American Journal of Theoretical and Applied Statistics
2013; 2(6): 191-201
Published online November 10, 2013 (http://www.sciencepublishinggroup.com/j/ajtas)
doi: 10.11648/j.ajtas.20130206.16

Method of maximum likelihood estimation of optimal number of factors: An information criteria approach

Nwosu F. Dozie[1], Onyeagu, Sidney I.[2], Osuji, George A.[2], Ekezie Dan Dan[3]

[1]Department of Mathematics and Statistics Federal Polytechnic Nekede, Owerri Imo State
[2]Department of Statistics Nnamdi Azikiwe University Awka, Anambra State
[3]Department of Statistics, Imo State University, Owerri, PMB 2000, Owerri Nigeria

Email address:
fedocon2003@yahoo.com(N. F. Dozie), danekezie@yahoo.com(E. D. Dan)

To cite this article:
Nwosu F. Dozie, Onyeagu I. Sidney, Osuji A. George, Ekezie Dan Dan. Method of Maximum Likelihood Estimation of Optimal Number of Factors: An Information Criteria Approach. *American Journal of Theoretical and Applied Statistics.* Vol. 2, No. 6, 2013, pp. 191-201. doi: 10.11648/j.ajtas.20130206.16

Abstract: In this paper, we established the number of factors (k) to retain in a factor analysis for different sample sizes for the method of maximum likelihood estimation using StataSE 9 package. Using simulated values, the Akaike (AIC), the Schwarz (SIC) and Hannan Quinne (HQIC) Information criteria values are obtained for samples of 30, 50, and 70 when the number of variables considered is 10 and the number of factors to retain are 2,3, and 5. It was discovered that the AIC, SIC and HQIC values are smallest when k is 5 and highest when k is 2. This implies that the optimal number of factors to retain is 5. Also, as the sample size increases, the AIC, SIC and HQIC for all the k values increases.

Keywords: Factor Analysis, Factor Rotation, Maximum Likelihood Estimation Method, Akaike, Schwarz, Hannan Quinne Information Criteria

1. Introduction

Factor analysis attempts to simplify complex and diverse relationships that exist among a set of observed variables by uncovering common dimensions or factors that link together the seemingly unrelated variables and consequently provides insight into the underlying structure of the data. The goal of factor analysis is to reduce a large number of variables to a smaller number of factors, to concisely describe the relationship among observed variables or to test theory about underlying processes (Comrey and Lee; 1992).

Factor analysis can be viewed as an extension of principal components analysis. Both are attempts to approximate the covariance matrix, Σ. principal components analysis and factor-analytic model often yield solutions that are very similar. For this reason, many authors treat principal components analysis as just another many authors type of factor analysis. Under the factor model, each response variates will be represented as a linear function of a small number of unobservable common-factor variates and a single latent specific variates. The common factors generate the covariances among the observable responses, while the specific terms contribute only to the variances of their particular responses.

1.1. The Mathematical Model for Factor Structure

Suppose that the multivariate system consists of p responses described by the observable random variables $X_1, \ldots X_p$. The Xi have a non-singular multinormal distribution. Since only the covariance structure will be of interest, we can assume without loss of generality that the population means of X_i are zero (Onyegu;2003).

Let,

$$Y_1 - \mu_1 = \lambda_{11}X_1 + \cdots + \lambda_{1m}X_m + e_1$$

$$\vdots$$

$$Y_p - \mu_p = \lambda_{p1}X_1 + \cdots + \lambda_{pm}X_m + e_p$$

where,

X_j = j-th common-factor variates

λ_{ij} = parameter reflecting importance of the j-th factor in composition of ith response

$e_1 =$ ith specific factor variates.

In the usage of factor analysts, λ_{ij} is called the loading of the ith response on the jth common factor.

For matrix version of the model;

Let,

$$X' = (x_1, x_2, \ldots, x_m), \quad y' = (y_1, y_2, \ldots, y_p)$$

$\varepsilon' = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)$, $\mu' = (\mu_1, \mu_2, \dots, \mu_p)$,

and $\Lambda = \begin{bmatrix} \lambda_{11} & .. & \lambda_{1m} \\ \vdots & \vdots & \vdots \\ \lambda_{p1} & .. & \lambda_{pm} \end{bmatrix}$.

Then, the factor model can be written as

$(y_i - \mu_y) = \Lambda X_i + \varepsilon_i, i = 1, \dots, q.$

Where q is the number of observations. Y_i is an observed vector with p components with mean μ_y and Λ is a pxq matrix called the factor loadings. X_i is an observed vector with q components, $q < p$, the components of which are called common factors and ε_i is an unobserved error with mean $0_{(px1)}$ and is called the specific factor. X_i is assumed to follow a normal distribution with

$E(X_i) = 0_{(qx1)}$, $Var(X_i) = E(X_i X'_i) = I_{(qxq)}$, that is the orthogonal case.

Then, Y_i/X_i follows $N\left(\mu_y + \Lambda X_i, Cov(\varepsilon_i)\right)$, where $Cov(\varepsilon_i) = \psi^2 = diag\left(\psi_1^2, \dots, \psi_p^2\right)$.

ψ^2 *is* called uniqueness or specific variance.
Consequently, a covariance structure for Y is

$$\Sigma = Cov(Y_i) = E(Y_i - \mu_y)(Y_i - \mu_y)'$$

$= E(\Lambda X_i + \varepsilon_i)(\Lambda X_i + \varepsilon_i)' = E(\Lambda X_i X'_i \Lambda' + \varepsilon_i X'_i \Lambda' + \Lambda X_i \varepsilon'_i + \varepsilon_i \varepsilon'_i) = E(\Lambda X_i X'_i \Lambda') + E(\varepsilon_i X'_i \Lambda') + E(\Lambda X_i \varepsilon'_i) + E(\varepsilon_i \varepsilon'_i)$ Recall that, $E(X_i) = 0_{(qx1)}$, $E(X_i X'_i) = I_{(qxq)}$, $Cov(\varepsilon_i) = E(\varepsilon_i \varepsilon'_i) = \psi^2$,

Then, $E(\Lambda X_i X'_i \Lambda') + E(\varepsilon_i X'_i \Lambda') + E(\Lambda X_i \varepsilon'_i) + E(\varepsilon_i \varepsilon'_i) = \Lambda\Lambda' + \psi^2$, and

$$Cov(Y_i, X_i) = E(Y_i - \mu_y)X'_i = E(\Lambda X_i + \varepsilon_i)X'_i .$$

Introducing communality, $Cov(Y_i) = \Lambda\Lambda' + \psi^2$ can be written as

$Var(Y_{ij}) = \Lambda_{1j}^2 + \dots + \Lambda_{qj}^2 + \Psi_j^2$; $Cov(Y_{ij}, Y_{ik}) = \Lambda_{1j}\Lambda_{1k} + \dots + \Lambda_{qj}\Lambda_{qk}$ and

$Cov(Y_{ik}, X_{jk}) = \Lambda_{jk}$.

