Penalized varimax

Nickolay T. Trendafilov and Doyo Gragn
Department of Mathematics and Statistics, The Open University,
Walton Hall, Milton Keynes MK7 6AA, UK

Abstract

A common weakness of all analytical methods for simple structure rotation is that the rotated factors typically have unequal sum of squared loadings, which may spoil their interpretation. In this paper, a modified varimax criterion is introduced by attaching a penalty term to the original varimax objective function. The penalized varimax solutions have equal sum of squared loadings for all factors. The method is applied to well studied data sets and the results are compared to the classical varimax solutions.

Keywords: simple structure, orthogonal rotation, projected gradients.

1 Introduction

The analytic rotation methods have a long history in exploratory factor analysis. Browne (2001) gives a very complete and comprehensive overview of the field. Details can be found in the cited papers there and in the standard texts on factor analysis, e.g. (Harman, 1976; Mulaik, 1972).

A common weakness of all analytical methods for simple structure rotation is that the rotated factors are usually unequally loaded, which may spoil their interpretation. For instance, the quartimax rotation tends to produce solutions with a dominating factor (Harman, 1976; Knüsel, 2008). Such factor is overloaded by larger loadings, and hence has much higher sum of squared loadings, compared to the remaining factors. On the other hand, the varimax solution has a tendency towards equal sum of squared loadings for all factors. This probably explains the great success of the varimax criterion. Unfortunately, the existing varimax algorithms do not try to achieve this optimal property explicitly. Recently Knüsel (2008) showed that, indeed, in theory the varimax solution should have equal sum of squared loadings for all factors. As it is well-known (Harman, 1976; Mulaik, 1972), ideally the varimax criterion has maximum value when there is a single unit loading per factor and all the rest are 0s, which also implies equal (to 1) sum of squares per factor. The Thurstone's simple structure criteria (Thurstone, 1947, p.335) also suggest for equidistributed zeros across the rows and the columns of the rotated loading matrix.

In this paper, a modified varimax criterion is introduced by attaching a penalty term to the original varimax objective function. The penalty term explicitly controls the size of the column sums of squared loadings, by 'equi-distributing the load' from the overloaded factors to the less-loaded factors. As a result, the penalized varimax solution has equal sum of squared loadings for all factors. The penalized varimax is designed as a supplement to the classical varimax for treating problems with unsatisfactory simple structure possibly caused by uneven sum-of-squares per factor.

The paper is organized as follows. A formulation of the varimax rotation problem and a list of the most popular algorithms for its solution is given in Section 2. Definitions of the penalized varimax problem is proposed in Section 3. It is solved by a matrix algorithm making use of projected gradient method (Jennrich, 2001; Trendafilov, 2006). The matrix algorithm directly finds an orthogonal rotation matrix to produce the penalized varimax solution and is considered in Section 3.2. If the penalty term is switched off, the algorithm simply turns into a standard varimax rotation.

The method is applied on three benchmark data sets: the Five Socio-economic variables (Harman, 1976, p.135), the 24 Psychological Tests data (Harman, 1976, p.123 and p.215) and the Thurstone's Box data (Thurstone, 1947, p.370). The results are compared to the classical varimax solutions. It is demonstrated that if the application of the penalized varimax is reasonable, it can provides clearer simple structure than the standard varimax solution.

2 Varimax criterion

Varimax (Kaiser, 1958) is the most popular method for analytical rotation in factor analysis. Let $\mathbf{A}$ be the initial $p \times q$ factor loadings matrix and $\mathbf{B} = \mathbf{AQ}$ be an orthogonally rotated factor loadings matrix. The variance of the squared loadings of the jth rotated factor is:

$$V_j = \sum_{i=1}^{p} b_{ij}^4 - \frac{1}{p}\left(\sum_{i=1}^{p} b_{ij}^2\right)^2 . \tag{1}$$

The variance V_j will be large when there are few large squared loadings and all the rest are near zero. The variance V_j will be small when all squared loadings have nearly same value. The varimax rotation problem (Kaiser, 1958) is to find a $q \times q$ orthogonal matrix $\mathbf{Q}$ such that the total variance of all q factors is maximized , i.e. maximize

$$V = \sum_{j=1}^{q} V_j = \sum_{j=1}^{q} \left[\sum_{i=1}^{p} b_{ij}^4 - \frac{1}{p} \left(\sum_{i=1}^{p} b_{ij}^2 \right)^2 \right] \ . \qquad (2)$$

The original algorithm to find the varimax rotation $\mathbf{Q}$ proposed by Kaiser (1958) makes use of successive planar rotations of all possible $q(q-1)/2$ pairs of factors, such that each pair has maximum variance.

