EE2100: Matrix Analysis

Review Notes - 35

Topics covered :

1. Singular Value Decomposition

---

1. Singular Value Decomposition: A rectangular matrix (say $\mathbf{A} \in \mathcal{R}^{m \times n}$) can be decomposed as

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{1}$$

where $\mathbf{U} \in \mathcal{R}^{m \times m}$ and $\mathbf{V} \in \mathcal{R}^{n \times n}$ are orthonormal matrices whose column vectors are referred to as left singular vectors and right singular vectors respectively. The matrix $\mathbf{\Sigma} \in \mathcal{R}^{m \times n}$ is a matrix that has only diagonal entries (a few of which can be zero). The diagonal entries of $\mathbf{\Sigma}$ are referred to as the singular values.

Proof: For any $\mathbf{A} \in \mathcal{R}^{m \times n}$, the Gram matrix $\mathbf{A}^T\mathbf{A}$ (say $\mathbf{G} \in \mathcal{R}^{n \times n}$) is always real and Symmetric. Accordingly, by Spectral Theorem, the Gram matrix can be expressed as

$$\mathbf{G} = \mathbf{V}\mathbf{D}\mathbf{V}^T \tag{2}$$

where $\mathbf{D} \in \mathcal{R}^{n \times n}$ is a diagonal matrix whose entries are Eigen values of $\mathbf{G}$ (some of which can be zero, if $\mathbf{G}$ is singular). The matrix $\mathbf{V} \in \mathcal{R}^{n \times n}$ is an orthonormal matrix whose column vectors are Eigen vectors of $\mathbf{G}$. Let $\mathbf{v}_i \in \mathcal{R}^n$ denote the i^{th} column vector (or Eigen vector associated with Eigen value λ_i) of $\mathbf{V}$.

Similarly, for any $\mathbf{A} \in \mathcal{R}^{m \times n}$, the matrix $\mathbf{A}\mathbf{A}^T$ (say $\mathbf{H} \in \mathcal{R}^{m \times m}$) is always real and Symmetric. Accordingly, by Spectral Theorem, $\mathbf{H}$ can be expressed as

$$\mathbf{H} = \mathbf{U}\mathbf{D}_1\mathbf{U}^T \tag{3}$$

where $\mathbf{D}_1 \in \mathcal{R}^{m \times m}$ is a diagonal matrix whose entries are Eigen values of $\mathbf{H}$ (some of which can be zero, if $\mathbf{H}$ is singular). The matrix $\mathbf{U} \in \mathcal{R}^{m \times m}$ is an orthonormal matrix whose column vectors are Eigen vectors of $\mathbf{H}$. Let $\mathbf{u}_i \in \mathcal{R}^m$ denote the i^{th} column vector (or Eigen vector associated with Eigen value λ_i) of $\mathbf{U}$.

Further, it is also known that, the rank and nullity of the Gram matrix is the same as the rank and nullity of $\mathbf{A}$ respectively (i.e., $\mathbf{rank}(\mathbf{G})$=$\mathbf{rank}(\mathbf{A})$ and $\mathbf{nullity}(\mathbf{G})$=$\mathbf{nullity}(\mathbf{A})$). Stated alternately, any vector in the null space of $\mathbf{A}$ is also in the null space of $\mathbf{G}$.

Let $r(\leq n)$ denote the rank of $\mathbf{A}$ (i.e., $\mathbf{rank}(\mathbf{A}) = r$). Accordingly, the matrix $\mathbf{G}$ has r non-zero and positive Eigen values (some of which can be repeated).

Let $\sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2, \underbrace{0, 0, \cdots, 0}_{n-r}$ denote the Eigen values of $\mathbf{G}$ (note that we are using a slightly different notation for Eigen values). Let the corresponding Eigen Vectors of $\mathbf{G}$ be denoted by $\mathbf{v}_1, \cdots, \mathbf{v}_r, \mathbf{v}_{r+1}, \cdots, \mathbf{v}_n$ (where $\mathbf{v}_i$). Further, assume that the Eigen values and Eigen vectors of $\mathbf{G}$ are indexed such that $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_r^2, 0 \cdots, 0$ (this is a commonly adopted indexing which we will follow in this course). Under this indexing scheme, the vectors

