202B hw4

Kun Zhou

2016-2-18

1)

a. Based on the factor analysis assumption, we have:

$$X = \Lambda F + U$$

where $Var(F) = I, Var(U) = diag(\psi_1, \ldots, \psi_q) = \Psi, Cov(F, U) = 0, Var(X) = \Sigma$ and Λ is the loading matrix. Then,

$$\begin{aligned}
&\Sigma = Var(X) = Var(\Lambda F + U) \\
&= Var(\Lambda F) + Var(U) \qquad \text{since independence of common factors and specific variances} \\
&= \Lambda Var(F)\Lambda^t + Var(U) \\
&= \Lambda\Lambda^t + \Psi
\end{aligned}$$

If factors are correlated, we assume $Var(F) = \Phi$ instead.

$$\Sigma = Var(X) = \Lambda Var(F)\Lambda^t + Var(U) = \Lambda\Phi\Lambda^t + \Psi$$

b) If $\Lambda^* = \Lambda M$, then $X = \Lambda^*(M^tF) + U$, where $Var(M^tF) = M^tIM = I$ and $Cov(M^tF, U) = M^tCov(F, U) = 0$
According to (a),

$$\begin{aligned}
&\Sigma = (\Lambda^*)(\Lambda^*)^t + \Psi \\
&= (\Lambda M)(\Lambda M)^t + \Psi = \Lambda\Lambda^t + \Psi
\end{aligned}$$

2)

$$tr\left[D_r^{-\frac{1}{2}}(P - \widehat{P})D_c^{-1}(P - \widehat{P})^tD_r^{-\frac{1}{2}}\right] = tr\left[\left(D_r^{-\frac{1}{2}}(P - \widehat{P})D_c^{-\frac{1}{2}}\right)\left(D_r^{-\frac{1}{2}}(P - \widehat{P})D_c^{-\frac{1}{2}}\right)^t\right]$$

By SVD, $D_r^{-\frac{1}{2}}PD_c^{-\frac{1}{2}} = U\Lambda V^t$.

By Eckart-Young Theorem, the best t-rank reduced matrix

$D_r^{-\frac{1}{2}}\widehat{P}D_c^{-\frac{1}{2}} = U\Lambda_tV^t$, where the first t diagonal elements of Λ_t are same with those of Λ and other elements are zeros. Thus, $\widehat{P} = D_r^{\frac{1}{2}}U\Lambda_tV^tD_c^{\frac{1}{2}}$.

$D_r^{-\frac{1}{2}}PD_c^{-\frac{1}{2}} = \left(\frac{n_{ij}}{\sqrt{n_i}\sqrt{n_j}}\right)_{ij}$. Now prove that $D_r^{-\frac{1}{2}}r = (\sqrt{\frac{n_{i+}}{n}})_i$, $\lambda = 1$ and $D_c^{-\frac{1}{2}}c = (\sqrt{\frac{n_{+j}}{n}})_j$ are corresponding singular vecotrs and singular values for $D_r^{-\frac{1}{2}}PD_c^{-\frac{1}{2}}$ through the following equation.

$$\left(D_r^{-\frac{1}{2}}PD_c^{-\frac{1}{2}}\right)\left(D_r^{-\frac{1}{2}}PD_c^{-\frac{1}{2}}\right)^t\left(D_r^{-\frac{1}{2}}r\right)=\left(\sum_{k,j}\frac{n_{ik}}{\sqrt{n_{i+}n_{+k}}}\frac{n_{jk}}{\sqrt{n_{j+}n_{+k}}}\sqrt{\frac{n_{j+}}{n}}\right)_i=\left(\sqrt{\frac{n_{i+}}{n}}\right)_i=D_r^{-\frac{1}{2}}r$$

$$\left(D_r^{-\frac{1}{2}}PD_c^{-\frac{1}{2}}\right)^t\left(D_r^{-\frac{1}{2}}PD_c^{-\frac{1}{2}}\right)\left(D_c^{-\frac{1}{2}}c\right)=\left(\sum_{k,i}\frac{n_{kj}}{\sqrt{n_{k+}n_{+j}}}\frac{n_{ki}}{\sqrt{n_{k+}n_{+i}}}\sqrt{\frac{n_{+i}}{n}}\right)_j=\left(\sqrt{\frac{n_{+j}}{n}}\right)_j=D_c^{-\frac{1}{2}}c$$

Now prove that $\lambda = 1$ is the largest eigenvalue for $D_r^{-\frac{1}{2}}PD_c^{-\frac{1}{2}}$.

