Journal Homepage: - www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

RESEARCH ARTICLE

A NOTE ON EFFECT OF OFF-DIAGONAL ENTRIES IN A VARIANCE-COVARIANCE MATRIX ON EIGEN ANALYSIS.

Bushra Shamshad.
Department of Statistics, University of Karachi.

………………………………………………………………………………………………………......

Manuscript Info	*Abstract*
……………………..	………………………………………………………………………

A positive definite symmetric variance covariance matrix with non-zero diagonal entries- plays an important role in multivariate analysis. In this paper we will try to explain the basic algorithm needed to calculate the Eigen values and Eigen vector using such variance covariance matrix (or standardized variance covariance matrix). It is to be noted that coefficients of Principal Components are simply Eigen vectors and Eigen values are there variances (See, Anderson (1984), Johnson & Wichern (1988) & Hotelling (1936)). These eigen values are calculated from variance covariance matrix (or standardized variance covariance matrix). In this paper we will be discussing the effect on eigen analysis when entries in variance covariance matrix are changed. This eigen value problem will be studied on 2×2 and 3×3 variance covariance matrix, which then be carried out in higher order matrices (i.e. on 4×4) in order to generalize the coefficients of characteristic polynomial. With some proportion, eigen values vary, when off diagonal entries in variance covariance matrix are changed. This direct or indirect relationship between covariance terms of a variance covariance matrix and its respective eigen values then be deeply studied by regression analysis.

………………………………………………………………………………………………………......

Introduction:-

1.1. Eigenvalues And Eigenvectors

The eigenvalues, also called characteristic or latent roots of a square $p \times p$ matrix *A,* the *p* solutions for λ of the characteristic equation, given as

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

(1)

The roots can be complex-valued and some roots can have the same value, since λ is the root of polynomial of the degree *p*. Associated with each eigen value is a (right) eigenvector, which is a non-zero vector *x* with the property $Ax = \lambda$ x, and a left eigenvector, which is a non-zero vector *y* with the property $y'x = \lambda$ y'. If *x* is a (left or right) eigenvector *a x* is also an eigen vector for any non-zero scalar *a*, so eigen vectors are usually normalized to length 1, i.e. $x'x = 1$, (see, Hadely (1961), Anton (1994)) the two main properties of eigen values are

- The product of the eigenvalues of *A* is equal to |A|.
- The sum of the eigenvalues of *A* is equal to the trace of *A*.

Corresponding Author:- Bushra Shamshad.
Address**:-** Department of Statistics, University of Karachi.

Behavior of off diagonal entries in 2×2 matrix:
Consider a variance covariance matrix of order 2×2,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{2}$$

In order to obtain the eigenvalues and eigenvector we have to use the characteristic equation, that is,

$$(A - \lambda I) x = 0 \tag{3}$$

Where λ are the unknown constants or Eigen values which has to be determined and I is the identity matrix of order 2×2. After taking the determinant, we get

$$\lambda^2 - (\text{tr } A) \lambda + |A| = 0, \tag{4}$$

This is a second degree polynomial or the Second order characteristic polynomial of variance covariance matrix A. The behavior is obvious from the equation (4), when the off diagonal entries in a variance covariance matrix are changed, there will be no effect on the second part of equation (4), because it is the sum of the diagonal entries. Only the last term of polynomial equation of variance covariance matrix is effected by the change, in the off diagonal entries. For example, consider a symmetric variance covariance matrix of order 2×2,

$$A = \begin{bmatrix} 10 & 0 \\ 0 & 15 \end{bmatrix} \tag{5}$$

The corresponding characteristic polynomial of matrix A is

$$\lambda^2 - 25 \lambda + 150 = 0 \tag{6}$$

The eigen roots of the variance covariance matrix are $\lambda = \{15, 10\}$. These eigen values are exactly equal to the diagonal entries of the variance covariance matrix. The reason is that, in calculating the determinant (that is, the third term in equation (6)), we subtract the product of diagonal entries with the off diagonal entries. Here the off diagonal entries of the variance covariance matrix are zero. So, the negative part of the determinant is eliminated. That is,

$$[(10 - \lambda) (15 - \lambda) - 0] = 0$$

In other words, we are not subtracting any value form $(10 - \lambda) (15 - \lambda)$. Consider the same variance covariance matrix of order 2×2, with a change in off diagonal entries. That is, in place of zero covariance term we use covariance term equal to 1, the characteristic polynomial becomes

$$\lambda^2 - 25 \lambda + 149 = 0 \tag{7}$$

with the corresponding eigenvalues.

