ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS DE MATHEMATICA
Volume 21, Number 2, December 2017
Available online at http://acutm.math.ut.ee

Comparison of estimators of variance parameters in the growth curve model with a special variance structure

Rastislav Rusnačko and Ivan Žežula

Abstract. Three different estimators of the variance parameters in the growth curve model with generalized uniform correlation structure are compared on the basis of mean square error. Since the situation in general depends on specific choice of the structure matrix, we investigate two important special cases.

1. Introduction

The growth curve model, introduced by Potthoff and Roy in [6], is one of fruitful models of multivariate analysis. It is of the form

$$Y = XBZ + \varepsilon, \qquad \mathrm{E}(\varepsilon) = 0, \qquad \mathrm{Var}(\mathrm{vec}\, \varepsilon) = \Sigma \otimes I,$$

where $Y_{n\times p}$ is the matrix of observations, $X_{n\times m}$ is an ANOVA matrix, $B_{m\times r}$ is a matrix of unknown parameters, $Z_{r\times p}$ is a matrix of regression constants, $\varepsilon_{n\times p}$ is a matrix of normally distributed random errors, $I_{n\times n}$ is the identity matrix, $\Sigma_{p\times p}$ is the variance matrix of rows of matrix Y, and vec operator transforms a matrix into a vector by stacking the columns one underneath the other. Usually, p is the number of time points, n is the number of subjects studied, and m is the number of groups.

Since the number of variance parameters in Σ grows quickly with p, the idea of using some special variance structures appeared. The first one to deal with such modification of the model was Khatri [3], who derived maximum likelihood estimators (MLE) of model parameters and likelihood ratio tests for block independence, sphericity, and intraclass model (in this article denoted by GUCS, see later). The next one was Lee [5], who considered the

Received June 29, 2015.
2010 *Mathematics Subject Classification.* Primary 62H12; Secondary 62J05.
Key words and phrases. Growth curve model, uniform correlation structure, generalized uniform correlation structure, estimators of variance parameters.
http://dx.doi.org/10.12697/ACUTM.2017.21.11

uniform correlation structure (UCS, see (1)) and the autoregressive correlation structure. He derived the MLEs for the unknown parameters in the UCS. Although MLEs in general have many optimal properties, these are assured only asymptotically, and their small sample properties can be much worse. That is why another principles of estimation of the UCS parameters, more suitable to small sample situation, were later used by Žežula [10] and Ye and Wang [8]. Both articles investigated also generalized uniform correlation structure (GUCS), see (2).

Žežula in [10] compared his estimators with MLEs. Later, Klein and Žežula [4] proved that Žežula's estimators and Ye and Wang's estimators are equivalent in the UCS model. However, the situation remains unclear in the GUCS model. In this article, we compare estimators of Khatri, Žežula, and Ye and Wang on the basis of their variances and mean square errors in the GUCS model.

2. The uniform correlation structure

If p, the dimension of observations made on a single subject, is not large, and the number of observations n is substantially bigger, there are more suitable estimators of Σ at our disposition. However, when the number of unknown parameters in the variance matrix Σ is close to or even bigger than sample size, the situation gets complicated. One of the possible solutions is to reduce the number of unknown parameters by considering a simpler variance structure. Such a structure may also be naturally implied by the nature of the data, e.g., when the observations are autoregressive time series or a mixture of several populations. One of such simple models is the model with the uniform correlation structure given as

$$\Sigma = \sigma^2 \left[(1-\varrho) I + \varrho \mathbf{1}\mathbf{1}' \right], \tag{1}$$

where $\sigma^2 > 0$ is the common variance of all observations, $\rho \in (-1/(p-1); 1)$ correlation coefficient of any pair in a row, and $\mathbf{1}$ vector of ones.

Under normality, we know that the uniformly minimum variance unbiased invariant estimator of unstructured Σ is

$$S = \frac{1}{n - r(X)} Y' M_X Y,$$

where $r(X)$ is the rank of X and $M_X = I - P_X = I - X(X'X)^- X$, see [9]. We establish estimator of structured matrix Σ on the basis of estimator of unstructured matrix Σ. Since $\mathrm{E}(S) = \Sigma$ holds regardless of the true value of Σ, it can be reasonable to base estimators of the parameters of the simplified model also on S.

Using the moment method, we easily get unbiased estimating equations

$$\mathrm{Tr}(S) - p\widehat{\sigma}^2 = 0 \quad \text{and} \quad \mathbf{1}'S\mathbf{1} - \mathrm{Tr}(S)[1 + (p-1)\widehat{\varrho}] = 0$$

for the parameters σ^2 and ϱ. Thus, natural simple estimators of the two unknown variance structure parameters are

$$\widehat{\sigma}^2 = \frac{\mathrm{Tr}(S)}{p} \quad \text{and} \quad \widehat{\varrho} = \frac{1}{p-1}\left(\frac{\mathbf{1}'S\mathbf{1}}{\mathrm{Tr}(S)} - 1\right).$$

It holds that $\widehat{\sigma}^2 \geq 0$ and $\widehat{\rho} \in \langle -1/(p-1); 1\rangle$, which must hold also for their true values. These estimators have been derived by Žežula in [10]. Later Ye and Wang [8] derived seemingly different estimators by the decomposition of the model into two parts, one parallel and the other perpendicular to vector $\mathbf{1}$, again using the moment method. Equivalence of these estimators with the previous ones was proved by Klein and Žežula in [4].

3. The generalized uniform correlation structure

To keep the simplicity, but to allow for different variances and correlations, if we know their ratios, the structure of Σ can be considered in the form

$$\Sigma = \theta_1 G + \theta_2 ww', \tag{2}$$

called the generalized uniform correlation structure, where $G_{p\times p} \geq 0$ is a known symmetric matrix, $w \in \mathbb{R}^p$ is a known vector, and the variance-covariance parameters θ_1 and θ_2 are unknown.

