Isparta Uygulamalı Bilimler Üniversitesi
Ziraat Fakültesi Dergisi
ISSN 1304-9984

Research Article/Araştırma Makalesi

Ziraat Fakültesi Dergisi,
Cilt 18, Sayı 2,
Sayfa 134-142, 2023

Journal of the Faculty of Agriculture
Volume 18, Issue 2,
Page 134-142, 2023

Factor Rotation Methods in Factor Analysis: An Application on Agricultural Data

Ayşe Sümeyye CAN[1], Özgür KOŞKAN[1], Malik ERGİN*[1]

Abstract: In this study, the rotation stage of factor analysis, which is one of the multivariate analysis methods, was examined. All stages of factor analysis have been defined. The material of the study consisted of a data set obtained from barley planted in 20 plots (replication) having 9 variables. In each plot, the average of 6 plants selected from that plot was used. The variables emphasized in the study were plant height, number of leaves, spike length, spike weight, grain yield, flowering period (days), harvest index, yield, and 1000-grain weight. Factors were obtained by principal component analysis, which is a factor extraction method, from the data set that met the prerequisites of the analysis. The criteria used in different factor rotations are given and based on these criteria, the formula that gives the optimum rotation angle for each data set was obtained. As a result, the formulas obtained for orthomax, varimax, quartimax, and equamax were applied to the factors obtained from the data set and the results were interpreted. As a result of factor rotation, when varimax, quartimax, and equamax methods were used, the values of the variables in terms of factor loads differed in each factor. This is a desirable situation for factor analysis results.

Keywords: Equamax, factor analysis, orthomax, rotation methods, varimax, quartimax

Faktör Analizinde Faktör Döndürme Yöntemleri: Ziraat Verisi Üzerinde Bir Uygulama

Öz: Bu çalışmada, çok değişkenli analiz yöntemlerinden biri olan faktör analizinde döndürme aşaması incelenmiştir. Faktör analizinin tüm aşamaları tanımlanmıştır. Çalışmanın materyali, 9 değişken içeren 20 parsele (tekrarlama) ekilen arpadan elde edilen bir veri setinden oluşmaktadır. Her parselde, o parselde seçilen 6 bitkinin ortalaması kullanılmıştır. Çalışmada vurgulanan değişkenler bitki boyu, yaprak sayısı, başak uzunluğu, başak ağırlığı, tane verimi, çiçeklenme periyodu (gün), hasat indeksi, verim ve 1000 tane ağırlığıdır. Analizin ön koşullarını sağlayan veri setinden, faktör çıkarma yöntemi olan temel bileşen analizi ile faktörler elde edilmiştir. Farklı faktör döndürmeler için kullanılan kriterler verilmiş ve bu kriterlere dayanarak her veri seti için en uygun döndürme açısını veren formül elde edilmiştir. Sonuç olarak, orthomax, varimax, quartimax ve equamax için elde edilen formüller veri setinden elde edilen faktörlere uygulanmış ve sonuçlar yorumlanmıştır. Faktör döndürme sonucunda varimax, quartimax ve equamax yöntemleri kullanıldığında değişkenlerin faktör yükleri açısından değerleri her faktörde farklılık göstermiştir. Bu, faktör analizi sonuçları için arzu edilen bir durumdur.

Anahtar Kelimeler: Equamax, faktör analizi, orthomax, döndürme yöntemleri, varimax, quartimax

***Sorumlu yazar (Corresponding author)**
malikergin@isparta.edu.tr

Alınış (Received): 02/10/2023
Kabul (Accepted): 22/11/2023

[1]Isparta Uygulamalı Bilimler Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü, Isparta, Türkiye.

1. Introduction

The purpose of factor analysis is to explain the underlying structure of a multivariate data matrix. As it is a stand-alone analysis, it can also serve as a precursor to many multivariate analysis techniques. It is a set of methods that allows explaining the structure explained with the data obtained from related p variables, with a smaller number

DOI: 10.54975/isubuzfd.1370165

of k variables that are not related. The new variables obtained by this method are also called factors/components. Thus, the variation explained by a large number of variables can be seen whether it can be explained by a smaller number of variables. In summary, the two main purposes of factor analysis are to reduce the variable (size) and to investigate the relation between the variables, in other words, to classify the variables (Alpar, 2013). The steps of analysis are; investigation of the suitability of the data set, factor creation stage, deciding factor numbers, factor rotation, interpretation of results, and naming factors. In order for factor analysis to be used, the data set should be in a structure that the Pearson correlation coefficient can be applied. At the same time, one of the most important requirements is that the correlations should be large enough to enable factorization, and the correlation coefficient range is accepted as between 0.30 and-0.90 (Didia and Idenedo, 2021). Besides the correlation coefficients, another element to look at is the partial correlation coefficient. Partial correlation is the correlation coefficient between two variables after the effect of other variables has been held out on both other variables (Brown and Hendrix, 2005).

Therefore, it is not recommended to apply factor analysis if the partial correlation is high. The relationship between the variables should also be sufficient for factorization. One way to examine this is to look at the determinant of the correlation matrix. The closer this value is to zero, the greater the dependency is. In addition to these, the other most widely used criterion is the Kaiser-Mayer-Olkin (KMO) measure (Khalaf, 2007). The KMO value is a measure of how well data is suited for factor analysis. It also indicates the suitability of the sample size. The KMO value measures the sampling adequacy for each variable in the model and the sampling adequacy for the overall model (Shrestha, 2021). The result obtained with KMO and the sample adequacy is interpreted with the help of information given in Table 1.

