DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

A MODEL FOR BAYESIAN FACTOR ANALYSIS WITH JOINTLY DISTRIBUTED MEANS AND LOADINGS

Daniel B. Rowe

SOCIAL SCIENCE WORKING PAPER 1108

January 2001

A Model for Bayesian Factor Analysis with Jointly Distributed Means and Loadings

Daniel B. Rowe[1]

Abstract

In the Bayesian approach to factor analysis, available prior knowledge regarding the model parameters is quantified in the form of prior distributions and incorporated into the inferences along with the data. The incorporation of prior knowledge has the added consequence of eliminating the ambiguity of rotation and the need for model constraints found in the traditional factor analysis model. A focus of recent work (Rowe, 2000a and Rowe, 2000b and Rowe, 2000c) has been on quantifying and incorporating available prior knowledge when estimating the population mean. This previous work has considered, vague, conjugate and generalized conjugate distributions for the population mean. In this paper, unlike previous work, the population mean vector and the factor loading matrix are taken to be jointly distributed which allows available interrelated prior information to be quantified and incorporated with the data. The model parameters are estimated by Gibbs sampling and iterated conditional modes algorithms.

[1]Correspondence and requests for reprints should be addressed to Daniel B. Rowe, Division of Humanities and Social Sciences, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA (e-mail: drowe@hss.caltech.edu).

1 Introduction

In the Bayesian approach to statistics, available prior knowledge regarding the model parameters is quantified in the form of prior distributions. Information contained in the data is quantified in the form of a likelihood distribution. The priors and the likelihood are combined by Bayes' rule so that knowledge from both sources is incorporated into the inferences. Bayesian statistical methods not only incorporate available prior information either from substantive experts or previous data, but allow the knowledge regarding the parameter values to accumulate as subsequent data is acquired.

In the non-Bayesian factor analysis model, the factor loading matrix is determinate up to an orthogonal rotation. Typically after a non-Bayesian factor analysis, an orthogonal rotation is performed on the factor loading matrix according to one of many subjective criteria. This is not the case in Bayesian factor analysis. The incorporation of prior knowledge has the added consequence of eliminating the ambiguity of rotation and the need for model constraints found in the traditional factor analysis model (Press & Shigemasu, 1989).

A focus of recent work (Rowe, 2000a and Rowe, 2000b and Rowe, 2000c) has been on quantifying and incorporating available prior knowledge when estimating the population mean. This previous work has considered, vague, conjugate and generalized conjugate distributions for the population mean when quantifying prior knowledge. However, in the aforementioned previous work, the prior distributions for the population mean vector and the factor loading matrix were taken to be independent. In this paper, the population mean vector and the factor loading matrix are taken to be jointly distributed which allows available interrelated prior information to be quantified and incorporated with the data.

The model parameters are estimated by both Gibbs sampling (Geman & Geman, 1984 and Gelfand & Smith, 1990) and iterated conditional modes (Lindley & Smith, 1972 and O'Hagen, 1994) algorithms which find posterior marginal mean and posterior joint modal (maximum a posteriori) estimates respectively.

The plan of the paper is to review the model and to adopt prior distributions in Section 2. Present the conditional posterior distributions along with the Gibbs sampling and ICM algorithms in Section 3. In Section 4 an example is detailed, and estimates from both the Gibbs sampling and the ICM estimation methods are presented.

2 Model

2.1 Likelihood Function

The Bayesian factor analysis model is:

$$\underset{(p\times 1)}{(x_j|\mu,\Lambda,f_j)} = \underset{(p\times 1)}{\mu} + \underset{(p\times m)}{\Lambda} \; \underset{(m\times 1)}{f_j} + \underset{(p\times 1)}{\epsilon_j} \quad , \quad m < p, \tag{2.1}$$

for $j = 1, \ldots, n$, where x_j is the j^{th} observation, μ is the overall population mean, Λ is a matrix of constants called the factor loading matrix; f_j is the factor score vector for subject j; and the ϵ_j's are observation errors.

In order to incorporate jointly distributed prior knowledge regarding the mean vector and factor loading matrix, the model is rewritten as

$$\underset{(p \times 1)}{(x_j|C, f_j)} = \underset{p \times (m+1)}{C} \quad \underset{(m+1) \times 1}{g_j} + \underset{p \times 1}{\epsilon_j}, \quad m < p, \tag{2.2}$$

where $C = (\mu, \Lambda)$ and $g_j' = (1, f_j')$.