Communality is the portion of the variance of the variable contributed by the q common factors.

Suppose the jth communality is h_j^2, then

$$Var(Y_{ij}) = \sigma_{ij} = \text{communality} + \text{specific variance} = h_j^2 + \psi_j = (\lambda_{1j}^2 + \lambda_{2j}^2 + \dots + \lambda_{qj}^2) + \psi_j, j = 1, \dots, p.$$

The j-th communality is the sum of square of the loading of the jth variable on the q common factor. When the number of factors q >1, there are multiple factor loadings that generate the same covariance matrix.

The loading in the model above can be multiply by an orthogonal matrix without impairing their ability to reproduce the covariance matrix in $\Sigma = \Lambda\Lambda' + \psi^2$.

Let β be any qxq orthogonal matrix. If we let $\Lambda^* = \Lambda\beta$ and $X_i^* = \beta' X_i$, then X_i^* has the same statistical properties as X_i since

$$E(X_i^*) = E(\beta' X_i) = \beta' E(X_i) = 0.$$

$$Cov(X_i^*) = Cov(\beta' X_i) = \beta' Cov(X_i)\beta = \beta' \beta = I_{mxm}$$

Λ and Λ' yield the same covariance because $\Lambda\Lambda' = \Lambda^*\Lambda^{*\prime}$. The factor model

$(Y_i - \mu_y) = \Lambda X_i + \varepsilon_i = \Lambda\beta\beta' X_i + \varepsilon_i = \Lambda^* X_i^* + \varepsilon_i$,

produces the same covariance matrix Σ, since $\Sigma = \Lambda\Lambda' + \psi^2 = \Lambda\beta\beta' \Lambda + \psi^2 = \Lambda^*\Lambda^{*\prime} + \psi^2$.

A particular set of loadings needs to be chosen. A good set is one that is easily interpreted. This means sparse solution with many zero. Factor rotation is one approval to finding the solution, as it rotates the coordinates system for Y on X .

In this work, we consider four kinds of orthogonal factor rotations namely, varimas, equmax, quartimax and orthomax.

2. The Maximum Likelihood Method

If the distribution of X and the specific factors, ε_i, are assumed to be normal, the estimates of the factor loadings and the uniqueness can be obtained using the maximum likelihood method (Johnson and Winchern, 2007). The distribution of Y given Λ, ψ^2 is normal with mean 0 and covariance matrix $\Lambda\Lambda' + \psi^2$, and the likelihood is

$$L(\Lambda, \psi^2) \propto \det(\Sigma)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n}(y_i - \mu_y)\Sigma^{-1}(y_i - \mu_y)'} .$$

The resulting log likelihood (LL) is

$$LL(\Lambda, \Psi) \propto -\frac{n}{2} \text{In} \det(\Sigma) - \frac{1}{2}\sum_{i=1}^{n}(y_i - \mu_y)\Sigma^{-1}(y_i - \mu_y)'$$

The maximum likelihood estimates of Λ, ψ^2 (called $\widehat{\Lambda}$ *and* $\widehat{\psi}^2$) can be obtained by maximizing the LL. It can be shown that the estimates $\widehat{\Lambda}$ *and* $\widehat{\psi}^2$ satisfy the following;

$$S\widehat{\Psi}^2\widehat{\Lambda} = \widehat{\Lambda}\left(I + \widehat{\Lambda}' (\widehat{\Psi}^2)^{-1}\widehat{\Lambda}\right).$$

$\widehat{\psi}^2 = \text{diag}(S - \widehat{\Lambda}\widehat{\Lambda}')$, $\widehat{\Lambda}' (\widehat{\psi}^2)^{-1}\widehat{\Lambda}$ is diagonal.

Based on the invariance property of maximum likelihood estimates, the maximum likelihood estimate of the communality due to j-th factor is $\hat{\lambda}_{1j}^2 + \hat{\lambda}_{2j}^2 + \dots + \hat{\lambda}_{qj}^2$..

As a result, the proportion of the total sample variance due the j-th factor is $\frac{\hat{\lambda}_{1j}^2 + \hat{\lambda}_{2j}^2 + \dots + \hat{\lambda}_{qj}^2}{s_{11} + s_{22} + \dots + s_{qq}}$, where s_{ii} is the (i,i)-th entry of the sample covariance matrix which is an estimated of the unknown population covariance matrix, Σ

If y_i is standardized to be

$$Z_i = V^{-\frac{1}{2}}(y_i - \mu_y),$$

the covariance matrix q is given as

$Q = V^{-\frac{1}{2}}\,\Sigma\,V^{-\frac{1}{2}} = \left(V^{-\frac{1}{2}}\Lambda\right)\left(V^{-\frac{1}{2}}\Lambda\right)' + V^{-\frac{1}{2}}\,\widehat{\psi}^2\,V^{-\frac{1}{2}} = \Lambda_z\Lambda'_z + \widehat{\psi}^2{}_z$, where $V^{-\frac{1}{2}}$ is the diagonal matrix with the reciprocal of the sample standard deviation on the main diagonal of y_i. Based on the invariance property of maximum likelihood estimators, the maximum likelihood estimator of Q is

$Q = \widehat{\Lambda}_z\widehat{\Lambda}'_z + \widehat{\psi}^2{}_z$, where $\widehat{\Lambda}_z = V^{-\frac{1}{2}}\widehat{\Lambda}$.