$$\lambda = \{15.1925, 9.8074\}$$

As there is a unit increase in the covariance term, the largest eigen values also increases with proportion 0.1925. This indicates that as the covariance term increases it accumulates with the maximum variation present in the data and this mount up of maximum variation is represented by highest eigen values. The non-zero off diagonal entries plays an important role. Whereas, both the properties (given in section (1)) are also verified. Consider variance covariance matrix A, given below. Where, 'z' indicates the covariance term.

$$A = \begin{bmatrix} 10 & z \\ z & 15 \end{bmatrix} \tag{8}$$

The behavior of eigen values are studied and presented in Table (1), for z = -11 to 11.
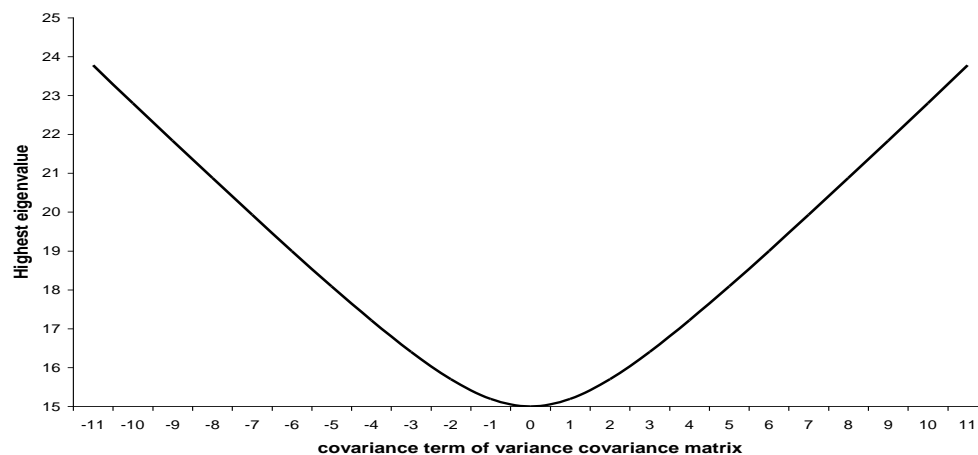
Table no: 1. Behavior of eigen values

Off diagonal Entry 'z'	determinant	λ_1	λ_2	increment	eigen vector
-11	29	23.78	1.219		$\begin{bmatrix} -0.6238 & .7815 \\ .7815 & .6238 \end{bmatrix}$
-10	50	22.8	2.192	-0.98	$\begin{bmatrix} -0.615 & 0.788 \\ 0.788 & 0.615 \end{bmatrix}$
-9	69	21.84	3.159	-0.96	$\begin{bmatrix} -0.605 & 0.7961 \\ 0.7961 & 0.605 \end{bmatrix}$
-8	96	20.88	4.1118	-0.96	$\begin{bmatrix} -0.592 & 0.805 \\ 0.805 & 0.592 \end{bmatrix}$
-7	101	19.93	5.06	-0.95	$\begin{bmatrix} -0.576 & 0.817 \\ 0.817 & 0.576 \end{bmatrix}$
-6	114	19	6	-0.93	$\begin{bmatrix} -0.554 & 0.832 \\ 0.832 & 0.554 \end{bmatrix}$
-5	125	18.09	6.9	-0.91	$\begin{bmatrix} -0.525 & 0.850 \\ 0.850 & 0.525 \end{bmatrix}$
-4	134	17.217	7.783	-0.873	$\begin{bmatrix} -0.4847 & 0.874 \\ 0.874 & 0.4847 \end{bmatrix}$
-3	141	16.4051	8.591	-0.8119	$\begin{bmatrix} -0.4241 & 0.9055 \\ 0.9055 & 0.4241 \end{bmatrix}$
-2	146	15.7016	9.2998	-0.7035	$\begin{bmatrix} -0.3310 & 0.943 \\ 0.943 & 0.3310 \end{bmatrix}$
-1	149	15.1926	9.8074	-0.509	$\begin{bmatrix} -0.189 & 0.9819 \\ 0.9819 & 0.189 \end{bmatrix}$
0	150	15	10	-0.1926	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
1	149	15.1926	9.8074	0.1926	$\begin{bmatrix} 0.189 & 0.9819 \\ 0.9819 & -0.189 \end{bmatrix}$
2	146	15.7016	9.2998	0.509	$\begin{bmatrix} 0.3310 & 0.943 \\ 0.943 & -.3310 \end{bmatrix}$
3	141	16.405	8.591	0.7034	$\begin{bmatrix} 0.4241 & 0.9055 \\ 0.9055 & -0.4241 \end{bmatrix}$
4	134	17.217	7.783	0.812	$\begin{bmatrix} 0.4847 & 0.874 \\ 0.874 & -0.4847 \end{bmatrix}$

