MODULE 1.1
Spectral decomposition of a real symmetric matrix

by

S.Lakshmivarahan
School of Computer Science
University of Oklahoma
Norman, OK-73019, USA
varahan@ou.edu

Eigenvalue and eigenvector pair of a matrix

- Let $A \in R^{n \times n}$ be a real matrix of order n
- If there exist a scalar, $\lambda(real/complex)$ and a vector, v(real/complex) such that

$$Av = \lambda v \tag{1}$$

  then λ is the eigenvalue and v is the corresponding eigenvector of A
- The pair (λ, v) satisfying (1) is called an eigenpair of A
- The set of all eigenvalues of A is called the spectrum of A

Invariant subspace of A

- Let $S_k = \{v_1, v_2, ..., v_k\}$ be a set of linearly independent vectors in R^n
- $SPAN(S_k)$ denotes the set of all linear combinations of the vectors in S_k
- $SPAN(S_k)$ is a K-dimensional subspace of R^n
- If $AX \in SPAN(S_k)$ for any $X \in SPAN(S_k)$, then S_k is said to be A-invariant
- From (1), since $Av \in SPAN(v)$, every eigenvector defines an invariant subspace of dimension 1.

- Rewrite(1) as a linear homogeneous system:

$$(A - \lambda I)v = 0 \qquad (2)$$

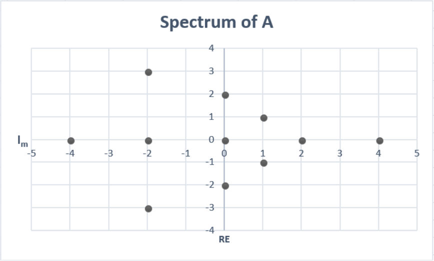
- Equation (2) has a non-null solution, exactly when $(A - \lambda I)$ is singular, that is

$$p(\lambda) = |A - \lambda I| = 0 \qquad (3)$$

- The n eigenvalues of A are given by the n roots of the characteristic polynomial, $p(\lambda)$ of A

Distribution of eigenvalues of A

- Since A is real, the coefficients of $p(\lambda)$ are also real
- An n^{th} degree polynomial of degree n has n roots
- The roots are real or complex and the complex roots occur in conjugate pairs
- A typical distribution of eigenvalues



Eigenpairs of a real symmetric matrix

- Let $A \in R^{n\times n}$ and $A^T = A$, that is, A is symmetric
- SM1: The eigenvalues of a real symmetric matrix are real
- SM2: Eigenvectors corresponding to distinct eigenvalues are orthogonal
- SM3: If λ as a root of $p(\lambda) = 0$ in (3) is of (algebraic) multiplicity $1 \leq k \leq n$, then there exists a set of k mutually orthogonal vectors $v_1, v_2, v_3, ..., v_k$ such that (λ, v_i) is an eigenpair of A for $1 \leq i \leq k$, that is, k is also the geometric multiplicity which is the dimension of the invariant subspace spanned by $\{v_1, v_2, ...v_k\}$ where $Av_i = \lambda v_i$ for $1 \leq i \leq k$

- Let (λ_i, v_i) such that

$$Av_i = \lambda_i v_i \tag{4}$$

- Define

$$V = [v_1, v_2, ..., v_n] \in R^{n \times n}$$

$$\Lambda = Diag[\lambda_1, \lambda_2, ..., \lambda_n] \in R^{n \times n}$$

- Then (4) becomes:

$$AV = V\Lambda \tag{5}$$

- The eigenvectors are mutually orthogonal (see appendix)

$$v_i^T v_j \neq 0 \quad for \quad i = j$$

$$= 0 \quad \text{otherwise}$$

Orthonormality of eigenvectors

- Since $Av = \lambda v \implies A(\alpha v) = \lambda(\alpha v)$ for any α, non-zero constant, we need to only consider unit vector for eigenvectors.
- Consequently, assume that the vectors v_i in (4) are orthonormal:

$$\begin{aligned} {v_i}^T v_j &= 1 \quad \textit{if } i = j \\ &= 0 \quad \textit{otherwise} \end{aligned} \tag{6}$$

