times m, \mathbf{C})(U^+ = U^{-1})$ and a unitary matrix $V \in M(n \times n, \mathbf{C})(V^+ = V^{-1})$ such that $U^+ \, AV = \Sigma$ is a $m \times n$ "diagonal" matrix of the following form:

$$\Sigma \; = \; \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \;\; D \; := \; diag \; (\sigma_1, \; \cdots, \sigma_r); \; \sigma_1 \geq \sigma \geq \; \cdots \; \geq \sigma_r > 0$$

   where $\sigma_1, \; \cdots, \sigma_r$ are the nonvanishing singular values of $A$ and $r = rank A$.

2. The nonvanishing singular values of $A^+$ are also given by $\sigma_1, \; \cdots, \sigma_r$.

   The decomposition

$$\mathbf{A} = \mathbf{U} \; \mathbf{\Sigma} \; \mathbf{V}^+$$

   is called **singular value decomposition**.

**Preliminaries**:

1. We know that for every Hermitian matrix $A \in M(n \times n, \mathbf{C})$ there is a unitary matrix $U \in M(n \times n, \mathbf{C})$ with

$$U^{-1} A U = \begin{pmatrix} \lambda_1 & & 0 \\ & \cdot & \\ & \cdot & \\ 0 & & \lambda_n \end{pmatrix}, \quad \lambda_i \in \mathbf{R}$$

   and a Hermitian matrix $A$ is positive (positive semidefinite) if and only if the eigenvalues of $A$ are positive (non-negative).

2. An arbitrary matrix $A \in M(n \times n, \mathbf{C})$ is called **normal** if there exists a unitary matrix $U \in M(n \times n, \mathbf{C})$ such that

$$U^{-1}AU = \begin{pmatrix} \lambda_1 & & 0 \\ & \cdot & \\ & \cdot & \\ 0 & & \lambda_n \end{pmatrix}$$

So, Hermitian matrices are normal.

What can be done if $A \in M(m \times n, \mathbf{C})$?

Then

$AA^+ \ \in M(m \times m, \mathbf{C})$ to Hermitian.

**Proof**: $(AA^+)^+ = AA^+$

and $A^+A \in M(n \times n, \mathbf{C})$ is Hermitian

$$(A^+A)^+ = A^+A$$

and $A^+A$ is positive semidefinite via construction.

$\rightarrow$ Eigenvalues $\lambda_i \leq 0$ can be written as $\lambda_i = \sigma_i^2$. The numbers $\sigma_i$ are **called singular values of** $A$.

Proof of the theorem by induction on $m$ and $n$:

1. $m = 0, n = 0 \ \rightarrow$ nothing to prove.

2. Assume that theorem holds for matrices $\tilde{A}, \tilde{\Sigma} \in M((m-1) \times (n-1), \mathbf{C})$, i.e., there exists unitary matrix $\tilde{U} \in M((m-1) \times (m-1), \mathbf{C})$; unitary matrix $\tilde{V} \in M((n-1) \times (n-1), \mathbf{C})$ such that

$$\tilde{U}^+\tilde{A}\tilde{V} = \ \tilde{\Sigma} \ = \begin{pmatrix} \tilde{D} & 0 \\ 0 & 0 \end{pmatrix}$$

with $\tilde{D} = diag\ (\sigma_2, \ \cdots, \sigma_r)$ and $\sigma_2^2 \geq \sigma_3^2 \geq \ \cdots \sigma_r^2 \geq 0$.

3. Show under assumption (2) that theorem holds for $A \in M(m \times n, \mathbf{C})$.

Let largest singular value of $A$ be $\sigma_1 > 0$ (for $\sigma_1 = 0 \;\rightarrow\; A = 0$, since $\sigma_1$ should be largest).

Let $x_1 \neq 0$ be eigenvector of $A^+A$ with eigenvalue $\sigma_1^2$, i.e., $A^+Ax_1 = \sigma_1^2 x_1$ and $\| x_1 \| = 1$.

Now find additional $n-1$ vectors $x_2, x_3, \cdots, x_n \in \mathbf{C}^n$ such that matrix

$X := (x_1, x_2, \cdots, x_n) \in M(n \times n, \mathbf{C})$ is unitary; i.e., $X^+X = \mathbf{1}$.

$X_1$ is the column $X_1 = \begin{pmatrix} X_1^1 \\ X_1^2 \\ \cdot \\ \cdot \\ x_1^n \end{pmatrix}$

$$\| Ax_1 \|^2 = \langle Ax_1 \mid Ax_1 \rangle = x_1^+ A^+ A x_1 = x_1^+ \sigma_1^2 x_1 = \sigma_1^2 \;\; \| x_1 \| = \sigma_1 > 0$$

Define $y_1 = \frac{1}{\sigma_1} Ax_1 \in \mathbf{C}^m$, with $\| y_1 \|^2 = \frac{\|Ax_1\|}{\sigma_1^2} = \mathbf{1}$.

$y_1$ are well defined.

