Linear Algebra:

Characteristic Value Problem

1. The Characteristic Value Problem

- Let $\Re$ be the set of real numbers and $\mathbb{C}$ be the set of complex numbers.
- Given an $n \times n$ real matrix A, does there exist a number $\lambda \in \mathbb{C}$ and a non-zero vector $x \in \mathbb{C}^n$ such that

$$Ax = \lambda x? \tag{1}$$

This is known as the *characteristic value problem* or the *eigenvalue problem*.

– If $x \neq 0$ and λ satisfy the equation $Ax = \lambda x$, then λ is called a *characteristic value* or *eigenvalue* of A, and x is called a *characteristic vector* or *eigenvector* of A.

– Clearly, (1) holds if and only if

$$(A - \lambda I)\, x = 0. \tag{2}$$

- But (2) holds for non-zero x if and only if the column vectors of $(A - \lambda I)$ are linearly dependent, that is,

$$|A - \lambda I| = 0. \tag{3}$$

This equation is called the *characteristic equation* of A.

– Consider the expression

$$f(\lambda) = |A - \lambda I| \,. \tag{4}$$

- f is a polynomial of degree n in λ. It is called the *characteristic polynomial* of A.

- **Example:** Consider the 2×2 matrix

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} .$$

– The characteristic equation is:

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0, \Rightarrow (2-\lambda)^2 - 1 = 0, \Rightarrow (\lambda - 1)(\lambda - 3) = 0.$$

Thus, the characteristic roots are $\lambda = 1$ and $\lambda = 3$.

– Putting $\lambda = 1$ in (2), we get

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which yields

$$x_1 + x_2 = 0.$$

- Thus the general solution of the characteristic vector corresponding to the characteristic root $\lambda = 1$ is given by

$$(x_1, x_2) = \theta\,(1, -1)\,, \text{ for } \theta \neq 0.$$

- Similarly, corresponding to the characteristic root $\lambda = 3$, we have the characteristic vector given by

$$(x_1, x_2) = \theta\,(1, 1)\,, \text{ for } \theta \neq 0.$$

- In general, the characteristic equation will have n roots in the complex plane (by the "Fundamental Theorem of Algebra"), since it is a polynomial equation (in λ) of degree n. (Of course some of these roots might be repeated.)
  - In general, the corresponding eigenvectors will also have their components in the complex plane.
- **Example:** Consider the 2×2 matrix

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}.$$

  - Check that the eigenvalues are complex numbers and the eigenvectors have complex components.

2. Characteristic Values, Trace & Determinant of a Matrix

- If A is an $n \times n$ matrix, the *trace* of A, denoted by tr(A), is the number defined by

$$\text{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

#1. Consider the 2×2 matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.

- Write down the characteristic equation.
- Suppose λ_1 and λ_2 are two characteristic values of A. Prove that

  (a) $\lambda_1 + \lambda_2 = \text{tr}(A)$; and (b) $\lambda_1 \lambda_2 = |A|$.

- In general, for an $n \times n$ matrix A, with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$, we have

$$\sum_{i=1}^{n} \lambda_i = \text{tr}(A), \text{ and } \prod_{i=1}^{n} \lambda_i = |A|.$$

3. Characteristic Values & Vectors of Symmetric Matrices

- There is considerable simplification in the theory of characteristic values if A is a *symmetric* matrix.
- **Theorem 1:**

  If A is an $n \times n$ symmetric matrix, then all the eigenvalues of A are real numbers and its eigenvectors are real vectors.

  – We will develop the theory of eigenvalues and eigenvectors *only for symmetric matrices*.
- **Normalized Eigenvectors:**

  – Note that if x is an eigenvector corresponding to an eigenvalue λ, then so is tx, where t is any non-zero scalar.

  – So we normalize the eigenvectors.

  – A *normalized eigenvector* is an eigenvector with (Euclidean) norm equal to 1.

– Let $x = (x_1, x_2, ..., x_n)$ be an eigenvector with norm $\|x\| = \left(x_1^2 + x_2^2 + ... + x_n^2\right)^{\frac{1}{2}}$.

- Since x is a non-zero vector, $\|x\| > 0$. Define

$$y = \left(\frac{x_1}{\|x\|}, \frac{x_2}{\|x\|}, ..., \frac{x_n}{\|x\|}\right).$$

Now we have $\|y\| = 1$, that is, y is a normalized eigenvector.

4. Spectral Decomposition of Symmetric Matrices

- (To proceed, we specialize further – we consider *symmetric* matrices with *distinct* eigenvalues, that is, $\lambda_i \neq \lambda_j$.)

- **Orthogonal Matrix:** An $n \times n$ matrix C is called an *orthogonal* matrix if C is invertible and its inverse equals its transpose, that is, $C^T = C^{-1}$.