The GUCS is indeed a generalization of UCS. E.g., if we have design with constant correlation ρ but different individual variances $c_1^2\sigma^2, \ldots, c_p^2\sigma^2$, then

$$G = \mathrm{diag}\left\{c_1^2, \ldots, c_p^2\right\}, \;\; w = (c_1, \ldots, c_p)', \;\; \theta_1 = \sigma^2(1-\rho), \;\; \theta_2 = \sigma^2\rho.$$

Thus, the GUCS allows inhomogeneity of variances and covariances, while saving parameter parsimony.

Should Σ be positive-semidefinite, one of the following conditions must hold:

$$\theta_1 \geq 0 \;\&\; \theta_2 \geq 0\,,$$

or

$$\theta_1 \geq 0, \theta_2 < 0, w \in \mathcal{R}(G) \;\&\; \theta_1 \geq |\theta_2| w'G^+w\,, \tag{3}$$

where G^+ is the Moore–Penrose g-inverse (MP-inverse) of G.

Matrix G expresses our knowledge on ratios of variances of different row elements, up to the the diagonal multiples of w_i^2 (all rows of Y have the same variance matrix). Usually we take G diagonal, and our knowledge of correlations ratios is expressed by vector w.

If we have general $G \geq 0$, we can use its spectral decomposition $G = U\Lambda U'$ and transform the original model into $YU = XBZU + \varepsilon U$. Since matrix of eigenvectors U is orthogonal, this model is equivalent to the original one (all distances and vector angles are preserved), and the rows of the transformed Y have variance matrix $U'\Sigma U = \theta_1\Lambda + \theta_2 vv'$ with $v = U'w$. Thus, it is sufficient to consider G diagonal.

Notice that in (2) we can restrict both norm of G and w arbitrarily, since

$$\theta_1 G = (a_1\theta_1)\left(\frac{1}{a_1}G\right) = \theta_1^* G^*, \quad \text{and} \quad \theta_2 ww' = (a_2\theta_2)\left(\frac{1}{a_2}ww'\right) = \theta_2^* w^* w^{*\prime}.$$

We will make use of it later.

Similarly to the UCS model, from

$$\mathrm{E}\left[\mathrm{Tr}(S)\right] = \theta_1 \,\mathrm{Tr}(G) + \theta_2 w'w \quad \text{and} \quad \mathrm{E}(\mathbf{1}'S\mathbf{1}) = \theta_1 \mathbf{1}'G\mathbf{1} + \theta_2(\mathbf{1}'w)^2\,,$$

we get unbiased estimating equations, which lead to the estimators

$$\widehat{\theta}_1^Z = \frac{(\mathbf{1}'w)^2\,\mathrm{Tr}(S) - \mathbf{1}'S\mathbf{1}w'w}{(\mathbf{1}'w)^2\,\mathrm{Tr}(G) - \mathbf{1}'G\mathbf{1}w'w} \quad \text{and} \quad \widehat{\theta}_2^Z = \frac{\mathbf{1}'S\mathbf{1}\,\mathrm{Tr}(G) - \mathbf{1}'G\mathbf{1}\,\mathrm{Tr}(S)}{(\mathbf{1}'w)^2\,\mathrm{Tr}(G) - \mathbf{1}'G\mathbf{1}w'w}$$

(see [10]). We will call them Z-estimators. Note that if G is diagonal, then $\mathrm{Tr}(G) = \mathbf{1}'G\mathbf{1}$ and the second estimator gets free of G:

$$\widehat{\theta}_{2d}^Z = \frac{\mathbf{1}'S\mathbf{1} - \mathrm{Tr}(S)}{(\mathbf{1}'w)^2 - w'w}\,.$$

Using similar principle in an orthogonal decomposition of the model, assuming $G > 0$, Ye and Wang in [8] derived alternative estimators of the form

$$\widehat{\theta}_1^{YW} = \frac{w'G^{-1}w\,\mathrm{Tr}(G^{-1}S) - w'G^{-1}SG^{-1}w}{(p-1)(w'G^{-1}w)},$$
$$\widehat{\theta}_2^{YW} = \frac{p\,w'G^{-1}SG^{-1}w - w'G^{-1}w\,\mathrm{Tr}(G^{-1}S)}{(p-1)(w'G^{-1}w)^2}.$$

We will call them YW-estimators.

Khatri in [3] derived ML estimators, which under condition that $w \in \mathcal{R}(Z')$ have the form

$$\widehat{\theta}_1^{ML} = \frac{n - r(X)}{n} \cdot \frac{wG^{-1}w\,\mathrm{Tr}\left(G^{-1}S\right) - wG^{-1}SG^{-1}w}{(p-1)wG^{-1}w} + \frac{\mathrm{Tr}\left(G^{-1}M_{Z'}^{G^{-1}}Y'P_XY\right)}{n(p-1)},$$

$$\widehat{\theta}_2^{ML} = \frac{n - r(X)}{n} \cdot \frac{p\,wG^{-1}SG^{-1}w - wG^{-1}w\,\mathrm{Tr}\left(G^{-1}S\right)}{(p-1)\left(wG^{-1}w\right)^2} - \frac{\mathrm{Tr}\left(G^{-1}M_{Z'}^{G^{-1}}Y'P_XY\right)}{n(p-1)wG^{-1}w}.$$

Unbiasedness of S implies that both Z- and YW-estimators are unbiased. Using results of [2], we easily get that MLEs are biased, and both underestimate the true value:

$$E\widehat{\theta}_1^{ML} = \theta_1\left(1 - \frac{r(X)(r(Z')-1)}{n(p-1)}\right),$$
$$E\widehat{\theta}_2^{ML} = \theta_2\frac{n-r(X)}{n} - \theta_1\frac{r(X)(p-r(Z'))}{n(p-1)}.$$

That is not surprising. However, there is a bigger problem with $\widehat{\theta}_2^{ML}$ – the mean depends also on θ_1, which can lead to severe bias especially in small samples.