Find additional $(m-1)$ vectors $y_1, \cdots, y_m$ (orthogonal) such that

$$Y := (y_1, y_2, \cdots y_m) \in M(m \times m, \mathbf{C})$$

is unitary, i.e., $Y^+Y = \mathbf{1}$.

Look at components of this matrix equation:

$$^t y_1 y_1 = \delta_{\mathbf{1}}, \quad in\ general \sum_{\ell} y_{ie}^+ y_{ej} = \sum_{\ell} y_{ei}^* y_{ej} = \delta_{ij}$$

Take unit vectors $e_1^n \in \mathbf{C}^n$ and $e_1^m \in \mathbf{C}^m$ and take 'matrix elements' of $Y^+AX$

$$\begin{aligned} Y^+AXe_1^n &= Y^+A_1x_1 = \sigma_1 Y^+ y_1 = \sigma_1 e_1^m \in \mathbf{C}^m \qquad (3.93) \\ (Y^+AX)^+ e_1^m &= X^+A^+Ye_1^m = X^+A^+y_1 = \frac{1}{\sigma_1} X^+A^+Ax_1 \\ &= \sigma_1 X^+ x_1 = \sigma_1 e_1^n \in \mathbf{C}^n \\ \rightarrow \;\; \langle e_1^m \mid Y^+AX \mid e_1^n \rangle &= \sigma_1 \end{aligned}$$

or

$$Y^+AX = \begin{pmatrix} \sigma_1 & \cdots & 0 \\ \cdot & & \\ \cdot & \tilde{A} & \\ 0 & & \end{pmatrix}$$

with $\tilde{A} \in M((m-1) \times (n-1), \mathbf{C})$.

According to induction hypothesis, there are unitary matrices $\tilde{U}, \tilde{V}$ such that $\tilde{A}$ can be written as $\tilde{U}^+\tilde{A}\tilde{V} = \tilde{\Sigma} = \begin{pmatrix} \tilde{D} & 0 \\ 0 & 0 \end{pmatrix}$ with $\tilde{D} = diag(\sigma_2, \cdots, \sigma_r)$ and $\sigma_2^2 \geq \sigma_3^2 \geq \cdots \geq \sigma_r^2 \geq 0$.

Now define

$$\begin{aligned} U &:= Y \cdot \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U} \end{pmatrix} \in M(m \times m, \mathbf{C}) \\ V &:= X \cdot \begin{pmatrix} 1 & 0 \\ 0 & \tilde{V} \end{pmatrix} \in M(n \times n, \mathbf{C}) \end{aligned} \tag{3.94}$$

$$\begin{aligned} \Rightarrow \quad U^+AV &= \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U}^+ \end{pmatrix} Y^+AX \begin{pmatrix} 1 & 0 \\ 0 & \tilde{V} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U}^+ \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \tilde{A} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{V} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U}^+ \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \tilde{A}\tilde{V} \end{pmatrix} \\ &= \begin{pmatrix} \sigma_1 & 0 \\ 0 & \tilde{U}^+\tilde{A}\tilde{V} \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \tilde{\Sigma} \end{pmatrix} = \begin{pmatrix} \sigma_1 & & 0 \\ & \tilde{D} & \\ 0 & & 0 \end{pmatrix} \\ &:= \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} = \Sigma \end{aligned} \tag{3.95}$$

with $D = diag(\sigma_1, \sigma_2, \cdots, \sigma_r)$, $\Sigma \in M(m \times n, \mathbf{C})$ and $\sigma_2^2 \geq \sigma_3^2 \geq \cdots \geq \sigma_r^2 \geq 0$.

$\sigma_1^2 = \lambda_{max}(A^+A)$ according to choice at beginning.

$$rank\ A = rank\ (U^+AV) = rank\ \Sigma\ = r$$

Still to prove that $\sigma_1^2 \geq \sigma_2^2$ and that $\sigma_i$ are the singular values of $A$. Consider

$$\Sigma\ ^+\ \Sigma\ = V^+A^+UU^+AV = V^+A^+AV = diag(\sigma_1^2, \sigma_2^2,\ \cdots, \sigma_r^2)$$

Since $V$ is the unitary matrix diagonalizing $A^+A$ and $A^+A$ Hermitian, thus the unitary matrix $V$ exists.

$\rightarrow\ \ \sigma_1^2, \sigma_2^2,\ \cdots, \sigma_r^2$ are eigenvalues of $A^+A\ \Rightarrow$

$\sigma_1^2, \sigma_2^2,\ \cdots, \sigma_r$ are the nonvanishing singular values of $A$.

Because

$$\sigma_1^2 = \lambda_{max}(A^+A)\ \Rightarrow\ \sigma_1^2 \geq \sigma_2^2$$

Shown that there exists a decomposition

$$U^+AV =\ \Sigma \quad or \quad A = U\Sigma\ V^+$$

with $U^{-1} = U^+,\ V^{-1} = V^+$ and

$$U \in M(m \times m, \mathbf{C});\quad V \in M(n \times n, \mathbf{C});\quad A, \Sigma\ \in r(m \times n, \mathbf{C})$$

Columns of $U$ represent $m$ orthogonal eigenvectors of $AA^+ \in M(m \times m, \mathbf{C})$.