The proportion of total standardized sample variance due to j-th factor is

$\frac{\Sigma_{i=1}^{q}\hat{\lambda}_{ij}^2}{tr(S)} = \frac{\hat{\lambda}_{1j}^2+\hat{\lambda}_{2j}^2+\cdots+\hat{\lambda}_{qj}^2}{S_{11}+S_{22}+\cdots+S_{qq}} = \frac{\hat{\lambda}_{1j}^2+\hat{\lambda}_{2j}^2+\cdots+\hat{\lambda}_{qj}^2}{q}$, where $S_{ii} = \hat{\lambda}_{1j}^2 + \widehat{\Psi}_{1j}^2$.

2.1. Information Criteria

The necessity of introducing the concept of model evaluation has been recognized as one of the important technical areas and the problem is posed on the choice of the best approximating model among a class of competing models by a suitable model evaluation criterion given a data set. Model evaluation criteria are figures of merit, or performance measures for competing models. Factor analysis can be characterized as multivariate technique for analyzing the internal relationship among a set of variables. Based on the usual factor analysis model, choosing a model with too few parameters can involve making unrealistically simple assumptions and lead to high bias, poor prediction, and missed opportunities for insight. Such models are not flexible enough to describe the sample or the population well. A model with too many parameters can fit the observed data very well, but be too closely tailored to it; such models may generalize poorly. Penalized-likelihood information criteria, such as Akaike's information criterion (AIC), the Schwarz's information criterion (SIC), the Hannan-Quinn information criterion (HQIC) and so on are widely used for model selection. The comparison of the models using information criterion can be viewed as equivalent to a likelihood ratio test and understanding the differences among the criteria may make it easier to compare the results and to use them to make informed decisions (Akaike; 1973).

AKAIKE'S information criterion is probably the most relevant and famous as for the comparison and selection between different models and is constructed on log likelihood

$$\text{AIC} = -2logmax\ L + 2k$$

where L denotes the likelihood function of the factor model and k is the number of the model's parameter/factors. $logmax\ L(k) = -\frac{1}{2}N\left|log|\Sigma_k| + tr\Sigma_k^{-1}S\right|$, where S denotes the sample covariance matrix of Y and $\Sigma_k = \Lambda_k\Lambda_k^{-1} + \psi^2$; Λ_k is the matrix factor of factor loading. The first term can be interpreted as a goodness-of-fit measurement, while the second gives a growing penalty to increasing numbers of parameters according to the parsimony principle. In the choice of model, a minimization rule is used to select the model with the minimum Akaike information criterion value.

Still in the likelihood based procedures, Schwarz (1978) proposed the alternative information criterion given by

$$SIC = -logmax\ L + \frac{1}{2}k\log N.$$

Unlike the AIC. SIC considers the number of N of observations and is therefore less favourable to factors inclusion.

Finally, the third criteria are the Hannan-Quinn information criterion (HQC); it is a criterion for model selection. It is an alternative to Akaike information criterion (AIC) and Bayesian information criterion (BIC)(Harman;1976). It is given as

$$HQIC = -logmax\ L + 2k\log\log N$$

where k is the number of parameters, N is the number of observations..

3. Comparison of AIC and SIC after Rotation at Different Sample Sizes and Different Retained Number of Factors (k)

3.1. For n=30, p =10 and k = 2

3.1.1. Rotated Factor Loadings

Varimax Loadings		Equamax Loadings		Quartimax Loadings		Orthomax Loadings	
I	II	I	II	I	II	I	II
-0.3702	-0.2308	-0.3577	-0.2347	-0.3609	-0.2297	-0.3552	-0.2385
-0.4130	0.2369	-0.4155	0.2311	-0.4122	0.2369	-0.4179	0.2267
0.1917	0.9809	0.1811	0.9835	0.1948	0.9809	0.1707	0.9853
0.9773	0.2086	0.9750	0.2222	0.9780	0.2086	0.9726	0.2325
0.3106	-0.2791	0.3136	-0.2747	0.3098	-0.2786	0.3165	-0.2714

Varimax Loadings		Equamax Loadings		Quartimax Loadings		Orthomax Loadings	
I	II	I	II	I	II	I	II
0.2343	-0.1618	0.2361	-0.1586	0.2338	-0.1618	0.2377	-0.1561
0.0613	0.1670	0.0595	0.1678	0.0618	0.1670	0.0577	0.1684
-0.1235	-0.1640	-0.1217	-0.1657	-0.1240	-0.1640	-0.1200	-0.1670
0.1354	0.2694	0.1324	0.2713	0.1362	0.2694	0.1296	0.2727
0.1303	-0.1413	0.1318	-0.1395	0.1299	-0.1413	0.1333	-0.1381

3.1.2. Factor Rotation Matrix

Varimax

	Factor I	Factor II
Factor I	0.4587	0.8886
Factor II	0.8886	-0.4587

Equamax

	Factor I	Factor II
Factor I	0.4490	0.8935
Factor II	0.8935	-0.4490

Quartimax

	Factor I	Factor II
Factor I	0.4614	0.8872
Factor II	0.8872	-0.4614

Orthomax

	Factor I	Factor II
Factor I	0.4396	0.8982
Factor II	0.8982	-0.4396

3.1.3. Information Criteria

Information Criteria	Values
Log Likelihood	-116.6265
Akaike	237.2530
Schwarz	118.1036
Hannan Quinne	117.3042