- **Theorem 2:**

  Let A be an $n \times n$ symmetric matrix with n distinct eigenvalues, $\lambda_1, \lambda_2, ..., \lambda_n$. *If $y^1, y^2, ..., y^n$ are (normalized) eigenvectors corresponding to the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$, then the matrix B such that B^i, the i-th column vector of B, is the vector y^i* ($i = 1, 2, ..., n$), *is an orthogonal matrix.*

  – Proof: To be discussed in class.

– Hints: 3 steps:

- **Step 1.** If $i \neq j$, then $\left(y^i\right)^T y^j = 0.$

  Use the definition (equation (1)) to prove that $\left(y^j\right)^T A y^i = \lambda_i \left(y^j\right)^T y^i$ and, at the same time, $\left(y^j\right)^T A y^i = \lambda_j \left(y^j\right)^T y^i.$

- **Step 2.** B is invertible.

  Show that the vectors $y^1, y^2, ..., y^n$ are linearly independent.

- **Step 3.** $B^{-1} = B^T$.

  Show directly that $B^T B = I$ (the identity matrix).

- **Theorem 3 (Spectral Decomposition):**

  Let A *be an* $n \times n$ *symmetric matrix with* n *distinct eigenvalues,* $\lambda_1, \lambda_2, ..., \lambda_n$. *Let* $y^1, y^2, ..., y^n$ *are (normalized) eigenvectors corresponding to the eigenvalues* $\lambda_1, \lambda_2, ..., \lambda_n$. *Let* B *be the* $n \times n$ *matrix such that* B^i, *the* i*-th column vector of* B, *is the vector* y^i ($i = 1, 2, ..., n$). *Then*

  $$A = BLB^T$$

  where L *is the diagonal matrix with the eigenvalues of* A ($\lambda_1, \lambda_2, ..., \lambda_n$) *on its diagonal.*

  – **Remark:** The above expression shows that matrix A can be "decomposed" into a matrix L consisting of its eigenvalues on the diagonal and the matrices B and B^T which consists of its eigenvectors.

  – **Proof:** To be discussed in class.

– **Hints:** Consider any two $n \times n$ matrices, C and D.

- Let $(C_1, C_2, ..., C_n)$ be the set of row vectors of C and $\left(D^1, D^2, ..., D^n\right)$ be the set of column vectors of D.

- To prove Theorem 3 use the following observation on matrix multiplication:

$$CD = \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_4 \end{pmatrix} \begin{pmatrix} D^1 & D^2 & \cdots & D^n \end{pmatrix}$$

$$= \begin{pmatrix} C_1D^1 & C_1D^2 & \cdots & C_1D^n \\ C_2D^1 & C_2D^2 & \cdots & C_2D^n \\ \vdots & \vdots & \ddots & \vdots \\ C_nD^1 & C_nD^2 & \cdots & C_nD^n \end{pmatrix}$$

$$= \begin{pmatrix} CD^1 & CD^2 & \cdots & CD^n \end{pmatrix}.$$

5. Quadratic Forms

- A *quadratic form* on $\Re^n$ is a real-valued function of the form

$$\begin{aligned} Q(x_1, x_2, ..., x_n) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j \\ &= a_{11}x_1x_1 + a_{12}x_1x_2 + ... + a_{1n}x_1x_n \\ &\quad + a_{21}x_2x_1 + a_{22}x_2x_2 + ... + a_{2n}x_2x_n \\ &\quad + ... + a_{n1}x_nx_1 + a_{n2}x_nx_2 + ... + a_{nn}x_nx_n. \end{aligned}$$

- A quadratic form Q can be represented by a matrix A so that

$$Q(x) = x^T A x$$

where

$$A_{n\times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \text{ and } x_{n\times 1} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

- Note that a_{ij} and a_{ji} are both coefficients of $x_i x_j$ when $i \neq j$.
  - The coefficient of $x_i x_j$ is $(a_{ij} + a_{ji})$ when $i \neq j$.
  - If $a_{ij} \neq a_{ji}$, we can uniquely define new coefficients

$$b_{ij} = b_{ji} = \frac{a_{ij} + a_{ji}}{2}, \text{ for all } i, j$$

    so that $b_{ij} + b_{ji} = a_{ij} + a_{ji}$, and $B = (b_{ij}) = B^T$, that is B is a symmetric matrix.
  - This redefinition of the coefficients does not change the value of Q for any x.
  - Thus we can always assume that the matrix A associated with the quadratic form $x^T A x$ is symmetric.

- Examples:
  - The general quadratic form in two variables is

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2.$$

In matrix form this can be written as

$$(x_1 \quad x_2)\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

  - The general quadratic form in three variables

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3$$

can be written as

$$(x_1 \quad x_2 \quad x_3)\begin{pmatrix} a_{11} & \frac{1}{2}a_{12} & \frac{1}{2}a_{13} \\ \frac{1}{2}a_{12} & a_{22} & \frac{1}{2}a_{23} \\ \frac{1}{2}a_{13} & \frac{1}{2}a_{23} & a_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

- **Definiteness of Quadratic Forms:**

  Let A be a symmetric $n \times n$ matrix. Then

  – A is *positive definite* if $x^T Ax > 0$ for all x in $\Re^n$, $x \neq 0$.