The varimax rotation problem can be defined in matrix form as follows (Magnus and Neudecker, 1988; Sherin, 1966). Compose the matrix:

$$\mathbf{S}(\mathbf{Q}) = \mathbf{C}^\top \mathbf{M}_p \mathbf{C} \text{ with } \mathbf{M}_p = \left(\mathbf{I}_p - \frac{\mathbf{1}_p \mathbf{1}_p^\top}{p} \right) \ , \qquad (3)$$

where $\mathbf{C} = \mathbf{B} \odot \mathbf{B}$ and $\odot$ denotes the Hadamard (elementwise) matrix product. If $\mathbf{S}$ in (3) is divided by $p-1$ it presents the (sample) covariance matrix of the squared orthogonally transformed factor loadings $\mathbf{B}$. Then the varimax problem is to maximize the following objective function (criterion):

$$V(\mathbf{Q}) = \text{trace } \mathbf{S}(\mathbf{Q}) \ , \qquad (4)$$

over all possible orthogonal rotations $\mathbf{Q} \in \mathcal{O}(q)$, i.e.:

$$\max_{\mathbf{Q} \in \ \mathcal{O}(q)} \text{trace } \mathbf{S}(\mathbf{Q}) \ . \qquad (5)$$

A number of different algorithms are available for solving the varimax problem, e. g. (Jennrich, 1970; Kaiser, 1958; Magnus and Neudecker, 1988; Mulaik, 1972; Sherin, 1966; ten Berge, 1984).

3 Varimax with equal column sums of squares

For orthogonal rotation $\mathbf{Q}$, the sum of the squared initial loadings equals the sum of the squared rotated loadings. Indeed, trace$(\mathbf{A}^\top \mathbf{A})$ = trace$(\mathbf{A}\mathbf{A}^\top)$ = trace$(\mathbf{A}\mathbf{Q}\mathbf{Q}^\top \mathbf{A}^\top)$ = trace$(\mathbf{B}\mathbf{B}^\top)$ = trace$(\mathbf{B}^\top \mathbf{B})$, i.e. the value of trace$(\mathbf{B}^\top \mathbf{B})$ is constant.

The aim of the paper is to construct an algorithm that finds loadings both maximizing the varimax criterion and having equal sums of squares across the factors. In other words, the rotated loadings should posses the property:

$$\mathbf{b}_1^\top \mathbf{b}_1 = ... = \mathbf{b}_q^\top \mathbf{b}_q \ . \qquad (6)$$

To achieve this one should solve the original varimax problem (5) subject to the additional constraint (6). Alternative way to impose the additional constraint (6) on the varimax solution is to modify the varimax criterion by adding a term penalizing the deviation from (6). This modified varimax problem is called the penalized varimax. The next Section deals with the construction of such penalty term.

3.1 Penalizing unequal column sums of squares of B

Consider the Lagrange's identity:

$$q\sum_{j=1}^{q} x_j^2 = \left(\sum_{j=1}^{q} x_j\right)^2 + \sum_{1\leq j<k\leq q} (x_j - x_k)^2 \ , \tag{7}$$

where $x_1, x_2, ...x_q$ are some non-negative numbers. Apparently, if $\sum_{j=1}^{q} x_j$ is constant, then $\sum_{j=1}^{q} x_j^2$ is minimized when $x_j = x_k$ for any $1 \leq j < k \leq q$, and vice verse. By substituting $x_j = \mathbf{b}_j^\top \mathbf{b}_j$ in (7), one can see that the minimization of the following expression

$$\sum_{j=1}^{q} (\mathbf{b}_j^\top \mathbf{b}_j)^2 = \mathbf{1}_p^\top \mathbf{C}\mathbf{C}^\top \mathbf{1}_p \geq 0 \ , \tag{8}$$

equates the column sums of squares of the rotated loading matrix $\mathbf{B}$. It also follows from (7), that the $\mathbf{1}_p^\top \mathbf{C}\mathbf{C}^\top \mathbf{1}_p$ is bounded below by $\frac{(\text{trace}\mathbf{A}^\top\mathbf{A})^2}{q}$ and it is achieved when $\mathbf{b}_1^\top \mathbf{b}_1 = ... = \mathbf{b}_q^\top \mathbf{b}_q$.