3.2. For $n = 30$, $p = 10$ and $k = 3$

3.2.1. Rotated Factor Loadings

Varimax Loadings			Equamax Loadings			Quartimax Loadings			Orthomax Loadings		
Factor I	Factor II	Factor III	Factor I	Factor II	Factor III	Factor I	Factor II	Factor III	Factor I	Factor II	Factor III
-0.3596	-0.2388	0.1169	-0.3566	-0.2458	0.1116	-0.3605	-0.2369	0.1181	-0.3534	-0.2497	0.1128
-0.4125	0.2337	0.0431	-0.4154	0.2275	0.0481	-0.4117	0.2354	0.0418	-0.4186	0.2222	0.0456
0.1929	0.9791	0.0647	0.1805	0.9798	0.0859	0.1962	0.9788	0.0596	0.1671	0.9828	0.0788
0.9776	0.2104	0.0079	0.9748	0.2226	0.0127	0.9783	0.2071	0.0070	0.9717	0.2358	0.0129
0.3121	-0.2970	0.2662	0.3159	-0.2987	0.2598	0.3111	-0.2967	0.2678	0.3193	-0.2925	0.2626
0.2339	-0.1580	-0.0552	0.2358	-0.1538	-0.0586	0.2333	-0.1591	-0.0544	0.2380	-0.1511	-0.0570
0.0656	0.1231	0.6598	0.0641	0.1096	0.6623	0.0660	0.1263	0.6592	0.0613	0.1154	0.6616
-0.1223	-0.1792	0.2167	-0.1200	-0.1854	0.2127	-0.1229	-0.1777	0.2176	-0.1179	-0.1854	0.2139
0.1375	0.2498	0.3046	0.1344	0.2449	0.3100	0.1383	0.2509	0.3033	0.1304	0.2490	0.3084
0.1267	-0.1041	-0.5539	0.1280	-0.0905	-0.5560	0.1264	-0.1074	-0.5533	0.1303	-0.0930	-0.5550

3.2.2. Factor Rotation Matrix

Varimax

	Factor I	Factor II	Factor III
Factor I	0.9598	0.2805	0.0128
Factor II	-0.2807	0.9575	0.0662
Factor III	0.0063	-0.0671	0.9977

Equamax

	Factor I	Factor II	Factor III
Factor I	0.9562	0.2922	0.0191
Factor II	-0.2928	0.9522	0.0868
Factor III	0.0072	-0.0886	0.9960

Quartimax

	Factor I	Factor II	Factor III
Factor I	0.9607	0.2772	0.0115
Factor II	-0.2774	0.9588	0.0612
Factor III	0.0059	-0.0620	0.9981

Orthomax

	Factor I	Factor II	Factor III
Factor I	0.9521	0.3052	0.0187
Factor II	-0.3057	0.9488	0.0791
Factor III	0.6452	-0.0810	0.9967

3.2.3. Information Criteria

Information Criteria	Values
Log Likelihood	-112.6245
Akaike	231.2490
Schwarz	114.8402
Hannan Quinne	109.9072

3.3. For n=30, p =10, and k = 5

3.3.1. Rotated Factor Loadings

Varimax Loadings					Equamax Loadings				
1	2	3	4	5	1	2	3	4	5
-0.3048	-0.1789	-0.1510	-0.3019	0.3626	-0.2696	-0.2996	-0.1599	-0.1268	0.4077
0.2245	0.0814	0.0675	-0.6254	0.0872	0.2542	-0.6117	0.0910	0.0727	0.0889
0.9738	-0.1255	-0.0744	0.0788	-0.0250	0.9665	0.1128	-0.1273	-0.0839	-0.0799
0.3850	0.2427	0.1088	0.7501	0.0204	0.3552	0.7662	0.2363	0.1027	-0.0492
-0.1048	0.9879	-0.0852	0.0734	-0.0219	-0.1122	0.0757	0.9852	-0.0880	-0.0570
-0.1445	-0.1293	0.1336	0.3524	0.2666	-0.1405	0.3567	-0.1207	0.1479	0.2594
0.1798	0.1513	-0.3317	0.0477	0.3284	0.1985	0.0681	0.1652	-0.3154	0.3250
-0.2144	-0.2273	-0.1283	-0.0367	0.4503	-0.1843	-0.0283	-0.2068	-0.1011	0.4801
0.3184	0.1505	-0.0711	0.0362	0.3408	0.3364	0.0629	0.1651	-0.0557	0.3147
-0.0542	-0.0761	0.9944	0.0391	-0.0304	-0.0524	0.0376	-0.0778	0.9917	-0.0796
Quartimax Loadings					Orthomax Loadings				
1	2	3	4	5	1	2	3	4	5
-0.3124	-0.1831	-0.1571	-0.2980	0.3545	-0.2443	-0.1536	-0.1189	-0.1712	0.4935
0.2156	0.0778	0.0654	-0.6289	0.0893	0.2799	0.0927	0.0769	-0.5509	0.2527
0.9749	-0.1263	-0.0731	0.0670	-0.0170	0.9550	-0.1374	-0.1013	0.0990	-0.1598
0.3945	0.2448	0.1101	0.7439	0.0306	0.3308	0.2290	0.0843	0.7248	-0.2960
-0.1024	0.9885	-0.0848	0.0711	-0.0145	-0.1102	0.9851	-0.0904	0.0549	-0.0789
-0.1430	-0.1304	0.1301	0.3537	0.2669	-0.1388	-0.1192	0.1441	0.4174	0.1492
0.1775	0.1481	-0.3362	0.0444	0.3269	0.2039	0.1641	-0.3217	0.1611	0.2781
-0.2197	-0.2314	-0.1351	-0.0342	0.4438	-0.1648	-0.2022	-0.0989	0.1111	0.4775
0.3155	0.1470	-0.0758	0.0307	0.3445	0.3485	0.1635	-0.0644	0.1597	0.2621
-0.0544	-0.0758	0.9947	0.0385	-0.0170	-0.0411	-0.0748	0.9924	0.1269	-0.0848

3.3.2. Factor Rotation Matrix

Varimax

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	-0.0891	0.9550	-0.2756	0.0622	-0.0149
Factor 2	-0.0897	0.2632	0.9577	0.0640	-0.0377
Factor 3	0.9825	0.0973	0.0541	0.1444	-0.0376
Factor 4	-0.1320	-0.0911	-0.0521	0.9853	0.0263
Factor 5	0.0357	0.0302	0.0355	-0.0171	0.9981