5	125	18.0902	6.909	0.8732	$\begin{bmatrix} 0.525 & 0.850 \\ 0.850 & -0.525 \end{bmatrix}$
6	114	19	6	0.9098	$\begin{bmatrix} 0.554 & 0.832 \\ 0.832 & -0.554 \end{bmatrix}$
7	101	19.9336	5.067	0.9336	$\begin{bmatrix} 0.576 & 0.817 \\ 0.817 & -0.576 \end{bmatrix}$
8	96	20.88	4.1185	0.9464	$\begin{bmatrix} 0.592 & 0.805 \\ 0.805 & -0.592 \end{bmatrix}$
9	69	21.84	3.159	0.96	$\begin{bmatrix} 0.605 & 0.7961 \\ 0.7961 & -0.605 \end{bmatrix}$
10	50	22.8	2.192	0.96	$\begin{bmatrix} 0.615 & 0.788 \\ 0.788 & -0.615 \end{bmatrix}$
11	29	23.78	1.219	0.98	$\begin{bmatrix} 0.6238 & 0.7815 \\ 0.7815 & -0.623 \end{bmatrix}$

In table (1), the first column shows the z values, i.e. at the positions 12 and 21 of a variance covariance matrix (2.7), the eigenvalues are only affected by the determinant of matrix A. So, when the covariance term changes in matrix A, the determinant cause the change in the eigen values. It is clear that, when the off diagonal entries increase, determinant of a variance covariance matrix decreases, which further causes an increase in the magnitude of the largest eigen value, with some proportion.

Figure (1)

Figure: 1. Behavior of highest eigen value with an increase in off diagonal entries.

Another important point, that the eigen values are not affected by the sign of covariance term, can be seen in table (1), which is also presented in Figure (1). That is, the eigen values corresponds to covariance term equals to +1 is exactly the same as the eigen values corresponds to covariance term equal to -1. This is due to the fact that in a variance covariance matrix of order 2×2, the determinant is the value obtained by subtracting the product of two covariance terms from the product of two variances. At that point when two covariances are multiplied with each other the negative sign is eliminated and we get the same value of determinant as of the positive signed covariance term. The behavior of the negative and positive covariance's can be further explained by the Figure (1), i.e., the curve is symmetric about zero.
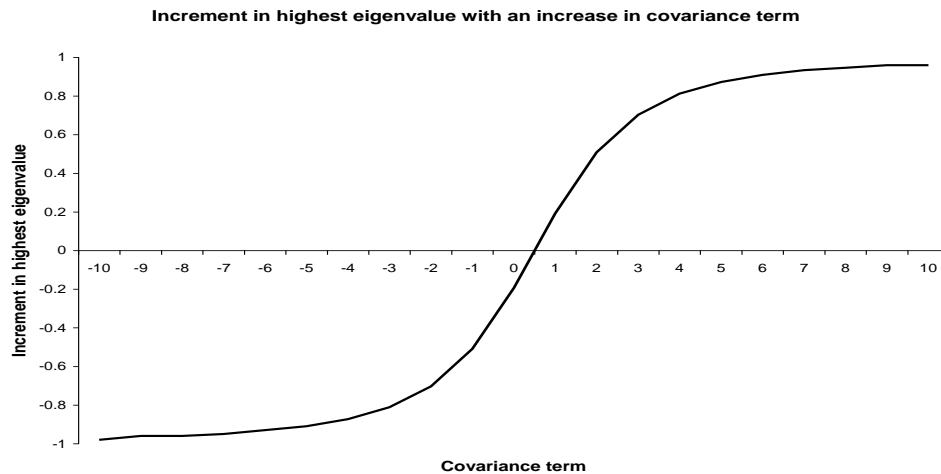
Figure: 2. Investment in highest eigenvalue with an increase in covariance term.