As in the traditional factor analysis model, the errors are specified to be mutually uncorrelated and normally distributed $N(0, \Psi)$ variables. With the error specification, the distribution of each x_j can be written as

$$p(x_j|C, f_j, \Psi) = (2\pi)^{-\frac{p}{2}} |\Psi|^{-\frac{1}{2}} e^{-\frac{1}{2}(x_j - Cg_j)'\Psi^{-1}(x_j - Cg_j)}. \tag{2.3}$$

If proportionality is denoted by "$\propto$" then the likelihood for (C, F, Ψ) is

$$p(X|C, F, \Psi) \propto |\Psi|^{-\frac{n}{2}} e^{-\frac{1}{2} tr \Psi^{-1}(X - GC')'(X - GC')} \tag{2.4}$$

where the p-variate observation vectors on n subjects are $X' = (x_1, \ldots, x_n)$, the factor scores are contained in $G' = (g_1, \ldots, g_n)$, and the errors of observation are $E' = (\epsilon_1, \ldots, \epsilon_n)$. The notation $p(\cdot)$ will generically denote a probability distribution which is distinguished by its argument whose proportionality constant does not depend on its argument.

2.2 Priors

Available prior knowledge is incorporated into the inferences in terms of prior distributions for the model parameters. The joint prior distribution for the parameters is:

$$p(C, F, \Psi) \propto p(C|\Psi)p(\Psi)p(F), \tag{2.5}$$

where

$$
\begin{aligned}
p(C|\Psi) &\propto |\Psi|^{-\frac{m+1}{2}} e^{-\frac{1}{2} tr \Psi^{-1}(C-C_0)M^{-1}(C-C_0)'}, &(2.6)\\
p(\Psi) &\propto |\Psi|^{-\frac{\nu}{2}} e^{-\frac{1}{2} tr \Psi^{-1} Q}, &(2.7)\\
p(F) &\propto e^{-\frac{1}{2} tr F'F} &(2.8)
\end{aligned}
$$

with M, Q, and Ψ positive definite matrices. The matrix C conditional on Ψ has elements which are jointly normally distributed, and hyperparameters (C_0, M) are to be assesssed. The hyperparameter C_0 contains both the prior means for the population mean vector and the factor loading matrix as $C_0 = (\mu_0, \Lambda_0)$. The population mean vector and the loading matrix are jointly distributed just as the intercepts and variable coefficients are in a multivariate regression model. The matrix Ψ follows an inverted Wishart distribution, with hyperparameters (ν, Q) which are to be assessed. It is specified that $E(\Psi)$ is diagonal, in order to represent traditional views of the factor model containing "common" and "specific" factors (Press & Shigemasu, 1989). The factor scores, the rows of F are taken to be normally distributed with mean zero and identity covariande which is the model part of the traditional orthogonal factor analysis model and thus Q is taken to be a diagonal.

2.3 Joint Posterior

Using Bayes rule, combine Equations (2.4)–(2.8), to get the joint posterior distribution of the parameters

$$
p(C, F, \Psi | X) \propto e^{-\frac{1}{2} tr F'F} |\Psi|^{-\frac{(n+\nu+m+1)}{2}} e^{-\frac{1}{2} tr \Psi^{-1} U} \qquad (2.9)
$$

where

$$U = (X - GC')'(X - GC') + (C - C_0)M^{-1}(C - C_0)' + Q.$$

3 Estimation

As stated earlier, marginal mean and joint modal posterior estimates are found by the Gibbs sampling (Geman & Geman, 1984 and Gelfand & Smith, 1990) and iterated conditional modes (Lindley & Smith, 1972 and O'Hagen, 1994) algorithms. For both methods of parameter estimation, the conditional posterior distributions are needed.

3.1 Conditional Posterior Distributions

The posterior conditional distributions are as follows.

$$\begin{aligned} p(C|F,\Psi,X) &\propto p(C|\Psi)p(X|C,F,\Psi) \\ &\propto |\Psi|^{-\frac{m+1}{2}} e^{-\frac{1}{2}tr\Psi^{-1}(C-C_0)M^{-1}(C-C_0)'} \\ &\cdot\ |\Psi|^{-\frac{n}{2}} e^{-\frac{1}{2}tr\Psi^{-1}(X-GC')'(X-GC')} \\ &\propto e^{-\frac{1}{2}tr\Psi^{-1}(C-\tilde{C})(M^{-1}+G'G)(C-\tilde{C})'} \end{aligned} \tag{3.1}$$

where $\tilde{C} = (X'G + C_0M^{-1})(M^{-1} + G'G)^{-1}$.