Quality of all these estimators can be compared by their mean square errors (MSE), which for Z- and YW-estimators are equal to corresponding variances. According to [4], it holds

$$\operatorname{Var}\widehat{\theta}_1^Z = \frac{2\Big[(\mathbf{1}'w)^4\operatorname{Tr}(\Sigma^2) - 2(\mathbf{1}'w)^2w'w\mathbf{1}'\Sigma^2\mathbf{1} + (w'w)^2(\mathbf{1}'\Sigma\mathbf{1})^2\Big]}{(n-r(X))\left[(\mathbf{1}'w)^2\operatorname{Tr}(G) - \mathbf{1}'G\mathbf{1}w'w\right]^2},$$

$$\operatorname{Var}\widehat{\theta}_2^Z = \frac{2\Big[\left[\operatorname{Tr}(G)\mathbf{1}'\Sigma\mathbf{1}\right]^2 - 2\operatorname{Tr}(G)\mathbf{1}'G\mathbf{1}\mathbf{1}'\Sigma^2\mathbf{1} + (\mathbf{1}'G\mathbf{1})^2\operatorname{Tr}(\Sigma^2)\Big]}{(n-r(X))\left[(\mathbf{1}'w)^2\operatorname{Tr}(G) - \mathbf{1}'G\mathbf{1}w'w\right]^2}.$$

Inserting the formula for Σ, we get alternative expressions

$$\begin{aligned}\operatorname{Var}\widehat{\theta}_1^Z = {} & \frac{2\theta_1^2\left[(\mathbf{1}'w)^4\operatorname{Tr}\left(G^2\right) - 2(\mathbf{1}'w)^2(w'w)\mathbf{1}'G^2\mathbf{1} + (w'w)^2(\mathbf{1}'G\mathbf{1})^2\right]}{(n-r(X))\left[(\mathbf{1}'w)^2\operatorname{Tr}(G) - \mathbf{1}'G\mathbf{1}w'w\right]^2} \\ & + \frac{4\theta_1\theta_2\left[(\mathbf{1}'w)^4w'Gw - 2(\mathbf{1}'w)^3(w'w)\mathbf{1}'Gw + (\mathbf{1}'w)^2(w'w)^2\mathbf{1}'G\mathbf{1}\right]}{(n-r(X))\left[(\mathbf{1}'w)^2\operatorname{Tr}(G) - \mathbf{1}'G\mathbf{1}w'w\right]^2}\end{aligned}$$

and

$$\begin{aligned}\operatorname{Var}\widehat{\theta}_2^Z = {} & \frac{2\theta_2^2}{n-r(X)} \\ & + \frac{2\theta_1^2\left[(\mathbf{1}'G\mathbf{1})^2\operatorname{Tr}^2(G) - 2(\mathbf{1}'G\mathbf{1})\operatorname{Tr}(G)\mathbf{1}'G^2\mathbf{1} + \operatorname{Tr}(G^2)(\mathbf{1}'G\mathbf{1})^2\right]}{(n-r(X))\left[(\mathbf{1}'w)^2\operatorname{Tr}(G) - \mathbf{1}'G\mathbf{1}w'w\right]^2} \\ & + \frac{4\theta_1\theta_2}{(n-r(X))\left[(\mathbf{1}'w)^2\operatorname{Tr}(G) - \mathbf{1}'G\mathbf{1}w'w\right]^2}\Big[(\mathbf{1}'G\mathbf{1})^2\operatorname{Tr}^2(G)(\mathbf{1}'w)^2 \\ & \qquad -2(\mathbf{1}'w)(\mathbf{1}'Gw)(\mathbf{1}'G\mathbf{1})\operatorname{Tr}(G) + (\mathbf{1}'G\mathbf{1})^2w'Gw\Big].\end{aligned}$$

Similarly,

$$\operatorname{Var}\widehat{\theta}_1^{YW} = \frac{2}{(n-r(X))(p-1)^2(w'G^{-1}w)^2}\Big[(w'G^{-1}w)^2\operatorname{Tr}(G^{-1}\Sigma G^{-1}\Sigma)$$

$$+(w'G^{-1}\Sigma G^{-1}w)^2 - 2(w'G^{-1}w)(w'G^{-1}\Sigma G^{-1}\Sigma G^{-1}w)\Big]$$
$$= \frac{2\theta_1^2}{(n-r(X))\,(p-1)},$$
$$\operatorname{Var}\widehat{\theta}_2^{YW} = \frac{2}{(n-r(X))\,(p-1)^2(w'G^{-1}w)^4}\Big[(w'G^{-1}w)^2\operatorname{Tr}(G^{-1}\Sigma G^{-1}\Sigma)$$
$$+p^2(w'G^{-1}\Sigma G^{-1}w)^2 - 2p(w'G^{-1}w)(w'G^{-1}\Sigma G^{-1}\Sigma G^{-1}w)\Big]$$
$$= \frac{2\left[p\theta_1^2 + 2(p-1)\left(w'G^{-1}w\right)\theta_1\theta_2 + (p-1)\left(w'G^{-1}w\right)^2\theta_2^2\right]}{(n-r(X))\,(p-1)\,(w'G^{-1}w)^2}.$$

Since the last term in MLEs uses projector P_X, the other two M_X, and $M_X P_X = 0$, one can quickly find that the last term is independent of the previous ones. Thus, using results of [2], we get

$$\operatorname{Var}\widehat{\theta}_1^{ML} = \left(\frac{n-r(X)}{n}\right)^2 \operatorname{Var}\widehat{\theta}_1^{YW} + 2r(X)\frac{\operatorname{Tr}\left(G^{-1}M_{Z'}^{G^{-1}}\Sigma G^{-1}M_{Z'}^{G^{-1}}\Sigma\right)}{n^2(p-1)^2}$$
$$= \frac{2\theta_1^2\left[(n-r(X))\,(p-1) + r(X)(p-r(Z'))\right]}{n^2(p-1)^2},$$
$$\operatorname{Var}\widehat{\theta}_2^{ML} = \left(\frac{n-r(X)}{n}\right)^2 \operatorname{Var}\widehat{\theta}_2^{YW} + 2r(X)\frac{\operatorname{Tr}\left(G^{-1}M_{Z'}^{G^{-1}}\Sigma G^{-1}M_{Z'}^{G^{-1}}\Sigma\right)}{n^2(p-1)^2\left(w'G^{-1}w\right)^2}.$$

