

Optimal design for linear regression models in the presence of heteroscedasticity caused by random coefficients^a

Ulrike Graßhoff^b, Anna Doebler^c, Heinz Holling^c, Rainer Schwabe^{b,d}

Abstract

Random coefficients may result in a heteroscedasticity of observations. For particular situations, where only one observation is available per individual, we derive optimal designs based on the geometry of the design locus.

1 Introduction

In the social sciences and biosciences random effects play a growing role, whenever different individuals are involved in a study. While in fixed effects models typically only additive errors are taken into account, the situation has to change, when the coefficients of the regression function may vary randomly across those individuals. Our approach here is motivated by a validation problem in intelligence testing, when only one observation is made available per individual. Freund and Holling (2009) analyzed the impact of reasoning and creativity on GPA based on data from the standardization sample of the Berlin Structure of Intelligence Test for Youth: Assessment of Talent and Giftedness (BIS-HB). Since the effect of both variables on school performance may vary between different classrooms, a random effects model incorporating the explanatory variables on two levels (level 1: students and level 2: classrooms) was specified. Thus, the results allow for a more detailed interpretation of the role of different variables in the context of predicting scholastic achievement.

Such “sparse” observations have also been considered by Patan and Bogacka (2007) in a population pharmacokinetics setup. In those applications the response is usually non-linear. However, to solve the corresponding design problem it is advisable to have some knowledge of the influence of random coefficients already in the linear model setup. This is a further motivation of the present investigation.

In the linear setup the random coefficient model with single observations can be reformulated as a heteroscedastic fixed effects model with a specific structure of

^aWork supported by grant HO 1286/6-1 of the Deutsche Forschungsgemeinschaft.

^bOtto von Guericke University, Institute for Mathematical Stochastics, PF 4120, D-39 016 Magdeburg, Germany

^cWestfälische Wilhelms-Universität, Psychologisches Institut IV, Fliednerstr. 21, D-48 149 Münster, Germany

^dcorresponding author, E-mail: rainer.schwabe@ovgu.de

the variance function. If all coefficients are substantially random it can be easily verified that the corresponding standardized design locus is included in the surface of an ellipsoid generated by the covariance matrix of the random coefficients. In the case of high variability this ellipsoid may coincide with the smallest circumscribing one. Then multiple solutions to the design problem are possible as indicated in the discussion by Silvey (1972) and Sibson (1972). This phenomenon will be demonstrated by some simple but illustrative examples.

2 Model description

We consider a random coefficient regression model $Y_i(\mathbf{x}_i) = \mathbf{f}(\mathbf{x}_i)^\top \mathbf{b}_i$. The dependence of the observations Y_i on the experimental settings \mathbf{x}_i is given by the p -dimensional vector of known regression functions \mathbf{f} and the independent vectors \mathbf{b}_i of random coefficients, which come from a normal distribution, $\mathbf{b}_i \sim N_p(\boldsymbol{\beta}, \mathbf{D})$, with mean vector $\boldsymbol{\beta}$ and dispersion matrix \mathbf{D} . The design problem is to choose the experimental settings \mathbf{x}_i from the design region \mathcal{X} for estimating the location parameters $\boldsymbol{\beta}$, while the dispersion matrix \mathbf{D} is assumed to be known.

In this note we assume that all observations Y_i are independent, i. e. only one observation is made for each realization \mathbf{b}_i of the random coefficients. Moreover, we assume here that an intercept is included in the model ($f_1(\mathbf{x}) \equiv 1$) such that additive observational errors may be subsumed into the random intercept.