$$\begin{aligned} p(\Psi|C,F,X) &\propto p(\Psi)p(C|\Psi)p(X|C,F,\Psi) \\ &\propto |\Psi|^{-\frac{\nu}{2}} e^{-\frac{1}{2}tr\Psi^{-1}Q} |\Psi|^{-\frac{m+1}{2}} e^{-\frac{1}{2}tr\Psi^{-1}(C-C_0)M^{-1}(C-C_0)'} \\ &\cdot\ |\Psi|^{-\frac{n}{2}} e^{-\frac{1}{2}tr\Psi^{-1}(X-GC')'(X-GC')} \\ &\propto |\Psi|^{-\frac{(n+\nu+m+1)}{2}} e^{-\frac{1}{2}tr\Psi^{-1}U} \end{aligned} \tag{3.2}$$

where $U = (X - GC')'(X - GC') + (C - C_0)M^{-1}(C - C_0)' + Q$.

$$\begin{aligned} p(F|\mu, \Lambda, \Psi, X) &\propto p(F)p(X|\mu, \Lambda, F, \Psi) \\ &\propto e^{-\frac{1}{2}trF'F}|\Psi|^{-\frac{n}{2}}e^{-\frac{1}{2}tr\Psi^{-1}(X-e_n\mu'-F\Lambda')'(X-e_n\mu'-F\Lambda')} \\ &\propto e^{-\frac{1}{2}tr(F-\tilde{F})(I_m+\Lambda'\Psi^{-1}\Lambda)(F-\tilde{F})'} \end{aligned} \tag{3.3}$$

where $\tilde{F} = (X - e_n\mu')\Psi^{-1}\Lambda(I_m + \Lambda'\Psi^{-1}\Lambda)^{-1}$.

The modes of these conditional distributions are $\tilde{C}$, $\tilde{F}$, (as defined above), and

$$\tilde{\Psi} = \frac{U}{n + \nu + m + 1}, \tag{3.4}$$

respectively.

3.2 The Gibbs Sampling Algorithm

For Gibbs estimation of the posterior, start with initial values for F and Ψ say $\bar{F}_{(0)}$ and $\bar{\Psi}_{(0)}$. Then cycle through

$$\begin{aligned} \bar{C}_{(l+1)} &= \text{a random variate from } p(C|\bar{F}_{(l)}, \bar{\Psi}_{(l)}, X) \\ \bar{\Psi}_{(i+1)} &= \text{a random variate from } p(\Psi|\bar{F}_{(l)}, \bar{C}_{(l+1)}, X) \\ \bar{F}_{(l+1)} &= \text{a random variate from } p(F|\bar{C}_{(l+1)}, \bar{\Psi}_{(l+1)}, X). \end{aligned}$$

An initial number of random variates called the "burn in" are discarded and the remaining L variates are kept. The means of the remaining random variates

$$\bar{F} = \frac{1}{L}\sum_{l=1}^{L}\bar{F}_{(l)} \quad \bar{C} = \frac{1}{L}\sum_{l=1}^{L}\bar{C}_{(l)} \quad \bar{\Psi} = \frac{1}{L}\sum_{l=1}^{L}\bar{\Psi}_{(l)}.$$

are the marginal posterior mean estimates of the parameters.

3.3 The ICM Algorithm

For iterated conditional modes estimatiom of the posterior, start with an initial value for $\tilde{F}$, say $\tilde{F}_{(0)}$, form $\tilde{G}_{(0)} = (e_n, \tilde{F}_{(0)})$, and cycle through

$$\begin{aligned}
\tilde{C}_{(l+1)} &= (X'\tilde{G}_{(l)} + C_0 M^{-1})(M^{-1} + \tilde{G}'_{(l)}\tilde{G}_{(l)})^{-1} \\
\tilde{\Psi}_{(l+1)} &= [(X - \tilde{G}_{(l)}\tilde{C}'_{(l+1)})'(X - \tilde{G}_{(l)}\tilde{C}'_{(l+1)}) + \\
&\quad (\tilde{C}_{(l+1)} - C_0)M^{-1}(\tilde{C}_{(l+1)} - C_0)' + Q]/(n + \nu + m + 1) \\
\tilde{F}_{(l+1)} &= (X - e_n\tilde{\mu}'_{(l)})\tilde{\Psi}^{-1}_{(l+1)}\tilde{\Lambda}_{(i+1)}(I_m + \tilde{\Lambda}'_{(l+1)}\tilde{\Psi}^{-1}_{(l+1)}\tilde{\Lambda}_{(l+1)})^{-1}
\end{aligned}$$

until convergence is reached with the joint posterior modal (maximum a posteriori) estimator $(\tilde{C}, \tilde{F}, \tilde{\Psi})$.