Equamax

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	-0.0966	0.0647	0.9528	-0.2778	-0.0387
Factor 2	-0.0905	0.0633	0.2607	0.9540	-0.0986
Factor 3	0.9717	0.1802	0.0942	0.0428	-0.1124
Factor 4	-0.1691	0.9791	-0.0990	-0.0535	0.0041
Factor 5	0.0984	0.0253	0.0744	0.0894	0.9880

Quartimax

	Factor 1	Factor 2	Factor 3	Factor 4	factor 5
Factor 1	-0.0868	0.9555	-0.2753	0.0601	-0.0105
Factor 2	-0.0892	0.2638	0.9582	0.0626	-0.0219
Factor 3	0.9847	0.0969	0.0558	0.1315	-0.0252
Factor 4	-0.1197	-0.0874	-0.0510	0.9873	0.0263
Factor 5	0.0251	0.0206	0.0208	-0.0206	0.9990

Orthomax

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	-0.0968	0.9521	-0.2801	0.0470	-0.0585
Factor 2	-0.0786	0.2636	0.9539	0.0455	-0.1112
Factor 3	0.9605	0.0838	0.0237	0.1564	-0.2130
Factor 4	-0.2026	-0.1028	-0.0651	0.9310	-0.2780
Factor 5	0.1442	0.0801	0.0826	0.3232	0.9282

3.3.3. Information Criteria

Information Criteria	Values
Log Likelihood	-108.2130
Akaike	226.4260
Schwarz	111.9058
Hannan Quinne	109.9072

3.4. For n =50 , p =10, and k =2.

3.4.1. Rotated Factor Loadings

Varimax Loadings		Equamax Loadings		Quartimax Loadings		Orthomax Loadings	
I	II	I	II	I	II	I	II
0.9922	-0.1243	0.9881	-0.1540	0.9933	-0.1159	0.9934	-0.1149
-0.2951	0.7780	-0.2717	0.7865	-0.3017	0.7755	-0.3025	0.7752
0.1655	0.1709	0.1705	0.1659	0.1640	0.1723	0.1638	0.1725
0.0827	0.0269	0.0835	0.0244	0.0825	0.0276	0.0825	0.0276
-0.0489	-0.0194	-0.0494	-0.0179	-0.0487	-0.0198	-0.0487	-0.0199
0.0562	-0.1489	0.0517	-0.1505	0.0575	-0.1484	0.0576	-0.1483
0.2249	-0.2948	0.2160	-0.3014	0.2274	-0.2929	0.2277	-0.2926
-0.2756	-0.3872	-0.2871	-0.3787	-0.2723	-0.3895	-0.2719	-0.3898
0.0914	0.2652	0.0993	0.2630	0.0891	0.2666	0.0888	0.2667
0.1112	-0.0118	0.1108	-0.0152	0.1113	-0.0109	0.1113	-0.0108

3.4.2. Factor Rotation Matrix

Varimax

	Factor I	Factor II
Factor I	0.9922	-0.1243
Factor II	0.1243	0.9922

Equamax

	Factor I	Factor II
Factor I	0.9881	-0.1540
Factor II	0.1540	0.9881

Quartimax

	Factor I	Factor II
Factor I	0.9933	-0.1159
Factor II	0.1159	0.9933

Orthomax

	Factor I	Factor II
Factor I	0.9934	-0.1149
Factor II	0.1149	0.9934

3.4.3. Information Criteria

Information Criteria	Values
Log Likelihood	-237.8175
Akaike	479.6350
Schwarz	239.6635
Hannan Quinne	238.7382

3.5. For n =50 , p =10, and k =3.

3.5.1. Rotated Factor Loadings

Varimax Loadings			Equamax Loadings			Quartimax Loadings			Orthomax Loadings		
Factor I	Factor II	Factor III	Factor I	Factor II	Factor III	Factor I	Factor II	Factor III	Factor I	Factor II	Factor III
0.9998	-0.0199	-0.0026	0.9998	-0.0200	0.0047	0.9998	-0.0198	-0.0044	0.9996	-0.0202	0.0176
-0.0792	0.1098	-0.1916	-0.0778	0.1101	-0.1920	-0.0796	0.1097	-0.1915	-0.0753	0.1087	-0.1937
0.0384	0.0163	0.1952	0.0369	0.0160	0.1955	0.0387	0.0164	0.1951	0.0344	0.0164	0.1959
-0.0707	-0.2210	0.0189	-0.0709	-0.2210	0.0180	-0.0707	-0.2210	0.0191	-0.0711	-0.2209	0.0187
-0.2750	0.0242	-0.2080	-0.2735	0.0246	-0.2100	-0.2754	0.0241	-0.2076	-0.2708	0.0232	-0.2137
-0.1164	-0.0678	-0.0494	-0.1161	-0.0677	-0.0504	-0.1165	-0.0679	-0.0492	-0.7754	-0.0681	-0.0514
-0.0205	0.9995	0.0251	-0.0205	0.9994	0.0266	-0.0205	0.9995	0.0248	-0.0206	0.9996	0.0193
0.1947	0.0172	0.1319	0.1937	0.0169	0.1334	0.1949	0.0172	0.1316	0.1920	0.0178	0.1357
0.1810	0.1431	-0.1012	0.11818	0.1433	-0.0997	0.1808	0.1431	-0.1016	0.1830	0.1425	-0.0983
-0.0067	0.0419	0.8396	-0.0129	0.0405	0.1896	-0.0052	0.0422	0.8396	-0.0236	0.0465	0.8391

3.5.2. Factor Rotation Matrix

Varimax

	Factor I	Factor II	Factor III
Factor I	-0.1719	0.9848	0.0251
Factor II	0.9851	0.1719	0.0022
Factor III	0.0021	-0.0251	0.9997

Equamax

	Factor I	Factor II	Factor III
Factor I	-0.1720	0.9848	0.0254
Factor II	0.9851	0.1718	0.0097
Factor III	-0.0052	-0.0267	0.9996

Quartimax

	Factor I	Factor II	Factor III
Factor I	-0.1719	0.9848	0.0250
Factor II	0.9851	0.1720	0.0004
Factor III	0.0039	-0.0247	0.9997