We may conclude that eigen values are not effected by the sign of the covariance term in matrix of order 2×2. But have a slight change in the eigenvector. The negative sign with the first entry of the first eigenvector corresponds to the largest eigen value, will now transferred to the second entry of the other eigenvector. Figure (2), shows the increment in the first eigenvalues, with an increase in covariance term, shown in second last column of the table (1). Plotting these increment against off diagonal entries of matrix A, we observe an 'S' shaped graph. For both negative and positive signed covariance term, the increment is similar but the direction is different. If we neglect the sign, we get a symmetric plot. We know that solution of equation (4), mainly depends on the third term. That is, the determinant of matrix A, which is directly depending on the covariance term in the variance covariance matrix. In spite of varying the covariance term in matrix A and observe the change in the eigen values, we constructed a table (2) in which we considered (unit) change in the determinant of a variance covariance matrix to observe the change in eigen values. The purpose is to take into account with what increment, change in eigen values occurs with a change in covariance term.

Table no: 2

Determinant	Highest eigen value (λ_1)	increment in λ_1	Covariance(x_1, x_2)	Increment in Cov(x_1, x_2)
149	15.19258	0	1	-
148	15.37228	0.1797	1.4142135	0.414214
147	15.5413	0.16902	1.732	0.317787
146	15.7015	0.1637	2	0.268
145	15.854	0.153	2.236	0.236
144	16	0.146	2.44	0.204
143	16.14005	0.14005	2.645	0.205
142	16.2749	0.13405	2.82	0.175
141	16.40512	0.13022	3	0.18
140	16.53112	0.126	3.1622	0.1622
139	16.6533	0.12218	3.3166	0.1544
138	16.772	0.1187	3.464	0.1474
137	16.887	0.115	3.605	0.141
136	17	0.113	3.741	0.136
135	17.109	0.109	3.872	0.131
134	17.216	0.107	4	0.128
133	17.3218	0.1058	4.123	0.123
132	17.424	0.1022	4.242	0.119
131	17.524	0.1	4.35	0.108
130	17.632	0.108	4.472	0.122

Considering increment in the highest eigen value as response variable (Y) and increment in covariance term as predictor variable (X), linear regression was applied on the data. That resulted in 79.5% of the coefficient of determination. Whereas, when quadratic regression is applied on the same variables, the coefficient of determination is 98.9%. The output of the second degree fit is given below.

$Y = 7.19E-03 + 0.941852X - 1.30122X**2$
R-Sq = 0.989

Analysis of Variance

SOURCE	DF	SS	MS	F	P
Regression	2	2.58E-02	1.29E-02	792.863	0.00
Error	17	2.77E-04	1.63E-05		
Total	19	2.61E-02			

SOURCE	DF	Seq SS	F	P
Linear	1	2.10E-02	74.6677	8.04E-08
Quadratic	1	4.80E-03	294.317	3.61E-12

The F-ratio for the quadratic regression is also very high as compare to the F value for linear model. Where as, the fitted plot shows almost all points are on the curve. Therefore, it may be concluded that there is a quadratic relationship between the increment in covariance term and the increment in the magnitude of highest eigen value.
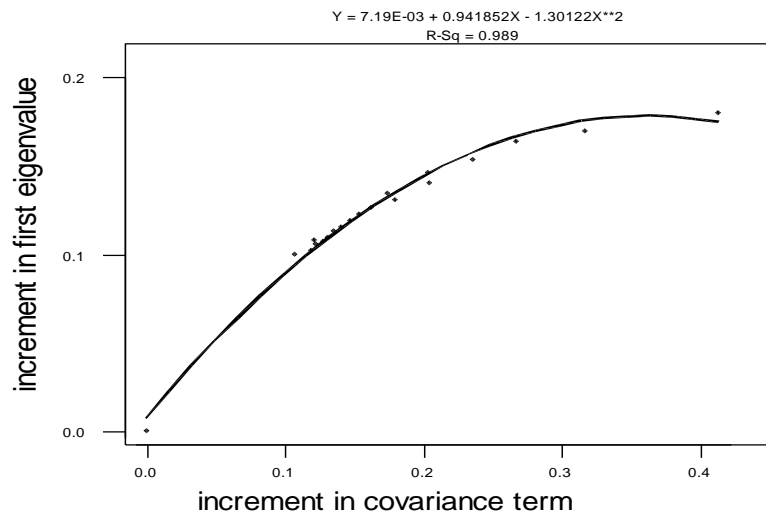

