[Academic Script]

[Eigenvalues and Quadratic Forms]

Subject:	Business Economics
Course:	B.A., 6^{th} Semester, Undergraduate
Paper No. & Title:	631 (Six Three One) Advanced Mathematical Techniques
Unit No. & Title:	Unit - 1 Linear Algebra
Lecture No. & Title:	1 (One) Eigenvalues and Quadratic Form

Eigenvalues and Quadratic Forms

Introduction: In this module we learn some very important and useful concepts of matrix algebra like eigenvalues, eigenvectors and quadratic forms. These concepts play a vital role in the theory of economics. The use of eigenvalues and eigenvectors is eminent in the study of various statistical and economic models especially those dealing with solving system of linear differential equations. The problems of finding maxima or minima of functions of two or more variables occur frequently in economics. In such problems we need to deal with finding the signs of the quadratic forms which arise through the second-order conditions for maximality or minimality of functions at a point. In this sense study of quadratic forms is quite important for economists.

With this little motivation we now introduce eigenvalues through the eigenvalue problem.

The Eigenvalue Problem

Given a square matrix A of order n, the eigenvalue problem is to determine a non-zero (column) vector $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ in $\mathbb{R}^n$ and a real scalar λ satisfying

$$\boldsymbol{Ax = \lambda x}.$$

We may interpret the eigenvalue problem as a problem of solving the system of n linear equations in two unknowns; a vector x and a scalar λ. Such situations occur frequently in economics and econometrics. This leads us to the definition of eigenvalues and eigenvectors.

Definition: Eigenvalues and Eigenvector

$\mathbb{R}$

Let A be a square matrix of order n. A real number λ is called an eigenvalue of the matrix A if there exists a non-zero (column) vector $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ in $\mathbb{R}^n$ satisfying $Ax = \lambda x$. The vector x is called an eigenvector of the matrix A associated to the eigenvalue λ.

Later we shall see through an example that more than one eigenvectors can associate with the same eigenvalue λ.

Eigenvalues and Eigenvectors are also known as characteristic roots and characteristic vectors of a matrix respectively.

Let us now discuss the method of determining the eigenvalues and eigenvectors.

Note that the equation $Ax = \lambda x$ is same as $(A - \lambda I)x = 0$ where I denotes the identity matrix of order n. But we know that the system of equations $(A - \lambda I)x = 0$ has a non-zero solution if and only if the determinant of the matrix $A - \lambda I$ is zero. Thus if we denote the determinant of $A - \lambda I$ by $|A - \lambda I|$ then the eigenvalues of the matrix are simply the roots of the equation

$$|A-\lambda I\,|=0$$

This equation is known as the characteristic equation of the matrix A whereas $|A-\lambda I|$, which is a polynomial in λ of degree n, is known as the characteristic polynomial of the matrix A.

Here it is worth mentioning that every matrix satisfies its characteristic equation.

In other words,

If $f(\lambda)=|A-\lambda I|$ denotes the characteristic polynomial of the matrix A, then $f(A)=0$. This result is known as Cayley-Hamilton theorem and it is often useful in computing the higher powers of matrix A or its inverse if it exists.

Now since $|A-\lambda I|$ is a polynomial in λ of degree n the characteristic equation of a matrix A of order n has at the most n real roots in general.

Note that it is very much possible that a polynomial of an even degree has no real roots and so a square matrix of an even order may not have any real eigenvalues. However if the matrix is symmetric then it can be proved that its characteristic equation has exactly n real roots, which may not be necessarily distinct. Thus we can say that if A is a square matrix of order n then it has at the most n eigenvalues and further if A is symmetric then it has exactly n eigenvalues which are usually denoted by $\boldsymbol{\lambda_{1,} \lambda_2, \dots, \lambda_n}$ but not necessarily distinct.

Let us now consider some examples about computing the eigenvalues and eigenvectors.