  – A is *negative definite* if $x^T Ax < 0$ for all x in $\Re^n$, $x \neq 0$.

  – A is *positive semi-definite* if $x^T Ax \geq 0$ for all x in $\Re^n$.

  – A is *negative semi-definite* if $x^T Ax \leq 0$ for all x in $\Re^n$.

  – A is *indefinite* if $x^T Ax > 0$ for some x in $\Re^n$ and $x^T Ax < 0$ for some other x in $\Re^n$.

  – **Remarks:**

  Consider, for example, the definition of positive definiteness. Note that the relevant inequality must hold for *every* vector $x \neq 0$ in $\Re^n$.

  - Observation 1: If we know that A is positive definite, then we should be able to infer some useful properties of A quite easily.

#2. Prove that if a symmetric $n \times n$ matrix A is positive definite then all its diagonal elements must be positive.

- Observation 2: On the other hand, if we do not know that A is positive definite, then the above definition by itself will not be very easy to check to determine whether A is positive definite or not.

$\rightarrow$ This observation leads one to explore convenient characterizations of quadratic forms.

6. Characterization of Quadratic Forms

- **Theorem 4:**

  Let A be a symmetric $n \times n$ matrix with distinct eigenvalues, $\lambda_1, \lambda_2, ..., \lambda_n$. *Then*

(a) A *is positive (negative) definite if and only if every eigenvalue of A is positive (negative)*;

(b) A *is positive (negative) semi-definite if and only if every eigenvalue of A is non-negative (non-positive)*;

(c) A *is indefinite if and only if A has a positive eigenvalue and a negative eigenvalue.*

  – Proof: To be discussed in class.

  – Hints: Use the spectral decomposition of A given in Theorem 3.

#3. **Examples:** Consider the following matrices:

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}; B = \begin{pmatrix} -1 & 1 \\ 1 & -3 \end{pmatrix}; C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Find out their eigenvalues, and characterize them (in terms of definiteness).

7. Alternative Characterization of Quadratic Forms

- An alternative way to characterize quadratic forms is in terms of the signs of the "principal minors" of the corresponding matrix.
- If A is an $n \times n$ matrix, **a principal minor of order** r is the determinant of the $r \times r$ submatrix that remains when $(n-r)$ rows and $(n-r)$ columns *with the same indices* are deleted from A.
- **Example:** Consider the 3×3 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

  – The principal minors of order 2 are:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}; \quad \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}; \quad \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}.$$

– The principal minors of order 1 are:

$$a_{11};\ a_{22};\ a_{33}.$$

– The principal minors of order 3 is: $|A|$.

- If A is an $n \times n$ matrix, the **leading principal minor of order** r is defined as

$$A_r = \begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}.$$

- **Example:** For the 3×3 matrix A, the three leading principal minors are

$$A_1 = a_{11};\ A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix};\ A_3 = |A|\ .$$

- **Theorem 5:**

  Let A be a symmetric $n \times n$ matrix. Then

(a) *A is positive definite if and only if all its n leading principal minors are positive.*

(b) *A is negative definite if and only if all its n leading principal minors alternate in sign, starting with negative.* (*That is*, *the r-th leading principal minor*, A_r, $r = 1, 2, ..., n$, *has the same sign as* $(-1)^r$.)

(c) *If some r-th leading principal minor of A (or some pair of them) is non-zero but does not fit into either of the two sign patterns in* (a) *and* (b)*, then A is indefinite.*

(d) *A is positive semi-definite if and only if every principal minor of A of every order is non-negative.*

(e) *A is negative semi-definite if and only if every principal minor of A of odd order is non-positive and every principal minor of even order is non-negative.*

– Proof: See section 16.4 (pages 393 – 395) of the textbook.

• **Examples:** Consider the matrices studied above:

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}; \ B = \begin{pmatrix} -1 & 1 \\ 1 & -3 \end{pmatrix}; \ C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

#4. Characterize the matrices A, B and C (in terms of definiteness) using the principal minors criteria.

References

- Must read the following sections from the textbook:
  - Section 23.1 (pages 579 – 585): Definitions and Examples of Eigenvalues and Eigenvectors;
  - Section 23.3 (pages 597 – 601): Properties of Eigenvalues;
  - Section 23.7 (pages 620 – 626): Symmetric Matrices;
  - Sections 16.1 and 16.2 (pages 375 – 386): Quadratic Forms and Definiteness of Quadratic Forms;
  - Section 23.8 (pages 626 – 627): Definiteness of Quadratic Forms.
- Most of the materials covered can be found in

1. Hohn, Franz E., *Elementary Matrix Algebra*, New Delhi: Amerind, 1971 (chapters 9, 10),
2. Hadley, G., *Linear Algebra*, Massachusetts: Addison-Wesley, 1964 (chapter 7).