Figure: 3. Fitted regression line on the original covariance terms

Behavior Of Eigen Value In Variance Covariance Matrix Of Order 3×3
In the previous section we studied the behavior of variance covariance matrix of order 2×2, and observe the effects on the eigen roots, for different values of covariance term. In variance covariance matrix of order 3×3, it is quite complicated to observe the change in the eigen values, as the covariance terms of a variance covariance matrix are changed. Consider a variance covariance matrix of order 3×3, in its general form;

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \qquad (9)$$

This variance covariance matrix can be written in the form of characteristic equation given in equation (3), given as

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = 0 \qquad (10)$$

$$\{(a_{11} - \lambda)[(a_{22} - \lambda)(a_{33} - \lambda) - a_{23}a_{32}] - a_{12}[a_{21}(a_{33} - \lambda) - a_{31}a_{23}] + a_{13}[a_{21}a_{32} - a_{31}(a_{22} - \lambda)]\} = 0$$

$$\lambda^3 - (trA)\lambda^2 + (a_{11}a_{22} + a_{11}a_{33} + a_{22}a_{33} - a_{12}a_{21} - a_{13}a_{31} - a_{23}a_{32})\lambda - (\det A) = 0 \quad (11)$$

From equation (11), the coefficient of λ^3 is unity, where as the coefficient of λ^2 is the trace of variance covariance matrix (9). The coefficient of λ can be partitioned into two parts, one containing the three combination of pairs of diagonal entries and the other part containing the negative effect of the cross product term of the off diagonal entries (that is, product of entries at position 12 and 21, at position 13 and 31 and so on). Hence, in general we can say that we are subtracting the effect of product of covariances form the possible pairs of product of diagonal element.

Here the point should be noted that the negative sign in the covariance term does not affect the coefficient of λ. So, whether the value of covariance terms are negative or positive, the coefficient of λ remains the same. But this negative sign of covariance terms affect the constant term in the equation (11). It is also applicable when standardized variance covariance matrix (i.e. correlation matrix) is used for eigen analysis. Consider the following correlation matrix, (Everitt and Dunn, 1991), of order 3×3.

$$A = \begin{bmatrix} 1 & 0.83 & 0.78 \\ 0.83 & 1 & 0.67 \\ 0.78 & 0.67 & 1 \end{bmatrix} \qquad (12)$$

Characteristic equation for a standardized variance covariance matrix is same as that of a variance covariance matrix. From equation (11), we can easily write the characteristic polynomial of order 3×3.

$$\lambda^3 - 3\lambda^2 + [(1 - (0.83)^2) + (1 - (0.78)^2) + (1 - (0.67)^2)]\lambda - (0.121316) = 0 \quad (13)$$

with the following roots. $\lambda_1 = 2.521912$, $\lambda_2 = 0.334108$ and $\lambda_3 = 0.143979$

Characteristic polynomial for a correlation matrix can be more easily explained as compare to variance covariance matrix. From equation (13), it is clear that coefficients of λ^3 and λ^2, of any standardized variance covariance matrix of order 3×3, will always be the same as in (13), i.e., 1 and 3 respectively. Whereas, the coefficient of λ is the sum of coefficients of non-determination and the last term (i.e. the constant term) will always be the determinant of the standardized variance covariance matrix.

The behavior of eigen values for higher order of matrix A is similar to that of variance covariance matrix of order 2×2. That is, the largest eigenvalue is formed by accumulating some proportion of covariance into the variance of the variable containing the maximum variation. Thus the largest eigenvalue explain the maximum variation of the data set. In case of variance covariance matrix (or standardized variance covariance matrix) of order 3×3 the behavior is similar but quite complicated. If we carry out eigen analysis of correlation matrix (12), by changing the correlation at different positions one by one, we can observe the behavior of eigenvalues, as given in Table (3).

First column in Table (3), shows the change (by 0.1) in the value of correlation, starting from -1. Whereas, rest of the columns represent the highest eigen values, when correlation at different position are being swapped. For example, the second column entitled "at position 13", "at position 12" and "at position 23", represent the largest eigen value (i.e. 2.02711, 2.01235 and 2.00351 respectively), when the entity at position 13 in a correlation matrix has been replaced by -1 and so on.