Define the following penalty term as

$$\mathcal{P}(\mathbf{Q}) = \mathbf{1}_p^\top \mathbf{C}\mathbf{C}^\top \mathbf{1}_p - \frac{(\text{trace}\mathbf{A}^\top\mathbf{A})^2}{q} \ , \tag{9}$$

which is a nonnegative continuous function of the rotation matrix $\mathbf{Q} \in \mathcal{O}(q)$. As $\mathcal{O}(q)$ is compact, then there exists $\mathbf{Q}$ at which $\mathcal{P}(\mathbf{Q})$ achieves its minimum 0. Thus, $\mathcal{P}(\mathbf{Q})$ penalizes unequal column sums of squares of the rotated loading matrix $\mathbf{B}$ and vanishes when $\mathbf{b}_1^\top \mathbf{b}_1 = ... = \mathbf{b}_q^\top \mathbf{b}_q$. One can easily see that in fact $\mathcal{P}(\mathbf{Q})$ penalizes the total deviation of all column sums of squares of $\mathbf{B}$ from their mean value. Indeed, let for $j = 1, ..., q$ denote $b_{.j}^2 = \sum_i^p b_{ij}^2$ and $b_{..}^2 = \frac{1}{q}\sum_j^q b_{.j}^2 = \frac{1}{q}\sum_j^q (\sum_i^p b_{ij}^2)$. Then

$$\sum_{j=1}^{q} \left(b_{.j}^2 - b_{..}^2\right)^2 = \sum_{j}^{q} \left(\sum_{i}^{p} b_{ij}^2 - \frac{\sum_i^q \sum_j^p b_{ij}^2}{q}\right)^2$$

$$
\begin{aligned}
&= \sum_j^q \left(\mathbf{b}_j^\top \mathbf{b}_j - \frac{\mathrm{trace}\mathbf{B}^\top\mathbf{B}}{q} \right)^2 \\
&= \sum_j^q \left(\mathbf{b}_j^\top \mathbf{b}_j \right)^2 - 2\frac{\mathrm{trace}\mathbf{B}^\top\mathbf{B}}{q} \sum_j^q \mathbf{b}_j^\top \mathbf{b}_j + q \left[\frac{\mathrm{trace}\mathbf{B}^\top\mathbf{B}}{q} \right]^2 \\
&= \mathbf{1}_p^\top \mathbf{C}\mathbf{C}^\top \mathbf{1}_p - \frac{1}{q} \left(\mathrm{trace}\mathbf{A}^\top\mathbf{A} \right)^2
\end{aligned}
$$

since $\sum_j^q \mathbf{b}_j^\top \mathbf{b}_j = \mathrm{trace}\mathbf{B}^\top\mathbf{B} = \mathrm{trace}\mathbf{A}^\top\mathbf{A}$.

3.2 Penalized varimax criterion

Consider the following penalized varimax criterion:

$$PV(\mathbf{Q}) = \mathrm{trace}\ \mathbf{C}^\top \mathbf{M}_p \mathbf{C} - \mu \mathcal{P}(\mathbf{Q}) \ , \tag{10}$$

where μ is a large positive number and the penalty term $\mathcal{P}(\mathbf{Q})$ was derived in (9) of the previous Section. As with the standard varimax, by maximizing the PV criterion (11) the loadings are forced to get either small values around 0 or values near 1 or -1, but having as equal as possible sums of squared loadings of all factors. The importance of the penalty term is controlled by varying μ. Low values of μ will result in solutions close to the original variamax ones, while large values of μ can suppress entirely the varimax maximization and result in $\mathbf{B}$ with equal column sums of squares.