Orthomax

	Factor I	Factor II	Factor III
Factor I	-0.1721	0.9850	0.0162
Factor II	0.9849	0.1717	0.0211
Factor III	-0.0180	-0.0196	0.9996

3.5.3. Information Criteria

Information Criteria	Values
Log Likelihood	-234.1150
Akaike	474.2300
Schwarz	236.6635
Hannan Quinne	235.4961

3.6. For n=30, p =10, and k = 5.

3.6.1. Rotated Factor Loadings

Varimax Loadings					Equamax Loadings				
1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
-0.2795	0.0571	-0.0315	0.8106	0.0238	-0.2534	0.0564	-0.0282	0.8188	0.0379
0.9894	-0.0043	-0.0118	-0.1384	-0.0412	0.9817	-0.0036	-0.0120	-0.1680	-0.0893
0.1089	0.0651	-0.1572	0.1962	-0.0836	0.1107	0.0641	-0.1569	0.1937	-0.0880
-0.0070	0.9996	-0.0084	0.0240	0.0014	-0.0078	0.9996	-0.0083	0.0250	-0.0090
-0.0104	-0.0081	0.9997	-0.0210	0.0008	-0.0122	-0.0082	0.9996	-0.0248	-0.0066
-0.0680	0.0455	0.0348	0.0569	0.6576	-0.0343	0.0526	0.0403	0.0571	0.6594
-0.2466	-0.1801	-0.1940	0.2114	0.2730	-0.2260	-0.1775	-0.1911	0.2187	0.2884
-0.2962	0.1142	0.0306	-0.3878	-0.0122	-0.3082	0.1140	0.0287	-0.3786	0.0003
0.1796	0.1563	0.1081	0.1366	-0.3720	0.1649	0.1523	0.1058	0.1317	-0.3826
-0.0854	-0.1362	-0.0856	0.1554	-0.1272	-0.0864	-0.1378	-0.0860	0.1585	-0.1205
Quartimax Loadings					Orthomax Loadings				
1	2	3	4	5	1	2	3	4	5
-0.2872	0.0573	-0.0325	0.8079	0.0210	-0.2816	0.0631	-0.0282	0.8083	0.0505
0.9911	-0.0047	-0.0120	-0.1292	-0.0295	0.9815	-0.0059	-0.0130	-0.1312	-0.1384
0.1081	0.0653	-0.1573	0.1970	-0.0823	0.0989	0.0642	-0.1581	0.1970	-0.0925
-0.0068	0.9997	-0.0085	0.0237	0.0039	-0.0102	0.9994	-0.0085	0.0172	-0.0254
-0.0098	-0.0080	0.9997	-0.0197	0.0025	-0.0122	-0.0085	0.9994	-0.0246	-0.0182
-0.0763	0.0438	0.0335	0.0559	0.6570	-0.0033	0.0641	0.0484	0.0562	0.6588
-0.2519	-0.1807	-0.1947	0.2087	0.2693	-0.2189	-0.1711	-0.1876	0.2116	0.3048
-0.2923	0.1143	0.0313	-0.3905	-0.0156	-0.1936	0.1110	0.0290	-0.3907	0.0133
0.1828	0.1572	0.1085	0.1386	-0.3691	0.1402	0.1469	0.1010	0.1363	-0.3942
-0.0854	-0.1358	-0.0855	0.1546	-0.1286	-0.0980	-0.1386	-0.0876	0.1560	-0.1125

3.6.2. Factor Rotation Matrix

Varimax

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	0.9894	-0.0058	-0.0127	-0.1384	-0.0412
Factor 2	0.0087	0.9994	-0.0252	0.0222	0.0008
Factor 3	0.0098	0.0257	0.9994	-0.0230	-0.0000
Factor 4	0.1395	-0.0226	0.0220	0.9894	0.0242
Factor 5	0.0374	-0.0004	-0.0010	-0.0297	0.9989

Equamax

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	0.9816	-0.0051	-0.0129	-0.1680	-0.0893
Factor 2	0.0079	0.9993	-0.0251	0.0228	-0.0103
Factor 3	0.0078	0.0256	0.9992	-0.0275	-0.0088
Factor 4	0.1704	-0.0229	0.0265	0.9846	0.0185
Factor 5	0.0850	0.0105	0.0069	-0.0334	0.9957

Quartimax

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	0.9911	-0.0062	-0.0129	-0.1292	-0.0295
Factor 2	0.0089	0.9994	-0.0253	0.0220	0.0034
Factor 3	0.0104	0.0257	0.9994	-0.0216	0.0021
Factor 4	0.1300	-0.0225	0.0206	0.9907	0.0264
Factor 5	0.0257	-0.0031	-0.0029	-0.0300	0.9992

Orthomax

	Factor 1	Factor 2	Factor 3	Factor 4	factor 5
Factor 1	0.9815	-0.0074	-0.0139	-0.1312	-0.1384
Factor 2	0.0054	0.9992	-0.0253	0.0155	-0.0272
Factor 3	0.0077	0.0251	0.9991	-0.0267	-0.0220
Factor 4	0.1342	-0.0152	0.0261	0.9904	0.0106
Factor 5	0.1361	0.0272	0.0192	-0.0291	0.9897

3.6.3. Information Criteria

Information Criteria	Values
Log Likelihood	-230.0950
Akaike	470.1900
Schwarz	234.3424
Hannan Quinne	232.3969

3.7. For n=70, p =10 and k = 2

3.7.1. Rotated Factor Loadings

Varimax Loadings		Equamax Loadings		Quartimax Loadings		Orthomax Loadings	
I	II	I	II	I	II	I	II
1.0000	-0.0057	1.0000	-0.0048	1.0000	-0.0059	1.0000	-0.0037
-0.0818	-0.1573	-0.0816	-0.1574	-0.0818	-0.1573	-0.0815	-0.1575
0.0385	0.1685	0.0383	0.1685	0.0385	0.1685	0.0381	0.1686
-0.0663	0.0006	-0.0663	0.0005	-0.0663	0.0006	-0.0663	0.0004
-0.2759	-0.1688	-0.2757	-0.1691	-0.2759	-0.1688	-0.2755	-0.1694
-0.1151	-0.0390	-0.1151	-0.0391	-0.1152	-0.0390	-0.1151	-0.0393
-0.0401	0.0629	-0.0401	0.0629	-0.0401	0.0630	-0.0402	0.0629
0.1946	0.1182	0.1945	0.1184	0.1947	0.1182	0.1944	0.1186
0.1780	-0.0731	0.1780	-0.0730	0.1779	-0.0732	0.1781	-0.0728
-0.0041	1.0000	-0.0050	1.0000	-0.0038	1.0000	-0.0061	1.0000