4 Example

In this section the Gibbs Sampling and ICM procedures for estimating the parameters of the aforementioned Bayesian factor analysis model are used and the resulting estimators are presented. The data is extracted from an example in Kendall 1980, p.53. The problem (Press & Shigemasu, 1989) is the following.

> There are 48 applicants for a certain job, and they have been scored on 15 variables regarding their acceptability. They are:

(1)	Form of letter application	(9)	Experience
(2)	Appearance	(10)	Drive
(3)	Academic ability	(11)	Ambition
(4)	Likeabiliy	(12)	Grasp
(5)	Self-confidence	(13)	Potential
(6)	Lucidity	(14)	Keenness to join
(7)	Honesty	(15)	Suitability
(8)	Salesmanship		

Table 1: Raw scores of 48 applicants scaled on 15 variables.

Person	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	6	7	2	5	8	7	8	8	3	8	9	7	5	7	10
2	9	10	5	8	10	9	9	10	5	9	9	8	8	8	10
3	7	8	3	6	9	8	9	7	4	9	9	8	6	8	10
4	5	6	8	5	6	5	9	2	8	4	5	8	7	6	5
5	6	8	8	8	4	5	9	2	8	5	5	8	8	7	7
6	7	7	7	6	8	7	10	5	9	6	5	8	6	6	6
7	9	9	8	8	8	8	8	8	10	8	10	8	9	8	10
8	9	9	9	8	9	9	8	8	10	9	10	9	9	9	10
9	9	9	7	8	8	8	8	5	9	8	9	8	8	8	10
10	4	7	10	2	10	10	7	10	3	10	10	10	9	3	10
11	4	7	10	0	10	8	3	9	5	9	10	8	10	2	5
12	4	7	10	4	10	10	7	8	2	8	8	10	10	3	7
13	6	9	8	10	5	4	9	4	4	4	5	4	7	6	8
14	8	9	8	9	6	3	8	2	5	2	6	6	7	5	6
15	4	8	8	7	5	4	10	2	7	5	3	6	6	4	6
16	6	9	6	7	8	9	8	9	8	8	7	6	8	6	10
17	8	7	7	7	9	5	8	6	6	7	8	6	6	7	8
18	6	8	8	4	8	8	6	4	3	3	6	7	2	6	4
19	6	7	8	4	7	8	5	4	4	2	6	8	3	5	4
20	4	8	7	8	8	9	10	5	2	6	7	9	8	8	9
21	3	8	6	8	8	8	10	5	3	6	7	8	8	5	8
22	9	8	7	8	9	10	10	10	3	10	8	10	8	10	8
23	7	10	7	9	9	9	10	10	3	9	9	10	9	10	8
24	9	8	7	10	8	10	10	10	2	9	7	9	9	10	8
25	6	9	7	7	4	5	9	3	2	4	4	4	4	5	4
26	7	8	7	8	5	4	8	2	3	4	5	6	5	5	6
27	2	10	7	9	8	9	10	5	3	5	6	7	6	4	5
28	6	3	5	3	5	3	5	0	0	3	3	0	0	5	0
29	4	3	4	3	3	0	0	0	0	4	4	0	0	5	0
30	4	6	5	6	9	4	10	3	1	3	3	2	2	7	3
31	5	5	4	7	8	4	10	3	2	5	5	3	4	8	3
32	3	3	5	7	7	9	10	3	2	5	3	7	5	5	2
33	2	3	5	7	7	9	10	3	2	2	3	6	4	5	2
34	3	4	6	4	3	3	8	1	1	3	3	3	2	5	2
35	6	7	4	3	3	0	9	0	1	0	2	3	1	5	3
36	9	8	5	5	6	6	8	2	2	2	4	5	6	6	3
37	4	9	6	4	10	8	8	9	1	3	9	7	5	3	2
38	4	9	6	6	9	9	7	9	1	2	10	8	5	5	2
39	10	6	9	10	9	10	10	10	10	10	8	10	10	10	10
40	10	6	9	10	9	10	10	10	10	10	10	10	10	10	10
41	10	7	8	0	2	1	2	0	10	2	0	3	0	0	10
42	10	3	8	0	1	1	0	0	10	0	0	0	0	0	10
43	3	4	9	8	2	4	5	3	6	2	1	3	3	3	8
44	7	7	7	6	9	8	8	6	8	8	10	8	8	6	5
45	9	6	10	9	7	7	10	2	1	5	5	7	8	4	5
46	9	8	10	10	7	9	10	3	1	5	7	9	9	4	4
47	0	7	10	3	5	0	10	0	0	2	2	0	0	0	0
48	0	6	10	1	5	0	10	0	0	2	2	0	0	0	0