$\mathbf{v}_i \forall i > r$ are in the null space of $\mathbf{A}$ and $\mathbf{G}$ and the vectors $\mathbf{v}_i \forall i \leq r$ satisfy the equation

$$\mathbf{G}\mathbf{v}_i \ = \ \sigma_i^2 \mathbf{v}_i \tag{4}$$

Pre-multiplying on the either sides of (4) by $\mathbf{A}$ results in

$$\mathbf{A}\mathbf{A}^T\mathbf{A}\mathbf{v}_i \ = \ \sigma_i^2 \mathbf{A}\mathbf{v}_i \implies \mathbf{H}\left(\mathbf{A}\mathbf{v}_i\right) \ = \ \sigma_i^2\left(\mathbf{A}\mathbf{v}_i\right) \tag{5}$$

Equation (5) indicates that $\mathbf{A}\mathbf{v}_i$ (for $i \leq r$) is an Eigen Vector of $\mathbf{H}$ with the associated Eigen value of σ_i^2. Thus $\mathbf{G}$ and $\mathbf{H}$ have the same non-zero Eigen values. The norm of the computed Eigen vectors of $\mathbf{H}$ is

$$\|\mathbf{A}\mathbf{v}_i\|_2 \ = \ \sqrt{\mathbf{v}_i^T\mathbf{A}^T\mathbf{A}\mathbf{v}_i} = \sqrt{\sigma_i^2\mathbf{v}_i^T\mathbf{v}_i} = \sigma_i \tag{6}$$

In order to choose the Eigen vectors of unit norm for $\mathbf{H}$, the Eigen vectors are scaled by σ_i. Accordingly, $\mathbf{u}_i$ as defined in (**??**) are the Eigen vectors i.e.,

$$\mathbf{u}_i = \frac{\mathbf{A}\mathbf{v}_i}{\sigma_i} \tag{7}$$

The r Eigen vectors of $\mathbf{G}$ and the r Eigen vectors of $\mathbf{H}$ satisfy the set of equations (**??**).

$$\left.\begin{aligned} \sigma_1\mathbf{u}_1 &= \mathbf{A}\mathbf{v}_1 \\ \sigma_2\mathbf{u}_2 &= \mathbf{A}\mathbf{v}_2 \\ \vdots \quad &= \vdots \\ \sigma_r\mathbf{u}_r &= \mathbf{A}\mathbf{v}_r \end{aligned}\right\} r \text{ equations} \tag{8}$$

Equation **??** can be alternately represented as

$$[\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_r]\begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix} = \mathbf{A}\left[\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_r\right] \tag{9}$$

Noting that $\mathbf{A}\mathbf{v}_i = \mathbf{0} \ \forall i > r$, (9) can be modified as

$$[\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_r]\begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \end{bmatrix} = \mathbf{A}\left[\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_r \ \mathbf{v}_{r+1} \cdots \ \mathbf{v}_n\right] \tag{10}$$

In order to account for the other possible Eigen vectors of $\mathbf{H}$, (10) can be further modified as

$$[\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_r\ \mathbf{u}_{r+1} \cdots\ \mathbf{u}_m] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} = \mathbf{A}\,[\mathbf{v}_1\ \mathbf{v}_2\ \cdots\ \mathbf{v}_r\ \mathbf{v}_{r+1} \cdots\ \mathbf{v}_n] \tag{11}$$

Equation (11) can be written as

$$\mathbf{U}\boldsymbol{\Sigma} = \mathbf{A}\mathbf{V} \tag{12}$$

where $\mathbf{U} \in \mathcal{R}^{m\times m}$ is an orthonormal matrix that contains the Eigen Vectors of $\mathbf{A}\mathbf{A}^T$, $\mathbf{V} \in \mathcal{R}^{n\times n}$ is an orthonormal matrix that contains the Eigen Vectors of $\mathbf{A}^T\mathbf{A}$ and $\boldsymbol{\Sigma} \in \mathcal{R}^{m\times n}$ is a matrix whose diagonal entries are σ_i. Since $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices, the matrix $\mathbf{A}$ can be written as

$$\mathbf{A} \ = \ \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T \tag{13}$$