Table 1. Sampling adequacy according to KMO value (Alpar, 2013).

KMO	Sampling adequacy
0.90-1.00	*Excellent*
0.80-0.89	Very well
0.70-0.79	Well
0.60-0.69	Mediocre
0.50-0.59	Poor
<0.50	Unacceptable

The most basic step in factor analysis is factor creation. This step involves different but also related techniques. These techniques are: principal component analysis, basic axis factor, unweighted least squares, generalized least squares, maximum likelihood, alpha factor, and image factor (Süzülmüş, 2005). Of these techniques, the most commonly used one is principal component analysis. This method tries to summarize the structure of the secondary data matrix derived from the original data matrix. Secondary data matrix can be a variance-covariance matrix or correlation matrix. Thus, the total variance, which is equal to the number of variables, is explained. In this method, the first factor is calculated to explain the maximum variance between the variables, and the second factor is calculated to explain the maximum remaining variance. This situation is repeated for each factor. The important thing that should be considered is that the factors obtained as a result of the analysis should be orthogonal. The factor load matrix provides an answer to the question of how much each variable contributes to which factor; in other words, the factor loads in each factor are a measure of each variable's contribution to that factor. Similarly, the correlation coefficient between the related variable and the related factor is the factor load. The loads (weights) here are between -1 and +1 since they are correlation coefficients and can be classified as.

- 0.30-0.40 acceptable
- 0.50-0.70 reasonable to use
- >0.70 loads that explain the factor well

Another data obtained here is the explained variance, in another word, the eigenvalues. These eigenvalues are equal to the sum of the squares of the factor loads in each factor. When a single factor load is squared, the variance of that variable explained by the relevant factor is obtained. In addition, the sum of the multiplication of loads of any two variables gives the correlation coefficient between the two variables. After the factors are established, an important step is to decide which factors to consider. Some methods are used to decide which factors to consider or how many factors to choose (Khalaf, 2007). These methods are:

Eigenvalue criterion: It is based on the principle that only factors with an eigenvalue greater than 1 are taken into account.

Explained variance: after eigenvalues are found, the smallest m value for which $\left(\sum_{j=1}^{m} \lambda_i / p\right) \geq 2/3$, the condition is met is determined as the number of significant principal components.

Scree plot approach: A plot is drawn with the factor number on the x-axis and the eigenvalue of the relevant factor on the y-axis. In this plot, the number of factors up to the factor where the slope decreases steeply is taken into account. Jolliffe criterion: factors with an eigenvalue of 0.7 and higher are taken into account.

Percentage of total variance: factors are considered until the contribution of each additional factor to the explained total variance falls below 5%. In the literature, there are

 DOI: 10.54975/isubuzfd.1370165

many studies that used factor analysis in various disciplines. For instance, Sadek et al. (2006) used factor analysis with promax rotation for each gender to derive fewer independent common factors in Arabian horses. They found that three key factors were extracted which explained a significant portion of the overall variation in mares and stallions, accounting for 66% and 67%, respectively. Beniston et al. (2014) developed a soil quality index using factor analysis.

The main focus of this study is factor rotations. The factors obtained by factor analysis method can sometimes be difficult to interpret and have a complex structure. In this case, in order to provide clarity and independence in interpretation, axis rotation increases the load of the variables on one factor and decreases the loads on the other factors. As a result, variables that are predominantly effective in each factor are determined (Polat, 2012).

The rotated factor load matrix is the final result of the factor analysis. The factor with a high load on a variable is closely related to that variable. For instance, a common title can be given to variables A, B, and C with the greatest weight under the first factor. Likewise, a naming can be made for the 2nd and 3rd factors.

The purpose of this study was to show the calculation steps of factor rotation methods on a data set obtained from barley planted in 20 plots (replication) having 9 variables and to compare the results of the methods with each other.

2. Material and Method

2.1. Materials

The material of the study consisted of a data set obtained from barley planted in 20 plots (replication) having 9 variables. In each plot, the average of 6 plants selected from that plot was used. The variables emphasized in the study were plant height (PH), number of leaves (NL), spike length (SL), spike weight (SW), grain yield (GY), flowering period (days) (FP), harvest index (HI), yield (Y), and 1000-grain weight (1000-GW).

2.2. Methods

2.2.1. Rotation of factor loads matrix

Researchers may choose to rotate an axis to provide "independency of variables, clarity in interpretation, and significance" to the important factors they obtained.

As a result of rotating the axis, the load of the variables on one factor will increase while the load on the other factors will decrease so that there will be predominantly effective variables in each factor and the factors can be interpreted more easily (Polat, 2012; İlhan, 2007).

This rotation process can be explained using the graphical representation of a two-factor structure in Figure 1. Here, the axes are the factor and the coordinates of the variables are the load value of that factor.