Then, their MSEs are

$$\mathrm{MSE}\widehat{\theta}_1^{ML} = \left(\frac{n-r(X)}{n}\right)^2 \operatorname{Var}\widehat{\theta}_1^{YW}$$
$$+\ \frac{2r(X)\operatorname{Tr}\left(G^{-1}M_{Z'}^{G^{-1}}\Sigma G^{-1}M_{Z'}^{G^{-1}}\Sigma\right) + r(X)^2(r(Z')-1)^2\theta_1^2}{n^2(p-1)^2},$$
$$\mathrm{MSE}\widehat{\theta}_2^{ML} = \left(\frac{n-r(X)}{n}\right)^2 \operatorname{Var}\widehat{\theta}_2^{YW}$$
$$+\ 2r(X)\frac{\operatorname{Tr}\left(G^{-1}M_{Z'}^{G^{-1}}\Sigma G^{-1}M_{Z'}^{G^{-1}}\Sigma\right)}{n^2(p-1)^2\left(w'G^{-1}w\right)^2} + \frac{r(X)^2}{n^2}\left(\theta_2 + \frac{p-r(Z')}{p-1}\theta_1\right)^2.$$

Again, $\operatorname{Tr}\left(G^{-1}M_{Z'}^{G^{-1}}\Sigma G^{-1}M_{Z'}^{G^{-1}}\Sigma\right)$ depends on θ_1, which is a problem for $\widehat{\theta}_2^{ML}$. To get the MSEs, we need to add squared biases, which in case of $\widehat{\theta}_2^{ML}$ will strengthen the dependence on θ_1. We see now that while $\widehat{\theta}_1^{ML}$ can be comparable with the YW-estimator (two more terms can be compensated by $(n-r(X)/n)^2$ coefficient at the first term), $\widehat{\theta}_2^{ML}$ is a dangerous estimator in small samples. That is why we will compare only Z- and YW-estimators in the next.

For the comparison, we define the following two functions:

$$r_1 = \mathrm{Var}\ \widehat{\theta}_1^Z - \mathrm{Var}\ \widehat{\theta}_1^{YW} \quad \text{and} \quad r_2 = \mathrm{Var}\ \widehat{\theta}_2^Z - \mathrm{Var}\ \widehat{\theta}_2^{YW}.$$

We will consider them not as functions of the unknown variance components, but – for fixed values of these components – as functions of the elements of G and w.

There cannot be said much about these differences in general. That is why we will look at some special cases in the following subsections. In order to keep the structure as simple as possible, in accordance with our remarks above, we choose G diagonal in all cases.

3.1. The GUCS model with G diagonal and $w = \mathbf{1}$. Choosing $w = \mathbf{1}$ means that we suppose uniform correlation structure, but not necessarily with equal variances in all components. Due to multiplication by an unknown parameter, this model is equivalent with the one using $w = c\mathbf{1}$ for some c.

Let us suppose that measurements in different times have all potentially different precisions/variances $x_i > 0$, $i = 1, \ldots, p$. Then we have $G = \mathrm{diag}(\boldsymbol{x})$, where $\boldsymbol{x} = (x_1, \ldots, x_p)'$ is the vector of variances. It is easy to get the relations

$$\begin{aligned}
\mathrm{Tr}\,(\Sigma) = \boldsymbol{x}'\mathbf{1}\theta_1 + p\theta_2,\ \ \mathrm{Tr}\left(\Sigma^2\right) &= \boldsymbol{x}'\boldsymbol{x}\theta_1^2 + 2\boldsymbol{x}'\mathbf{1}\theta_1\theta_2 + p^2\theta_2^2,\\
\mathbf{1}'\Sigma\mathbf{1} = \boldsymbol{x}'\mathbf{1}\theta_1 + p^2\theta_2,\ \ \mathbf{1}'\Sigma^2\mathbf{1} &= \boldsymbol{x}'\boldsymbol{x}\theta_1^2 + 2p\boldsymbol{x}'\mathbf{1}\theta_1\theta_2 + p^3\theta_2^2.
\end{aligned}$$

Denoting by $\boldsymbol{x}^{-1} = (1/x_1, \ldots, 1/x_p)'$ the vector of reciprocals of $\boldsymbol{x}$, we have $G^{-1} = \mathrm{diag}(\boldsymbol{x}^{-1})$, and

$$\begin{aligned}
\mathbf{1}'G^{-1}\mathbf{1} = \mathbf{1}'\boldsymbol{x}^{-1},\ \ \mathbf{1}'G^{-1}\Sigma G^{-1}\mathbf{1} &= \mathbf{1}'\boldsymbol{x}^{-1}\theta_1 + (\mathbf{1}'\boldsymbol{x}^{-1})^2\theta_2,\\
\mathbf{1}'G^{-1}\Sigma G^{-1}\Sigma G^{-1}\mathbf{1} &= \mathbf{1}'\boldsymbol{x}^{-1}\theta_1^2 + 2(\mathbf{1}'\boldsymbol{x}^{-1})^2\theta_1\theta_2 + (\mathbf{1}'\boldsymbol{x}^{-1})^3\theta_2^2,\\
\mathrm{Tr}\left(G^{-1}\Sigma G^{-1}\Sigma\right) &= p\theta_1^2 + 2\,\mathbf{1}'\boldsymbol{x}^{-1}\theta_1\theta_2 + (\mathbf{1}'\boldsymbol{x}^{-1})^2\theta_2^2.
\end{aligned}$$

