Tatra Mt. Math. Publ. **26** (2003), 247–254

TATRA MOUNTAINS
Mathematical Publications

ABOUT THE INFLUENCE OF INTERBLOCK VARIATION ON GROWTH CURVE ESTIMATORS OBTAINED BY POTTHOFF–ROY'S METHOD

Mirosława Wesołowska-Janczarek

ABSTRACT. The objective of this paper is an attempt at a theoretical solution of the problem of the influence of block structure in the experimental material in Potthoff–Roy's method of estimation of growth curves. Due to the presence of matrix T in the model, the estimators of the coefficients in the fitted polynomials, calculated with and without block effects in the model, are, generally, different. The scholarly considerations have not yet yielded a unique solution. It has only been established that the differences between estimators of coefficients depend on respective covariance matrices and their difference G. The condition of block effects insignificance in growth curve model is shown, too.

1. Introduction

This problem arose when Potthof–Roy's method was used to describe the dynamics of raspberries fruitbearing. The yielding of 16 cultivars was studied in an experiment carried out in 4 complete blocks. Fruit was picked over a period of about 30 days. The estimators of growth curves were calculated. If there are block effects in the model, interblock variation is eliminated from the error, and then the estimators of growth curves for treatment and for blocks are obtained too and polynomial degree must be the same for all groups. But block curves are not interesting for the experimenter. In these circumstances a problem arises: is it possible to ignore block effects in the model, and what are consequences of this omission? The objective of the consideration of this problem is to obtain negative value estimators for the raspberry yield in all the studied time points, if a model with block effects was taken. An explanation of this problem by the application of determination coefficients, adapted for growth curves and mean determination coefficients was undertaken by Wesołowska-Janczarek

2000 Mathematics Subject Classification: Primary 62J12; Secondary 62H12.
Keywords: growth curves, estimators of coefficients, completely randomized design, randomized block design, influence of block effects.

and Fus [7]. The theoretical approach to this problem is presented here. It is shown, that the estimators of growth curves for treatments obtained in both cases, without and with block effects in the model, differ only in the covariance matrix that is used. In the latter case the covariance matrix is diminished by matrix G, whose form is complicated because of matrix T included in the model. Therefore, the graph of estimated curve for each of the cultivars with blocks is moved in the direction of negative values. It is still an open question, why the estimated values of yields are negative throughout the studied period. This may by caused be a great changeability of yielding of some cultivars.

2. The problem

In the growth curve estimation by Potthoff–Roy's [3] method for the model

$$\underset{np}{Y} = \underset{na}{A}\,\underset{aq}{B}\,\underset{qp}{T} + \underset{np}{E} \tag{1}$$

under the assumptions: $\varepsilon(Y) = ABT$ and $\Sigma_Y = \underset{pp}{\Sigma} \otimes I_n$, the structure of experimental material is taken into consideration through the choice of a suitable matrix A of the experimental design. In consequence, the curves for treatments and blocks are assessed simultaneously, while only the former are of interest to the experimenter. Could we then ignore block effects in the model? Baksalary, Caliński and Kala [1] say yes, as long as the block effects are insignificant. This question is especially crucial in the particular case of assessing the dynamics of the fruitbearing for different raspberry cultivars. A real-life experiment conducted in the complete block design, for some cultivars, has yielded functions, whose values were negative, sometimes throughout the considerable time span.

A similar estimation using the same data set, but without the blocks in the model, has yielded polynomials which were translated in the direction of the positive values. However, they were not generally parallel to those obtained with the block effects. The examples of the curves obtained for 2 out of the 16 tested cultivars with (Bpr) and without block (pr) designs are presented on Fig. 1 and Fig. 2.

The details about the experiment and the analysis of the fruitbearing of 16 cultivars of raspberries are contained in Wesołowska-Janczarek and others [8]. At this point a problem arises: what are the consquences of ignoring block effects in the estimation of growth for the treatment in Potthoff–Roy's method. The preliminary study of this problem has been carried out using as the measure of fit, the determination coefficients R_j^2 adapted to the growth curve problem and the mean determination coefficient $\bar{R}^2$. Additional information on the coefficients can be found in Wesołowska-Janczarek [6].