This model can be rewritten as a heteroscedastic linear fixed effects model

$$Y_i(\mathbf{x}_i) = \mathbf{f}(\mathbf{x}_i)^\top \boldsymbol{\beta} + \varepsilon_i, \quad (1)$$

where $\varepsilon_i = \mathbf{f}(\mathbf{x}_i)^\top (\mathbf{b}_i - \boldsymbol{\beta}) \sim N(0, \sigma^2(\mathbf{x}))$ and the variance function is defined by $\sigma^2(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \mathbf{D} \mathbf{f}(\mathbf{x})$. Within this heteroscedastic linear model for each single setting $\mathbf{x} \in \mathcal{X}$ the information equals $\mathbf{M}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) \mathbf{f}(\mathbf{x})^\top / \sigma^2(\mathbf{x})$. Then for a design ξ , the standardized information matrix is defined by $\mathbf{M}(\xi) = \sum_{j=1}^m \xi(\mathbf{x}_j) \mathbf{M}(\mathbf{x}_j)$, where $\xi(\mathbf{x}_j)$ is the proportion of observations at setting \mathbf{x}_j , $\sum_{j=1}^m \xi(\mathbf{x}_j) = 1$. Note that the covariance matrix for the weighted least squares estimator $\hat{\boldsymbol{\beta}}$, which is the best unbiased estimator for $\boldsymbol{\beta}$ and coincides with the maximum likelihood estimator in the present setting, equals the inverse of the information matrix. Hence, maximizing the information matrix is equivalent to minimizing the covariance matrix of $\hat{\boldsymbol{\beta}}$.

To compare different designs we consider the most popular criterion, the D -criterion, with respect to which a design ξ^* is D -optimal, if it maximizes the determinant of the information matrix. This is equivalent to the minimization of the volume of a confidence ellipsoid for $\boldsymbol{\beta}$. In the setting of approximate designs, for which the proportions $\xi(\mathbf{x})$ are not necessarily multiples of $1/n$, where n denotes the sample size, the D -optimality of a design ξ^* can be established by the well-known Kiefer-Wolfowitz equivalence theorem (see Fedorov, 1972, for a suitable version): A design ξ^* is D -optimal, if $\mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi^*)^{-1} \mathbf{f}(\mathbf{x}) / \sigma^2(\mathbf{x}) \leq p$, uniformly in $\mathbf{x} \in \mathcal{X}$. When we substitute $\sigma^2(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \mathbf{D} \mathbf{f}(\mathbf{x})$ into this relation and rearrange terms, D -optimality is achieved, if

$$\delta(\mathbf{x}; \xi^*) \geq 0 \quad (2)$$

for all $\mathbf{x} \in \mathcal{X}$, where $\delta(\mathbf{x}; \xi) = \mathbf{f}(\mathbf{x})^\top (p\mathbf{D} - \mathbf{M}(\xi)^{-1})\mathbf{f}(\mathbf{x})$ is the suitably transformed sensitivity function. Moreover, equality is attained in (2) for design points, where $\xi^*(\mathbf{x}) > 0$.

3 Linear regression on the standard interval

In the situation of linear regression we have observations $Y_i = b_{i0} + b_{i1}x$. The vector of regression functions \mathbf{f} is given by $\mathbf{f}(x) = (1, x)^\top$, and we assume that the setting x may be chosen from the symmetric standard interval $\mathcal{X} = [-1, 1]$.

For $x_1, x_2 \in [-1, 1]$ we consider the uniform two-point design ξ_{x_1, x_2} on x_1 and x_2 , which is defined by $\xi_{x_1, x_2}(x_1) = \xi_{x_1, x_2}(x_2) = 1/2$, with information matrix

$$\mathbf{M}(\xi_{x_1, x_2}) = \frac{1}{2\sigma^2(x_1)\sigma^2(x_2)} \begin{pmatrix} \sigma^2(x_1) + \sigma^2(x_2) & x_1\sigma^2(x_2) + x_2\sigma^2(x_1) \\ x_1\sigma^2(x_2) + x_2\sigma^2(x_1) & x_1^2\sigma^2(x_2) + x_2^2\sigma^2(x_1) \end{pmatrix} \quad (3)$$

and corresponding determinant $\det(\mathbf{M}(\xi_{x_1, x_2})) = (x_1 - x_2)^2 / (4\sigma^2(x_1)\sigma^2(x_2))$.