Using the above formulas, after short calculation we get

$$\begin{aligned}
\mathrm{Var}\,\widehat{\theta}_1^Z &= \frac{2\theta_1^2\left[p(p-2)\boldsymbol{x}'\boldsymbol{x} + (\mathbf{1}'\boldsymbol{x})^2\right]}{(n-r(X))(p-1)^2(\mathbf{1}'\boldsymbol{x})^2},\\
\mathrm{Var}\,\widehat{\theta}_2^Z &= \frac{2\left[\left((\mathbf{1}'\boldsymbol{x})^2 - \boldsymbol{x}'\boldsymbol{x}\right)\theta_1^2 + 2\left(p-1\right)^2(\mathbf{1}'\boldsymbol{x})\,\theta_1\theta_2 + p^2(p-1)^2\theta_2^2\right]}{(n-r(X))p^2(p-1)^2},\\
\mathrm{Var}\,\widehat{\theta}_1^{YW} &= \frac{2\theta_1^2}{(n-r(X))(p-1)},\\
\mathrm{Var}\,\widehat{\theta}_2^{YW} &= \frac{2\left[p\theta_1^2 + 2(p-1)\mathbf{1}'\boldsymbol{x}^{-1}\theta_1\theta_2 + (p-1)(\mathbf{1}'\boldsymbol{x}^{-1})^2\theta_2^2\right]}{(n-r(X))(p-1)(\mathbf{1}'\boldsymbol{x}^{-1})^2}.
\end{aligned}$$

Then, the differences r_1 and r_2 as functions of $\boldsymbol{x}$ are

$$r_1(\boldsymbol{x}) = \frac{2\theta_1^2(p-2)\left(p\boldsymbol{x}'\boldsymbol{x} - (\mathbf{1}'\boldsymbol{x})^2\right)}{(n-r(X))(p-1)^2(\mathbf{1}'\boldsymbol{x})^2}, \tag{4}$$

$$\begin{aligned} r_2(\boldsymbol{x}) = {} & \frac{2\theta_1^2\left[\left(\mathbf{1}'\boldsymbol{x}^{-1}\right)^2\left(\left(\mathbf{1}'\boldsymbol{x}\right)^2 - \boldsymbol{x}'\boldsymbol{x}\right) - p^3(p-1)\right]}{(n-r(X))p^2(p-1)^2\left(\mathbf{1}'\boldsymbol{x}^{-1}\right)^2} \\ & + \frac{4\theta_1\theta_2(p-1)^2\left(\mathbf{1}'\boldsymbol{x}^{-1}\mathbf{1}'\boldsymbol{x} - p^2\right)}{(n-r(X))p^2(p-1)^2\mathbf{1}'\boldsymbol{x}^{-1}}. \end{aligned} \tag{5}$$

Lemma 3.1. *Under the above assumptions, it is $r_1(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} > 0$.*[1]

Proof. The denominator of $r_1(\boldsymbol{x})$ is positive for $\boldsymbol{x} > 0$, so that we have only to look at the numerator. It must be $p \geq 2$ for the model to have sense. The rest immediately follows from Schwarz inequality $(\mathbf{1}'\boldsymbol{x})^2 \leq \mathbf{1}'\mathbf{1} \cdot \boldsymbol{x}'\boldsymbol{x} = p\,\boldsymbol{x}'\boldsymbol{x}$. □

Thus, YW-estimator of θ_1 is always at least as good as Z-estimator in this model. Note that $r_1(\boldsymbol{x}) \equiv 0$ for $p = 2$.

Since equality occurs in Schwarz inequality if and only if $\boldsymbol{x} = c\mathbf{1}$ for some c, $r_1(\boldsymbol{x})$ attains zero values on $M = \{\boldsymbol{x} = c\mathbf{1}; c > 0\}$ for arbitrary p.

Let us now denote S_k the arithmetic average of all possible products of k elements out of $x_1, \ldots, x_p$ (k-th symmetric polynomial mean) for $k = 1, 2, \ldots, p$. We express the terms in $r_2(\boldsymbol{x})$ by means of them. Because $\boldsymbol{x} > 0$, $S_k > 0$ for all k. Then we have

$$\mathbf{1}'\boldsymbol{x} = \sum_{i=1}^{p} x_i = pS_1\,,$$

$$\mathbf{1}'\boldsymbol{x}^{-1} = \sum_{i=1}^{p} x_i^{-1} = \frac{1}{\prod_{j=1}^{p} x_j}\sum_{i=1}^{p}\frac{\prod_{j=1}^{p} x_j}{x_i} = \frac{1}{\prod_{j=1}^{p} x_j}\sum_{i=1}^{p}\prod_{\substack{j=1\\ j\neq i}}^{p} x_j = \frac{pS_{p-1}}{S_p}\,,$$

$$\left(\mathbf{1}'\boldsymbol{x}\right)^2 - \boldsymbol{x}'\boldsymbol{x} = \left(\sum_{i=1}^{p} x_i\right)^2 - \sum_{i=1}^{p} x_i^2 = \sum_{\substack{i=1\\ j\neq i}}^{p}\sum_{j=1}^{p} x_i x_j = p(p-1)S_2\,.$$

Lemma 3.2. *Under the above assumptions, $r_2(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} > 0$.*

Proof. As before, the denominators of both fractions in (5) are positive for $\boldsymbol{x} > 0$, so that it is sufficient to look at the numerators.

a) Let us first consider the case $\theta_1, \theta_2 \geq 0$. We will look at the two terms separately. Since

$$\left(\mathbf{1}'\boldsymbol{x}^{-1}\right)^2\left(\left(\mathbf{1}'\boldsymbol{x}\right)^2 - \boldsymbol{x}'\boldsymbol{x}\right) - p^3(p-1) = p^3(p-1)\left(\frac{S_{p-1}^2 S_2}{S_p^2} - 1\right),$$

[1] Taken element-wise.

we need to prove that $S_{p-1}^2 S_2 - S_p^2 \geq 0$. It is known that $S_k^2 \geq S_{k-1}S_{k+1}$ for $1 \leq k \leq p-1$ (Theorem 1, p. 324 in [1]), and $S_p \leq S_k S_{p-k}$ for $1 \leq k \leq p-1$ (Corollary 2(c), p. 326 in [1]). Using these inequalities, we get

$$S_{p-1}^2 S_2 - S_p^2 \geq S_{p-2}S_p S_2 - S_p^2 = S_p\left(S_{p-2}S_2 - S_p\right) \geq 0.$$

