CrossMark

Erratum – A Two-step PLS Inspired Method for Linear Prediction with Group Effect

Ying Li and Peter Udén
Swedish University of Agricultural Sciences, Uppsala, Sweden
Dietrich von Rosen
Swedish University of Agricultural Sciences, Uppsala, Sweden and
Linköping University, Linköping, Sweden

In the original article cited above, it is claimed that maximum likelihood estimators (MLEs) have been obtained. Unfortunately, the proof of this statement is not correct and the proposed estimators will not be MLEs.

However, our estimator of $\boldsymbol{\Sigma}$ satisfies

$$n\boldsymbol{\Sigma} = (\mathbf{X} - \mathbf{A}\mathbf{B}_{\boldsymbol{\Sigma}}\mathbf{C})(\mathbf{X} - \mathbf{A}\mathbf{B}_{\boldsymbol{\Sigma}}\mathbf{C})', \tag{1}$$

$$\mathbf{A}\mathbf{B}_{\boldsymbol{\Sigma}}\mathbf{C} = \mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A})^{-}\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{X}\mathbf{C}'(\mathbf{C}\mathbf{C}')^{-}\mathbf{C}.$$

Therefore, given (1), all the results presented in the article are correct. In the following, we motivate that (1) is a reasonable basis for estimating $\boldsymbol{\Sigma}$.

The space of interest can be decomposed as

$$\mathbb{R}^p \otimes \mathbb{R}^n = \mathcal{C}_{\boldsymbol{\Sigma}}(\mathbf{A}) \otimes \mathcal{C}(\mathbf{C}') \boxplus \mathbb{R}^p \otimes \mathcal{C}(\mathbf{C}')^{\perp} \boxplus \mathcal{C}_{\boldsymbol{\Sigma}}(\mathbf{A})^{\perp} \otimes \mathcal{C}(\mathbf{C}'),$$

where $\mathbb{R}^p$ ($\mathbb{R}^n$) represent the whole space, $\mathcal{C}_{\boldsymbol{\Sigma}}(\bullet)$ denotes column vector space with inner product defined via $\boldsymbol{\Sigma}^{-1}$, $\mathcal{C}(\bullet)$ if the standard inner product is used, $\otimes$ denotes the tensor product of linear spaces and $\boxplus$ represents the orthogonal sum of linear spaces. The decomposition is illustrated in Fig. 1.

The estimation of the mean $\mathbf{ABC}$ is performed through projection of the data on the subspace $\mathcal{C}_{\boldsymbol{\Sigma}}(\mathbf{A}) \otimes \mathcal{C}(\mathbf{C}')$. The subspace of $\mathbb{R}^p \otimes \mathcal{C}(\mathbf{C}')^{\perp}$ and $\mathcal{C}_{\boldsymbol{\Sigma}}(\mathbf{A})^{\perp} \otimes \mathcal{C}(\mathbf{C}')$ are the spaces where residuals are constructed via projections, i.e.

$$\mathbf{R}_1 \;=\; \mathbf{X}(\mathbf{I} - \mathbf{P}_{\mathbf{C}'}), \qquad \mathbf{P}_{\mathbf{C}'} = \mathbf{C}'(\mathbf{C}\mathbf{C}')^{-}\mathbf{C},$$
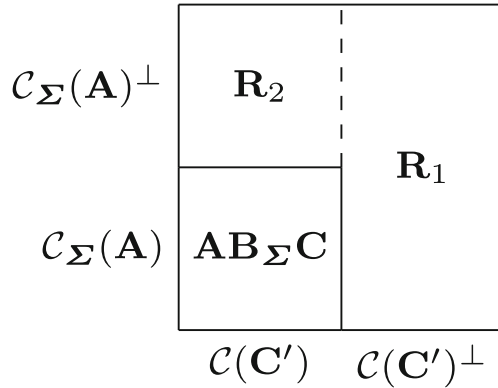
Figure 1: Decomposition of the tensor space of interest

$$\mathbf{R}_2 \;\;=\;\; (\mathbf{I} - \mathbf{P}_{\mathbf{A},\boldsymbol{\Sigma}})\mathbf{X}\mathbf{P}_{\mathbf{C}'} \qquad \mathbf{P}_{\mathbf{A},\boldsymbol{\Sigma}} = \mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A})^{-}\mathbf{A}'\boldsymbol{\Sigma}^{-1}.$$

By squaring the residuals a measure of variation can be obtained:

$$\mathbf{R}_1\mathbf{R}_1' + \mathbf{R}_2\mathbf{R}_2' = (\mathbf{X} - \mathbf{A}\mathbf{B}_{\boldsymbol{\Sigma}}\mathbf{C})(\mathbf{X} - \mathbf{A}\mathbf{B}_{\boldsymbol{\Sigma}}\mathbf{C})', \qquad (2)$$

which summarizes most of the variation in the data. Thus, the following estimating equation is proposed,

$$n\boldsymbol{\Sigma} = \mathbf{R}_1\mathbf{R}_1' + \mathbf{R}_2\mathbf{R}_2' \qquad (3)$$

which is identical to (1). We may note that $E[\mathbf{R}_1\mathbf{R}_1' + \mathbf{R}_2\mathbf{R}_2'] = n\boldsymbol{\Sigma}$. In the bilinear model (growth curve model), MANOVA model or Anova model, due to normality the variance is estimated through the residual space, i.e. (2), is identical to the maximum likelihood estimator. It turns out that in a special case we can show that the maximum likelihood estimator of $\boldsymbol{\Sigma}$ in the two-step method will be different from the one we have proposed since there is variation in the mean space $\mathcal{C}_{\boldsymbol{\Sigma}}(\mathbf{A}) \otimes \mathcal{C}(\mathbf{C}')$, due to that $\mathbf{A}$ is a function of $\boldsymbol{\Sigma}$, which has been neglected.

It is argued that our estimator, satisfying (1), is only based on the residual space, which is reasonable. Besides this, the numerical results presented in the article have also supported that our estimators works well.

Ying Li and Peter Udén
Swedish University of Agricultural
Sciences, Box 7032, 75007, Uppsala
Sweden
E-mail: ying.li@slu.se

Dietrich von Rosen
Swedish University of Agricultural
Sciences, Box 7032, 75007, Uppsala
Sweden and Department of
Mathematics, Linköping University
Linköping, Sweden