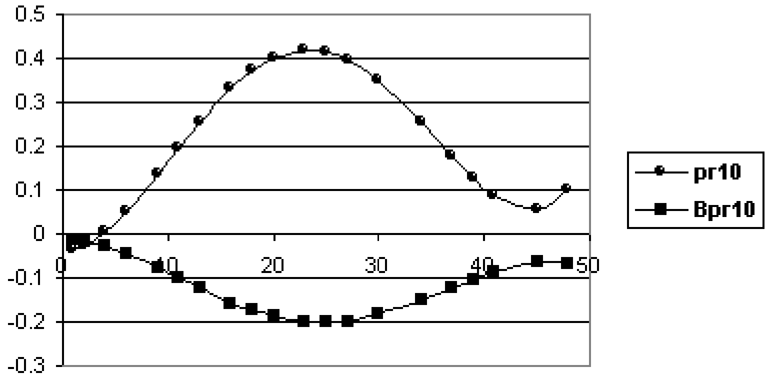

Figure 1.

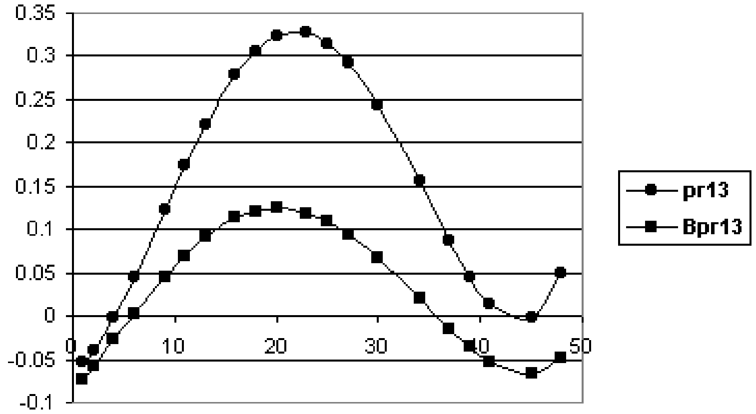

Figure 2.

Using experimental data from the years of 1987 and 1989 where the blocks were insignificant and significant respectively, it was verified in analysis of variance, that the following statements are true:

1. The goodness of fit of assessed curves for all cultivars in both cases when the block effects were present in the model and when they were omitted differ less in the case of the insignificant block effects than in the case when they were significant.
2. In both cases, when blocks were significant and insignificant, the values of R_j^2 $(j = 1, \ldots, 16)$ were greater in general if the block effects were omitted.
3. In the following four cases
   a) significant block, block effects present in the model,
   b) significant block, block effects omitted,

c) nonsignificant block, block effects present,

d) nonsignificant block, block effects omitted,

a larger variation of the calculated determination coefficients R_j^2 ($j = 1, \dots, 16$) was observed in the items a and d.

4. From the calculated values of mean determination coefficients $\bar{R}^2$ it follows, that goodness of fit of the assessed functions for the cultivars obtained by Potthoff–Roy's method is almost the same if the blocks are insignificant in both cases, when block effects are appeared and omitted, in the model. However if the blocks are significant, $\bar{R}^2$ is greater when the block effects were omitted.

These conclusions were given in Wesołowska-Janczarek and Fus [7]. In this paper the theoretical solution of this problem was undertaken.

3. Growth curve estimators for treatments with and without block effects

The known maximum likelihood estimators of growth curve coefficients take the form (see Potthoff, Roy [3] or Kshirsagar, Smith [2]):

$$\hat{B} = (A'A)^{-}A'Y\hat{\Sigma}^{-1}T'(T\hat{\Sigma}^{-1}T')^{-1}, \tag{2}$$

with

$$\hat{\Sigma} = \frac{1}{n}Y'\left[I_n - A(A'A)^{-}A'\right]Y. \tag{3}$$

For the same data in matrix Y collected from each of n units in p time point, the estimators of B are considered when blocks were in the model and when they were omitted. In the former case, matrix B is divided in two subblocks $\underset{v+r,q}{B} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$, where $\underset{vq}{B_1}$ consists of growth curve coefficients for treatments and $\underset{rq}{B_2}$—the same coefficients for blocks, v is a number of treatments, r—number of blocks, $q-1$ is the polynomial degree and q is a number of coefficients in each of the curves. Likewise, matrix $\underset{n,v+r}{A} = \left[\underset{nv}{A_1} \vdots \underset{nr}{A_2}\right]$, where $A_1 = I_v \otimes \mathbf{1}_r$, $A_2 = \mathbf{1}_v \otimes I_r$ and I_a is an $a \times a$ unit matrix, $\mathbf{1}_b$ is a vector of b ones, n is a number of the studied units and $n = v \cdot r$. In this case model (1) can be written as