Table no: 3 "Eigen Values of Matrix A in equation (12)"

Correlation	at position 13	at position 12	at position 23
-1	2.02711	2.01235	2.00351
-0.9	1.9391	1.91782	1.90661
-0.8	1.86674	1.83217	1.83245
-0.7	1.83169	1.78247	1.84655
-0.6	1.83694	1.78891	1.88011
-0.5	1.86162	1.81718	1.91739
-0.4	1.89476	1.85261	1.95719
-0.3	1.93279	1.89208	1.99929
-0.2	1.97444	1.9347	2.04363
-0.1	2.01915	1.98013	2.0902
0	2.06668	2.02825	2.13899
0.1	2.11686	2.07897	2.19
0.2	2.16961	2.13222	2.24322
0.3	2.22483	2.18796	2.29862
0.4	2.28246	2.24612	2.35617
0.5	2.34241	2.30662	2.41582
0.6	2.4046	2.3694	2.47753
0.7	2.46895	2.43436	2.54123
0.8	2.53535	2.50141	2.60684
0.9	2.60373	2.57044	2.67431
1	2.67398	2.64137	2.74354

Figure (4) indicates the similar pattern for any position (say, 12, 13 or 23) in a correlation matrix. It is also clear that, when a positive correlation (say, at position 13) is replaced by a strong negative (say -1 to -0.7), there will be a decline in the magnitude of largest eigen value. As, it move towards moderate to week negative correlation, the other two positive correlations (i.e., at 12 and 23) pulls the magnitude of largest eigen value, and thus increase its magnitude. If we take the correlation matrix similar to an identity matrix and then we change the correlation at different position one by one, the behavior of eigen values at either position will exactly the same, making a V-shaped graph that is symmetric at correlation equals to zero.
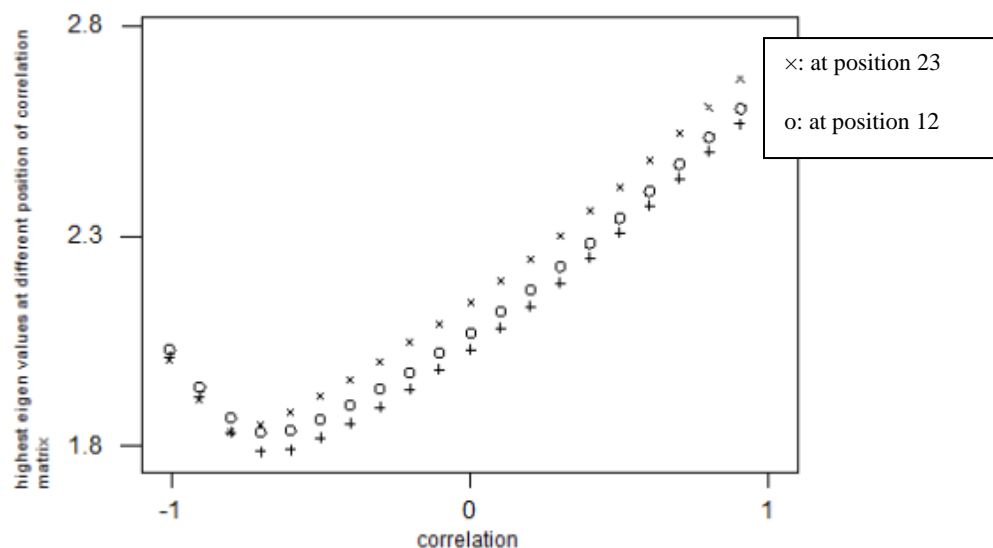

Figure: 4. Behavior of highest eigen values with a change in correlation at position 12, 13 and 23.

Behavior of eigen value in higher order variance-covariance matrix.
In previous sections we have deeply observe the behavior of eigen values with variance covariance matrix (or standardized variance covariance matrix) of order 2×2 and 3×3. The characteristic polynomial, given in equation (15), of a variance covariance matrix, given in (14), of order 4×4 can be generalized for higher order matrices.