2.2.2. Vertical rotation methods

In this set of methods, rotation is done by finding an optimum angle according to criteria. Since both axes rotate in the same direction at the same angle, orthogonality is preserved, so they are called orthogonal rotation methods. The differentiation of these methods is due to the accepted criteria when deciding on the optimum angle. The criteria of the methods, in other words, the functions that obtable to give the maximum value after the rotation process *a* load, are as in Table 2 (Finch, 2011; Kaiser, 1958).

Table 2. Criteria of the methods

Orthomax	$\sum\left(\sum(a^4) - \gamma\left(\sum a^2\right)^2\right)$
Varimax	$\sum\left(\left\{n\sum(a^4) - \left(\sum a^2\right)^2\right\}/n^2\right)$
Quartimax	$\sum\left(\sum a^4\right)$
Equamax	$\sum\left(\sum a^4 - \frac{k}{2}\left(\sum a^2\right)^2\right)$

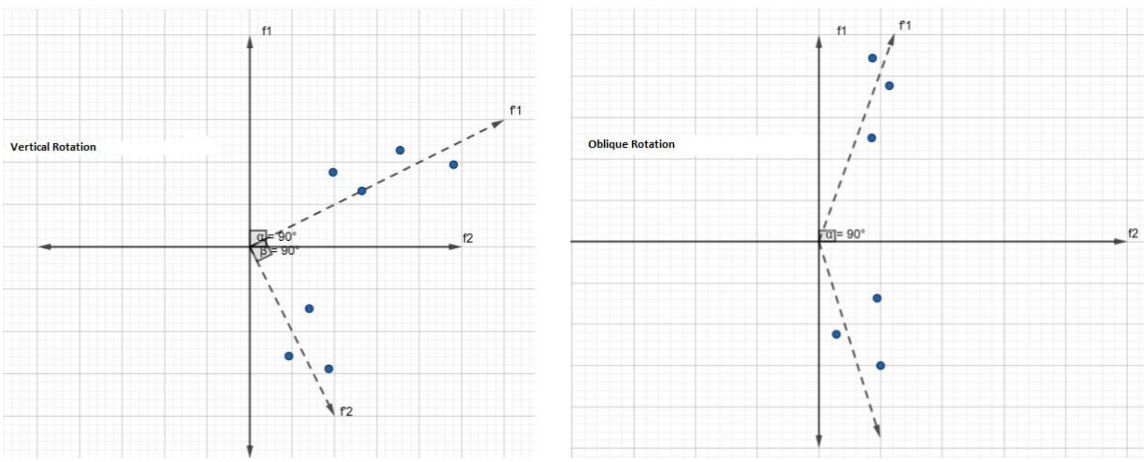

Figure 1. Vertical and oblique rotation method.

DOI: 10.54975/isubuzfd.1370165

When the functions are examined, the difference and similarity between the criteria show that the orthomax criterion is a general form depending on γ, it gives the criteria of quartimax when γ=0, varimax when γ=1, and equamax when γ=k/2 (Browne, 2001).

These functions, each developed by different researchers, are called orthogonal rotation criteria. The rotation process is repeated many times for the entire two-factor combination. When rotating both factors, the angle of rotation that maximizes these criteria is calculated separately. The functions determined for the criteria are functions that depend on the factor loads. How to reach a rotation angle from these criteria is explained in general and then the formulas for each criterion are given below.

x_i: The factor load of the 1st factor in the ith row, which has not yet been rotated,
y_i: The factor load of the 2nd factor in the ith row, which has not yet been rotated,
X_i: The factor load of the 1st factor in the ith row, which has been rotated,
Y_i: The factor load of the 2nd factor in the ith row, which has been rotated.

The relationship between them can be expressed mathematically (Kaiser, 1958) in Equation (1):

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} \times \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} = \begin{bmatrix} X_1 & Y_1 \\ X_2 & Y_2 \\ \vdots & \vdots \\ X_n & Y_n \end{bmatrix} \tag{1}$$

where, $\begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$ is a rotation matrix with an angle of ϴ

So from this matrix multiplication, the following formulas are obtained in Equation (2);

$$\begin{aligned} X_i &= x_i cos\theta + y_i \sin\theta \\ Y_i &= -x_i sin\theta + y_i cos\theta \end{aligned} \tag{2}$$

From here also differentiation formulas with respect to θ gives the equation (3);

$$\begin{aligned} dX_i &= Y_i \\ dY_i &= -X_i \end{aligned} \tag{3}$$

In the orthogonal rotation criteria,

$$\text{instead of } a_{1i}\, X_i = x_i cos\theta + y_i \sin\theta$$
$$\text{instead of } a_{2i} Y_i = -x_i sin\theta + y_i cos\theta$$

is written. Since these rotation criteria specify a function, if it is differentiated with respect to θ and set to zero, the value of θ, that is, the rotation angle, will be the angle value that makes our criterion maximum or minimum.