Example 1:

Determine the eigenvalues and eigenvectors of the matrix A of order 2 whose rows are $\begin{bmatrix} \mathbf{9} & \mathbf{2} \\ \mathbf{2} & \mathbf{6} \end{bmatrix}$

Solution: Here the characteristic equation of the matrix A **is**

$$|\boldsymbol{A} - \boldsymbol{\lambda I}| = \mathbf{0}$$

which on expansion of the determinant is same as $\boldsymbol{\lambda}^2 - \mathbf{15}\boldsymbol{\lambda} + \mathbf{50} = \mathbf{0}$. **But** $\boldsymbol{\lambda}^2 - \mathbf{15}\boldsymbol{\lambda} + \mathbf{50} = (\boldsymbol{\lambda} - \mathbf{5})(\boldsymbol{\lambda} - \mathbf{10})$ **and hence the eigenvalues of** A **are** $\boldsymbol{\lambda_1} = \mathbf{5}$ **and** $\boldsymbol{\lambda_2} = \mathbf{10}$.

In order to determine an eigenvector associated to the eigenvalue $\boldsymbol{\lambda_1} = \mathbf{5}$ **we have to solve the matrix equation** $(\boldsymbol{A} - \mathbf{5}\boldsymbol{I})\boldsymbol{x} = \mathbf{0}$ **which is same as solving the pair of equations** $\mathbf{4}\boldsymbol{x_1} + \mathbf{2}\boldsymbol{x_2} = \mathbf{0}$ **and** $\mathbf{2}\boldsymbol{x_1} + \boldsymbol{x_2} = \mathbf{0}$**. Since these two equations are identical any non-zero vector** $(\boldsymbol{x_1}, \boldsymbol{x_2})$ **in which** $\boldsymbol{x_2} = -\mathbf{2}\boldsymbol{x_1}$ **is an eigenvector of** A **associated with the eigenvalue** $\mathbf{5}$.

Similarly for finding an eigenvector associated to the eigenvalue $\lambda_2 = 10$ we have to solve the matrix equation $(A - 10I)x = 0$

which is same as solving the pair of equations $-x_1 + 2x_2 = 0$ and $2x_1 - 4x_2 = 0$. Thus any non-zero vector (x_1, x_2) in which $x_1 = 2x_2$ is an eigenvector of A associated with the eigenvalue 10.

We saw that there are infinitely many eigenvectors associated with the eigenvalues 5 and 10 of matrix A. However, if we insist on eigenvectors of unit length then they are $\left(\frac{1}{\sqrt{5}}, \frac{-2}{\sqrt{5}}\right)$ and $\left(\frac{-1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$ associated to the eigenvalue 5 and those associated to the eigenvalue 10 are $\left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$ and $\left(\frac{-2}{\sqrt{5}}, \frac{-1}{\sqrt{5}}\right)$. Now we consider an example of a matrix of order 3.

Example 2:

Determine the eigenvalues and eigenvectors of the matrix A of order 3 whose rows are $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Solution:

It is easily observed that the characteristic equation of the matrix A here is

$$(1 - \lambda)(2 - \lambda)(1 - \lambda) = 0$$

and hence the eigenvalues of A are $1, 2$ and 1.

Note that although we have a matrix of order 3 we have only two distinct eigenvalues. For eigenvectors corresponding to the eigenvalue 1, we solve the matrix equation $(A - I)x = 0$ and this gives $x_2 = 0$.

Thus any non-zero vector (x_1, x_2, x_3) in which $x_2 = 0$,

is an eigenvector of A associated with the eigenvalue 1.

On the other hand it may be verified that the eigenvectors corresponding to the eigenvalue 2 are non-zero vectors (x_1, x_2, x_3) in which $x_1 = x_3 = 0$.

It is clear from Example 2 that the eigenvalues of a diagonal matrix are simply its diagonal elements. We now record few more interesting results regarding the eigenvalues and eigenvectors.

Theorem:

If λ is an eigenvalue of a square matrix A then it is also an eigenvalue of its transpose matrix A^T.