Columns of $V$ represent $n$ orthogonal eigenvectors of $A^+A \in M(n \times n, \mathbf{C})$.

Easy to see:

$U^+AA^+U = U^+AVV^+A^+U = \Sigma \sum^+$,

i.e., $U$ diagonalizes $AA^+$ with eigenvalues $\sigma_1^2, \cdots, \sigma_r^2$

$V^+A^+AV = V^+A^+UU^+AV = \Sigma^+ \sum$,

i.e., $V$ diagonalizes $A^+A$ with eigenvalues $\sigma_1^2, \cdots, \sigma_r^2$.

**Side Remark**:

**Definition**: The **Pseudoinverse** (or Moore-Penrose inverse) of an arbitrary matrix $A \in M(m \times n, \mathbf{C})$ is a matrix $A_+ \in M(n \times m, \mathbf{C})$ with

1. $A_+A = P$, where $P$ is the orthogonal projector

$$P : \mathbf{C}^n \to N(A);\ N(A) := \{x \in \mathbf{C}^n; Ax = 0\}$$

   $AA_+ = \bar{P}$, where $\bar{P}$ is the projector

$$\bar{P} : \mathbf{C}^m \to \mathbf{R}(A);\ \mathbf{R}(A) := \{Ax \in \mathbf{C}^m, x \in \mathbf{C}^n\}$$

2. (a) $A_+A = (A_+A)_+$
   (b) $AA_+ = (AA_+)_+$
   (c) $AA_+A = A$
   (d) $A_+A\ A_+ = A_+$

Build the 'pseudoinverse' $A_+$ of matrix $A \in M(m \times n, \mathbf{C})$:

We have $U^+AV = \Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$, $D = diag(\sigma_1, \cdots, \sigma_r)$

$\to \Sigma_+ \in M(n \times m, \mathbf{C})$ easy to obtain:

$$\Sigma_+ = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

for $A = U\ \Sigma\ V$ define $A_+ \ := \ V\ \Sigma_+\ U^+$

and in principle the properties of a pseudoinverse have to be checked.

$\rightarrow$ for $m = n$ (square matrices)

$A = U\ \Sigma\ V^+$ can be inverted to give

$$A^{-1} = V \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^+ \quad with \quad D^{-1} = diag \left( \frac{1}{\sigma_1}\ ,\ \cdots\ ,\ \frac{1}{\sigma_r} \right)$$

**Application of SVD**:

Solve system of homogeneous or inhomogeneous linear equations:

$$Ax = b \quad \rightarrow \quad x = A^{-1}b$$

Since one can calculate $A^{-1}$, even if $A$ singular or ill-determined $\rightarrow$ good way to solve $x = A^{-1}b$. Considerations about solution space are easy:

Look at $A \cdot x = 0 \quad = \quad (U\ \Sigma\ V^+)x = 0$.

Solution are in Ker $A$ : then any column of $V$ which corresponds to $\sigma_i = 0$ are $\in$ Ker $A$.

For $A \cdot x = b$, has only solution if $b \in Im(A)$.

If $b \in Im(A)$, then one still can construct a "solution" vector, which will not solve $A{\cdot}x = b$, but be closest possible solution in a least square sense, i.e., one finds $X$ which minimizes

$$\begin{aligned} r \quad &:= \quad \mid A \cdot x - b \mid \qquad (3.96) \\ r \quad &\triangleq \quad residual\ solution \end{aligned}$$

(See Numerical Recipes, p. 54.)

**Construct Orthonormal Basis**:

$A \in M(m \times n, \mathbf{C})$ or $M(m \times n, \mathbf{R})$

Matrix $U \in M(m \times m, \mathbf{C}(\mathbf{R}))$ represent the orthogonal eigenvectors of $AA^+ \rightarrow$

Columns of $U$ are desired orthonormal system of eigenvectors.

$\rightarrow$ If some of the $\sigma_i^2 = 0$, then the space spanned by the column vectors of $U$ has dimension $< m$.

**Approximation of Matrices**:

$A = U \, \Sigma \, V^+$

with elements

$$A_{ij} = \sum_{km} U_{ik} \sum_{km} \delta_{km} V_{jm} \tag{3.97}$$

$$= \sum_{k=1}^{m} \sigma_k \, U_{ik} V_{jk} \tag{3.98}$$

If only $\sigma_r$ with $r < m$ are $\neq 0$, then matrix $A$ can be approximated by 'smaller' matrix.

Then consider

$$A \cdot X = \sum_j A_{ij} x_j = \sum_j \sum_k \sigma_k u_{ik} v_{ik} \cdot x_j \tag{3.99}$$

$$= \sum_k \sigma_k u_{ik} \sum_j v_{jk} x_j \tag{3.100}$$

If $k \leq m$, then the evaluation of $A \cdot x$ requires $k(m+n)$ multiplications which is in that sense smaller than $m \cdot n$ multiplications needed to evaluate $\sum_j A_{ij} x_j$ for all $i$.