2.2.3. Orthomax

The optimum θ value satisfying the Orthomax criterion is:

$$\varrho = n.\sum (X^2)^2 - \gamma\left(\sum X^2\right)^2 + n.\sum (Y^2)^2 - \gamma\left(\sum Y^2\right)^2$$

The derivative of the expression can be taken by using equation (2.2.) in the derivative operation.

$$n.\left(\sum 4X_i^3 Y_i\right) - 2\gamma\left(\sum X_i^2\right).\left(\sum 2X_i Y_i\right) + n.\left(\sum -4Y_i^3 X_i\right) - 2.\gamma\left(\sum Y_i^2\right).\left(\sum -2Y_i X_i\right)$$

Let set the expression equaling to zero, and put the common multipliers in parentheses:

$$4.n.\left(\sum X_i^3 Y_i\right) - 4\gamma\left(\sum X_i^2\right).\left(\sum X_i Y_i\right) + 4.n.\left(\sum -Y_i^3 X_i\right) - 4.\gamma\left(\sum Y_i^2\right).\left(\sum -Y_i X_i\right) = 0$$
$$n.\sum XY(X^2 - Y^2) - \gamma\sum XY \sum (X^2 - Y^2) = 0$$

This equation is written in place of equation (2).

$$n.\sum \left((x_i cos\theta + y_i sin\theta).(-x_i sin\theta + y_i cos\theta).((x_i cos\theta + y_i sin\theta)^2) - (-x_i sin\theta + y_i cos\theta)^2\right) =$$
$$\gamma.\left(\sum (x_i cos\theta + y_i sin\theta).(-x_i sin\theta + y_i cos\theta)\right).\left(\sum ((x_i cos\theta + y_i sin\theta)^2 - (-x_i sin\theta + y_i cos\theta)^2)\right)$$

If the right and left sides of this equation are arranged separately. For the left side of the equation;

$$n.\sum\left(-x_i^2 cos\theta.sin\theta + x_i y_i cos^2\theta - x_i y_i sin^2\theta + y_i^2 sin\theta.cos\theta\right).\left(x_i^2 cos^2\theta + y_i^2 sin^2\theta + 2x_i y_i cos\theta\, sin\theta - x_i^2 \sin^2\theta + 2x_i y_i cos\theta\, sin\theta - y_i^2 \cos^2\theta\right)$$

$$n.\sum\left(\left(\frac{sin2\theta}{2}(y_i^2 - x_i^2) + x_i y_i cos2\theta\right).\left(2x_i y_i sin2\theta + (x_i^2 - y_i^2)cos2\theta\right)\right) =$$
$$n.\sum\left(x_i y_i(y_i^2 - x_i^2)sin^2 2\theta + sin4\theta\,\frac{(x_i^2 - y_i^2)^2}{-4} + x_i^2 y_i^2 sin4\theta + cos^2 2\theta\, x_i y_i(x_i^2 - y_i^2)\right) = n\sum\left(x_i y_i(x_i^2 - y_i^2).cos4\theta + sin4\theta\left(x_i^2 y_i^2 - \frac{(x_i^2 - y_i^2)^2}{4}\right)\right)$$

results are obtained. Now let's arrange the right side of the equation:

$$\gamma.\left(\sum\left(-x_i^2\frac{sin2\theta}{2} + x_i y_i cos2\theta + y_i^2\frac{sin2\theta}{2}\right)\right).\left(\sum\left(cos^2\theta(x_i^2 - y_i^2) + sin^2\theta(y_i^2 - x_i^2) + 2x_i y_i sin2\theta\right)\right) = \gamma.\left(\sum\left(\frac{sin2\theta}{2}(y_i^2 - x_i^2)\right) + \sum(x_i y_i cos2\theta)\right).\left(\sum\left((x_i^2 - y_i^2).cos2\theta\right) + \sum(2x_i y_i sin2\theta)\right) = \gamma\left(\frac{sin4\theta}{4}\left(\left(\sum(y_i^2 - x_i^2)\right)\left(\sum(x_i^2 - y_i^2)\right)\right) + cos4\theta\left(\left(\sum(x_i^2 - y_i^2)\right)\left(\sum x_i y_i\right)\right) + sin4\theta\left(\left(\sum x_i y_i\right)^2\right)\right) = \gamma\left(\frac{sin4\theta}{4}\left[\left(4\left(\sum x_i y_i\right)^2\right) - \left(\sum(x_i^2 - y_i^2)\right)^2\right] + cos4\theta\left(\left(\sum(x_i^2 - y_i^2)\right)\left(\sum x_i y_i\right)\right)\right)$$

results are obtained. Now let's equate the right and left sides and arrange them in Equation (4):

$$n\sum\left(x_i y_i(x_i^2 - y_i^2).cos4\theta + sin4\theta\left(x_i^2 y_i^2 - \frac{(x_i^2 - y_i^2)^2}{4}\right)\right) = \gamma\left(\frac{sin4\theta}{4}\left[\left(4\left(\sum x_i y_i\right)^2\right) - \left(\sum(x_i^2 - y_i^2)\right)^2\right] + cos4\theta\left(\left(\sum(x_i^2 - y_i^2)\right)\left(\sum x_i y_i\right)\right)\right)$$

$$sin4\theta\left[\left(n.\sum x_i^2 y_i^2 - \frac{(x_i^2 - y_i^2)^2}{4}\right) - \frac{\gamma}{4}\left(4\left(\sum x_i y_i\right)^2 - \left(\sum(x_i^2 - y_i^2)\right)^2\right)\right] = cos4\theta\left[\gamma\left(\sum(x_i^2 - y_i^2)\right)\left(\sum x_i y_i\right) - n\sum x_i y_i.(x_i^2 - y_i^2)\right]$$