The raw scores of the applicants on these 15 variables, measured on the same scale, are presented in Table 1. The question is, Is there an underlying subset of factors that explain the variation observed in the scores? If so, then the applicants could be compared more easily.

The underlying structure which is postulated (Press & Shigemasu, 1989) is a model with 4 factors. This choice is based upon a principal components analysis which found that 4 factors accounted for 81.5% of the variance. Based upon underlying theory (Press & Shigemasu, 1989) the prior factor loading matrix

$$\Lambda_0' = \begin{pmatrix} 0 & 0 & 0 & 0 & .7 & .7 & 0 & .7 & 0 & .7 & .7 & .7 & .7 & 0 & 0 \\ 0 & 0 & .7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .7 & 0 & 0 & 0 & 0 & 0 & .7 \\ 0 & 0 & 0 & .7 & 0 & 0 & .7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

is assessed.

The hyperparameter M is assessed as $M = \frac{1}{10} I_5$, Q is assessed as $Q = 0.2 I_{15}$, and ν is assessed as $\nu = 33$. The prior population mean is assessed as $\mu_0 = 7.5 e_{15}$ where the 15 dimensional unit vector has been denoted by e_{15}. The population mean, factor loadings, factor scores, and disturbance covariance matrix may now be estimated. The rows of X along with the columns of μ_0 were scaled by the variances of the columns X for estimation and the estimated means were rescaled. It was found that a burn in period of $5,000$ samples worked well, so then the next $25,000$ samples were taken for the Gibbs estimates.

Table 2 displays the Gibbs sampling and ICM estimates of the population mean along with the prior and sample means.

Table 3 displays the Gibbs sampling and ICM estimates of the factor loadings. For enhanced interpretability, the rows of the factor loading matrices have been rearranged. It is seen that factor 1 loads heavilly for

Table 2: Gibbs Sampling and ICM estimates of the mean.

p	Gibbs Mean	ICM Mean	Sample Mean	Prior Mean
1	7.0536	7.3959	6.0000	7.5000
2	7.4084	7.5305	7.0833	7.5000
3	7.2871	7.3538	7.0833	7.5000
4	6.6636	6.8584	6.1458	7.5000
5	7.6118	7.6941	6.9375	7.5000
6	7.4313	7.6225	6.3333	7.5000
7	7.9234	7.9147	8.0417	7.5000
8	6.3472	6.6418	4.7917	7.5000
9	5.8173	6.3038	4.2292	7.5000
10	6.6362	6.9201	5.3125	7.5000
11	7.0960	7.3092	5.9792	7.5000
12	7.4049	7.6515	6.2500	7.5000
13	6.9856	7.2853	5.6875	7.5000
14	6.3746	6.5979	5.5625	7.5000
15	7.3620	7.8412	5.9583	7.5000

Table 3: Gibbs (left) and ICM (right) Estimates of Factor Loadings.

p	1	2	3	4	1	2	3	4
5	.7916	-.0419	-.1290	.0077	.7828	-.0472	-.1676	-.0111
6	.7563	-.0342	.0280	.0981	.7315	-.0212	-.0087	.0742
8	.7982	-.0722	.0624	-.0623	.7869	-.0664	.0645	-.0738
10	.7078	-.0507	.1737	.0320	.6839	-.0439	.1716	.0120
11	.7946	-.0711	.0053	-.0756	.7885	-.0630	.0004	-.0844
12	.7114	.0090	.1404	.1245	.6792	.0356	.1070	.1103
13	.6612	.0561	.1824	.2038	.6247	.0887	.1609	.1898
3	.0224	.7251	.0891	.0249	.0277	.7191	.0638	.0270
1	.0666	-.1008	.7739	.0622	.0049	-.0942	.7606	.0622
9	.0430	.0094	.8079	.0126	-.0090	.0423	.8184	-.0002
15	.2411	-.0863	.7242	.0463	.1817	-.0526	.7233	.0453
4	.1198	-.0815	.1523	.7177	.0557	-.0544	.1834	.7227
7	.1043	-.0169	-.1260	.7332	.0499	-.0016	-.1483	.7339
2	.2951	-.0292	.0557	.1054	.2822	.0111	.0836	.1548
14	.3279	-.2928	.1739	.2971	.2903	-.2993	.2279	.3115