Similarly,

$$\mathbf{1}'\boldsymbol{x}^{-1}\mathbf{1}'\boldsymbol{x} - p^2 = p^2\left(\frac{S_1 S_{p-1}}{S_p} - 1\right) = p^2\frac{S_1 S_{p-1} - S_p}{S_p} \geq 0.$$

Thus, both terms are non-negative and the result holds.

b) Let us now consider the case $\theta_1 \geq 0, \theta_2 < 0$. According to (3), Σ is p.s.d. for $\theta_2 < 0$ if and only if $\theta_1 \geq |\theta_2|\,\mathbf{1}'\operatorname{diag}\left(\boldsymbol{x}^{-1}\right)\mathbf{1} = |\theta_2|\,\mathbf{1}'\boldsymbol{x}^{-1}$. Since

$$\begin{aligned} r_2(\boldsymbol{x}) = \frac{2\theta_1^2}{(n-r(X))p^2(p-1)^2\left(\mathbf{1}'\boldsymbol{x}^{-1}\right)^2} &\Big[\left(\mathbf{1}'\boldsymbol{x}^{-1}\right)^2\left(\left(\mathbf{1}'\boldsymbol{x}\right)^2 - \boldsymbol{x}'\boldsymbol{x}\right) \\ &- p^3(p-1) - \frac{|\theta_2|\,\mathbf{1}'\boldsymbol{x}^{-1}}{\theta_1}2(p-1)^2\left(\mathbf{1}'\boldsymbol{x}^{-1}\mathbf{1}'\boldsymbol{x} - p^2\right)\Big] \end{aligned}$$

and the fraction $\frac{|\theta_2|\mathbf{1}'\boldsymbol{x}^{-1}}{\theta_1} \leq 1$ can be arbitrarily close to 1, it is clear that $r_2(\boldsymbol{x}) \geq 0$ if and only if

$$\left(\mathbf{1}'\boldsymbol{x}^{-1}\right)^2\left(\left(\mathbf{1}'\boldsymbol{x}\right)^2 - \boldsymbol{x}'\boldsymbol{x}\right) - p^3(p-1) - 2(p-1)^2\left(\mathbf{1}'\boldsymbol{x}^{-1}\mathbf{1}'\boldsymbol{x} - p^2\right) \geq 0\,.$$

Thus, we need to prove that

$$\left(\mathbf{1}'\boldsymbol{x}^{-1}\right)^2\left(\left(\mathbf{1}'\boldsymbol{x}\right)^2 - \boldsymbol{x}'\boldsymbol{x}\right) + p^2(p-1)(p-2) - 2(p-1)^2\mathbf{1}'\boldsymbol{x}^{-1}\mathbf{1}'\boldsymbol{x} \geq 0, \quad \boldsymbol{x} > 0.$$

It is easy to verify that the left-hand side of the above inequality is equal to 0 for $p = 2$. That means that the non-negativity of $r_2(\boldsymbol{x})$ is equivalent to the positive-semidefiniteness of Σ. Thus, the inequality holds for $p = 2$.

Now take $p > 2$. The desired inequality is of the form

$$p^3(p-1)\frac{S_{p-1}^2 S_2}{S_p^2} + p^2(p-1)(p-2) - 2p(p-1)^2\frac{S_{p-1}S_1}{S_p} \geq 0, \quad \boldsymbol{x} > 0\,,$$

which is equivalent to

$$p^2 S_{p-1}^2 S_2 + p(p-2)S_p^2 - 2(p-1)S_{p-1}S_1 S_p \geq 0, \quad \boldsymbol{x} > 0\,.$$

Since Corollary 2(a) on page 326 in [1] implies that $S_2 S_{p-1} \geq S_1 S_p$ for $p > 2$, and the left-hand side of the previous inequality can be written in the form

$$p^2 S_{p-1}\left(S_{p-1}S_2 - S_1 S_p\right) + p(p-2)S_p^2 + \left[(p-1)^2 + 1\right]S_{p-1}S_1 S_p\,,$$

the non-negativity of all three terms for all $p > 2$ is now obvious. □

As a consequence, the YW-estimator of θ_2 is always at least as good as Z-estimator. It is easy to verify that also here $r_2(\boldsymbol{x})$ attains zero values only on $M = \{\boldsymbol{x} = c\mathbf{1}; c > 0\}$.

3.1.1. *A special case.* We can also be interested in the question, how big can be the differences $r_1(\boldsymbol{x})$ and $r_2(\boldsymbol{x})$. Let us investigate it on a simple special case of the above model. Let

$$\boldsymbol{x} = (\underbrace{a, \dots, a}_{k \text{ times}}, \underbrace{1, \dots, 1}_{p-k \text{ times}})', \;\; 1 \le k < p.$$

It means that the variance matrix Σ contains diagonal elements $\theta_1 a + \theta_2$ and θ_2, and all off-diagonal elements are θ_2. Formulas (4) and (5) give after some algebra

$$r_1(\boldsymbol{x}) = r_1(a) = \frac{2k(p-k)(p-2)\theta_1^2}{(n-r(X))(p-1)^2(ka+p-k)^2}(a-1)^2$$

and

$$\begin{aligned} r_2(\boldsymbol{x}) = r_2(a) = {} & \frac{2k\theta_1}{(n-r(X))\left[p(p-1)\left[(k-p)a-k\right]\right]^2}\Big\{(p-k)^2(k-1)\theta_1 a^4 \\ & + 2(p-k)\left[\theta_1\left[(p-k)^2+k(k-1)\right]+\theta_2(p-k)(p-1)^2\right]a^3 \\ & + \left[\theta_1\left[p^2(3-4p)+kp(10p-12k-3)+6k^3\right]\right. \\ & \left. -2\theta_2(p-1)^2(2p^2-5kp+3k^2)\right]a^2 \\ & + 2\left[\theta_1(p-k)\left[(p-k)^2-(p-k)+k^2\right]+\theta_2(p-1)^2\left[(p-2k)^2-k^2\right]\right]a \\ & + k(p-k)\left[\theta_1(p-k-1)+2\theta_2(p-1)^2\right]\Big\}. \end{aligned}$$

