EE2100: Matrix Analysis

Review Notes - 30

Topics covered :

1. Spectral Theorem and its Applications

---

1. Spectral Theorem: Let $\mathbf{A} \in \mathcal{R}^{n\times n}$ be a Symmetric matrix (i.e., $\mathbf{A}^T = \mathbf{A}$). The matrix $\mathbf{A}$ can always be represented as $\mathbf{QDQ}^{-1}$, where $\mathbf{Q}$ is an orthonormal matrix (i.e., matrix whose column vectors are of unit norm and are mutually orthogonal) and $\mathbf{D}$ is a diagonal matrix with Eigen values (considering algebraic multiplicities) as its entries.

   Proof: If $\mathbf{A}$ is a real symmetric matrix that has n distinct Eigen values, then it is known that (shown in Lecture 33) the n associated Eigen vectors (picking the vector with unit norm as the Eigen vector associated with a given Eigen Value) are orthogonal. Let $\lambda_1, \cdots, \lambda_n$ denote the distinct Eigen values of a real symmetric matrix with the corresponding Eigen vectors (with unit norm) $\mathbf{v}_1, \cdots, \mathbf{v}_n$ respectively. Since $\mathbf{v}_1, \cdots, \mathbf{v}_n$ are orthogonal (and hence linearly independent) and are the Eigen vectors of $\mathbf{A}$, the product $\mathbf{AQ}$ (where Q is the matrix whose column vectors are $\mathbf{v}_1, \cdots, \mathbf{v}_n$)

$$\mathbf{AQ} = \mathbf{QD} \implies \mathbf{A} = \mathbf{QDQ}^{-1} \tag{1}$$

   Equation (1) proves the spectral decomposition theorem for a real symmetric matrix for n distinct Eigen values. However, the theorem also states that, such a decomposition exists even when the Eigen values are repeated. In this notes, an proof is outlined for a specific scenario, which can be used to prove the theorem for a generic case as well.

   Let $\lambda_1, \cdots, \lambda_{n-2}, \lambda_{n-1}, \lambda_{n-1}$ denote the Eigen values of $\mathbf{A}$ (i.e., it has $n-1$ distinct Eigen values with λ_{n-1} having an algebraic multiplicity of 2). Let $\mathbf{v}_1, \cdots \mathbf{v}_{n-1}$ denote the $n-1$ orthogonal vectors associated with the corresponding distinct Eigen values. Since we are dealing with an n dimensional space, we can find a vector $\mathbf{x}_n$ such that the set $\{\mathbf{v}_1, \cdots, \mathbf{v}_{n-1}, \mathbf{x}_n\}$ that spans $\mathcal{R}^n$ and form the orthogonal basis. Let $\mathbf{Q} = [\mathbf{v}_1, \cdots, \mathbf{v}_{n-1}, \mathbf{v}_n]$ (where $\mathbf{v}_n = \frac{\mathbf{x}_n}{\|\mathbf{x}_n\|_2}$). Notice that $\mathbf{Q}$ is an orthonormal matrix. Then

$$\begin{aligned}\mathbf{AQ} &= [\mathbf{Av}_1, \cdots, \mathbf{Av}_{n-1}, \mathbf{Av}_n] \\ &= [\lambda_1\mathbf{v}_1, \cdots, \lambda_{n-1}\mathbf{v}_{n-1}, \mathbf{Av}_n]\end{aligned} \tag{2}$$

   Since $\{\mathbf{v}_1, \cdots, \mathbf{v}_{n-1}, \mathbf{x}_n\}$ span entire $\mathcal{R}^n$, the term $\mathbf{Av}_n$ can be computed as

$$\mathbf{Av}_n \;=\; \alpha_1\mathbf{v}_1 + \cdots + \alpha_{n-1}\mathbf{v}_{n-1} + \alpha_n\mathbf{v}_n \tag{3}$$

   Since the basis is orthonormal basis, the coefficients α_i (for $i \le n-1$) can be computed as

$$\begin{aligned}\alpha_i =& \frac{<\mathbf{v}_i, \mathbf{Av}_n>}{<\mathbf{v}_i, \mathbf{v}_i>} \text{ for } i \le n-1 \\ =&\mathbf{v}_n^T\mathbf{A}^T\mathbf{v}_i \;=\; \mathbf{v}_n^T\mathbf{Av} \text{ (since } \mathbf{A} \text{ is symmetric)} \\ =&\mathbf{v}_n^T\lambda_i\mathbf{v}_i \;=\; 0 \text{ (since } \mathbf{v}_i \perp \mathbf{v}_n)\end{aligned} \tag{4}$$