First we consider the case that the random intercepts b_{i0} and the random slopes b_{i1} are uncorrelated, where an easy geometric interpretation can be achieved. The associated variances will be denoted by $d_0 = \text{Var}(b_{i0})$ and $d_1 = \text{Var}(b_{i1})$, respectively, and the covariance matrix $\mathbf{D} = \text{diag}(d_0, d_1)$ of the random coefficients is diagonal. Maximizing the determinant of $\mathbf{M}(\xi_{x_1, x_2})$ with respect to x_1 and $x_2 \in [-1, 1]$ leads to the following solutions.

If $d_0 \geq d_1$, the endpoints $x_1^* = 1$ and $x_2^* = -1$ are optimal. In this case the information matrix results in $\mathbf{M}(\xi_{1, -1}) = (d_0 + d_1)^{-1}\mathbf{I}_2$, where \mathbf{I}_2 denotes the 2×2 identity matrix. Since $2\mathbf{D} - \mathbf{M}(\xi_{1, -1})^{-1} = \text{diag}(d_0 - d_1, d_1 - d_0)$ we obtain $\delta(x; \xi_{1, -1}) = (d_0 - d_1)(1 - x^2)$, and the inequality (2) is satisfied, which proves the D -optimality of the design $\xi_{1, -1}$.

In the case $d_0 < d_1$ the solutions of the optimization problem are characterized by the hyperbolic equation $d_0 + d_1 x_1 x_2 = 0$, $x_1, x_2 \in [-1, 1]$. For each such pair x_1^*, x_2^* the information matrix reduces to $\mathbf{M}(\xi_{x_1^*, x_2^*}) = \frac{1}{2}\mathbf{D}^{-1}$ and the inequality (2) is obviously satisfied with $\delta(x; \xi_{x_1^*, x_2^*}) = 0$ for all x , which proves the D -optimality of the design $\xi_{x_1^*, x_2^*}$. Note that for $d_0 < d_1$ the optimal choice is not unique. In particular, we may choose a symmetric solution $x_1^* = \sqrt{d_0/d_1}$ and $x_2^* = -x_1^*$.

In Figure 1 we exhibit the standardized design locus together with the smallest circumscribing ellipse related to $\mathbf{M}(\xi^*)^{-1}$. The standardized design locus is itself an arc of an ellipse generated by $p\mathbf{D}$. For $d_0 \geq d_1$ this arc touches the circumscribing ellipse only at its endpoints, while for $d_0 < d_1$ the arc is part of that ellipse, which results in multiple solutions in accordance with the discussion by Silvey (1972) and Sibson (1972).

Remark. The present design locus may be identified as a segment of the design locus corresponding to a rescaled trigonometric regression without intercept. For the latter model equally spaced design points would be optimal on the entire ellipse, if the number of design points is, at least, three. Due to the lack of an intercept in that model also equidistant design points on a semi-circle (modulo its length) will be optimal, where now the number of design points has to be, at least,

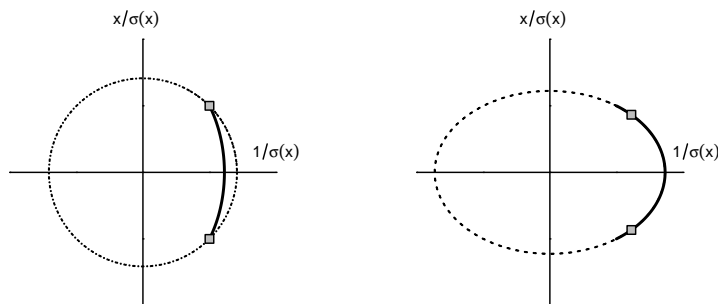


Figure 1: Standardized design locus (solid line), smallest circumscribing ellipse (dotted line) and optimal design points (\square) for $d_0 > d_1$ (left panel) and $d_0 < d_1$ (right panel)

two. This means, in particular, that in the original model, any two “equidistant” design points will be optimal, if the length of the segment is sufficiently large.