- Hence, V is an orthogonal matrix, that is, using (6):

$$V^T V = \begin{bmatrix} v_1^T \\ v_2^T \\ v_3^T \\ \vdots \\ v_n^T \end{bmatrix} \begin{bmatrix} v_1, v_2, v_3, \ldots, v_n \end{bmatrix}$$

$$= \begin{bmatrix} v_1^T v_1 & v_1^T v_2 & \ldots & v_1^T v_n \\ v_2^T v_1 & v_2^T v_2 & \ldots & v_2^T v_n \\ v_3^T v_1 & v_3^T v_2 & \ldots & v_3^T v_n \\ & \vdots & & \\ v_n^T v_1 & v_n^T v_2 & \ldots & v_n^T v_n \end{bmatrix}$$

$$= I_n = VV^T \tag{7}$$

Spectral or eigen decomposition of a symmetric matrix

- Multiplying both sides of (5) by V^T and using (7), we obtain

$$A = AVV^T = V\Lambda V^T \tag{8}$$

- This multiplicative decomposition in (8) is called the eigen decomposition of A

- Expanding V and Λ in (8): A =

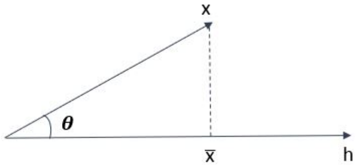
$$\begin{bmatrix} v_1, v_2, v_3, \ldots, v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix}$$

$$= \sum_{i=1}^{n} \lambda_i V_i V_i^T \tag{9}$$

- Since $V_i V_i^T$ is a rank-1 (outer product) matrix, (9) expresses A as a sum of n linearly independent rank-1 matrices

A digression

- Consider:



- Let $\hat{h} = \frac{h}{||h||}$ be the unit vector along h
- Orthogonal projection, $\bar{x}$ of x along h is given by

$$\bar{x} = (x^T\hat{h})\hat{h} = (\hat{h}^T x)\hat{h} = \hat{h}(\hat{h}^T x) = (\hat{h}\hat{h}^T)x \tag{10}$$

- The rank-1 matrix

$$P_h = (\hat{h}\hat{h}^T) \tag{11}$$

is called an orthogonal projection matrix and (10) becomes:

$$\bar{x} = P_h x$$

- Consequently, the rank-1 matrix $v_i v_i^T$ in (9) is an orthogonal projection matrix along v_i
- That is, (9) expresses A as a linear combination of orthogonal projection matrices

A-symmetric and positive definite (SPD)

- In this case, the eigenvalues of A are all real and positive
- That is, we can express

$$\Lambda = \Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}} \tag{12}$$

where

$$\Lambda^{\frac{1}{2}} = Diag(\lambda_1^{\frac{1}{2}}, \lambda_2^{\frac{1}{2}}, \dots, \lambda_n^{\frac{1}{2}})$$

- $A = V\Lambda V^T = V\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}V^T$

$$= (V\Lambda^{\frac{1}{2}})(V\Lambda^{\frac{1}{2}})^T = \bar{V}\bar{V}^T \tag{13}$$

is the another form of the eigen decomposition for A

Why SPD matrices?

- In multivariate statistical analysis, SPD matrices arise naturally as covariance matrices
- In fact, the many well known methods in multivariate statistical analysis such as
    - Principal Component Analysis (PCA)
    - Singular Value Decomposition (SVD)
    - Cononical Correlation (CC)

  are based on the spectral or eigen decomposition of SPD matrices

APPENDIX

- The goal of this appendix is to provide a proof of various properties of real symmetric matrices used in the development of this module
- The final result is to prove that every real symmetric matrix is diagonalizable using orthogonal transformation

- Let A be a real symmetric matrix of order $n \geq 2$
- The characteristic polynomial equation

$$p(\lambda) = |A - \lambda I| = 0 \tag{14}$$

of degree n must have at least one solution, say, α

- Then, there is atleast one real eigenvector that lies in the null space of $(A - \lambda I)$ or the kernel of $(A - \lambda I)$

A factorization of p(λ)

- In general, the monic polynomial can be expressed as

$$p(\lambda) = \prod_{i=1}^{k} (\lambda - \lambda_i)^{n_i} \tag{15}$$

  where n_i is the algebraic multiplicity of λ_i and $(n_1 + n_2 + \ldots + n_k) = n$
- The number of distinct eigenvectors m_i corresponding to a given eigenvalue λ_i is called the geometric multiplicity
- In general, $1 \leq m_i \leq n_i$, when A is symmetric, $m_i = n_i$ for $1 \leq i \leq k$