$$Y = (I_v \otimes \mathbf{1}_r)B_1T + (\mathbf{1}_v \otimes I_r)B_2T + E_1, \tag{4}$$

however, if blocks in model are omitted, matrix $\underset{nv}{A} = [I_v \otimes \mathbf{1}_r]$ and then model (1) is as follows:

$$Y = (I_v \otimes \mathbf{1}_r)BT + E_2. \tag{5}$$

Matrix T takes the form

$$T = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_p \\ \cdots & \cdots & \cdots & \cdots \\ t_1^{q-1} & t_2^{q-1} & \cdots & t_p^{q-1} \end{bmatrix},$$

where $t_1, t_2, \ldots, t_p$ are time points in which observations were taken. In both models, E_1 and E_2 are matrices of random errors and $\otimes$ denotes the Kronecker product of matrices. The estimator of covariance matrix in (4) is

$$\hat{\Sigma}_{BL} = \frac{1}{n} Y' \left[I_n - \frac{1}{r}\left(I_v \otimes \underset{rr}{E}\right) - \frac{1}{v}\left(\underset{vv}{E} \otimes I_r\right) + \frac{1}{n} \underset{nn}{E} \right] Y, \tag{6}$$

where $\underset{nn}{E}$ is an $n \times n$ matrix of ones and in (5):

$$\hat{\Sigma}_{CR} = \frac{1}{n} Y' \left[I_n - \frac{1}{r}\left(I_v \otimes \underset{rr}{E}\right) \right] Y. \tag{7}$$

Both covariance matrix estimators (6) and (7) are not unbiased.

From (6) and (7) it follows that:

$$\hat{\Sigma}_{CR} = \hat{\Sigma}_{BL} + G, \tag{8}$$

where

$$G = \frac{1}{n} Y' \left[\frac{1}{v}\left(\underset{vv}{E} \otimes I_r\right) - \frac{1}{n} \underset{nn}{E} \right] Y. \tag{9}$$

If matrix B_1 in the model (4) is denoted by $B_{1(BL)}$, B by B_{BL} and B_1 in the model (5) by B_{CR}, then $B_{BL} = \begin{bmatrix} B_{1(BL)} \\ B_2 \end{bmatrix}$ and the estimators of coefficients matrix are under the model (4):

$$\hat{B}_{BL} = \begin{bmatrix} \hat{B}_{1(BL)} \\ \hat{B}_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{r}(I_v \otimes \mathbf{1}_r') Y \hat{\Sigma}_{BL}^{-1} T' \left(T \hat{\Sigma}_{BL}^{-1} T'\right)^{-1} \\ \frac{1}{v}\left[(\mathbf{1}_v' \otimes I_r) - \frac{1}{r} \underset{rn}{E}\right] Y \hat{\Sigma}_{BL}^{-1} T' \left(T \hat{\Sigma}_{BL}^{-1} T'\right)^{-1} \end{bmatrix} \tag{10}$$

and under the model (5)

$$\hat{B}_{CR} = \frac{1}{r}(I_v \otimes \mathbf{1}_r') Y \hat{\Sigma}_{CR}^{-1} T' \left(T \hat{\Sigma}_{CR}^{-1} T'\right)^{-1}. \tag{11}$$

4. Comparison of estimators

From (10) and (11) it follows, that the estimators of growth curve coefficients for the treatments differ only in the applied estimators of covariance matrices.

Then difference between the estimators of parameters matrices equals:

$$\hat{B}_{CR} - \hat{B}_{1(BL)} = \frac{1}{r}(I_v \otimes \mathbf{1}'_r)Y\left[\hat{\Sigma}_{CR}^{-1}T'(T\hat{\Sigma}_{CR}^{-1}T')^{-1} - \hat{\Sigma}_{BL}^{-1}T'(T\hat{\Sigma}_{BL}^{-1}T')^{-1}\right]$$
$$= \frac{1}{r}(I_v \otimes \mathbf{1}'_r)YZ\,. \tag{12}$$

