Factor Analysis

Section 5.2 of Drton and Plummer (2017) considers a factor analysis model applied to the data analysed by Lopes and West (2004) originally published in West and Harrison (1997, pp.610-618). The data concern average monthly exchange rates of six currencies relative to the pound sterling (GBP) from January 1975 to December 1986.

The data are read in as a data frame then converted to a matrix X containing only the currency values. The data to be passed to the factor analysis model take the form of a matrix Y of lag-1 differences in which each column is standardized to have mean 0 and standard deviation 1.

```
X = read.table("lopeswest.dat", header = TRUE)
X = as.matrix(X[, c("U.S.A.", "Canada", "Yen", "Franc", "Lira", "Mark")])
Y = scale(diff(X))
```

Under a factor analysis model $\mathrm{Var}(Y) = \Sigma + \beta\beta^T$ where Σ is a diagonal matrix (*uniquenesses*) and β is a $6 \times i$ matrix (*factor loadings*). Model singularities occur when the rank of β is assumed to be i, but the true data-generating model has rank $j < i$.

Factor analysis models are fitted with the `stats::factanal()` function. For example, the two-factor model is given by

```
factanal(Y, factors=2)
```

```
##
## Call:
## factanal(x = Y, factors = 2)
##
## Uniquenesses:
## U.S.A. Canada    Yen  Franc   Lira   Mark
##  0.005  0.162  0.621  0.016  0.260  0.266
##
## Loadings:
##        Factor1 Factor2
## U.S.A. 0.216   0.974
## Canada 0.250   0.881
## Yen    0.503   0.354
## Franc  0.976   0.177
## Lira   0.829   0.229
## Mark   0.826   0.229
##
##               Factor1 Factor2
## SS loadings     2.685   1.985
## Proportion Var  0.448   0.331
## Cumulative Var  0.448   0.778
##
## Test of the hypothesis that 2 factors are sufficient.
## The chi square statistic is 2.05 on 4 degrees of freedom.
## The p-value is 0.727
```

The frequentist analysis suggests that 2 factors are sufficient for these data. For Bayesian model choice using sBIC the `FactorAnalyses()` function creates an object representing a collection of factor analysis models

with a given number of covariates, up to a maximum number of factors. For the Lopes and West (2004) data, the number of covariates is 6, i.e. the number of columns of `Y`. The maximum number of factors that can be fitted to 6 variables is 3.

```
library(sBIC)
factorAnalyses = FactorAnalyses(numCovariates=ncol(Y), maxNumFactors=3)
```

The learning coefficients λ_{ij} can be extracted from the resulting object. Note that the minimal model is $i = 0$ corresponding to independence of the columns of Y.

```
lambda <- matrix("", 4, 4, dimnames=list(paste0("i=",0:3), paste0("j=",0:3)))
for (i in 1:4) {
  for (j in 1:i) {
    lambda[i,j] <- factorAnalyses$learnCoef(i, j)$lambda
  }
}
knitr::kable(lambda, caption="Learning coefficients $\\lambda_{ij}$ for sBIC")
```

Table 1: Learning coefficients λ_{ij} for sBIC

	j=0	j=1	j=2	j=3
i=0	3			
i=1	4.5	6		
i=2	6	7.25	8.5	
i=3	7.5	8.5	9.5	10.5

Table 1 reproduces Table 2 of Drton and Plummer (2017) showing learning coefficients for model with j factors embedded in a model with $i \geq j$ factors. The multiplicity is equal to 1 for all models (not shown).

The singular BIC is calculated by passing to the `sBIC()` function the data `Y` and the object `factorAnalysis` representing the model collection.

```
results = sBIC(Y, factorAnalyses)
print(results)
```

```
## $logLike
## [1] -429.00000 -191.61082  -70.04375  -68.93844
##
## $sBIC
## [1] -443.8885 -221.3879 -112.2279 -116.1058
##
## $BIC
## [1] -443.8885 -221.3879 -112.2279 -121.0483
##
## $modelPoset
## [1] "FactorAnalyses: 0x7fde170241c8"
```

If the singular BIC is considered to be an approximation to the log of the marginal likelihood, then it can be converted to an approximate posterior distribution by exponentiating and then normalizing the sBIC values. The convenience function `BICpost` carries out this transformation.

```
BICpost <- function(BIC) {
  w = exp((BIC - max(BIC)))
  zapsmall( w/ sum(w) )
}
```

The approximate posterior probabilities according to BIC and sBIC for models $i = 0$ up to $i = 3$ are as follows.

BICpost(results$BIC)

[1] 0.0000000 0.0000000 0.9998523 0.0001477

BICpost(results$sBIC)

[1] 0.0000000 0.0000000 0.9797249 0.0202751

These values can be compared with the posterior distribution of the number of components derived by Lopes and West (2004) using reversible jump MCMC. They give two posterior distributions based on two different models.

Table 2: Posterior distributions for the number of classes (i) from Lopes & West (2004)

	i = 1	i = 2	i = 3
Table 3	0.00	0.88	0.12
Table 5	0.00	0.98	0.02

The approximate posterior derived from sBIC resembles the posterior from Table 5 of Lopes and West (2004). In contrast, the approximate posterior derived from BIC puts extremely low weight on the model i = 3.

Bibliography

- Drton M. and Plummer M. (2017), A Bayesian information criterion for singular models. J. R. Statist. Soc. B; 79: 1-38.
- Lopes, HF and West M (2004). Bayesian model assessment in factor analysis, Statistica Sinica, 14, 41-67.
- West, M. and Harrison, P.J. (1997). Bayesian Forecasting and Dynamic Model. Springer-Verlag, New York.