Next we consider the design problem for linear regression on the standard interval with an underlying general covariance matrix

$$\mathbf{D} = \begin{pmatrix} d_0 & d_{01} \\ d_{01} & d_1 \end{pmatrix},$$

where $d_{01} = \text{Cov}(b_{i0}, b_{i1})$ denotes the covariance of the two random coefficients. In this situation the geometric interpretation is slightly less persuasive. As before we maximize the determinant of the information matrix (3) for the uniform two-point designs ξ_{x_1, x_2} , $x_1, x_2 \in [-1, 1]$.

For $d_0 \geq d_1$ again the endpoints $x_1^* = 1$ and $x_2^* = -1$ are optimal. As in the uncorrelated case we obtain $2\mathbf{D} - \mathbf{M}(\xi_{1, -1})^{-1} = \text{diag}(d_0 - d_1, d_1 - d_0)$, which does not depend on d_{01} , and the D -optimality follows from (2).

In the case $d_0 < d_1$ the solutions for the optimization problem are now characterized by the hyperbolic equation $d_0 + d_{01}(x_1 + x_2) + d_1 x_1 x_2 = 0$. Some easy but tedious computations yield that for each such solution x_1^*, x_2^* the information matrix equals $\mathbf{M}(\xi_{x_1^*, x_2^*}) = \frac{1}{2}\mathbf{D}^{-1}$ also in this situation, which proves the D -optimality of the design $\xi_{x_1^*, x_2^*}$ in view of (2). In total we obtain the following result.

Theorem 1. *In the heteroscedastic model (1) of linear regression on the standard interval $\mathcal{X} = [-1, 1]$ with dispersion matrix \mathbf{D} the design $\xi_{x_1^*, x_2^*}$ is D -optimal, where $x_1^* = 1$ and $x_2^* = -1$ for $d_0 \geq d_1$, and x_1^*, x_2^* satisfy $d_0 + d_{01}(x_1 + x_2) + d_1 x_1 x_2 = 0$ for $d_0 < d_1$, respectively.*

The optimal settings are not unique for $d_0 < d_1$, but it turns out that the symmetric pair x^* and $-x^*$, where $x^* = \sqrt{d_0/d_1}$, constitutes a solution, whatever the magnitude of the covariance d_{01} is.

Corollary 1. *In the heteroscedastic model (1) of linear regression on the standard interval $\mathcal{X} = [-1, 1]$ with dispersion matrix \mathbf{D} the design $\xi_{x^*, -x^*}$ is D -optimal, where $x^* = 1$, if $d_0 \geq d_1$, and $x^* = \sqrt{d_0/d_1}$, if $d_0 < d_1$.*

4 Linear regression: General case

Within the setup of the previous section we pass now to the situation where the design region may be an arbitrary interval, $\mathcal{X} = [a, b]$, where $a < b$. The model equation and all notations are as in the previous section.

For obtaining D -optimal designs we use the standard linear transformation technique to map the standard interval $[-1, 1]$ onto $[a, b]$: Define $g : [-1, 1] \rightarrow [a, b]$ by $g(x) = \frac{a+b}{2} + \frac{b-a}{2}x$. This mapping induces a linear transformation of the regression function, $\mathbf{f}(g(x)) = \mathbf{A}\mathbf{f}(x)$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ (a+b)/2 & (b-a)/2 \end{pmatrix}.$$