We see that function $r_1(a)$ has one zero point $a = 1$. It is because in that case the model turns into usual uniform correlation structure model with $\theta_1 = \sigma^2(1-\varrho)$ and $\theta_2 = \sigma^2\varrho$, so that Z-estimators and YW-estimators are the same. We can see a typical behaviour of $r_1(a)$ in Figure 1. Let us notice that this difference is bounded from above, since

$$\lim_{a\to\infty} r_1(a) = \frac{2(p-k)(p-2)\theta_1^2}{(n-r(X))k(p-1)^2}.$$

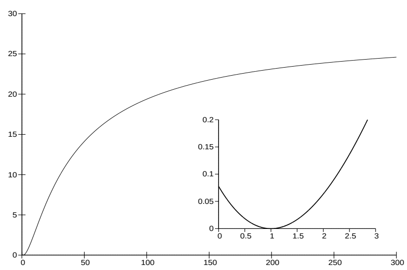

Figure 1: behaviour of function $r_1(a)$ (with detail of neighbourhood of 1).

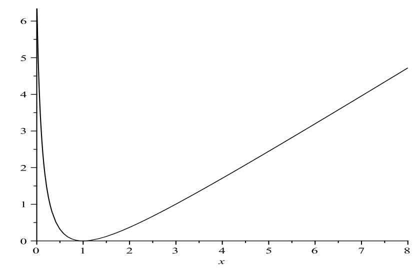

Figure 2: behaviour of function $r_2(a)$.

Also function $r_2(a)$ has one zero point $a = 1$ on the interval $\langle 0, \infty)$. Its typical shape is depicted on Figure 2. In this case $\lim\limits_{a\to\infty} r_2(a) = \infty$, so that the difference is not bounded. Thus, the difference in precision can be much bigger in the case of θ_2.

If $\theta_1 \geq 0$ and $\theta_2 < 0$, then condition (3) turns into $\theta_1 > |\theta_2|(p-k)$, and the range of a is restricted to $a \in \left\langle k\left(\theta_1/|\theta_2| - p + k\right)^{-1}, \infty\right)$.

Some other special cases were investigated by Rusnačko in [7].

3.2. The GUCS model with $\boldsymbol{G = I}$ and non-constant $\boldsymbol{w}$. Let us now consider the GUCS model with $G = I_p$ and any w such that $w'w = 1$. Let us denote $w_0 = \mathbf{1}'w$. Since $\mathrm{Tr}(G) = \mathbf{1}'G\mathbf{1} = \mathbf{1}'\mathbf{1} = p$, it is easy to get the relations

$$\begin{aligned} \mathrm{Tr}\,(\Sigma) = p\theta_1 + \theta_2,\ \ \mathrm{Tr}\left(\Sigma^2\right) = p\theta_1^2 + 2\theta_1\theta_2 + \theta_2^2, \\ \mathbf{1}'\Sigma\mathbf{1} = p\theta_1 + w_0^2\theta_2,\ \ \mathbf{1}'\Sigma^2\mathbf{1} = p\theta_1^2 + 2w_0^2\theta_1\theta_2 + w_0^2\theta_2^2. \end{aligned}$$

Also, it holds

$$\begin{aligned} w'G^{-1}w = w'w = 1,\ \ w'G^{-1}\Sigma G^{-1}w = w'\Sigma w = \theta_1 + \theta_2, \\ \mathrm{Tr}\left(G^{-1}\Sigma G^{-1}\Sigma\right) = \mathrm{Tr}\left(\Sigma^2\right), \\ w'G^{-1}\Sigma G^{-1}\Sigma G^{-1}w = w'\Sigma^2 w = \theta_1^2 + 2\theta_1\theta_2 + \theta_2^2 = \left(\theta_1 + \theta_2\right)^2. \end{aligned}$$

Then, after some algebra we get

$$r_1(w) = r_2(w) = \frac{2\left(p - w_0^2\right)\left[p\left(p - 2 + w_0^2\right)\theta_1^2 + 2(p-1)w_0^2\theta_1\theta_2\right]}{(p-1)(n - r(X))\left(w_0^2 - 1\right)^2}.$$

Lemma 3.3. *Under the above assumptions, it holds $r_1(w) = r_2(w) \geq 0$ for all w and for all admissible values of θ_1, θ_2.*

Proof. Since $p \geq 2$, the denominator is clearly positive if and only if $n > r(X)$ and $w_0 \neq \pm 1$. Hölder's inequality gives $w_0^2 = (w'\mathbf{1})^2 \leq w'w \cdot \mathbf{1}'\mathbf{1} = p$, so that the first term of the numerator is also non-negative.

If both $\theta_1, \theta_2 \geq 0$, then also the second term of the numerator is clearly non-negative for any $p \geq 2$, and $r_1(w) = r_2(w) \geq 0$.

If $\theta_1 \geq 0$ and $\theta_2 < 0$, then (3) with $w'w = 1$ implies $\theta_1 \geq |\theta_2|$. Thus we get

$$\begin{aligned} & p\left(p - 2 + w_0^2\right)\theta_1^2 - 2(p-1)w_0^2\theta_1\left|\theta_2\right| \geq \\ & \quad \geq \left[p\left(p - 2 + w_0^2\right) - 2(p-1)w_0^2\right]\theta_1^2 = (p-2)(p - w_0^2)\theta_1^2 \geq 0, \quad p \geq 2. \end{aligned}$$

It has to be noted that $\lim_{w_0^2\to 1} r_1(w) = +\infty$, so that this possible complication does not threaten the non-negativity. $\square$

We can conclude that also in this case YW-estimators are always better than Z-estimators.