Theorem:

If λ is an eigenvalue of an invertible (non-singular) matrix A then λ^{-1} is an eigenvalue of its inverse matrix A^{-1}.

Theorem:

If λ is an eigenvalue of a square matrix A then λ^k is an eigenvalue of the matrix A^k for every positive integer k.

Theorem:

Let A be a square matrix of order n. Then the eigenvectors corresponding to the distinct eigenvalues of A are linearly independent vectors.

In addition to these results, if A is symmetric then we have some additional information regarding its eigenvalues and eigenvectors.

Theorem:

Let A be a symmetric matrix of order n. Then real eigenvalues of A exist and further if $\lambda_{1,}\lambda_2, \dots, \lambda_r$ are its distinct real eigenvalues which are repeated $k_{1,}k_2, \dots, k_r$ times respectively then

$$k_1+k_2+ \dots + k_r = n.$$

Theorem:

Let A be a symmetric matrix of order n. If q_1 and q_2 are the eigenvectors corresponding to the distinct eigenvalues λ_1 and λ_2 of A then q_1 and q_2 are orthogonal vectors.

Theorem:

Let A be a symmetric matrix of order n and suppose $\lambda_{1,}\lambda_2,\dots,\lambda_r$ are its distinct real eigenvalues which are repeated $k_{1,}k_2,\dots,k_r$ times respectively. Then for each $1 \leq j \leq r$, there exist k_j mutually orthonormal eigenvectors associated with the eigenvalue λ_j. Further, the matrix P formed by these n mutually orthonormal eigenvectors is orthogonal and $P^{-1}AP$=Λ is a diagonal matrix whose diagonal elements are simply the eigenvalues of A.

In the above theorem since $P^{-1}AP$ produces a diagonal matrix we often say that P diagonalizes the symmetric matrix A. As diagonal matrices are easy to deal with in many scenarios, the last theorem leads us to the topic of matrix diagonalization.

The Diagonalization of a Square Matrix

Definition: Similar Matrices

Let A and B be square matrices of the same order. If there exists an invertible matrix P such that $P^{-1}AP = B$ then we say that A is similar to B.

Note that if $P^{-1}AP = B,$ then by taking $Q = P^{-1}$, we have $Q^{-1}BQ = A.$ Thus A is similar to B if and only B is similar to A and so we also use the terminology A and B are similar matrices instead of A is similar to B or B is similar to A. Similar matrices share many properties and perhaps that is the reason they are called similar. It is not difficult to prove that similar matrices have same determinant, trace, eigenvalues and rank. Also since matrix multiplication is associative, for every positive integer n

$$B^n = P^{-1}A^nP \text{ and } A^n = PB^nP^{-1}.$$

Now for a diagonal matrix, since the computation of its determinant, rank, powers and other such things, is a very simple task; we are often interested in knowing whether the given matrix is similar to a diagonal matrix or not. This leads us to the definition of diagonalizable matrix.

Definition: Diagonalizable Matrix

Let A be a square matrix. Then it is called a diagonalizable matrix if it is similar to a diagonal matrix or in other words if we can find an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix. The process of finding such P for the matrix A is called matrix diagonalization.

The above definition raises an immediate question. Is every square matrix diagonalizable?

The answer to this question is given by the following technical theorem.

Theorem: Let A be a square matrix of order n. Then A is diagonalizable if and only if it has n linearly independent eigenvectors.

Since we have seen in one of our earlier theorems that the eigenvectors corresponding to distinct eigenvalues are linearly independent we have the following nice result.

Theorem: Let A be a square matrix of order n. Then A is diagonalizable if it has n distinct real eigenvalues.

In view of the last theorem about symmetric matrix above, we can say that every symmetric matrix is similar to a diagonal matrix whose diagonal elements are its eigenvalues. Since the trace and determinant of a diagonal matrix are respectively the sum and product of its diagonal elements, it follows that the trace and determinant of every symmetric matrix is respectively the sum and product of its eigenvalues.