3.7.2. Factor Rotation Matrix

Varimax

	Factor I	Factor II
Factor I	-0.0279	0.9996
Factor II	0.9996	0.0279

Equamax

	Factor I	Factor II
Factor I	-0.0288	0.9996
Factor II	0.9996	0.0288

Quartimax

	Factor I	Factor II
Factor I	-0.0277	0.9996
Factor II	0.9996	0.0277

Orthomax

	Factor I	Factor II
Factor I	-0.0299	0.9996
Factor II	0.9996	0.0299

3.7.3. Information Criteria

Information Criteria	Values
Log Likelihood	-326.9945
Akaike	657.9890
Schwarz	328.8396
Hannan Quinne	328.0586

3.8. For n =70 , p =10, and k =3

3.8.1. Rotated Factor Loadings

Varimax Loadings			Equamax Loadings			Quartimax Loadings			Orthomax Loadings		
Factor I	Factor II	Factor III	Factor I	Factor II	Factor III	Factor I	Factor II	Factor III	Factor I	Factor II	Factor III
0.9989	-0.0199	-0.0026	0.9998	-0.0200	0.0047	0.9998	-0.0198	-0.0044	0.9996	-0.0202	0.0176
-0.0792	0.1089	-0.1916	-0.0778	0.1101	-0.1920	-0.0796	0.1097	-0.1915	-0.0753	0.1087	-0.1937
0.0384	0.0163	0.1952	0.0369	0.1060	0.1955	0.0387	0.1064	0.1951	0.0344	0.1074	0.1959
-0.0707	-0.2210	0.0189	-0.0709	-0.2210	0.0180	-0.0707	-0.2210	0.0191	-0.0711	-0.2209	0.0187
-0.2750	0.0242	-0.2080	-0.2735	0.0246	-0.2100	-0.2754	0.0241	-0.2076	-0.2708	0.0232	-0.2137
-0.1164	-0.0678	-0.0494	-0.1161	-0.0677	-0.0504	-0.1165	-0.0679	-0.4092	-0.1154	-0.0681	-0.0514
-0.0205	0.9995	0.0251	-0.0205	0.9994	0.0266	-0.0205	0.9995	0.0248	-0.0206	0.9996	0.0109
0.1947	0.0172	0.1319	0.1937	0.0169	0.1334	0.1949	0.0172	0.1316	0.1920	0.0178	0.1357
0.1810	0.1431	-0.1012	0.1818	0.1433	-0.0997	0.1808	0.1431	-0.1016	0.1830	0.1425	-0.0983
-0.0067	0.0419	0.8396	-0.0129	0.0405	0.8396	-0.0052	0.0422	0.2933	-0.0236	0.0465	0.8391

3.8.2. Factor Rotation Matrix

Varimax

	Factor I	Factor II	Factor III
Factor I	-0.179	0.9848	0.0251
Factor II	0.9851	0.1719	0.0022
Factor III	0.0021	-0.0251	0.9997

Equamax

	Factor I	Factor II	Factor III
Factor I	-0.1720	0.9848	0.0254
Factor II	0.9851	0.1718	0.0097
Factor III	-0.0052	-0.0267	0.9996

Quartimax

	Factor I	Factor II	Factor III
Factor I	-0.1719	0.9848	0.0250
Factor II	0.9851	0.1720	0.0004
Factor III	0.0039	-0.0247	0.9997

Orthomax

	Factor I	Factor II	Factor III
Factor I	-0.1721	0.9850	0.0162
Factor II	0.9849	0.1717	0.0211
Factor III	-0.0180	-0.0196	-0.9996

3.8.3. Information Criteria

Information Criteria	Values
Log Likelihood	-323.7360
Akaike	653.4720
Schwarz	326.5036
Hannan Quinne	325.3321

3.9. For n=70, p =10, and k = 5

3.9.1. Rotated Factor Loadings

Varimax Loadings					Equamax Loadings				
1	2	3	4	5	1	2	3	4	5
0.9966	-0.0598	-0.0005	0.0554	-0.0076	0.9949	-0.0623	0.0128	0.0775	-0.0141
-0.0639	0.0848	-0.2353	-0.1824	0.2870	-0.0544	0.0854	-0.2368	-0.1856	0.2854
0.0365	-0.0451	0.2044	-0.0216	0.0411	0.0346	-0.0441	0.2050	-0.0222	0.0405
-0.0659	-0.0353	-0.0008	-0.0990	-0.3618	-0.0664	-0.0346	-0.0020	-0.0967	-0.3624
-0.2767	-0.0161	-0.2359	0.0071	0.0808	-0.2732	-0.0166	-0.2393	0.0006	0.0830
-0.0573	0.9977	-0.0142	0.0324	-0.0085	-0.0559	0.9975	-0.0193	0.0387	-0.0082
-0.0487	-0.0687	0.0752	0.1504	0.5611	-0.0490	-0.0690	0.0758	0.1423	0.5631
0.2154	0.0999	0.1252	-0.2546	0.0801	0.2203	0.1019	0.1268	-0.2502	0.0755
0.1457	0.0920	-0.0539	0.7086	0.0813	0.1312	0.0862	-0.0501	0.7116	0.0882
-0.0080	-0.0265	0.7554	-0.0507	0.0269	-0.0164	-0.0227	0.7552	-0.0540	0.0258
Quartimax Loadings					Orthomax Loadings				
1	2	3	4	5	1	2	3	4	5
0.9969	-0.0592	-0.0036	0.0504	-0.0063	0.9943	-0.0633	0.0358	0.0766	-0.0165
-0.0660	0.0847	-0.2348	-0.1817	0.2878	-0.0484	0.0799	-0.2312	-0.1917	0.2887