DOI: 10.54975/isubuzfd.1370165

$$tan4\theta = \frac{\left[\gamma\left(\sum\left(x_i^2 - y_i^2\right)\right)\left(\sum x_i y_i\right) - n\sum x_i y_i.\left(x_i^2 - y_i^2\right)\right]}{\left[\left(n.\sum x_i^2 y_i^2 - \frac{(x_i^2 - y_i^2)^2}{4}\right) - \frac{\gamma}{4}\left(4\left(\sum x_i y_i\right)^2 - \left(\sum\left(x_i^2 - y_i^2\right)\right)^2\right)\right]}$$

$$\theta = \frac{1}{4}arctan\frac{2\left[n\sum\left(x_i^2-y_i^2\right).\left(2x_iy_i\right)-\gamma\sum\left(x_i^2-y_i^2\right).\sum\left(2x_iy_i\right)\right]}{n.\left\{\sum\left(\left(x_i^2-y_i^2\right)^2-\left(2x_iy_i\right)^2\right)\right\}-\gamma\left\{\left(\sum\left(x_i^2-y_i^2\right)\right)^2-\left(\sum 2x_iy_i\right)^2\right\}} \quad (4)$$

The Equation (4) that gives an optimum angle value for the Orthomax criterion is as above. As can be seen in Table 2, the orthomax criterion is accepted as a general form. The values that can be substituted for the parameter γ here allow us to reach other orthogonal criteria

2.2.4. Varimax

As seen in Table 2, the varimax criterion is obtained when γ=1 in the orthomax criterion (Kaiser, 1958). In this case, if γ=1 is written in the formula (6) for θ, the formula that gives the optimum angle θ for varimax was explained in Equation (5):

$$\theta = \frac{1}{4}arctan\frac{2\left[n\sum\left(x_i^2-y_i^2\right).(2x_iy_i)-\sum\left(x_i^2-y_i^2\right).\sum(2x_iy_i)\right]}{n.\left\{\sum\left(\left(x_i^2-y_i^2\right)^2-(2x_iy_i)^2\right)\right\}-\left\{\left(\sum\left(x_i^2-y_i^2\right)\right)^2-(\sum 2x_iy_i)^2\right\}} \quad (5)$$

2.2.5. Quartimax

As seen in Table 2, the quartimax criterion is obtained when γ=0 in the orthomax criterion (Kaiser, 1958). In this case, the formula that gives the optimum angle θ for quartimax was expressed in Equation (6), if γ is written as zero in Equation (4) for θ.

$$\theta = \frac{1}{4}arctan\frac{-4\sum(x_i^2 - y_i^2)\,x_iy_i}{\sum(4x_i^2y_i^2 - (x_i^2 - y_i^2)^2)} \quad (6)$$

2.2.6. Equamax

As seen in Table 2, the equamax criterion is reached by writing γ=k/2, where k is the number of factors in the orthomax criterion. In this case, the formula obtained for θ in the equamax method was represented in Eqaution (7);

$$\theta = \frac{1}{4}arctan\frac{2\left[n\sum\left(x_i^2-y_i^2\right).(2x_iy_i)-\frac{k}{2}\sum\left(x_i^2-y_i^2\right).\sum(2x_iy_i)\right]}{n.\left\{\sum\left(\left(x_i^2-y_i^2\right)^2-(2x_iy_i)^2\right)\right\}-\frac{k}{2}\left\{\left(\sum\left(x_i^2-y_i^2\right)\right)^2-(\sum 2x_iy_i)^2\right\}} \quad (7)$$

3. Results

In order to exemplify the stages of factor analysis over numerical values, a data set with 9 variables and 20 observations for each variable is given below. The data meet the preconditions for the analysis and are given in Table 3. In addition, the Pearson correlation coefficients between variables in the present dataset and their descriptive statistics were tabulated in Tables 4 and 5, respectively. When the data is examined according to the KMO precondition, it is seen that it has reached a sufficient level. KMO value was found as 0.718. From these data, the factor load matrix obtained that has not been rotated yet, as seen in Table 6. Let us take the factor load matrix obtained from our example and apply rotations with ϴ angles that we will obtain from orthogonal rotation methods..

3.1. Varimax

Let us calculate the required angle of rotation for our example from the formula for θ in Equation (5). Since the explained variances will decrease from the first factor, the

Table 3. Data set belonging to barley plant

Plant Height (PH)	Number of Leaves (NL)	Spike Length (SL)	Spike Weight (SW)	Grain Yield (GY)	Flowering Period in Days (FP)	Harvest Index (HI)	Yield (Y)	1000-Grain Weight (1000-GW)
93,2	10,2	20,1	184	38,4	89	21	160,86	24,43
71,6	12	17,2	133,2	26,4	95	20	164,86	21,42
66,9	10	16,1	127,2	25,2	89	20	150,86	28,82
81,1	11	18,2	184,4	36,8	90	20	164,28	22
69,6	12,4	17,1	124,4	18	94	15	62	24,45
71,4	10,4	16,4	131,2	24,2	91	19	89,43	18,8
85,6	10,6	20	236	42,4	92	18	124,28	20,67
71	12	16,1	126,8	20	93	16	72,57	23,2
76,6	10	17,4	180,4	34,8	89	19	82,28	26,02
81,1	10,2	20,1	208,4	38	90	18	126,57	19,97
66.6	10.8	14.9	102	12.8	92	13	85.71	22.6
68.9	10.6	15.5	141.2	27.6	91	20	96.86	24.02
70.7	10.2	17.2	136.8	24.8	89	18	130.28	19.2
74.9	12.4	17.8	136.4	28.4	95	21	86.29	22.45
77.3	10	17.7	165.2	29.6	90	18	124.28	22.55
81.1	9.8	18.7	158.4	38.4	91	24	133.57	22.65
66.7	11.4	15.9	118	25.6	91	22	92.86	20.62
71.3	10.8	16.5	148.8	23.6	93	16	74	21.9
78	9.8	18.8	152.4	34.4	90	23	125.42	20.05
77.6	9.8	17.6	146.2	32.2	88	13	93.4	23.6

