HUDM6122 Homework_05

Chenguang Pan

2023-03-20

0.1 Github Address

All my latest homework can be found on Github: https://github.com/cgpan/hudm6122_homeworks . Thanks for checking if interested.

0.2 Ex 5.1

Show how the result rises from the assumptions of uncorrelated factors, independence of the specific variates, and independence of common factors and specific variances. What form does take if the factors are allowed to be correlated?

MY SOLUTION:
Based on the assumption of Exploratory Factor Analysis(EFA), a set of observed variables **x** assumed to be linked to a set of latent variables **f**. Therefore, we can have a regression model in matrix form

$$\boldsymbol{x} = \boldsymbol{\Lambda}\boldsymbol{f} + \boldsymbol{u}$$

, where $\boldsymbol{\Lambda}$ is a $q \times k$ matrix of factor loadings (a.k.a., the coefficients of the regression model), and the u is the vector of unexplained error of each observed variables.

Let's take the variance of the formula above

$$V(\boldsymbol{x}) = V(\boldsymbol{\Lambda}\boldsymbol{f} + \boldsymbol{u})$$

. Based on the operation rule of variance, like

$$V(a + b) = V(a) + V(b) + 2Cov(ab)$$

, we combined the two formulas above, then

$$V(\boldsymbol{x}) = V(\boldsymbol{\Lambda}\boldsymbol{f} + \boldsymbol{u}) = V(\boldsymbol{\Lambda}\boldsymbol{f}) + V(\boldsymbol{u}) + 2Cov(\boldsymbol{\Lambda}\boldsymbol{f}\boldsymbol{u})$$

. Since the we assumed that the error terms are uncorrelated with the factors, therefore the $Cov(\boldsymbol{\Lambda}\boldsymbol{f}\boldsymbol{u}) = 0$. Then, we can continue to drive the variance formula as

$$V(\boldsymbol{x}) = V(\boldsymbol{\Lambda}\boldsymbol{f}) + V(\boldsymbol{u}) = \boldsymbol{\Lambda}V(\boldsymbol{f})\boldsymbol{\Lambda}^T + \Psi$$

. In addition, we assumed that the factors are uncorrelated with each other. The $V(\boldsymbol{f})$ is actually an identity matrix. Therefore, the formula can be written as

$$V(\boldsymbol{x}) = \boldsymbol{\Lambda}V(\boldsymbol{f})\boldsymbol{\Lambda}^T + \Psi = \boldsymbol{\Lambda}\boldsymbol{\Lambda}^T + \Psi$$

. Finally, the formula can be written as

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}^T + \Psi$$

.

If we allow the factors to be correlated with each other, then the $V(\boldsymbol{f})$ is not an identity matrix. Let's use the greek letter Φ to represent the variance matrix of loadings **f**. Thus, the formula should be

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^T + \Psi$$

.

0.3   Ex 5.2

Show that the communalities in a factor analysis model are unaffected by the transformation ...

MY SOLUTION:
This question mentioned that we need to use the transformed factor loadings $\mathbf{\Lambda}^* = \mathbf{\Lambda M}$. Let's assume that $\mathbf{M}$ is an $k \times k$ orthogonal matrix. We can re-write the the basic regression equation linking the observed and the factors as:

$$\mathbf{x} = (\mathbf{\Lambda M})(\mathbf{M}^T \mathbf{f}) + \mathbf{u}$$

.

Using the rule of variance, we can have

$$\mathbf{\Sigma} = (\mathbf{\Lambda M})(\mathbf{\Lambda M})^T + \Psi$$

. Since the $\mathbf{M}$ is a orthogonal matrix and $\mathbf{MM}^T = \mathbf{I}$. Therefore, the variance equation can be written as

$$\mathbf{\Sigma} = \mathbf{\Lambda \Lambda}^T + \Psi$$

. That is, the transformed factor loadings $\mathbf{\Lambda}^* = \mathbf{\Lambda M}$ will not influence the communalities (i.e., $\mathbf{\Lambda \Lambda}^T$) in the a factor analysis model.

0.4   Ex 5.3

Give a formula for the proportion of variance explained by the jth factor estimated by the principal factor approach.

MY SOLUTION:
The proportion of variance explained by the jth factor represents the proportion of the total variance in the observed variables that is accounted for by that factor alone. Therefore, the formula could be

$$Proportion_j = \frac{\sum_{i=1}^{q} \lambda_{ij}^2}{\mathbf{\Lambda \Lambda}^T}$$