Linear Algebra: Some Useful Existing Results

Soumyabrata Talukder

July 2020

1 Eigen-values and Left/Right Eigen-vectors

For a matrix $A \in \mathbb{R}^{n \times n}$, $v \in \mathbb{C}^n$ is called its eigen-vector if the product $A.v$ evaluates to be a scaled version of v, say $\lambda.v$, where $\lambda \in \mathbb{C}$ is called the corresponding eigen-value.

For the purpose of understanding modal decomposition, we will consider A to have distinct eigen-values. This is realistic in the sense that the state matrix of a linearized practical system, represented by A, is unlikely to have two eigen-values exactly equal (however, complex conjugate pairs of eigen-values are common). In other words, we are assuming the matrix A to be of full rank, having n distinct eigen-values denoted by $\lambda_1, \lambda_2, \dots, \lambda_n$, and the corresponding linearly independent eigen-vectors are denoted by $v_1, v_2, \dots, v_n$. Note that the eigen-values of A and A^T are identical, and let $u_1, u_2, \dots, u_n$ denote the three linearly independent eigen-vectors of A^T. So, we have:

$$A.v_i = \lambda_i.v_i, \forall i \in \{1, \dots, n\}, \tag{1}$$

$$A^T.u_i = \lambda_i.u_i, \forall i \in \{1, \dots, n\}, \tag{2}$$

which can be written more compactly as:

$$A.V = V.\Lambda \Rightarrow A = V\Lambda.V^{-1}, \tag{3}$$

$$A^T.U' = U'.\Lambda \Rightarrow A = [U'^T]^{-1}\Lambda.U'^T, \tag{4}$$

where Λ is a diagonal matrix with the diagonal elements being $\lambda_1, \lambda_2, \ldots, \lambda_n$ in order, $V := [v_1, v_2, \ldots, v_n]$ and $U' := [c_1.u'_1, c_2.u'_2, \ldots, c_n u'_n]$ with c_is being arbitrary scalars. (3) is called the *eigen-decomposition* of the matrix A. From (2), we see that $u_i'^T.A = u_i'^T.\lambda_i$ holds for all $\{1, \ldots, n\}$, which justifies u'_is being termed as the *left* eigen-vector of A, in contrast to v_is, called as the *right* eigen-vectors. Comparing (3) and (4), we also see that by choosing appropriate $c_i \in \mathbb{R}$ for all $i \in \{1, \ldots, n\}$, one can establish the following relation:

$$V^{-1} = U^T, \tag{5}$$

where $U := [u_1 u_2 \ldots u_n]$ with $u_i := c_i.u'_i$ for all $\{1, \ldots, n\}$. Next, from (3) and (5) we have the following:

$$\begin{aligned}
A &= V.\Lambda.U^T \\
&= \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix} \\
&= \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 u_1^T \\ \lambda_2 u_2^T \\ \vdots \\ \lambda_n u_n^T \end{bmatrix} \\
&= \sum_{i \in \{1, \ldots, n\}} \lambda_i v_i u_i^T
\end{aligned} \tag{6}$$

2 Special Case of Symmetric Matrices

Consider the special case when A is a symmetric matrix, i.e. $A^T = A$. If we perform eigen-decomposition of A and A^T, clearly it follows from (1)-(2) that $U = V$. This along with the relation in (5) yields $V^{-1} = U^T = V^T$, implying that V (and hence, U) is a unitary matrix and that $A = V.\Lambda.V^T$. Now, we have:

$$\begin{aligned}
A^T.A &= [V.\Lambda.V^T]^T.[V.\Lambda.V^T] \\
&= V.\Lambda.V^T.V.\Lambda.V^T \\
&= V.\Lambda^2.V^T
\end{aligned} \tag{7}$$

Note that (7) is the eigen-decomposition of $A^T.A$ since Λ^2 is diagonal. Also, (7) is the *singular value decomposition* (SVD) of A since V is unitary. Hence, if A is symmetric, its SVD and the eigen-decomposition of $A^T.A$ are identical.

SVD is not covered in this write up but it is useful to explore and understand. It finds various applications and (i) principal component analysis (PCA) of multi-variate data, (ii) low-rank estimation of a matrix $A \in \mathbb{R}^{m\times n}$, (iii) pseudo-inverse of a matrix $A \in \mathbb{R}^{m\times n}$ are possibly the most significant ones among those.

3 Maximum and Minimum Eigen-values

The maximum and the minimum eigen-values (denoted by λ_{max} and λ_{min} resp.) of a real symmetric matrix A are shown to be expressed as follows:

$$\begin{aligned} \lambda_{max} &= \sup_{x\neq 0} \frac{x^T A x}{x^T x} \\ \lambda_{min} &= \inf_{x\neq 0} \frac{x^T A x}{x^T x} \end{aligned} \tag{8}$$

The above result can be used to show a few necessary conditions of positive-semidefiniteness of A (denoted by $A \succeq 0$). From (8), the following can be deduced:

$$x^T x.\lambda_{max} \geq x^T A x \geq x^T x.\lambda_{min} \forall x \neq 0. \tag{9}$$

If $A \succeq 0$, we have $x^T A x \geq 0$ for all $x \neq 0$. Also noting that $x^T x > 0$ for all $x \neq 0$, it is concluded from (9) that (i) $\lambda_{max} \geq \lambda_{min} \geq 0$ and hence, (ii) $det(A) = \prod_{i\in\{1,\ldots,n\}} \lambda_i \geq 0$, where $det(A)$ denotes the determinant of A. These are the two necessary conditions of $A \succeq 0$. In this context, the only sufficient condition of $A \succeq 0$ is the existence of a $B \in \mathbb{R}^{n\times n}$ such that $A = B^T.B$.