DOI: 10.54975/isubuzfd.1370165

Table 4. Pearson correlation coefficient of the variables in the study

	PH	NL	SL	SW	GY	FP	HI	Y	1000-GW
PH	1								
NL	-0.356	1							
SL	0.911	-0.289	1						
SW	0.818	-0.368	0.839	1					
GY	0.849	-0.481	0.856	0.872	1				
FP	-0.308	0.857	-0.217	-0.289	-0.412	1			
HI	0.271	-0.142	0.324	0.180	0.506	-0.063	1		
Y	0.483	-0.331	0.496	0.392	0.542	-0.316	0.499	1	
1000-GW	-0.082	-0.038	-0.213	-0.128	-0.091	-0.160	-0.092	-0.016	1

Table 5. Descriptive statistics of the variables

Variable	Mean	Minimum	Maximum	Standard Deviation
PH	75.06	66.6	93.2	6.97
NL	10.72	9.8	12.4	0.87
SL	17.46	14.9	20.1	1.51
SW	152.07	102	236	32.78
GY	29.08	12.8	42.4	7.75
FP	91.1	88	95	2.05
HI	18.7	13	24	2.99
Y	112.03	62	164.86	32.53
1000-GW	22.47	18.8	28.82	2.4

Table 6. Non-rotational factor loads matrix

	F1	F2	F3	F4	F5	F6	F7	F8	F9
PH	-0.89371	0.18259	0.210248	-0.16877	0.029005	0.2545	0.033444	0.165908	-0.03152
NL	0.589313	0.703287	-0.03953	-0.29236	0.078643	0.026821	-0.25321	0.006352	-0.00072
SL	-0.89029	0.315936	0.178741	-0.07606	0.029145	0.157054	0.025808	-0.20881	-0.00738
SW	-0.86091	0.176007	0.333949	-0.11343	-0.04522	-0.30024	-0.02723	0.016024	-0.10144
GY	-0.95622	0.093428	0.000228	-0.06713	-0.17474	-0.10445	-0.05324	0.028634	0.165271
FP	0.52052	0.792154	-0.04514	-0.18471	0.005479	-0.07652	0.241814	0.017089	0.027265
HI	-0.4611	0.203807	-0.78067	0.067879	-0.35724	0.024179	-0.00917	0.000428	-0.05926
Y	-0.65859	-0.01459	-0.51027	-0.02713	0.547827	-0.0678	0.009292	0.009134	0.007218
1000-GW	0.115903	-0.49265	-0.12496	-0.85006	-0.06279	-0.00477	0.032403	-0.02481	-0.00324

first two factors were considered here in order to do the operations manually.

The first two factors of the factor loads matrix are given in Table 7. The common factor variance used to standardize the factor loads given in Table 7.

Table 7. Loads matrix

	F1	F2
PH	0.894	-0.183
NL	-0.589	-0.703
SL	0.89	-0.316
SW	0.861	-0.176
GY	0.956	-0.093
FP	-0.521	-0.792
HI	0.461	-0.204
Y	0.659	0.015
1000-GW	-0.116	0.493

The common factor variance for the ith row is h_i^2, where h_i^2 is calculated as;

$$h_i^2 = a_{i1}^2 + a_{i2}^2 + \cdots + a_{ip}^2$$

h_i^2 matrix is calculated from Table (7) and given in Table (8).

Table 8. Calculated h_i^2 matrix

	h_i^2
PH	0.832
NL	0.842
SL	0.892
SW	0.772
GY	0.923
FP	0.898
HI	0.254
Y	0.434
1000-GW	0.256

In the next step, the square root of the sum of the squares of the calculated loads (h_i^2) for each row is calculated, and each load in this row is weighted by dividing it by this calculated value. The weighted loads can now be substituted in the formula. For convenience in calculations, the required values are found for each row in Table 9 where,

$f_{i1}z$: weighted factor loading of ith row and 1st factor

Table 9. Values calculated over weighted loads

	$f_{i1}z$	$f_{i2}z$	u	v	u^2-v^2	$u.v$
PH	0.980112	-0.20063	0.920369	-0.39327	0.692415	-0.36196
NL	-0.64189	-0.76612	-0.17493	0.983532	-0.93674	-0.17205
SL	0.94234	-0.33458	0.776058	-0.63058	0.204632	-0.48937
SW	0.979929	-0.20031	0.920136	-0.39258	0.692531	-0.36123
GY	0.995078	-0.0968	0.980809	-0.19265	0.924873	-0.18895
FP	-0.54979	-0.83577	-0.39624	0.919002	-0.68756	-0.36414
HI	0.914711	-0.40477	0.672854	-0.7405	-0.09561	-0.49825
Y	1.000324	0.022769	1.000129	0.045553	0.998183	0.045559
1000-GW	-0.22927	0.974377	-0.89685	-0.44678	0.604722	0.400695
Total			3.802343	-0.84829	2.397447	-1.98969