3.2.1. *A special case.* Similarly to the previous subsection, let us consider a special case given by

$$w = \frac{1}{\sqrt{kb^2 + p - k}} (\underbrace{b, \dots, b}_{k \text{ times}}, \underbrace{1, \dots, 1}_{p-k \text{ times}})', \ 1 \le k < p.$$

In this case the variance matrix Σ contains two types of diagonal elements, variances $\theta_1 + \theta_2 b^2$ and $\theta_1 + \theta_2$, and three types of off-diagonal elements, covariances $\theta_2 b^2$, $\theta_2 b$, and θ_2. We can call such structure block-wise uniform. Note that b can be negative.

As stated in Section 3, for $\theta_1, \theta_2 \ge 0$, matrix Σ is positive semi-definite for any value of b. In the case $\theta_1 \ge 0$ and $\theta_2 < 0$, the condition for semi-definiteness of matrix Σ is $\theta_1 \ge |\theta_2|$. It is easy to verify that, for $1 \le k \le p$ and $b \in \mathbb{R}$,

$$\frac{|b|}{\sqrt{kb^2 + p - k}} < \lim_{b \to \pm\infty} \frac{|b|}{\sqrt{kb^2 + p - k}} = \frac{1}{\sqrt{k}}.$$

Then

$$w_0(b) = \mathbf{1}' w = \frac{kb + p - k}{\sqrt{kb^2 + p - k}}$$

assumes values from $(-\sqrt{k}; +\sqrt{p}\rangle$, and it holds $w_0^2(b) = 1$ in one or two points. This causes that Z-estimators are not defined at some points, having zero in the denominator. It is another reason why we have to recommend using YW-estimators.

We show a typical behaviour of function $r_1(b)$ $(= r_2(b))$ for some specific values of parameters in the following figures.

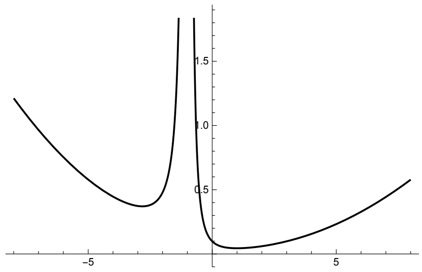

Figure 3: function $r_1(b)$ with one discontinuity.

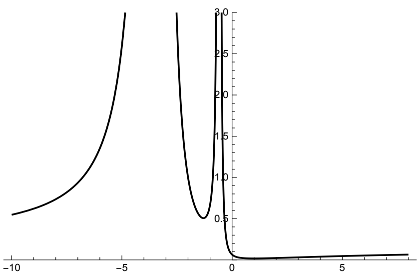

Figure 4: function $r_1(b)$ with two discontinuities.

We see that the difference can grow to infinity with increasing value of $|b|$, but also can asymptotically remain low. More typical is the second case, when the difference is asymptotically constant.

4. Conclusions

We have compared ML-, Z-, and YW-estimators in the growth curve model with GUC structure. We have showed that MLEs and YW-estimators are similar, but use of $\widehat{\theta}_2^{ML}$ is dangerous in small samples, since both mean and MSE of it depend not only on θ_2 but also on θ_1.

We have proved that in the model with positive definite diagonal matrix G and $w = \mathbf{1}$ or $G = I$ and arbitrary w, YW-estimators of variance parameters are always better than Z-estimators. YW-estimators typically have also much smaller variance than Z-estimators.

Moreover, Z-estimators suffer of problem of not being defined in some situations. As a result, we have to recommend using YW-estimators as a good small-sample alternative to MLEs. However, when matrix G is singular, the YW-estimators are not defined and closed form of MLEs is not known. Z-estimators, if they exist, can be considered as a reasonable alternative in such a case.

It has to be noted that our results can also be applied to standard multivariate regression model (i.e., regression model with multivariate responses) with the GUCS, since the Z matrix is not used for the variance components estimation.

Acknowledgements

The authors thank to an anonymous referee for his/her criticism which led to substantial improvement of the quality of the paper. This work was supported by the grants VEGA MŠ SR 1/0344/14 and VEGA MŠ SR 1/0073/15.

References

[1] P. S. Bullen, *Handbook of Means and Their Inequalities*, Kluwer Academic Publishers, Dordrecht, 2003.

[2] G. A. Ghazal and H. Neudecker, *On second-order and fourth-order moments of jointly distributed random matrices: a survey*, Linear Algebra Appl. **321** (2000), 61–93.

[3] C. G. Khatri, *Testing some covariance structures under a growth curve model*, J. Multivariate Anal. **3** (1973), 102–116.

[4] D. Klein and I. Žežula, *Orthogonal decompositions in growth curve models*, Acta Comment. Univ. Tartu. Math. **14** (2010), 35–44.

[5] J. C. Lee, *Prediction and estimation of growth curves with special covariance structures*, J. Amer. Statist. Assoc. **83** (1988), No. 402, 432–440.

[6] R. F. Potthoff and S. N. Roy, *A generalized multivariate analysis of variance model useful especially for growth curve problems*, Biometrika **51** (1964), 313–326.

[7] R. Rusnačko, *Porovnanie odhadov variančných parametrov v špeciálnom modeli rastových kriviek*, Forum Statisticum Slovacum **5** (2010), 204–209.

[8] R.-D. Ye and S.-G. Wang, *Estimating parameters in extended growth curve model with special covariance structures*, J. Statist. Plann. Inference **139** (2009), 2746–2756.

[9] I. Žežula, *Covariance components estimation in the growth curve model*, Statistics **24** (1993), 321–330.

[10] I. Žežula, *Special variance structure in the growth curve model*, J. Multivariate Anal. **97** (2006), 606–618.

Institute of Mathematics, P.J. Šafárik University, Jesenná 5, 041 54 Košice, Slovak Republic

E-mail address: rasto.rusnacko@gmail.com

E-mail address: ivan.zezula@upjs.sk