variables 5, 6, 8, 10, 11, 12, and 13; factor 2 on variable 3; factor 3 heavily on variables 1, 9, and 15; while factor 4 loads heavily on variables 4 and 7. These factors in terms of the original

Table 4: Gibbs (left) and ICM (right) Estimates of the Factor Scores.

Person	1	2	3	4	1	2	3	4
1	0.2960	-3.2654	-0.1959	-0.5257	0.1606	-3.5584	-0.3575	-0.5349
2	0.7632	-1.5504	0.2340	0.1894	0.7575	-1.6808	0.2195	0.3713
3	0.4924	-2.6609	0.0333	-0.1107	0.3921	-2.9206	-0.1091	-0.0416
4	-0.7529	0.4371	-0.1179	0.1661	-0.8869	0.5279	-0.4046	0.0312
5	-1.1055	0.0273	-0.1270	0.6511	-1.0895	0.2790	-0.1725	0.6625
6	-0.2011	-0.1394	0.4296	0.5563	-0.3319	-0.1462	0.1647	0.4682
7	0.3281	0.1819	0.7599	0.0752	0.3905	0.2531	0.8603	0.1704
8	0.5781	0.9024	0.7009	0.0136	0.6483	0.9351	0.8176	0.1400
9	0.0926	-0.4091	0.7961	0.1165	0.1116	-0.3768	0.8034	0.2049
10	1.6741	2.0134	-0.0900	-1.2662	1.4590	2.0559	-0.4261	-1.3752
11	1.2500	1.8020	-0.6335	-2.7076	1.0815	1.9658	-0.9212	-2.9114
12	1.3113	1.9370	-0.5688	-0.6862	1.0971	2.0434	-0.9648	-0.8000
13	-1.2942	0.1815	-0.5897	0.9755	-1.2880	0.3323	-0.6155	1.0498
14	-1.1869	0.2541	-0.0622	0.6158	-1.2535	0.4241	-0.2638	0.6192
15	-1.2105	0.1900	-0.3958	0.8499	-1.2975	0.4498	-0.6034	0.7697
16	0.2014	-1.1197	0.1396	-0.1016	0.2113	-1.0312	0.1744	-0.0367
17	-0.1031	-0.0094	0.2347	0.0640	-0.2017	-0.1779	0.0809	0.0221
18	-0.5328	0.7953	-0.8235	-1.2201	-0.7106	0.6368	-1.1290	-1.2534
19	-0.5331	0.6843	-0.5756	-1.3078	-0.7226	0.6290	-0.9075	-1.4296
20	0.1655	-0.1471	-0.8550	0.6893	0.0829	-0.1970	-1.0102	0.8057
21	0.1664	-0.9140	-0.8296	0.9003	0.0443	-0.8392	-1.0544	0.9045
22	0.8550	-0.0303	-0.2515	0.5729	0.8245	-0.2741	-0.3225	0.7138
23	0.7007	-0.3255	-0.9118	0.6038	0.7634	-0.4100	-0.8217	0.8613
24	0.5499	-0.1482	-0.4692	1.0635	0.5545	-0.3338	-0.4949	1.2228
25	-1.4395	-0.2626	-1.0962	0.3713	-1.5303	-0.2132	-1.3128	0.3743
26	-1.1870	-0.2669	-0.4081	0.4448	-1.2976	-0.2034	-0.6532	0.3946
27	-0.1562	-0.3893	-1.5819	1.0424	-0.2345	-0.2079	-1.7420	1.0966
28	-1.9049	-0.7866	-1.3147	-1.1532	-2.2280	-1.2265	-1.8600	-1.5166
29	-2.4161	-1.9105	-2.1853	-2.5681	-2.5958	-2.1770	-2.4250	-2.9581
30	-1.2047	-0.8843	-1.4650	0.4613	-1.4420	-1.2948	-1.8597	0.3869
31	-0.9543	-1.6733	-1.2850	0.7778	-1.1459	-2.0788	-1.6079	0.6640
32	-0.2949	-1.0632	-1.2234	1.2409	-0.5964	-1.2671	-1.7646	0.9369
33	-0.5818	-0.9921	-1.4000	1.1941	-0.8973	-1.2243	-1.9410	0.9053
34	-1.8264	-0.4977	-1.6063	-0.1801	-2.0689	-0.6754	-2.0180	-0.4497
35	-2.3990	-1.8693	-0.8847	-0.4015	-2.6429	-2.0414	-1.3609	-0.5105
36	-1.0635	-1.3135	-0.3380	-0.2968	-1.2520	-1.4339	-0.7874	-0.3159
37	0.5058	-0.5549	-1.5516	-0.6353	0.2833	-0.6483	-1.9382	-0.6854
38	0.3374	-0.6981	-1.9067	-0.5610	0.2040	-0.7895	-2.1215	-0.5485
39	0.8581	1.1046	1.0410	1.2921	0.8801	1.0039	1.0739	1.3098
40	1.0071	1.1324	1.0180	1.2752	1.0284	1.0084	1.0633	1.2868
41	-2.4526	0.3591	2.1495	-2.7291	-2.6775	0.6079	1.7383	-3.0415
42	-2.7562	0.6324	2.3316	-2.9559	-3.0521	0.7185	1.8513	-3.4691
43	-2.0919	0.9520	-0.4439	0.0251	-2.2026	1.1547	-0.5726	-0.2761
44	0.5611	-0.1663	-0.0036	-0.0265	0.4625	-0.1958	-0.1921	-0.1354
45	-0.2254	2.1088	0.1468	1.5599	-0.4851	2.0908	-0.4240	1.3945
46	0.1684	1.8709	-0.2032	1.6362	-0.0210	1.9787	-0.6763	1.5681
47	-2.0743	2.1837	-2.2167	0.1476	-2.3946	2.2267	-2.7794	-0.1245
48	-1.9674	2.3437	-2.0335	-0.2506	-2.3430	2.3154	-2.6884	-0.5719