Given (8), matrix Z can be written as:

$$Z = (\hat{\Sigma}_{BL} + G)^{-1}T'\left[T(\hat{\Sigma}_{BL} + G)^{-1}T'\right]^{-1} - \hat{\Sigma}_{BL}^{-1}T'(T\hat{\Sigma}_{BL}^{-1}T')^{-1} \tag{13}$$

or after taking the relation $(U+V)^{-1} = U^{-1} - U^{-1}(U^{-1}+V^{-1})^{-1}U^{-1}$,

$$\begin{aligned} Z = & -\hat{\Sigma}_{BL}^{-1}(\hat{\Sigma}_{BL}^{-1} + G^{-1})^{-1}\hat{\Sigma}_{BL}^{-1}T'(T\hat{\Sigma}_{BL}^{-1}T')^{-1} \\ & -\left[\hat{\Sigma}_{BL}^{-1}T' - \hat{\Sigma}_{BL}^{-1}(\hat{\Sigma}_{BL}^{-1} + G^{-1})^{-1}\hat{\Sigma}_{BL}^{-1}T'\right](T'\hat{\Sigma}_{BL}^{-1}T')^{-1}\times \\ & \times\left\{(T\hat{\Sigma}_{BL}^{-1}T')^{-1} - [T\hat{\Sigma}_{BL}^{-1}(\hat{\Sigma}_{BL}^{-1} + G^{-1})^{-1}\hat{\Sigma}_{BL}^{-1}T']^{-1}\right\}^{-1}(T\hat{\Sigma}_{BL}^{-1}T')^{-1}\,. \end{aligned}$$

From (12) and (13) it is seen that the difference between coefficients values will be smaller if matrix $\hat{\Sigma}_{CR}$ is closer to $\hat{\Sigma}_{BL}$ or if the generalized quadratic form G is closer to the zero matrix. Moreover, it is clear, that the estimators of polynomial coefficients for the treatments in both models (4) and (5) are the same if the measurements in all the time points $t_1, t_2, \ldots, t_p$ are uncorrelated and the variances of the observations in all the time points are equal. These conditions can be written as:

$$\hat{\Sigma}_{CR} = \hat{\sigma}_{CR}^2 I_p\,, \qquad \hat{\Sigma}_{BL} = \hat{\sigma}_{BL}^2 I_p\,, \qquad G = (\hat{\sigma}_{CR}^2 - \hat{\sigma}_{BL}^2)I_p\,.$$

Then the matrix $Z = 0$.

The estimators $\hat{B}_{1(BL)}$ and $\hat{B}_{CR}$ are identical too, if matrix T is of the order $p \times p$.

5. Conditions of block insignificance

It is not simple to determine the size of the difference between the matrices $\hat{\Sigma}_{BL}$ and $\hat{\Sigma}_{CK}$. The confirmation of the insignificance of block effects could provide a useful criterion, but the analysis of variance for the sum of observations in p time points, that is for $\underset{n1}{y} = Y\mathbf{1}_p$, is sometimes not possible, as that sum of observation can sometimes make no sense. Nonetheless, in our experiment each of the sums will represent a combined yield of every unit. In the combined analysis of variance for the block experiment under $y = Y\mathbf{1}_p$ the sums of square are the following

$$SS_T = y'C_1y\,, \quad SS_{BL} = y'C_2y \quad \text{and} \quad SS_E = y'C_3y\,,$$

where

$$C_1 = \frac{1}{r}(I_r \otimes \underset{vv}{E}) - \frac{1}{n}\underset{nn}{E},$$
$$C_2 = \frac{1}{v}(\underset{vv}{E} \otimes I_v) - \frac{1}{n}\underset{nn}{E},$$
$$C_3 = I_n - \frac{1}{r}(I_r \otimes \underset{vv}{E}) - \frac{1}{v}(\underset{vv}{E} \otimes I_v) + \frac{1}{n}\underset{nn}{E}.$$

The blocks are insignificant if the calculated value $F^o_{BL} < F_{\alpha,r-1,(v-1)(r-1)}$ where

$$F^o_{BL} = \frac{y'C_2y}{r-1} \bigg/ \frac{y'C_3y}{(v-1)(r-1)} = \frac{(v-1)y'C_2y}{y'C_3y}. \tag{14}$$

As $\hat{\Sigma}_{BL} = \frac{1}{n}Y'C_3Y$, $G = \frac{1}{n}Y'C_2Y$ and $\hat{\Sigma}_{CR} = \frac{1}{n}Y'(C_3 - C_2)Y$, then

$$F^o_{BL} = \frac{(v-1)\mathbf{1}'_pG\mathbf{1}_p}{\mathbf{1}'_p\hat{\Sigma}_{BL}\mathbf{1}_p} \tag{15}$$

or given (8)

$$F^o_{BL} = \frac{(v-1)\mathbf{1}'_p(\hat{\Sigma}_{CR} - \hat{\Sigma}_{BL})\mathbf{1}_p}{\mathbf{1}'_p\hat{\Sigma}_{BL}\mathbf{1}}. \tag{16}$$