The information under the mapping g then results in

$$\mathbf{M}(g(x)) = \frac{\mathbf{f}(g(x))\mathbf{f}(g(x))^\top}{\sigma^2(g(x))} = \frac{\mathbf{A}\mathbf{f}(x)\mathbf{f}(x)^\top\mathbf{A}^\top}{\mathbf{f}(x)^\top\mathbf{A}^\top\mathbf{D}\mathbf{A}\mathbf{f}(x)} = \mathbf{A}\tilde{\mathbf{M}}(x)\mathbf{A}^\top,$$

where $\tilde{\mathbf{M}}(x)$ denotes the corresponding information for the linear regression model on the standard interval with dispersion matrix $\tilde{\mathbf{D}} = \mathbf{A}^\top\mathbf{D}\mathbf{A}$. Since the transformation matrix \mathbf{A} does not depend on x the design problem on $[a, b]$ can be solved by applying the results of the previous section to the heteroscedastic model (1) with induced dispersion matrix

$$\tilde{\mathbf{D}} = \frac{1}{4} \begin{pmatrix} 4d_0 + 4(a+b)d_{01} + (a+b)^2d_1 & (b-a)(2d_{01} + (a+b)d_1) \\ (b-a)(2d_{01} + (a+b)d_1) & (b-a)^2d_1 \end{pmatrix}.$$

and then transforming the optimal design by the mapping g .

In order to apply Theorem 1 we have to compare the diagonal entries of the dispersion matrix $\tilde{\mathbf{D}}$, which leads to the following result.

Theorem 2. *In the heteroscedastic model (1) of linear regression on $[a, b]$ and dispersion matrix \mathbf{D} the design $\xi_{a,b}$ is D -optimal, if $d_0 + (a+b)d_{01} + abd_1 \geq 0$. Otherwise the design $\xi_{x_1^*, x_2^*}$ is D -optimal, where the design points x_1^* and x_2^* satisfy $4d_0 + 4(a+b)d_{01} + (a+b)^2d_1 + (b-a)(2d_{01} + (a+b)d_1)(x_1^* + x_2^*) + (b-a)^2d_1x_1^*x_2^* = 0$.*

Remark. In the latter case particular solutions x_1^* and x_2^* are given by

$$\frac{1}{2} \left(a + b \pm \sqrt{4(d_0 + (a+b)d_{01})/d_1 + (a+b)^2} \right).$$

Note that for diagonal \mathbf{D} the design $\xi_{a,b}$ concentrated on the endpoints a and b is D -optimal as long as $d_0 + abd_1 \geq 0$. This condition is automatically satisfied, if the endpoints a and b have the same sign, i. e. if the design region is completely contained in the positive (or the negative) half axis.

5 Multilinear regression

In this section we will extend the linear regression model of Section 3 to K factors which can be chosen from the design region of the symmetric standard hypercube

$\mathcal{X} = [-1, 1]^K$ in the special case of a diagonal dispersion matrix \mathbf{D} . In this situation we have observations $Y_i(\mathbf{x}) = b_{i0} + \sum_{k=1}^K b_{ik}x_k$ with $\mathbf{x} = (x_1, \dots, x_K)^\top \in \mathcal{X} = [-1, 1]^K$. The vector of regression functions is given by $\mathbf{f}(\mathbf{x}) = (1, x_1, \dots, x_K)^\top$.

The covariance matrix $\mathbf{D} = \text{diag}(d_0, d_1, \dots, d_K)$ of the random coefficients \mathbf{b}_i is assumed to be diagonal. Hence, the variance of each design point is equal to $\sigma^2(\mathbf{x}) = d_0 + \sum_{k=1}^K d_k x_k^2$. Without loss of generality we may also assume that the factors are arranged according to an ascending order in the magnitude of the variance components, i.e. $d_1 \leq \dots \leq d_K$. This can be achieved by a suitable relabeling of the indices of the factors.

As candidates for optimal designs we consider uniform full factorial 2^K - designs $\xi_{\mathbf{x}} = \xi_{x_1, \dots, x_K}$ on the points $= (\pm x_1, \dots, \pm x_K)$ generated by $\mathbf{x} = (x_1, \dots, x_K)$, which assign equal proportions $1/2^K$ to each of these 2^K design points.

In order to characterize D -optimal designs for this model we introduce the cumulative averages

$$c_m = \frac{1}{m+1} \sum_{k=0}^m d_k$$

of the variance components, $m = 0, \dots, K$, for which the following lemma holds.