4 Real Vector-space Decomposition Induced by a Matrix

For a matrix $A \in \mathbb{R}^{m \times n}$, the range-space $\mathcal{R}(A)$ is defined as:

$$\mathcal{R}(A) := \{A.x | x \in \mathbb{R}^n\}, \tag{10}$$

which is a subspace of $\mathbb{R}^m$ of dimension equal to the rank of A ($\leq$ minimum of m and n). The nullspace $\mathcal{N}(A)$ is defined as:

$$\mathcal{N}(A) := \{x \in \mathbb{R}^n | A.x = 0\}, \tag{11}$$

which is a subspace of $\mathbb{R}^n$ of dimension equal to the minimum of m and n less the rank of A.

For a set $\mathcal{V} \subseteq \mathbb{R}^n$, its vertical complement is defined as:

$$\mathcal{V}^{\perp} := \{z \in \mathbb{R}^n | z^T.x = 0, x \in \mathcal{V}\}. \tag{12}$$

By basic algebra it can be shown that $\mathcal{R}(A^T)$ is the verical complement of $\mathcal{N}(A)$, and consequently we have:

$$\mathbb{R}^n = \mathcal{N}(A) \oplus \mathcal{R}(A^T), \tag{13}$$

which is called the decomposition of $\mathbb{R}^n$ induced by A.

5 Modal Decomposition of Linearized Systems

We discuss modal decomposition of transient response of linearized autonomous systems of the form $\dot{x} = A.x$, where $x \in \mathbb{R}^n$ denotes the dynamic state vector. Considering x_0 as the initial state, the transient response of the system is known to be as follows:

$$x(t) = e^{At}.x_0, \tag{14}$$

where $t \in \mathbb{R}_{\geq 0}$ denotes the progressing time.

Note that using (6), the system state equation can be rewritten as:

$$\begin{aligned} &\dot{x} = V.\Lambda.U^T.x, \\ \Rightarrow &V^{-1}\dot{x} = \Lambda.U^T.x, \\ \Rightarrow &\dot{z} = \Lambda.z, \qquad \text{(using (5) and letting } z := U^T.x = V^{-1}x\text{)}, \end{aligned} \tag{15}$$

and the transient response in the changed coordinate $z \in \mathbb{R}^n$ is given by:

$$
\begin{aligned}
& z(t) = e^{\Lambda t}.z_0, \\
\Rightarrow & V^{-1}.x(t) = e^{\Lambda t}.U^T x_0, \qquad \text{(back to the original coordinate system)} \\
\Rightarrow & x(t) = V.e^{\Lambda t}.U^T x_0, \\
\Rightarrow & x(t) = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n . t} \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix} x_0 \\
\Rightarrow & x(t) = \sum_{i \in \{1,\dots,n\}} [e^{\lambda_i t} u_i^T x_0] v_i.
\end{aligned}
\tag{16}
$$

The above expression of $x(t)$ is called its *modal decomposition*.

The modal decomposition provides useful insight about the system's transient behavior. Each term under the summation of the modal decomposition corresponds to a distinct eigen-value and hence, the i^{th} term is called as the i^{th} *mode* of the transient response $x(t)$, natural frequency of oscillation and attenuation rate of which are governed by the i^{th} eigen-value alone. In particular, if the i^{th} eigen-value is $\lambda_i = \sigma_i + j\omega_i$, then the natural frequency of oscillation of the i^{th} mode of $x(t)$ is ω_i rad./sec and the exponential attenuation rate is σ_i. Hence, $\sigma_i \leq 0$ (resp. $\sigma_i < 0$) for all $i \in \{1, \dots, n\}$ guarantee *stability* (resp. *asymptotic stability*), and $\sigma_i > 0$ for any i implies unstable behavior of the system. This argument regarding stability holds *globally* for a linear system, whereas that for a nonlinear system linearized around an equilibrium holds *locally* within an infinitesimally small neighborhood around the equilibrium.

For each mode i of the transient response, the term within the square bracket (i.e. $e^{\lambda_i t} u_i^T x_0$) is a scalar which determines the amplitude (with an exponential attenuation with time) of oscillation of the mode. However, not all states equally exhibit this amplitude; rather the i^{th} right eigen-vector determines the relative effect of the amplitude on the n states. Higher (resp. lower) the value of the j^{th} element of v_i indicates that higher is the impact of the i^{th} mode on the j^{th} state. This could be insightful to identify the states that are important to be controlled, under a given unstable mode of oscillation.