variables are factor 1: Self-confidence, Lucidity, Salesmanship, Drive, Ambition, Grasp, Potential; factor 2: Academic ability; factor 3: Form of letter application, Experience, Suitability; and factor 4: Likeabiliy, Honesty. These factors may be loosely interpreted as factor 1 being personality, factor 2 being academic ability, factor 3 being position match, and factor 4 being charisma.

In Table 4, the Gibbs sampling and ICM estimates of the factor scores are presented. An employer may now decide on a criteria to select a person. For example, if the employer wished to hire a person that is "very" hard working with a "very good" academic record and a "fair" match for the position, person 10 might be selected.

Table 5: Gibbs (top) and ICM (bottom) Estimates of the Disturbance Covariance Matrix.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.2243	0.0681	-0.0045	0.0887	0.0087	0.0100	-0.0283	0.0312	-0.0363	0.0238	0.0596	0.0189	0.0291	0.1516	-0.0321
2		0.4832	0.0590	0.0779	0.0234	-0.0198	0.0698	0.0571	0.0317	-0.0182	0.1010	0.0591	0.0635	0.0251	0.1021
3			0.0515	0.0317	-0.0196	0.0015	-0.0126	0.0149	0.0425	0.0153	0.0160	0.0226	0.0393	0.0022	0.0266
4				0.1662	-0.0217	0.0361	-0.0654	0.0606	0.0572	0.0425	0.0607	0.0350	0.0665	0.1266	0.0640
5					0.0937	-0.0144	0.0345	0.0001	-0.0014	-0.0177	0.0144	-0.0353	-0.0383	0.0145	-0.0249
6						0.0991	-0.0305	0.0093	0.0047	-0.0349	-0.0278	0.0436	-0.0013	0.0214	-0.0001
7							0.1215	-0.0122	-0.0189	-0.0094	-0.0146	-0.0073	-0.0287	-0.0133	0.0036
8								0.1297	0.0622	0.0520	0.0336	-0.0089	0.0088	0.0836	0.0678
9									0.2493	0.0660	0.0491	0.0270	0.0516	0.0869	0.0411
10										0.1890	0.0308	-0.0226	0.0463	0.1142	0.0726
11											0.1170	0.0076	0.0326	0.0754	0.0219
12												0.1131	0.0464	0.0413	0.0139
13													0.1283	0.0487	0.0375
14														0.3312	0.0527
15															0.1848