As the penalty term $\mathcal{P}(\mathbf{Q})$ in (10) contains a constant term which will not be affected by the maximization process, it seems more reasonable and cheaper to work with the following penalized varimax criterion:

$$PV(\mathbf{Q}) = \mathrm{trace}\ \mathbf{C}^\top \mathbf{M}_p \mathbf{C} - \mu \mathbf{1}_p^\top \mathbf{C}\mathbf{C}^\top \mathbf{1}_p \ . \tag{11}$$

The penalized varimax criterion $PV(\mathbf{Q})$ in (11) is in matrix form. The penalized varimax problem requires solving the following constrained maximization problem:

$$\max_{\mathbf{Q} \in \ \mathcal{O}(q)} PV(\mathbf{Q}) \ . \tag{12}$$

The problem (12) can be readily solved by orthogonal rotation algorithm `varimaxP` based on the dynamical system approach proposed by Trendafilov (2006). For this reason the gradient of $PV(\mathbf{Q})$ is needed, which, in turn, requires a smooth approximation of the penalty term. Same results are obtained by iterative implementation of the gradient projection algorithm of (Jennrich, 2001). As the penalized varimax function $PV(\mathbf{Q})$ is more complicated, alternatively one can rely on the derivative free version of the gradient projection algorithm (Jennrich, 2004).

Straightforward manipulations (Magnus and Neudecker, 1988) give the gradient of trace $\mathbf{C}^\top \mathbf{M}_p \mathbf{C}$ as:

$$4\mathbf{A}^\top \left(\mathbf{B} \odot (\mathbf{M}_p \mathbf{C})\right) \ , \tag{13}$$

and the gradient of the penalty in (11), as:

$$4\mathbf{A}^\top \left(\mathbf{B} \odot (\mathbf{1}_p \mathbf{1}_p^\top \mathbf{C})\right) \ . \tag{14}$$

Then, the gradient of the objective function $PV(\mathbf{Q})$ is:

$$4\mathbf{A}^\top \left\{\mathbf{B} \odot \left[\left(\mathbf{M}_p - \mu \mathbf{1}_p \mathbf{1}_p^\top\right) \mathbf{C}\right]\right\} \ . \tag{15}$$

4 Numerical examples and comparisons

An idea about the behavior of the penalized varimax can be given with the following small artificial example. Consider the loading matrix given in the first two columns of Table 1. Strictly speaking, such loading matrix has nearly perfect simple structure, and does not need rotation at all. The only unsatisfied condition for perfect simplicity is that the second column has lass 0s than factors (Thurstone, 1947, p.335). Applying the standard varimax algorithm has no effect, the loadings are left unrotated. Then the `varimaxP` algorithm is applied for different μ. For $\mu \in [0, 2.5]$, `varimaxP` also leaves the loadings unrotated. After further increasing of μ, the penalty term becomes more important than the varimax term, as seen in the next columns of Table 1. Finally, one ends up with the worst possible simple structure solution given in the last two columns of Table 1. This example is artificial and unlikely to happen in practice, but it shows that the penalized varimax should not be applied automatically. Using inappropriate μ may lead to unsatisfactory loadings. In reality, loading matrices composed by 0s and ± 1s only are very hard to find and impossible to achieve by orthogonal rotation. In general, the penalized varimax is expected to produce factors with balanced contributions to the total variance of the solution, while keeping well its simple structure.

Insert Table 1 about here.

The simple structure rotation of the first two principal components of the Five Socio-economic variables (Harman, 1976, p.135) gives a more realistic example illustrating the same potential problem with the penalized varimax. The standard varimax solution is given in the first two columns of Table 2. The loadings have pretty good simple structure. The column sums of squares are of quite similar magnitude (2.15/2.52 = .85). After these observations are made, the application of the penalized varimax seems unreasonable: there is no much room to neither improve nor spoil the varimax solution. Nevertheless, for illustration purposes the `varimaxP` algorithm is applied with three different $\mu = 1, 5, 10$ and they are depicted in Table 2.

Insert Table 2 about here.

Three varimax rotations (without Kaiser's normalization) are applied to the maximum likelihood solution for the 24 Psychological Tests (Harman, 1976, p.215), called for short 24HH data. The value of the varimax criterion for this initial solution is 0.6249. The first four columns of Table 3 are obtained by the classical varimax rotation algorithm based on plane rotations and implemented in `MATLAB` (MATLAB, 2009). No local maxima are found for 100 random runs. The value of the varimax criterion is 2.5110. For 100 runs, the `varimaxP` algorithm without penalty ($\mu = 0$) produces exactly the same loadings (not depicted) as the `MATLAB` one and no local maxima.