Claim 1: Eigenvalues of a real symmetric matrix are real

- Let (λ, v) be an eigenpair of A. That is

$$Av = v\lambda \tag{16}$$

- Taking complex conjugates of both sides:

$$A\bar{v} = \bar{v}\bar{\lambda} \tag{17}$$

- Multiplying both sides of (16) by $\bar{v}^T$ on the left and that of (17) by v^T on the left and subtracting

$$0 = \bar{v}^T Av - v^T A\bar{v} = \lambda\bar{v}^T v - \bar{\lambda} v^T \bar{v} = v^T \bar{v}(\lambda - \bar{\lambda}) \tag{18}$$

- Since $v^T\bar{v}{>}0, \implies \lambda = \bar{\lambda}$ and hence the claim

Claim 2: Eigenvectors corresponding to different eigenvalues of a real symmetric matrix are orthogonal

- Let (λ, v) and (μ, u) be two eigenpairs of a symmetric matrix A and let $\lambda \neq \mu$
- $$Then \quad Av = \lambda v \quad and \quad Au = \mu u \tag{19}$$
- Multiplying both sides of the first equation on the left by u^T and that of the second by v^T and subtracting:

$$0 = u^T Av - v^T Au = \lambda u^T v - \mu v^T u = (\lambda - \mu) u^T v \tag{20}$$

- Since $\lambda \neq \mu$, it is immediate that $u^T v = 0$ and the claim follows

A notation

- Let $D \subset R^n$ denote an A-invariant subspace of A. That is, $Av \in D$ when $v \in D$
- Let $D^\perp$ denote the subspace of R^n that is orthogonal to D. That is $u^T v = 0$ whenever $u \in D$ and $v \in D^\perp$

Claim 3: If $D \subseteq R^n$ is A-invariant, then so is $D^\perp$

- For any $u, v \in R^n$
$$v^T A u = (Av)^T u \tag{21}$$
- If $u \in D$, then $Au \in D$. If $v \in D^\perp$, then $v^T A u = 0$
- From $(Av)^T u = 0$, it follows that $Av \in D^\perp$, and the claim is true.

Claim 4: Every (non-null) A-invariant subspace D of A contains a real eigenvector of A

- Let k be the dimension of D. Then there exists a $n \times k$ matrix B whose columns constitute an orthogonal basis for D.
- Since D is A-invariant, it is immediate that

$$AB = BE \tag{22}$$

  for some $E \in R^{k \times k}$
- Then,

$$B^T AB = B^T BE = E \tag{23}$$

  where E is a real symmetric matrix

Proof of claim 4 (Continues)

- Since E is real and symmetric, there exists atleast one eigenpair (λ, x) for E: $Ex = \lambda x$ where $x \in R^k$
- Then $(AB)x = A(Bx) = (BE)x = B(Ex) = \lambda(Bx)$
- Since $x \neq 0$ and the columns of B are orthogonal and hence linearly independent, it follows that $Bx \neq 0$
- Hence, Bx is an eigenvector A contained in D

Claim 5: The set of all n eigenvectors of a real symmetric matrix $A \in R^{n \times n}$ form an orthogonal basis for R^n

- Recall that every real symmetric matrix A is endowed with at least one eigen pair
- Hence, for some $m \geq 1$, let $\{v_1, v_2, \ldots, v_m\}$ be the (orthonormal) eigenvector basis for a subspace D of R^n
- Clearly, D and $D^\perp$ are A-invariant.Hence, there is a vector $v_{m+1} \in D^\perp$ such that $\{v_1, v_2, \ldots, v_{m+1}\}$ are the eigenvectors of A.
- Starting with m=1 and using this inductive argument, we obtain an orthonormal basis for R^n which are eigenvectors of A

Claim 6: Every real symmetric matrix A is diagonalizable

- Given A, let v=$[v_1, v_2, \ldots, v_n] \in R^{n\times n}$ be the matrix of eigenvectors of A, that is $Av_i = v_i\lambda_i$ and $\Lambda = Diag(\lambda_1, \lambda_2, \ldots, \lambda_n) \in R^{n\times n}$
- Then $AV = V\Lambda$ and $V^TV = VV^T = I$
- Hence, $V^TAV = \Lambda$

References

- Appendix follows the developments in Chapter 8 of C. Godsil and G. Royle (2001) Algebraic Graph Theory, Springer Verlag
- G.Golub and C. Van Loan (1989) Matrix Computations, Johns Hopkins University Press contains a wealth of information on computing the eigen pairs of real matrices