We now consider an example to understand the matrix diagonalization process.

Example 3: Diagonalize the matrix $A = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix}$.

Solution: The characteristic equation of the matrix A is

$$\begin{vmatrix} 3-\lambda & -1 & 1 \\ -1 & 5-\lambda & -1 \\ 1 & -1 & 3-\lambda \end{vmatrix} = 0.$$

Expanding the determinant and simplifying it gives

$$\lambda^3 - 11\lambda^2 + 36\lambda - 36 = 0$$

whose roots are $6,3$ and 2 and these are the eigenvalues of A. In order to diagonalize the matrix A, we have to determine the orthogonal matrix P whose 3 columns are the orthonormal eigenvectors corresponding to the eigenvalues $6,3$ and 2. For the eigenvector corresponding to the eigenvalue 6, we solve

$$(A-6I)x = \begin{bmatrix} -3 & -1 & 1 \\ -1 & -1 & -1 \\ 1 & -1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

It may be verified that one of the solution of this system of equations is $x_1 = 1, x_2 = -2$ and $x_3 = 1$. Thus $v_1 = (1,-2,1)$ is one of the eigenvector of A associated to the eigenvalue $6.$ Similarly it may be verified that $v_2 = (1,1,1)$ and $v_3 = (1,0,-1)$ are the eigenvectors of A associated to the eigenvalues 3 and 2

respectively. Therefore the orthogonal matrix P which diagonalizes A is given by

$$P = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{3} & 1/\sqrt{2} \\ -2/\sqrt{6} & 1/\sqrt{3} & 0 \\ 1/\sqrt{6} & 1/\sqrt{3} & -1/\sqrt{2} \end{bmatrix}.$$

It may now be verified that

$$\begin{aligned} &P^{-1}AP = P^{T}AP \\ &= \begin{bmatrix} 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{3} & 1/\sqrt{2} \\ -2/\sqrt{6} & 1/\sqrt{3} & 0 \\ 1/\sqrt{6} & 1/\sqrt{3} & -1/\sqrt{2} \end{bmatrix} \\ &= \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}. \end{aligned}$$

Let us now discuss the role of eigenvalues in the study of quadratic forms.

Quadratic Forms

Definition: Quadratic Forms

Consider a square matrix $A = [a_{ij}]$ of order n and a column vector $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ in $\mathbb{R}^n$. Then the quadratic form associated with the matrix $A = [a_{ij}]$ is an expression of the form

$$q(x) = x^T A x = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i x_j.$$

Remark: If A is not symmetric then by defining

$$a_{ij}^* = \frac{a_{ij} + a_{ji}}{2} = a_{ji}^*$$

we see that the quadratic form $q^*(x)$ associated with the symmetric matrix $A^* = [a_{ij}^*]$ is same as the quadratic form $q(x)$ associated with the non-symmetric matrix A. This is the reason we may confine our attention to the study of quadratic forms associated with the symmetric matrix only.

Example 4: Determine the quadratic form associated with the symmetric matrix $A = \begin{bmatrix} 3 & -1 & 2 \\ -1 & 1 & 4 \\ 2 & 4 & 5 \end{bmatrix}$.

Solution: Here

$$q(x) = [x_1 \quad x_2 \quad x_3] \begin{bmatrix} 3 & -1 & 2 \\ -1 & 1 & 4 \\ 2 & 4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= 3x_1^2 + x_2^2 + 5x_3^2 - 2x_1x_2 + 4x_1x_3 + 8x_2x_3.$$

We now introduce the terminology related to the sign of quadratic forms.