Lemma 1. *For $d_0 \geq 0$ and $0 \leq d_1 \leq \dots \leq d_K < d_{K+1} = \infty$ there exists a unique index m such that*

$$d_m \leq c_m < d_{m+1}. \quad (4)$$

The proof of Lemma 1 is deferred to the appendix. We are now ready to specify optimal design in the present situation.

Theorem 3. *In the heteroscedastic model (1) of multilinear regression on the unit hypercube $[-1, 1]^K$ with diagonal dispersion matrix $\mathbf{D} = \text{diag}(d_0, d_1, \dots, d_K)$, where $d_1 \leq \dots \leq d_K$, the design $\xi^* = \xi_{\mathbf{x}^*} = \xi_{x_1^*, \dots, x_K^*}$ is D -optimal, if $x_k^* = 1$ for $k \leq m$ and $x_k^* = \sqrt{\frac{c_m}{d_k}}$ for $k > m$, respectively, and where m satisfies $d_m \leq c_m < d_{m+1}$.*

Proof. First we note that $\sigma^2(\mathbf{x}) = d_0 + \sum_{k=1}^K d_k x_k^2$ for any $\mathbf{x} = (x_1, \dots, x_K)$ and that for the corresponding full factorial design $\xi_{\mathbf{x}}$ the information matrix $\mathbf{M}(\xi_{\mathbf{x}}) = \sigma^2(\mathbf{x})^{-1} \text{diag}(1, x_1^2, \dots, x_K^2)$ is diagonal. Inserting the values x_1^*, \dots, x_K^* yields $\sigma^2(\mathbf{x}^*) = \sum_{k=0}^m d_k + (K-m)c_m = (K+1)c_m$ and consequently

$$\mathbf{M}(\xi^*)^{-1} = (K+1)c_m \text{diag}(1, 1, \dots, 1, d_{m+1}/c_m, \dots, d_K/c_m),$$

which implies

$$p\mathbf{D} - \mathbf{M}(\xi^*)^{-1} = (K+1) \text{diag}(d_0 - c_m, d_1 - c_m, \dots, d_m - c_m, 0, \dots, 0)$$

as $p = K+1$. Thus we obtain for the sensitivity function

$$\begin{aligned} \delta(\mathbf{x}; \xi^*) &= (K+1) \left(d_0 - c_m + \sum_{k=1}^m (d_k - c_m) x_k^2 \right) \\ &\geq (K+1) \left(d_0 - c_m + \sum_{k=1}^m (d_k - c_m) \right) = 0, \end{aligned}$$

since $d_k - c_m \leq 0$, for $k = 1, \dots, m$, and $x_k^2 \leq 1$, which proves the D -optimality of ξ^* in view of (2). \square

Note that in Theorem 3 the optimal settings are not unique, if $m < K$, and may be replaced by non-symmetric solutions. Moreover, for $K > 2$ the full factorials may be substituted by fractional factorials such that the number of required design points reduces to $O(K)$.

As an example we will derive the design points of the D -optimal design ξ^* in case of the bilinear regression, i.e. in the case $K = 2$.

Corollary 1. *In the heteroscedastic model (1) of bilinear regression on the unit square $[-1, 1]^2$ with diagonal dispersion matrix $\mathbf{D} = \text{diag}(d_0, d_1, d_2)$ the design $\xi_{x_1^*, x_2^*}$ is D -optimal, where*

	$x_1^* =$	$x_2^* =$	
$(1, 1)$	1	1	if $d_0 + d_1 \geq 2d_2$ and $d_0 + d_2 \geq 2d_1$,
$(1, x_2^*)$	1	$\sqrt{\frac{d_0 + d_1}{2d_2}}$	if $d_0 \geq d_1$ and $d_0 + d_1 < 2d_2$,
$(x_1^*, 1)$	$\sqrt{\frac{d_0 + d_2}{2d_1}}$	1	if $d_0 \geq d_2$ and $d_0 + d_2 < 2d_1$,
(x_1^*, x_2^*)	$\sqrt{\frac{d_0}{d_1}}$	$\sqrt{\frac{d_0}{d_2}}$	if $d_0 < d_1$ and $d_0 < d_2$.