$$
\begin{bmatrix}
a_{11} & a_{12} & a_{13} & a_{14} \\
a_{21} & a_{22} & a_{23} & a_{24} \\
a_{31} & a_{32} & a_{33} & a_{34} \\
a_{41} & a_{42} & a_{43} & a_{44}
\end{bmatrix}
\tag{14}
$$

$$
\lambda^4 - (trA)\lambda^3 + (a_{11}a_{22} + a_{11}a_{33} + a_{11}a_{44} + a_{22}a_{33} + a_{22}a_{44} + a_{33}a_{44}
$$
$$
- a_{12}a_{21} - a_{13}a_{31} - a_{14}a_{41} - a_{23}a_{32} - a_{24}a_{42} - a_{34}a_{43})\lambda^2 + (-a_{11}a_{22}a_{33} - a_{11}a_{22}a_{44} - a_{11}a_{33}a_{44} - a_{22}a_{33}a_{44}
$$
$$
+ a_{11}a_{23}a_{32} + a_{11}a_{24}a_{42} + a_{11}a_{34}a_{43} + a_{22}a_{13}a_{31} + a_{22}a_{14}a_{41} + a_{22}a_{34}a_{43} + a_{33}a_{12}a_{21} + a_{33}a_{14}a_{41} + a_{33}a_{24}a_{42}
$$
$$
+ a_{44}a_{12}a_{21} + a_{44}a_{13}a_{31} + a_{44}a_{32}a_{23} - a_{12}a_{24}a_{41} - a_{13}a_{34}a_{41} - a_{42}a_{21}a_{14} - a_{23}a_{34}a_{42} - a_{14}a_{43}a_{31} - a_{24}a_{43}a_{32}
$$
$$
- a_{12}a_{23}a_{31} - a_{13}a_{32}a_{21})\lambda - (detA) = 0
\tag{4.2}
$$

The pattern is clear from equation (15), that the coefficient of the highest power of λ (i.e. λ^4) will always be unity. With the second highest power of λ, (i.e. λ^3), the coefficient will be the trace of the corresponding matrix. The coefficient of third highest power of λ will contain possible pairs of variances and a negative effect of their squared covariances. The coefficients of the next higher power of λ will be having possible pairs of entities, of size three, of a variance covariance matrix. Where as, the last term will always be the determinant of the corresponding matrix (having pairs of entries of size 4). In general we can write the characteristic polynomial as,

$$
\lambda^4 - \sum_{i=1}^{4} a_{ii}\lambda^3 + \left(\sum_{i \neq j} \sum a_{ii}a_{jj} - \sum_{i \neq j} \sum a_{ij}a_{ji} \right)\lambda^2 - \left(\sum_{i \neq j \neq k} \sum \sum a_{ii}a_{jj}a_{kk} \right.
$$
$$
\left. - \sum_{i \neq j \neq k} \sum \sum a_{ii}a_{jk}a_{kj} + \sum_{i \neq j \neq k} \sum \sum a_{ij}a_{jk}a_{ki} \right)\lambda - (detA) = 0
$$

Conclusion:-

From the above analysis we can conclude that, there is a direct relationship between covariance term of a variance covariance matrix and its respective eigen values. Now, one can understand the phenomenon of eigen value problem, which is the base of principal component analysis (See, Anderson (1984), Johnson & Wichern (1988) & Hotelling (1936)). The characteristic polynomial of order n can be given as,

$$
= \lambda^n - \sum_{i=1}^{4} a_{ii}\lambda^{n-1} + \left(\sum_{i \neq j} \sum a_{ii}a_{jj} - \sum_{i \neq j} \sum a_{ij}a_{ji} \right)\lambda^{n-2} - \left(\sum_{i \neq j \neq k} \sum \sum a_{ii}a_{jj}a_{kk} - \sum_{i \neq j \neq k} \sum \sum a_{ii}a_{jk}a_{kj} \right.
$$
$$
\left. + \sum_{i \neq j \neq k} \sum \sum a_{ij}a_{jk}a_{ki} \right)\lambda^{n-3} + \dots - (detA)\lambda^0 = 0
\tag{16}
$$

Using (16) the behavior of eigen values and eigen vectors can be studied for any value of n.

References:-

1. Anderson, T. W. (1984). An Introduction to multivariate statistical Analysis. 2nd Edition, John Wiley, New York.
2. Anton, H. (1994). Elementary Linear Algebra. John Wiley and Sons, United States of America.
3. Everitt. B. S. and Dunn. G. (1991). Applied Multivariate Data Analysis. Edward Arnold, London, pp: 244.
4. Hadely. (1961): Linear Algebra. 7th Printing, Addison Wesley Publishing Company, pp: 236-249.
5. Hotelling, H. (1936): Simplified Calculation of Principal Components. Psycometrika, Vol: 1, pp: 27-35.
6. Johnson, R. A. and Wichern, D. W., (1988): Applied Multivariate Statistical Analysis", Prentice-Hall, New Jersey, pp: 340-366