Substituting (4) in (3) shows that $\mathbf{v}_n$ is an Eigen vector of matrix $\mathbf{A}$ i.e.,

$$\mathbf{A}\mathbf{v}_n = \alpha_n \mathbf{v}_n \tag{5}$$

Since $\mathbf{v}_n$ is also an Eigen vector (which is linearly independent of $\mathbf{v}_1, \cdots \mathbf{v}_{n-1}$), the matrix $\mathbf{Q}$ becomes a matrix with Eigen vectors as its columns and is orthonormal. Further it can be shown that $\alpha_n = \lambda_{n-1}$ (a trivial proof, left as a part of exercise).

Substituting (5) in (2) results in

$$\begin{aligned}\mathbf{AQ} &= [\lambda_1 \mathbf{v}_1, \cdots, \lambda_{n-1}\mathbf{v}_{n-1}, \lambda_{n-1}\mathbf{v}_n] \\ &= \mathbf{QD} \implies \mathbf{A} = \mathbf{QDQ}^{-1}\end{aligned} \tag{6}$$

2. Representing a real symmetric square matrix as a summation of Rank-1 matrices: Let $\mathbf{A} \in \mathcal{R}^{n\times n}$ denote a real symmetric matrix. Using spectral decomposition theorem, $\mathbf{A}$ can be represented as

$$\begin{aligned}\mathbf{A} &= \mathbf{QDQ}^T = [\mathbf{v}_1, \cdots, \mathbf{v}_n]\,\mathbf{D}\begin{bmatrix}\mathbf{v}_1^T \\ \cdots \\ \mathbf{v}_n^T\end{bmatrix} = [\mathbf{v}_1, \cdots, \mathbf{v}_n]\begin{bmatrix}\lambda_1\mathbf{v}_1^T \\ \cdots \\ \lambda_n\mathbf{v}_n^T\end{bmatrix} \\ &= \lambda_1\mathbf{v}_1\mathbf{v}_1^T + \cdots + \lambda_n\mathbf{v}_n\mathbf{v}_n^T \\ &= \sum_{i=1}^{n}\lambda_i\mathbf{v}_i\mathbf{v}_i^T\end{aligned} \tag{7}$$

The term $\mathbf{v}_i\mathbf{v}_i^T$ is commonly referred to as taking the outer product of a vector with itself. The result of the outer product is a matrix whose Rank is 1. Equation (7) indicates that every real and symmetric square matrix can be represented as a summation of Rank 1 matrices.

3. Rayleigh quotient: Let $\mathbf{A}$ be a real symmetric matrix. We are interested in computing the value $\mathbf{x}^T\mathbf{Ax}$. By the spectral decomposition theorem,

$$\begin{aligned}\mathbf{x}^T\mathbf{Ax} &= \mathbf{x}^T\mathbf{QDQ}^T\mathbf{x} \\ &= \mathbf{y}^T\mathbf{Dy} \text{ where } \mathbf{y} = \mathbf{Q}^T\mathbf{x} \\ &= \sum_{i=1}^{n}\lambda_i y_i^2 \text{ (since } \mathbf{D} \text{ is a diagonal matrix with Eigen values as its entries)}\end{aligned} \tag{8}$$

It is interesting to note that

$$\|\mathbf{y}\|_2 = \sqrt{<\mathbf{Q}^T\mathbf{x}, \mathbf{Q}^T\mathbf{x}>} = \sqrt{\mathbf{x}^T\underbrace{\mathbf{Q}^T\mathbf{Q}}_{\mathbf{I}}\mathbf{x}} = \|\mathbf{x}\|_2 \tag{9}$$

Accordingly (8) be simplified as

$$\mathbf{x}^T\mathbf{Ax} = \sum_{i=1}^{n}\lambda_i x_i^2 \tag{10}$$

If λ_{max} and λ_{min} denote the maximum and minimum Eigen values of $\mathbf{A}$, then it can be deduced that

$$\begin{aligned}\lambda_{min}\sum_{i=1}^{n}x_i^2 \leq \mathbf{x}^T\mathbf{A}\mathbf{x} &\leq \lambda_{max}\sum_{i=1}^{n}x_i^2 \\ \lambda_{min}\|\mathbf{x}\|_2 \leq \mathbf{x}^T\mathbf{A}\mathbf{x} &\leq \lambda_{max}\|\mathbf{x}\|_2 \\ \lambda_{min} \leq \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\|\mathbf{x}\|_2} &\leq \lambda_{max}\end{aligned} \tag{11}$$