Note that we have, here, refrained from the assumption $d_1 \leq d_2$, which leads to the distinction of the cases $(1, x_2^*)$ and $(x_1^*, 1)$.

For the four alternative situations in Theorem 3 the parameter regions for the variance ratios d_1/d_0 and d_2/d_0 are depicted in Figure 2.

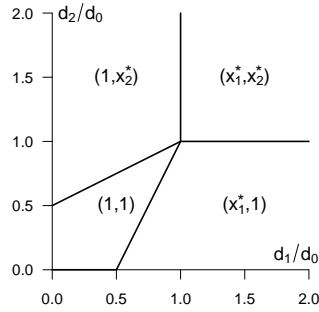


Figure 2: Parameter regions of the variance ratios d_1/d_0 and d_2/d_0 for optimal designs in bilinear regression

Whether these results can be extended to more general dispersion matrices involving correlations is not, yet, clear and requires further investigations.

6 Discussion

In the heteroscedastic regression models considered here the standard optimal designs for the homoscedastic case remain optimal, if the random variations of the slopes are small compared to that of the intercept. If, however, the variability of the slopes are large, the optimal design points may become narrower, and the optimal design is no longer unique.

This phenomenon occurs due to the fact that design points, where the observations have high variability, are penalized. It is also observable in more complicated models, for example in quadratic regression as indicated by Mielke (2009).

7 Appendix

Proof of Lemma 1. Denote by m the smallest index ℓ such that $c_\ell < d_{\ell+1}$. Note that $m \leq K$ since $c_K < d_{K+1} = \infty$. Now $c_0 = d_0$ and $c_{m-1} \geq d_m$ for $m > 0$, which implies

$$c_m = \frac{\sum_{k=0}^m d_k}{m+1} = \frac{m c_{m-1} + d_m}{m+1} \geq d_m.$$

Hence, in either case $d_m \leq c_m < d_{m+1}$ as required.

To see that m is unique, suppose that there exist $m < n$, such that m and n both satisfy (4). From (4) we deduce that $c_m < d_m \leq d_n \leq c_n$. Hence, by the definition of the cumulative averages, we obtain

$$c_n = \frac{(m+1)c_m + \sum_{k=m+1}^n d_k}{n+1} < \frac{(m+1)d_n + (n-m)d_n}{m+1} = d_n,$$

which leads to a contradiction since $d_n \leq c_n$, and, thus, establishes the uniqueness of m . \square

References

- [1] Fedorov V.V. (1972) *Theory of Optimal Experiments*. Academic Press, New York.
- [2] Freund Ph. A., Holling, H. (2008). Creativity in the classroom: A multilevel analysis investigating the impact of creativity and reasoning ability on GPA. *Creativity Research Journal*, **20**, 309–318.
- [3] Mielke, T. (2009) D -optimal designs for paired observations in quadratic regression. *J. Statist. Plann. Inference* (submitted).
- [4] Patan M., Bogacka B. (2007) Efficient sampling windows for parameter estimation in mixed effects models. In *mODa 8 – Advances in Model-Oriented Design and Analysis* (López-Fidalgo L., Rodríguez-Díaz J.M., Torsney B., eds.). Physica, Heidelberg, 147–155.

- [5] Sibson, R. Discussion of the paper “Results in the theory and construction of D -optimum experimental designs.” by Henry P. Wynn. *J. Roy. Statist. Soc. Ser. B*, **34**, 181–183
- [6] Silvey, S.D. (1972) Discussion of the paper “Results in the theory and construction of D -optimum experimental designs.” by Henry P. Wynn. *J. Roy. Statist. Soc. Ser. B*, **34**, 174–175, 181–183.