$f_{i2}z$: weighted factor loading of i^th row and 2^nd factor

$$u: (f_{i1}z)^2-(f_{i2}z)^2$$
$$v: 2.(f_{i1}z).(f_{i2}z)$$

The optimum angle is found by substituting the sums calculated from Table 9 in the angle formula:

$$\theta = \frac{1}{4} arctan\left(\frac{2.\left(9.\sum uv - \left(\sum u \sum v\right)\right)}{9.\left(\sum u^2 - v^2\right) - \left(\left(\sum u\right)^2 - \left(\sum v\right)^2\right)}\right)$$

(k=9; k is the number of variables)

$$\Theta = \frac{1}{4} arctan \frac{-29.3636}{7.838799} = -0.32748$$

so the rotation matrix T:

$$T = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \text{ if so}$$

$$T = \begin{bmatrix} 0.946776 & 0.321894 \\ -0.321894 & 0.946776 \end{bmatrix}$$

Now, according to equation 7, If the matrix (Table 7) multiplied with the matrix T, the varimax rotation completed matrix is found (Table 10).

Table 10. Varimax rotation completed matrix

	F1	F2
PH	0.905324	0.114513
NL	-0.33136	-0.85518
SL	0.944349	-0.0127
SW	0.871827	0.110518
GY	0.935054	0.21968
FP	-0.23833	-0.91755
HI	0.50213	-0.04475
Y	0.619097	0.22633
1000-GW	-0.26852	0.429421

3.2. Quartimax

Let us rotate our data with the quartimax method. The weighing of data will not differ from varimax. Let us consider the table of values given in Table 11 by Equation (8).

$$\theta = \frac{1}{4} arctan \frac{-4\sum(x_i^2-y_i^2)x_i y_i}{\sum\left(4x_i^2 y_i^2 - \left(x_i^2-y_i^2\right)^2\right)} = \frac{1}{4} arctan \frac{2\sum uv}{\sum(u^2-v^2)} \quad (8)$$

$$\Theta = \frac{1}{4} arctan \frac{2 * -1{,}98969}{2{,}39745} = -0{,}257141$$

$$T = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \text{ if so}$$

$$T = \begin{bmatrix} 0{,}967121 & -0{,}254317 \\ -0{,}254317 & 0{,}967121 \end{bmatrix}$$

According to equation 8, multiplying the matrix (Table 7.) with the matrix T results in the quartimax rotation completed matrix (Table 11).

Table 11. Quartimax rotation completed matrix

	F1	F2
PH	0.911135	0.050577
NL	-0.39067	-0.82977
SL	0.941119	-0.07906
SW	0.87744	0.048947
GY	0.948185	0.153394
FP	-0.30225	-0.89853
HI	0.497741	-0.07994
Y	0.633478	0.182241
1000-GW	-0.23766	0.447237

3.3. Equamax

Two factors were chosen in the present example, γ=k/2=1. Therefore, the Equamax method will give the same rotational factor load as varimax.

4. Discussion

Although there are examples of the use of factor analysis in agricultural and animal research in the literature (Goddard and Beilharz, 1984; Tan and Corke, 2002; Sadek et al., 2006; Beniston et al., 2016), the rotation methods, which is one of the stages of this analysis, have not been adequately explained. Although there are methods other than the rotation methods used in package programs, they are not widely used because they are not documented in the literature. In addition, the lack of clear algorithms of the methods used in the program both makes it difficult to discover new methods and causes the preferred method to be chosen by trial and error or randomly. Our examples illustrate how the value for the angle of rotation was obtained. As a result, the angle of rotation that will increase the variance of the data is found with the help of formulas that depend on the parameter γ. Since our example was solved manually, two factors were taken into account. Therefore, the results of equamax and varimax were the same. A clearer distinction is made by the varimax rotation when compared to the quartimax rotation, making it more suitable for this dataset. According to Wrigley et al. (1958), although varimax is better at reaching Thurstune's simple structure criterion, quartimax is more useful in terms of operation simplicity. In a study examining rotation methods, Saraçlı (2011) found that the Equamax results were closest to the Varimax values and the Quartimax results were the farthest to varimax. Karaman et al. (2017) reported that principal component analysis was the method that explains the total variance best in all analyzed steps, and according to the comparisons made in terms of factor loads, the principal component analysis yielded the highest factor load for each step. In this study, the principal component analysis method was used while calculating factor loads. Osborne (2015), reported that in the modern era of high-power computing, vertical rotations are probably not the best practice because oblique rotations can accurately model unrelated and correlated factors, whereas orthogonal rotations cannot effectively address correlated factors. Thus, it is reported that there is little cost to using oblique rotations, regardless of the underlying relationship of the factors.