- If $q(x) = x^T A x > 0$ for all $x \neq 0$, then we say that the quadratic form $q(x)$ is positive definite and the matrix A is said to be a positive definite matrix.
- If $q(x) = x^T A x \geq 0$ for all x, then we say that the quadratic form $q(x)$ is positive semidefinite and the matrix A is said to be a positive semidefinite matrix.
- If $q(x) = x^T A x < 0$ for all $x \neq 0$, then we say that the quadratic form $q(x)$ is negative definite and the matrix A is said to be a negative definite matrix.
- If $q(x) = x^T A x \leq 0$ for all x, then we say that the quadratic form $q(x)$ is negative semidefinite and the matrix A is said to be a negative semidefinite matrix.

If a quadratic form takes positive as well as negative values then it is said to be indefinite. The following theorem gives a necessary and sufficient condition for determining the definiteness of a given quadratic form.

Theorem: Let $q(x)$ be a quadratic form associated with a symmetric matrix A. Then

1. $q(x)$ is positive definite if and only all the eigenvalues of A are positive.
2. $q(x)$ is positive semidefinite if and only all the eigenvalues of A are non-negative.
3. $q(x)$ is negative definite if and only all the eigenvalues of A are negative.
4. $q(x)$ is negative semidefinite if and only all the eigenvalues of A are non-positive.
5. $q(x)$ is indefinite if some of the eigenvalues of A are positive and some of them are negative.

We now see one application of quadratic forms.

Theorem: For a real-valued function $y = f(x_1, x_2, \ldots, x_n)$, if a point $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ is a point of local minimum or local maximum then $\nabla f(x^*) = 0$. Further, if $H = [f_{ij}(x^*)]$ is the Hessian matrix evaluated at the point x^*, then the point x^* is the point of local minimum or maximum accordingly as the quadratic form

$$d^{\,2}y = dx^T H dx$$

is positive or negative definite. On the other hand if the quadratic form $dx^T H dx$ is indefinite then x^* is neither the point of local minimum nor the point of local maximum.

Finally let us consider one example to understand this.

Example 5: Determine the points of local extremum for the function $y = f(x_1, x_2) = \frac{x_1^3}{3} + 3x_1^2 + x_1\, x_2 + \frac{x_2^2}{2} + 6x_2$.

Solution: Here $f_1 = \frac{\partial y}{\partial x_1} = x_1^2 + 6x_1 + x_2$ and $f_2 = \frac{\partial y}{\partial x_2} = x_1 + x_2 + 6$ and so $\nabla f(x_1, x_2) = (0,0)$ gives

$$x_1^2 + 6x_1 + x_2 = 0 \text{ and } x_1 + x_2 + 6 = 0$$

whose solutions are $(1, -7)$ and $(-6,0)$. Now the Hessian matrix

$$H = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} = \begin{bmatrix} 2x_1 + 6 & 1 \\ 1 & 1 \end{bmatrix}$$

and so H at the point $(1, -7)$ is the matrix $\begin{bmatrix} 8 & 1 \\ 1 & 1 \end{bmatrix}$. The eigenvalues of this matrix are roots of the equation

$$\begin{vmatrix} 8-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = 0.$$

But this is same as solving the quadratic equation $\lambda^2 - 9\lambda + 7 = 0$ whose roots are $\frac{9+\sqrt{53}}{2}$ and $\frac{9-\sqrt{53}}{2}$. Since both these roots are positive we conclude that H is positive definite at the point $(1, -7)$ and so $(1, -7)$ is the point of local minimum for the function f. It may be verified that the Hessian matrix at the point $(-6,0)$ has one positive and one negative eigenvalue and hence the point $(-6,0)$ is neither the point of local minimum nor the point of local maximum.

Summary

We introduced the concept of eigenvalues and eigenvectors through the eigenvalue problem and discussed the method of computing eigenvalues and eigenvectors of a square matrix. We showed how a symmetric matrix can be diagonalized through an orthogonal matrix which is constructed with the help of its orthonormal eigenvectors and discussed the importance of such

diagonalization process. In the final part of this module we tried to understand what quadratic forms are and how they are linked with the concept of eigenvalues while studying the nature of extreme points of a multivariate function.

THANK YOU.