The first factor of the varimax solution is overloaded. Then the `varimaxP` algorithm with $\mu = 20$ is applied and the solution is given in the last four columns in Table 3. The value of the penalized varimax criterion for this solution is -654.9085, and the value of the varimax criterion is 2.2326. For the HH24 data, $\frac{1}{4}(\text{trace}\mathbf{A}^\top\mathbf{A})^2 = 32.8570$, which value is achieved by the penalty term for the depicted solution, as $\mathbf{1}_q^\top\mathbf{C}\mathbf{C}^\top\mathbf{1}_q = 32.8570$. For 100 runs, the `varimaxP` algorithm with $\mu = 20$ produces no local maxima.

Insert Table 3 about here.

The loadings greater than .4 of the solutions depicted in Table 3 are given in bold typeface. The simple structure of the `varimaxP` ($\mu = 20$) solution is clearer that the one following from the standard varimax (`MATLAB`). In fact, the `varimaxP` simple structure almost exactly matches (except $b_{21,3}$) the simple structure obtained in (Browne, 2001, p.133) from *oblique* rotation.

Next, the same varimax rotation algorithms (without Kaiser's normalization) are applied to the Thurstone's 26 Box problem (Thurstone, 1947), called for short 26 Box data. The initial solution to be analyzed in the sequel is composed by the first three principal components extracted by Cureton and Mulaik (1975) from the correlation matrix of the 26 Box data (Thurstone, 1947, p.370). The value of the varimax criterion for this initial principal component solution is 6.1017.

The first three columns in Table 4 are the varimax solution for the 26 Box problem obtained by the `MATLAB` algorithm (MATLAB, 2009). The maximum of the varimax objective function is 6.2365. No local maxima are found within 100 random runs. The `varimaxP` algorithm without penalty ($\mu = 0$) produces exactly the same loadings (not depicted) as the `MATLAB` ones and no local maxima for 100 runs.

The first factor of the standard varimax solution of the 26 Box data (first three columns of Table 4) is considerably overloaded. This is a clear indication to apply the penalized varimax. The next three columns in Table 4 are obtained by the `varimaxP` algorithm with $\mu = 20$. The value of the penalized varimax criterion for this solution is -4298.6981, and the value of the varimax criterion is 5.5309. For the 26 Box data, $\frac{1}{4}(\text{trace}\mathbf{A}^\top\mathbf{A})^2 = 215.2114$, which value is achieved by the penalty term for the depicted solution, i.e. $\mathbf{1}_q^\top\mathbf{C}\mathbf{C}^\top\mathbf{1}_q = 215.2115$. For 100 runs, the `varimaxP` algorithm with $\mu = 20$ produces no local maxima.

The standard varimax solution does not reveal the simple structure of the 26 Box data: the loadings are complete mess. It is easily noticeable that the penalized varimax solution has much more structured loadings. Moreover, it provides kind of a 'negative' (as in photography) of the simple structure of the 26 Box data.

The experienced difficulties with revealing the simple structure of the 26 Box data is a notorious problem with the varimax criterion. This is not surprising because the weighted varimax solution of Cureton and Mulaik (1975) which reveals the simple structure in the 26 Box data, has varimax value of only 5.3746. The orthogonal minimum entropy solution of the 26 Box problem reported by Browne (2001) also reveals its simple structure and has varimax value 5.4370. Clearly, all these 'successful' solutions are local maxima for the varimax criterion.

Such local maxima are also obtained by the penalized varimax. While experimenting with `varimaxP`, it was observed that new local maxima emerge if using very short integration step. One can get rid of them by increasing the required convergence accuracy, say from 10^{-4} to 10^{-7} between two consecutive varimax values. For 100 random runs, only one local maximum of the penalized varimax criterion was observed with value of -4298.86, which for the varimax criterion gives 5.37. Ironically, just this solution reconstructs well the simple structure of the 26 Box data and is given in the last three columns of Table 4.

Insert Table 4 about here.

5 Conclusion

A penalized version of the well known varimax orthogonal rotation method is proposed which produces loadings with squares having equal sums for all factors. Such factors are balanced and may give more adequate interpretation. The penalized varimax is proposed as a supplement/companion procedure to the standard varimax especially for factor solutions with considerably different sum of squares.

References

Browne, M. W. (2001) An overview of analystic rotation in exploratory factor analysis, *Multivariate Behavioral Research*, 36, 111–150.