For three real datasets, Akhtar-Danesh (2017) used principal component and principal axis factoring methods for factor extraction, as well as varimax, equamax, and quartimax factor rotation techniques. Akhtar-Danesh (2017) compared these techniques according to the number of Q-types loaded on each factor, the number of distinctive expressions in each factor, and the excluded Q-types and reported that there was not much difference between the principal component and the principal axis factoring factor inferences. The main findings of Akhtar-Danesh (2017) were the emergence of a general factor and fewer excluded Q-types based on quartimax rotation. Another interesting finding was that there were fewer discriminative expressions for factors based on quartimax rotations than for varimax and equamax rotations and it was reported that these findings were not conclusive and that further analysis on more datasets was required.

5. Conclusion

In this study, explanations about rotation methods are given and how, and how much rotation each method performs is explained. Although some studies provide the criteria for the most commonly used rotation methods, it is not stated how many degrees of angle should be used based on these criteria. Rather than specifying the ideal method by subjecting a limited number of data to various methods and generalizing according to the results, knowing the basics of the methods, the researcher should choose the most appropriate method for the data. Since all calculations are made manually in the study, the calculation of a third factor requires repeating the same calculation steps many times. For this reason, calculations were done for only two factors. Similar calculation steps can be performed for more factors using the same algorithm.

Acknowledgements

This manuscript was generated from first author’s master thesis.

Author Contributions

Ayşe Sümeyye Can: Investigation, Methodology, Validation, Original Draft Writing, Conceptualization. Özgür Koşkan: Supervision, Project Administration, Validation. Malik Ergin: Data Curation, Formal Analysis, Review and Editing.

Conflict of Interest

As the authors of this study, we declare that we do not have any conflict of interest statement.

Ethics Committee Approval

As the authors of this study, we declare that we do not have any ethics committee approval.

References

Akhtar-Danesh, N. (2017). A comparison between major factor extraction and factor rotation techniques in q-methodology. *Open Journal of Applied Sciences*, *7*(4), 147-156. https://10.4236/ojapps.2017.74013

Alpar, R. (2013). *Çok Değişkenli istatistiksel yöntemler*. Detay Yayıncılık, Ankara.

Beniston, J. W., Lal, R., & Mercer, K. L. (2016). Assessing and managing soil quality for urban agriculture in a degraded vacant lot soil. *Land Degradation & Development*, *27*(4), 996-1006.

Brown, B. L., & Hendrix, S. B. (2005). Partial correlation coefficients. *Encyclopedia of Statistics in Behavioral Science*, *3*, 1518-1523. https://doi.org/10.1002/0470013192.bsa469

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *36*(1), 111-150. https://doi.org/10.1207/S15327906MBR3601_05

Darton, R. A. (1980). Rotation in factor analysis. *The Statistician*, 167-194. https://doi.org/10.2307/2988040

Didia, J. U. D., & Idenedo, O. W. (2021). Customer orientation and loyalty behaviour of users of global system of mobile telecommunication services. *International Journal of Marketing and Management*, *7*(3), 34-44.

Finch, W. H. (2011). A comparison of factor rotation methods for dichotomous data. *Journal of Modern Applied Statistical Methods*, *10*(2), 549-570. https://10.22237/jmasm/1320120780

Goddard, M. E., & Beilharz, R. G. (1984). A factor analysis of fearfulness in potential guide dogs. *Applied Animal Behaviour Science*, *12*(3), 253-265.

İlhan, F. (2007). A study on factor analysis and its application to agricultural data. Master thesis, Selçuk Üniversitesi, Konya, 63 p.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 1-14. https://10.1007/BF02289233

Karaman, H., Atar, B., & Çobanoğlu, A. D. (2017). The comparison of factor extraction methods used in exploratory factor analysis. *Gazi University Journal of Education Faculty*, *37*(3), 1173-1193.

Khalaf, K. (2007). Factor analysis and an application. Master Thesis. *Gazi University Institute of Science*, Ankara, 63 p.

Osborne, J. W. (2015). What is rotating in explor what is rotating in exploratory factor analysis? or analysis? *Practical Assessment, Research, and Evaluation*, *20*(2), 1-7. https://10.7275/hb2g-m060

Polat, Y. (2012). Examining factor analyses methods comparatively and applying to animal husbandry trial. PhD thesis, Çukurova Üniversitesi, Adana, 206 p.

Sadek, M. H., Al-Aboud, A. Z., & Ashmawy, A. A. (2006). Factor analysis of body measurements in Arabian horses. *Journal of Animal Breeding and Genetics*, *123*(6), 369-377.

Saraçlı, S. (2011). An application on comparison the extracting methods in factor analysis. *Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi*, *1*(3), 22-26.

Shrestha, N. (2021). Factor analysis as a tool for survey analysis. *American Journal of Applied Mathematics and Statistics*, *9*(1), 4-11. http:// 10.12691/ajams-9-1-2

Süzülmüş, S. (2005). Identifying models of factor analysis and using of generalized inverses. PhD Thesis. Çukurova Üniversitesi, Adana, 159 p.

Tan, Y., & Corke, H. (2002). Factor analysis of physicochemical properties of 63 rice varieties. *Journal of the Science of Food and Agriculture*, *82*(7), 745-752.

Wrigley, C., Saunders, D. R., & Neuhaus, J. O. (1958). Application of the quartimax method of rotation to thurstone's primary mental abilities study. *Psychometrika*, *23*(2), 151-170. https://10.1007/BF02289011