If the number of treatments v and the number of blocks r are fixed, the condition of the block insignificance is

$$\mathbf{1}'_pG\mathbf{1}_p < \frac{F_\alpha}{v-1}\mathbf{1}'_p\hat{\Sigma}_{BL}\mathbf{1}_p \qquad \text{or} \qquad \mathbf{1}'_p\hat{\Sigma}_{CR}\mathbf{1}_p < \frac{v-1+F_\alpha}{v-1}\mathbf{1}'_p\hat{\Sigma}_{BL}\mathbf{1}_p. \tag{17}$$

As the present results are not complete, the problem calls for a further study of the difference between $\hat{B}_{CR}$ and $\hat{B}_{1(BL)}$ and their relation with $\hat{\Sigma}_{BL}$ and $\hat{\Sigma}_{CR}$.

6. Some alternative methods

Another approach to the problem of block effects estimation was shown by von Rosen [5] who has modified Pothoff–Roy's method. He assumes that for a certain group of parameters the polynomials are not taken into consideration. Therefore he considers a model with covariates, which takes the following form:

$$Y = A_1B_1^*T + A_2B_2^* + E. \tag{18}$$

In such a model the suitable estimators of the matrices of parameters B_1^* and B_2^*, where B_1^* contains the coefficients in the polynomials for the treatments and B_2^*-block effects in p time points, are as follows:

$$\underset{vq}{\hat{B}_1^*} = (C'C)^-C'QS^{-1}T'(TS^{-1}T')^{-1} \tag{19}$$

where

$$\underset{nv}{C} = \left[I_n - A_2(A_2'A_2)^{-1}A_2'\right]A_1\,,$$
$$\underset{np}{Q} = \left[I_n - A_2(A_2'A_2)^{-1}A_2'\right]Y,$$
$$\underset{pp}{S} = Y'\left[I_n - A(A'A)^{-}A'\right]Y,$$

A, A_1 and A_2 are the same as in Section 3 and

$$\underset{rp}{\widehat{B}_2^*} = (A_2'A_2)^{-1}A_2'(Y - A_1\widehat{B}_1^*T)\,. \tag{20}$$

The curves for treatment assessed by this method are not better fitted than those considered previous.

REFERENCES

[1] BAKSALARY, J.—CALIŃSKI, T.—KALA, R.: *Growth curve estimation in complete block design*, Biuletyn Oceny Odmian, T. X, Z.1 **15** (1983), 105–117. (In Polish)

[2] KSHIRSAGAR, A. M.—SMITH, B. W.: *Growth Curves*, Marcel Dekker, Inc., New York, 1995.

[3] POTTHOFF, R. F.—ROY, S. N.: *A generalized multivariate analysis of variance model useful especially for growth curve problems*, Biometrika **51** (1964), 313–326.

[4] SRÍVASTAVA, M. S.: *Methods of Multivariate Statistics*, John Wiley & Sons, Inc., New York, 2002.

[5] VON ROSEN, D.: *Multivariate linear normal models with special references to the growth curve model*, Dissertation, University of Stockholm, Stockholm, Sweden, 1985.

[6] WESOŁOWSKA-JANCZAREK, M.: *On the use of determination coefficient to describe goodness of fit of assessed growth curves*, Biometrical Letters **37** (2000), 13–19.

[7] WESOŁOWSKA-JANCZAREK, M.—FUS, L.: *Influence of block effects on growth curve fitting obtained by three different methods of estimation. Part I and II.*, Colloq. Biometryczne (Lublin), **30** (2000), 101–128. (In Polish)

[8] WESOŁOWSKA-JANCZAREK, M.—FUS, L.—OSYPIUK, Z.: *The application of the growth curve method with concomitant variables in the studies of the raspberry fruitbearing*, Colloq. Biometryczne (Lublin), **28** (1997), 269–281. (In Polish)

Received March 27, 2002

Institute of Applied Mathematics
Agricultural University of Lublin
Akademicka 13
PL–20-950 Lublin
POLAND

E-mail: janczar@ursus.ar.lublin.pl