Clarkson, D. B., and Jennrich, R. I. (1988) Quartic rotation criteria, *Psychometrika*, 53, 251–259.

Cureton, E. E. and Mulaik, S. A. (1975). The weighted varimax rotation and the promax rotation, *Psychometrika*, 40, 183-195.

Jennrich, R. I. (1970) Orthogonal rotation algorithms, *Psychometrika*, 35, 229–235.

Jennrich, R. I. (2001) A simple general procedure for orthogonal rotation, *Psychometrika*, 66, 289-306.

Jennrich, R. I. (2004) Derivative free gradient projection algorithms for rotation, *Psychometrika*, 69, 475-480.

Harman, H. H. (1976) *Modern faclor analysis*, 3rd Edition. Chicago: University of Chicago Press.

Kaiser, H. F. (1958) The varimax criterion for analytic rotation in factor analysis, *Psychometrika*, 23, 187–200.

Knüsel, L. (2008) Chisquare as a rotation criterion in factor analysis, *Computational Statistics and Data Analysis*, 52, 4243-4252.

Magnus, J.R., and Neudecker, H. (1988) *Matrix Differential Calculus with Application in Statistics and Econometrics*, New York: Wiley.

MATLAB R2009a (2009) *The MathWorks, Inc.*

Mulaik, S. A. (1972) *The Foundations of Factor Analysis*, New York: McGraw-Hill.

Sherin, R. J. (1966) A matrix formulation of Kaiser's varimax criterion, *Psychometrika*, 31, 535–538.

ten Berge, J.M.F. (1984) A joint treatment of varimax rotation and the problem of diagonalizing symmetric matrices simultaniously in the least-squares sense, *Psychometrika*, 49, 347–358.

Trendafilov, N. T. (2006) The dynamical system approach to multivariate data analysis, a review, *Journal of Computational and Graphical Statistics*, 15, 628–650.

Thurstone, L. L. (1947). *Multiple Factor Analysis.* Chicago, IL: University of Chicago Press.

Table 1: Limitations of the penalized varimax.

	Initial loadings		varimax (MATLAB)		varimaxP ($\mu = 2.6659$)		varimaxP ($\mu = 2.6669$)		varimaxP ($\mu = 2.7$)	
Var	I	II	I	II	I	II	I	II	I	II
1	1	0	1	0	1.00	.06	.86	-.50	.71	-.71
2	1	0	1	0	1.00	.06	.86	-.50	.71	-.71
3	0	1	0	1	.06	1.00	.50	.86	-.71	-.71
s.s.	2	1	2	1	2	1	1.75	1.25	1.5	1.5
varimax	1.33		1.33		1.31		.33		.00	

Table 2: Factor loadings for Five Socio-economic variables from two varimax algorithms.

	varimax (MATLAB)		varimaxP ($\mu = 1$)		varimaxP ($\mu = 5$)		varimaxP ($\mu = 10$)	
Var	I	II	I	II	I	II	I	II
1	.01	.99	-.03	.99	-.10	.99	-.12	.99
2	.94	-.00	.94	.04	.94	.10	.93	.12
3	.13	.98	.09	.99	.02	.99	-.00	.99
4	.82	.45	.80	.49	.77	.54	.75	.56
5	.97	-.00	.97	.04	.96	.11	.96	.13
s.s.	2.52	2.15	2.47	2.20	2.40	2.27	2.37	2.30
varimax	1.8684		1.8560		1.7885		1.7496	

Table 3: Factor loadings for HH24 data from two varimax algorithms.