Varimax Loadings					Equamax Loadings				
1	2	3	4	5	1	2	3	4	5
0.0370	0.0454	0.2043	-0.0214	0.0412	0.0300	-0.0447	0.2066	-0.0267	0.0317
-0.0659	-0.0355	-0.0005	-0.0996	-0.3616	-0.0677	-0.0317	-0.0182	-0.0839	-0.3651
-0.2775	-0.0160	-0.2351	0.0085	0.0804	-0.2675	-0.0171	-0.2421	0.0028	0.0925
-0.0577	0.9977	-0.0129	0.0308	-0.0085	-0.0533	0.9980	-0.0207	0.0266	0.0045
-0.0485	-0.0686	0.0751	0.1524	0.5606	-0.489	-0.0732	0.0974	0.1230	0.5637
0.2143	0.0994	0.1249	-0.2555	0.0811	0.2174	0.0976	0.1294	-0.2566	0.0628
0.1490	0.0934	-0.0549	0.7079	0.0795	0.1338	0.0938	-0.0375	0.7077	0.1137
-0.0060	-0.0274	0.7554	-0.0499	0.0271	-0.0332	-0.0230	0.7539	-0.0680	-0.0034

3.9.2. Factor Rotation Matrix

Varimax.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	0.9966	-0.0907	-0.0000	0.0542	-0.0073
Factor 2	0.0885	0.9951	-0.0143	0.0407	-0.0096
Factor 3	0.0236	0.0266	0.9130	-0.3987	-0.0788
Factor 4	-0.0446	-0.0208	0.3988	0.8350	0.3759
Factor 5	0.0290	0.0204	-0.0846	-0.3731	0.9232

Equamax

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	0.9926	-0.0932	0.0133	0.0759	-0.0138
Factor 2	0.0896	0.9945	-0.0175	0.0503	-0.0104
Factor 3	0.0204	0.0336	0.9118	-0.4000	-0.0842
Factor 4	-0.0657	-0.0249	0.4012	0.8280	0.3853
Factor 5	0.0453	0.0233	-0.0848	-0.3822	0.9188

Quartimax

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	0.9947	-0.0901	-0.0032	0.0493	-0.0060
Factor 2	0.0882	0.9952	-0.0135	0.0384	-0.0095
Factor 3	0.0244	0.0249	0.9133	-0.3982	-0.0776
Factor 4	-0.0397	-0.0198	0.3981	0.8367	0.3735
Factor 5	0.0255	0.0197	-0.0844	-0.3708	0.9243

Orthomax

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	0.9919	-0.0942	0.0363	0.0754	-0.0166
Factor 2	0.0922	0.9949	-0.0155	0.0380	0.0021
Factor 3	-0.0008	0.0302	0.9006	-0.4137	-0.1298
Factor 4	-0.0722	-0.0183	0.4289	0.8078	0.3975
Factor 5	0.494	0.0083	-0.0584	-0.4114	0.9082

3.9.3. Information Criteria

Information Criteria	Values
Log Likelihood	-319.8686
Akaike	649.7370
Schwarz	324.4814
Hannan Quinne	322.5287

4. Discussion of Results and Conclusion

When the sample size (n) is 30 and the number of variables (p) is 10, the values of the Akaike (AIC) and Schwarz (SIC) Information criteria for different number of retained factors are as follows: for k = 2, AIC, SIC and HQIC values are 237.2530, 118.1036, and 117.304 respectively; for k = 3 AIC is 231.2490, SIC is 114.8402 and HQIC is 109.9072; for k = 5, AIC is 226.4260, SIC is 111.9058 and HQIC is 109.9072.

On increasing the sample size to 50 and retaining the same number of variables i.e. 10; for k =2, AIC is 479.6350, SIC is 239.6635 and HQIC is 238,7380; for k = 3, AIC is 474.2300, SIC is 236.6635 and HQIC is 235.4961 and finally if k = 5, AIC is 470.1910, SIC is 234.3424 and HQIC is 232.3969.

Finally, when the sample size is increased to 70 and the same number of variables maintained, for k = 2, AIC is 657.9890, SIC is 328.8396 and HQIC is 328.0586; for k = 3, AIC is 653.4720. SIC is 326.5036 and HQIC is 325.3321; then for k = 5, AIC is 649.7370, SIC is 324.4814 and HQIC is 322.5287.

The factor rotation matrix for all the sample sizes and the number of parameters retained as considered is almost the same for all the four methods of rotation considered here.

In conclusion based on the results above, it shows that the values for both AIC, SIC and HQIC all decreases for all the sample sizes considered as k increases. Again, as we moved from on sample size to the other, the values of the AIC, SIC and HQIC increases as the sample size increases for all the number of retained factors considered. Also, since in the information criteria model selection, the model with the least values of AIC, SIC and HQIC is considered as the best model. Here, the AIC, the SIC and the HQIC values are least for k =5 for all the sample sizes considered. Hence, the optimal number of factors to using the Maximum Likelihood Estimation method based on the simulated data used is five (5).

References

[1] Akaike, H. (1973). Information Theory and Extension of the Maximum Likelihood Principle; Second International Symposium on Information Theory (B.N. Petrov and F. Csaki, Eds.). *Budapest Hungary: Akademia Kiado,* 267-281

[2] Comrey, A.L., and Lee, H.B. (1992). *A First Course in Factor Analysis* (2nd Ed.). Hillsdale, NJ: Erlbaum.

[3] Harman, H.H. (1976). *Modern Factor Analysis (3rd Ed.).* Chicago: The University of Chicago Press.

[4] Johnson, R.A., and Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis* (*6th Ed.)*. Prentics Hall, Englewood Cliffs, New Jersey.

[5] Onyeagu, S.I. (2003). *A First Course in Multivariate Statistical Analysis (1st Ed.)*. Mega Concept Publishers.

[6] Schwarz,G.(1978). Estimating the Dimension of a Model. *Annals of Statistics 6*, 461-464.