Let A $= D_r^{-1}PD_c^{-1}P^t = (a_{ij})$. Each element of a_{ij} is nonnegative. And

$$A1_r = D_r^{-1}PD_c^{-1}P^t1_r = D_r^{-1}PD_c^{-1}c = D_r^{-1}P1_c = D_r^{-1}r = 1_r$$

So each row sum of A is zero, namely $\sum_j a_ij = 1$ Assume $\exists\lambda,\quad x = (x_1, \ldots, x_r)^t$ s.t. $Ax = \lambda x$. Let $x_k = \max|x_i|$.

$$|\lambda x_k| = |\sum_i a_{ki}x_i| \iff |\lambda||x_k| = |\sum_i a_{ki}x_i| \iff |\lambda| = |\sum_i a_{ki}\frac{x_i}{|x_k|}| \Rightarrow$$

$$|\lambda| \leq \sum_i a_{ki}|\frac{x_i}{|x_k|}| \leq \sum_i a_{ki} = 1$$

Thus eigenvalues of A are no greater than 1. So eigenvalues of $(D_r^{-\frac{1}{2}}PD_c^{-\frac{1}{2}})(D_r^{-\frac{1}{2}}PD_c^{-\frac{1}{2}})^t$ are no greater than 1. Finally, eigenvalues of $D_r^{-\frac{1}{2}}PD_c^{-\frac{1}{2}}$ are no greater than 1. $\lambda = 1$ is the largest.

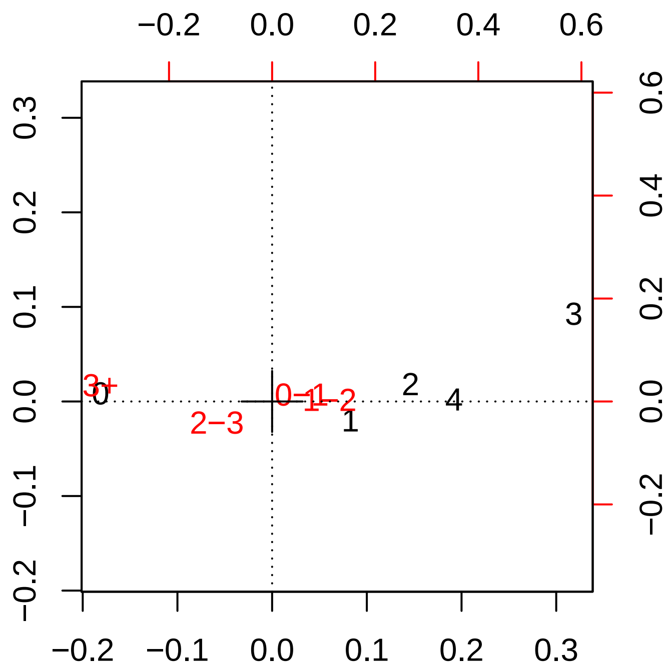
so when rank $t = 1$, $D_r^{-\frac{1}{2}}\widehat{P}D_c^{-\frac{1}{2}} = D_r^{-\frac{1}{2}}rc^tD_c^{-\frac{1}{2}} \Rightarrow \widehat{P} = rc^t$.

3)

```
data = dget("http://www.stat.ucla.edu/~handcock/202B/datasets/Table17.11.dput")
library(MASS)
chisq.test(data)
```

```
##
##  Pearson's Chi-squared test
##
## data:  data
## X-squared = 568.57, df = 12, p-value < 2.2e-16
```

```
data.cor = corresp(data, nf = 2)
biplot(data.cor)
abline(v = 0, h = 0, lty = 3)
```



From χ^2 test, the p-value is 2.2e-16, which means the number of children and income are dependent.

For those families whose yearly income is 0-1k or 1-2k, their children distribution is similar. When yearly income is 0-1k, the ratio of number of children is 1: 1.27: 0.43: 0.10: 0.02. When yearly income is 1-2k, the ratio is 1: 1.42: 0.49: 0.12: 0.27. And the number of those families having 1 children is largest.

For families whose yearly income is 3+k, the number of those having no children is largest.

For those families which have 2 or 4 children, their yearly income distribution is similar. When they have 2 children, the ratio of amount of groups for different income is 1: 1.87: 0.68: 0.33. When they have 4 children, the ratio is 1 : 2.51: 0.79: 0.36.