	varimax (`MATLAB`)				`varimaxP` ($\mu = 20$)			
Var	I	II	III	IV	I	II	III	IV
1	.25	.15	**.68**	.13	.08	**.68**	.17	.25
2	.17	.06	**.43**	.08	.07	**.43**	.07	.16
3	.21	-.05	**.55**	.10	.08	**.56**	-.04	.19
4	.30	.07	**.50**	.05	.18	**.53**	.09	.16
5	**.76**	.21	.12	.07	**.69**	.23	.27	.23
6	**.80**	.07	.12	.16	**.72**	.23	.13	.32
7	**.83**	.15	.12	-.01	**.76**	.25	.21	.17
8	**.61**	.23	.29	.06	**.51**	.37	.28	.21
9	**.84**	.05	.11	.15	**.76**	.23	.11	.32
10	.17	**.85**	-.08	.08	.09	-.07	**.85**	.13
11	.22	**.53**	.13	.31	.09	.11	**.54**	.38
12	.05	**.70**	.26	.03	-.07	.24	**.70**	.08
13	.24	**.50**	**.45**	.02	.10	**.47**	**.52**	.13
14	.25	.12	.03	**.53**	.14	-.00	.12	**.57**
15	.18	.11	.11	**.50**	.06	.06	.10	**.53**
16	.16	.08	.40	**.51**	-.01	.35	.07	**.57**
17	.20	.26	.06	**.54**	.07	.01	.25	**.58**
18	.10	.35	.31	**.42**	-.07	.25	.34	**.47**
19	.20	.17	.23	.34	.08	.21	.18	.40
20	**.44**	.12	.36	.26	.31	.39	.14	.37
21	.23	**.43**	.39	.17	.09	.39	**.44**	.26
22	**.43**	.12	.36	.26	.30	.39	.14	.37
23	**.44**	.23	**.47**	.18	.29	**.50**	.26	.32
24	**.41**	**.51**	.15	.23	.29	.17	**.53**	.33
ss	4.35	2.69	2.62	1.81	2.87	2.86	2.86	2.86
varimax	2.5110				2.2326			

Table 4: Factor loadings for 26 Box data from two varimax algorithms.

	varimax(`MATLAB`)			`varimaxP`($\mu = 20$)			`varimaxP`($\mu = 20$)		
Vars	I	II	III	I	II	III	I	II	III
x_1	.61	-.22	.74	-.25	.71	.64	.98	-.04	.14
x_2	.69	.68	-.04	.63	-.09	.73	.28	.92	.12
x_3	.83	-.33	-.42	.73	.66	-.04	.14	.23	.94
x_1x_2	.81	.35	.44	.26	.34	.89	.77	.60	.13
x_1x_3	.90	-.38	.17	.32	.88	.33	.68	.10	.72
x_2x_3	.91	.22	-.34	.86	.32	.38	.20	.71	.66
$x_1^2x_2$	.78	.11	.59	.06	.54	.83	.91	.35	.16
$x_1x_2^2$	.80	.52	.22	.46	.16	.85	.57	.78	.14
$x_1^2x_3$	.83	-.35	.41	.10	.86	.46	.83	.03	.53
$x_1x_3^2$	.00	-.41	-.04	.52	.91	.23	.56	.17	.91
$x_2^2x_3$	.86	.41	-.26	.82	.17	.51	.22	.83	.49
$x_2x_3^2$	.91	.03	-.39	.85	.45	.24	.18	.57	.79
x_1/x_2	-.06	-.79	.61	-.70	.69	-.13	.55	-.83	.07
x_2/x_1	.06	.79	-.61	.70	-.69	.13	-.55	.83	-.07
x_1/x_3	-.15	.15	.96	-.77	.01	.61	.70	-.17	-.68
x_3/x_1	.15	-.15	-.96	.77	-.01	-.61	-.70	.17	.68
x_2/x_3	-.10	.95	.28	-.02	-.71	.70	.08	.68	-.73
x_3/x_2	.10	-.95	-.28	.02	.71	-.70	-.08	-.68	.73
$2x_1 + 2x_2$	.80	.43	.37	.32	.26	.89	.70	.67	.11
$2x_1 + 2x_3$	.90	-.40	.12	.34	.89	.28	.64	.09	.76
$2x_2 + 2x_3$	.91	.22	-.32	.85	.33	.40	.22	.72	.65
$(x_1^2 + x_2^2)^{1/2}$	.79	.42	.36	.32	.26	.87	.69	.66	.12
$(x_1^2 + x_3^2)^{1/2}$	.88	-.38	.10	.36	.86	.27	.61	.10	.74
$(x_2^2 + x_3^2)^{1/2}$	.90	.23	-.29	.82	.32	.41	.24	.71	.63
$x_1x_2x_3$	.98	.08	.11	.54	.57	.61	.62	.54	.55
$(x_1^2 + x_2^2 + x_3^2)^{1/2}$	.96	.14	.01	.61	.49	.57	.52	.59	.56
s.s.	14.79	5.55	5.08	8.47	8.47	8.47	8.47	8.47	8.47
varimax	6.2365			5.5309			5.3700		