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	.1731	.0287	.0009	.0359	.0097	.0028	-.0265	-.0017	-.0799	-.0100	.0315	.0043	.0040	.0808	-.0715
2		.4147	.0209	.0184	.0087	-.0425	.0318	.0275	-.0107	-.0481	.0718	.0214	.0166	-.0128	.0551
3			.0078	.0159	-.0093	-.0035	-.0101	.0085	.0123	.0091	.0086	.0029	.0098	.0185	.0054
4				.1006	-.0288	.0166	-.0830	.0271	.0104	.0100	.0306	.0042	.0263	.0643	.0090
5					.0817	-.0155	.0320	-.0088	.0002	-.0231	.0036	-.0352	-.0398	-.0033	-.0214
6						.0855	-.0250	-.0075	-.0081	-.0459	-.0404	.0312	-.0142	-.0035	-.0110
7							.0838	-.0087	-.0047	-.0029	-.0130	-.0086	-.0289	-.0300	.0048
8								.0882	.0186	.0162	.0047	-.0297	-.0198	.0384	.0271
9									.1531	.0186	.0147	.0000	.0099	.0300	-.0251
10										.1397	.0006	-.0419	.0140	.0634	.0277
11											.0827	-.0127	.0046	.0366	-.0070
12												.0867	.0189	.0085	-.0087
13													.0839	.0107	.0005
14														.2327	-.0080
15															.1136

Table 5 displays the Gibbs sampling and ICM estimates of the disturbance covariance matrix.

5 Conclusion

A Bayesian statistical model which allowed the poplation mean vector and the factor loading matrix to be jointly distributed was detailed. Available prior information either from substantive experts or previous experiments was incorporated. An added feature of the Bayesian factor analysis model is that the there is no need to rotate the factor loading matrix. The rotation is automatically found. Available prior information regarding the population mean was incorporated along with the other parameters through a prior distribution. By incorporating prior knowlege regarding the mean, estimation of it and the other parameters may be improved.

References

[1] A. E. Gelfand and A. F. M Smith. Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85:398–409, 1990.

[2] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, 6:721–741, 1984.

[3] K. Hayashi. *The Press-Shigemasu Bayesian Factor Analysis Model*

With Estimated Hyperparameters. PhD thesis, Department of Psychology, University of North Carolina, Chapel Hill, NC 27599, 1997.

[4] M. Kendall. *Multivariate Analysis.* Charles Griffin & Company LTD, London, second edition, 1980.

[5] D. V. Lindley and A. F. M. Smith. Bayes estimates for the linear model. *Journal of the Royal Statistical Society B*, 34(1), 1972.

[6] A. O'Hagen. *Kendalls' Advanced Theory of Statistics, Volume 2B Bayesian Inference.* John Wiley and Sons Inc., New York, 1994.

[7] S. J. Press. *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference.* Robert E. Krieger Publishing Company, Malabar, Florida, 1982.

[8] S. J. Press. *Bayesian Statistics: Principles, Models, and Applications.* John Wiley and Sons, New York, 1989.

[9] S. J. Press and K. Shigemasu. Bayesian inference in factor analysis. In *Contributions to Probability and Statistics*, chapter 15. Springer-Verlag, 1989.

[10] S. J. Press and K. Shigemasu. Bayesian inference in factor analysis-Revised. Technical Report No. 243, Department of Statistics, University of California, Riverside, CA 92521, May 1997.

[11] D. B. Rowe. On estimating the mean in Bayesian factor analysis. Social Science Working Paper 1096, Division of Humanities and Social Sciences, Caltech, Pasadena, CA 91125, July 2000a.

[12] D. B. Rowe. Incorporating prior knowledge regarding the mean in Bayesian factor analysis. Social Science Working Paper 1097, Division of Humanities and Social Sciences, Caltech, Pasadena, CA 91125, July 2000b.

[13] D. B. Rowe. A Bayesian factor analysis model with generalized prior information. Social Science Working Paper 1099, Division of Humanities and Social Sciences, Caltech, Pasadena, CA 91125, August 2000c.

[14] D. B. Rowe and S. J. Press. Gibbs sampling and hill climbing in Bayesian factor analysis. Technical Report No. 255, Department of Statistics, University of California, Riverside, CA 92521, May 1998.

[15] D. M. Shera. *Bayesian Factor Analysis.* PhD thesis, Harvard University, School of Public Health, Boston, MA 02115, March 2000.